SYM-AM-25-298

Proceedings of the Twenty-Second Annual Acquisition Research Symposium and Innovation Summit

Wednesday, May 7, 2025
Volume I

Creating Synergy for Informed Change: Transitioning Technology to the Warfighter

May 7–8, 2025

Published: May 5, 2025

Disclaimer: The views represented in this report are those of the author and do not reflect the official policy position of the Navy, the Department of Defense, or the federal government.

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

Preface & Acknowledgements

The purpose of the "Creating Synergy for Informed Change: Transitioning Technology to the Warfighter, NPS 22nd Annual Acquisition Research Symposium and Innovation Summit" is to provide a forum for the presentation of scholarly acquisition research, as well as for dialogue between scholars and acquisition policy-makers and practitioners. Research papers and presentations are given on recently completed and on-going Departments of Defense and US Navy (DoD/DON)-sponsored projects conducted by researchers at a variety of research institutions. Senior DoD/DON acquisition officials serve as panelists or keynote speakers to present their critiques and comments on research papers and priorities.

This year our symposium is coupled with an Innovation Summit and takes up the theme of "Transitioning Technology." The goal of this dual event is to explore and promote innovative ways to transition technology from research and development to programs of record to support the warfighter.

Although attendees come from many U.S. locations, as well as from some international locales, a large number are from Naval Postgraduate School (NPS) where faculty members and graduate students engage in acquisition-related research. In particular, NPS graduate students are an integral component of the research and dialogue. The Symposium serves an essential part of their graduate learning experience and provides them the opportunity to meet with senior policy-makers, practitioners, and distinguished scholars.

We gratefully acknowledge the ongoing support and leadership of our sponsors, whose foresight and vision have assured the continuing success of the Acquisition Research Program:

- Assistant Secretary of the Navy for Research, Development and Acquisition (ASN(RDA))
- Program Executive Officer, Integrated Warfare Systems
- Director, Acquisition Career Management, U.S. Navy
- Office of the Under Secretary of Defense (Acquisition & Sustainment)

The research presented in this report was supported by the Acquisition Research Program, Graduate School of Defense Management at the Naval Postgraduate School.

To request defense acquisition research, please contact:

Acquisition Research Program
Department of Defense Management
Naval Postgraduate School
E: arp@nps.edu
www.acquisitionresearch.net

Copies of Symposium Proceedings and Presentations; and Acquisition Sponsored Faculty and Student Research Reports and Posters may be printed from the **NPS Defense Acquisition & Innovation Repository** at **https://dair.nps.edu/**

Table of Contents

WELCOME: ANN E. RONDEAU, ED.D, VADM, U.S. NAVY (RET.), PRESIDENT, NAVAL POSTGRADUATE SCHOOL

Ann E. Rondeau, Ed.D, Vice Admiral, U.S. Navy (Ret.), was appointed as President, Naval Postgraduate School on January 29, 2019. She brings to the assignment an unparalleled record of leadership and achievement within the military and academia in the areas of education, training, research, executive development, change management, and strategic planning. Prior to her appointment, Adm. Rondeau served as the sixth president of the College of DuPage. Her most recent military position was as the President of the National Defense University, a consortium of five colleges and nine research centers in Washington, DC.

Rondeau has extensive leadership experience in significant military and educational roles. In 1985, she was selected and served as a White House Fellow in the Reagan Administration and went on to serve as the Deputy Commander of the U.S. Transportation Command in Illinois, Pentagon Director/Chief of Staff for the U.S. Navy Staff, Commander of the Navy Personnel Development Command in Virginia, Commander of the Naval Service Training Command at Great Lakes, Ill., Pacific Fleet Staff Chief of Staff in Hawaii, Commanding Officer of Naval Support Activity in Tennessee and other staff and commanding responsibilities with policy, planning, Fleet support, joint logistics, training and education. Rondeau retired from the U.S. Navy as a three-star admiral in 2012 and was the second woman to have achieved that rank in the Navy. She then served as a partner and later an independent consultant with the IBM Watson group.

President Rondeau's leadership has served many, both past and present, to include: Board of Directors, United States Institute of Peace; Board of Directors, German Marshall Fund; Board of Directors, The Atlantic Council; Board of Directors, National Museum of the American Sailors; Board of Directors, Council of Higher Education Accreditation; Board of Directors, Chicago Regional Growth Corporation; Board of Directors, Choose DuPage (regional development organization for Chicago northwest suburbs); Tennessee/Mid-South Economic Development Board; DoD liaison to the Center for the Study of the Presidency; Military Advisory Board (studying energy and environment impacts on national security); Flag Officer Advisory Council for Arizona State University, the National Naval Officers Association Senior Advisory Panel, the Eisenhower Memorial Commission and the National Cold War Veterans Memorial Design Steering Committee among others.

Rondeau holds a B.A. from Eisenhower College (NY), an M.A. from Georgetown University (DC) and an Ed.D. from the College of Education at Northern Illinois University in DeKalb. She also holds an honorary Doctorate in Public Service from Carthage College (Kenosha, WI) and an honorary Doctorate in Humane Letters from Rosalind Franklin University of Medicine and Science (Chicago, IL)

WEDNESDAY KEYNOTE: MR. STEVEN J. MORANI, PERFORMING DUTIES OF THE UNDER SECRETARY OF DEFENSE FOR ACQUISITION AND SUSTAINMENT

Mr. Steven J. Morani—a member of the Senior Executive Service, is performing the duties of Under Secretary of Defense for Acquisition and Sustainment (USD(A&S)). In this role, he is responsible to the Secretary of Defense for all matters pertaining to acquisition; contract administration; logistics and materiel readiness; installations and environment; operational energy; nuclear, chemical, and biological defense; the acquisition workforce; and the defense industrial base. He assumed his current position in January 2025.

Mr. Morani is concurrently serving as the Acting Assistant Secretary of Defense for Sustainment. In this capacity he provided oversight of the Department's $190 billion sustainment enterprise. His portfolio included maintenance, supply, distribution, international logistics, weapon system product support, and logistics workforce development.

Mr. Morani previously served as the Principal Deputy Assistant Secretary of Defense for Sustainment, the principal advisor to the Assistant Secretary of Defense for Sustainment.

Mr. Morani is a retired U.S. Air Force Colonel with 28 years of military service. He entered federal civilian service in June 2011.

PANEL 1. INNOVATIVE ACQUISITION: DRIVING TECHNOLOGY TRANSITION WITH RESULTS-ORIENTED APPROACHES

Wednesday, May 7, 2025	
0715 – 0815 PT 0915 – 1015 CT 1015 – 1115 ET	**Chair: Hon. David Berteau, President & CEO, Professional Services Council** **Panelists:** **Mr. Patrick Mason**, Senior Official Performing the Duties of the Assistant Secretary of the Army, (Acquisition, Logistics and Technology) **Ms. Maria A. Proestou,** Deputy Assistant Secretary of the Navy, Acquisition Policy & Budget and Executive Director, Science and Technology (S&T) Board federal advisory committee, Department of the Navy) **Mr. Peter Modigliani**, Senior Advisor, Govini

Honorable David Berteau—became the President and Chief Executive Officer of the Professional Services Council (PSC) on March 28, 2016. With more than 400 members, PSC is the premier advocate of and resource for the federal services industry. As CEO, Mr. Berteau focuses on legislative and regulatory issues related to government acquisition, budgets, and requirements by helping to shape public policy, leading strategic coalitions, and working to improve communications between government and industry, focusing on outcomes and results for the government.

Prior to PSC, Mr. Berteau was confirmed in December 2014 as the Assistant Secretary of Defense for Logistics and Materiel Readiness. He oversaw the management of the $170 billion in Department of Defense logistics funding.

Previously, Mr. Berteau served as Senior Vice President at the Center for Strategic and International Studies (CSIS), where his research and analysis covered federal budgets, national security, management, contracting, logistics, acquisition, and industrial base issues.

Mr. Berteau is a Fellow of the National Academy of Public Administration and a Director of the Procurement Round Table. He also served as an adjunct professor at Georgetown University, at the Lyndon B. Johnson School of Public Affairs, and at Syracuse University's Maxwell School.

Mr. Patrick Mason—is currently performing the duties of the Assistant Secretary of the Army (Acquisition, Logistics and Technology). He also serves as Deputy Assistant Secretary of the Army for Defense Exports and Cooperation (DASA (DE&C)), where he is the Army principal responsible for Security Assistance and armaments cooperation, export policies, direct commercial sales of Army defense articles, and international cooperative research, development, and acquisition. These programs employ over 3100 Army Soldiers and Civilians and exceed $15B annually in sales and cooperative efforts with over 150 foreign countries.

Prior to arriving at DASA (DE&C), he had served as the Deputy Program Executive Officer for Aviation and as the PEO for Aviation since 2017. Previously, Mr. Mason was the Chief of Staff for the Aviation Development Directorate of the U.S. Army Aviation and Missile Research, Development and Engineering Center. In this capacity he supported execution of the Army's $1.2 billion aviation science and technology investment portfolio. Earlier, he was the Director of the U.S. Army Redstone Test Center.

Mr. Mason has also served as the Project Manager, Technology Applications Program Office, U.S. Army Special Operations Aviation Command. In this position, he directed the life cycle management of Army Special Operations rotary wing aircraft and associated mission systems supporting the 160th Special

Operations Aviation Regiment (Airborne). In 2008, he was one of eight individuals selected from across DoD to serve as a Secretary of Defense Corporate Fellow.

His other acquisition positions include Deputy Program Manager, CH-53K; Director, Flight Test Directorate, U.S. Army Aviation Technical Test Center; and Chief, Rapid Prototyping and Integration, U.S. Army Aviation Applied Technology Directorate. Mr. Mason began his acquisition career after graduating from the Naval Test Pilot School and served as an experimental test pilot supporting aircraft development programs.

Mr. Mason's education includes a Bachelor of Industrial Engineering from the Georgia Institute of Technology; a Master of Aeronautical Engineering (with Distinction) from the Naval Postgraduate School, the U.S. Naval Test Pilot School (Distinguished Graduate), and the U.S. Army War College. He was selected as the 2013 Army Project Manager of the Year for his work with U.S. Army Special Operations Aviation and has received "Best Paper Awards" by the Society of Experimental Test Pilots and the American Helicopter Society. He was awarded the Robert N. Turk Award in 2002 for his efforts in advancing Army engineering flight tests.

Ms. Maria A. Proestou—serves as the Deputy Assistant Secretary of the Navy for Acquisition, Policy and Budget (DASN(AP&B)) under the Assistant Secretary of the Navy for Research, Development and Acquisition (ASN(RD&A)). Ms. Proestou isthe principal advisor and coordinator on matters pertaining to the Planning, Programming, Budgeting, and Execution of the Acquisition and Sustainment enterprise; she oversees the development, review and implementation of Acquisition and Sustainment Policies, including acquisition reporting, cost analysis and associated data steward responsibilities.

Ms. Proestou also serves as the Strategic Acquisition Advisor and provides executive leadership and expertise to strategic acquisition efforts across the Naval Research, Development and Acquisition enterprise. Her primary focus is on expanding the traditional and non-traditional industrial base to increase capacity and harness technological innovation. As the Executive Director of the Department of Navy (DON), Science and Technology (S&T) Board federal advisory committee, she represents the Secretary of the Navy's desire for independent assessments from national experts.

An accomplished former defense industry executive, she managed strategic contracts spanning the DON's portfolio of acquisition programs across every DON Buying Command, as well as with DoD Agencies. Prior to joining ASN(RD&A), she integrated DELTA Resources, Inc., the company she founded and led for 20 years, into a successful mid-tier C5ISRT-focused Engineering and Sustainment business. While at DELTA, she was active in promoting industry's role in enabling DOD's Joint All-Domain Command and Control (JADC2) objectives. As a founder and CEO, she successfully grew her business from a woman-owned small business to a large business earning "Great Places to Work" awards for 12 consecutive years. A strong advocate for workforce engagement, culture, and workplace flexibility; Ms. Proestou has served as a keynote speaker at Fortune Magazine's "Best Places to Work" Annual Conference and was invited to participate in the Obama White House Forum on Workplace Flexibility.

Ms. Proestou received her Bachelor of Arts degree from The George Washington University's Elliott School of International Affairs focusing on security studies. She attended the Fletcher School of Law and Diplomacy at Tuft's University receiving her Master of Arts in Law and Diplomacy following completion of a dual thesis in national security and finance.

Mr. Pete Modigliani—is currently a Senior Advisor to the USD(A&S) leading on a series of defense acquisition reforms. Pete helped shape the Adaptive Acquisition Framework, Middle Tier of Acquisition, and Software Acquisition Pathways. He's been a longtime champion of reforms from portfolio management, JCIDS, and PPBE reforms. He's been on multiple Atlantic Council Commissions and the Section 809 Panel. A former Air Force program manager

PANEL 2. ADVANCING DEFENSE ACQUISITION: INSIGHTS ON INNOVATION, NEGOTIATION, AND TECHNOLOGY READINESS

Wednesday, May 7, 2025	
0825 – 0940 PT 1025 – 1140 CT 1125 – 1240 ET	**Chair: Maj. Gen. Alice Treviño, Deputy Assistant Secretary for Contracting, Office of the Assistant Secretary of the Air Force for Acquisition, Technology and Logistics** ***Sustained Innovation Through Composable Systems*** Hon. Nickolas Guertin, former Assistant Secretary of the Navy for Research, Development, and Acquisition (ASN RD&A) ***Beyond the Table: Insights on Negotiated Terms, Synthetic Scenario Simulations, and Future Competencies in Contract Management*** Tim Cummins, President, World Commerce and Contracting ***From R&D to Readiness: Navigating Technology Transitions with the Naval Power and Energy Systems Technology Development Roadmap*** Victor Sorrentino, Director, Herren Associates

Maj. Gen. Alice W. Treviño—serves as the Deputy Assistant Secretary for Contracting, Office of the Assistant Secretary of the Air Force for Acquisition, Technology and Logistics, the Pentagon, Arlington, Virginia. She is responsible for all aspects of contracting relating to the acquisition of weapons systems, logistics, operational and enterprise efforts for the Air Force and provides contingency contracting support to the geographic combatant commanders. She leads a highly skilled staff of mission-focused business leaders and acquisition change agents to deliver $825 billion in United States Air Force and Space Force platforms. Additionally, she is the Contracting Functional Manager for nearly 9,000 professionals, who execute programs worth approximately $99 billion annually for the Department of the Air Force.

Maj. Gen. Treviño received her commission from the U.S. Air Force Academy in 1993 and is a joint qualified officer with extensive deployment experience in support of combat, humanitarian and peacekeeping/enforcement operations to Croatia, Turkey, Oman, Kuwait and Afghanistan.

Prior to this assignment, she was the Commander of the Air Force Installation Contracting Center. Maj. Gen. Treviño has also served as the Deputy Secretary of Defense's Principal Military Assistant; an unlimited dollar warranted Procuring Contracting Officer for major defense programs; and the Senior Contracting Official-Afghanistan for U.S. Central Command. She has commanded two Air Force units at the squadron level, joint units at both the group and wing levels and an Air Force unit at the wing level.

Maj. Gen. Treviño is a graduate of the U.S. Air Force Academy. She also holds dual M.A. degrees from Webster University, an M.B.A. from the Naval Postgraduate School, and an M.S. from The Eisenhower School at National Defense University.

Sustained Innovation Through Composable Systems

Hon. Nickolas H. Guertin, PE—performs consulting and is also part-time faculty at Virginia Tech's National Security Institute as a Senior Research Fellow. He is on the board of Southern New England Defense Industrial Association. He most recently served as an Assistant Secretary of the Navy for Research, Development and Acquisition (RDA), to include sustainment and oversight of the contracting community for total of $130 billion. Prior to that role, Mr. Guertin served as the Director, Operational Test and Evaluation, a senior advisor to the Secretary of Defense, reporting independently to Congress on Department of Defense weapon systems. Mr. Guertin has an extensive four-decade combined military and civilian career in submarine operations; ship construction and maintenance; development and testing of weapons, sensors, combat management products including the improvement of systems engineering; and defense acquisition. Prior to his confirmation positions, he was performing applied research for government and academia at Carnegie Mellon University's Software Engineering Institute. He received a Bachelor of Science in Mechanical Engineering from the University of Washington and an MBA from Bryant University. He is a retired Navy Reserve Engineering Duty Officer and is a licensed Professional Engineer (Mechanical). [nickolashg@vt.edu / nickolas.guertin@transformus.us]

CAPT Gordon Hunt USN (Ret.)—is a Naval Reserve Engineering Duty Officer recently retired as a naval combat system software solutions and design principles. He is the co-founder and vice president of the company Skayl focusing on system of systems integration and semantic data architectures, which enable increased systems flexibility, interoperability, and cyber security. Hunt's experience in building real systems with current and emerging infrastructure technologies is extensive. His technical expertise spans embedded systems, distributed real-time systems, data modeling and data science, system architecture, and robotics & controls. He is a recognized expert in industry on Open Architecture and data-centric systems and as a regular author and presenter, he speaks frequently on modern combat-system and command and control architectures. As a CAPT Engineering Duty Officer in the U.S. Navy, he supported combat system development and system integration scalability challenges. Hunt earned his BS in Aeronautical Engineering from Purdue University and his MS in Aerospace Engineering & Robotics from Stanford University. [gordon@skayl.com]

Robert Matthews—is a Technical Fellow and Advanced Concepts Engineer at L3Harris Technologies, advocating for and architecting MOSA developments for critical pursuits. Bob has a Bachelor of Science in Aerospace Engineering from the University of Maryland and earned his MBA at Florida Tech. He joined the U.S. Army Reserves and served 8 years as a CH-47 Chinook Helicopter Mechanic and Flight Engineer. Before joining L3Harris, he enjoyed a 20-year career as a civil servant with the Naval Air Systems Command. While at NAVAIR, Bob led the Avionics Architecture Team that coordinated with Army Aviation and The Open Group to found the FACE Consortium. Bob was elected the inaugural FACE Steering Committee Chair, serving in that role for 5 years. During that time, Bob and his team also initiated the HOST standard and collaborated with the Air Force, The Open Group, and the FACE Consortium to establish the SOSA Consortium. Bob has received multiple Navy and DoD-level awards for his work on Open Architecture. [bob_matthews@live.com]

Abstract

The enemy gets a vote. Our adversaries seek to outpace us as we seek to win in any clime and place. Winning future conflicts isn't just about innovation, it's about operational excellence and delivering useful innovation to the warfighter faster. The long-standing paradigm of building expensive, highly complex, monolithically integrated weapons systems that take years to plan and upgrade, are extremely vulnerable to asymmetric innovation (Schmidt, 2016). A simple zero-day hack can undo a decade of development and billions of dollars of taxpayer investment, leaving warfighters exposed and our economy irreparably harmed. The Hollywood climaxes of a Jedi against a Death Star or a small cell of rebels injecting a virus into networked alien attacker remain far too plausible an outcome against our inflexible and increasingly networked systems. Commercial technology cycles are outpacing DoD's ability to integrate, giving our opponents the critical time needed to exploit the same technology against us. There is no doubt that emerging technologies like AI, autonomy, quantum computing, and others just entering our imagination, will be critical to overmatch, but only if we can field it first and change it faster. Technology Superiority only worked as a strategy when seismic innovation was generational and we could ensure disproportionate access. The next conflict won't be decided by a particular technology, but by how quickly it is adapted for military use. We must shift our strategy from technology superiority to implementation superiority. This paper investigates the common pain points that have often impeded programs and proposes a set of acquisition, design, and deployment practices that shift toward composable systems to foster sustained, disruptive, and rapid innovation that outpaces our adversaries despite increasingly egalitarian access to technology.

Recent worldwide activities where American firepower has been put to the test show that the products we have built so far have been equal to the challenges presented in limited engagements, regional conflicts, or by unsophisticated opponents (Bath, 2025) but with increasingly smaller margins. In addition, the Department of Defense (DoD) is continually challenged by delays in capability delivery as well as program cost overruns. Rigidity in the current systems and the patterns of delivering improvements have been responsive to evolving operational engagements under only the most extreme and extraordinary circumstances, not as a matter of course and a reflection of purposeful design. Clearly, we do not lack technological innovation; it is the acquisition process that is broken. The authors have observed a wide variety of design teams that have been hampered by insufficient focus or funding two fundamental aspects of executing any large-scale cyber-physical system: the speed of integration and up-front considerations for future adaptability. Projects are often structured as end-item completion tasks, or a "one and done" approach to design. This is antithetical to an environment where capability is delivered by complex systems that need to be periodically improved as a part of their life cycle (Shenker, 2021). The paper highlights the detrimental impact of tightly coupled, monolithic products that end up being fragile to changing capability requirements and highlights the need to establish a requirements strategy predicated on flexibility and long-term growth. In this context, the full range of acquisition architecture must include approaches and design patterns for future-proofing systems where designs are purpose-built to change over time. A design pattern for continuous improvement (McCarthy et al., 2024).

Keywords: MOSA, Integration, Open Architecture, MBSE, Rapid-Fielding, Frameworks, Rapid, Composable, Flexible, Artificial Intelligence, lock, integrated, weapon systems, implementation superiority, acquisition process, loose coupling, cyber-physical, interoperability, composable systems, key architecture drivers

Introduction

This paper argues for the adoption of composability requirements and a strategy that prioritizes flexible architectures. It does so in the context using newly evolving methods that take advantage of selected standards that have evolved over the past few years to facilitate fluid integration and capability improvement that has long been sought, but few times achieved. Such an approach has long been projected to mitigate the pain points we will discuss, and to deliver innovation faster, more reliably, and at scale (Boydston et al., 2019). The maturation of the development models and supporting environments are now mature enough to, if properly applied, to ensure the rapid deployment of capabilities to warfighters while also maintaining a competitive edge through the ability to swiftly integrate cutting-edge technologies across diverse systems. The focus, therefore, is not merely on the technologies themselves but on architecting two different classes of systems needed for capability transition; the pipeline that creators use to build (how we identify and adapt innovation) and the products that get fielded to the users (how we design for composability and rapidly integrate). Together, these will seamlessly and affordably accommodate future innovations (Hughes & Jackson, 2021).

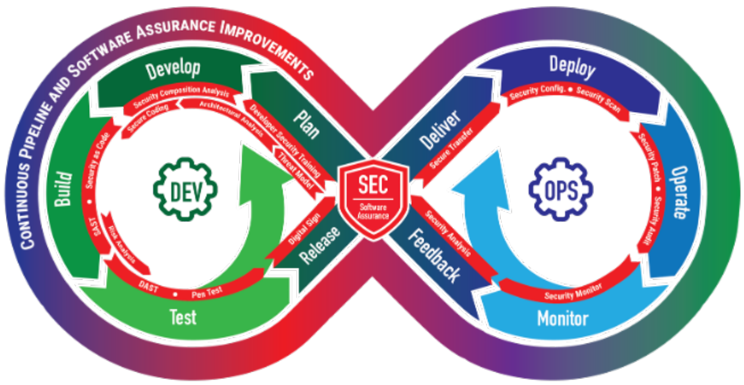


Figure 1: Continuous Evolution of Development and Capability Deployment (Chick et al., 2023)

In 2024, the three secretaries of the armed services jointly signed a memo advocating the use of MOSA (Del Toro et al., 2024). In it, they attest that the DoD Armed Forces face rapidly evolving threats across the world.

> "The dynamic and rapid change of adversary capabilities observed in current conflicts necessitates a critical warfighting capacity to integrate advanced capabilities to counter and maintain a warfighting advantage. To meet this threat, Modular Open Systems Approach (MOSA) shall be implemented and promulgated among the Military Services to facilitate rapid transition and sharing of advanced warfighting capability to keep pace with the dynamic warfighting threat."
>
> Memorandum for Service Acquisition Executives and Program Executive Officers, 2024

While this is good for guidance at the very top, we must defend it and implement it effectively. National Defense Authorization Act legislation has in the past prioritized the foundations of composable systems as critical to winning future conflicts. There is a tension with those in industry who have maintained market share by controlling key aspects of integration for major acquisition program and are profit-motivated to undermine a strategy where integration is engineered to the point where vendor-lock procurement practices can no longer be justified.

The problem is dauntingly complex. The concept is simple, the standards are written, the technology is rapidly maturing, but the work to modernize architectures and adopt cohesive and composable acquisition strategies will be hard. Utilizing the practices discussed here will have an impact to the market is as monumental as Ford's moving assembly line to cars, standard electrical outlets to appliances and the market app stores to smartphones. Innovation will skyrocket for DoD if we can muster the will to unlock it. The greatest challenge will be overcoming the business model environment that establishes long-term vendor lock conditions that makes sources of alternatives and innovation difficult to bring into the solution space for our national security.

Problem Statement

Current DoD tactical edge solutions that are highly interactive (cohesive) software systems and closely tied to physical events (cyber-physical), are not delivering capabilities affordably or fast enough to maintain overmatch (Eckstein, 2024).

Integration issues and latent defects are found and corrected through intense manual intervention, taking years and millions of dollars and repeated many times over a system lifecycle, across most DoD ACAT programs. Over time, the result has become a collection of exquisite, one-off designs that have limited interoperability and are not interchangeable at any level with other systems in the battlespace. We don't have the time or money for that anymore. In order to change this state to a new environment fielding high-quality products fast, the foundational elements of good design must be in place as a linchpin element of success. Going fast, and staying fast, takes upfront work to set the right conditions for success and continue to evolve. Instant success can happen by chance, sustained success takes deliberate design (Flyvbjerg & Gardner, 2023).

Thinking through different aspects of how the portfolio of products came to be, to include mission, procurement, business, intellectual property, technical, data, deployment and improvement, have all evolved in an ad-hoc fashion. At the very core of our acquisition process, we seek to identify capability gaps and derive new material solution requirements independent of any other solution or financial solution. Title 10 authority is delegated to program managers who often share institutional bias to avoid shared risk, common components, or interdependency on another program. These manifest in contracts to a single provider who is inherits that bias to support or create as many economic barriers to disruptive competition as possible. Our current acquisition process is focused on a 50-year-old problem that is no longer relevant and our current product architectures reflect that process. If we truly want a different outcome we must rethink our process, analyze the unintended consequences that resulted, and do the hard work of thinking through in advance a new and adaptable process that is more explicit and creates incentives for composable systems, better manages interdependencies and is more measurable to ensure we produce not just repeatable but positive outcomes. The result is a more flexible acquisition process that ensures we solve today's challenges and adapts itself to address tomorrow's. This will create an enabling environment for sustainable innovation.

Research Focus

Failure analysis, while often painful, is an indiscernible tool to practical engineers. This research draws upon the more than 100 years of collective military, civilian, and industry experience of the authors. This includes in-depth knowledge of hundreds of DoD programs of record and career perspectives from enlisted maintainer to senior officer, GS-7 to Presidential Appointee, and junior software developer to technical fellow. Despite the vastly different experiences between authors, the failure modes discovered are extensively repeated and almost universal across programs. Our investigation resulted in six common program "pain points" that have driven DoD to program ever increasingly expensive, late and brittle systems. While many initiatives and projects espouse the benefits of modular open systems, integrability, portability and many other positive attributes, most fail to demonstrate any measurable benefit at scale because they failed to address the same common pain points. This paper and our research conclusions address how composable systems can address these frequent program pain points:

1. **The "one-and-done" design trap**: Acquisition training is replete with case studies of failed programs and examples of what not to do. The culture in almost every new program office includes a mantra of needing to "do it right the first time." They have a strategy, a list of (most times) overly ambitious key performance parameters and have in the past focused their energy in designing a point solution that does not address capacity for future growth and implementation discovery. But the operational needs of these systems are too complex to ever get it right the first time and because our enemies are never static, neither can be our requirements. Even if they created that perfect solution, the cyber-physical elements would likely be obsolete before operational evaluation (OPEVAL) is complete and the cyber-security

environment is ever-changing. Sadly, many of these programs do not include an upgrade strategy or are able to get those strategies put into their initial budget plans. So even programs that start with MOSA requirements and good architectural design patterns often traded them off to achieve marginal improvements in initial capability.

Instead, every program must start with the premise that the only constant is change and therefore all systems must be cheap and replaceable, or resilient to change through adaptable architectures. Key architecture drivers (KAD) must explicitly address enterprise, domain and system composability requirements, must be measurable and made equal to, or of greater importance, than any initial system KPP values. This will ensure programs prioritize how a system is put together for iterative improvement as opposed to initial capability.

2. 10-pounds of requirements in a 5-pound budget: While “one-and-done” is principally a schedule-driven acquisition fault mode, the drive to overachieve in initial implementation is a cost driven failure mode. That said, the symptoms are often identical. At some point, almost every development program faces significant cost challenges, forcing program leadership to choose between preparing for the future or achieving as much as can be done in the near-term. Where there is program cost pressure, fear of program cancellation always follows. This drives program leaders to focus on key program measures to demonstrate success and without KAD requirements, program leadership focuses near-term capability while allowing designs to become less open, more tightly coupled and more rigid, making it even harder and more costly to add or change capability in the future. While conventional wisdom dictates that larger programs are less susceptible to pressure to pursue short-term gains at the expense of long-term flexibility, practical experience shows the opposite to be true.

The all-or-nothing uber-programs frequently become “too big to fail” which almost universally becomes the next Defense Acquisition University (DAU) lesson in what not to do (Flyvbjerg & Gardner, 2023). Sadly, programs called “too big to fail” have proven time and again to be too big to ever fully succeed. Loose coupling, modularity, and open systems provides the means to replace underperforming, low-value solutions with competitive alternatives. If an underperforming component is truly critical, it can be spun off as a separate program, developed and matured independently, while a less capable but more mature alternate “80% solution” is substituted to allow the rest of the program to proceed at lower risk. If instead of spinning-off problem elements, we use composability to segregate an uber program into smaller programs with lower risk, we have a more resilient enterprise portfolio, with a higher probability of each program success and higher probability of overall mission success. This does create new risks for interdependency between the different programs that make up the system, but this risk is manageable as long as alternatives are available. Interdependencies between programs often present themselves as Government Furnished Equipment (GFE) to the program responsible for systems integration. The programs that spin up to provide modules for larger systems must also employ composability to offer tailored, modular, and competitive solutions to better manage risk.

3. Overbuilt by design, not need: This is an insidious failure mode. In proprietary, monolithic and tightly coupled systems, it can be difficult, if not impossible, to identify and isolate critical safety and security boundaries. During requirements refinement, significant cost and schedule impacts can be identified that last the life of the program. For example, many safety related requirements require deterministic time behaviors that necessitate a Real-Time Operating System (RTOS). When only 10–20% may be related to deterministic time behaviors, the requirement for an RTOS is applied to 100% of the code. This prevents the use of much more flexible and affordable software environments that can rely on modular software techniques using mainstream development and deployment approaches (e.g., run on Linux, built with containers). This means 80–90% of the code is over-built and over-tested using an RTOS and associated development environment that can drive cost and schedule by a factor of 2-5x during initial design and every time the code is modified for the life of the system. If, however, the systems team had architected with composability and put the malice of forethought to use that to segregate and isolate safety concerns, the stringent safety requirements and RTOS usage could be constrained to a single module within the system with additional requirements to ensure a high reliability (HIREL) design that requires few upgrades over the life of the system. This minimizes the impact of the stringent requirements over the life, while also allowing the rest of the system to progress much more efficiently. Security-related concerns follow a nearly identical design pattern of needing to be physically and logically isolated to the smallest possible security boundary.

4. Custom for custom's sake: This failure mode is frequently coupled with "overbuilt by design" by merging requirements inflation and tight coupling to the point where the only feasible answer to any make-buy decision drives into custom make. Seasoned MOSA architects recognize this as rationalization based on false pretenses where ample opportunities to use Commercial-Off-The-Shelf (COTS) or Military-Off-The-Shelf (MOTS) (M/COTS) abound. The rapid evolution of commercial processing has dramatically increased the amount of high-performance M/COTS processors that are suitable for military environments. As a result, DoD processing hardware modules are commoditizing, yet first and second tier defense suppliers are remarkably resistant to the trend. Based on the author's experience across many programs, a second 80/20 rule has emerged. Despite 80% of our processing is now being general purpose, we still custom build 80% of our hardware. Application of composability to isolate the 20% of our system with specialized processing needs is a simple concept technically, but overcoming the bias to put all solutions into the same pipeline turn out to be incredibly challenging. To optimize DoD systems to maximize use of M/COTS solutions, our Composability KADs must align military system interfaces with M/COTS interface standards. The solution isn't hard, it's just hard to accept.

5. Trouble letting go: As discussed throughout the paper, initial program requirements tend to focus on technical superiority, which leads to setting overly optimistic performance of key and/or emerging technologies. Fervent adherence to overly ambitious key performance parameters (KPPs) can and has killed programs. Early performance improvements are frequently stalled by diminishing returns as the dreaded 80/20 rule takes hold. In this context it means you tend to get 80% improvement with the first 20% of the budget, then spend 80% of the budget trying to achieve the last 20% of the performance required. A lot of architectural sins are committed trying to squeeze out the last margin of performance improvement. For example, a program team may start with good design patterns for modularity and abstraction to decouple hardware from software, but performance improvements can be had if you allow hardware to be directly controlled by the software. So as pressure mounts to achieve performance requirements, we have seen engineers skip abstraction layers to directly manage hardware to gain marginal improvement but incur long-term technical debt. Now, any time you try to upgrade the hardware or port the software, you will have to modify the software for the specific hardware. In many cases the marginal performance gains that cost so much and eroded composability gets "baked in" to the next technology upgrade.

If the program team had instead defined their minimum-viable-product to the 80% solution and allowed technology to mature independently to achieve the last 20%, overall cost and schedule would have been reduced while maintaining system composability. Here again, an Agile approach to KPPs combined with composability KADs produces better, faster and lower risk life-cycle performance improvement than attempts to achieve a performance measure ahead of the where the current technology is actually capable of going. Using Agile processes, we can lock in early gains and continuously evaluate incremental gains until we reach the point of diminishing returns. Investments can then be more efficiently and effectively focused on the next technology that will provide performance with a much greater return. This approach also maintains system composability, so the next technology can be affordable integrated into the system.

6. Silos of excellence: In defense systems integration, the term "silos of excellence" refers to highly capable yet isolated systems that are engineered for performance within a single program but built with tightly coupled architectures that resist reuse, extension, or interoperability. These systems often contain custom-built interfaces, undocumented assumptions, and implementation-specific dependencies that form a "walled garden" of functionality. These silos enshrine some of our most exquisite, extraordinary and technologically superior capabilities we have today, but they are too costly and rigid. While they may perform exceptionally within their intended context, adapting them for use in other platforms, missions, or domains typically requires reengineering, code rewrites, and extensive testing, an effort that undermines agility and inflates life-cycle cost. Well-defined, open and modular interfaces are the key to breaking down these silos. Standards enable systems to communicate through shared, semantically-rich contracts that clearly define what data is exchanged, how it should be interpreted, and how components behave. By separating the "what" from the "how," systems become loosely coupled but can have high cohesion. Each module can evolve independently, be reused in new contexts, and be integrated more rapidly without deep knowledge of the internal implementation. This interface-centric architecture supports versioning, validation, and automation, making it possible to test, deploy, and update components with confidence. As MOSA efforts mature, the defense ecosystem shifts from fragile, bespoke integrations to composable, interoperable systems of systems, enabling faster innovation and mission readiness.

The Government Accountability Office (GAO) got it wrong: As we conclude our analysis of pain points and pivot toward our path forward, it is important to question how prior failure analysis missed the mark. The GAO's conclusion to many troubled programs points to a lack of early requirements definition/understanding and recommends more upfront analysis to better define requirements. That answer is at best, insufficient. Spending more time, better defining a more refined set of overly optimistic KPPs is just going to lead to more fervent application of the six pain points. Their conclusion actually makes matters worse, because it presupposes that the tomorrow's requirements are knowable and relatively stable when neither condition is true. More requirements refinement upfront is only better if the emphasis is placed on KADs that allow us specify system architectures and requirements that are resilient to changing performance and capability requirements. While the concept of composable architectures is simple, actually architecting for composability and doing it well for our complex and highly cohesive military systems is quite difficult. To be clear, good architecture is not free and integration of composable military systems is not simple. The good news is that the tools and standards needed to make composability a tractable problem are rapidly maturing.

Aspects of Architecture – Implications on Design

The story of CAD (Computer-Aided Design) modeling begins in the 1960s with simple wireframe graphics, evolving rapidly in the 1970s and 1980s into parametric solid modeling. Early tools like Sketchpad, CATIA, and Pro/ENGINEER transformed engineering from drafting boards to digital workstations, improving productivity and enabling 3D visualization of complex assemblies. However, these models remained largely isolated and focused on geometric representation and disconnected from simulation, manufacturing, or systems-level design.

In the 1990s and early 2000s, the integration of CAD with CAM (Computer-Aided Manufacturing) began a digital manufacturing revolution. Parametric modeling and associative design meant that design changes automatically updated tooling paths, reducing errors and shortening production cycles. This era saw the emergence of Product Lifecycle Management (PLM) tools and standards like STEP (ISO 10303) that enabled better sharing of product models across design and manufacturing systems. Still, CAD-CAM workflows were largely mechanical in focus, with limited integration of electrical, software, or systems-level logic.

The growing complexity of modern systems—particularly in aerospace, automotive, and defense—pushed engineers to think beyond parts and assemblies. As products became mechatronic (blending hardware, electronics, and software), the limitations of traditional CAD-centric workflows became apparent. This gave rise to Model-Based Systems Engineering (MBSE): an approach where a *digital model of the entire system* becomes the authoritative source of truth across life-cycle phases—from concept through disposal.

MBSE expands on CAD by incorporating behavioral models, functional logic, data flows, and inter-domain dependencies. Tools that use related modeling standards, like SysML (Systems Modeling Language) and UML (Unified Modeling Language) provide the backbone for modeling logical architecture, requirements, interfaces, and verification pathways. These models are not only descriptive but executable, allowing for early validation through simulation and integration with analysis tools.

The transformation from CAD-centric to model-based digital engineering has been powered by a constellation of standards and interfaces designed to facilitate interoperability and automation. Keys standards and interfaces enabling integration and automation include:

- STEP (ISO 10303): The foundational standard for exchanging 3D product data, especially geometry and product structure. Recent updates (like AP242) support PMI (Product Manufacturing Information), kinematics, and electrical harnesses.
- JT (Jupiter Tessellation): A lightweight 3D visualization format widely used in PLM systems for CAD data exchange and digital mock-up.
- SysML (OMG): The primary modeling language for MBSE, enabling modeling of system requirements, structure, behavior, and parametric constraints. Its extension into SysML v2 aims to bridge semantics more closely with engineering analysis and software execution.
- PLCS (Product Life Cycle Support – ISO 10303-239): Focused on long-term support and configuration management of complex systems, tying together engineering data over the entire life cycle.
- QIF (Quality Information Framework): Standard for metrology and quality data, enabling closed-loop quality control from design through inspection.
- MTConnect and OPC UA: Standards for machine-to-machine communication in manufacturing environments, enabling real-time integration between design models, MES (Manufacturing Execution Systems), and shop floor equipment.
- Functional Mock-up Interface (FMI): Enables co-simulation of models from different engineering domains, such as thermal, control, and mechanical systems—critical for virtual integration and digital twin development.

At the heart of today’s efforts is the Digital Thread—a traceable, integrated chain of data that connects requirements, design models, simulation results, manufacturing plans, and operational feedback. This concept builds on decades of CAD and PLM progress, now extended by MBSE to enable automation and agility across the entire engineering ecosystem (AIAA, 2023).

In this environment, model updates ripple through simulation, requirements, and even machine tool paths without manual translation. The result is faster development, fewer integration errors, and enhanced ability to manage complexity and change. As AI and generative design begin to influence engineering processes, the robust digital infrastructure enabled by

CAD, PLM, and MBSE standards is what makes the next wave of intelligent, adaptive systems development possible.

Software & System Architecture Consideration

Why is software different? It doesn't have to be. Like hardware we need to architect software for integration, standardizing integration surfaces for agility and adaptability. Unlike hardware, however, software implementation can be performed without the necessary up-front work of preliminary design and management for future change.

Modern software-intensive systems operate in environments where rapid capability delivery and frequent technology refreshes are paramount. Yet, these very systems are often shackled by integration challenges that consume resources and slow progress. To overcome these constraints, software architects must shift their perspective, treating integration not as an afterthought, but as a primary design objective. This requires careful and intentional definition of integration surfaces that enable modularity, upgradability, and resilience in software architectures.

Integration surfaces, traditionally thought of as application program interfaces (APIs), must be broadened in scope. APIs represent only one form of interface; data representation, protocol adherence, timing dependencies, and the internal use of data also form critical aspects of the integration landscape. By considering all these surfaces, architects can create infrastructures that not only reduce the complexity of connecting components but also enable those connections to evolve independently. For instance, aligning application-level data models with protocol-level and signal-level representations creates an architectural clarity that minimizes the impact of change.

The goal is to reduce the brittleness typically associated with integration. Too often, systems are designed with tight coupling between components, where even minor modifications to a single module ripple across the entire ecosystem, requiring widespread retesting, recertification, and costly engineering effort. This "tyranny of commonality" constrains flexibility and undermines the very goal of rapid capability insertion (Lunde, 2023). Instead, by clearly defining software mating surfaces—boundaries at which change can be isolated and controlled—systems become more adaptable and maintainable.

A helpful analogy is that of mechanical assembly: in a well-designed machine, parts are interchangeable and interfaces well characterized. A transmission upgrade should not require recasting the entire engine block. In software, this principle is rarely followed. Without disciplined architectural separation, changes in business logic, transport protocols, or data models often require rebuilding and redeploying the entire system. This leads to high update costs and extended downtimes.

To counter this, integration surfaces must be designed to support software update flexibility as a first-class objective. We do this by enabling software updates to occur through architected integration design. Software update flexibility is a critical enabler of operational agility. In domains like defense, aerospace, and industrial systems, the ability to upgrade individual components without disrupting the entire platform is essential—not just for capability growth, but also for sustainment and compliance. This demands an infrastructure-first architecture where integration surfaces are purposefully constructed to absorb change rather than propagate it.

Such an approach requires architects to decouple infrastructure, protocols, and data semantics from application logic. Software components should not be statically compiled with assumptions about their runtime environment, dependencies, or mission set. Instead, interfaces should be abstracted, configurable, and well-documented, allowing components to be replaced or upgraded without altering the underlying infrastructure.

A significant source of integration cost today stems from the repeated manual mapping of messages, fields, units, and primitives between systems. This repetitive activity, often buried in engineering documentation, consumes enormous effort and leads to brittle integration points. Automating these mappings through model-driven architectures, mediation layers, or canonical data models can eliminate large portions of non-value-added work.

Moreover, placing integration boundaries at appropriate levels of abstraction allows teams to manage risk more effectively. Moving integration surfaces closer to the application core—rather than at the infrastructure edge—can contain changes within a module and reduce recertification burdens. For example, modifying a data transport mechanism (e.g., switching from shared memory to TCP) should not require changes to the business logic, provided the integration surface is clearly delineated.

Ultimately, the hallmark of a robust software architecture is not only its initial design, but how gracefully it accommodates change. Architects must deliberately plan for evolution—not by hardening interfaces to resist change, but by making them flexible, modular, and expressive. This means rethinking how systems are integrated: using standards wisely, avoiding over-reliance on uniformity, and focusing on *interoperability*, not just compatibility (Carlton et al., 2021).

By taking an architecture-driven approach to integration—one that acknowledges all the software surfaces where change occurs—teams can break the cycle of rework and build systems that are ready for the demands of today and tomorrow (Allport et al., 2016).

DevOps and the Illusion of Integration: Automation Alone Doesn't Solve Architectural Interoperability

Over the past decade, DevOps has transformed software engineering by enabling rapid, automated build, test, and deployment cycles. Tools such as Jenkins, Kubernetes, GitLab CI, Docker, and Terraform have become the backbone of continuous integration and continuous delivery (CI/CD) in modern workflows (Kim et al., 2016). However, while DevOps enables faster deployment and operational responsiveness, it does not address the deeper architectural challenges required for true system-of-systems (SoS) integration, particularly in defense, aerospace, and other complex, multi-domain systems.

As emphasized in the 2021 ASNE Intelligent Ships paper (Hunt et al., 2021), DevOps automates the *Build and Deploy* stages of development, but defers critical integration work related to system behavior, semantic meaning, and interface design. These challenges are not about speed of deployment, but about cohesion, interoperability, and composability—aspects that DevOps, by itself, does not solve.

Where DevOps Ends: The Limits of Tooling in System Integration

DevOps tools provide value through automation but make key assumptions; that component interfaces are well-specified, that semantics are shared across systems, and that integration logic is known ahead of time. These assumptions rarely hold in large, evolving systems. For example:

- CI/CD Pipelines automate building and testing but don't resolve semantic mismatches or behavioral alignment (Fitzgerald & Stol, 2017).
- Kubernetes and other orchestration frameworks deploy containers but are blind to how those containers exchange and interpret data.

- Infrastructure-as-Code (e.g., Terraform) provisions compute resources, but doesn't enforce or even describe functional interactions between services.

Even within DevOps, the emphasis is on reducing time-to-deploy and increasing testing coverage. The DoD DevOps Reference Design (DoD CIO, 2019) defines Continuous Integration as automated testing and security scanning but does not prescribe methods for aligning system interfaces or documenting data semantics.

The Architectural Gap: Why Modularity and Interface Rigor Still Matter

To achieve scalable interoperability, systems must be designed around modular components with precise, semantically clear interfaces. This level of rigor is absent from most DevOps workflows. The Interface Documentation Maturity Levels (IDML) framework (Hunt & Allport, 2018) identifies levels of interface documentation maturity, showing that most DevOps efforts operate at levels 2 or 3 (human-readable or syntactic interface specifications). However, true integration requires IDML 5–7, where semantic context and compositional behavior are machine-readable and automatable.

Similarly, the Levels of Conceptual Interoperability Model (LCIM; Tolk, 2004) identifies that true composability is not achieved until LOI 5 (Conceptual)—where both the meaning and use of data and behavior are explicitly modeled. DevOps only addresses LOI 1 (Technical) and LOI 2 (Syntactic), leaving the most challenging and impactful aspects of integration unaddressed.

Beyond Pipelines: Architectures for Integration

To fill the architectural gap left by DevOps, the defense and aerospace communities are investing in open architecture frameworks that promote interface standardization, semantic modeling, and composable system design. Notable efforts include:

- FACE™ Technical Standard – A modular approach to airborne software architecture (The Open Group, 2021).
- SOSA™ (Sensor Open Systems Architecture): A collaborative standard that defines modular hardware and software interfaces for sensor systems, ensuring plug-and-play interoperability and rapid integration across vendors (The Open Group, 2023).
- Air Force GRA (Government Reference Architecture): A formal architecture effort aimed at ensuring that Air Force systems follow reference models for modularity, data exchange, and scalability across platforms and domains (Department of the Air Force, 2022).
- UDDL (Universal Data Description Language): UDDL is a machine-readable language designed to support data model documentation and semantic interoperability across systems. It provides a structured way to describe data entities, attributes, relationships, and semantics beyond traditional schema formats like XML or JSON Schema. UDDL is particularly useful in model-based integration, where precise, reusable, and shareable data definitions are essential for automating the composition of system interfaces. Unlike syntax-only representations, UDDL enables semantic layering, allowing systems to not only exchange data but also understand the context and intended use of that data.
- OMS Open Mission Systems (OMS): The standard is a U.S. Air Force initiative that defines common interfaces for integrating mission systems across different aircraft. It promotes a modular, open architecture that makes it easier and faster to add or upgrade capabilities. By separating software from hardware, OMS allows components to be reused and updated without full system redesign. It often works with the Universal Command and Control Interface (UCI) to improve communication between subsystems.

- UMMA (Unmanned Maritime Mission Architecture): The Navy's evolving reference architecture for modular and interoperable unmanned maritime systems, designed to enable faster capability insertion across unmanned surface and undersea platforms (Naval Sea Systems Command, 2022).

These frameworks address what DevOps cannot: the architectural separation of concerns, explicit documentation of behavior, and semantic mapping of interfaces required for composable and evolvable systems.

DevOps is necessary but not sufficient. While it automates deployment and life-cycle management, it does not resolve integration at the architectural level. For system-of-systems environments, the most expensive and brittle parts of development occur not in deployment, but in the misalignment of data semantics, undefined interface behavior, and manual integration rework. Addressing these challenges demands a shift toward explicit, machine-readable models of interface and behavior, guided by open architecture frameworks like SOSA, Air Force GRA, and UMMA.

Without this architectural rigor, DevOps simply accelerates the delivery of disjointed components. With architectural rigor, we can build adaptive, interoperable systems at scale.

Applying MOSA

Modular Open Systems Approach is well recognized as a strategic pilar of improving defense acquisition. This has been a multi-year approach with many laudable achievements and deserved accolades. However, there is much progress to be made, especially in achieving the business goals of MOSA (Guertin et al., 2015). There have been several attempts to assess compliance or to characterize a maturity model for what it means to achieve the business and technical goals of MOSA (Schenker et al., 2024).

When a contract includes the phrase "do MOSA to the maximum extent practical," it often implies a lack of genuine commitment to the Modular Open Systems Approach (MOSA). Essentially, it indicates that the contracting party is not willing to invest the necessary resources or funding to fully implement MOSA principles. Instead, they are opting for a more superficial or minimal compliance, rather than embracing the full potential and benefits of MOSA.

When a contract states "shall comply with MOSA policy," it often means that the contracting party is merely fulfilling a requirement for the sake of formality. From this language, it is not clear if the acquiring party is genuinely interested in the practical application or benefits of MOSA. Instead, it would appear that they are including the phrase to meet a bureaucratic or regulatory requirement, absent a measurable commitment for implementing MOSA.

The Open Systems Management Plan (OSMP) is a critical document in defense contracting that outlines the strategy for implementing the Modular Open Systems Approach (MOSA). By documenting the MOSA strategy in the OSMP, defense contractors can demonstrate their commitment to delivering innovative, cost-effective, and sustainable solutions that can evolve with changing mission requirements and technological advancements

However, when a contract includes the phrase "shall document MOSA strategy in an Open Systems Management Plan," it often implies a lack of clarity or understanding about how to implement the Modular Open Systems Approach (MOSA). Essentially, it indicates that the contracting party may lack the knowledge or implementation guidance on how to effectively request these strategies. This can lead to confusion and or superficial compliance, rather than a traceable commitment.

If the OSMP needs to be delivered with the proposal, it indicates that the contracting party is genuinely interested in the MOSA strategy and its implementation. This shows a commitment to understanding and integrating MOSA principles from the outset, making the answer to the MOSA strategy question significant and impactful.

Instead, if the OSMP is delivered after the contract award, it implies that the MOSA strategy is not a critical factor in the decision-making process for awarding the contract. In this scenario, the contracting party is not prioritizing MOSA principles during the initial stages of the contract, and the answer to the MOSA strategy question becomes less relevant and less impactful.

These sentences reflect a broader issue within the DoD acquisition process, where the true potential of MOSA is often overlooked or underutilized due to superficial compliance and lack of genuine commitment.

Real changes in achieving the business objectives of MOSA would be most likely to be achieved if the program crafting the contract solicitation also had sufficient depth of understanding of the available standards and technical approaches needed to perform to the desired outcome of the contracts. Key indicators that this is the case would be seen if the requirements specified a technical compliance to standards that are known to be useful for the intended purpose. Figure 2 graphically depicts a set of MOSA-related standards and how they can be applied in system design decisions.

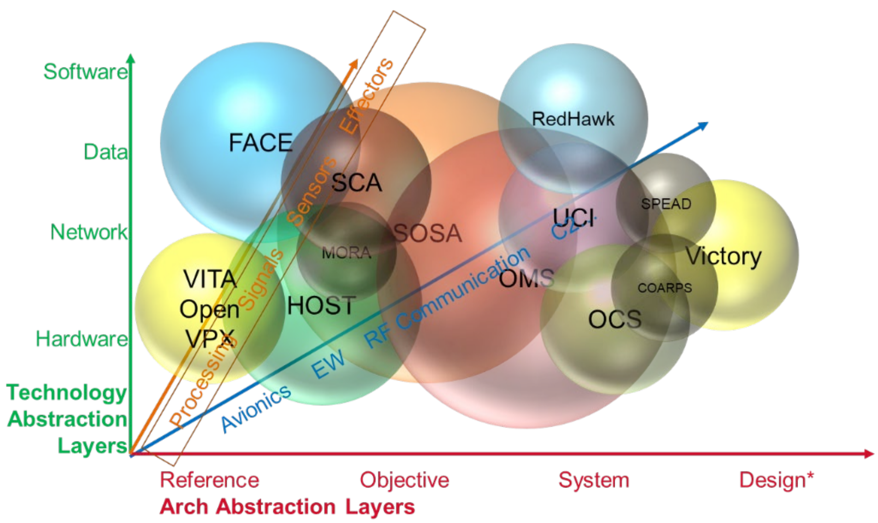


Figure 2. Mapping of Standards to Abstraction Layers for Technical MOSA Strategies

Test objectives and demonstrations of modular design criteria or interchange of components should be clearly spelled out and unambiguous to the reader. One of the critical enablers of this approach would be an open dialogue with industry about the goals of the product for life-cycle performance improvement and compliance to open standards. Industry/Government collaboration prior to release of the request for proposal (RFP) is key to achieving the Governments objectives.

FORGE Act Connection

The time is ripe for rethinking the transformation of effort and resource allocation into battlefield advantage. The crescendo of calls for achieving different outcomes is at a fever pitch and the willingness to act on this confluence is heartening; however, in our drive to be different, we risk culling protections that would lead us back to the pit of the six pain points discussed above.

The most recent draft National Defense Authorization Act (NDAA) seeks to remove barriers to innovation and puts a heavy premium on commercial products and flexible acquisition approaches. While the thrust of this reform effort is laudable, it runs the risk of making it easy to present tight vertical integration. Knitted carefully into the FORGE Act language are key sentences that would promote large “one and done” procurements that would create unbreakable barriers to entry to all but the largest companies and once again risk vendor-lock solutions.

DoD risks losing access to American innovation if it does not become a better customer and carefully cull the counterproductive language from the FORGE Act. Doing so requires process improvements and investing in the management of the architecture of architectures it needs to create and maintain competitive pressure from large and small businesses. These are the methods by which we ensure the product is both sound and has many market alternatives. In doing so, it can also facilitate the use of commercial product development strategies while preserving fair opportunities with industry. In the years after World War II, it was common for DoD to create austere prototypes of several platforms across organizations. Then, it would put into production only the very best of the bunch. However, in doing so, we must preserve the learning that came from the long road we have traveled to ensure the government can make sound procurement decisions and acquisition flexibility to ensure we attract competitive alternatives and reward innovation on an open playing field.

Changes Needed for Certification for Use and Operational Test and Evaluation

We have made the assertion that the operational community needs to have products updated and re-fielded, fast. We also need to achieve our objectives against a thinking enemy who brings new things to the fight. To be employed at the speed of need, the system must also be composed to reduce fragility through modularity and loose coupling, facilitate rapid testing to find defects, and then quickly push corrections out to operational employment. Also, if a latent problem is found in the field, the system must be built to fall back to a known good state, report the issues, and be receptive to a subsequent update (McQuade et al., 2019). In future conflicts, a zero-day attack cannot wait months or even years for a proven remedy, yet that is routinely the case today.

Not fully addressed in the pain points above, the changing nature of software-intensive and cyber-physical systems use requires that testing have a higher level of prominence in the cost-trades associated with advancing a new and innovative design (Fields, 2018).

The concept of a “testable architecture” necessarily involves a contextual awareness of how the system performs its objectives and an approach to include the perspectives of a wider array of participants than the developers alone. This must be achieved by considering the testability of architecture as a first principle (Guertin & Hunt, 2017). With systems undergoing regular updates and deployment, the full range of stakeholder needs have to be present all along the path of creativity.

The application of artificial intelligence (AI) and machine learning (ML) is being considered for use across the defense portfolio. These designs require their own attention to architecture precepts for testing and monitoring of in-use behaviors that comport to structured frameworks for how testing will be performed as a life-cycle consideration, not just for producing improvements, but also for assessing how performance is changing in the deployed state (McCarthy et al., 2024).

- The application of AI is done in the context of solving a problem.

- Data has to be curated and managed to be effectively applied and integrated.
- Identification of training data and segregating of test data along with the best-fit use of algorithms is critical to success.
- Understanding of the operational environment and how the trained algorithm will behave in the use-case can be better understood through modeling and simulation.
- Once the AI/ML model is integrated into the system, then full useability and operational workflow can be assessed for effectiveness, survivability, and suitability.
- Lastly, how an AI-enabled system is continuously updated, recalibrated, evaluated for behavior drift and the need to trigger re-assessment needs to be established.

The tools and methodologies for testing must incorporate a cost-risk balance between automation and defect detection. Automation of test does not come for free, but advancing tools employed in this environment can identify areas ripe for investment. An area that calls for further study is to use the integrated decision support key (IDSK) as a possible framework for identifying parameters that will need to be evaluated as a life-cycle continuum as the system being built is improved over time.

Summary and Recommendations

Winning the future fight requires more than innovation and demands implementation superiority. This paper has explored how tightly coupled systems, siloed development practices, and legacy acquisition approaches limit our ability to respond with speed and agility. The solution lies in composable architectures, where modular design, rigorous interface definition, and semantic interoperability enable continuous evolution, rapid fielding, and affordable upgrades. But composability doesn't happen by accident; it must be deliberately architected, measured, and enforced.

To realize these benefits, we must elevate architecture to a first-class engineering discipline. Architecture is not just about structure—it is the mechanism that makes complexity manageable and change possible. It is not easier than design, but it is more essential. Without well-formed architectural strategies, programs fall into familiar traps: brittle integration, vendor lock, and systems that are too costly to adapt or sustain.

Crucially, we must stop trying to align system-specific designs to standards in isolation. Instead, we must align architecture concepts across standards, establishing a normalized foundation of abstraction, separation of concerns, and modularity. Each standard should clearly document where architectural principles are preserved and where they are refined into constraining design specifications. This allows for interoperability between standards, not just between systems.

The tools, models, and frameworks are now mature enough to support this approach. The opportunity is real, but success depends on clear leadership intent, acquisition reform, and a shared understanding that composability is not a technical preference, but a strategic necessity. Systems built today must be designed to evolve—because if we don't build to change, we won't be able to compete.

The recommendations are clear and need courage and commitment to realize and achieve our goals. We are on a similar path to the decades long model-based transformation which has occurred for hardware and material systems manufacturing. To see this to the end, we must:

- Architect first. Elevate architecture to drive integration, not lag behind it.
- Prioritize composability as a Key Architecture Driver (KAD) equal to performance and cost.

- Align architectures across standards and use standards to refine—not redefine—those concepts.
- Measure and enforce interface rigor through appropriately architected interfaces boundaries.
- Reject superficial compliance. Build systems that are testable, modular, and ready to evolve.

Architecture is not a checkbox. It is the battlefield where flexibility, speed, and superiority are won.

References

Allport, C., Hunt, G., & Revill, G. (2016). Rethinking system integration. *Electronic Engineering Journal*. http://www.eejournal.com/archives/articles/20160609-interoperability

American Institute of Aeronautics and Astronautics. (2023). *Digital thread: Definition, value and reference model*.

Bath, A. (2025, January 16). Navy fired more than 200 missiles to fight off Red Sea shipping attacks, admiral says. *Stars and Stripes*. https://www.stripes.com/branches/navy/2025-01-16/houthis-navy-red-sea-missiles-drones-16500246.html

Boydston, A., Feiler, P., Vestal, S., & Lewis, B. (2019, September). *Architecture centric virtual integration process (ACVIP): A key component of the DoD digital engineering strategy*. U.S. Army.

Carlton et al. (2021). *Architecting the future of software engineering, a national agenda for software engineering research & development*. Carnegie Mellon University, Software Engineering Institute. https://insights.sei.cmu.edu/documents/1309/2021_014_001_741214.pdf

Chick, Pavetti, & Shevchenko. (2023). *Using model-based systems engineering (MBSE) to assure a DevSecOps pipeline is sufficiently secure* (Technical Report CMU/SEI-2023-TR-001). Carnegie Mellon University, Software Engineering Institute. https://doi.org/10.11184/R1/22592884

Colclough et al. (2024). Software development practices are changing the character of the test and evaluation enterprise. *ASNE Naval Engineer's Journal*, *136*(1 & No. 2).

Del Toro, Kendall, & Wormouth. (2024). *Modular open systems approach for Department of Defense weapon systems*. https://www.cto.mil/wp-content/uploads/2024/12/Tri-Service-Memo-Signed-17Dec2024.pdf

Department of Defense Chief Information Officer. (2019). *DoD enterprise DevSecOps reference design*. https://dodcio.defense.gov

Department of the Air Force. (2022). *Digital engineering playbook: Government reference architecture (GRA)*. Air Force Digital Transformation Office.

Eckstein. (2024, March 21). US Navy making Aegis updates, training changes based on Houthi attacks. *Defense News.* https://www.defensenews.com/naval/2024/03/21/us-navy-making-aegis-updates-training-changes-based-on-houthi-attacks/

Fields. (2018). *Designing and acquisition of software for defense systems*. Defense Science Board.

Fitzgerald, B., & Stol, K. J. (2017). Continuous software engineering: A roadmap and agenda. *Journal of Systems and Software*, *123*, 176–189. https://doi.org/10.1016/j.jss.2015.06.063

Flyvbjerg, B., & Gardner, D. (2023). *How big things get done: The surprising factors that determine the fate of every project*. Currency.

Guertin & Hunt. (2017). *Transformation of test and evaluation: The natural consequences of model-based engineering and modular open systems architecture*. Naval Postgraduate School, Acquisition Research Program.

Guertin, N., & Schenker, F. (2021). Using value engineering to propel cyber-physical systems acquisition. In *Proceedings of the 18th Annual Acquisition Research Symposium*. Naval Postgraduate School.

Guertin, N., Schmidt, D. C., & Sweeney, R. (2015). How the Navy is using open systems architecture to revolutionize capability acquisition. In *Proceedings of the 12th Annual Acquisition Research Symposium*. Naval Postgraduate School.

Hand, S., Lombardi, D., Hunt, G., & Allport, C. (2018). *Interface documentation maturity levels (IDML): An introduction*. Open Group FACE™/U.S. Army Technical Interchange Meeting.

Hughes & Jackson. (2021). *A framework for DevSecOps evolution and achieving continuous-integration/continuous-delivery (CI/CD) capabilities*. Carnegie Mellon University, Software Engineering Institute.

Hunt, G., Allport, C., Foley, S., & Hand, S. (2021). Achieving warfighting lethality via system and semantic system-of-systems model-based engineering integration. *Proceedings of the ASNE Intelligent Ships Symposium*.

Kim, G., Humble, J., Debois, P., & Willis, J. (2016). *The DevOps handbook: How to create world-class agility, reliability, & security in technology organizations*. IT Revolution Press.

Lunde. (2023). *The true cost of software complexity*. https://ronlunde.com/post/costofcomplexity/

McCarthy et al. (2024). Key steps to fielding combat credible AI-enabled systems. *ASNE Naval Engineer's Journal*, *136*(1 & 2).

McQuade et al. (2019). *Software is never done: Refactoring the acquisition code for competitive advantage* [The Software Acquisition and Practices [SWAP] Study].

Naval Sea Systems Command. (2022). *Unmanned maritime mission architecture (UMMA) overview*. Presented at NDIA Expeditionary Warfare Conference.

The Open Group. (2023). *Sensor open systems architecture (SOSA) technical standard edition 1.0*. https://www.opengroup.org/sosa

Schenker, A., Guertin, H., & Schmidt, D. (2024). *A model for evaluating the maturity of a modular open systems approach.* Carnegie Mellon University, Software Engineering Institute.

Schmidt, D. C. (2016, July 11). *A design for maintaining maritime superiority: A naval perspective on open-systems architecture* [Blog post]. SEI Blog. https://insights.sei.cmu.edu/sei_blog/2016/07/a-naval perspective-on-open-systems-

Tolk, A. (2004). Composable mission spaces and M&S repositories—applicability of open standards. *Spring Simulation Interoperability Workshop*. IEEE.

Beyond the Table: Insights on Negotiated Terms, Synthetic Scenario Simulations, and Future Competencies in Contract Management

Tim Cummins—is President of World Commerce & Contracting and a Professor in the School of Law at the University of Leeds. For 25 years, he has led research and development of standards for Commercial and Contract Management, recognized with the Financial Times "Market Shaper of the Year" award in 2019. He founded the world's only non-profit association dedicated to Commercial and Contract Management, working with its 80,000 members to improve trading relationships, and serves as the first Executive Director of the Commerce and Contract Management Institute.

Kraig Conrad—is the Chief Executive Officer of NCMA, a role he assumed in November 2018. Under his leadership, NCMA has adopted an ambitious strategic plan to expand reach, relevance, and value for contract management professionals. Conrad brings over 20 years of association leadership experience, previously serving as CEO of Professional Risk Managers' International Association and holding leadership positions at the National Investor Relations Institute and other organizations.

Daniel J. Finkenstadt, PhD—is an expert in contract management and negotiation strategies, focusing on innovative approaches to acquisition planning and supply chain resilience. He is the owner and principal at Wolf Stake Consulting LLC and an active duty officer. His research explores supply chain management and the application of generative AI in procurement, with publications in Harvard Business Review, California Management Review, and various peer-reviewed journals. He is the author of the books Supply Chain Immunity (Springer, 2023) and Bioinspired Strategic Design (Production Press, 2024).

Abstract

This paper examines the critical disconnect between heavily negotiated contract terms and those that actually drive successful performance outcomes. Through a comprehensive 2024 study involving over 600 contracting professionals, we demonstrate that government procurement practices remain overly focused on administrative details rather than performance-based terms, potentially costing $100 billion annually in inefficiencies. We present findings from an exploratory study comparing AI-generated versus human-authored negotiation training scenarios, revealing that generative AI can produce comparable quality materials in minutes rather than hours. Finally, we outline essential competencies for modern contract managers, emphasizing the need for skills in strategic negotiation planning, performance-focused drafting, risk management, and AI-augmented decision-making. This research underscores the importance of aligning negotiation practices with operational realities to foster adaptive, collaborative business relationships that create sustainable value.

Introduction

Contracts are foundational to business relationships, defining obligations, allocating risks, and providing mechanisms for conflict resolution. However, research consistently shows that the terms most heavily negotiated, such as pricing, payment schedules, and indemnification clauses, do not always align with the factors that most influence contract performance (World Commerce & Contracting [WorldCC], 2022). This disconnect points to a gap in negotiation strategies and underscores the need for a more performance-focused approach to contracting.

Agency theory provides an illuminating framework to analyze this phenomenon. As Eisenhardt (1989) explains, negotiated terms often function as tools for addressing risks inherent in agency relationships, such as asymmetric information and moral hazard. For example, pricing terms and limitation-of-liability clauses are frequently used to mitigate financial risk and manage expectations. This focus also reflects a mindset identified in prior research among legal and contracting professionals, one of "preventism," in which practitioners see their primary role as preventing failure rather than enabling success (Bauman et al., 2019). This mentality results in contracts that prioritize rigid control mechanisms at the expense of adaptability and long-term performance.

However, these provisions often fail to address operational risks, such as coordination breakdowns or delivery challenges, which are far more likely to impact contract outcomes.

Moreover, the prioritization of financial terms over operational details may stem from the bargaining dynamics between parties. Negotiators may focus on highly visible and quantifiable elements, such as cost and penalties, to satisfy immediate concerns or demonstrate value to stakeholders. In doing so, they may neglect less tangible yet equally critical elements, such as the clarity of roles, expectations, and dispute-prevention mechanisms.

Such negotiation practices often result in wasted resources, as companies invest significant time and money haggling over terms that rarely come into play, while overlooking the operational details that truly drive success and value creation. This risk-focused approach not only stifles innovation by prioritizing rigid standards over creative solutions but also damages relationships by fostering an atmosphere of competition rather than cooperation. In this environment, transparency and openness are scarce, leading to contracts that divide rather than unify. Ultimately, by concentrating on self-protection, organizations frequently miss opportunities for mutual gain and long-term value (Cummins & Finkenstadt, 2024).

To bridge this gap, negotiators could adopt strategies that align better with long-term performance objectives. For example, contingency clauses and flexible pricing models could address uncertainties while promoting collaboration and adaptability (Eisenhardt, 1989). Emphasizing coordination provisions, such as shared milestones and transparent reporting structures, can also reduce risks associated with asymmetric information and moral hazard. Value creation in negotiations requires a paradigm shift. As Bazerman (2025) explains, negotiators must identify and prioritize issues that are critical to both parties, enabling trades across those issues to maximize overall gains.

Causes of Contract Conflicts and Disputes

While well-negotiated terms can mitigate risk, the root causes of disputes often lie beyond the surface of contract language. Disputes frequently arise from poorly defined roles, vague performance standards, and unforeseen circumstances. As MacMahon (2018) notes, ambiguity in contract language is a persistent issue, creating misunderstandings and disagreements between parties.

Agency theory adds another layer of understanding to these challenges. Asymmetric information, when one party has more knowledge or control than the other, can lead to opportunistic behavior, where one party manipulates terms to their advantage (Eisenhardt, 1989). Similarly, moral hazard occurs when a party takes risks or makes decisions that impose costs on others, particularly in the absence of effective monitoring. Recent work in relational contracting, especially the articulation of nine core relational principles, provides a complementary lens for understanding how contracts should facilitate long-term collaboration and adaptive administration (WorldCC, 2016).

Strategies for Minimizing Disputes

To prevent disputes, negotiators must go beyond traditional contract terms and address the root causes of conflict. Unforeseen events, such as economic disruptions or supply chain breakdowns, further complicate matters. Incorporating contingency clauses can account for uncertainties, specifying actions and remedies for various scenarios. These measures not only enhance the resilience of the contract but also promote a spirit of partnership between parties. However, even meticulously negotiated contracts cannot anticipate every contingency, highlighting the importance of incorporating flexible, adaptive terms. For example, dispute resolution clauses, such as mediation or arbitration provisions, provide structured avenues for addressing disagreements before they escalate (Amoah & Nkosazana, 2023).

Building trust through collaborative negotiation is another key strategy. Eisenhardt (1989) emphasizes the importance of aligning incentives and reducing information asymmetry. Transparency and open communication are critical, as are performance-monitoring mechanisms that hold parties accountable. Contracts that prioritize shared goals and clearly define responsibilities are less prone to conflict.

Implications for Contract Negotiation Practices

The misalignment between heavily negotiated terms and contract performance highlights the need for a paradigm shift in negotiation practices. Agency theory suggests that negotiators should prioritize terms that reduce information asymmetry, foster accountability, and promote shared outcomes (Eisenhardt, 1989). For instance, incorporating performance-based incentives or penalties can align parties' interests and improve outcomes. Bazerman (2025) outlines several practical approaches, including building trust, asking targeted questions, and sharing information strategically to foster reciprocity.

By focusing on long-term performance rather than immediate gains, negotiators can craft contracts that are both flexible and robust. Coordination mechanisms, such as regular progress reports and shared decision-making, can further enhance collaboration and mitigate risks. These strategies not only improve the likelihood of successful performance but also strengthen the foundation for future partnerships.

Conclusion

Despite the foundational role of contracts in business relationships, our review reveals a critical misalignment: the terms most heavily negotiated often differ from those that drive conflicts during contract performance. Agency theory offers valuable insights into this phenomenon, highlighting the impact of asymmetric information, moral hazard, and misaligned incentives. Additionally, operational risks, coordination breakdowns, and unforeseen circumstances often exacerbate disputes, underscoring the need for a shift in negotiation priorities and strategies.

In the next section, we present findings from the *2024 Most Negotiated Terms* study, a collaborative research endeavor by the Commerce and Contracting Institute (World Commerce & Contracting and NCMA). This study reveals the stark disconnect between the most negotiated contract terms and those that generate the greatest conflict during performance. Following this, we share results from a short exploratory study examining the potential of generative AI to enhance negotiation scenario planning and prepare teams at scale, aiming to proactively address these challenges.

We conclude with a discussion of the competencies negotiators and contract managers must develop to better align their negotiation focus with contract performance outcomes. By fostering skills in scenario planning, risk management, and collaboration, we can bridge the gap between negotiation practices and performance realities, ensuring contracts serve as tools for sustainable and successful business relationships.

Section 2: Most Negotiated Terms 2024 Study Findings and Analysis

Introduction

Contract negotiations represent the foundation of business relationships, yet recent research reveals a persistent disconnect between what organizations negotiate most frequently and what actually matters for successful business outcomes. The 2024 World Commerce & Contracting study was conducted in collaboration with the National Contract Management Association as part of the new Commerce and Contract Management Institute. The findings, from over 600 contracting professionals with a primary focus on U.S. government procurement and contracting, illuminates this paradox while highlighting the particular challenges faced in negotiations between government buyers and suppliers.

The study reveals that despite significant shifts in the business environment, government procurement remains surrounded by detailed rules and regulations intended to achieve cost control and value while protecting against misuse of funds. In the United States, this has resulted in relatively complicated and inflexible procedures, which constrain freedom of action in the acquisition process, such as limiting negotiation and imposing costly bureaucracy.

This persistence of traditional negotiation focuses becomes particularly problematic when examining the relationship between government agencies and suppliers. The research indicates that approximately 70% of government buyers would welcome greater freedom to negotiate, yet many procurement practices remain adversarial and risk-focused, driven by process rather than outcomes. This dynamic not only affects individual business relationships but has broader implications for supply chain resilience and innovation, as well as substantial cost implications.

The data demonstrates a potential to modernize procurement practices that could reduce costs by as much as 13.3%, translating into $100 billion in savings. This misalignment between negotiation focus and operational needs suggests an opportunity for fundamental reform in how government approaches contract negotiations.

Methodology and Sample Characteristics

The study's findings are based on responses from more than 600 professionals involved in government procurement and contracting, with a primary focus on the U.S. federal government. The demographic composition of respondents provides a balanced perspective across organization sizes, roles, and sectors.

Data Collection

The survey captured insights from both buyers and suppliers, shedding light on their experiences, challenges, and practices, and the impact these have on contract outcomes.

Sample Demographics

Figure 1: Sample Demographics from Most Negotiated Terms Study

Comparative Analysis: Government Buyer and Supplier Negotiation Dynamics

The study reveals fundamental differences in how government agencies and suppliers approach contract negotiations, both in terms of priorities and preparation capabilities. The data demonstrates a complex dynamic where regulatory frameworks significantly influence not just negotiating power, but the entire approach to contract formation and risk management.

Negotiation Priorities and Power Dynamics

Government buyers consistently prioritize cost reduction or budget management (72%), ensuring compliance with regulations or legal requirements (65%), improving service or product quality (62%), and being able to demonstrate value-for-money outcomes (44%).

This contrasts notably with suppliers, who place greater emphasis on protecting proprietary information, intellectual property rights, and ensuring clarity of scope and obligations. This divergence reflects the fundamental power imbalance in these relationships, with suppliers reporting that government agencies tend to adhere strictly to established terms and regulations, resulting in less room for negotiation.

Most Discussed Terms vs. Most Important Terms

The research reveals a critical disconnect between the terms that are most frequently discussed and those considered most important. Government buyers most frequently discuss:

For government buyers, "Amendments/Changes to Contract" tops the negotiation frequency list (60%), yet ranks only 7th in importance (31%). Similarly, while they frequently negotiate on "Price/Charge/Price Changes" (58%), it ranks just 6th in importance (32%). Most strikingly, "Acceptance, Inspection and Quality Assurance" doesn't appear in the top 10 most negotiated terms but is considered the most important term (58%) by government buyers. This suggests government buyers may be spending negotiation time on administrative matters rather than focusing on their highest-value concerns related to quality assurance and project outcomes. It also suggests that negotiation behaviors may be driven less by outcome optimization and more by the need for control and risk avoidance. The reluctance to engage on cost or performance issues likely reflects a fear of opportunism, which ironically may reinforce the very adversarial behaviors procurement seeks to avoid.

For suppliers, there's better alignment between negotiation focus and importance. "Scope of Work" tops both their negotiation focus (3.0 mean score) and importance ranking (52%). However, "Indemnification" ranks 4th in negotiation frequency but doesn't appear in their top 10 most important terms. Meanwhile, "Intellectual Property and Data Rights" ranks 7th in negotiation frequency but jumps to 6th in importance (41%). Both buyers and suppliers rank "Contract Type" highly in importance (3rd for both) despite it not appearing in either's top 10 most negotiated terms, indicating a crucial aspect that may be predetermined or underaddressed in actual negotiations.

Preparation and Resource Utilization

The research reveals that government personnel work with a wide variety of suppliers and consequently encounter different levels of sophistication. The two top challenges reported by government buyers relate to supplier knowledge and skills, specifically:

1. Supplier understanding of contract terms (76%)
2. Skills and knowledge of their negotiators (46%)

Only 35% of government participants provide tools or resources to help suppliers understand or negotiate contract terms, primarily in the form of written guides or manuals (24%), templates or standard form contracts (19%), websites (17%), and in-person training sessions (17%).

Impact of Contract Type on Negotiation Practices

Government buyers acknowledge that contract negotiations are impacted by the contract type, due to the influence this has on allocation of risks, responsibilities, and incentives between the parties. While suppliers also acknowledge many of these points, they have different perspectives and concerns.

Risk Allocation Perspectives

Government Buyers: Different contract types distribute risks differently. For instance, in a Firm-Fixed-Price (FFP) contract, the contractor assumes the majority of the cost risk, which may lead to negotiations focusing on higher pricing to mitigate potential losses. Conversely, in Cost-Reimbursement contracts, the government bears more risk, prompting discussions on cost control and oversight mechanisms.

Suppliers: Suppliers emphasize the importance of indemnification clauses, especially in T&M contracts, to mitigate potential liabilities. They express heightened concern over FFP contracts due to the increased risk they bear, necessitating meticulous negotiation of terms like scope and pricing.

Flexibility and Scope Changes

Government Buyers: Contracts like Time & Materials (T&M) or Labor-Hour agreements offer flexibility to accommodate changes in scope. Negotiations for these contracts often involve detailed discussions on hourly rates, labor categories, and mechanisms for managing scope changes to prevent cost overruns.

Suppliers: Suppliers stress the need for precise Statements of Work (SOW) in FFP contracts to prevent ambiguities that could lead to unforeseen costs. They are also attentive to payment structures, particularly in T&M contracts, to ensure timely compensation for services rendered.

Disagreements and Disputes

Both buyers and suppliers concur that roughly 25% of contract negotiations face significant disagreements during performance, though their perspectives on the causes differ somewhat. Government buyers identify the primary dispute sources as changes and modification of terms (51%), acceptance, inspection, and quality assurance (45%), and amendments to contracts (42%), while suppliers similarly rank changes and modification of terms highest but at a lower rate (48%), followed by amendments to contracts (44%) and acceptance/inspection issues (40%). Notable in both rankings is the consistency of the top five issues, with delivery dates and payment terms completing both lists, though suppliers consistently report higher rates of disagreement on payment options (34% vs. 28%) and delivery terms (36% vs. 31%). A significant disparity exists in the area of intellectual property and data rights, with only 17% of government buyers but 41% of suppliers saying this generates disagreements during performance.

Supplier Cost Impact and Potential Savings

When asked about the impact of non-negotiable terms and complex contracting processes, suppliers estimate that if government agencies were more open to negotiation and simplified their processes, they could reduce their overall transaction costs (pre- and post-award) by an average of 13.3% across all contract types. Based on 2023 Federal spend, this translates to approximately $100 billion in potential savings. Estimated cost reduction potential varies by contract type, with the highest savings opportunities in Public–private partnerships (18%), Cooperative research and grants (15%), and Other contract types (15%), followed by OTA (14%) and International agreements (13%), while IDIQ requirements, FFP, Single contracts, and FAR-based contracts all show 12% potential savings, and Ordering agreements offer 11% savings potential.

Suppliers identified the following areas for significant cost savings:

1. Lengthy approval and procurement processes (highest concern in Ordering Agreements at 73.7% and FFP at 67.9%)
2. Rigid contract terms with limited negotiation room (particularly challenging in Other at 83.3% and Cooperative Research and Grants at 71.4%)
3. Complex regulatory and compliance requirements (notably problematic in International Agreements at 51.7% and Cooperative Research & Grants at 50.0%)
4. Requirement for extensive documentation and reporting (a major issue in International at 48.3% and Ordering Agreements at 50.9%)
5. Intellectual property rights and data security concerns (prominent in Cooperative Research and Grants at 57.1% and Public-Private Partnerships at 57.1%)

Addressing Power Imbalances

When facing power imbalances, suppliers use various strategies to approach negotiations, though these vary by contract type. The most common strategies include:

1. Applying industry standards or benchmarks in negotiations (50–83% frequency)
2. Emphasizing unique strengths or exclusive advantages of offerings (50–70%)
3. Establishing firm alternatives and clear walk-away thresholds (47–76%)
4. Highlighting potential long-term collaborations or partnership benefits (41–57%)
5. Leveraging personal or professional relationships in the public sector (21–55%)

Importantly, only 7–25% of suppliers report agreeing to terms with the intention to renegotiate post-award, challenging the common perception that suppliers often accept unfavorable terms planning to recover through changes after award. It's also important to note that these dynamics are not unique to the public sector. Similar tensions exist in commercial negotiations, but they are intensified in government procurement by inflexible, rules-based frameworks that limit adaptive problem-solving.

Recommendations and Future Directions

The research findings point to several critical areas for improvement in contract negotiations between government agencies and suppliers, with particular emphasis on practical strategies for creating more balanced and effective relationships.

Contract Simplification Strategies

Government employees recognize several areas for improvement where costs and cycle times could be reduced and outcomes improved:

1. Lengthy approval and procurement processes (66%)
2. Complex regulatory and compliance requirements (51%)
3. Requirement for extensive documentation and reporting (42%)
4. Rigid contract terms with little room for negotiation (40%)
5. Budget constraints or volatility within public sector organizations (28%)

Balanced Risk Allocation and Technology Adoption

The path forward lies in streamlining approval processes, fostering trust-based relationships, and shifting focus from risk avoidance to collaborative value creation. The imposition of standard models fosters a compliance culture in which acquisition professionals equate risk mitigation with conformity. This discourages early-stage engagement on fit-for-purpose models and undermines opportunities for innovation. By balancing regulatory oversight with operational flexibility, agencies can reduce costs and grow the supply market, while improving existing supplier partnerships and delivering greater public value.

Conclusion

The 2024 study reveals a critical disconnect between traditional negotiation practices and successful government procurement relationships, highlighting an urgent need for reform to capture $100 billion in potential savings. As agencies face pressure to deliver more value with limited resources, success requires balancing risk management with operational practicality through simplified contracts and improved technology. Rather than rigid adherence to traditional terms, the future lies in creating flexible, clear agreements that benefit both government agencies and suppliers.

Section 3: Generative AI in Negotiation Scenario Development: A Comparative Analysis Study

Negotiation is consistently ranked among the most important skills for purchasing and supply management PSM professionals, appearing in 73% of PSM skills studies (Heunis et al., 2024; Stek & Schiele, 2021), and is crucial for both internal organizational relationships and external buyer-supplier interfaces (Saorín-Iborra & Cubillo, 2019). Negotiations are counted as one of the most complex and demanding areas of PSM activity and essential for sustaining competitive advantage for organizations (Carr & Pearson, 2002; Ramsay, 2007). Recent research has highlighted the importance of adaptability in negotiation preparation within PSM contexts using scenario-based training (Heunis et al., 2024). While traditional approaches to developing negotiation scenarios rely on expert knowledge and significant time investment, emerging generative AI technologies may offer new possibilities for creating diverse, adaptable training materials efficiently. Recently published managerial studies indicate that AI-generated scenarios hold significant potential for both general business scenario planning and contingency planning (Finkenstadt et al., 2023; Finkenstadt et al., 2024). This section presents findings from an exploratory study specifically examining the comparative efficacy of AI-generated versus human-authored contract negotiation scenarios.

Research Questions and Study Design

The primary research question addressed in this study was: ***How do AI-generated negotiation scenarios compare to human-authored scenarios in terms of perceived quality, realism, and practical utility?*** Secondary questions explored the relationship between evaluator characteristics (such as industry experience and scenario planning familiarity) and scenario assessments.

The study employed a three-phase approach to develop and evaluate negotiation scenarios. In Phase 1, researchers developed synthetic initial negotiation dialogue and background materials, which underwent expert review by professionals in contracting, supply chain management, and academia who did not participate in the subsequent survey. This review process ensured the foundational materials were robust and relevant for scenario development.

Phase 2 focused on scenario generation, utilizing both AI systems and human authors from the National Contract Management Association (NCMA) Contract Leadership and Management Development Program (CLMDP). Two AI platforms (ClaudeOpus and GPT-4o) were used to generate scenarios, while five CLMDP members with varying levels of experience participated in human scenario development. The CLMDP participants included directors, contract managers, and senior administrators, with experience ranging from 10 to 21 years across government, commercial, defense, and non-profit sectors. Their educational backgrounds varied from bachelor's degrees to advanced graduate degrees, and two participants had prior scenario development experience.

From the five human-developed scenarios, researchers selected two for the comparative study based on time investment and scenario detail level. The selected scenarios (designated as Scenarios 2 and 4) represented the most detailed scenarios produced, taking 3 hours 17 minutes and 2 hours 32 minutes to develop, respectively. This selection process ensured that the AI-generated scenarios would be compared against the most robust human-authored scenarios, providing a meaningful benchmark for quality assessment.

Phase 3 comprised the evaluation of these scenarios. Professional conference participants were randomly assigned to evaluate either AI-generated or human-authored versions of the scenarios. Prior to evaluation, participants received primer coursework on negotiation styles, scenario planning, and if/then analysis to establish a baseline understanding of key concepts. Participants were not aware that the scenarios they were provided may have been generated by AI. There were no distinguishing features of the materials provided that alluded to the source creators of the content to ensure such potential biases were mitigated.

Methodology

Scenario Background and Development

The study utilized a complex negotiation scenario involving a high-stakes satellite receiver upgrade project between a satellite development firm (seller) and a government agency (buyer). This scenario was chosen for its multifaceted nature and representation of real-world negotiation complexity. The background materials provided to scenario developers included detailed information about a sole-source arrangement for upgrading three satellites with 12 new receivers, where both parties lacked alternative options and faced a two-month deadline for agreement.

The scenario incorporated multiple negotiation elements requiring consideration, including total project pricing, technical data deliverables, production efficiency targets, and profit structures. The background materials included comprehensive stakeholder perspectives, detailed historical context of the buyer–seller relationship, and specific constraints facing each party. Initial negotiation exchanges were provided to establish the tone and starting positions of both parties. The scenario was enriched with information about the broader context, including the fact that the satellites represented units 6-8 in a historical series of exchanges between the parties, adding depth to the relationship dynamics.

Both AI systems and human scenario developers from the NCMA CMLDP were tasked with generating additional dialogue exchanges and "if/then" planning statements based on their analysis of the scenario. This parallel approach allowed for direct comparison of how human and AI developers would extend and elaborate upon the initial scenario framework. Developers were provided with a framework of five negotiation styles (competing, avoiding, accommodating, collaborating, and compromising) and were asked to incorporate these styles into their scenarios in specific ways. For instance, developers were instructed to create scenarios where parties might take a collaborative approach to certain technical aspects while maintaining a more competitive stance on intellectual property matters. All developers were required to generate practical planning guidance and anticipate potential negotiation exchanges while deliberately incorporating these varied negotiation styles across different aspects of the discussion.

This scenario design allowed for the evaluation of how different scenario authors (human or AI) would approach complex, multi-variable negotiation planning. The scenario's structure enabled the development of detailed "if/then" planning approaches while maintaining real-world applicability through its incorporation of common negotiation elements such as pricing, intellectual property concerns, efficiency targets, and payment terms. The comprehensive background enabled developers to create realistic negotiation planning materials that accounted for multiple variables, relationship dynamics, and strategic considerations.

Measurement Development

The study employed a 7-point Likert agreement scale (1 = Strongly Disagree to 7 = Strongly Agree) across eight dimensions of scenario quality. For assessing perceived realism and real-world applicability, the study adapted elements from the Perceived Realism Scale developed by Cho et al. (2012). This established scale was chosen for its validated approach to measuring narrative realism, with modifications made to specifically address negotiation scenario contexts. The measurement instrument was organized into three primary categories, each containing specific items designed to capture different aspects of scenario quality:

1. Realism and Real-World Applicability (derived from Cho et al.'s (2012) Perceived Realism Scale):

    - "The events in the negotiation scenario portrayed possible real-life situations" (Q12)
    - "What happened to the people in the negotiation scenario is what happens to people in the real world" (Q13)
    - "The negotiation scenario was realistic" (Q17)

These items were adapted from the original scale's typological realism and narrative consistency dimensions, modified to focus specifically on negotiation contexts. The adaptation maintained the core focus on perceived authenticity while adjusting language to reflect professional negotiation situations rather than general narrative contexts.

2. Structural Quality:

    - "The negotiation scenario was coherent" (Q14)
    - "The negotiation scenario was consistent" (Q15)

These measures evaluated the internal logic and narrative flow of the scenarios. The items were designed to assess how well the various elements of each scenario worked together to create a cohesive training tool. While related to Cho et al.'s (2012) narrative consistency dimension, these items were specifically crafted to evaluate the structural elements necessary for effective negotiation training scenarios.

3. Practical Value:

    - "This negotiation scenario would be useful for preparing a person or team for a real world negotiation" (Q16)
    - "The negotiation scenario provided valuable information" (Q18)
    - "The negotiation scenario adds value in decision-making" (Q19)

These items were developed specifically for this study to assess the practical utility of the scenarios in professional development contexts. They were designed to evaluate both immediate training value and broader applicability to decision-making processes in negotiation contexts.

The scale development process included careful consideration of item wording to ensure relevance to the professional negotiation context while maintaining measurement integrity. The 7-point scale was chosen to provide sufficient granularity in responses while maintaining ease of use. The neutral midpoint (4 = Neither Agree nor Disagree) allowed respondents to express uncertainty or ambivalence, particularly important given the novel nature of AI-generated content.

Each category was designed to capture distinct but related aspects of scenario quality, enabling analysis of how AI-generated and human-authored scenarios might differ across these dimensions. The realism measures, adapted from Cho et al.'s (2012) validated scale, provided a theoretical foundation for assessing perceived authenticity, while the structural and practical value measures addressed specific requirements for negotiation training materials.

Data Collection

The study collected evaluation data from 36 professionals across government, aerospace, defense, and related sectors. These evaluators were distinct from the CLMDP members who participated in scenario development during Phase 2. Demographic data included education level, years of experience, industry sector, scenario planning training, and familiarity with various planning tools. Each participant evaluated one of four scenarios (two AI-generated, two human-authored), rating them across the eight quality dimensions.

Results

Scenario Development Efficiency

A notable finding emerged regarding the time investment required for scenario creation. The AI-generated scenarios were developed in 3–4.5 minutes, while human-authored scenarios required 2.5–3.25 hours. This substantial difference in development time did not necessarily correlate with proportional differences in quality ratings.

Quality Ratings and Exploratory Findings

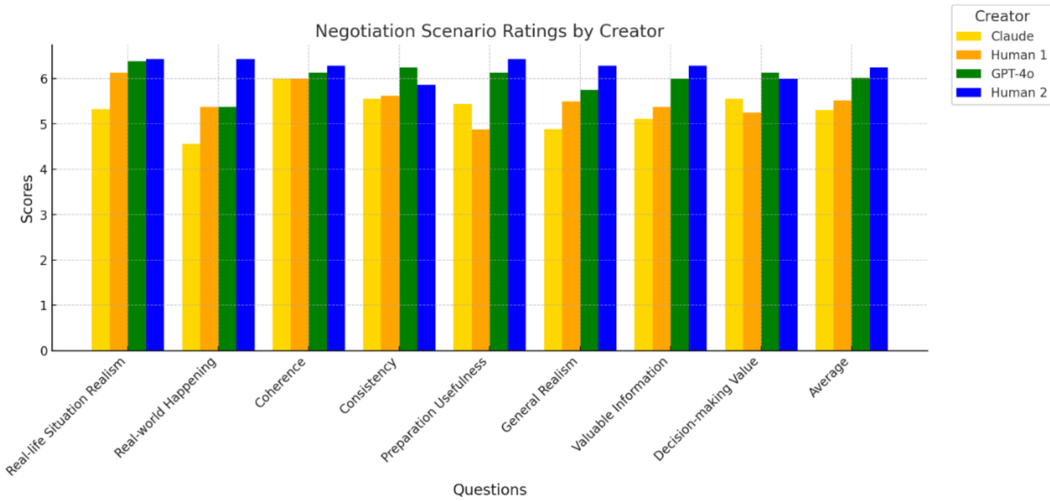


Figure 2: Negotiation Scenario Quality Ratings by Author Type

Table 1: Average Respondent Scores Per Scenario Item by Creator Source

Question/Item	Scenario 1, Claude	Scenario 2, Human 1	Scenario 3, GPT4o	Scenario 4, Human 2
Real-life Situation Realism	5.33	6.13	6.38	6.43
Real-world Happening	4.56	5.38	5.38	6.43
Coherence	6	6	6.13	6.29
Consistency	5.56	5.63	6.25	5.86
Preparation Usefulness	5.44	4.88	6.13	6.43
General Realism	4.89	5.5	5.75	6.29
Valuable Information	5.11	5.38	6	6.29
Decision-making Value	5.56	5.25	6.13	6
Average	5.31	5.52	6.02	6.25

The study revealed several intriguing patterns in scenario evaluation scores. While the sample size precludes definitive statistical conclusions, the exploratory findings suggest notable trends in perceived scenario quality across different dimensions.

The human-authored scenario created by an experienced professional (Scenario 4) achieved the highest overall average rating (6.25 on the 7-point scale), indicating strong agreement with positive quality attributes across all dimensions. This scenario particularly excelled in real-world applicability (6.43) and usefulness for preparation (6.43), suggesting that the author's professional experience may have contributed to creating highly practical training materials.

Notably, the AI-generated scenario using GPT-4o (Scenario 3) achieved the second-highest average rating (6.02), performing particularly well in coherence (6.13) and consistency (6.25). This strong performance is especially remarkable given the scenario's creation time of just 3 minutes, compared to 2.5 hours for the top-rated human-authored scenario. This finding, while preliminary, suggests potential for AI tools to generate quality training materials with unprecedented efficiency. These findings are consistent with previous research by Jensen and Cummins (2023), which showed that generative AI use in contracting improved negotiation speed and economic value, indicating real-time applicability beyond training and planning.

Analysis of Specific Quality Dimensions Revealed Interesting Patterns

Real-world applicability demonstrated the largest variance among quality dimensions (scores ranging from 4.56 to 6.43), with the greatest differentiation between AI-generated and human-authored scenarios. The lower scores for AI-generated scenarios in this dimension (4.56 for Scenario 1, 5.38 for Scenario 3) might reflect limitations in AI systems' ability to fully capture nuanced real-world dynamics, though the GPT-4o scenario still achieved relatively strong ratings.

Coherence showed remarkable consistency across all scenarios (scores ranging from 6.00 to 6.29), suggesting that both AI and human authors could create logically structured scenarios. This finding is particularly noteworthy for AI-generated content, as it indicates strong capability in maintaining narrative consistency even in complex negotiation scenarios.

Usefulness for preparation showed notable variation (scores ranging from 4.88 to 6.43), with an interesting pattern where scenario development time did not necessarily correlate with perceived utility. The longest development time (Scenario 2, 3 hours 17 minutes) received the lowest usefulness rating (4.88), while the AI-generated Scenario 3 (3 minutes) received a strong rating (6.13).

Decision-making value ratings (ranging from 5.25 to 6.13) suggested that both AI and human-authored scenarios could provide valuable strategic insights, though human-authored scenarios maintained a slight edge in this dimension.

Demographic Influences

Analysis of demographic data revealed several potential influences on scenario evaluations. Mid-career professionals (11–20 years of experience) generally provided higher ratings across scenarios. Industry alignment appeared to influence perceptions of realism and applicability, with scenarios closely matching the evaluator's industry background receiving higher ratings in these dimensions.

Interestingly, formal training in scenario planning did not consistently correlate with higher ratings, suggesting the scenarios' effectiveness transcended specialized training. However, familiarity with various scenario planning tools showed a positive correlation with higher appreciation of scenarios' informational and decision-making value.

Discussion

While the sample size (n=36) limits the statistical power of the findings and precludes definitive conclusions, several compelling patterns emerged that warrant further investigation. The exploratory results suggest that AI-generated scenarios can achieve quality ratings comparable to human-authored scenarios in multiple dimensions, particularly in areas such as coherence and informational value. The efficiency advantage of AI generation (3–4.5 minutes versus 2.5–3.25 hours) combined with strong quality ratings suggests significant potential for rapid development of diverse training materials.

The pattern of ratings across dimensions offers interesting insights into the relative strengths of AI and human scenario development. While human-authored scenarios maintained an edge in real-world applicability and practical utility, AI-generated scenarios demonstrated strong performance in structural elements such as coherence and consistency. This suggests that AI tools might be particularly valuable for quickly generating well-structured baseline scenarios that could then be refined with human expertise to enhance real-world relevance.

The success of the experienced human author's scenario (Scenario 4) highlights the continued value of expert knowledge in scenario development. However, the strong performance of the GPT-4o generated scenario (Scenario 3) suggests that AI tools might serve as effective supplements to human expertise, particularly when time constraints are significant.

Limitations

Several important limitations should be considered when interpreting these results:

The small sample size (n=36) limits statistical power and the ability to detect subtle differences between scenarios. This also constrains the generalizability of findings and increases sensitivity to individual responses. The distribution of participants across scenarios (ranging from 8–10 per scenario) further limits the ability to control for confounding variables and conduct meaningful subgroup analyses.

Research Design Efficacy

Despite these limitations, the study demonstrated the effectiveness of its three-phase research design for comparative analysis of AI and human-authored scenarios. The combination of expert review, controlled scenario development, and structured evaluation provides a robust framework for larger-scale investigations. The incorporation of demographic data collection enabled preliminary exploration of factors influencing scenario perceptions, though larger samples would be needed for definitive conclusions.

Future Research Directions

This exploratory study suggests several promising areas for future research:

1. Large-scale replication studies with sufficient sample sizes for statistical validation of preliminary findings.
2. Investigation of optimal human–AI collaboration methods in scenario development, potentially combining the efficiency of AI generation with human expert refinement.
3. Examination of how different AI models perform in generating scenarios for specific industries or negotiation contexts, with additional focus on the use of text-to-video generative models, such as OpenAI’s recently released Sora, to enhance realism.
4. Exploration of how various demographic and experience factors influence perceptions of AI-generated versus human-authored scenarios.
5. Development of standardized quality metrics for negotiation scenarios, building on the adapted narrative realism scale used in this study.
6. Investigation of the long-term effectiveness of AI-generated scenarios in actual negotiation training programs.

Practical Implications

The findings suggest potential for using generative AI to supplement traditional scenario development methods, particularly in situations requiring rapid development of diverse training materials. Building on recent work highlighting the importance of strategic adaptability in negotiation training (Heunis et al., 2024), AI-generated scenarios might offer a scalable approach to creating varied, adaptable training materials. The recent developments in immersive AI-generated video offer additional opportunities to explore realism using visual simulations and should be explored.

The comparable quality ratings between AI-generated and human-authored scenarios, particularly when considering the dramatic difference in development time, suggest potential for expanding access to high-quality negotiation training materials. This could be especially valuable for organizations lacking extensive expert resources for scenario development.

Conclusion

While preliminary in nature, this study provides valuable insights into the potential role of generative AI in negotiation scenario development. The research design demonstrated effectiveness for comparative analysis of AI and human-authored scenarios, establishing a foundation for larger-scale investigations. The dramatic efficiency advantage of AI generation, combined with promising quality ratings, suggests significant potential for expanding access to diverse, adaptable negotiation training materials. Future research with larger samples will be crucial for validating these initial findings and exploring optimal methods for integrating AI tools into negotiation training development.

Section 4: The Competencies We Need

The findings of this research underscore the need to treat negotiation as a recurring process rather than a discrete event. Successful outcomes require upfront planning for ongoing renegotiation and change, supported by frameworks that manage, rather than resist, evolution. This section explores the competencies organizations need to navigate modern contracting challenges and opportunities, aligning them with the NCMA Contract Management Body of Knowledge (CMBOK) framework.

Competencies Mapped to CMBOK

1. Pre-Award (Acquisition & Sales) Competencies

Strategic negotiation planning is essential for modern contract managers, requiring an understanding that goes beyond traditional price and liability discussions. Negotiators must adopt a holistic approach that incorporates operational risks, coordination mechanisms, and performance-driven contract terms. The ability to conduct scenario-based planning is critical, leveraging tools like generative AI to simulate potential negotiation dynamics and outcomes. Additionally, expertise in market and risk analysis ensures informed negotiation strategies that align with both short-term and long-term business objectives.

Stakeholder collaboration and engagement are also pivotal during the pre-award phase. Contract managers must navigate complex relationships between large enterprises and SMEs, working toward balanced agreements that foster long-term partnerships rather than short-term wins. Strong communication and negotiation skills are required to advocate for contract simplification strategies, ensuring that terms are clear, fair, and effective. In this context, professionals must also manage multi-stakeholder interests, ensuring that contractual goals align with broader business outcomes. Modern contract managers must function as integrationists. As commercial and operational complexity increases, so too does the need to reconcile competing interests across legal, technical, financial, and mission-focused stakeholders. This shift demands a move from rules-based execution to judgment-based navigation. A guiding principle in this transformation might well be: "From rules-based to judgment-based."

2. Award (Formation & Execution) Competencies

Once a contract is awarded, performance-focused contract drafting becomes crucial. Organizations need professionals who can develop terms that emphasize clear performance expectations, shared milestones, and collaborative problem-solving mechanisms. The inclusion of well-crafted contingency clauses helps account for operational uncertainties, allowing for flexibility and resilience in contract execution. Moreover, structuring performance-based incentives fosters alignment between parties, ensuring that all stakeholders remain committed to achieving mutual success while minimizing disputes.

Legal and regulatory acumen plays a significant role during contract formation and execution. A thorough understanding of global legal frameworks is necessary, especially when dealing with international negotiations that require compliance with various Common Law and Civil Law traditions. Additionally, with the rise of AI-assisted contract creation and automated contract analysis, contract managers must remain vigilant about the legal implications of these technologies. Compliance with industry-specific regulations and standards further adds to the complexity, necessitating deep regulatory expertise.

While legal and regulatory expertise is foundational, it is no longer sufficient. Success now requires integration of market intelligence, finance, and economic insights. Knowing the rules is necessary, but understanding the market, supplier drivers, and opportunity costs is what differentiates effective contract outcomes. Ironically, these legal and compliance elements are the most likely to be automated by AI in the coming years.

3. Post-Award (Performance & Closeout) Competencies

Risk management and dispute resolution are essential competencies for ensuring contract success post-award. Contract managers must be adept at identifying and mitigating operational risks that could impact contract execution. A proactive approach to dispute resolution, including expertise in mediation and arbitration techniques, can help resolve conflicts before they escalate. Effective contract monitoring is also critical, enabling professionals to track performance and make necessary adjustments to maintain value creation and compliance.

Technology plays an increasingly important role in contract life-cycle management, making proficiency in contract life-cycle management (CLM) tools valuable competency. These tools allow contract professionals to track performance metrics, assess risk exposure, and ensure regulatory compliance. Furthermore, AI-powered contract monitoring and predictive analytics can help identify potential issues before they become major problems, enabling organizations to maintain efficiency and transparency in their contract management processes.

Emerging Competencies Beyond CMBOK

While the CMBOK provides a strong foundation for contract management competencies, emerging trends necessitate additional skill sets beyond traditional models. One such area is AI-augmented negotiation and decision-making. Contract professionals must develop the ability to critically assess and integrate AI-generated insights into their negotiation strategies. Understanding the strengths and limitations of AI in contract scenario development and risk analysis is crucial for leveraging technology effectively without compromising human judgment.

Behavioral economics and negotiation psychology are also becoming increasingly relevant. Negotiators must be aware of behavioral biases that can impact contract structuring and negotiation dynamics. Applying principles from behavioral economics enables professionals to craft agreements that drive long-term success rather than short-term gains.

Ethical contracting and ESG (Environmental, Social, and Governance) integration represent another critical area of emerging competencies. Organizations must align contract structures with ESG goals, ensuring that business practices remain sustainable and ethical. Contracts should promote fair labor conditions, responsible sourcing, and sustainability initiatives, requiring contract professionals to incorporate ESG considerations into their negotiation and contract management processes.

Professional Development and Certification Pathways

To develop these competencies, professionals can pursue specialized training and certifications offered by NCMA and WorldCC. Structured training provides a foundation, a license to practice, but cannot keep pace with change alone. AI will increasingly deliver baseline technical knowledge on demand. What will differentiate professionals is emotional and adaptability intelligence (EQ and AQ), which drive sustained value through relationship and outcome management.

The NCMA Certified Professional Contract Manager (CPCM) program provides advanced knowledge in contracting, including negotiation strategy and risk management. The Certified Federal Contract Manager (CFCM) certification is particularly useful for professionals working in government contracting, emphasizing regulatory compliance and legal frameworks. Similarly, the Certified Commercial Contract Manager (CCCM) certification focuses on best practices for private-sector negotiations.

WorldCC also offers several programs designed to enhance contract management expertise. The Contract and Commercial Management (CCM) Certification equips professionals with the skills needed to align contracts with business strategy while improving negotiation effectiveness. The Supplier Relationship Management (SRM) Certification addresses collaboration and risk-sharing strategies essential for contract performance. Additionally, WorldCC provides Negotiation Masterclass Programs that deliver practical training in advanced negotiation skills, including scenario-based planning and behavioral negotiation techniques.

Conclusion

As organizations navigate complex business landscapes, developing competencies that align contract negotiation practices with performance outcomes is essential. The findings from this research highlight the need for strategic negotiation planning, performance-focused contracting, effective risk management, and the integration of AI and behavioral economics into contract processes. Through professional development opportunities offered by NCMA and WorldCC, contract professionals can enhance their ability to craft agreements that drive value, mitigate risks, and foster long-term, sustainable business relationships. Ongoing research ensures that practitioners receive the up-to-date knowledge and methods they need to navigate constant change and uncertainty. The joint venture between WorldCC and NCMA is intended to become the platform that delivers these evolving insights, helping the profession anticipate rather than follow change.

Research and Authorship Disclaimer

This report was produced by the Commerce and Contract Management Institute, a collaborative initiative between World Commerce & Contracting and NCMA. The research employed a hybrid methodology combining human expertise with AI capabilities, with all sources and findings independently validated by human researchers. The report was co-authored through collaboration between human experts and AI tools, with final human editorial oversight.

© 2024 Commerce and Contract Management Institute A collaboration between World Commerce & Contracting and the National Contract Management Association

References

Amoah, C., & Nkosazana, H. (2023). Effective management strategies for construction contract disputes. *International Journal of Building Pathology and Adaptation*, *41*(6), 70–84. https://doi.org/10.1108/IJBPA-01-2022-0004

Bauman, C., Weber, P., & Cummins, T. (2019). *Optimism vs. prevention*. World Commerce & Contracting. https://www.worldcc.com/Portals/IACCM/resources/files/10916_iaccm-optimismprevention.pdf

Bazerman, M. H. (2025). What people still get wrong about negotiations. *Harvard Business Review*, *103*(1–2). https://hbr.org/2025/01/what-people-still-get-wrong-about-negotiations

Carr, A. S., & Pearson, J. N. (2002). The impact of purchasing and supplier involvement on strategic purchasing and its impact on firm's performance. *International Journal of Operations & Production Management*, *22*(9), 1032–1053. https://doi.org/10.1108/01443570210440528

Cho, H., Shen, L., & Wilson, K. (2012). Perceived realism: Dimensions and roles in narrative persuasion. *Communication Research*, *41*(6), 828–851. https://doi.org/10.1177/0093650212450585

Cummins, T., & Finkenstadt, D. J. (2024, December 18). Contract negotiations should be collaborative, not adversarial. *Harvard Business Review*. https://hbr.org

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, *14*(1), 57–74.

Finkenstadt, D. J., Eapen, T. T., Sotiriadis, J., & Guinto, P. (2023, November 30). Use GenAI to improve scenario planning. *Harvard Business Review*. https://hbr.org

Finkenstadt, D. J., Sotiriadis, J., Guinto, P., & Eapen, T. T. (2024, January 22). Contingency scenario planning using generative AI. *California Management Review Insights*.

Heunis, H., Pulles, N. J., Giebels, E., Kollöffel, B., & Sigurdardottir, A. G. (2024). Strategic adaptability negotiation training in purchasing and supply management: A multi-method instructional approach. *Journal of Purchasing and Supply Management*, Article 100968. https://doi.org/10.1016/j.pursup.2024.100968

Jensen, R., & Cummins, T. (2023). Realizing value through AI-augmented contracting. *Journal of Strategic Contracting*.

Kirgis, F. L. (2019). The coercive use of leverage in negotiation. *Harvard Negotiation Law Review*, *19*, 69–123.

MacMahon, J. A. (2018). Conflict and contract law. *London School of Economics Law, Society and Economy Working Papers, 19/2018*.

Ramsay, J. (2007). [Review of the book *Great negotiators: How the most successful business negotiators think and behave*, by T. Beasor]. *Journal of Purchasing and Supply Management*, *13*(1), 84–85.

Saorín-Iborra, M. C., & Cubillo, G. (2019). Supplier behavior and its impact on customer satisfaction: A new characterization of negotiation behavior. *Journal of Purchasing and Supply Management*, *25*(1), 53–68. https://doi.org/10.1016/j.pursup.2018.03.002

World Commerce & Contracting. (2016). *Unpacking relational contracting*. https://www.worldcc.com/Portals/IACCM/resources/files/9487_unpacking-relational-contractingv19-oct-11-2016.pdf

World Commerce & Contracting. (2022). *Most negotiated terms 2022*. https://www.worldcc.com/Portals/IACCM/Resources/11463_0_Most-Negotiated-Terms-2022.pdf

From R&D to Readiness: Navigating Technology Transitions with the Naval Power and Energy Systems Technology Development Roadmap

Henry L. Jones III—is a lead model-based systems engineer at Herren Associates, Inc., focusing on advanced naval power and energy systems. He began his career specializing in aircraft carrier design before transitioning into increasingly complex roles in marine engineering. He holds a bachelor's degree in mechanical engineering and is currently pursuing a master's degree in systems engineering from Naval Postgraduate School.

Jeffrey M. Voth—is the President of Herren Associates, an engineering and technology firm. He has written extensively on the strategic, economic, and business implications of technology investment across the aerospace and defense sector. Voth holds an MBA from Georgetown University and an MSc from the University of Oxford.

CDR Victor Sorrentino, USN (Ret.)—served over two decades as a Surface Warfare Officer, holding many roles including Deputy Director, Operational Energy for the Secretary of the Navy Staff. He is now Director of Energy Programs at Herren Associates, supporting innovation in naval power and energy systems.

Abstract

Power and energy will remain fundamental to maintaining the U.S. Navy's decisive maritime advantage, enabling advanced sensors, electronic warfare, directed-energy weapons, resilient power and propulsion systems, and operationally dominant integrated combat system capabilities. In an increasingly competitive and rapidly evolving threat environment, the Navy will chart a course to strengthen today's Fleet and accelerate capability delivery for next-generation surface ships and systems. The Naval Power and Energy Systems Technology Development Roadmap (NPES TDR) should serve as a strategic mechanism to synchronize research and development (R&D) across the acquisition community, ensuring that emerging capabilities will mature in lock step with the operational requirements.

By applying insights from established roadmapping theory, this paper demonstrates how the next NPES TDR should guide gap analyses, stakeholder collaboration, and iterative technology readiness evaluations. Through an illustrative case study, a laser weapon system, part of the Navy's solid-state laser technology maturation effort, it explains how the roadmap could streamline technology transition timelines, minimize risk, and align with complex budget cycles. The analysis also addresses enduring challenges, such as bridging the extended expected service life of naval platforms. Concluding with targeted recommendation—such as conducting regular roadmap updates, adopting scenario-based planning, and deepening public-private partnerships—this paper asserts that technology roadmaps such as the NPES TDR are essential to increasing lethality, accelerating warfighting capabilities, and improving readiness amidst fast-changing technical and strategic conditions.

Keywords: defense industrial base, requirements management, technology transition, adaptive acquisition framework/rapid acquisition, engineering and technical management

Introduction

"The versatility of our surface force deters adversaries globally and enables rapid, coordinated responses to emerging threats. Our ships must be prepared to engage the full spectrum of threats, from existing capabilities to emerging ballistic and hypersonic missiles."

Admiral James W. Kilby, Vice Chief of Naval Operations
Statement on the Readiness of the U.S. Navy before the Senate Armed Services Committee, Subcommittee on Readiness and Management Support, March 12, 2025

In an increasingly complex and contested global environment (Kilby, 2025), power and energy systems will serve as a cornerstone of the Navy's combat effectiveness. These systems will provide the power necessary for lethal effects, accelerate warfighting capabilities, and sustain Fleet readiness. However, despite investments in advanced electrical power systems since the early 1990s, the Navy will continue to face challenges integrating advanced power electronic equipment—such as high-current semiconductor devices—into current and future ship systems (Doerry & Amy, 2024). Modern power and energy solutions will demand rigorous attention throughout the development life cycle to preempt costly redesigns if adequate size, weight, and power (SWaP) margins are not allocated for the platform's expected service life (Doerry & Amy, 2020; IEEE, 2023). This complexity—combined with uncertainties in linking early-stage research to real-world fleet adoption—risks perpetuating the valleys of death (see, for example, Figure 1) experienced where promising technologies fail to transition at various points of development before transitioning to a formal program of record (Letts, 2024). Technology roadmaps, such as the 2019 *Naval Power and Energy Systems Technology Development Roadmap* (NAVSEA, 2019), will require continuous updates and function as a "living document." Through iterative updates, the Fleet should maintain access to robust, scalable power solutions aligned with mission capabilities—from surface warfare and conventional strike to integrated air and missile defense to assert dominance and project power.

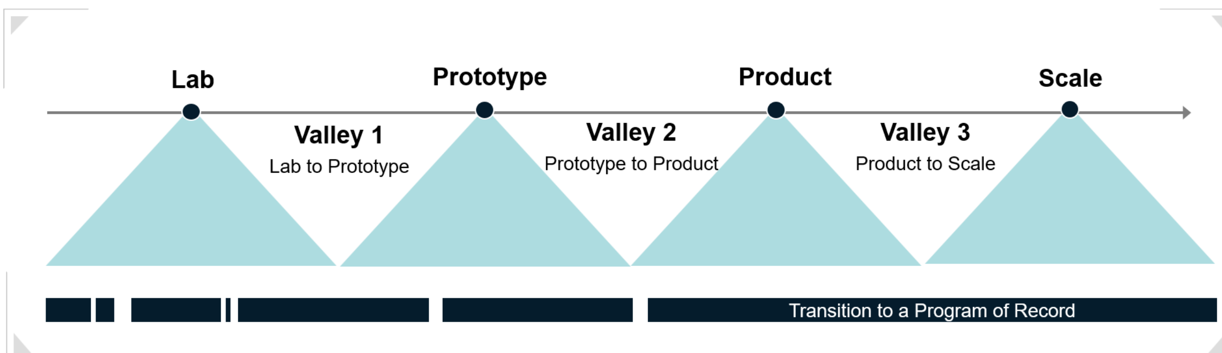


Figure 1. Crossing the U.S. Department of Defense Valleys of Death (Adapted from McEntush and Hay, 2025)

To address these challenges, updates to the Naval Power and Energy Systems Technology Development Roadmap (NPES TDR) should coordinate closely with government laboratories, technical authorities, program offices, industry partners, Fleet stakeholders, and resource sponsors, aligning technology transition efforts with the Navy's longer-range strategy. By embedding risk reviews, readiness thresholds, and cross-functional collaboration, the NPES TDR will incorporate a mission-led capability perspective, where iterative feedback loops replace linear innovation (Moore, 2024). This paper examines how the NPES TDR could further streamline the path from early-stage research to Fleet adoption of vital power and energy systems. Specifically, it shows how roadmapping principles can help structure phased testing, optimize resource allocation, and adjust technical priorities when necessary (Phaal et al., 2024).

This paper begins with an examination of technology roadmapping, defined by Kerr and Phaal (2022, p. 13) as "the application of a temporal–spatial structured lens" to support research and development decisions by identifying critical technologies and gaps (Garcia & Bray, 1997). The Roadmap as a Strategic Planning Tool section positions the NPES TDR within broader strategic planning frameworks used within the naval acquisition community, underscoring the theoretical underpinnings of the roadmap approach. The NPES TDR in Action: Key Processes section then explores the roadmap's operational processes—such as data gathering, industry engagement, gap analysis, and cross-functional collaboration—and outlines how they will keep the NPES TDR adaptive and outcome-focused. The Research to Readiness: Key Outcomes and Lessons Learned section presents tangible outcomes and lessons learned, including the successful demonstration of directed-energy weapons enabled by energy storage technologies progressing from early laboratory research to real-world applications. Discussion: Challenges and the Way Ahead addresses ongoing challenges, such as the extended expected service lives of surface ships, budget constraints, and the integration of emergent digital engineering practices. Finally, the Conclusion offers forward-looking recommendations to ensure the NPES TDR remains a living document, retaining practical relevance amid evolving technological and operational demands.

The Roadmap as a Strategic Planning Tool

Definition and Scope of Technology Roadmaps

Technology roadmaps will become integrative planning tools that coordinate upcoming ship and system-level milestones, resource allocations, and mission scenarios within a unified, time-phased framework (Phaal et al., 2021). By fusing “technology push”—the outputs of laboratories, industry, and academia—with “requirements pull”—the operational needs of the surface navy—this roadmap will offer a holistic strategy for directing innovation. In contrast to linear Gantt charts, they will incorporate iterative readiness gates, stakeholder engagement points, and forward-looking force development objectives (Garcia & Bray, 1997).

For the U.S. Navy, these attributes will prove vital. Complex warfighting capabilities require synchronization across propulsion, ship-service power generation systems configuration, any available energy storage systems (Araujo et al., 2024; Doerry & Amy, 2017; McCoy, 2025). A well-structured roadmap will help senior decision-makers envision how one emerging area of research and development—such as a modular universal converter building blocks (Lawson et al., 2024)—might intersect with broader modernization initiatives or doctrinal shifts. By combining near-term readiness checkpoints with longer-range objectives, technology roadmaps could accelerate capability maturation and better align R&D (Kerr, 2023).

The Value Proposition of NPES TDR

The NPES TDR will represent a mission capability-led adaptation of general roadmapping principles (Figure 2). By placing naval power and energy solutions within acquisition timelines and ship modernization availabilities, the NPES TDR will align R&D milestones with operational readiness (Markle et al., 2021). Moreover, it will integrate structured risk assessments and validation trials, ensuring no technology proceeds into Fleet integration without targeted readiness reviews. This sequential approach—comparable to standard Technology Readiness Levels (TRLs)—is expected to promote transparency, allowing acquisition milestone decision authorities to authorize procurement only after a system meets prescribed maturity thresholds (Olechowski, 2020).

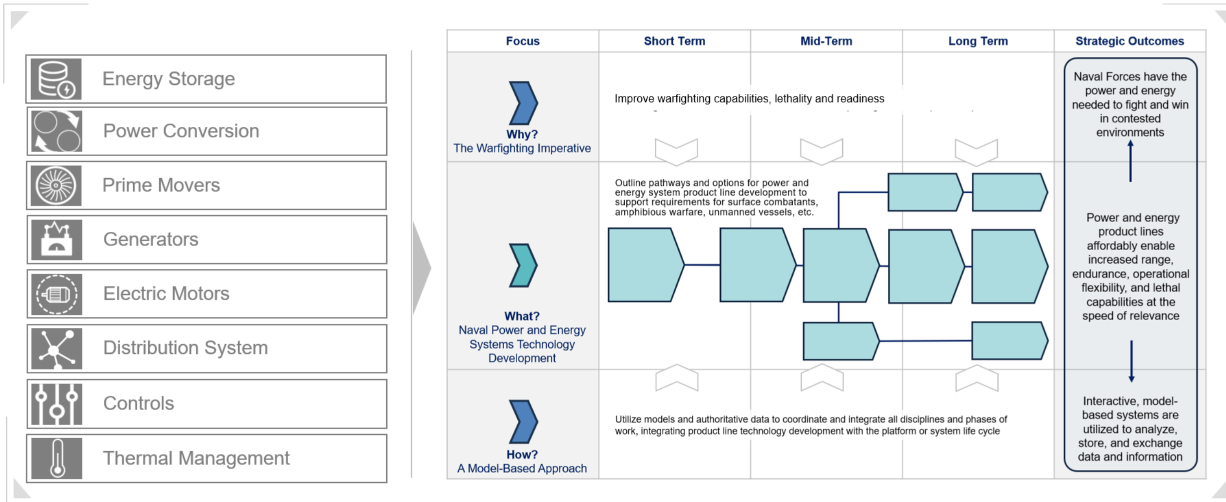


Figure 2: Governing Technology Roadmapping Framework (Adapted from Phaal, 2024)

Another key feature of NPES TDR will be its capacity for recalibration. Although the initial roadmap will define near-, mid-, and long-term objectives, it will incorporate feedback loops that respond swiftly to newly emerging technologies or threats (Chakraborty et al., 2022). In so doing, NPES TDR will replace static planning documents and stand-alone technical reports and act instead as a dynamic guide that balances established milestones with agile responsiveness to evolving needs (Ding & Ferràs Hernández, 2023).

Aligning Research with Acquisition Cycles

Historically, misalignment between R&D progress and formal acquisition steps within the Department of Defense (DoD) has often stemmed from organizational siloes (Kotila et al., 2023). Naturally occurring due to distinct functional areas, specialized expertise, and separate budgetary streams, siloes create communication gaps, narrative drift, and divergent timelines (Jeske & Olson, 2024). As a result, promising technologies developed in research laboratories frequently face delays or fail to transition effectively into acquisition programs, leading to underutilized capabilities and diminished operational advantage (Wong et al., 2022).

The NPES TDR aims explicitly to bridge these siloes by embedding clearly defined readiness-level checkpoints within the strategic roadmap. By stipulating specific maturity benchmarks (Ma, 2021), such as achieving TRL 6 or higher—which indicates successful demonstration of power and energy systems/subsystems in relevant operational environments—the roadmap systematically aligns technological advancements with formal acquisition processes. This alignment ensures that once these systems reach a designated readiness threshold, they become candidates for immediate consideration within funded programs of record (Stotts et al., 2010), specifically outlined in the Future Years Defense Program (FYDP).

Supported by this structured approach, the NPES TDR facilitates better synchronization between technology developers, program managers, and acquisition officials. Consequently, it reduces the historical gap between cutting-edge laboratory developments and tangible fleet capabilities (Tuinstra, 2022). The strategic, iterative decision gates built into the roadmap provide a mechanism for continuous evaluation and refinement, further enhancing communication across departmental boundaries (Cilli, 2015). Ultimately, the structured, strategic roadmapping methodology inherent to the NPES TDR—characterized by rigorous, proactive assessments and timely interventions—promotes more efficient technology adoption, optimizes resource allocation, and enhances overall defense readiness.

NPES TDR in Action: Key Processes

The NPES TDR should serve as a continually updated strategic plan, aligning power and energy system development with new and emerging warfighting requirements. It should be structured to provide a time-phased trajectory (2025–2035+) for the evolution critical power and energy systems, with executive steering group governance—consisting of SYSCOM stakeholders, Program Executive Offices (PEOs), Resource Sponsor, Fleet representatives, and all associated technical authorities—reviewing it regularly. This governance approach would integrate analytical foresight (Garcia & Bray, 1997; Hussain et al., 2017; Phaal et al., 2021) with acquisition imperatives, ensuring that the NPES TDR guides R&D activities within the Navy, industry, and academia. Four interrelated processes: (1) data gathering and requirements analysis, (2) industry engagement, (3) gap analysis and prioritization, and (4) cross-functional collaboration, will keep the NPES TDR relevant and flexible (Kerr & Phaal, 2021).

Data Gathering and Requirements Analysis

Data gathering and requirements analysis forms the foundation of the NPES TDR, aligning strategic objectives with technological feasibility. First, top-level directives (e.g., the National Security Strategy [NSS], National Defense Strategy [NDS], CNO's Navigation Plan [NAVPLAN]) form a strategic framework for U.S. national security, with the NSS outlining broad national security goals (Anderson & Karambelas, 2024), the NDS detailing how the DoD will contribute to those goals (Harman et al., 2024), and the NAVPLAN focusing on the Navy's role in achieving those objectives (Ullman, 2024). Next, fleet force structure reviews identify platforms and their operational profiles. Simultaneously, PEOs, SYSCOM technical authorities, Navy Surface Warfare Center (NSWC), and the broader Naval Research Enterprise will supply technical data, including power margins and load growth forecasts.

Mission-driven scenario modeling and digital engineering approaches (Ames et al., 2024; Voth & Sturtevant, 2022) should reveal challenges to future power distribution or availability for specific platforms (Figure 3). Upon validation, these shortfalls can be identified as potential capability gaps. Data from naval technical authorities, subject matter experts, RFIs and market surveys will refine assumptions about technology readiness, enabling near-real-time revisions of performance targets. This adaptive roadmapping approach (Phaal, 2024) will incorporate diverse sources of data, preventing NPES TDR from stagnating or becoming outdated.

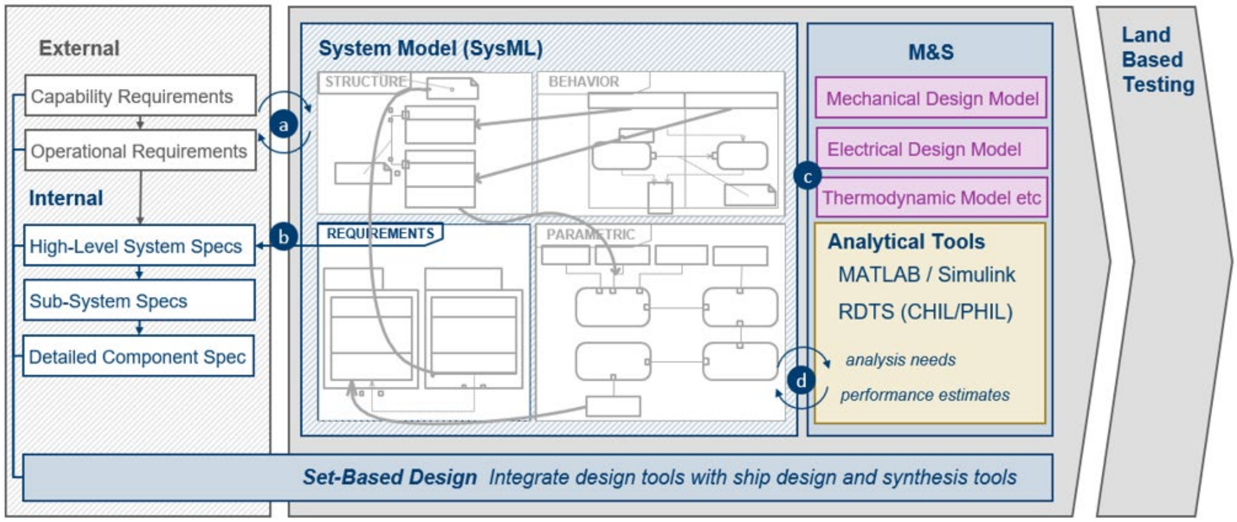


Figure 3. Cut Through Complexity with a Formalized Application of Modeling to Support System Requirements, Design, Analysis, Verification, and Validation ctivities for all Naval Power and Energy Systems (NPES)

Requests for Information and Industry Engagement

NPES TDR will rely on robust collaboration with industry, acknowledging that naval power and energy innovations frequently emerge from commercial and academic research. Regular requests for information (RFIs) and industry days will allow the Navy to identify technology maturation and research breakthroughs in key areas (e.g., prime movers, power electronics, battery systems). These exchanges will help shape both Navy requirements and future industry investment considerations. Purpose-driven technical meetings will delve into specific technologies tied to roadmap milestones (e.g., advanced wide-bandgap semiconductors). The roadmap will also track developments from Allied partners through existing channels (e.g., government-to-government agreement that provides official mechanisms for the exchange of research and development information). By actively incorporating commercial and global expertise, the NPES TDR will remain relevant and leverage the broader marketplace to emphasize solutions that meet short- and long-term naval requirements.

Gap Analysis and Prioritization

Gap analysis will be central to the NPES TDR. Future demands—such as pulsed power for directed-energy weapons—will be compared with the limitations of current shipboard systems, highlighting challenges in capacity, endurance, or speed of power delivery. These gaps will be quantified, and their impacts and time horizons will be delineated. The roadmap will use urgency, strategic value, and feasibility as prioritization criteria. Key needs often include the need for hybrid power systems, integrated power system technology architectures or modular energy storage. Once adopted in the roadmap, these priorities inform broader R&D goals and acquisition strategies. By delineating near-, mid-, and long-term objectives, the NPES TDR will ensure that emerging technologies progress methodically through the Planning, Programming, Budgeting, and Execution (PPBE) process.

Stakeholder Collaboration

Cross-functional collaboration will complete the process. Through formal governance forums, SYSCOMs, PEOs, Warfare Centers, the Naval Research Enterprise, and the Fleet will merge progress updates, re-sequence milestones as needed, and integrate emergent insights (Kerr & Phaal, 2022). If new operational data reveals a capability shortfall—like potential reliability issues under stressing combat loads—teams will be positioned to adjust the roadmap's emphasis. This method will coordinate technology investment, ship design, and acquisition timeframes. For example, once the roadmap highlights a key technology, relevant programs will incorporate it into prototype development. Ongoing feedback from technical, operational, and program personnel promote transparency and sustain a capability-focused culture. This collaborative structure will maintain the NPES TDR's "living" quality, allowing the Navy to introduce advanced power and energy capabilities when they are needed to support warfighting requirements.

Research to Readiness: Key Outcomes and Lessons Learned

Framework for Technology Transition

By systematically applying recognized roadmapping frameworks (Garcia & Bray, 1997), updates to the NPES TDR can continue to accelerate the transition of technologies onto naval platforms. Early identification of critical enablers—such as integrated energy storage and power controls—and well-coordinated efforts across government, industry, and academia have produced concrete advances, including at-sea demonstrations of directed energy weapons. These outcomes confirm the roadmap's effectiveness in forecasting and driving technology maturation under operational constraints. Equally important, the NPES TDR has reinforced that a roadmap must be adaptive (Kerr & Phaal, 2022). Periodic revisions will accommodate unexpected technologies or changes in mission need, maintaining strategic coherence and ensuring tangible results.

Identification of Critical Processes

Three intertwined processes will be crucial for bridging the traditional "valley of death" (Moore, 2024). First, comprehensive data analysis and modeling will yield more accurate projections and power load profiles. Second, multi-stakeholder gap evaluations will clarify priorities for bridging technology transition challenges. Third, iterative readiness reviews will align transitional technologies with established acquisition checkpoints and decision gates. This synergy will be important to mature advanced energy storage systems, which could progress from laboratory bench tests to land-based demonstrations and eventually to system-level integrations (Markle et al., 2021). By explicitly tracking each enabler, the NPES TDR will account for technical, organizational, and workforce factors to expedite transitions while minimizing risk.

Collaboration Models for Sustained Innovation

Implementing the NPES TDR will involve extensive cross-sector engagement. Public–private partnerships will leverage industry's expertise in power electronics, energy storage, and advanced controls, with the Navy providing robust tactical testbeds and clear operational requirements. Research collaborations, such as the Navy's partnership with the Electric Ship Research and Development Consortium, unites the combined programs and resources of leading electric power research institutions, including Florida State University's Center for Advanced Power Systems (FSU CAPS) and the University of Texas at Arlington's Pulsed Power and Energy Laboratory (UT Arlington PPEL), to advance near to mid-term electric ship concepts.

Inter-agency and joint service efforts will also broaden this ecosystem, standardizing best practices and accelerating lessons-learned exchanges across other high-power platforms. Fleet participation from the outset will anchor technology evolution in real operational experiences, thereby shaping design refinements and fostering user acceptance. By integrating these distinct collaboration paths, the Navy will build a resilient innovation network extending beyond individual programs and accelerating the pace of technology adoption.

Case Study: Directed Energy Weapons

A significant example of the 2019 NPES TDR's research-to-readiness approach is the Navy's deployment of shipboard laser weapons. Identified in the roadmap as a transformative capability, Directed Energy Weapons (DEW) require robust power generation, energy storage, and thermal controls (see, for example, highly stochastic loads provided in Figure 4). Although the Office of Naval Research (ONR) spearheaded the overall technology maturation effort, the NPES TDR synchronized energy storage technology development and defined the appropriate testing venues.

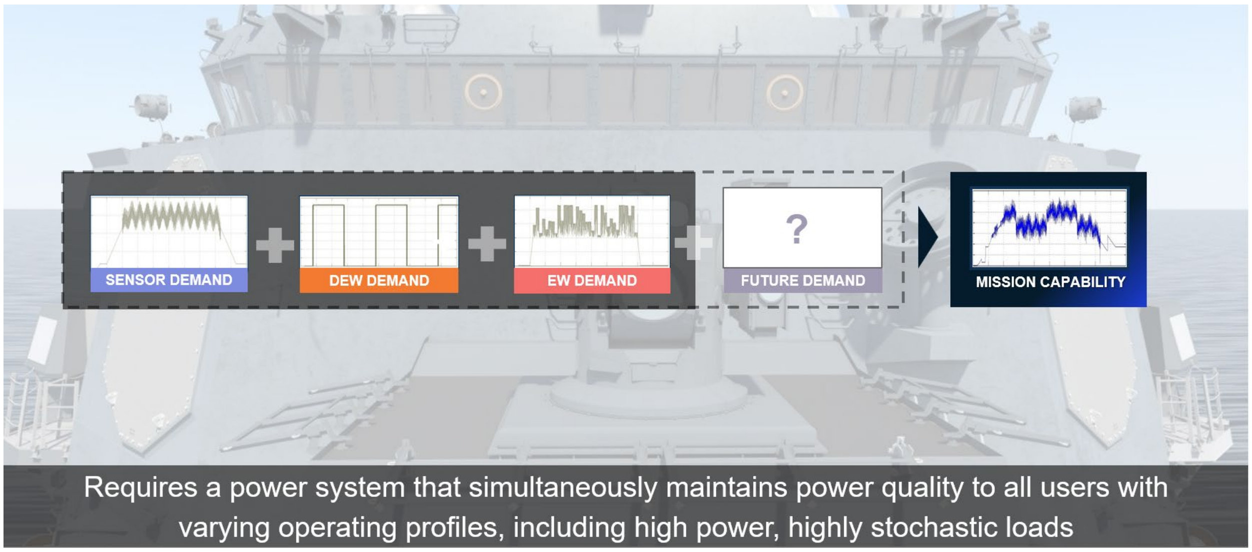


Figure 4. Rapidly Integrating Naval Power and Energy Systems to Enable Emerging Warfighting Capabilities

An indirect descendant of these developments is the AN/SEQ-3 LaWS: a solid-state laser system with variable power, designed specifically to combat unmanned aerial vehicles (UAVs) and small maritime threats (Bernatskyi et al., 2024). Initially installed on the U.S. Navy destroyer USS *Ponce* in 2014, LaWS successfully demonstrated operational effectiveness during annual testing (Chandler, 2014; LaGrone, 2014). Building upon the LaWS experience, USS *Portland* (LPD 27) served as the most recent demonstration of the Solid-State Laser Technology Maturation (SSL-TM) system, integrating a more powerful 150 kW solid-state laser coupled with appropriate pulsed-power energy storage system. The SSL-TM system validated its operational feasibility through successful at-sea tests in 2020, during which it disabled a UAV, marking a critical step toward the integration of directed energy systems across the Fleet (5th Fleet Public Affairs, NAVCENT, 2021).

Subsequent efforts shaped by these demonstrations include the High Energy Laser with Integrated Optical-Dazzler and Surveillance (HELIOS) currently installed on Arleigh Burke-class destroyer USS *Preble* (DDG 88), illustrating how NPES TDR–guided demonstrations feed into strategic acquisition decisions (Johnson, 2025). Additionally, other concurrent laser system developments within the Navy Laser Family of Systems (O'Rourke, 2022), such as the Optical Dazzling Interdictor, Navy (ODIN)—are leveraging lessons learned from SSL-TM's integration and operational employment (O'Rourke, 2022; O'Rourke, 2024).

The staged "crawl–walk–run" progression—moving methodically from lower-power prototypes, such as LaWS on USS *Ponce*, to higher-power and more advanced systems like SSL-TM—effectively manages stakeholder expectations and refines technical solutions. Early Fleet input on operational employment, energy storage requirements for pulsed loads and additional cooling requirements helps mitigate risk. Ultimately, the successful demonstration of directed-energy weapon capabilities underscores the NPES TDR's increasingly important role in guiding complex systems effectively from laboratory concept to the Fleet.

Discussion: Challenges and the Way Ahead

Risk Mitigation Across Life Cycles

The Navy will need to hold periodic NPES TDR reviews aligned with new construction and major ship modernization availabilities to address the challenges between a 35+ year surface ship Expected Service Life (ESL) and more rapid technological and warfighting capability evolution. Embedding digital engineering and distributed test environments early will help validate systems, well before shipboard integration occurs.

Adapting to Technological Shifts

Scenario-based planning and routine market assessments should identify potential breakthroughs—such as innovative supercapacitors or newer battery chemistries—and shift resources accordingly. Existing RFIs and industry workshops will feed into NPES TDR updates, ensuring that possible technologies receive near- and mid-term evaluations without unsettling acquisition timelines.

Enhanced Collaboration and Funding Alignment

To realize the full potential of the NPES TDR, consistent funding and cross-program cooperation will be vital, especially when large-scale integrated power solutions transcend conventional boundaries. Close and consistent coordination with broader DoD initiatives—like directed energy weapons—may augment the NPES TDR's influence. Strengthened public-private partnerships, including those with FSU CAPS and UT Arlington's PPEL will only further accelerate prototype validation for emerging capabilities. Sustaining financial support often proves difficult across multiple budget cycles. Balancing near-term achievements with longer-term research will be key to align efforts across portfolios in similar mission-focused areas. The roadmap's ability to more effectively linking platform and new warfighting capability schedules with key technology power and energy system developments should also help mitigate potential funding shortfalls.

The Roadmap as a Continuous Learning Ecosystem

Maintaining the NPES TDR as a "living" roadmap will require proactive data collection from fleet demonstrations, wargaming, and concurrent R&D projects. Annual or biennial workshops at major milestones could serve to blend current operational findings with industry forecasts, reinforcing the roadmap's adaptive nature. This iterative structure would enable the Navy to continuously refine power and energy priorities in alignment with real-world operational demands.

Conclusion

The updated NPES TDR should serve to showcase how a systematically constructed roadmap can continue to help guide naval power and energy systems from R&D to Fleet operations. By encouraging collaboration among government and industry stakeholders and embedding iterative readiness reviews, NPES TDR updates will facilitate early risk mitigation and help optimize resource prioritization. Ultimately, these processes will ensure high-impact technologies achieve timely integration into acquisition pipelines, essential to increasing lethality, accelerating warfighting capabilities, and improving readiness.

References

Ames, R., Doerry, N., Koerner, M., & Parsons, M. (2024). An overview of digital engineering methods for platform integration of power and energy systems. *Proceedings of the 15th International Marine Design Conference (IMDC 2024)*. https://doi.org/10.59490/imdc.2024.916

Anderson, D., & Karambelas, H. (2024). *Through the joint, interagency, and multinational lens: Challenges posed by the national security strategy, volume 4*. US Army Command and General Staff College Press. https://www.armyupress.army.mil/Portals/7/Research%20and%20Books/2024/May/Through-the-Joint-Interagency-and-Multinational-Lens.pdf

Araujo, C., Gross, D., Steurer, M., Schegan, C., Ali, N., Bosworth, M., & Song, S. (2024). Baselining a functional architecture for a power electronic power distribution system for navy vessels. *IEEE Transactions on Transportation Electrification*. https://doi.org/10.1109/TTE.2024.1234567

Bernatskyi, A., Lukashenko, V., Siora, O., & Sokolovskyi, M. (2024). Analysis of the application of lasers for counter-UAV purposes. *History of Science and Technology, 14*(2), 487–512. https://doi.org/10.32703/2415-7422-2024-14-2-487-512

Chakraborty, S., Nijssen, E. J., & Valkenburg, R. (2022). A systematic review of industry-level applications of technology roadmapping: Evaluation and design propositions for roadmapping practitioners. *Technological Forecasting and Social Change, 179*, 121141. https://doi.org/10.1016/j.techfore.2021.121141

Chandler, D. L. (2014). Fiber lasers mean ray guns are coming. *IEEE Spectrum, 51*(3), 34–39. https://spectrum.ieee.org/aerospace/military/fiber-lasers-mean-ray-guns-are-coming

Cilli, M.V. (2015). *Improving defense acquisition outcomes using an integrated systems engineering decision management (ISEDM) approach* [Doctoral dissertation, Stevens Institute of Technology].

Ding, B., & Ferràs Hernández, X. (2023). Case study as a methodological foundation for technology roadmapping (TRM): Literature review and future research agenda. *Journal of Engineering and Technology Management, 67*, 101731. https://doi.org/10.1016/j.jengtecman.2023.101731

Doerry, N., & Amy, J. (2017). Electric ship power and energy system architectures. *IEEE Electric Ship Technologies Symposium (ESTS)*. https://ieeexplore.ieee.org/document/7991540

Doerry, N., & Amy, J. (2020). Key requirements for surface combatant electrical power system and propulsion system design. *Advanced Machinery Technology Symposium (AMTS), American Society of Naval Engineers*. https://www.doerry.org/norbert/papers/20200815-AMTS-Doerry-Amy-DistroA.pdf

Doerry, N. H., & Amy, J.V. Jr. (2024). Integrating power electronic equipment into shipboard power systems. *Proceedings of the SNAME Maritime Convention 2024*. https://onepetro.org/SNAMESMC/proceedings/SMC24/SMC24/D021S005R002/563290

Garcia, M., & Bray, O. (1997). *Fundamentals of technology roadmapping* (SAND97-0665). Sandia National Laboratories. https://doi.org/10.2172/471364

Harman, J., Edelman, E. S., Keane, J. M., Mahnken, T. G., Rudman, M., Sixkiller, M., ... & Wasser, B. (2024). *Report of the commission on the national defense strategy*.

Commission on the National Defense Strategy. https://www.armed-services.senate.gov/imo/media/doc/nds_commission_final_report.pdf

Hussain, M., Tapinos, E., & Knight, L. (2017). Scenario-driven roadmapping for technology foresight. *Technological Forecasting and Social Change, 124*, 160–177. https://doi.org/10.1016/j.techfore.2017.05.005

IEEE. (2023). *IEEE recommended practice for electrical installations on shipboard—design* (IEEE Std 45.1-2023). https://standards.ieee.org/ieee/45.1/10718/

Jeske, D. & Olson, D. (2024). Silo mentality in teams: emergence, repercussions and recommended options for change. *Journal of Work-Applied Management*. https://doi.org/10.1108/JWAM-07-2023-0064

Johnston, C. (2025, February 4). U.S. Navy HELIOS laser test underscores greater advancements in directed energy weapons. *Naval News*. https://www.navalnews.com/naval-news/2025/02/u-s-navy-helios-laser-test-underscores-greater-advancements-in-directed-energy-weapons/

Kerr, C. (2023). Technology roadmapping: A framework for planning and coordination in complex systems. *Systems Research and Behavioral Science*, *40*(5), 745–759. https://doi.org/10.1002/sres.2953

Kilby, J. (2025). *Statement on the readiness of the U.S. Navy*, testimony before the Senate Armed Services Committee, Subcommittee on Readiness and Management Support, 12 March. https://www.armed-services.senate.gov/imo/media/doc/statement_of_admiral_james_wkilby1.pdf

Kotila, B., Drezner, J. A., Bartels, E. M., Hill, D., Hodgson, Q. E., Huilgol, S. S., Manuel, S., Simpson, M., & Wong, J. P. (2023). *Strengthening the defense innovation ecosystem*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1352-1.html

LaGrone, S. (2014, December 10). U.S. Navy allowed to use Persian Gulf laser for defense. *USNI News*. https://news.usni.org/2014/12/10/u-s-navy-allowed-use-persian-gulf-laser-defense

Lawson, M., Moaz, T., Rajagopal, N., Mitrovic, V., Dong, D., & DiMarino, C. (2024). Model-based design and analysis of a Navy integrated power electronics building block (NiPEBB). *IEEE Transactions on Transportation Electrification*. https://doi.org/10.1109/TTE.2024.3454554

Letts, O., Hyatt, E., & McGinn, J. (2024). *PPBE, technology transition, and "the valley of death."* Acquisition Research Program, Naval Postgraduate School. https://dair.nps.edu/bitstream/123456789/5146/1/SYM-AM-24-083.pdf

Ma, J. (2021). Data-driven TRL transition predictions for early technology development in defence. *Defence Science Journal, 71*(6), 1–8. https://doi.org/10.14429/dsj.71.17230

Markle, S. P., Steurer, M. E., Gross, D. C., Bosworth, M. D., Ammeen, E. S., & Voth, J. M. (2021). The fundamental shift in US Navy warship power and energy system design. *Proceedings of the Engine as a Weapon International Symposium IX (EAAW IX)*, 15–17 November.

McCoy, T. (2015). Integrated power systems—An outline of requirements and functionalities for ships. *Proceedings of the IEEE, 103*(12), 2276–2284. https://doi.org/10.1109/JPROC.2015.2494212

McEntush, R., & Hay, L. (2023, March 13). *DoD contracting for startups 101*. Andreessen Horowitz. https://a16z.com/dod-contracting-for-startups-101/

Moore, D. (2024). *Conceptualizing the next-generation DoD innovation ecosystem*. Acquisition Research Program, Naval Postgraduate School. https://dair.nps.edu/handle/123456789/5273

Naval Sea Systems Command. (2019). *Naval power and energy systems technology development roadmap*. https://www.navsea.navy.mil/Portals/103/Documents/2019_NPES_TDR_Distribution_A_Approved_Final.pdf

O'Rourke, B. (2022). Now arriving: High-power laser competition. *Proceedings*, *148*(7), Article 1,433. https://www.usni.org/magazines/proceedings/2022/july/now-arriving-high-power-laser-competition

O'Rourke, R. (2024). *Navy shipboard lasers: Background and issues for Congress* (CRS Report No. R44175). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R44175

Olechowski, A. L., Eppinger, S. D., Joglekar, N., & Tomaschek, K. (2020). Technology readiness levels: Shortcomings and improvement opportunities. *Systems Engineering, 23*(4), 395–408. https://doi.org/10.1002/sys.21533

Phaal, R. (2024). What they should tell you about roadmapping at business school. *IEEE Engineering Management Review, 52*(3), 10–16. https://doi.org/10.1109/EMR.2024.3390160

Stotts, L., Paterson, R. & Greenberg, J. (2010). Command post of the future: Successful transition of a science and technology initiative to a program of record. *Defense Acquisition Review Journal, 17*(1), 3–20.

Tuinstra, J. D. (2022) *Speed through flexibility: Shortening the acquisition timeline of U.S. defense capabilities using flexible systems* [Master's thesis, MIT]. https://hdl.handle.net/1721.1/143304

Ullman, H.K. (2024). Navigation plan for America's warfighting navy: Aspirational or actionable?. *Proceedings*, *150*(9), Article 1,459.

Voth, J. & Sturtevant, G. (2022). Digital engineering: Expanding the advantage. *Journal of Marine Engineering & Technology, 21*(6), 355–363. https://doi.org/10.1080/20464177.2021.2024382

PANEL 3. EMPOWERING THE ACQUISITION WORKFORCE: PERSPECTIVES FROM SERVICE DIRECTORS

Wednesday, May 7, 2025	
0825 – 0940 PT 1025 – 1140 CT 1125 – 1240 ET	**Chair: Mr. Frank Kelley, BGen, USMC (Ret.), Vice President of Defense Acquisition University (DAU)** **Panelists:** **Ronald R. Richardson, Jr.,** Director, U.S. Army Acquisition Support Center and Director, Acquisition Career Management (DACM) **Renee King,** U.S. Navy Deputy Director, Acquisition Talent Management (DATM) **Otis Lincoln,** 4th Estate Director Acquisition Career Management (DACM)

Mr. Frank Kelley, BGen, USMC (Ret.)—is the Vice President of Defense Acquisition University (DAU). In this position, he is responsible for aligning DAU strategic plans to the goals of both the Secretary of Defense and Under Secretary of Defense for Acquisition and Sustainment, while continuing to build the outstanding reputation of DAU as the Department of Defense primary learning institution for acquisition. He oversees the development and expansion of acquisition curriculum and learning opportunities and the delivery of those learning assets throughout the five DAU regional campuses, the Defense Systems Management College, and the College of Contract Management.

Prior to this assignment, Mr. Kelley served as the Deputy Assistant Secretary of the Navy Unmanned Systems from October 2015 – June 2018. In this capacity, he was the principal advisor to the Assistant Secretary of the Navy for Research, Development and Acquisition on matters relating to unmanned systems across all domains—land, sea, and air.

Mr. Kelley joined the civil service in 2015 following a 32-year career as a United States Marine.

In 1983, Mr. Kelley graduated from the University of Notre Dame with a degree in Aerospace Engineering and was the recipient of the Naval ROTC Donald R. Bertling Award. He was commissioned as a Second Lieutenant after completing Officer Candidate School.

He attended the Marine Corps War College and taught at the Command and Staff College. Mr. Kelley transferred to Marine Corps Systems Command in Quantico, VA, where he was the Program Manager for Unmanned Systems. His next assignment was Military Assistant to Dr. Delores Etter, then the Assistant Secretary of the Navy for Research, Engineering, and Acquisition.

In August 2007, Mr. Kelley was assigned to the position of Marine Corps Systems Command Program Manager for Training Systems in Orlando, FL. In August 2009, he was reassigned as the command's Chief of Staff before being promoted to the rank of Brigadier General and assuming command from July 2010 to July 2014. He then served in the position of the Vice Commander, Naval Air Systems Command, preceding his last military assignment as Director for Prototyping, Experimentation and Transition in the Office of the Deputy Assistant Secretary of the Navy.

Ronald R. Richardson, Jr.—currently serves as the Director of the Army Acquisition Support Center. In this role, he oversees the Army Acquisition Corps (AAC) and the Army Acquisition Workforce (AAW), and supports the Army's Program Executive Offices in the areas of human resources, resource management, program structure, acquisition information management, and program protection.

Mr. Richardson has over 30 years of medical, information, and weapon system acquisition experience as both a Department of Defense (DoD) civilian and a U.S. Army Officer. Before coming to ASC, he served as the Director of Acquisition and Operations for Program Executive Office Soldier. Prior to joining PEO Soldier, he was the Deputy Project Manager for the DoD Healthcare Management System Modernization (DHMSM®) Program, a $14B Major Automated Information System (MAIS) acquisition to replace the legacy Military Health System (MHS) Electronic Health Record (EHR) with an off-the-shelf (OTS) system now known as MHS GENESIS. Before that, he was the Product Lead for Increment 3 of the Integrated Electronic Health Record (iEHR) Program in the DoD/Department of Veterans Affairs Interagency Program Office (IPO). Prior to joining the DoD/VA IPO, he served as the Director of Acquisition Review and Analysis for the Office of the Assistant Secretary of the Army, Acquisition, Logistics and Technology (ASA(ALT)). Before joining ASA(ALT), Mr. Richardson served in a multitude of Military, Civilian, and Private Sector positions culminating in his selection for Senior Service College.

Mr. Richardson received his M.S. in Biomedical Engineering from Duke University, and his M.S. in National Resource Strategy from the Industrial College of the Armed Forces (ICAF). He is also a graduate of the U.S. Army Command and General Staff College.

He is the recipient of the Superior Civilian Service Medal with Two Oak Leaf Clusters, the Meritorious Civilian Service Medal, the Civilian Service Achievement Medal, the Army Staff Identification Badge, and the Order of Military Medical Merit. Mr. Richardson also holds multiple professional memberships and certifications, including membership in both the Army and Defense Acquisition Corps, and Level III Defense Acquisition Workforce Improvement Act (DAWIA) Certification in Program Management, Science and Technology Management, and Systems Engineering.

Ms. Renee King is the Chief of Staff for the Director, Acquisition Talent Management office under the Assistant Secretary of the Navy Research, Development and Acquisition (ASN RDA). Ms. King assumed this role in March 2024. In this role, she leads and synchronizes a staff of ten action officers and contractors, who support the Navy's 54,000-member Acquisition Workforce and the leadership and developmental programs and initiatives that enhance the capabilities of the Navy Warfighter.

In 2021, Ms. King competed and was selected to participate in the Naval Sea Systems Command (NAVSEA) Commander's Executive Fellows (CEFP) program as part of Cadre VII. This distinguished leadership initiative provided her with the chance to assist multiple commands, including the Program Executive Office (PEO) Ships, United States Fleet Forces Command, the NAVSEA Transformation Office (NTO), the Shipyard Infrastructure Optimization Program (SIOP) and Industrial Operations in various roles from an Acquisition Program Manager (APM) to Chief of Staff.

Prior to her selection into CEFP, Ms. Renee King served as the Director for the NAVSEA Acquisition Workforce Program Office in SEA 10. Ms. King led three enterprise-wide programs, the Defense Acquisition Workforce Development Account (DAWDA), the Defense Acquisition Workforce Improvement Act (DAWIA) Program and the Naval Acquisition Development Program (NADP). These programs collectively supported more than 20,000 Acquisition Workforce members, funded approximately $18M in incentives and training to those employees (annually), and created a pipeline of more than 400 NADP Interns and Professionals to NAVSEA' s acquisition workforce. In September 2016, Ms. King graduated from the NAVSEA Journey Level Leader (JLL) Program, in which her 90- day rotation supported Surface

Maintenance Engineering Planning Programs (SURFMEPP) Third Party Planning division, engaging in the technical management of three major contracts.

Before her employment with NAVSEA in 2010, Ms. King's background included various positions at institutions of higher education. Ms. King worked as a Human Resource Specialist with the University of South Carolina. She also worked as an administrative assistant at Trident Technical College, in North Charleston, SC supporting the Veterans Upward Bound Program, a program designed to transition Service members from active duty to prepare for successful postsecondary education completion.

Ms. King particularly enjoyed her role at Trident Technical College, supporting veterans, as she is a United States Marine Corps veteran herself. She continues to support and volunteer her time with veterans and veterans-based programs. Ms. King is an active member of the Women Marines Association (WMA), Federally Employed Women (FEW), and Blacks In Government (BIG). Ms. King holds a Bachelor of Science degree in Business Management from Park University and is currently pursuing a master's degree in organizational leadership from Southern New Hampshire University.

Mr. Otis Lincoln—serves as the Director, Acquisition Career Management (DACM) for the 4th Estate (31 defense agencies/field activities) with oversight of statutory training, professional credentialing, continuous learning, and career development for more than 31,000+ acquisition workforce members. He entered federal service in 2009 as a Contract Specialist within the Office of the Chief Financial Officer (CFO) of the Defense Intelligence Agency (DIA). After serving as a warranted Contracting Officer on several procurements supporting multiple Directorates across DIA for several years, he continued to expand his aperture within the acquisition community, moving into the project and program management realm supporting Small Business Programs and Human Resources analytics. In various capacities, he was responsible for the successful planning and execution of several multi-million-dollar programs that included increasing acquisition exposure to industry, training, and career development of the agency's acquisition workforce. He also played a fundamental part in the hiring and placement of new acquisition members and oversaw the development of career paths in the Finance and Acquisition professional fields.

Mr. Lincoln supported the Navy Systems Management Activity (NSMA) as their DAWIA Program Director, expanding their training, certification, and career development programs for the acquisition workforce. Following his tenure at NSMA, Mr. Lincoln assumed a senior leadership position as a Section Chief in the Contracting Office within CFO supporting the Mission Service's and Command Element's global procurement requirements. Prior to his government service, he spent 10 years in the financial sector, which provided him with the background to enter into the contracting field, first with industry and then in the Federal sector.

Mr. Lincoln is DAWIA certified in Contracting Professional and Program Management Advanced. He also holds a Bachelor of Science in Marketing from Virginia Commonwealth University and a Master of Business Administration from the University of Phoenix.

.

PANEL 4. BALANCING INNOVATION AND PROTECTION: DECISION FRAMEWORKS AND DIGITAL MATURITY IN DEFENSE ACQUISITION

Wednesday, May 7, 2025	
0825 – 0940 PT 1025 – 1140 CT 1125 – 1240 ET	**Chair: Mark E. Krzysko, Principal Deputy Director, Acquisition Policy and Innovation (API), OUSD A&S** ***Balancing Access and Protection: A Decision Framework for Additive Manufacturing Intellectual Property Rights in Defense Acquisition*** Waterloo Tsutsui, Senior Research Associate, Purdue University ***Integrated Digital Maturity Pathway for Technical Data Packages*** Darryl Draper-Amason, Research Assistant Professor, Old Dominion University Virginia Digital Maritime Center (VDMC) ***Time Value of Data Decision Modeling for Major Defense Acquisition Programs*** Frank Goertner, Director, Technology Management Programs, University of Maryland

Mr. Mark E. Krzysko—is the Deputy Director in Acquisition Policy and Innovation (API) and oversees Enterprise Information, Acquisition Analytic Support, the Acquisition Innovation Research Center (AIRC) and the Defense Civilian Training Corps (DCTC). In this senior leadership role, Mr. Krzysko directs acquisition data governance, data access, and data science to enable the Department to make sound business decisions with data. He is leading a philosophical and technical transformation within the Department to make timely, authoritative acquisition information available to support insight and decision-making on the Department of Defense's major programs—a portfolio totaling approximately $2 trillion of investment funds over the lifecycle of the programs—as well as smaller programs and nontraditional acquisition approaches.

Mr. Krzysko holds a Bachelor of Science Degree in Finance and a Master of General Administration and Financial Management, from the University of Maryland University College, and numerous certificates from Harvard University.

Balancing Access and Protection: A Decision Framework for Additive Manufacturing Intellectual Property Rights in Defense Acquisition

Waterloo Tsutsui—is a Senior Research Associate in the School of Aeronautics and Astronautics at Purdue University, Indiana. Tsutsui received his PhD in aeronautics and astronautics from Purdue University in 2017. Before Purdue, Tsutsui practiced engineering in the automotive industry for more than a decade, with his final position focused on research and development of lithium-ion battery systems for electric vehicles. Tsutsui's research interests include systems engineering, structures & materials, product design & advanced manufacturing, and engineering education. Tsutsui is the recipient of the 2023 Engineering Education Excellence Award from the National Society of Professional Engineers (NSPE).

Qian (Alex) Shi—is a PhD candidate at the School of Aeronautics and Astronautics at Purdue University. Her research involves developing methods and tools for addressing challenges in space sustainability, with a focus on the impact of large satellite constellations. She also has experience developing decision support tools and frameworks for tackling complex, multi-disciplinary problems in a range of aerospace systems. Shi obtained her bachelor's and master's degrees in mechanical engineering from the University of Cambridge, UK, on a Singapore Public Service Commission scholarship. Prior to joining Purdue University, she spent several years as a policymaker in the Singapore government, where she worked in diverse fields including economics, finance, and infrastructure policies.

Dalia Bekdache—is a PhD candidate in the School of Aeronautics and Astronautics at Purdue University, majoring in aerospace systems with a minor in astrodynamics. Her research focuses on developing methods and support tools for digital mission engineering, including the integration of commercial technologies into civil space missions. Bekdache holds a Bachelor of Engineering in Mechanical Engineering with a minor in applied mathematics from the American University of Beirut in Lebanon and a Master of Science in Aeronautics and Astronautics from Purdue University, along with a computational science and engineering certification. She has also completed a systems engineering certification as part of her PhD training. Bekdache is currently a Graduate Research Assistant affiliated with both the Center for Integrated Systems in Aerospace and the Digital Enterprise Center at Purdue's Indiana Manufacturing Institute, where she applies systems and system-of-systems engineering methods to support digital thread implementation in manufacturing environments.

Prajwal Balasubramani—is a Research Scientist at Amazon Fulfillment Technologies & Robotics. He earned his PhD from the School of Aeronautics and Astronautics at Purdue University in 2024. He has worked on diverse projects ranging from operations research, multi-agent system and simulation, system-of-systems modeling and analysis, machine learning, and explainable AI. His PhD dissertation focused on assessing and increasing trustworthiness in machine learning models by leveraging ensembled explanations and reasoning. His research interests include modeling & simulation, machine/deep learning, explainable AI, system design, multi-agent systems, operations research, and gamification.

Jitesh H. Panchal—is a Professor of Mechanical Engineering at Purdue University. He received his BTech from the Indian Institute of Technology (IIT) Guwahati and MS and PhD from the Georgia Institute of Technology. He is a member of the Systems Engineering Research Center (SERC) Council. He is a recipient of the NSF CAREER award, Young Engineer Award and three best paper awards from ASME, and was recognized by the Schaefer Outstanding Young Faculty Scholar Award, the Ruth and Joel Spira Award from Purdue University. He is a co-author of two books and a co-editor of one book on systems design.

Stephan Biller—is the Harold T. Amrine Distinguished Professor in the School of Industrial Engineering and the Mitchell E. Daniels, Jr. School of Business at Purdue University and serves as Director of the Dauch Center for the Management of Manufacturing Enterprises at the Daniels School of Business and the Co-director and Founder of Purdue's national initiative in eXcellence in Manufacturing and Operations (XMO). His expertise includes Smart Manufacturing, Digital Twin, Industry 4.0, and Resilient Supply Chain Management. He focuses on AI and IoT integrated with lean Six Sigma manufacturing processes to facilitate the Digital Transformation of large and especially small and medium manufacturing enterprises scale. Previously, he served as Founder and CEO of Advanced Manufacturing International, Vice President of Product Management for AI Applications & Watson IoT at IBM, Chief Manufacturing Scientist & Manufacturing Technology Director at General Electric, and Tech Fellow & Global Group Manager for Manufacturing Systems at General Motors. He is a Six Sigma Master Black Belt, an IEEE Fellow, and an elected member of the National Academy of Engineering.

Daniel A. DeLaurentis—is the Vice President for Discovery Park District (DPD) Institutes and the Bruce Reese Professor of Aeronautics and Astronautics at Purdue University. He directs the Center for Integrated Systems in Aerospace (CISA) researching modeling, design optimization and system engineering methods for aerospace systems and systems-of-systems, including urban and regional advanced aerial mobility and hypersonic systems. He is a Senior Research Fellow at the Krach Institute for Tech Diplomacy at Purdue. DeLaurentis served as Chief Scientist of the U.S. DoD's Systems Engineering Research Center (SERC) UARC from 2019 to 2023. He is an elected fellow of the American Institute of Aeronautics and Astronautics (AIAA) and the International Council on Systems Engineering (INCOSE).

Abstract

This research addresses the challenge of managing intellectual property (IP) rights in defense additive manufacturing (AM) acquisition. Specifically, the Department of Defense must balance operational requirements for IP access against defense industrial base companies' interests in protecting valuable IP assets. We introduce a decision framework to navigate IP complexities in AM applications, encompassing scenario screening, AM lifecycle analysis, IP asset identification, and strategic considerations for acquisition. The framework employs real options theory to provide acquisition professionals with structured guidance while maintaining strategic flexibility, demonstrated through a "Demand Surge" vignette examining a scenario where the DoD must rapidly increase the supply of a proprietary respirator mask beyond the original equipment manufacturer's production capacity during a crisis. Our results indicate that effective IP management in defense AM requires careful consideration of mission requirements, technological capabilities, and stakeholder interests, revealing critical decision points in the AM lifecycle where IP strategy significantly impacts program success. This research contributes a systematic approach to IP strategy development, promoting both fair compensation for IP holders and sustainable defense capabilities while identifying avenues for future research.

Introduction and Background

Intellectual property (IP) rights are a critical concern in Department of Defense (DoD) acquisitions. Obtaining and licensing the correct IP ensures that systems remain operational, sustainable, adaptable, and cost-effective (DoD Instruction 5010.44, 2019; GAO-22-104752, 2021). Thus, the DoD must obtain appropriate IP and technical data rights to operate, maintain, and sustain the capabilities it acquires from the defense industrial base (DIB). Without sufficient IP rights, the DoD may face issues like vendor lock, limited ability to source upgrades or repairs competitively, and surging sustainment costs (GAO-23-105850, 2023; Peters, 2022; Wydler, 2014). However, DIB companies view their IP as a valuable capital asset representing significant investments, thereby becoming a source of market competitiveness and future income. DIB entities aim to protect their IP rights to preserve their asset's monetary value (Hickey, 2022; Peters, 2022). As a result, the varying viewpoints on IP rights between the DoD (seeking access) and DIB entities (seeking protection) lead to tensions that require delicate handling (Tsutsui, Shi, et al., 2024). Therefore, the Purdue research team undertook the research, recognizing the pivotal role that IP rights play in DoD acquisitions and the impact on ongoing operations and sustainment (DeLaurentis, Biller, et al., 2024), which complements work on digital transformation in defense (Panchal et al., 2023, 2024; Tsutsui, Atallah, et al., 2024), as IP rights and digital implementation are intrinsically linked in acquisition.

Another opportunity involves the recent progress in additive manufacturing (AM) and three-dimensional (3D) scanning technologies. This includes addressing rights and compensation for IP holders in AM and determining suitable methods for identifying and incorporating these considerations into contractual agreements (Vogel, 2016; Widmer & Rajan, 2016). Therefore, effective IP management for AM is critical for ensuring a successful defense acquisition. However, the IP landscape within the AM domain presents significant challenges, necessitating a structured approach for effective navigation. Hence, we propose a greenfield approach to navigating negotiations for IP accounting for the uniqueness of AM. This effort was conducted with the notion of being able to work either in tandem or post-hoc integration with existing processes within the DoD.

Managing IP rights in AM for defense acquisitions presents challenges as government operational requirements often conflict with contractors' IP protection priorities. This paper addresses these challenges by presenting a decision framework tailored to the complexity of IP in AM applications, covering scenario screening, AM lifecycle analysis, IP asset identification, and strategy options. To demonstrate practical application, the framework is illustrated through a "Demand Surge" vignette, serving as a guide that blends theoretical insights with practical applications to strengthen AM IP management in the defense acquisition process.

Methodology: Proposed Decision Framework

The proposed decision framework aims to provide DoD users with a structured and informed decision-making process in IP acquisition for AM projects. An overview of the framework is shown in Figure 1. The framework consists of three steps as follows:

1. Scenario screening and scoping: to determine framework applicability and extract relevant use case information
2. IP asset identification and considerations: to ascertain the why (impetus), what (scope), and how (modality) of IP acquisition
3. IP strategy formulation: to consolidate the information and considerations and formulate the IP acquisition options and overall acquisition strategy

The rest of this section details each step and outlines the rationale and pertinent considerations in using the framework.

Scenario Screening and Scoping

The scenario screening process serves as the initial step to determine whether the proposed AM IP framework is appropriate. It is important to note that the framework is not meant to be used in cases where IP compensation is secondary to urgent operational needs or IP protection has expired or does not exist. For those cases, it may be more practical to consider other approaches like reverse engineering or leveraging the Defense Production Act where suitable.

Typically, the acquisition process encompasses several stages:

- Requirements definition
- Market research and supplier identification
- Supplier negotiations
- Contracting
- Development
- Production/Sustainment
- End-of-Life management

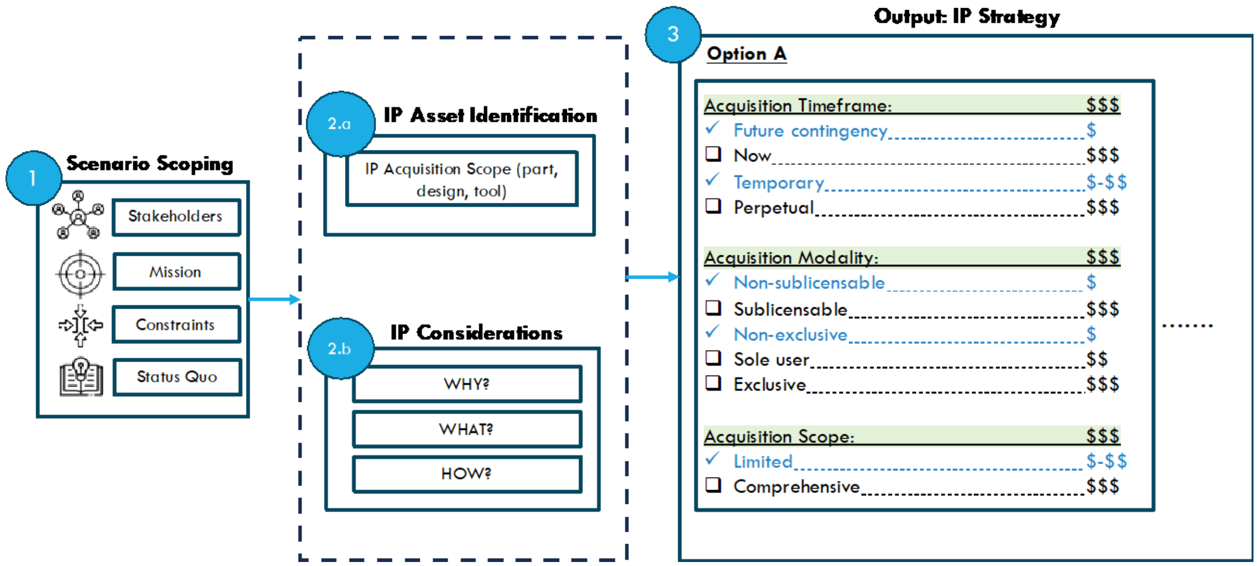


Figure 1. AM IP Strategy Decision Framework

The proposed framework is designed to be implemented during the supplier negotiations phase, ensuring that all negotiations made during this phase incorporate provisions for future IP acquisition rights. While the framework is expected to be applied during contracting, the actual scenarios it anticipates have not yet occurred at the time of framework application. Therefore, the framework should be used to develop preemptive strategies for swiftly acquiring existing IP of additively manufactured processes and parts in the future.

Once the scenario screening has been verified, the scoping process is the next step for extracting and synthesizing information from projected scenarios. This process involves: synthesizing the key elements of the scenario, clearly defining the problem statement, and listing any relevant assumptions and constraints that influence the decision-making process for IP strategy. To effectively address the scenario, the following scoping features must be defined:

1. OEM and Manufacturing Status: Anticipating the impact of future scenarios on OEM efficacy and potential manufacturing capabilities.
2. Part/Process sourcing: Identifying potential substitutes to the part/system to procure (if any).
3. IP Acquisition Requirements: Identifying DoD needs and requirements for IP acquisition.
4. Mission Time and Resource Constraints: Determining time-sensitive and resource-dependent factors.
5. Mission Criticality: Assessing the importance of additively manufactured parts or systems to the mission.
6. AM Capability Location: Identifying the need for in-theatre and/or out-of-theatre production and maintenance.
7. IP Rights Status: Identifying technical data AM parts or systems protected by IP, including ownership of rights.

With these scoping features delineated, a systematic approach was developed to identify the relevant details, which are cataloged in Table 1. These features set the baseline requirements that inform the management of IP assets and considerations and ultimately guide the final IP strategy formulation.

Table 1. Scenario Scoping: Systematic Method for Extracting Information to Inform IP Strategy

Scoping Category	Scenario features
OEM Status	Active or Inactive?
Manufacturing Status	Ongoing or discontinued?
Sourcing	Single-sourced or multi-sourced?
IP Acquisition Requirements	What are some needs/requirements that the IP acquisition strategy must fulfill?
Mission Status and Criticality	What are the timeline and criticality of the mission?
AM Capability Location	In-theatre or Out-of-theatre?
IP Rights Status	What parts, processes, and tools are protected by IP, and who owns the rights?

AM Lifecycle and IP Asset Identification

Surveying the vast landscape of protected assets and then identifying the relevant ones for the acquisition effort is the next step in configuring an acquisition strategy and compensation. To ensure comprehensive identification and reduce oversight, the product manufacturing lifecycle is utilized as a guiding mechanism and followed by a step-by-step vertical exploration in each of the lifecycle phases (i.e., design, build, post-process, testing, Maintenance/Repair/Operations [MRO], and end-of-life) to produce a portfolio of acquirable/needed assets.

Figure 2 depicts the lifecycle of an additively manufactured product and the possible IP assets involved at each phase of the lifecycle. For example, 3D models and digital design assets are identified in the Design/Planning stage of the product lifecycle. Similarly, unique maintenance processes and/or data assets are identified under the MRO phase of the lifecycle.

The assets are categorized (see Figure 2 legend) based on their nature into design, process, software, part, and tool IP. This enables modular acquisition strategies that are either demanded by the scenario or based on already acquired assets. Users have the freedom and flexibility to approach the asset grouping for acquisition either by manufacturing phase or IP area.

While the lifecycle outlined in Figure 2 details the stages and associated IP assets in additively manufactured products, it is crucial to recognize the practical scope of AM within larger systems. Typically, not all components of a product or system are suitable or cost-effective for AM. Instead, specific parts or components are identified as viable candidates for AM due to their design complexity or customization requirements. Therefore, negotiations for AM-related IP rights often represent just one facet of the broader strategy to acquire comprehensive data and IP rights.

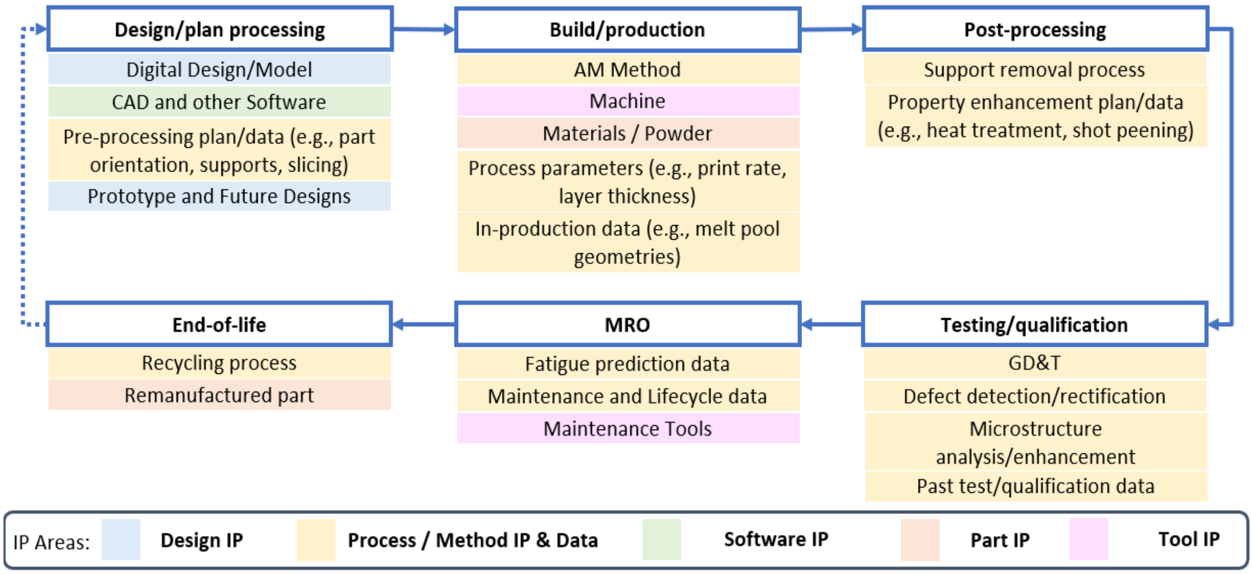


Figure 1. IP Assets in the AM Lifecycle

The assets listed in the framework schematic are not meant to be universal and/or exhaustive. Each acquisition effort will result in a unique asset portfolio based on the objective of the acquisition. First, if the objective is centered around in-use/deployed systems but needs new strategies for end-of-life management, the assets needed to carry out the operation are all identified within the MRO and end-of-life phases. The government can ignore the IP assets in the earlier phases while negotiating with the supplier. Contrarily, if the government is interested in a single-use product, it would buy assets concerning the first four phases without paying attention to MRO and end-of-life IP assets. Finally, suppose the government is interested in procuring a completely new product with no previous production and is planning on using said product for multiple years/cycles. In that case, all aspects of the lifecycle must be considered, and all assets must be carefully selected. The section titled What IP Assets Should Be Acquired? also discusses additional qualitative reasoning that helps choose from the identified assets based on mission needs and constraints.

It is also worth noting that identified assets may or may not have the same type of IP protection. The design of the framework is inspired by and accounts for the following types of IPs: Patents, Copyrights, Trademarks, and Trade Secrets. Knowing the type of IP informs negotiation and compensation strategies. Each type of IP has its unique strictness to usage; some can be more readily negotiated than others. For example, trade secrets are often more complicated to negotiate and procure when compared to buying a copyright license or licensing a patent. This a priori knowledge of distinguishable traits among the asset types helps optimally compensate during the acquisition.

AM IP Strategy Considerations

The next step is to use the relevant information from the scenario scoping and IP asset identification phases to ascertain the key considerations driving the AM IP acquisition strategy. The broad categories of considerations are:

1. Why should IP acquisition be considered? What is the value of the IP asset(s)?
2. What IP assets should be acquired?
3. How should IP acquisition be structured?

The rest of this section sets out a series of decision trees and guiding questions to help determine the features of an appropriate acquisition strategy.

Why Should IP Acquisition Be Considered?

In commercial IP trading, the value of an IP asset to the buyer usually refers to the expected benefits (often economic) from owning the asset. For example, it could include revenue growth from new product sales or increased market share from new customer segments. This usually forms the impetus for IP acquisition. Defense-related acquisitions by the government tend to differ from this aspect in that economic gain is not the primary aim. Instead, a more appropriate measure is the inverse “cost of inaction” (i.e., what the government stood to lose if IP acquisition were not carried out). Considerations to ascertain this cost of inaction (Figure 3) include:

1. How mission-critical are the systems/components that rely on this IP?
2. What are the alternatives to these systems/components, and how do their functionalities and costs compare?
3. What is the cost of ownership and opportunity costs of acquisition?

For example, suppose the systems/components that rely on the IP are mission-critical with few comparable alternatives. In that case, the impact of non-acquisition on mission success is likely to be high, resulting in a high cost of inaction. The converse would correspond to a low cost of inaction. In addition, the cost of ownership (e.g., data, system, and manpower upkeep) and the opportunity cost of acquiring the IP provide one benchmark for assessing whether the cost of the inaction provides a sufficient impetus to consider IP acquisition. If these costs far outweigh the cost of inaction, then acquisition may not be a good option, and there is no need to go through the rest of the framework to determine an acquisition strategy.

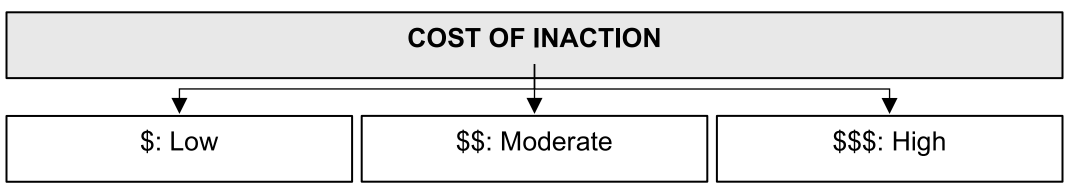


Figure 2. Decision Tree: Cost of Inaction[1]

What IP Assets Should Be Acquired?

Once the impetus for IP acquisition has been established, the next step is to determine the acquisition scope (Figure 4). This requires reviewing the list of relevant IP assets identified in the section titled AM Lifecycle and IP Asset Identification and prioritizing them according to their importance to the mission. Relevant considerations include:

1. What IP is necessary (cannot manufacture without) vs. good to have (makes manufacturing easier)?
2. What are the dependencies across the IP assets, if any?
3. Which parts of the AM lifecycle could be changed to lift the dependency on specific IP assets?

While having a thoroughly ranked list of IP assets will enable a more detailed calibration of acquisition scope, one should minimally aim to classify the assets into those necessary to

[1] In these decision tree figures, the number of dollar signs indicate the relative costs of the branches. With use case-specific information, these costs can be quantified to provide a more precise scale for decision-making.

enable manufacturing and those good-to-have. This will provide at least two acquisition options to suit different scenarios. For instance, a low-priority, low-budget mission may constrain the buyer to acquire the bare minimum IP. In contrast, higher priority/budget missions may require a more comprehensive acquisition scope.

In some cases, dependencies between IP assets (e.g., background IP) may further constrain which IP assets must be acquired together or even warrant all-or-nothing options. Consider whether parts of the AM lifecycle (for the specific system/component) can be adjusted to negate the requirement for one or more IP assets. This could be useful if certain IP assets are costly and/or have extensive dependencies on background IP. For example, the required AM machines, material, and process parameters are constrained by the choice of AM method. For very niche AM methods, there may be few suppliers with sufficient experience and expertise to work out the appropriate production parameters without having access to the relevant process IP. Hence, in evaluating IP acquisition for products that rely on niche AM methods, one may be compelled to consider process IP acquisition (in addition to product IP). One way to avoid being locked into acquiring a suite of IP may be to explore possible changes to the AM method. This would require additional input from stakeholders like engineering teams or AM experts and could thus be resource-intensive. For this reason, it would be prudent to explore options to adjust the AM lifecycle only if the expected IP cost is high or resources allow it.

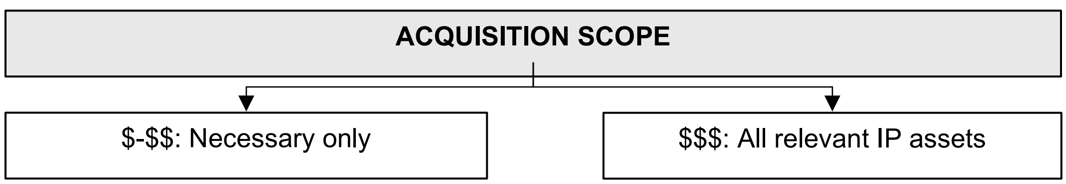


Figure 4. Decision Tree: IP Acquisition Scope

How Should the IP Acquisition Be Structured?

Finally, there are a series of considerations to determine the suitable modality of IP acquisition:

1. When should the IP be acquired (now or as a future contingency)?
2. How long is the data needed for (one-off vs. time-limited vs. in perpetuity)?
3. Are sublicensing rights required in addition to usage rights?
4. How sensitive is the IP – is there a need to limit distribution?

The consideration of the acquisition timing is twofold – whether one should acquire the actual IP now or buy an option to acquire the IP later (Figure 5), and if the latter when the option to acquire should be exercised. The first decision could depend on whether there is any use for the IP, such as creating redundancies in supply chains for strategic goods. The second decision hinges on the lead time required for users to develop the necessary skills and system literacy to use the IP effectively. More complex systems or components may necessitate earlier acquisition to allow sufficient time for capability building. There is, however, a tradeoff with cost – acquiring IP earlier may be a more expensive option due to the higher net present value of money and a longer tail of IP ownership expenses.

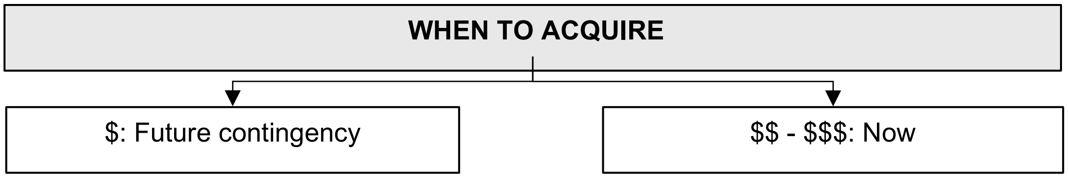


Figure 5. Decision Tree: IP Acquisition Timeline

Factors like the expected mission duration, OEM status, and system/component manufacturing status will drive the required IP use duration (Figure 6). Generally, one might expect the IP use duration to scale with the mission duration, resulting in temporary, time-bound licensing arrangements with the OEM. However, if there are uncertainties around the OEM's operational status or product sustainment capabilities, then a perpetual rights transfer might be a safer arrangement. A perpetual rights transfer option would generally cost more than licensing arrangements, with few exceptions (e.g., OEM liquidating assets below "market value" to manage cash flow/avoid bankruptcy).

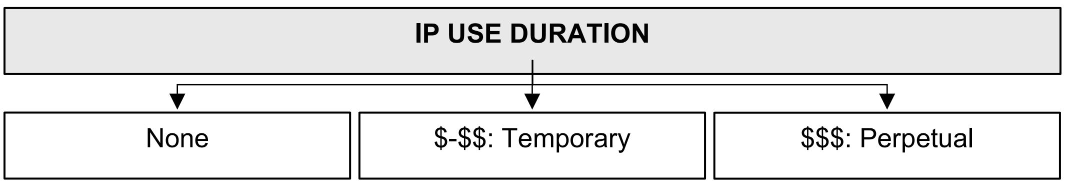


Figure 6. Decision Tree: IP Use Duration

Sublicensing rights (Figure 7) considerations are usually straightforward. It essentially comes down to whether the buyer requires the flexibility to distribute the IP rights to others in addition to using it in-house. This could be driven by the practice of outsourcing manufacturing functions or the need to tap into a wider supplier base to augment manufacturing capacity. In general, we would expect sublicensable rights to cost more since it could mean sharing what might have been the OEM's "trade secrets" with potential competitors, thus eroding some of the OEM's competitive advantage for similar future manufacturing contracts.

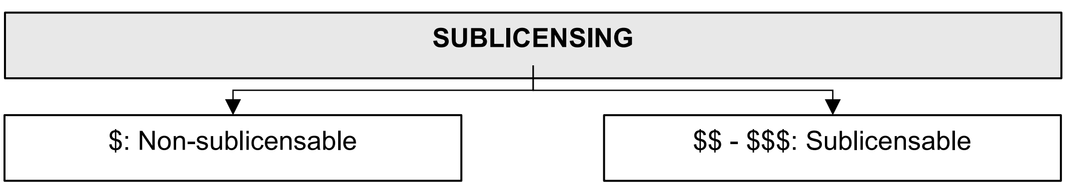


Figure 7. Decision Tree: IP Sublicensing Rights

Finally, we need to consider the level of exclusivity required in IP ownership (Figure 8). In this regard, there are three main categories of IP licenses (Halt et al., 2017; although more detailed calibration of the terms and conditions of IP ownership/use can be crafted using appropriate contractual clauses): non-exclusive, sole, and exclusive. Non-exclusive licenses allow multiple licensees, where the original IP owner can continue to own, use, and sublicense its IP to others. Sole licenses allow both the licensor and licensee to exploit the IP, but neither can sublicense it to others. Exclusive licenses offer the most flexibility to the IP buyer, who essentially would enjoy a monopoly on the IP rights. This option usually also prohibits the original IP owner from using the IP, except for any retained rights (usually non-commercial) provided for in the acquisition contract. The appropriate level of exclusivity will depend on how much control the buyer needs over the IP and the supply chain model. For example, potentially sensitive patents may require higher exclusivity, as information flow may need to be tightly controlled. However, exclusivity may not be a concern without chain outsourcing and sub-contracting restrictions.

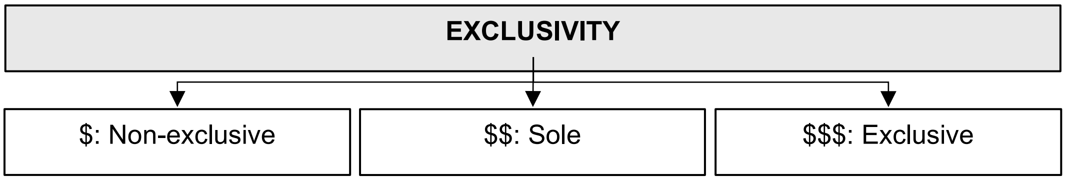


Figure 8. Decision Tree: IP Exclusivity

In practice, where there are sufficient grounds for secrecy (e.g., militarily sensitive information), the government can control IP rights via security classification, export controls, or secrecy orders under the Invention Secrecy Act. There may not be a compensation premium for IP exclusivity in those cases.

AM IP Strategy Formulation

The final step in the framework is to consolidate the information obtained from previous steps to construct an acquisition strategy, which may comprise multiple acquisition options. The underlying principle draws from real options theory, which applies the idea of financial options to quantify and account for the value of flexibility and delay in investment decisions (Trigeorgis & Reuer, 2017; Weeds, 2002). There are many case studies on how real options can be used to support investment analysis of a range of business decisions, such as franchise network expansion (Gorovaia & Windsperger, 2013; Nugroho, 2016) and firm merger & acquisition deals (Čirjevskis, 2024). Here, we borrow the same concept to aid IP acquisition decisions.

Figure 9 illustrates a template checklist (no check marks provided in the figure shown) that is used to consolidate the information obtained from previous steps in the framework. One must check the suitable features to create an IP acquisition option. For example, one IP acquisition option could include features like a comprehensive acquisition scope ($$$), which includes all necessary and good-to-have IP for a future contingency plan ($) for a temporary period (depending on mission duration, $-$$), and with non-exclusive IP rights that can be sublicensed to contractors. This option can be included in the acquisition contract with the appropriate legal language. Based on the flexibility of stakeholder and mission requirements, multiple acquisition options may be possible. In this case, the DoD may (or may not) want to prepare multiple acquisition options that make the IP acquisition strategy.

Acquisition Timeframe:	$$
❑ Future contingency	$
❑ Now	$$$
❑ Temporary (e.g., mission duration)	$-$$
❑ Perpetual	$$$

Acquisition Modality:	$$
❑ Non-sublicensable	$
❑ Sublicensable	$$$
❑ Non-exclusive	$
❑ Sole user	$$
❑ Exclusive	$$$

Acquisition Scope:	$$$
❑ Limited	$-$$
❑ Comprehensive	$$$

Figure 9. A Template Checklist for Acquisition Option: Selecting a Combination of Features for a Possible IP Acquisition Agreement

Results: Decision Framework Demonstration Using a Vignette

We have constructed a fictional vignette to demonstrate the framework application. The distinct features of this vignette aim to cover a specific mission condition for which the framework is applicable but is by no means meant to be exhaustive or representative of specific current or planned DoD activities. The vignette, which is elaborated in the rest of this section, is as follows:

> **Demand Surge**: The OEM may be unable to supply enough resources due to a potential demand surge (e.g., imminent threat and pandemic).

For this vignette, we have made specific assumptions to aid analysis. The recommended acquisition strategy depends on these assumptions and will likely change if these assumptions are different. For this reason, we also conduct a sensitivity analysis on some key assumptions to examine whether and how they affect the recommended acquisition strategy. Nevertheless, this section aims to set out the process of framework application rather than to recommend any specific output produced by the framework. The recommended option, which can be one of several acquisition options, is presented at the end as a checklist. It must be described using appropriate legal language for inclusion in the acquisition contract.

In our research, we also considered two additional vignettes that present interesting applications of the framework: (1) “Limited Access to Original Equipment Manufacturer (OEM),” which explores scenarios where the DoD has restricted access to OEMs due to manufacturing discontinuation or inactive suppliers; and (2) “MRO,” which addresses situations requiring urgent maintenance improvisation due to mission criticality or in-theatre capability requirements. However, due to space constraints in this conference paper, we will focus exclusively on the “Demand Surge” vignette, which provides a sufficient demonstration of the framework’s application.

Overview and Assumptions for the “Demand Surge” Vignette

This vignette explores a hypothetical future scenario of a demand surge for respirators. We assume that the DoD is currently at the supplier negotiations stage with a company called BestMasks (and other potential suppliers) and seeks to prepare an IP acquisition strategy for a possible future demand surge. The vignette sketch is as follows:

“Sometime in the future, US intelligence sources warn of an imminent chemical warfare threat from a large adversarial nation. To combat this threat, the DoD seeks to urgently ramp up the supply of personal protective equipment (PPE) for its troops. A core piece of PPE is a proprietary, best-in-class respirator mask that is additively manufactured by BestMasks. DoD has an ongoing contract with the OEM, BestMasks, for a small supply of respirators required for the Army’s Business-as-Usual (BAU) operations (e.g., regular training, emergency response), but the demand surge is expected to far outstrip current supply as well as the OEM’s maximum production capacity. A possible solution is to tap into the manufacturing capacity of other respirator suppliers to produce this mask, but these suppliers will need access to IP and other proprietary information owned by the OEM to achieve the high manufacturing precision required for the respirator to function.”

To apply the framework, we will make the following assumptions about the vignette:

- The demand surge is deemed temporary rather than a “new normal.”
- Although other respirator options exist, the OEM’s is deemed the best-in-class and most mission-appropriate model. Hence, the DoD wants to prioritize ramping up the supply of this specific product for maximum mission effectiveness.
- Both product and process IP exist and are required to enable high-precision production by alternative suppliers.
- The OEM owns all relevant IP.
- Alternative suppliers have worked with similar AM methods, materials, and products, such that:
    - They only require a short lead time to start production upon access to relevant IP and proprietary information.
    - With some trial and error, they can figure out the process, post-processing, and qualification parameters for the build.
- The IP required is not subject to invention secrecy protection or export control.
- The supply of filter cartridges is managed separately and not deemed an issue.
- Stockpiling masks is not favored due to high inventory and obsolescence costs.
- Reverse engineering will take too long due to the precision required and may also deter industry from developing IP for crisis-critical products since they risk losing it to the government.
- Hence, a fair IP compensation agreement upfront is desired to facilitate timely supply ramp-up and avoid stifling innovation for crisis-critical products during BAU operations.

Scenario Screening and Scoping

The relevant features of the acquisition scenario, as gleaned from the vignette setup and assumptions, are compiled in Table 2. In particular, we note that IP protection exists from the “IP Status” information. There is also latitude in considering IP compensation issues since the demand surge has yet to occur, and the DoD is not yet in crisis management mode. Hence, this use case has met the scenario screening conditions for framework applicability.

Table 2. Scenario Scoping: Demand Surge Vignette

Scoping Category	**Scenario features**
OEM Status	**Active** / ~~Inactive~~
Manufacturing Status	**Ongoing** / ~~Discontinued~~ • The production capacity of OEM alone is **sufficient during BAU operations** • **Additional capacity needed** to meet demand surge
Sourcing	~~Single-source~~ / **Multi-source** • Other respirator options exist, but this OEM is deemed **best-in-class** and the most **mission-appropriate**
IP Acquisition Requirements	All IP is required for alternative suppliers to **produce** and **qualify** the respirators. • Respirators require **high-precision manufacturing** to function properly • Not "new demand" but "demand surge": assume DoD already has required IP/knowledge on respirator use, maintenance, and proper disposal • The supply of filter cartridges is managed separately and is not an issue • Suppliers can figure out process, post-processing, and qualification parameters with some trial and error
Mission Status	National priority to ensure the safety and effectiveness of troops • Demand surge is deemed **temporary** rather than a "new normal." • Stockpiling of masks is not a favored option due to high inventory and obsolescence costs. • Fair IP compensation agreement upfront will facilitate **timely supply ramp-up.**
AM Capability Location	~~In-theatre~~ / **Out-of-theatre**
IP Rights Status	Both product and process IP exist, and OEM owns all relevant IP • General manufacturers require access to relevant IP and proprietary information to attempt production • Assume the lead time to start production is short once manufacturers have access to relevant IP • IP is **not subject to invention secrecy or export control** • Fair IP compensation agreement upfront will **avoid stifling innovation for crisis-critical products** during BAU operations

AM Lifecycle and IP Asset Identification

Next, we need to identify the parts of the AM lifecycle and IP assets that could be relevant for acquisition. From the "IP Acquisition Requirements" information in Table 2, we can infer that:

- IP supporting production and qualification of the respirator mask are required, while
- IP that only supports sustainment and/or end-of-life management is irrelevant since the DoD already has the required know-how from BAU operations.

This means that for this acquisition scenario, the relevant parts of the AM lifecycle are design/plan processing, production, post-processing, and testing/qualification (Figure 10).

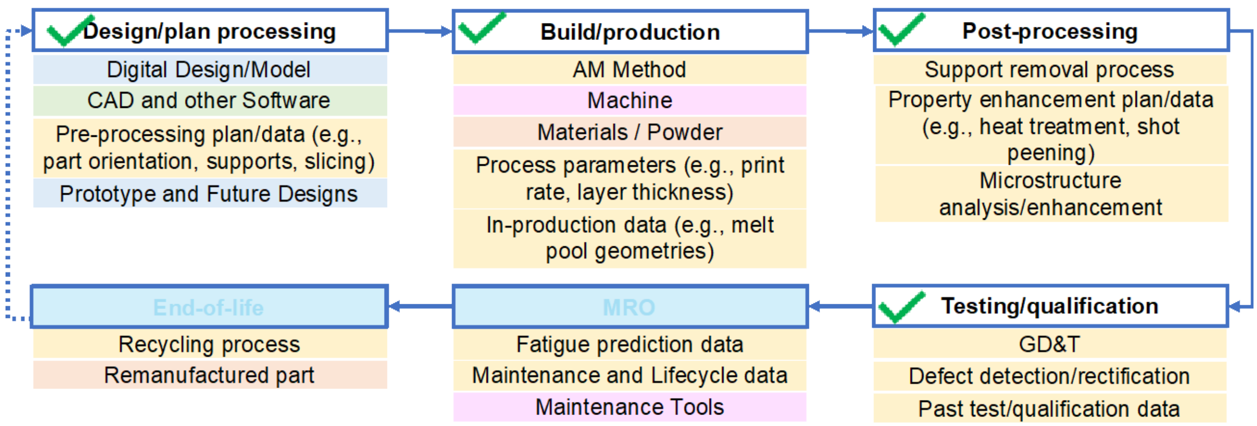


Figure 10. Relevant Parts of AM Lifecycle: Demand Surge Vignette

AM IP Strategy Considerations

The third step is to ascertain the key IP strategy considerations based on information from the scenario scoping step (Table 2).

Cost of Inaction

We start with the cost of inaction. From the “Sourcing,” “Mission Status,” and “IP Status” information, we can infer that if no IP acquisition is carried out:

- OEM may withhold critical IP/proprietary information from alternative manufacturers, causing delays and gaps in supply ramp-up.
- DoD can consider augmenting the supply with alternative but inferior respirators.

Hence, while the cost of inaction is not as high as it would be if there were no alternative respirator options, there can still be a non-negligible negative impact on troop health, safety, and mission success. This provides sufficient impetus to work through the rest of the considerations to construct an appropriate IP acquisition strategy.

Acquisition Scope

Next, we need to identify the necessary vs. good-to-have IP assets. For this vignette, we assumed that the new suppliers (e.g., other respirator makers) had worked with similar AM methods, materials, and products such that they were able to figure out the process and post-processing parameters, albeit with some trial and error. Hence, we could roughly classify the necessary vs. good-to-have IP as follows (Table 3):

Table 3. IP Classification: Demand Surge Vignette

Necessary	Good-to-have
Digital design/model, AM method/machine/materials	Design software, pre-processing plan/data, prototype designs, process parameters, in-production data, support removal process, property enhancement data, microstructure analysis/enhancement, Geometric Dimensioning and Tolerancing (GD&T), defect detection/rectification, past test/qualification data

From the “IP Acquisition Requirements” and “Mission Status” information, we know that:

- The respirators require high-precision manufacturing to function properly.
- A timely supply ramp-up is desired.

We thus infer that having more information – including the “good-to-have” proprietary production information – would enable new suppliers to achieve the required manufacturing precision more quickly (e.g., less trial and error with the process parameters). This could provide a competitive edge for the DoD, so a comprehensive acquisition scope would be preferable.

When to Acquire

From the “Manufacturing Status” and “IP Status” information, we know that:

- OEM’s production capacity is sufficient to meet demand during BAU operations today.
- Manufacturers only require a short lead time to start production once they have access to the relevant IP, so there is no need to acquire IP far ahead of time to build system capability or train manufacturers.

Hence, it would suffice to prepare a contingency option today to buy the IP when needed in the future rather than to acquire the IP right now.

IP Use Duration

From the “OEM Status,” “Manufacturing Status,” and “Mission Status” information, we anticipate that:

- The demand surge is assessed to be temporary.
- Additional manufacturing capacity is only required during the demand surge.
- OEM’s production capacity is sufficient to meet demand before and after the threat.

Hence, extra manufacturing capacity for the respirator masks is only needed if the chemical warfare threat and the need for U.S. countermeasures remain elevated. The DoD can rely on the OEM’s manufacturing capacity before and after the temporary threat. It would thus be appropriate to consider acquiring or leasing the relevant IP for the expected duration of the mission.

IP Sublicensing Rights

For a demand surge scenario, we can expect the government to want as much manufacturing capacity as possible (e.g., PPE needs during the COVID-19 pandemic). The “AM Capability Location” information also indicates the demand for outside-of-theatre manufacturing needs that the DoD itself can fulfill. It would thus make sense for the DoD to acquire IP sublicensing rights for the flexibility to contract as many alternative suppliers as necessary to ramp up supply.

IP Exclusivity

From the “OEM Status,” “Manufacturing Status,” and “IP Status” information, we can infer that:

- The OEM will continue to be active and contribute as a respirator supplier, requiring continued access to its own IP.
- The IP is not deemed so sensitive or secretive that the DoD needs to tightly monitor and control which entities have access.

Hence, a non-exclusive IP license should suffice in this case.

Summary

Figure 11 summarizes the IP considerations of this demand surge vignette.

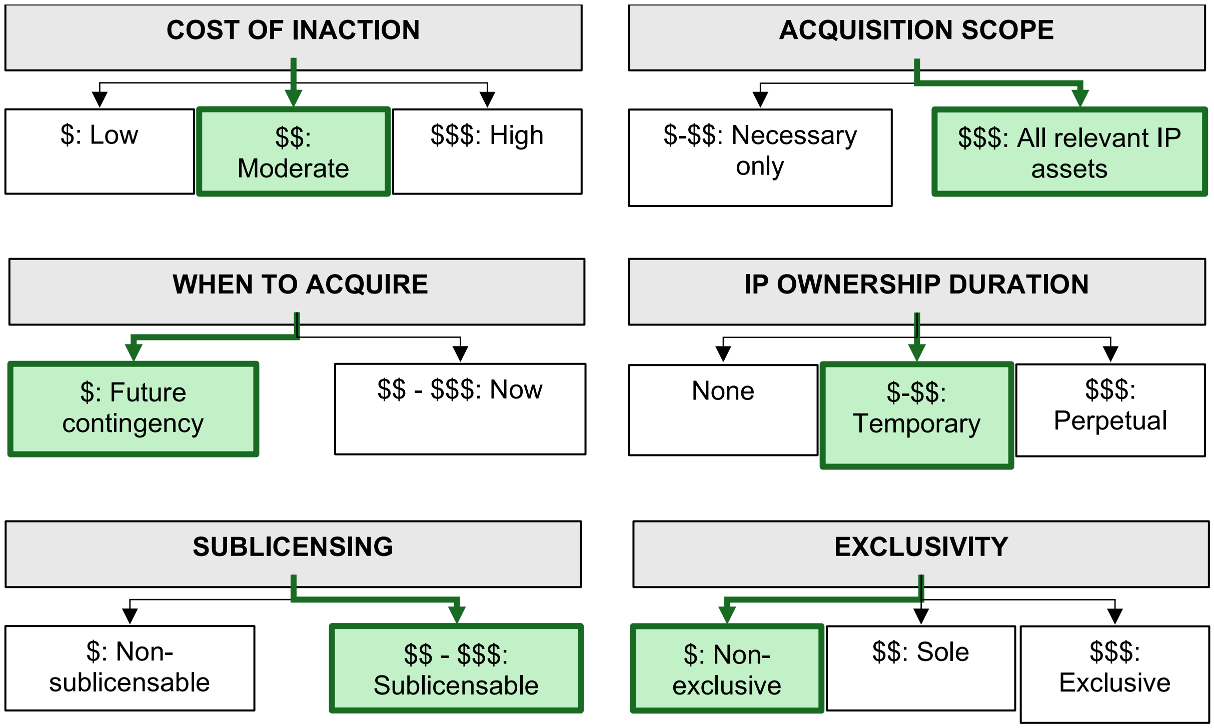


Figure 11. Decision Tree Output: Demand Surge Vignette

AM IP Strategy Formulation and Sensitivity Analysis

Figure 12 shows the acquisition option for the demand surge scenario based on the decision tree output in Figure 11. The recommendation here is to acquire from the OEM, as a buy option that can be exercised in the future, the temporary, sublicensable, and non-exclusive rights to a comprehensive set of IP and proprietary information pertaining to the production and qualification of respirator masks.

Finally, we examine the following variations in assumptions for sensitivity analysis:

Product Substitutability: If the respirator masks were not strictly best-in-class such that fully substitutable goods could be provided in sufficient quantity to meet the demand surge, then the cost of inaction could be significantly lower. The DoD could tap into other suppliers to augment the mask supply using other makes and models that also worked. In this case, there might be no need to own any IP, and thus, no need for an acquisition strategy.

Manufacturing Complexity: Suppose the manufacturing process involved complex and niche capabilities such that alternative suppliers needed a significant lead time to develop the human and system capabilities required for production; in that case, more planning might need to go into the timeline for acquisition. A straightforward option would be acquiring the required IP at the point of contracting and ramping up the strategic manufacturing capabilities required for these crisis-critical products in DIB companies. The tradeoff is higher upfront and retainer costs. Alternatively, suppose the required lead time can be reasonably estimated, and suitable signs that forewarn the onset of the demand surge can be identified. In that case, these can inform the DoD of the appropriate time to exercise the option. Admittedly, the telltale signs of chemical or biological warfare may occur too close to the threat to enable the development of significant manufacturing capabilities. Hence, this option may be more applicable to other demand surge scenarios where trends are more obvious, such as those induced by climate change or population growth.

IP Sensitivity: If the IP is deemed sensitive, such that the DoD needs to exert tighter control on its distribution and use, then a non-exclusive licensing arrangement might not work. Sole or exclusive licenses could then be considered. That said, the sensitivity of the IP will likely align with the sensitivity of the product, such that these patent secrecy and export control issues might be better dealt with outside the acquisition contract.

Acquisition Timeframe:	$$
❑ Future contingency	$
✓ Now	$$$
✓ Temporary (e.g., mission duration)	$-$$
❑ Perpetual	$$$
Acquisition Modality:	**$$**
❑ Non-sublicensable	$
✓ Sublicensable	$$$
✓ Non-exclusive	$
❑ Sole user	$$
❑ Exclusive	$$$
Acquisition Scope:	**$$$**
✓ Limited	$-$$

Figure 12. Acquisition Option: Demand Surge Vignette

Discussion: Additional Complexities and Research Implications

In this project, we have developed a decision framework to determine an IP acquisition strategy for AM systems. While the framework addresses the core considerations of IP acquisition and provides a qualitative decision-making approach, additional complexities merit further study. Future work in these areas could enhance the functionality and applicability of the framework. Specifically, an in-depth analysis of the following areas could significantly improve the framework's usability:

Interface With Existing Acquisition Rules and Processes

While this project was undertaken as a greenfield effort to develop an AM IP acquisition framework, it would be useful to examine how this framework could be adjusted to support and enhance existing defense acquisition frameworks, rules, processes, and decision support systems. For example, integration with the Planning, Programming, Budgeting, and Execution (PPBE) process could ensure adequate resource allocation and financial planning for AM IP acquisition strategies. The AM IP acquisition framework could support the Joint Capabilities Integration and Development System (JCIDS) process in identifying and prioritizing IP acquisition that fills critical AM capability gaps. The framework could also be refined to apply to different acquisition categories, in line with the Defense Acquisition System (DAS) classification, where more expensive programs are subject to more stringent oversight and consideration. Finally, future work should explore and ensure general framework alignment with the Federal

Acquisition Regulation (FAR) and the Defense Federal Acquisition Regulation Supplement (DFARS).

Portfolio-Level Acquisition Decisions

The decision-making process for IP acquisition often centers on their interconnected nature, where acquiring one type of IP can necessitate the acquisition of related IP to ensure full functionality. For example, the DoD might acquire manufacturing process IP that inherently requires the additional acquisition of software IP, like topology optimization. This interconnected acquisition strategy not only highlights the dependency of various IP on one another but also sets the stage for extending these capabilities across multiple missions and throughout the organization. The framework could also support Integrated Acquisition Portfolio Review (IAPR) processes by leveraging the model-based approach to defense portfolio management (DeLaurentis, Panchal, et al., 2024; Tsutsui, Guariniello, et al., 2023).

Sometimes, the interdependencies among different IP are not as straightforward as expected, forcing decision-makers to select from a constrained set of valuable IP due to limitations such as time, budget, capacity, or technological capabilities. For the DoD, this could mean choosing among various sensor technologies that vary significantly in cost, strategic value, or compatibility. For example, the DoD may be more inclined to rely on a single supplier for critical components, thereby reducing procurement flexibility and potentially leading to increased costs if cheaper or more advanced alternatives become available later. This situation is similar to that of consumers entrenched in Apple's ecosystem, where products like laptops, watches, and tablets are designed to work best together, encouraging continued investment within the same brand.

Uncertainty/Risk Quantification to Price Real Options

The current framework provides valuable inference from a qualitative decision-making perspective. However, there is a need to develop a specialized suite of software tools to streamline the IP management process quantitatively to make it more efficient and accessible.

One such tool is based on incorporating uncertainty quantification into the decision-making process to enhance the robustness of IP strategies, allowing for better risk management and more informed choices (Figure 13). These recommendations will contribute to refining the AM IP framework, ensuring the methodology remains adaptable and practical in various defense acquisition contexts. For example, utilizing the current decision-tree structure of the IP considerations and adding random events (e.g., risks, market, and change in status quo) relevant to the scenario allows the DoD to simulate and assign probabilities to outcomes. Simulating outcomes and their likelihood informs the DoD on which options strategy is the best suited or likely to lead to successful acquisition and deployment of the IP.

Additional Considerations for Framework Enhancement

Other considerations that would also be useful to explore are as follows:

IP and Data Qualification: Compared to physical assets, the quality of IP or digital assets (e.g., digital design) may be harder to verify. There may also be dependencies on human or system capabilities to use the assets effectively. This creates an impetus for an IP and data qualification process to validate the integrity of the IP and digital assets. One possible modality is a short post-acquisition "warranty" period where the IP seller must provide transitional support to ensure the usability of the acquired assets. The format and extent of this transitional support would need to be clearly defined upfront. It may also affect the pricing of the acquisition option.

Liability Implications: A related consideration is liability. When IP acquisition results in changes to the supply chain, it could also affect the traceability of liability. For example, suppose the products manufactured using the acquired IP were subpar. In that case, it might be challenging to ascertain whether the fault lay with the IP and digital assets (i.e., OEM's oversight) or how these assets were interpreted and used (i.e., new supplier's incompetence). In addition, a supply change could void any existing insurance policies on the equipment. The liability implications of IP acquisition should also be identified and, if possible, quantified as an acquisition consideration.

IP Compensation Quantification: The decision framework we have developed primarily considers the utility of acquisition to the government based on factors like cost of inaction. This value sets an upper bound for the acquisition price but may differ from the IP compensation that the government eventually pays. The IP compensation amount is dependent on several external factors. For example, the IP owners will have their valuation of the worth of their IP. Competition (or the lack of) from owners of similar IP assets could also affect the market value of the IP. The IP compensation amount may thus need to be determined through a negotiation process with one or more potential suppliers, considering the value to the government and these external factors. A thorough analysis may also be possible using game or auction theory concepts.

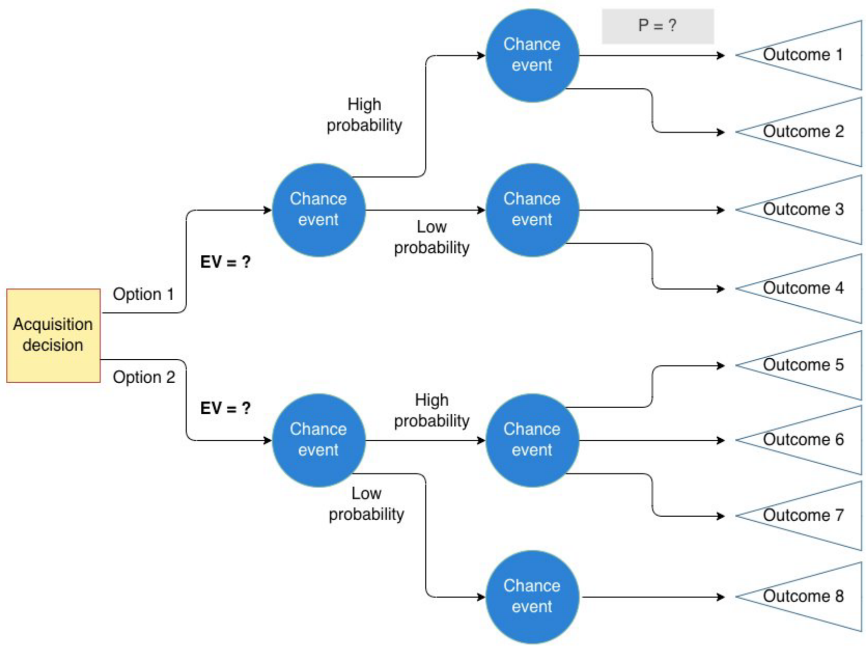


Figure 13. Integration of Uncertainty Quantification in Decision-Making to Enhance the Robustness of IP Strategies and Improve Risk Management

Future Research for Framework Validation and Integration

Future research could also include validating the framework for real-life defense acquisition processes. Implementing use cases based on previous acquisition processes, such as the Joint Light Tactical Vehicle (JLTV) acquisition, would provide valuable insights into what could be done differently. Also, integrating the IP decision framework with an AM decision framework based on real-life AM components (Shi et al., 2023; Tsutsui, Shi, et al., 2023) could provide a new dimension to acquisition research. Collaborating closely with the IP Cadre on current and future acquisition projects can enhance the framework's relevance and applicability. For instance, aligning and integrating the described framework with the DoD's mandatory acquisition pathways is crucial to ensure its practical implementation within the department. Consequently, additional research is required to integrate this framework into the broader Major Capability Acquisition process, ensuring it directly supports DoD acquisition strategies. In addition, examining the IP approaches of other nations, particularly those with innovative technologies developed under constrained budgets (Acquisition Innovation Research Center, 2021), and incorporating them as a part of the AM IP framework can offer helpful strategies and practices for the U.S. defense acquisition process.

Conclusions and Recommendations

IP rights are crucial for DoD acquisition, ensuring defense systems remain operational while avoiding vendor lock and rising sustainment costs. However, DIB companies view their IP as valuable assets, creating tension between the DoD's access needs and the industry's protection interests. The advancement of 3D scanning and AM technologies further complicates this dynamic, making effective IP management essential for defense acquisition and operational readiness.

The paper presented a comprehensive framework for navigating IP challenges in AM within defense acquisition, covering scenario screening, AM lifecycle analysis, IP asset identification, and strategy formulation. We explored the rationale for IP acquisition, identification of pertinent IP assets, and optimal structuring of IP agreements to ensure sustainable defense capabilities and competitive advantage. The framework's application was demonstrated through the "Demand Surge" vignette, examining IP management during extraordinary circumstances like imminent threats or pandemics.

Future recommendations emphasize portfolio-level acquisition decisions, uncertainty quantification, integration with existing DoD processes, and specialized software tools for better risk management. Implementing IP qualification processes, addressing liability implications, and establishing fair compensation frameworks through negotiation will ensure strategic value and successful deployment of acquired IPs across the DoD.

Acknowledgments

The authors acknowledge financial support from the U.S. Department of Defense (DoD) through the Systems Engineering Research Center (SERC)/Acquisition Innovation Research Center (AIRC) on research task WRT 1081.18, Contract No. HQ0034-19-D-0003, and Report No. AIRC-2024-TR-015. The authors are immensely grateful to Mr. Steve McKee for his vital support in advancing our research. Special thanks are also extended to Mr. George Winborne, Ms. Rosemary Solomon, Mr. Richard Gray, Mr. Mark Krzysko, and countless other DoD personnel (too many to enumerate) for their invaluable feedback. The authors thank the leaders of SERC/AIRC, including Dr. Dinesh Verma, Dr. Philip Anton, Dr. Douglas Buettner, Ms. Kara Pepe, and Ms. Tara Kelly, for their outstanding management of the funded project and their facilitation of the DoD-university partnerships.

References

Acquisition Innovation Research Center. (2021). *Acquisition innovation forum: Intellectual property (IP) valuation in the defense acquisition process*. https://acqirc.org/wp-content/uploads/2021/09/AIRC_IP-Forum_Summary_Final.pdf

Čirjevskis, A. (2024). Valuation of patent-based collaborative synergies under strategic settings with multiple uncertainties: Rainbow real options approach. *Journal of Risk and Financial Management*, *17*(4), 157.

DeLaurentis, D., Biller, S., Panchal, J., Shi, Q., Bekdache, D., Balasubramani, P., & Tsutsui, W. (2024). *Pilot program design to test innovative approaches in negotiating intellectual property* (AIRC-2024-TR-015).

DeLaurentis, D., Panchal, J., Raz, A., Tsutsui, W., Carpenter, D., Levin, W. Browne, D., Patterson, F., Arruda, J., Welz, Z., Inclan, E., Gray, M., & Geissler, T. (2024). *Advanced model-based tools for portfolio management and analytic* (AIRC-2024-TR-007).

DoD Instruction 5010.44. (2019). *Intellectual property (IP) acquisition and licensing*. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/501044p.pdf

GAO-22-104752. (2021). *DOD should take additional actions to improve how it approaches intellectual property*. https://www.gao.gov/assets/gao-22-104752.pdf

GAO-23-105850. (2023). *DOD needs department-wide guidance to inform acquisitions*. https://www.gao.gov/assets/830/827451.pdf

Gorovaia, N., & Windsperger, J. (2013). Real options, intangible resources and performance of franchise networks. *Managerial and Decision Economics*, *34*(3–5), 183–194.

Halt, G. B., Donch, J. C., Stiles, A. R., & Fesnak, R. (2017). Licensing of intellectual property rights for startups. In *Intellectual property and financing strategies for technology startups* (pp. 219–228). Springer International Publishing.

Hickey, K. J. (2022). *CRS IF10986: Intellectual property law: A brief introduction*. https://crsreports.congress.gov/product/pdf/IF/IF10986/3

Nugroho, L. A. (2016). Franchise ownership redirection: Real options perspective. *Financial Innovation*, *2*, 1–11.

Panchal, J., Atallah, M., Malak, R., Weaver-Rosen, J., Hartman, N., DeLaurentis, D., & Tsutsui, W. (2024). *Mission-aware integrated digital transformation for operational advantage* (AIRC-2024-TR-008).

Panchal, J., Tsutsui, W., Atallah, M., Malak, R., Hartman, N., & DeLaurentis, D. (2023). *Mission-aware integrated digital transformation for operational advantage* (AIRC-2023-TR-013).

Peters, H. M. (2022). *CRS IF12083: Intellectual property and technical data in DOD acquisitions*. https://crsreports.congress.gov/product/pdf/IF/IF12083

Shi, Q., Tsutsui, W., Walter, I., Panchal, J., & DeLaurentis, D. (2023). A decision support framework for additive manufacturing of space satellite systems. *IEEE Aerospace Conference Proceedings*.

Trigeorgis, L., & Reuer, J. J. (2017). Real options theory in strategic management. *Strategic Management Journal*, *38*(1), 42–63.

Tsutsui, W., Atallah, M., Malak, R., Hartman, N. W., DeLaurentis, D. A., & Panchal, J. H. (2024, May). Challenges and opportunities in enhancing Department of Defense ground vehicle capabilities through digital transformation. *2024 Annual Acquisition Research Symposium*.

Tsutsui, W., Guariniello, C., Mall, K., Patterson, F., Balestrini-Robinson, S., Panchal, J., & DeLaurentis, D. (2023). Model-based approach in defense portfolio management: Data preparation, analysis, and visualization of decision spaces. *2023 Annual Acquisition Research Symposium*.

Tsutsui, W., Shi, Q. A., Walter, I., Wei, A., Williams, C., DeLaurentis, D., & Panchal, J. (2023). Decision making for additive manufacturing in sustainable defense acquisition. *Naval Engineers Journal*, *135*(4), 47–58.

Tsutsui, W., Shi, Q., Bekdache, D., Balasubramani, P., Panchal, J. H., Biller, S., & DeLaurentis, D. (2024). *Intellectual property strategies for additive manufacturing in defense acquisitions*. 27th Annual Systems and Mission Engineering Conference, Norfolk, VA.

Vogel, B. J. (2016). Intellectual property and additive manufacturing/3D printing: Strategies and challenges of applying traditional IP laws to a transformative technology. *Minn. JL Sci. & Tech.*, *17*, 881. https://scholarship.law.umn.edu/mjlst/vol17/iss2/8

Weeds, H. (2002). Strategic delay in a real options model of R&D competition. *The Review of Economic Studies*, *69*(3), 729–747.

Widmer, M., & Rajan, V. (2016). *3D opportunity for intellectual property risk: Additive manufacturing stakes its claim*. https://www2.deloitte.com/content/dam/insights/us/articles/3d-printing-intellectual-property-risks/ER_2981-3D-opportunity-intellectual-property_MASTER.pdf

Wydler, V. (2014). *Gaining leverage over vendor lock to improve acquisition performance and cost efficiencies*. https://www.mitre.org/sites/default/files/publications/gaining-leverage-over-vendor-lock-14-1262.pdf

Integrated Digital Maturity Pathway for Technical Data Packages

Darryl Draper-Amason, PhD—is a leading expert in Performance Improvement, leveraging her PhD in instructional systems to enhance organizational effectiveness. Her research in maturity models, instructional design, and online education has transformed training effectiveness and readiness metrics. She specializes in Performance Improvement frameworks, guiding organizations through transitions and process optimizations. [ddraper@odu.edu]

Michael McGrath, DSc—is an independent consultant with broad government and industry experience in acquisition and technology management. He has served in VP positions and corporate boards at several technology companies. His government experience includes senior positions in OSD, DARPA, and in Navy as the DASN(RDT&E). He is a consultant for the University of Maryland Applied Research Lab for Intelligence and Security. His technology interests are in manufacturing, digital technical data and data analytics. Dr. McGrath holds a doctorate in operations research from George Washington University. [mmcgrat1@stevens.edu]

Abstract

A major challenge in technology transition is the Department of Defense's (DoD) requirement for a complete Technical Data Package (TDP) with intellectual property (IP) licenses for life-cycle use. Contractors often withhold proprietary details, including manufacturing trade secrets and sensitive IP, making TDPs incomplete or outdated. Negotiating royalties for items developed with mixed government and industry funding further complicates acquisition. As a result, industry seeks to protect its IP, while the DoD faces sustainment challenges due to limited access to critical data.

This research presents a case study from the National Defense Industrial Association's (NDIA) Digital Manufacturing Working Group (DMWG), exploring how digital transformation can address these challenges. The study applies the Integrated Digital Maturity Pathway (IDMP) framework, assessing how enterprises—suppliers, prime contractors, and government agencies—can achieve greater digital interoperability despite varying levels of digital maturity. The research evaluates use cases, process improvement roadmaps, and industry–government collaboration outcomes.

Findings highlight the potential for IDMP to enhance digital TDP practices, ensuring more effective data-sharing mechanisms. This study provides insights into the broader applicability of IDMP for digital transformation, with expected benefits in acquisition efficiency, data governance, and sustainment readiness across the DoD enterprise.

Introduction

The transition of technology in defense acquisition is contingent upon the availability of complete, interoperable, and life-cycle-ready Technical Data Packages (TDPs). These packages must reconcile contractor IP protection with DoD life-cycle needs. However, disparities in digital maturity among stakeholders, suppliers, prime contractors, and government entities complicate the effective development and use of TDPs.

This paper introduces the Integrated Digital Maturity Pathway (IDMP) framework as a novel approach to overcoming these challenges. By incorporating storytelling through user stories and leveraging the Air Force's VAULTIS framework, the research integrates diverse perspectives to address systemic issues in the development of TDPs and broader digital transformation efforts.

Research Problem and Question

Research Problem

The research issue is the challenge in transitioning technology due to the requirement for a complete Technical Data Package (TDP) and an IP license that allows life-cycle use. Contractors are often hesitant to provide proprietary or sensitive details, such as trade secrets, in the TDP, especially when items are developed with mixed funding sources (government and industry). These complexities, including royalty fee negotiations, create a protectionist stance from the industry, often resulting in the DoD receiving incomplete or outdated TDPs that fail to meet its needs.

Research Question

How can digital transformation, through an Integrated Digital Maturity Pathway (IDMP) framework, facilitate the creation of robust, interoperable TDPs that protect contractor IP while fulfilling DoD requirements?

Primary Research Sources: This study utilizes the National Defense Industrial Association's (NDIA) Digital Manufacturing Working Group (DMWG) data and analysis, focusing on digital transformation strategies for technical data interoperability. Primary sources include case studies of varied enterprises (suppliers, primes, government) at different stages of digital maturity. The research also incorporates industry and government feedback on implementing the IDMP framework to assess process improvements.

Background and Problem Statement

Technical Data Packages (TDPs) serve as the foundation of sustainment, spares manufacturing, and modernization efforts in defense acquisition. A TDP contains the essential technical information required to produce, maintain, and modify a system or component throughout its life cycle. These packages typically include engineering drawings, specifications, design models, and related technical documentation, ensuring that the DoD and its contractors have the necessary data to sustain critical systems independently of the original manufacturer. The TDP content is specified by MIL-STD-31000 (DoD, 2018) to include models, drawings, associated lists, engineering design data, specifications, standards, performance requirements, quality assurance provisions, software documentation, and packaging detail. Although the MIL-STD-31000 focus is on product definition data rather than manufacturing, the DoD often requires delivery of details of unique processes (i.e., not published or generally available to industry) when essential to design and manufacture. As used in this paper, the term TDP includes such details.

The availability of comprehensive, accurate, and up-to-date TDPs is essential for ensuring operational readiness, cost-effective maintenance, and long-term sustainment of defense assets. When properly structured and accessible, TDPs reduce dependency on original equipment manufacturers (OEMs), facilitate competitive procurement processes, and enable rapid response to mission-critical repair and sustainment needs. Furthermore, as the defense industry shifts toward digital engineering and model-based systems engineering (MBSE), the role of digitally mature TDPs becomes even more critical in integrating emerging technologies like additive manufacturing, artificial intelligence (AI), and predictive maintenance into sustainment strategies.

However, the efficacy of TDPs is frequently undermined by intellectual property (IP) challenges, incomplete documentation, and lack of interoperability across the DoD and its industrial partners. These issues present significant barriers to achieving the digital transformation necessary to modernize defense acquisition and sustainment operations.

Challenges of IP Protection and Contractor Reluctance to Share Proprietary Data

One of the most persistent challenges in TDP management is the tension between government data needs and contractor IP rights. Many defense systems are developed under mixed-funding models, where both private industry and the government contribute to research and development (R&D) efforts. When contractors invest private capital into product development, they seek to protect proprietary technical data, trade secrets, and competitive advantages by limiting the level of detail included in TDPs that might be disclosed to competitors.

This reluctance to share proprietary data often leads to incomplete or outdated TDPs, hindering the DoD's ability to maintain and sustain critical systems independently. Without access to comprehensive TDPs, the government is forced to rely on sole-source contracts, leading to higher costs, reduced competition, and increased lifecycle risks. Additionally, negotiating data rights and royalty fees for intellectual property developed under government contracts remains a contentious process, further complicating the development of interoperable, accessible, and lifecycle-ready TDPs.

The implications of inadequate TDPs extend beyond sustainment costs—they also impact the DoD's ability to leverage emerging manufacturing technologies such as additive manufacturing. Without sufficient technical data, DoD depots and sustainment centers cannot fabricate replacement parts, perform structural modifications, or integrate new capabilities, ultimately affecting mission readiness and supply chain resilience.

In this context, balancing contractor IP protection with the DoD's need for technical data remains a key challenge. Current acquisition policies and contract data requirements must evolve to ensure that contractors are incentivized to share critical data, while also safeguarding proprietary innovations that drive industry investment in new defense technologies.

The Need for the Integrated Digital Maturity Pathway (IDMP) to Address These Challenges

To navigate the complexities of TDP accessibility, IP protection, and interoperability, the Integrated Digital Maturity Pathway (IDMP) framework provides a structured digital maturity approach that systematically enhances the Department of Defense's (DoD) ability to manage, share, and utilize technical data. IDMP serves as a strategic roadmap for digital transformation within the TDP ecosystem, ensuring that data is Visible, Accessible, Trustworthy, and Interoperable across the defense industrial base while balancing the needs of contractors and government stakeholders.

The IDMP framework (see Figure 1) is designed to establish a progressive, scalable approach to digital maturity, allowing the DoD and industry partners to assess their current data capabilities and implement targeted improvements in TDP management. By integrating IDMP, the DoD can facilitate secure, interoperable, and life-cycle-ready TDPs that support sustainment, manufacturing, and next-generation digital engineering.

IDMP Level	Current State	Issues	Capabilities	Objectives
Initial Level: Current Process (Manual and Basic Processes)	The government specifies delivery of a Tech Data Package (TDP) for manufacturing components, but the TDP is often incomplete, causing issues for competitive procurement.	-Incomplete TDPs: Current processes result in insufficient technical data for competitive procurement. -Vendor lock: The government is often reliant on a single supplier due to incomplete TDPs. -Data Gaps: Vendors struggle with insufficient engineering details for manufacturing, leading to fewer competitive bids and higher risks of no-bid decisions.	-Establish basic TDP standards using MIL-STD-31000 to define minimum deliverables. -Initial efforts to incorporate basic TDP requirements into contracts to improve future procurement competitiveness.	Establish clarified TDP requirements and update contract deliverables to ensure complete, competitive TDPs for future procurement needs.
Initial Level: Decisions	-Determine minimum TDP requirements to avoid gaps that cause vendor lock. -Decide if manual processes for TDP validation are sufficient or if early digital initiatives should be introduced.			

Figure 1: Integrated Digital Maturity Pathway (IDMP) Level 1 User Story # 48 (Draper-Amason, 2024).

Key Benefits of IDMP in TDP Management

IDMP addresses data interoperability by enabling the development of standardized data formats, metadata protocols, and secure digital environments, ensuring that TDPs can be seamlessly integrated across government and contractor systems. It establishes a common digital maturity framework that highlights issues affecting interoperability between legacy and emerging digital systems.

When applied to TDPs, the IDMP framework highlights mitigating IP risks through secure data-sharing mechanisms by incorporating secure data-sharing models, including Digital Rights Management (DRM), role-based access control (RBAC), and blockchain-based verification. These mechanisms help contractors retain control over proprietary data while allowing the DoD to access necessary sustainment and life-cycle management information without compromising sensitive IP.

At higher levels of maturity, IDMP addresses additive manufacturing and AI-Driven Sustainment to ensure that TDPs are machine-readable and optimized for integration with additive manufacturing (AM), predictive maintenance, and AI-driven sustainment models. This enhances the DoD's ability to rapidly fabricate parts, conduct in-field repairs, and improve logistics planning through digital twin technologies.

The TDP IDMP provides a structured governance framework for TDPs, incorporating role-based access, digital audit trails, and compliance automation. This ensures that technical data remains authoritative, up-to-date, and protected across its entire life cycle. The framework also enables automated compliance verification to align with evolving DoD data policies and acquisition regulations.

Literature Review

The Role of Technical Data Packages (TDPs) in Defense Acquisition

Technical Data Packages (TDPs) are fundamental to the sustainment, manufacturing, and life-cycle management of defense systems. A TDP provides the necessary engineering documentation, including design specifications, drawings, manufacturing instructions, and quality assurance criteria, to support the production, modification, and maintenance of military assets. Without a complete and accessible TDP, the Department of Defense (DoD) and its sustainment partners face significant challenges in independently managing the life cycle of critical defense systems (McKay et al., 2021).

In defense acquisition, TDPs ensure that multiple vendors can competitively bid for contracts, thereby reducing sole-source dependencies and fostering a more resilient supply chain. For example, in the Amphibious Combat Vehicle (ACV) program, the U.S. Marine Corps has faced difficulties in identifying alternative manufacturers because the original contractor, BAE Systems, retains exclusive rights to the vehicle's technical data. As a result, prospective manufacturers must design and build vehicle variants without access to the TDP, increasing costs and technical risks (Wilson, 2023). This situation illustrates the critical role that TDPs play in ensuring that sustainment and future acquisitions remain cost-effective and adaptable to evolving operational requirements.

Existing Limitations in TDP Accessibility, Completeness, and Interoperability

Despite their importance, TDPs often suffer from accessibility and completeness issues, largely due to restrictions in intellectual property (IP) rights and the proprietary nature of many defense systems. A major limitation is that program managers (PMs) frequently do not procure sufficient data deliverables and associated data rights upfront, which can severely restrict the DoD's ability to sustain weapon systems over the long term (Harper, 2017). The Defense Acquisition University (DAU) emphasizes that acquiring the correct level of TDP access is essential for long-term sustainment but acknowledges that many program managers lack clear guidance on how to structure their procurement strategies for technical data rights. This has led to cases where the DoD is locked into long-term contracts with original equipment manufacturers (OEMs) who retain exclusive control over sustainment operations.

Another major issue is interoperability, particularly in large-scale programs that involve multiple stakeholders and contractors. Engineering design descriptions within TDPs are often developed in proprietary formats, making it difficult to integrate and modify data across different systems and sustainment environments (McKay et al., 2021). The complexity of maintaining configuration consistency across different versions of TDPs further complicates sustainment operations. Research highlights the need for improved digital product life-cycle management tools that can help engineers maintain TDP consistency while adapting designs to evolving mission requirements.

Additionally, many legacy defense systems were originally designed without consideration for future digital sustainment capabilities, making it difficult to apply modern manufacturing techniques such as additive manufacturing (AM) and AI-driven predictive maintenance. When TDPs are incomplete or not machine-readable, the ability to leverage advanced digital manufacturing techniques is significantly reduced. In some cases, defense sustainment depots have resorted to reverse engineering components due to the lack of available technical data, an expensive and time-consuming process that could have been avoided with proper TDP procurement strategies (Harper, 2017). The lack of accessibility, completeness, and interoperability in TDPs presents significant challenges to defense sustainment and manufacturing. As demonstrated in the USMC's ACV program, failure to secure full technical data rights from the outset can limit competitive procurement options and increase sustainment costs. Likewise, deficiencies in TDP structure and standardization hinder the adoption of modern digital sustainment technologies, making it harder for the DoD to fully leverage AI, additive manufacturing, and predictive analytics.

Digital Maturity and Data Interoperability Challenges

The increasing complexity of defense acquisition and sustainment demands a structured approach to digital maturity that enables organizations to assess their current capabilities, identify gaps, and implement targeted improvements. As the DoD advances toward digital transformation, the ability to integrate technical data, cybersecurity measures, and data-sharing protocols across different stakeholders remains a critical challenge. Without a standardized framework to guide this transformation, organizations risk data fragmentation, inconsistent governance, and cybersecurity vulnerabilities, ultimately undermining operational readiness and sustainment efficiency.

A structured digital maturity framework is essential for ensuring that data governance, security, and interoperability are addressed holistically. Kırmızı and Kocaoglu (2022) highlight that organizations benefit from a digital transformation maturity model that establishes measurable stages of digital adoption, allowing them to implement structured improvements rather than ad-hoc, reactionary changes. This model provides a roadmap for organizations to evolve from manual, siloed processes to fully integrated digital ecosystems that facilitate secure, efficient, and scalable data sharing. Within the DoD, the Integrated Digital Maturity Pathway (IDMP) serves as such a framework, guiding stakeholders through progressive stages of digital maturity that enhance TDP management, interoperability, and cybersecurity measures (Kırmızı & Kocaoglu, 2022).

Despite the potential benefits of digital maturity models, significant barriers to data interoperability persist across defense and industrial partnerships. Many legacy defense systems were not designed with modern digital architectures in mind, creating compatibility issues between older, proprietary formats and emerging model-based systems engineering (MBSE) standards. This lack of interoperability leads to data silos, where critical technical information remains locked within specific platforms, inaccessible to external stakeholders who require it for sustainment and modernization efforts. Research on data leakage prevention (DLP) and cybersecurity maturity further indicates that defense organizations struggle with balancing data accessibility with security, as proprietary data-sharing restrictions often hinder effective collaboration between DoD entities and private contractors (Domnik & Holland, 2024).

Compounding these challenges is the growing risk of cyber threats targeting defense networks and digital assets. Al Shidhani (2019) underscores the importance of cybersecurity maturity models, noting that organizations must progress through structured digital transformation phases to achieve secure and interoperable data environments. The DoD's digital maturity trajectory reflects these challenges, necessitating a comprehensive approach to data protection that ensures secure yet accessible technical data exchanges while mitigating the risks of data breaches, insider threats, and unauthorized access (Al Shidhani, 2019).

To address these issues, the IDMP framework provides a structured, security-focused roadmap for the DoD's digital transformation efforts. By incorporating cybersecurity best practices, governance policies, and data interoperability standards, IDMP enables the DoD to establish secure, standardized data-sharing mechanisms that protect technical data integrity while maintaining accessibility for authorized users. This framework not only enhances TDP management and supply chain resilience but also ensures that digital transformation efforts align with evolving cybersecurity and data governance regulations.

Through a structured approach to digital maturity and data interoperability, the DoD can transition toward a more resilient, secure, and data-driven sustainment strategy, leveraging IDMP to overcome the challenges of legacy system constraints, data silos, and cybersecurity vulnerabilities. As prior research suggests, the adoption of maturity models such as IDMP can serve as a critical enabler of digital transformation, ensuring that defense organizations remain adaptive, secure, and operationally effective in an increasingly digital landscape.

Addressing These Challenges with the IDMP Framework

The IDMP framework provides a structured approach to overcoming digital maturity and interoperability challenges within the DoD and its industry partners. By implementing progressive maturity levels, the framework systematically guides organizations through various phases of digital adoption, ensuring a measured and strategic transformation toward a fully integrated and interoperable data environment.

In addition to establishing a digital maturity roadmap, IDMP aligns with key data security standards by integrating cybersecurity principles and VAULTIS protocols, reinforcing data governance policies and risk mitigation strategies. This structured approach ensures that digital assets remain both accessible and secure against evolving cyber threats, allowing for controlled and compliant data management across different stakeholders.

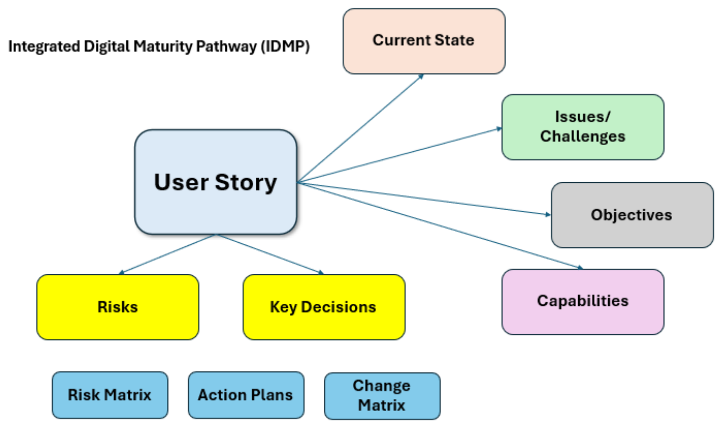


Figure 2: Integrated Digital Maturity Pathway (IDMP) Components (Draper-Amason, 2024).

To further enhance digital transformation efforts, the IDMP framework incorporates additional tools that aid in implementation, sustainability, and scalability. These tools include the Organizational Change Matrix, Risk Matrix, and Action Planning framework, each of which plays a critical role in guiding organizations through the challenges of digital transformation (Figure 2).

The Organizational Change Matrix assists organizations in assessing cultural and operational readiness for digital adoption, identifying potential barriers, change drivers, and necessary interventions to facilitate smooth transitions. By systematically analyzing stakeholder engagement, leadership support, and workforce adaptability, this tool enables organizations to strategically implement digital maturity models while fostering an environment conducive to long-term adoption and growth.

The Risk Matrix provides a structured method for identifying, assessing, and mitigating risks associated with digital transformation efforts. Given the complexities of technical data management, interoperability, and cybersecurity, this tool enables organizations to proactively address potential vulnerabilities, compliance risks, and integration challenges. Through risk prioritization and mitigation strategies, organizations can make informed decisions that minimize operational disruptions and sustainment risks during IDMP implementation.

Additionally, the Action Planning framework serves as a roadmap for execution, ensuring that digital transformation initiatives remain goal-oriented, measurable, and scalable. This structured approach supports milestone tracking, resource allocation, and performance evaluation, enabling organizations to sustain momentum and continuously improve digital capabilities over time. Action planning also helps to align IDMP implementation with broader DoD modernization objectives, ensuring that stakeholder coordination and policy compliance remain integral to digital maturity progress.

To further enhance secure data exchange, IDMP implements standardized, interoperable mechanisms that eliminate data fragmentation and security vulnerabilities. Through structured data-sharing protocols, the DoD and its partners can streamline access to critical technical data, improve collaboration between defense agencies and contractors, and strengthen supply chain resilience.

By leveraging IDMP as a structured digital maturity framework, the DoD can enhance TDP accessibility, improve supply chain efficiency, and enable secure, data-driven decision-making across defense acquisition and sustainment ecosystems. The incorporation of organizational change strategies, risk mitigation methodologies, and structured action planning ensures that digital transformation efforts are not only implemented effectively but also sustained and scalable. This comprehensive approach ensures that technical data remains both operationally effective and protected, supporting mission-critical sustainment activities while advancing the DoD's digital modernization goals.

User Stories

User stories capture end-user perspectives and operational challenges, providing a structured mechanism to define and prioritize capabilities within the IDMP framework. This approach allows for iterative refinement, ensuring that the IDMP remains adaptable to evolving technological and organizational needs.

As Cohn (2004) describes, user stories are simple, concise representations of system functionality expressed from the user's perspective, which help to drive agile development and ensure stakeholder alignment. They serve as a means of communication between developers, end-users, and decision-makers, helping to capture key functionalities in a format that is both understandable and actionable. The IDMP benefits from this approach by structuring digital maturity progression based on clearly articulated user needs, facilitating the alignment of technical capabilities with operational objectives.

User stories have been widely used in agile methodologies, such as Scrum and eXtreme Programming (XP), to facilitate a user-centered approach to software and systems development. Wautelet et al. (2017) emphasize that user stories serve as operational requirements representation models that drive transformation within a particular development paradigm. Their research highlights how user story's structure system development, allowing for an incremental and adaptable approach to digital transformation within frameworks like IDMP.

Beyond software engineering, user stories have also been used as a mechanism for measuring value in complex systems. For instance, research on nursing value user stories has demonstrated how this methodology can be applied to link specific actions to measurable outcomes, highlighting its broader applicability beyond software development. Similarly, within the IDMP framework, user stories can help quantify the impact of digital transformation initiatives by providing traceability from requirement definition to implementation outcomes.

The structured nature of user stories ensures that digital transformation initiatives, such as those encompassed by IDMP, are aligned with best practices in requirements engineering. Lucassen et al. (2016) outlines the INVEST framework (Independent, Negotiable, Valuable, Estimable, Small, and Testable), which defines key principles for writing high-quality user stories that improve clarity, prioritization, and testability. The integration of well-defined user stories into IDMP, organizations can enhance their ability to manage change, mitigate risks, and drive measurable improvements in digital capabilities (see Figure 2).

Ultimately, the integration of user stories within the IDMP framework provides a structured and scalable approach to digital transformation, ensuring that technological advancements are driven by stakeholder needs, operational requirements, and best practices in digital maturity modeling. The use of user stories as foundational elements of IDMP not only enhances requirements clarity and system adaptability but also ensures that digital maturity advancements remain user-driven, measurable, and aligned with operational goals.

These stories served as the foundation for the working group's application of the IDMP framework. The richness of these cases lies in their ability to highlight both commonalities and nuances, enabling the validation of the IDMP framework and the identification of targeted solutions.

The Role of the VAULTIS Framework

The Air Force's VAULTIS (Visible, Accessible, Understandable, Linked, Trustworthy, Interoperability and Secure) framework (Table 1) served as a guiding model for structuring and refining user stories. VAULTIS emphasizes the integration of advanced technologies, secure data-sharing protocols, and interoperability across diverse systems. Drawing from VAULTIS principles, this study adopted a structured approach to capturing user stories.

Table 1. VAULTIS Framework USAF

Visible	Ensuring that all critical data, processes, and dependencies are transparently available to stakeholders to support informed decision-making throughout the lifecycle of technical data management.
Accessible	Guaranteeing that authorized stakeholders can seamlessly retrieve necessary information when and where it is required, minimizing delays and barriers to effective use.
Understandable	Presenting data and processes in a format that is clear, consistent, and interpretable by both technical and non-technical stakeholders, ensuring alignment across diverse teams.
Linked	Establishing robust connections between datasets, systems, and processes to enable integration, reduce redundancies, and create a comprehensive digital thread for lifecycle management.
Trustworthy:	Building confidence in the integrity, accuracy, and authenticity of the data and systems to foster reliance on the framework for critical decision-making.
Interoperable:	Facilitating seamless communication and functionality across different systems, platforms, and stakeholders, regardless of varying digital maturity levels.
Secure:	Implementing rigorous safeguards to protect data and intellectual property from unauthorized access or misuse, ensuring compliance with contractual and regulatory requirements.

This alignment provided a robust foundation for developing user stories that address the specific challenges of TDP development and digital transformation.

Case Study Methodology in Digital Transformation Research

Qualitative research methods are used to uncover the direct actions and experiences of individuals in a social activity they carry out (Bryman, 2008; Mutch, 2005). Johnson and Christensen (2008) posit that these methods are valuable because they "view human behavior as dynamic and changing, and advocate studying phenomenon in depth and over an extended period of time" (p. 388). Qualitative approaches used in this study included working group discussions, a workshop, document review, user story contributions that include experiences of the working group participants and their direct actions associated with their experiences to enhanced understanding of a particular user story of the study inquiry.

Case studies are a powerful research methodology that allows for in-depth exploration of complex phenomena within real-world contexts. Yin (2014) posits that case studies are particularly valuable in addressing "how" and "why" questions, making them ideal for exploring the multifaceted challenges of digital transformation and TDP practices. According to Yin (2003), a case study design should be contemplated when four criteria are met. First, the answer to the “how” and “why” questions. Second, the individuals who are involved in the study cannot be manipulated. Third, to reveal contextual conditions with the belief that they are relevant to the phenomenon under investigation. Fourth, there are unclear boundaries between the phenomenon and context. For example, in this study, the case is the identification of user stories from the context of the government manufacturing community. It is within this setting that the user stories were developed and utilized. It would be impossible to have a correct picture of the development of user stories without considering the context.

NDIA and the Digital Manufacturing Working Group

The National Defense Industrial Association (NDIA) is a prominent 501(c)(3) educational nonprofit organization dedicated to promoting national security by facilitating collaboration among industry, government, and academia. The NDIA Manufacturing Division formed a Digital Manufacturing Working Group (DMWG) in 2024 to address digital transformation effects on manufacturing throughout the life cycle, with an emphasis on technology-enabled changes at the interface between industry and government.

DMWG User Stories Related to TDPs

To determine member interests, the DMWG captured user stories expressed as “As a (insert role), I want to (insert use of digital data), to achieve (insert benefit).” The initial set of 80 user stories ranged from interests in IP rights and TDP uses, to configuration management, to digital twins for manufacturing and supply chains, and to data analytics for Industry 4.0. An initial subset of five user stories was selected as the focus for a November 2024 workshop on manufacturing uses of technical data, as shown in Table 2.

Table 2. User Stories Related to TDP Delivery to the DoD

User story 16 & 20	As (16) a Supplier or (20) the Prime Integrator of a Supply Chain	I want to protect my proprietary data and that of my suppliers from disclosure or use outside specifically negotiated license provisions,	So that I can share data and models for collaboration within the supply chain and with the Government and ultimately provide better products and services.
User story 47	As a Prime Contractor	I want to offer Tech Data as a Service (TDaaS) as an alternative to data deliverables	So that my government customer will have secure access to up-to-date tech data for sustainment at the time of need.

User story 48	As a Government Product Support Manager	I want to get delivery of a complete tech data package with unlimited data rights	So that I can release build-to-print information for competitive procurement of spare parts
User story 49	As a Government Depot Manager	I want to have access to digital design and manufacturing data	So that I can accomplish depot repairs and fabricate parts as needed (additive and other mfg processes)

These user stories served as the foundation for the working group's application of the IDMP framework. The richness of these cases lies in their ability to highlight both commonalities and nuances, enabling the validation of the IDMP framework and the identification of targeted solutions. Workshop participants clarified user story boundaries, identified problems and mitigation strategies, and identified barriers to the widespread implementation of digital transformation ideas.

2024 Workshop Results

A central theme of the workshop was that DoD attempts to negotiate TDP deliverables and associated data rights years before specific sustainment needs are known. Consequently, the DoD tends to require that all relevant technical data be bought, ideally with unlimited rights, to be prepared for any need that may arise downstream. Industry lacks insight as to what the DoD will ultimately do with the TDP, makes the worst case assumption of disclosure to competitors, and fights hard to protect proprietary data. Digital engineering greatly expands the range of data that might be included in a TDP. At the same time, digital transformation offers new opportunities.

Ideas generated during the workshop centered on ways digital transformation can enable new ways of developing, protecting, delivering and using technical data. Examples include:

- User Stories 16 & 20: Start to implement role-based secure access and Digital Rights Management (like DRM in the music and movie industries) in Product Lifecycle Management (PLM) and other automated TDP systems. Evaluate prototype blockchain solutions to give prime contractors and suppliers more visibility and control over IP protection. Use generative artificial intelligence (Gen AI) to assist in proprietary markings in documents containing IP.
- User Story 48: Better define the TDP using MIL-STD-31000 Option Selection Worksheet and a 2-page Data Item Description (DID) rather than a 44-page DID. Include intended government TDP use cases in the RFP (as done by the Army Future Long Range Assault Aircraft) so the proposal can offer appropriate deliverables and data rights. Consider Data Escrow, with specified trigger events, to hedge against emergency use needs or loss of the original supplier. Allow spare parts bidders to negotiate use of IP directly with original manufacturers rather than relying on DoD TDPs. For software and firmware, use the commercial practice of Application Program Interfaces (APIs) rather than requiring source code.
- User Story 49: Update MIL-STD-31000 to include the minimum data needed for additive manufacturing. Build a centralized system for managing technical data packages (TDPs)

with version control, real-time updates, and integration with Additive Manufacturing (AM) platforms. Develop a database for AM material properties, tolerances, and stress thresholds specific to depot operations. Partner with manufacturers to enrich the database with certified material data. Use digital twins to simulate and test parts before fabrication and incorporate real-time feedback from digital twins to improve part accuracy and quality. Define and implement standards for data formats to ensure compatibility across all AM systems. Develop secure APIs for (1) seamless data exchange between depots and external suppliers, and (2) software/firmware associated with replacement parts to eliminate need for source code.

- User Story 47: Incorporate all these ideas into a future target called Technical Data as a Service (TDaaS). The TDaaS concept was well defined in a Naval Postgraduate School paper, "Technical Data as a Service (TDaaS) and the Valuation of Data Options" (Thompson & McGrath, 2019). Workshop participants agreed that the issues and ideas from the other TDP-related user stories could fit within the TDaaS framework, with centralized access to data (through PLM or other systems) as a near term implementation target and distributed access to data at its source as a longer-term target.

IDMP Roadmap

Post processing of the workshop results developed specific IDMP roadmaps for each user story and an overall roadmap for migration to the TDaaS concept (see Figure 3).

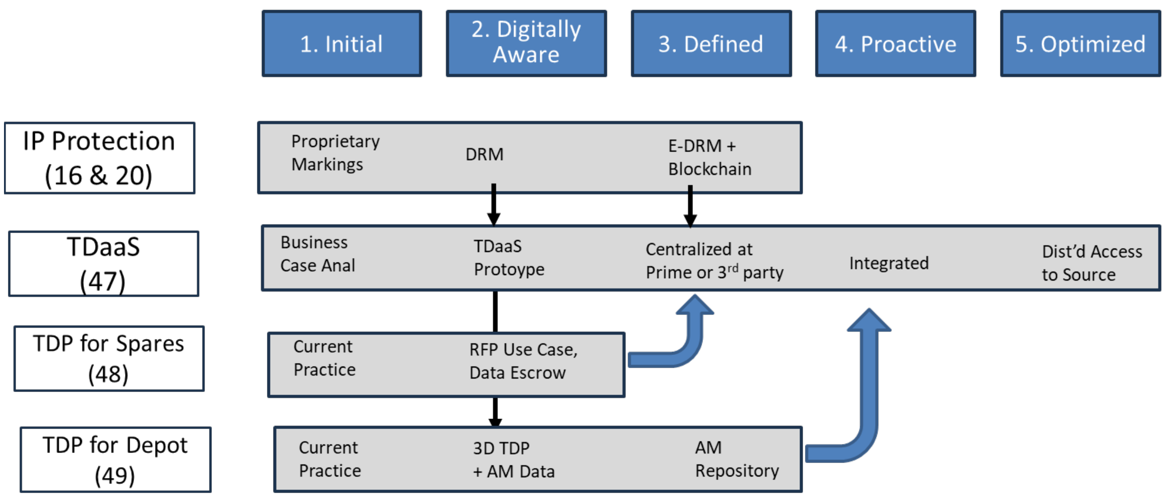


Figure 3: TDaaS Could Become the Target for Higher Maturity Levels

The NDIA DMWG is discussing this roadmap and the workshop ideas with government stakeholders to determine which ideas are ready for implementation and which need further demonstration and prototyping to resolve technical and business process uncertainties.

Findings

The findings of this research provide insights into the applicability of the Integrated Digital Maturity Pathway (IDMP) in addressing challenges associated with Technical Data Packages (TDPs), interoperability, and intellectual property (IP) protection within the Department of Defense (DoD) acquisition framework. Through an extensive analysis of case studies, working group discussions, and user story development, the research highlights three key areas of improvement: enhanced TDP practices, validation of the IDMP framework, and its broader applicability to digital transformation initiatives.

Enhanced TDP Practices

One of the most significant findings from this study is the need for improved TDP practices to facilitate DoD sustainment and acquisition. The Digital Manufacturing Working Group (DMWG) identified that current TDPs often suffer from incomplete documentation, lack of interoperability, and outdated technical data (DMWG, 2024). The analysis revealed that contractors frequently limit access to proprietary data due to IP concerns, forcing the DoD to negotiate for TDPs at various points in the product life cycle, often at higher costs and under restrictive conditions.

Findings from workshop discussions and user story development indicate that digital transformation, guided by the IDMP framework, can significantly improve TDP accessibility. The incorporation of secure digital rights management (DRM), blockchain verification, and structured data-sharing protocols allows for better protection of contractor IP while ensuring sufficient data availability for DoD sustainment and modernization efforts. Additionally, findings highlight the viability of a Tech Data as a Service (TDaaS) model, which shifts away from traditional upfront TDP acquisition toward a subscription-based access model that ensures the availability of up-to-date and relevant technical data as needed.

Validation of the IDMP Framework

The study validates the IDMP framework as an effective model for guiding digital maturity across government, contractors, and defense supply chain partners. Through its structured maturity levels, IDMP enables progressive digital transformation, allowing organizations to assess and enhance their technical data management capabilities. Case study findings show that IDMP's structured approach to interoperability, data governance, and security improves digital transformation efforts by addressing specific gaps in data-sharing policies and IP protections.

A key aspect of this validation is the role of user stories in shaping the IDMP framework. Monthly workgroup meetings generated 80 user stories, which were analyzed to identify common challenges in TDP management, digital rights, and additive manufacturing.

The application of user stories provided a direct mechanism for capturing stakeholder needs and aligning them with IDMP principles, reinforcing the practicality and scalability of the framework.

Broader Applicability of the IDMP Framework

The findings suggest that the IDMP framework has broad applicability beyond TDP management, extending to other areas of digital transformation within the defense industrial base. Case studies highlight its relevance in addressing challenges related to digital twins, data analytics for Industry 4.0, and configuration management. The analysis of user stories and working group discussions supports the scalability of IDMP to additional defense acquisition challenges, reinforcing its potential as a foundational framework for DoD digital modernization.

Additionally, the research identifies the importance of integrating emerging technologies such as AI, machine learning, and blockchain to further enhance digital maturity and sustainment capabilities. Applying IDMP principles to advanced manufacturing, predictive maintenance, and secure data exchange, the DoD and its industrial partners can create a more resilient, interoperable, and future-ready digital ecosystem.

Conclusion

The study's findings underscore the critical role of IDMP in transforming DoD technical data practices, validating its effectiveness in improving TDP accessibility, protecting contractor IP, and enabling scalable digital transformation efforts. The use of case studies and user stories demonstrates the practical applicability of the framework, while the integration of VAULTIS principles reinforces its alignment with broader DoD digital strategy initiatives. Furthermore, the research highlights the IDMP framework's potential for broader adoption, positioning it as a key enabler of future defense acquisition modernization efforts.

References

Al Shidhani, A. A. (2019). Cyber defense maturity levels and threat models for smart cities. *International Journal of Information Security and Privacy*, *13*(2), 32–46. https://doi.org/10.4018/IJISP.2019040103

Bertuca, T., & Judson, J. (2012). Legislation at play: Technical data package purchases re-emerge as important issue for Army. *Inside the Pentagon's Inside the Army*, *24*(14), 1–9.

Cohn, M. (2004). *User stories applied: For agile software development*. Addison-Wesley.

Department of Defense. (2018). *MIL-STD-31000B, Technical data packages*. http://everyspec.com/MIL-STD/MIL-STD-10000-and-Up/download.php?spec=MIL-STD-31000B.055788.pdf

Domnik, J., & Holland, A. (2024). On data leakage prevention maturity: Adapting the C2M2 framework. *Journal of Cybersecurity and Privacy*, *4*(2), 167–195. https://doi.org/10.3390/jcp4020009

Harper, J. (2017). Should *project managers buy technical data* (DTIC Technical Reports, AD1040333).

Kekeya, J. (2021). Qualitative case study research design: The commonalities and differences between collective, intrinsic and instrumental case studies. *Contemporary PNG Studies*, *36*, 28–37.

Kırmızı, M., & Kocaoglu, B. (2022). Digital transformation maturity model development framework based on design science: Case studies in manufacturing industry. *Journal of Manufacturing Technology Management*, *33*(7), 1319–1346. https://doi.org/10.1108/JMTM-11-2021-0476

Kuusisto, O., Kääriäinen, J., Hänninen, K., & Saarela, M. (2021). Towards a micro-enterprise–focused digital maturity framework. *International Journal of Innovation in the Digital Economy*, *12*(1), 72–85. https://doi.org/10.4018/IJIDE.2021010105

Lucassen, G., Dalpiaz, F., van der Werf, J. M. E. M., & Brinkkemper, S. (2016). Forging high-quality user stories: Towards a discipline for agile requirements. *Proceedings of the IEEE 24th International Requirements Engineering Conference (RE'16)*, 126–135.

Lucassen, G., Robeer, M., Dalpiaz, F., van der Werf, J. M. E. M., & Brinkkemper, S. (2017). Extracting conceptual models from user stories with Visual Narrator. *Requirements Engineering*, *22*(3), 339–358. https://doi.org/10.1007/s00766-017-0270-1

McKay, A., Rice, H. P., Chau, H. H., & de Pennington, A. (2021). *Maintaining consistency across design descriptions in engineering product development*. Cambridge University Press. https://doi.org/10.1017/pds.2021.460

Moon, L. , Clancy, G. , Welton, J. & Harper, E. (2019). Nursing value user stories. *Computers, Informatics, Nursing*, *37*(3), 161–170. https://doi.org/10.1097/CIN.0000000000000520

Poderi, G., Hasselqvist, H., Capaccioli, A., Bogdan, C., & D'Andrea, V. (2020). Matters of concerns and user stories: Ontological and methodological considerations for collaborative design processes. *CoDesign*, *16*(3), 220–232. https://doi.org/10.1080/15710882.2018.1557694

Ridder, H.-G. (2017). The theory contribution of case study research designs. *Business Research (Göttingen)*, *10*(2), 281–305. https://doi.org/10.1007/s40685-017-0045-z

Ross, N. J. (2015). Technical data packages: When can they reduce costs for the Department of Defense? *Defense AR Journal*, *22*(4), 450–471.

Shepherd, A. (2024). *Navy gains more access to F/A-18 technical data*. InsideDefense.Com's SitRep, http://proxy.lib.odu.edu/login?url=https://www.proquest.com/trade-journals/navy-gains-more-access-f-18-technical-data/docview/3031700409/se-2

Shepherd, A. (2024). Super Hornet technical data package access will allow Navy to be self-sufficient. *Inside the Pentagon's Inside the Navy*, *37*(32), http://proxy.lib.odu.edu/login?url=https://www.proquest.com/trade-journals/super-hornet-technical-data-package-access-will/docview/3091582955/se-2

Thompson, G. E., & McGrath, M. (2019). *Technical data as a service (TDaaS) and the valuation of data options*. Acquisition Research Program. https://dair.nps.edu/handle/123456789/2757

Vera-Rivera, F. H., Puerto Cuadros, E. G., Perez, B., Astudillo, H., & Gaona, C. (2023). SEMGROMI-a semantic grouping algorithm to identifying microservices using semantic similarity of user stories. *PeerJ. Computer Science*, *9*, e1380–e1380. https://doi.org/10.7717/peerj-cs.1380

Wautelet, Y., Heng, S., Kiv, S., & Kolp, M. (2017). User-story driven development of multi-agent systems: A process fragment for agile methods. *Computer Languages, Systems & Structures*, *50*, 159–176.

Wesley, R. D. (2021). Innovation at Mach5. *Army AL & T*, 37–42. http://proxy.lib.odu.edu/login?url=https://www.proquest.com/trade-journals/innovation-at-mach5/docview/2488113752/se-2

Wilson, N. (2023). USMC considers Add'l ACV maker, lacks technical data package rights. *Inside the Pentagon's Inside the Navy*, *36*(19). http://proxy.lib.odu.edu/login?url=https://www.proquest.com/trade-journals/usmc-considers-addl-acv-maker-lacks-technical/docview/2813539993/se-2

Time Value of Data Decision Modeling for Major Defense Acquisition Programs

Frank Goertner—is Director and Capstone Instructor for the Technology Management Graduate Program at the University of Maryland's Robert H. Smith School of Business. [fgoertne@umd.edu]

William Lucyshyn—is a Research Professor at the University of Maryland's School of Public Policy. [lucyshyn@umd.edu]

Contributors—Jon Crocker, Terrence O'Brien, Joseph Bailey, Wedad Elmaghraby, Harrison Hill, Ryan Huddleston

Abstract

This paper seeks to enhance the Department of Defense's (DoD) understanding of time-value associated with contracted data deliverables and intellectual property (IP), particularly as encapsulated in digital Technical Data Packages (TDPs) for Major Defense Acquisition Programs (MDAPs). Drawing on business contract theory, it examines "economic hold-up scenarios," where imbalances in transaction costs over terms, assets, or IP in controlled, specialized, or evolving markets create challenges. The Defining the Problem section defines the problem as it confronts DoD practitioners today, building on insights from past economic and business research; the How Hold-ups Have Been Addressed in Other Industries section investigates how other industries have tackled hold-ups, situating DoD's challenges within the wider U.S. market; and the Implications for DoD TDP Contracting in MDAPs section evaluates the implications for MDAPs, integrating theoretical frameworks with practical case studies. The paper concludes by proposing a decision model to implement mitigation strategies in future DoD MDAP contracts, accompanied by suggestions for further testing and research to refine this model.

Defining the Problem

Department of Defense's Time Value of Data Challenge

The pricing and procurement of parts for the acquisition and sustainment of the Department of Defense's (DoD) large, complex weapon platforms (MDAPs), regularly precipitate overlapping economic challenges. The specialized nature of components, regulatory structure of defense contracting, and competing interests around intellectual property (IP) rights within MDAPs are factors that commonly distort efficient market transactions between what is often a single DoD customer and a single available supplier.

These challenges are markedly evident in transactions involving *Technical Data Packages (TDPs),* and the resulting inefficiencies are particularly prevalent in their valuation over time. Transacted as a component of the original design, but then essential for both the remanufacture and maintenance of the parts they accompany, TDPs carry some economic value for both the supplier and customer for as long as the platform they comprise remains in service. However, the value they offer at the moment of production, and at any given time thereafter, can vary widely according to part type and fluctuate considerably due to operational conditions, shifting maintenance requirements, or supply chain developments.

This results in a multi-faceted dilemma for many MDAP acquisition strategies. The Defense Federal Acquisition Regulation Supplement (DFARS) mandates that DoD procuring activities develop an acquisition strategy for all major programs and weapons systems prior to solicitation that accounts for projected technical data use over the system's entire life cycle (DFARS 207.1, 2024). Furthermore, the DoD directs program offices to acquire essential IP deliverables and license rights at "fair and reasonable prices," ensuring that the DoD can sustain and upgrade systems throughout all program production, maintenance, and sustainment phases (OUSD A&S, 2019). Yet how should DoD executives value TDP ownership and IP access for millions of physical parts,

[1]each with its own predicted life cycle and idiosyncratic variables of operational necessity, produced in scattered manufacturing markets likely to change over the decades-long lifespan of a major weapons platform?

Pricing informed by the manufacturer alone can be costly—inviting maximalist estimates of future expenditures. However, a strategy informed by the DFARs alone may be too generic. It is absurd to think a TDP for a part like a data cable in an aircraft, ship, or submarine—used and replaced often but relatively simple to manufacture with relatively generic IP—should be priced and contracted equivalently to parts in the same platform's specialty propulsion unit—rarely replaced and extraordinarily complex to manufacture with IP potentially at the highest levels of national protection. It is equally absurd, though, to expect DoD procurement professionals to craft individualized valuation guidance for each part and TDP.

Today's DFARS dictates that DoD procuring activities strategize for price efficiency at scale. Yet it offers scant guidance for its decision-makers on execution-level tactics and tools that advance that goal. A look beyond defense contracting may help. What the DoD and its contractors routinely confront is what economic and business researchers refer to as a ***hold-up problem.*** Fortunately, they have also devised several options to mitigate it.

Dynamic or discriminative time-value modeling, real options contracts, and pooled IP access solutions are tools increasingly utilized in the private sector and select government agencies. Each, in its own way, seeks to remedy inflexible contracts and limited IP rights that impede long-term, cost-effective sustainment strategies.

What Is a Hold-up Problem?

According to contract economics theory, a *hold-up problem*[2] emerges when two parties refrain from efficient cooperation because of imbalances in their bargaining power. Hold-ups involve two factors: (1) a requirement for non-contractible specific investments prior to the transaction and (2) uncertainty between parties on the exact form of optimal transaction (e.g., quality, number of units, time of delivery; Rogerson, 1992). These conditions exacerbate the inherent challenges of incomplete contracts (Aghion & Holden, 2011), particularly in markets characterized by high levels of information asymmetry (Lofgren et al., 2002), monopoly power (Lerner, 1934), or consumer monopsony (Weintraub, 1949).

The products and services most at risk of hold-up problems are those with significant asset specificity, which refers to the degree to which investments in a specific transaction for a specific purpose retain value above and beyond their use for any other purpose (Williamson, 1981). Nonspecific assets, whether transacted recurrently or occasionally, typically have external competitive forces of supply and demand sufficient for the commercial market to govern price to some measure of certainty for all parties. However, the more idiosyncratic use a product or service has for a customer, the more asset specificity it assumes.

As products or services become more asset-specific, incentives toward more idiosyncratic and more hierarchical contract governance also grow. Outside markets may not exist for either supply or demand, meaning parties to the transaction only have each other as sources of information and price governance. When these asset-specific transactions are

[1] While challenges explored here may pertain to digital and physical parts, our scope is limited to physical hardware.

[2] Foundational works on *hold-ups* include *Markets and Hierarchies: Analysis and Antitrust Implications* (Williamson, 1975); *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process* (Klein et al., 1978); and *Transaction-Cost Economics: The Governance of Contractual Relations* (Williamson, 1979)

definitive and expected to occur only once or infrequently, both the customer and supplier tend to have clear incentives to share information and cooperate toward fair contractual pricing. Considering the stakes and cost of the process, private parties often acquire the assistance of a third-party arbiter. However, since it is expected to be a one-time cost with a high reward to both sides, such trilateral governance can often mitigate, or at least balance out, hold-up problems.

Yet when asset-specific transactions are recurrent or long-term, or if asset specificity of the goods or services fluctuates or is disputed due to changes in the external market, hold-up problems can become far more prevalent. Trilateral arbitration of each iterative transaction is often too costly, devolving into bilateral or unilateral contract governance with considerable space for opportunistic information asymmetries and pricing (Williamson, 1979).

Transactions in aerospace and defense industries are a prime example of such transactions. Many involve products or services with little to no value beyond the specific purpose for which they are designed. Their value is highly specific to a single monopsonistic government consumer. However, they demand extraordinary levels of pre-transaction information and investment, limiting viable suppliers to as few as one contractor with considerable monopoly power. Moreover, while the initial contract for the design and manufacture of the first version of a defense platform may be strictly governed, the recurrent transactions for maintenance and sustainment of its mission-specific parts may not. The resulting bargaining imbalances and information asymmetries incentivize contractual relationships riddled with uncertainties around information sharing and price adjustments over time.

Economic hold-ups are common across the DoD because the conditions favoring hold-ups are prevalent in the procurements it pursues. That said, not all hold-ups are created equal.

Types of Economic Hold-ups in the DoD

The highly specific missions supported by specialized platforms that define military procurement, the deep levels of pre-transaction investment in parts and TDPs by uniquely equipped defense contractors that enable it, and the inevitable uncertainty of future market disruptions for both sides offer favorable conditions for hold-ups in the DoD to occur and recur. As outlined above, these hold-ups can be expected to be most prevalent in defense contracts involving recurrent transactions for assets (i.e., products) of high or mixed specificity and insufficient governance of information sharing or market valuation over time.

The relationship between DoD and MDAP contractors in designing, producing, maintaining, and sustaining major weapons systems offers each of these conditions. However, that does not imply there is a single hold-up problem demanding a single analysis or solution. A major weapons platform, such as a Navy submarine or Air Force strategic bomber, is composed of a vast number of parts and processes (i.e., products) that have varied levels of uncertainty and governance. The result of hold-ups for each may be the same—transactional inefficiency—but the solution will likely differ according to its prime factor within it.

For this study, therefore, we identify **three types of hold-ups relevant to DoD contracting, categorized by the primary** factor underlying each (Figure 1).

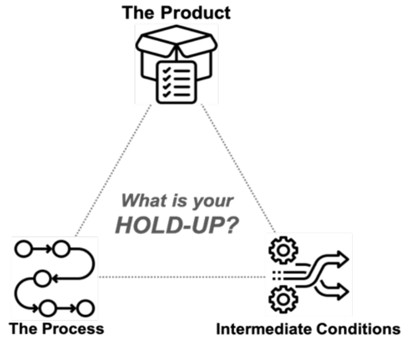


Figure 1.

Product Hold-up: A *Product Hold-up* is a hold-up that derives from asset specificity of a type of product anticipated to retain its specificity in transactions across all current markets and all envisioned future markets. This may be because the type of product is regulated (e.g., radar absorbing materials), the mission for which the product is useful is restricted (e.g., classified programs), or the product is used in a unitarily select purpose for which no alternative market is foreseen to ever exist (e.g., undersea nuclear deterrence). In this case, both a permanent monopoly and monopsony exist, fully disassociating free-market dynamics from influencing product life cycle valuation or transaction leverage between product customer and supplier.

Intermediate Hold-up: An *Intermediate Hold-up* exists when a type of product is transacted as asset-specific in current markets (i.e., there are no known commercial alternative suppliers or alternate uses today). However, the system or its components could be viably transacted under different competitive conditions in future or reimagined markets. In this case, the current market accommodates monopolistic and/or monopsonistic power. However, there are no technical or regulatory constraints to innovations or market disruptions that could rebalance future transactions. Perhaps a new, previously non-existent commercial market is emerging, other DoD systems could be designed to use the same part, or new technology like 3D printing would permit new supply alternatives.

Process Hold-up: A *Process Hold-up* is a hold-up that derives from the specificity of the process by which a type of product is transacted rather than any inherent product characteristics. There may be alternative market uses for the product (i.e., competitive demand) or viable alternatives for its manufacture (i.e., competitive supply). Still, the government's transaction process stipulates a unique and specific variant of the product or means by which it must be transacted. In this case, the government's monopsonistic demand for its product and transactions prompts the hold-up rather than the supplier's monopoly over the type of product itself. Perhaps the part is readily available in a commercial off-the-shelf version, or technology already exists for alternative manufacturing methods. However, government regulations either prohibit, fail to incentivize, or insufficiently describe accounting for them in contract pricing and negotiations.

These three categories are not exhaustive of all types of hold-ups, nor are they mutually exclusive. A hold-up can be driven by more than one factor. However, assessing and assigning the primary source and category are necessary first steps in devising potential solutions.

Equally important, time plays a role. A hold-up now may not last forever. Categorization of a hold-up type for a given product may not be static or permanent. For example, a part or its TDP subject to a *Process Hold-up* in an acquisition program's design and production phase could be recategorized as an *Intermediate Hold-up* as the program shifts into a maintenance and sustainment phase.

Time may even change the contractual parameters of the part or TDP across these phases. A bundle of components considered a single contractual product during the design and production phase could be disaggregated into multiple sub-components of contractual products during the maintenance and sustainment phase of the program. In other words, neither assets nor terms of specificity may be assumed to remain unchangeable across the entire life cycle of a platform and program. Any hold-up analysis should be considered a snapshot in time, subject to reframing and often open to disruption.

Options for Mitigating Hold-up Problems

The seemingly intractable nature of hold-ups is rooted in the economic reality that the value of products and services within a market is inherently dynamic. Valuation can rarely be perfectly determined by parties in advance or accurately represented in a static contract between them.

One approach to addressing this hold-up conundrum and rebalancing transaction asymmetries over time is the use of an options contract, an agreement between two parties that facilitates a *potential* transaction involving a contractually defined asset at a *preset* price and date (Corbin, 1914). Options reduce pre-contractual uncertainty while preserving flexibility for both the consumer and supplier to adjust toward an optimal transaction as information mediates risk over time. They offer a negotiated right, but not an obligation, to purchase or sell components of the transacted product or services, the value of which is projected to fluctuate.

While options contracts have been used in various industries to address hold-up problems, their effect derives from the balanced leverage of multiple parties in an otherwise open, competitive economic market. Consequently, this paper limits its exploration of options contracts and their utility to DoD hold-ups characterized as *Process or Intermediate Hold-ups.*

This does not imply that an asset in a *Product Hold-up* for which specificity of use or design is the prime hold-up factor could *never* be considered for an option contract or any other solution explored in this paper. As discussed previously, the nature of a product may change as time and acquisition phases shift. A part and its TDP may be a single highly specific asset subject to a *Product Hold-up* during platform design and production. With time, market evolution or progression into platform maintenance and sustainment can be recategorized as they become less specific. The part or TDP may have the same name, and a hold-up may remain, but the reason, type, and solutions to mitigate it will have changed. In other words, just because a part is not suitable for an option now does not mean it can never be considered for an option later; however, that alone would not invalidate our description of applicability. The key is to retain focus on categorizing the market around the part rather than fixating on any one label for the part itself.

The application of options to mitigate *Intermediate Hold-ups* is reviewed as regularly applied private sector use cases to draw lessons for challenges to the DoD scenarios.

A second approach to disrupt *Process Hold-ups*, particularly those involving IP, is patent-pooling: an agreement among patent owners to license their intellectual property as bundles to each other or third parties (Reisinger & Tarantino, 2019). In highly dynamic but regulated industries such as entertainment, information technology, and medicine, there is evidence that creative agreements among parties to pool patents can serve as a tool to counter innovator hold-up problems without curtailing technological progress (Baron & Pohlmann, 2015).

While U.S. government agencies have not traditionally established their government-sponsored patent pools, the National Institutes of Health has uniquely partnered in select patent pools to alleviate hold-ups for medical technologies critical to public health (National Institutes of

Health [NIH], 2022). The applicability of this approach beyond health technologies is also explored.

Finally, a third approach is summarized that could mitigate Intermediate Hold-ups and Process Hold-ups through a proposed shift in DoD policy to transact for access, rather than ownership, of TDPs. Termed Tech Data as a Service (TDaaS), this approach can be employed alone or in conjunction with an option or patent pooling arrangement.

However, it is important to acknowledge that since hold-up categories are not mutually exclusive, options to mitigate them may not be as well. An optimal contractual solution could incorporate more than one approach outlined above and below (Figure 2).

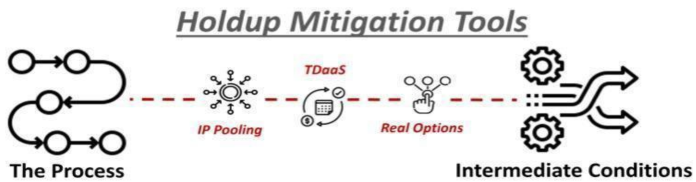


Figure 2.

How Hold-ups Have Been Addressed in Other Industries

Real Options: Pricing the Future of Tangible or Intangible Assets[3]

When considering how to use options contracts best to resolve hold-up problems, it is informative to first briefly review how options have been used in other applications and industries to solve similar problems. Traditionally, options contracts are financial instruments that provide individuals and institutions with opportunities to manage risk, speculate on market movements, and enhance portfolio performance over time. At their core, options contracts grant the holder the right, but not the obligation, to buy or sell an underlying asset at a predetermined price within a specified timeframe. This flexibility makes options contracts useful tools for mitigating transaction risk and planning for market uncertainty in the future. By offering the potential for leverage, diversification, and strategic positioning, options contracts empower market participants to tailor their risk exposure and optimize financial objectives with precision. Whether used by investors seeking to hedge against adverse price movements, traders aiming to capitalize on short-term fluctuations, or companies looking to mitigate future risks to their operations, options contracts have become a widely used and studied tool in the private sector.

Several key elements influence the pricing of options contracts, each playing a crucial role in determining the value of these derivatives. The first input is the *Current Price* of the underlying asset, as options derive most of their worth from an asset's performance as valued within current market conditions. The second input to the option's value is the *Strike Price*: the predetermined price at which the option holder can buy or sell the underlying asset. *Time Until Expiration* is the third input, with options typically losing value as expiration approaches due to diminishing time value. A fourth input, *Volatility*, reflects the market's expectation of future price fluctuations.

By analyzing these elements within shared pricing models, such as the Black-Scholes model (a mathematical model used for pricing options), investors and traders can align assumptions and calculations of time-value and risk across a diversity of products over a variety

[3] Foundational works on *Options Pricing* include *The Pricing of Options and Corporate Liabilities* (Black & Scholes, 1973); *Theory of Rational Option Pricing* (Merton, 1973a); and *An Intertemporal Capital Asset Pricing Model* (Merton, 1973b)

of time-spans (Black & Scholes, 1973). The math can be complicated, but its premise is simple. The agreed *Current Price* of any product transacted today is founded on shared assumptions about market dynamics, or *Volatility*, from now into the future. Yet there is also a reasonable probability that those assumptions are wrong, with the value of that probability corresponding to how well those assumptions match reality as the future nears the present. If market dynamics inflate the product's value beyond expectations, time offers a premium on behalf of the buyer. The seller enjoys the discount if market dynamics depress the product's value below expectations. However, if each agrees up-front to the *Strike Price* that adjusts value for both at a milestone between now and the future—the *Time Until Expiration*—each can split the risk according to their forecasts of how and when the market may change. The contractual option offers a path through an intermediate hold-up caused by uncertainty or doubt (Figure 3).

Though options contracts were born in real estate and financial markets, their applicability has proven useful beyond traditional investing and trading domains. One notable example of how non-financiers have adopted options principles is *Real Options.* In *Real Options*, options are considered more than investment opportunities, but rather *any* real-world opportunity. Analysis of *Real Options,* predicated on some version of the same four input elements—*Current Price, Strike Price, Time Until Expiration*, and *Volatility*—associated with capital projects or business ventures is employed to evaluate investment decisions in uncertain environments. By treating managerial choices as options, real options analysis enables decision-makers to assess the time-value of flexibility and adaptability in strategic planning.

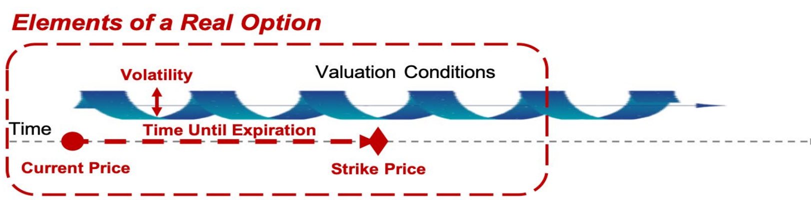


Figure 3.

Patent Pooling: Reimagining the Process of IP Stewardship

Alongside, or possibly between, tangible and intangible assets in the private sector market of valuation and trading stands IP. In an intangible sense, IP represents the valuation of creativity and access to innovation. Yet, a deeply rooted federal governance mechanism—IP patenting—has imbued it with a sense of tangible value nearly akin to real estate. Like a property deed, one can see, even frame, a patent.

In many ways, patents have evolved into their own class of products. A patent may be associated with a tangible or intangible asset it begets. However, it also represents some measure of value on its own—value to the inventor that expended specific investments prior to its award as well as value to a partner that wants to employ its utility, perhaps as or potentially beyond what that inventor envisioned. Both sides want the *current price* of access to the patented IP to reflect the hopes and needs they foresee for the future. However, by definition, future returns are unprecedented and, therefore, uncertain. Hold-ups in IP are common and complex.

One tool that can mitigate intellectual hold-up is embedded options in licensing contracts to access or use IP over time. These work similarly to any of the examples above, albeit with the product patented technology rather than land or a digital coin.

Particularly to patents, however, another tool has emerged to mitigate hold-ups in the transaction of intellectual rights: *patent pooling.*[4] We live in a period of breakneck discovery and high-stakes disruption. One result has been a massive increase in awarded patents in recent decades. Another has been an explosion of litigation between competitors with similar technologies. Such conflict in court is often a byproduct of Process Hold-ups that derive from governance or government procedures that struggle to keep up with the complexities and quantities of transactions a market demands.

As a type of technology advances in time, increasingly complex products need to draw on a widening span of IP for their design and production. This can drive *royalty stacking,* a market inefficiency where a single product must bear multiple royalty burdens to satisfy licensing requirements across complimentary but dispersed IP. This, in turn, can drive patent trolling or patent hoarding, whereby firms pursue strategies to monopolize patents to maximize their leverage in hold-ups.

Patent pooling preempts the incentive to prompt such hold-ups. Within patent pools, two or more patent owners agree to share access to their IP and the potential to license it to others jointly. This enables innovative technologies to be developed by more producers at less cost, accelerating the commercialization of the IP along with future market conditions of shared benefit to its owners, producers, and customers. If options contracts offer pit stops or offramps on the road from today's market to that of the future, patent pools repave the road to smooth and quicken the ride.

For example, within the biomedical industry, patent pools have been used successfully to advance technological progress in the creation of tests and medications for both HIV (Lampe & Moser, 2016) and COVID-19 (World Health Organization [WHO], 2023). The resulting market for new treatments and drugs has benefited consumers while profiting manufacturers and even seeding ground for new adjacent markets the technology can spur. That said, their efficacy on hold-ups depends on the motives of the poolers. If collectively, instead, they prefer the future to be slowed, they can also be used to defend the status quo. This was the case for movie film manufacturers in the United States in the early 20th century when Technicolor and Kodak pooled patents in collusion to inhibit the development of technology that could erode current pricing conditions for their high-revenue products (Lampe & Moser, 2016). In this case, the pool was a place for the hold-up to fester.

Tech Data as a Service (TDaaS): New Process and Options for Data Access

A final approach explored in and around the DoD with potential application in addressing hold-ups is a proposed shift in policy to transact for access, rather than ownership, of intangible assets: an approach coined by some as *Tech Data as a Service (TDaaS)*.[5] TDaaS aims to meet DoD challenges in the acquisition and management of digital assets such as TDPs at a sustainable cost that accounts for the value of those assets at the time they are needed.

Within the current DoD procurement process, government acquisition professionals are required to maximize the purchase of TDPs and associated data rights during the design and production phase of a platform. The goal is to assure availability to minimize risk in anticipation

---

[4] Scholarly works on the evolution and impact of patent pools include *Patent Pooling and the Anti-Trust Laws* (University of Chicago Law Review, 1950); *The Design of Patent Pools: The Determinants of Licensing Rules* (Lerner et al., 2007); and *Patent Pools, Competition, and Innovation—Evidence from 20 US Industries under the New Deal* (Lampe & Moser, 2016)

[5] For a seminal work exploring the adoption of TDaaS for the DoD, see *Technical Data as a Service (TDaaS) and the Valuation of Data Options* (Thompson & McGrath, 2019).

of future uncertain needs across the life cycle of the parts and platforms they support. Clearly, the ingredients are all there for a hold-up problem to endure.

What if, instead, DoD contracted with the inventors and suppliers for priced access and use of that data when needed? Similar to Software as a Service (SaaS), which has become ubiquitous among contracts in commercial IT (Mell & Grance, 2011), TDaaS contracts could break down bilateral hold-ups by changing the process governing the transactions in which they occur.

There are four specific potential advantages to the adoption of TDaaS. The first is to allow for quick and accurate purchase, lease, or access to TDPs and their digital subsets as the needs of the government customer arise. This “pick and choose” method of continuous procurement allows for monetary savings by avoiding an all-or-nothing up-front approach, as is current common practice. Second, the government gains adaptability to future needs by allowing a method to keep the door open with the contractor in an environment of uncertainty regarding future data needs. Third, it allows for more dynamic price modulation according to the changing conditions of shifting market or operational conditions over the lifespan of a platform and its parts. Finally, it incentivizes the contractor to maintain and update TDPs throughout the entire system’s life cycle, ensuring the DoD has access to the most current part specifications.

The importance of addressing such DoD hold-ups and the prospect of federal-level change as it relates to TDPs and data rights in general is evidenced in their inclusion in the National Defense Authorization Act for Fiscal Year 2024, H.R. 2670, 118 Cong. (2024).

This all points to one certainty on the path from today’s market to the ones that will follow. More tools are needed for the government and its partners to price their transactions over time amidst rapidly changing conditions and governance.

Implications for DoD TDP Contracting in MDAPs

Each of the industries, cases, and approaches presented in the previous section are unique. However, components of each offer insights into how the DoD may mitigate future hold-up costs as it negotiates rights over TDPs within MDAPs.

As described in the Defining the Problem section, the DoD must contend with a variety of economic hold-ups in the acquisition of parts and access to associated TDPs necessary to build and sustain major weapons platforms such as submarines or bombers over a multi-decade lifespan. The dynamics of rapidly evolving technologies, industries, and markets compound the inherent complexities of assessing net value and negotiating fair prices for a mix of tangible parts as well as intangible digital and IP assets within government contracts regulated by the DFAR. Standard practice is for the many hold-ups that derive to be confronted and cemented en masse in a few or single MDAP contracts for platform design and production.

Are Real Options an Option for DoD Intermediate Hold-ups?

The DFARS establishes that the DoD can include negotiated options in MDAP acquisition strategies. Despite this, real options are rarely employed. In fact, our research could not identify a DoD major weapons systems program to have used contracted real options for component parts to any considerable degree.

The most intractable obstacle DoD acquisition professionals face in employing options for parts or TDPs appears to be in their pricing. As outlined in the How Hold-ups Have Been Addressed in Other Industries section, options pricing in the private sector relies on quantifiable valuation measured in transparent, competitive markets. Within the Black-Scholes Model, shifts in market forces and factors impacting the *Current Price* can be combined with *Price Volatility* projections under calculated probabilities and assumptions across the *Time Until Expiration* to generate an optimal *Strike Price.*

In the monopoly-monopsony market of many DoD MDAPs, pricing works differently. For example, the DoD contract for the Navy's new Columbia Class Ballistic Missile Submarine (SSBN) is an Integrated Product and Process Development (IPPD) contract (DoD SAR, 2023) with a single standard for pricing, a Cost Plus Incentive Fee Approach (AFCEA, 2020). The *Current Price* is calculated as a sum of allowable supplier costs plus a negotiated fee, which is adjusted by a formula comparing total allowable costs to total target costs. However, identifying life cycle data needs early in the development of a program like this can be challenging. Deferring some amount of payment for TDP access and maintenance to an optionable future date would change the timing and accounting of contract deliverables, allowing the government to access only necessary TDPs when they are needed. That could generate efficiencies, but those may not translate into net dollars saved.

In order to adapt the hold-up mitigating potential of real options contracts to DoD contracting in this case, the Navy would need tools to account for forces and factors beyond supplier cost that may impact the price valuation of contracted components over time. One example of such forces and factors is monetary inflation. In the simplest construct of "risk-free" options pricing, interest rates alone are projected across the *Time Until Expiration* to establish the future *Strike Price*. This enables both parties to account for their own assumptions about inflationary pressure on the cost of capital and price over time.

The use of options by the DoD to hedge against inflation alone is unlikely. DoD acquisition policy for Cost Plus Incentive Contracts already controls for inflation versus profit (OUSD[A&S], 2022). Yet inflation is not the only force or factor to impact prices over time, especially in technology manufacturing and support industries such as this.

Trends in technology development and adoption suggests economic forces and technical factors can be expected to place both upward and downward pressure on per unit cost, and therefore price, over time. Even in closed, non-competitive monopoly-monopsony markets such as the current one between the Navy and its sole supplier of nuclear submarines, General Dynamic Electric Boat (GDEB), the benefits of scale should pressure marginal costs per unit down over time. One-time fixed design costs, such as TDP creation, along with recurring fixed production and sustainment costs, such as TDP maintenance, may be distributed across a larger set of priced transactions as the fleet of platforms grows and ages. Traditionally, this force derived from a scaling economy may have seemed irrelevant within a DoD program sourced solely by its own defined requirements, and U.S. government allocated funds. However, the potential for economic scale over time beyond pre-programmed requirements suddenly enters the equation as the manufacture and transfer of nuclear submarine parts to other countries, such as Australia and the UK, becomes possible (Australian Government, 2024). On the other hand, risks that the U.S. submarine industrial base cannot meet scaling demand or that demand for infrequently replaced parts is too sporadic to sustain subcontractor cash flow over time could inject new scarcities that pressure costs up.

Forces of technological change can add pressure as well. Complex systems with precisely engineered mechanical parts may be simplified or enhanced with new alloys and materials. Parts that must be molded and milled today could, in time, be additively manufactured either by contractors or directly by the Navy. Processes that demand skilled human labor today may be automated in the future. Various technological advances of various types are poised to disrupt wide swaths of the U.S. manufacturing industry in ways and at speeds we can foresee but not forecast with precision. In some cases, these changes may drastically reduce the cost of remanufacturing parts. In other cases, they may expand the breadth of eligible suppliers or supply techniques. They could even eliminate many of today's barriers and costs associated with forming a new business for a short-fused demand after an unanticipated market exit.

Each of these factors and forces is a source of potential asymmetric uncertainty about the future market between the Navy and GDEB or the DoD and any sole MDAP contractor. Were they in an open competitive market, these asymmetries would be balanced and distilled through competing bids by auction into an equitable *Strike Price* over an agreed *Time Until Expiration*. The contracted real option would then alleviate the hold-up by either side.

Elements of a Notional Real Option for MDAP Parts and TDPs

For the DoD and an MDAP supplier, bilateral negotiations between economically informed parties would need to set the price. To start, DoD acquisition professionals need variables they could independently quantify as cost-risk proxies in place of the *Price Volatility* used in Black-Scholes.

These variables would reflect the degree to which future forces and factors could drive the cost-value of a part or its TDP up or down in time. Next, they would need to assess the optimal *Time Until Expiration* for the real option according to their analysis of how rapidly those factors and forces will have a production or sustainment phase impact (Figure 4).

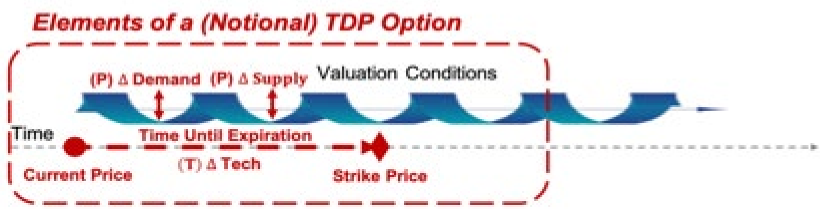


Figure 4.

Proposing specific sub-variables and equations is beyond the scope of this study and will be left to future research. However, extant research publications and private sector sources can offer a place to start. The potential monetary impact from changes in supply could be projected by assessing the savings the DoD could potentially realize in the maintenance and sustainment phase of a program, should they acquire both the rights and capabilities to reproduce the parts themselves. Additionally, scenarios for new manufacturing efficiencies or alternative suppliers could be modeled and assessed for their probability and impact on cost, either for part replacement or TDP reproduction. Like the private sector, a blend of historical data from predecessors or peer MDAPs and pro-forma analysis from industrial technologists and economists can be informative. In the case of Navy and GDEB, this could include data from the Ohio Class submarine program that the Columbia program is replacing.

The potential monetary impact from changes in demand could prove both simpler and harder. At a minimum, design and production phase executives would need to consult with maintenance and sustainment phase experts to assess the margin of error in replacement schedules for each major component and TDP. From our interviews with subject matter experts, this is something that already happens, but not always with the persistence, precision, and documented quantification this level of independent modeling would demand. The more challenging task may be forecasting potential changes in demand for the MDAP platforms over time. For example, the Navy's shipbuilding plan within any given budgetary cycle will always be the official projection of record for number of SSBNs demanded. However, the probability of future contract modification, such as the modification awarded in 2022 for Columbia Program expansion (General Dynamics, 2022), could theoretically be assigned with derivative effects projected on per-unit replacement price. As the United States reevaluates both force sizes and uses, such projections could be timely.

Finally, assessing the optimal *Time Until Expiration* could be the most complex variable to assign. Again, private sector business practices may offer some leads. Technology and market projections are available from both commercial and government sources. For the DoD, the Intelligence Community can help, as can technology consultants versed in enterprise transformation and industry-level valuation.[6] Summed and assessed, these variables could be used in options modeling prior to additional modification for MDAP design and production phase or additional contracts in maintenance and sustainment phases of a program.

Could a "TDP-Library" Circumvent Process Hold-ups?

In today's evolving technological landscape, many TDPs also represent commercially valuable IP, the rights for which can also lead to a process hold-up. When the DoD fails to acquire the necessary IP to operate and maintain its weapon systems, the hold-up often increases costs over time (GAO, 2021). The 2021 case of TransDigm offers a case study of how IP hold-ups can even be exploited to extreme ends (DoD-IG, 2021).

Applying private sector approaches to pooled IP management, three methods for consolidating and managing DoD TDPs and IP data rights merit attention: (1) IP pooling within the DoD; (2) establishing a non-profit IP consortium; and (3) delegation of IP governance to an independent commercial vendor.

IP Pooling within the DoD

One way to reconfigure the data-rights processes that prompt hold-ups in DoD transactions would be to establish a separate DoD program office tasked to pool shared IP management and TDP maintenance across all phases of MDAP design, production, maintenance, and sustainment. This is not a large deviation from the way the current process is supposed to function. However, the DoD office or agency, in this case, would be "pooling" IP and data access assurance as a "library" service independent of MDAP contract requirements. An advantage would be that familiarity with defense-specific requirements, protocols, and security measures would ensure compliance with federal regulations, reducing the risk of regulatory breaches and ensuring all data is managed according to defense and international trade standards.

Establishing a Non-Profit OTA Consortium

The Other Transactions Authority (OTA) framework has emerged as a flexible and streamlined approach to fostering innovation through partnerships between the DoD and the private sector. Through OTAs, a non-profit consortium could be established as a neutral entity dedicated to developing and implementing standardized data rights and TDP management processes. By bringing together key stakeholders from industry, government, and academia, the consortium would work to standardize processes, enhance data security, and facilitate innovation while ensuring the availability of the TDPs to all relevant partners.

This standardization would serve to level the playing field for current and future suppliers, reduce the administrative burden on industry and the DoD, and widen competitive innovation among small businesses and non-traditional contractors. It may be affiliated with or similar to the existing Defense Industrial Base (DIB) Consortium.

---

[6] This past year, the DoD announced a $2.4 billion contract for Deloitte, a leading accounting and consulting firm, to explore options to expand submarine workforce development as well as accelerate the development and adoption of more modern manufacturing supply chain techniques (Wilkers, 2024).

Delegation of Authority to a Commercial Vendor: An IP/TDP Escrow

Using a commercial vendor to manage the DIB IP and TDP library may be an effective alternate strategy. Contemporary vendors (e.g., Exostar) bring the latest technologies and specialized expertise, utilizing tools such as cloud computing, AI, and blockchain to enhance data management, security, and accessibility (Exostar, n.d.; Henderson, 2020).

This approach mirrors a tool widely used by both Amazon and Walmart—the IP escrow—in which a third party holds the vendor's data and data rights in an escrow account. If an original supplier goes out of business, discontinues the product, or fails to perform on the part of a contract, the buyer ensured sustained access to the data and data rights (Sander, 2022).

To implement an IP/TDP escrow for MDAPs, the DoD could require contractors to deposit comprehensive TDPs[7] into an escrow account when acquiring complex weapon systems (Figure 5). By securing these TDPs through escrow, the DoD ensures that the government can access the necessary information to sustain the system independently if the contractor is unable or unwilling to provide support. The DoD can protect its interests by adopting escrow agreements while fostering better collaboration with industry partners (Sander, 2022).

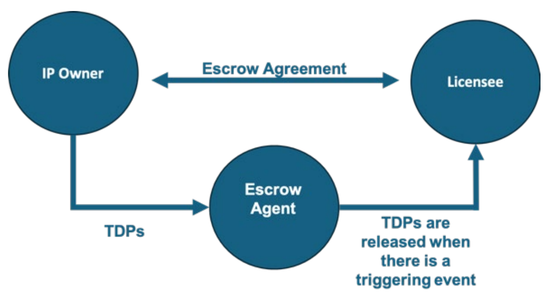


Figure 5.

Is TDaaS Worth Piloting in a Future MDAP?

The final approach introduced in the How Hold-ups Have Been Addressed in Other Industries section as a mitigation to some measures of both *Intermediate Hold-ups* and *Process Hold-ups* in MDAPs is a proposed shift to transact for access, rather than ownership, of intangible assets: considering *TDaaS*.

As outlined in our exploration of *TDP Libraries*, OTAs enable experimentation in contracting without requiring the rewriting of the DFARS. Yet, as discussed in the exploration of Real Options for MDAPs before, simply rescheduling payments of parts with prices locked by Cost Plus Incentive Fee contracts merely spreads out the impact of the hold-up rather than addressing it. TDaaS offers an "all of the above" approach that, in some cases, may prove to have the most effect.

Imagine a service-based contract for TDP access and maintenance that incentivizes the manufacturer or supplier and sustains their survival, but with a real option priced to account for the probability of use beyond the first most likely replacement period of the part as well as probability spread among new manufacturing techniques, supply chain efficiencies, and modified demand. If the part and its TDP include IP declared Government Purpose or Limited

[7] The TDPs can include models, drawings, specifications, performance requirements, and software documentation for system maintenance and future upgrades.

Rights, the option may account for the potential of a future OTA consortium or TDP library eclipsing the hold-up.

If the market or IP sharing arrangements evolve in the DoD's favor, the most efficient transaction for both sides could be for the Navy to execute the option to terminate the service agreement and either maintain and use the data itself or transfer it to the consortium. On the other hand, electing to forgo the option harms neither the mission nor the industrial base. The service contract would continue to economize both access and business support through the sustainment and maintenance phase of the program.

Of course, this could prove cost-prohibitive to negotiate and sustain for every part. However, for those with frequent replacement projections subject to hold up of acquisition processes known to be unsustainably unaffordable or risky, it is worth piloting to try. At worst, it will force cross-program collaboration and standards on acquisition and technological projections for critical parts and data as Columbia migrates from design to sustainment over the next half century or more. At best, it could pave the path to true innovation in MDAP acquisitions.

Which Tool for What Hold-up? A Proposed Decision Guide

Assuming the DoD adopts all of the above, the next question at hand is: Which one should apply for each type of hold-up originally identified within the Defining the Problem section? The challenge here is defining the dimensions of asset specificity with regards to part and TDP, and what approach is best suited to mitigate the hold-up to which it is subject.

First, it must be recognized that not all parts necessarily require special attention to the valuation of access or purchase to the rights to their design data or IP. Depending on the type of hold-up, the asset specificity of the part, and the economic and intellectual property considerations, a maximalist approach to government ownership or access rights is likely not practical. Second, it is important to acknowledge that a large number of parts in many MDAPs are likely best contracted for and acquired as currently done. Yet, if a subset of parts could benefit from the new Time Value of Data approaches outlined above, how can they be identified and matched to a solution? Another approach borrowed from corporate best practices may be the answer: a decision matrix.

Sketching a DoD Decision Matrix for Time Value of Data

Decision matrices serve as decisional guides rather than policies or procedures. Through a series of questions, the strategist is invited to dissect a complex, multi-faceted dilemma into addressable decision bins. Economists and corporate finance professionals often guide the way through acquisition decisions, financial modeling, and negotiations that could have enterprise-wide enduring effects.

A robust decision matrix for MDAP Time Value of Data decisions, along with more refined modeling of Real Options Pricing, will remain a rich area for research beyond this present project. We conclude, however, with some thoughts on what such a decision matrix could include. For the sake of this model, the matrix is represented as a "decision cube" (Figure 6).

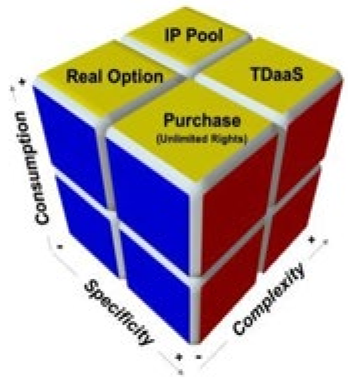


Figure 6.

Decision 1: What is the Hold-up?

This first question, more than any of those that follow, demands creative and consequential deliberation. At first pass, every part and TDP in an MDAP will likely seem fit for categorization as a Product Hold-up. If they are presently crafted and supplied in the monopoly-monopsony market and subject to previously defined data rights decisions, it is easy to assume that they could never be supplied or managed otherwise. The key is to think past current circumstances and ask if the current hold-up is grounded in forever exclusive and immutable conditions related to the type of part or its use.

<u>The first question is</u>: *Could you imagine any future in which the part or its components could be produced by alternate suppliers (including the DoD), or current suppliers could sell the part to alternate customers?* If YES, there is likely a hold-up at play, at least in part, on account of Intermediate Market Conditions. This would be an *Intermediate Hold-up*.

<u>The second question is</u>: *Could you imagine use cases or alternative supply options for the part or its components <u>today</u> if not for the present contract or data rights constraints?* If YES, there is likely a hold-up at play that derives from the DoD acquisition process as presently regulated or applied: a *Process Hold-up.*

Note that the answer to both questions could be YES, in which case you have components of both an *Intermediate Hold-up* and *Process Hold-up* sourcing contract inefficiencies potentially worth mitigating. However, if the answer to both questions is NO, then you are likely constrained by a *Product Hold-up,* in which case the remainder of this decision matrix is unlikely to help.

Decision 2: Is the Hold-up Worth Mitigating?

For *Process Hold-ups* and *Intermediate Hold-ups,* the decision matrix of mitigation tools can be thought of as an eight-binned cube (Figure 6). However, not all the bins will likely merit attention. For any DoD program, the key metric that tends to drive scale from production through maintenance and sustainment phases is the rate of consumption, or frequency of replacement, of parts. Because consumption, or replacement, in turn, drives the enduring value of those parts' TDPs, that can also serve as an indicator for the exigency of a new approach.

If a part contracted for acquisition in the design and production phase of the MDAP is intended to last the full lifespan of the platform, there is likely little value in dedicating decision time to contemplating new tools for better contracting and maintenance of its TDP. That is not to say low-consumption parts are not critical. On the contrary, they may be the utmost essential components for the military mission. However, if access to repetitive supply of either the part or its data is likely, it may not be worth isolating individually for focus.

Decision 3: What Solution-Bins Make Sense?

Narrowing the focus to the top four bins of the cube based on overall projected consumption, or usage rates, of the parts and TDPs across the life cycle of the program, the next decision becomes which tool, or combination of tools, outlined in this report may be best suited to mitigate the particular characteristics of the *Process Hold-up* or *Intermediate Hold-up* in question.

The next question is how to think across the bins to select the mitigation solution that fits best. Starting with an evaluation of two broad hold-up variables may help: Part *Specificity* and Part *Complexity*.

In general, parts that are less specific and less complex (i.e., more easily transferable for supply beyond initial contractors) offer the most opportunity to apply Real Options to transactions involving the parts' TDPs. For those with higher specificity but still low complexity, it may be more reasonable for the DoD to purchase or pursue Unlimited Rights to the TDP up front.

For parts that are more complex, either in terms of construction or IP, it is likely less favorable for the DoD to secure and maintain their TDPs independently. Therefore, collaborative approaches that share both cost and risk, such as IP pooling or TDaaS, could be better options.

Decision 4: What Tool Fits Best?

As discussed in the How Hold-ups Have Been Addressed in Other Industries section and the Implications for DoD TDP Contracting in MDAPs section, a cascading decision model should not be interpreted to imply that the categorization of hold-ups or choosing solutions to mitigate them is an exercise in checklists or mutual exclusion. The questions and decisions above could often lead to a place on the cube seemingly between or across two solution bins. How to proceed then?

In this case, zooming into the decision space to apply other variables introduced in the earlier sections could help.

For example, consider the case in which the hold-up over a part and its TDP lands DoD decision makers weighing whether to purchase or pursue unlimited rights to the data versus attempting to craft, negotiate, and price a TDaaS arrangement. The complexity of the part will be one variable to consider, but it would not be the only one (Figure 7).

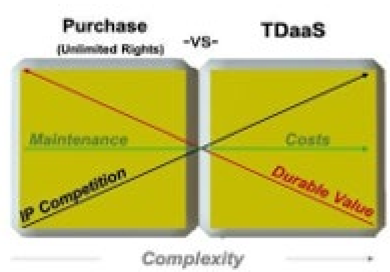


Figure 7.

Additional factors would be how competitive the contractor would consider the IP associated with the part and its TDP. The higher their competitive proprietary interests, the higher price they will likely seek to extract for unlimited rights. Another factor along the same axis would be the estimated cost to maintain and secure the TDP, as well as update it over time. High IP competition and/or data maintenance costs would tip the scales in favor of a TDaaS approach.

Thinking in reverse, however, durability of design and TDP relevance over the life cycle of sustainment could also make a difference. TDPs for parts likely to be wholly redesigned in time may have limited durable value, the lower of which the more favorable TDaaS could be.

Next, consider the seam between a Real Options approach and pursuing an IP Pool (Figure 8). Again, complexity is an important first variable. However, since the Black Scholes thinking is part of the equation, it may be worth breaking that complexity into sub-variables that make up or accompany it.

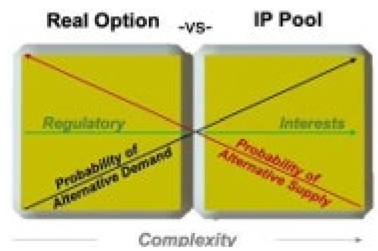


Figure 8

A byproduct of complexity will be Probability of Alternative Supply. The lower the likelihood that the market itself will change on account of new supply-side competitive pressures over time, the less a Real Option makes sense and the more an IP Pooling regime could be attractive.

On the other hand, the higher the likelihood of alternative demand, perhaps from other countries or other DoD programs, the higher the incentives for both the current supply and other parties to enter a pool. The same goes for regulatory interests. The more complex and widely impactful the regulatory interests embedded within a Process Hold-up, the less likely a bilaterally negotiated Real Option will drive meaningful change and the more likely a permanent consortium may be welcome.

Third, zoom into the Specificity axis for a look at the seam between IP pooling and TDaaS (Figure 9). Here, in scenarios more likely influenced by Process Hold-ups rather than Intermediate Hold-ups, Probability of Alternative Supply and Probability of Alternative Demand reorient their vectors. As the probability of future alternative demand options increases, the probability of downward price pressures increases, and DoD equities in a TDaaS approach grow (compared to a permanent IP pool). In the same direction, the higher the maintenance costs of the TDPs, the more sense it makes for a TDaaS subscription model that alleviates the DoD of those burdens.

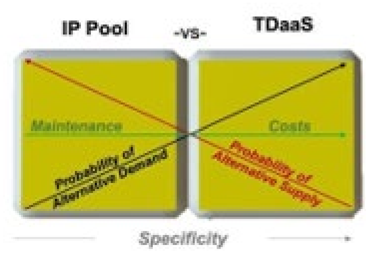


Figure 9.

Conversely, the higher the Probability of Alternative Supply emerging in the market over time, the more an IP pool makes sense, which can accommodate and even accelerate other suppliers, compared to a TDaaS arrangement with today's single source.

Fourth, on the same axis, is the decision space between a Real Option and the Pursuit or Purchase of Unlimited Rights up front (Figure 10). Here, as between other bins, Maintenance Costs of the TDPs over time need to be considered, with higher cost projections tipping the scales toward an Option.

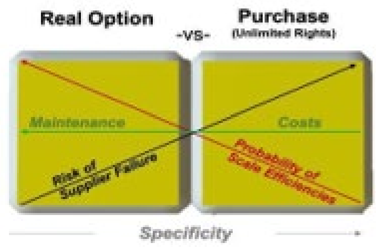


Figure 10.

Beyond that, two new variables merit inclusion. One is the Risk of Supplier Failure. The higher the risk that the market is insufficient to sustain the supplier in business from design and production through maintenance and sustainment, the more value there is to the DoD in securing the data upfront. On the other hand, the more likely technology or other developments could introduce scale efficiencies in production over time, either for the current supplier or new ones with new methods, the more it makes sense for the DoD to focus on valuing Real Options to buy time for those impacts to emerge.

Finally, it is worth thinking again of diagonal decisions across the top of the cube. Zooming into the seam between IP pooling and pursuing or purchasing unlimited rights for the data draws into relief the role of Complexity and Specificity in deciding between IP pools and unlimited data rights (Figure 11). Securing full rights and responsibilities makes the most sense when Specificity is high, but Complexity is low. Whereas IP pools are best when Complexity is high, but specificity is low.

Figure 11.

The last seam is that between Real Options and TDaaS (Figure 12). This diagonal gives a new perspective on something already explored. Real Options and TDaaS may complement each other and can go hand-in-hand. A Real Option could include TDaaS or vice-versa. However, where to start may depend on the same variables just discussed. Low specificity and complexity may suggest a Real Option base. Conversely, High Specificity plus Complexity may suggest TDaaS from the start.

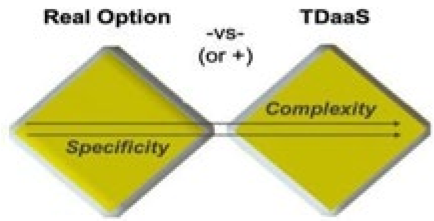


Figure 12.

Conclusion

This paper sought to apply contract economics theory and applied research on hold-up scenarios to challenges in DoD valuation of TDPs and IP in dynamic time-bound markets. The parallels may be imprecise, and the tools abstract, but the study elevates five points of insights and recommendations for the future:

1. **Many parts merit many tools.** In any complex MDAP, there is a risk of being overwhelmed and contractually paralyzed by mass. Millions of parts with varying values, projected lifespans, and data infrastructures cannot be transacted on their own terms. On the other hand, the risk of oversimplification must also be acknowledged. Assuming all components and their TDPs should be priced and acquired en masse, all priced as single type contract or bundled under broadly claimed usage rights carries considerable long-term consequences. Whether or not the tools and guidance outlined above are the right ones for the DoD to adopt, choosing among several will always beat "one size fits all."

2. **Cost-based pricing handicaps options.** If there is a single first step the DoD could make to improve its positioning vis-a-vis both *Process Hold-ups* and *Intermediate Hold-ups*, it is to start weaning wherever possible from Cost based pricing as the default approach. This will be neither immediate nor simple, but OTAs can help. Experimentation in this space may be the single most important foundational step toward further experimentation with Real Options or TDaaS on a measurable scale.

3. **Public-private IP pools are underexplored.** Challenges in IP management across the DoD are a topic of wide discourse. However, the majority of the discussion appears to be focused on policy and regulatory reform. These may overshadow the exploration of more collaboratively disruptive organizational solutions like IP pools. Additional investment and experimentation in his arena may be worthwhile.

4. **Further interdisciplinary study is warranted.** This project offers a theoretical decision framework derived from economics and business research as applied in other industries. The validity and functionality of the framework merits testing within real DoD acquisition scenarios. That should include both historical cases, from which assessments can be made on the impact it could have made, as well as an analysis of its feasibility in current and future MDAPs. If validated and summarized, examples of real-world applications would also serve to make a more robust decision guide more concrete and relatable to future acquisition professionals. In addition, the models and variables proposed deserve more mathematical attention. The use of real options and dynamic valuation models in the private sector has flourished because quantitative metrics and methodologies have been developed and accepted by both suppliers and consumers as fair and transparent. Sharpening future assessment tools for volatility factors like the risk of supplier failure, probability of alternate

demand, life cycle data maintenance costs, or technological obsolescence would advance decision-maker confidence in choosing the right tools.

5. **Even the best model is not enough on its own.** An assertion echoed throughout interviews for this research is that no model, tool, or guidance alone will prompt the cultural reform needed to drive change. In many cases, legislative authorities exist and the DFARS allows acquisition professionals to explore and experiment with many of the approaches proposed. The barriers to trying are both systemic and personal. DoD incentives for cost versus performance across all program phases, as well as expertise management through rotations of military personnel, are enduring issues in need of attention.

References

AFCEA Signal Media. (2020, November 5). *GD electric boat continues to support $9.5 billion Columbia class submarine construction*. https://www.afcea.org/signal-media/gd-electric-boat-continues-support-95-billion-columbia-class-submarine-construction

Aghion, P., & Holden, R. (2011). Incomplete contracts and the theory of the firm: What have we learned over the past 25 years? *Journal of Economic Perspectives 25*(2)*,* 181–197.

Australian Government. (2024, August 12). *Agreement strengthens AUKUS submarine partnership* [Press Release]. https://www.minister.defence.gov.au/media-releases/2024-08-12/agreement-strengthens-aukus-submarine-partnership

Baron, J., & Pohlmann, T. (2015, February). *The effect of patent pools on patenting and innovation—Evidence from contemporary technology standards*. Northwestern University. https://wwws.law.northwestern.edu/research-faculty/clbe/innovationeconomics/documents/baron_pohlmann_effect_of_patents.pdf

Black, F. & Scholes, M. (1973, May). The pricing of options and corporate liabilities. *Journal of Political Economy 81, 3*637–654.

Corbin, A. (1914, June). Option contracts. *The Yale Law Journal 23***(8)***,* 641–663.

Defense Federal Acquisition Regulation Supplement (DFARS), 48 C.F.R. ch 2 (2024).

DoD. (2023, May). *Selected Acquisition Report (SAR): SSBN 826 COLUMBIA Class Submarine (SSBN 826).* https://www.esd.whs.mil/Portals/54/Documents/FOID/Reading%20Room/Selected_Acquisition_Reports/FY_2022_SARS/SSBN%20826_SAR_DEC_2022_final.pdf

Department of Defense Inspector General. (2021, December 13). *Audit of the business model for TransDigm Group Inc. and its impact on Department of Defense spar*. https://www.oversight.gov/sites/default/files/documents/reports/2021-12/DODIG-2022-043-508.pdf

Exostar Website. (n.d.). *Accelerating digital transformation with the Exostar platform.* https://www.exostar.com/platform/

General Dynamics Electric Boat. (2024). *Columbia class*. https://www.gdeb.com/about/oursubmarines/columbia/

GAO. (2021, January 14). *Columbia class submarine: Delivery hinges on timely and quality materials from an atrophied supplier base* (Report GAO-21-257). https://www.gao.gov/products/gao-21-257

GAO. (2021, November 30). *Defense acquisitions: DOD should take additional actions to improve how it approaches intellectual property* (Report GAO-22-104752). https://www.gao.gov/products/gao-22-104752

Henderson, J. (2020, May 17). Exclusive interview: Exostar's supply chain lead, Kevin VanLowe. *Supply Chain Digital*. https://supplychaindigital.com/procurement/exclusive-interview-exostars-supply-chain-lead-kevin-vanlowe

Klein, B., Crawford, R., & Alchian, A. (1978, October). Vertical integration, appropriable rents, and the competitive contracting process. *The Journal of Law & Economics 21*(2), 297–326.

Lampe. R., & Moser, P. (2016, March). Patent pools, competition, and innovation—Evidence from 20 US industries under the New Deal. *Journal of Law, Economics, & Organization 32*(1), 1–36.

Lerner, A. (1934, June). The concept of monopoly and the measurement of monopoly power. *The Review of Economic Studies 1*(3), 157–175.

Lerner, J., Strowas, M., and Tirole, J. (2007, Autumn). The design of patent pools: The determinants of licensing rules. *The RAND Journal of Economics 38*(3), 610–625.

Lofren, K., Person, T., & Weibull, J. (2002, June). Markets with asymmetric information: The contributions of George Akerlof, Michael Spence and Joseph Stiglitz. *The Scandinavian Journal of Economics 104*(2), 195–211.

Mell, P. & Grance, T. (2011, September). *The NIST definition of cloud computing* (Special Publication 800-145). National Institute of Standards and Technology.

Merton, R. (1973a, Spring). Theory of rational option pricing. *The Bell Journal of Economics and Management Science 4*(1), 141–183.

Merton, R. (1973b, September). An intertemporal capital asset pricing model. *Econometrica 41(5),* 867–887.

National Defense Authorization Act for Fiscal Year 2024, H.R. 2670, 118 Cong. (2024).

National Institutes of Health. (2022, May). *NIH contributions to WHO COVID-19 technology access pool and Q&As*. https://www.techtransfer.nih.gov/policy/ctap

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2022, May 22). *Guidance on inflation and economic price adjustments* [Memorandum].

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2019, October 16). *DoD Instruction 5010.44 Intellectual property (IP) acquisition and licensing*. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/501044p.pdf

Reisinger, M. & Tarantino, E. (2019, Spring). Patent pools, vertical integration, and downstream competition. *The RAND Journal of Economics 50*(1), 168–200.

Rogerson, W. P. (1992, October). Contractual solutions to the hold up problem. *The Review of Economic Studies, 777–793.*

Sander, T. (2022, September 12). *Technical data escrow is the new black*. US Army Acquisition Support Center. https://asc.army.mil/web/news-technical-data-escrow-is-the-new-black/

Thompson, G. & McGrath, M. (2019). *Technical Data as a Service (TDaaS) and the valuation of data options*. Naval Postgraduate School, Acquisition Research Program.

Weintraub, S. (1949) The theory of consumer monopsony. *The Review of Economic Studies 17*(3), 168–178.

Wilkers, R. (2024, July 16). Deloitte wins $2.4B submarine industry development contract. *Washington Technology.* https://www.washingtontechnology.com/contracts/2024/07/deloitte-wins-24b-submarine-industry-development-contract/398078/

Williamson, O. E. (1975, January). *Markets and hierarchies: Analysis and antitrust implications, a study in the economics of internal organization*. The Free Press.

Williamson, O. E. (1979, October). Transaction-cost economics: The governance of contractual relations. *The Journal of Law & Economics 22*(2), 233–261.

Williamson, O. E. (1981, November). The economics of organization: The transaction cost approach. *American Journal of Sociology 87*(3), 548–577.

World Health Organization (WHO). *WHO initiative signs new licensing agreements on COVID-19 technologies* [Press release]. https://www.who.int/news/item/29-08-2023-who-initiative-signs-new-licensing-agreements-on-covid-19-technologies

PANEL 6. OPTIMIZING DEFENSE ACQUISITION: COMMUNICATION, INDUSTRIAL IMPACT, AND CAPITAL INTEGRATION

Wednesday, May 7, 2025	
0950 – 1105 PT 1150 – 1305 CT 1250 – 1405 ET	**Chair: Moshe Schwartz, President, Etherton & Associates** ***What We've Got Here Is Failure to Communicate: How Better Communication Can Improve DoD Acquisition Outcomes*** Moshe Schwartz, President, Etherton & Associates ***Assessing the Impact of Department of Defense Weapons Systems on the Defense Industrial Base*** Amanda Bresler, Chief Strategy Officer, PW Communications ***Tactics to Strengthen the U.S. Defense Industrial Base with Private Capital*** Sam Moyer, Research Fellow, Emerging Technologies Institute ***Shipbuilding and Repair: Navy Needs a Strategic Approach for Private Sector Industrial Base Investments*** Lindsey Cross, Senior Analyst, U.S. Government Accountability Office

Moshe Schwartz—is president of Etherton and Associates and holds positions as Senior Fellow for Acquisition Policy at NDIA and director at the Procurement Round Table. Previously, Moshe served as Executive Director of the congressionally mandated Advisory Panel on Streamlining and Codifying Acquisition Regulations (section 809 Panel), senior advisor to the Commission on Wartime Contracting in Iraq and Afghanistan, advisor at ISAF headquarters in Afghanistan, and spent 15 years providing analysis and legislative support to Congress on acquisition policy and industrial base issues, including as a specialist at the Congressional Research Service and senior analyst at GAO. He has testified before Congress and written extensively on a wide range of acquisition and industrial base issues.

Academically, Moshe was an adjunct instructor at Carnegie Mellon University's Heinz School of Public Policy, and at National Defense University's Eisenhower School for National Security and Resource Strategy. He holds an M.B.A. from Carnegie Mellon's Tepper School of Business, a Master of Science in Public Policy Management from Carnegie Mellon's Heinz School of Public Policy and Management, and a J.D. from Yeshiva University's Cardozo School of Law.

What We’ve Got Here Is Failure to Communicate: How Better Communication Can Improve DoD Acquisition Outcomes

Moshe Schwartz—is President of Etherton and Associates, a consulting and lobbying firm specializing in defense acquisition, industrial base, and cybersecurity policy. Prior to joining Etherton and Associates, Schwartz served as Executive Director of the congressionally mandated Advisory Panel on Streamlining and Codifying Acquisition Regulations and spent 15 years providing analysis and legislative support to Congress on acquisition policy and industrial base issues, including as a specialist at the Congressional Research Service and senior analyst at GAO. Schwartz also served as senior advisor to the Commission on Wartime Contracting in Iraq and Afghanistan, and as advisor at ISAF headquarters in Afghanistan. [moshe@ethertonandassociates.com]

Michelle Johnson—has researched and published about the defense acquisition process from multiple roles since 2017. Most recently, she was Communications Manager at the Acquisition Research Program at Naval Postgraduate School, where among other things she edited a weekly newsletter with updates and analysis of news in defense, acquisition, and innovation efforts within the DoD. Prior to joining NPS, she was a research associate and communications manager for the congressionally mandated Advisory Panel on Streamlining and Codifying Acquisition Regulations (the Section 809 Panel). Johnson has a PhD in English Literature from Michigan State University. [shel.veenstra@gmail.com]

Daniella Schwartz—is a senior at the University of Maryland, College Park, studying Psychology and on the Pre-Medical track. She previously interned at National Defense University and Children’s Medical Center, where her contributions were presented at academic conferences and will soon be submitted for publication. [dcschwartz02@gmail.com]

Abstract

This research identifies where poor communication between DoD and industry, between DoD and Congress, and internally within DoD is hampering acquisition outcomes. These challenges are widely acknowledged, having been cited by Deputy Secretary of Defense Kathleen Hicks, the DoD small business office, industry, and Congress. Poor communication leads to poor requirements, inefficient budgeting, a less effective public comment process for regulations and policy Requests for Information, distrust between Congress and DoD, and an increase in bid protests, and otherwise strains the DoD–industry relationship.

Our research provides a framework for identifying and defining communication challenges, including lack of clarity, withholding information, lack of trust, one-way communication, and communication processes that often lack the substantive discussions intended by the formal communication process. Our research also identifies where communications challenges exist, assesses the impact of these challenges on the acquisition process and its outcomes, identifies causes, and recommends approaches to improving communication and collaboration.

“A vibrant innovation ecosystem depends upon clear communication to ensure partners have accurate information and can build complementary processes to enable effective collaboration.”

(DoD Strategic Management Plan: FY24 Annual Performance Report, p. 85)

Introduction

Acquisition is a human endeavor, where success or failure depends primarily on the thoughts, beliefs, and foibles of the people who make up the acquisition workforce. Because of the human element in contracting, relationships matter. They are not all that matter; contracts and budgets also matter. But relationships, budgets, contracts, and other elements of acquisition share a common thread: the need for effective communication.

Effective communication is key to the success of any organization. Without effective communication, information is not shared, priorities and goals are not defined, culture suffers, and relationships are not as strong. The Department of Defense (DoD) is no exception. Breakdowns in communication—or the absence of robust communication—have led to subpar requirements, inefficient budgeting, increased bid protests, poor acquisition outcomes, increased costs, and a strained DoD–industry relationship.

This paper explores principles of organizational communication and identifies examples of how improved communication could help improve DoD acquisition, budget, and industrial base outcomes.

Let's Talk Communication

In our careers, we have conducted extensive research, written dozens of papers, and participated in numerous conferences on defense acquisition. Recently, in talking about acquisition and past research, we realized that communication is not an ancillary point but a core issue running throughout the entire acquisition process, serving as a key source of failure or a catalyst for success.

Congress and the federal government have made numerous efforts to enhance acquisition outcomes. For example, DoD and other agencies have turned to non-traditional contracting methods, such as other transactions (OT), and the use of consortia to promote greater communications and expand the defense industrial base. Research on the use of consortia and the (still) shrinking defense industrial base provide insight into the importance of communication.

Why Are Consortia So Popular? The Opportunity to Communicate

A 2022 analysis of consortia, *The Power of Many: Leveraging Consortia to Promote Innovation, Expand the Defense Industrial Base, and Accelerate Acquisition* (Halcrow & Schwartz), found that

> the consortia model supports government acquisition efforts by promoting *government–industry–academia communication*, facilitating industry partnerships and collaboration, providing critical surge capacity to government acquisition, offering a ready, pre-established network of potential suppliers who have expertise in specific areas, and helping government program offices that do not have the requisite skill and experience in executing OTs. (p. 1, emphasis added)

Communication is not just the first point in the paragraph but a running theme in the report. Specifically, the report shows that companies are attracted to consortia for two primary reasons:

- Business relationships are generally governed by other transaction authorities which are not bound by the FAR or many other regulatory and legislative requirements.
- Members of consortia enjoy more communication and collaboration both between government and industry, and within industry.

The importance of communication to the consortia model is no coincidence. The pioneers of consortia developed the model precisely to "develop a new approach to contracting that encouraged collaboration and communication between government and a diverse team of industry participants throughout the acquisition process" (p. 3). The focus on collaboration and communication is a major contributor to the success of consortia.

Why Is the Defense Industrial Base (Still) Shrinking? Actions Speak Louder than Words

Communication was also featured in a series of articles and reports we wrote on the shrinking defense industrial base, where we argued:

> Excessive regulation artificially constrains the potential of business relationships by reducing them to mechanistic processes focused on checklists and fear of legal action for compliance failures. … The first step to relational contracting is for DoD to develop a better understanding of how industry operates: what motivates companies, what drives business decisions, and, most importantly, what prompts companies to leave (or not enter) the National Security Innovation and Industrial Base. (Schwartz & Johnson, 2023)

At the time we did not appreciate the foundational import of communication to relational contracting, the full span of the acquisition system, and acquisition reform. Now we do. And as is the case with excess regulation, communication must be consistent with actions. Actions, after all, speak louder than words.

Others have recognized the importance of communication to improve acquisitions, including DoD and the Office of Management and Budget. OMB's 2019 *Myth-Busting Memo #4: Strengthening Engagement with Industry Partners through Innovative Business Practices*, reminded acquisition professionals to leverage all methods of communication available to them and asked each agency to appoint an industry liaison. One-on-one conversations with industry, for example, can "foster business partnerships while capturing industry feedback to improve acquisition planning and requirements definition." Building on the memo, on December 1, 2022, a Federal Acquisition Regulation final rule was published that made clear "agency acquisition personnel are permitted and encouraged to engage in responsible and constructive exchanges with industry, so long as those exchanges are consistent with existing law and regulation and do not promote an unfair competitive advantage to particular firms" (Department of Defense, 2022).

> "Our office has produced 5-year investment plans that we continuously share with industry. While we need to maintain a level of flexibility to respond to accidents and national emergencies, it's not fair to tell industry and other investors that we need their help and their investments but not disclose what our own 5-year plan is."
>
> Anthony R. Di Stasio,
> Dep. Assistant Secretary of Defense
> (Industrial Base Resilience)

Despite these and other mandates to prioritize effective communication, the message has not been received—or perhaps more accurately, the rules, regulations, culture of compliance, and existing incentives (and disincentives) serve as barriers to DoD communication. As Soraya Correa, former chief procurement officer at the Department of Homeland Security, notes, "Acquisition professionals still tend to be risk averse and limit or restrict communications" (personal communication, email with author, March 25, 2025).

Do Communication and Relational Contracting Work?

A comprehensive study on contracting found that "the best sourcing relationships apply what is known as 'relational' contracting principles, which create flexible contract frameworks and embody 'win-win'" behaviors" (Vitasek et al., 2022, p. 2). Such an approach has long been recognized by leading companies such as McDonald's, which famously relies on long-term relational contracting to manage its supply chain and subcontractor relationships. One of the foundational principles of relational contracting is communication. McDonald's uses a multi-level communication approach with its partners. As one analysis pointed out,

> McDonald's maintains communication with suppliers both in formal and informal styles. The goal is to encourage … an open culture in communication. An open culture and communication also ensure all decisions are based on the company's "System First" philosophy. … The McDonald's Supplier Management principles consider the basic tenets of human psychology. (Tabansi, 2023)

Communication as a management principle is not an end in itself but a catalyst for better contractual relationships and acquisition outcomes. Communication offers distinct benefits to an organization, including:

- Increased productivity and improved efficiency
- Reduced costs
- Improved outcomes (through better understanding of desired effects) (Olkkonen et al., 2000).

While McDonald's supplier management principles have been described as trying to achieve trust, freedom, clear and easy communication, and scalability and profitability (Tabansi, 2023), the end goal of the communications is the scalability and profitability—in other words, the desired outcomes of the supplier management policies.

DoD Communication Challenges

GAO's *Standards for Internal Control in the Federal Government* identifies five components of internal controls, the fourth of which is "Information and Communication" (2014). The GAO standards set down principles, including the need to

- *internally communicate* the necessary quality information to achieve the entity's objectives and
- *externally communicate* the necessary quality information to achieve the entity's objectives. (emphasis added)

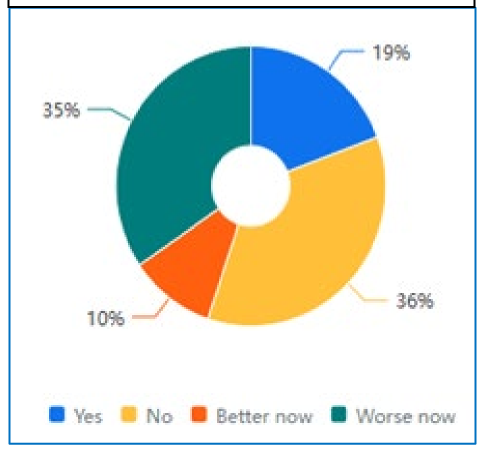
Barriers to communication are not surprising in a highly rigid and hierarchical organization such as DoD but are critical for internal controls and effective management. These barriers can be overcome. Below, we focus on two areas where DoD communications hamper acquisition outcomes:

- External communication with industry
- External communication with Congress.

Communication With Industry

As discussed above, a long-recognized weakness of the defense acquisition system is a lack of early, consistent, and effective communication with industry. The consequences of such insufficient communication include unclear requirements that do not attract industry interest, deter companies from working on government contracts, and increase legal challenges and contract disputes. These challenges plague small businesses, commercial tech companies, and even large traditional defense contractors. A recent NCMA poll of contracting professionals asked if there is good communication between government and industry. Of the more than 530 responses, more than 70% said there is not good communication or it is worse now (online poll conducted by author on March 10, 2025). This poll is consistent with other data addressing the issue of communication.

Do you believe there is good communication between government and industry?



A 2023 presentation by Khalil Mack, then director of APEX Accelerators in DoD, presented the results of a Federal Register Notice requesting industry input on barriers working with DoD. Based on 211 responses, 13 major barriers facing small businesses in contracting with DoD were identified, the first of which was "communications and outreach."

This is not just a small business challenge; it is an all-of-industry challenge. In its report *Vital Signs 2025: The Health and Readiness of the Defense Industrial Base*, NDIA polled 1,273 government and industry respondents. When asked to identify what is difficult about government acquisition processes, more people (58%) cited "unclear or changing requirements" than any other issue (National Defense Industrial Association, 2025, p. 12). Lack of clarity is often the result of poor communication. The report found communication challenges in a variety of areas, including:

- *Improving Relationships* - when asked what steps DoD could take to improve its ability to work with industry, the most common response was "provide clear, consistent demand signal through contract vehicles" (p. 13). Of the 12 responses listed, three relate to communication, including providing industry with timely updates as requirements evolve through OTs and providing clear identification of specific points of contact in program offices.
- *Cybersecurity* – When asked what challenges organizations face in implementing the security requirements in NIST SP 800-171 to manage Controlled Unclassified Information, the second and third most cited challenges were "insufficient guidance" on NIST SP compliance (32%) and "difficulty in understanding" the requirements (27%).
- *Foreign Sales* – When asked about barriers in selling to foreign customers, 39% of respondents cited "transparency with and communication from the US federal government"[1] (p. 36).

Part of the challenge appears to be poor communications strategies and writing. At the 18th Naval Postgraduate School Acquisition Research Symposium, a paper entitled *Why*

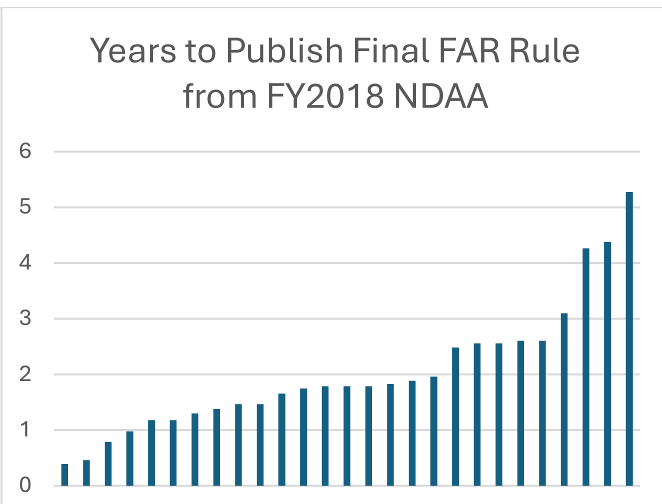
[1] This was the third highest factor cited out of 10

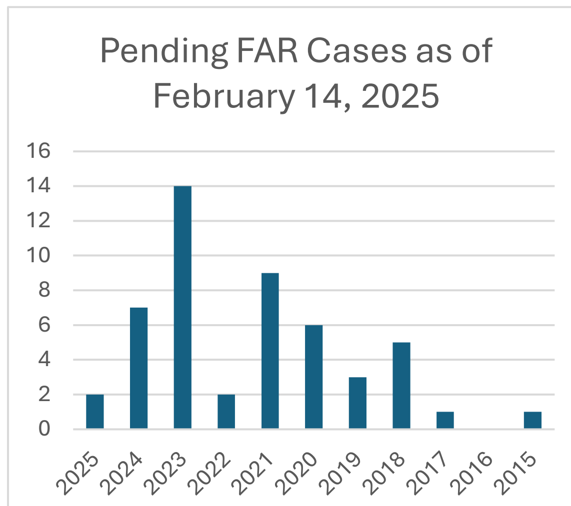
Marketing Matters: Strengthening the Defense Supplier Base Through Better Communications with Industry found that *"how* and *where* the DoD communicates with industry have contributed to" its failure to attract and engage a "significant number of new suppliers over the last decade" (Bresler & Bresler, 2021, p. 91).

Case Study: The Regulatory Process as Inefficient and Impersonal Communications

Every year, dozens of new or modified rules governing federal acquisition are added to the Federal Acquisition Regulation (FAR) and Defense Federal Acquisition Regulation Supplement (DFARS). The process by which these rules are crafted is designed to be deliberative, ensuring robust collaboration and public input. As executed, however, the process takes too long and uses communication strategies that are formalistic, asynchronous, and complicated. Final rules often reflect changes suggested by public comments,[2] but the process takes on average two to three years from initiation to final rule. In some cases, it takes much longer. This delay causes confusion and fatigue, and dissuades stakeholders from participating in the process. Some companies delay taking steps to implement enacted legislation because they know it could take years before regulations are issued.

As of February 14, 2025, there were 50 open FAR cases at different stages in the regulatory process, in one case dating back to 2015. Half of these cases (25 of 50) have been pending for over 4 years, since 2021 or earlier (Defense Acquisition Regulations System, 2025).





One of these pending cases has been in the works for over 7 years. In July 2018, the Section 809 Panel published a recommendation that cost accounting standards (CAS) applicability for indefinite delivery vehicles be determined at the time of the task order award, not the contract award, as has been standard practice. Three years later, in 2021, the CAS Board took up the proposal and assigned it a case number. After another three years, it was published in the Federal Register for public comment on July 18, 2024 (Office of Federal Procurement Policy, 2024). As of February 2025, no final rule has been published on this relatively straightforward change that makes it easier for businesses to get on contract with the federal government. In the meantime, the need for clarity on the issue has led to several legal cases, one of which is now referred to in lieu of an updated regulation. This 2020 legal case

[2] A notable example is the public comment process for the proposed Cybersecurity Maturity Model Certification process, which elicited scores of comments and led to more precise identification of roles and responsibilities.

established the same precedent as the proposed rule of determining applicability at the task order award (Ferrari, 2024).[3]

In addition to such delays, public meetings on proposed rules are now primarily held remotely, making it harder to hold meaningful conversations. Before COVID, public meetings on proposed rules were held in person and benefited from more back-and-forth during the presentation phase. We believe that the remote nature of the public meetings results in less relationship building, fewer informal conversations and data sharing, and less give-and-take during public presentations. Indeed, the theory of media richness used in organizational behavior explains this dynamic by considering how different forms of communication

> convey cues (e.g., tone of voice, nonverbal gestures) and allow for immediate feedback, personalization, and language variety. Media that convey more of these characteristics are considered to provide richer information and are theorized to be better at reducing ambiguity and uncertainty. Richer communication media have a greater capacity to facilitate *a sense of shared meaning or understanding of the information being relayed*. (Cordova et al., 2013, p. 3; emphasis added)

The more human the interaction, the richer the communication. “Face-to-face communication is considered the richest type of communication, because it allows for the reading of nonverbal cues, allows individuals to ask questions and verify a mutual understanding, and allows for personal interaction” (p. 3). To make the regulatory process richer and more successful at sharing information, public meetings should be held in-person or as hybrid events.

A January 2025 Memorandum from OMB, *Broadening Public Participation and Community Engagement with the Federal Government*, acknowledges the need for improved channels of communication between the federal government and members of the public, whether as private citizens or as representatives of industry or other communities. The memorandum goes on to offer guidance for how to achieve more open, richer, and synchronous channels of communication in a variety of formats, such as website portals, webinars, and listening sessions. What the memorandum drives home is that public participation must be an ongoing effort that meets the public where they are, not just a series of formal written notices posted in the Federal Register.

Communication with Congress

DoD communication with Congress, which is important for setting budgets that support acquisitions, has also experienced challenges with effective communication. Ahead of a classified oversight briefing with Secretary of the Navy Carlos Del Toro in September 2024, the chairman of the defense appropriations subcommittee released a statement that read, “This subcommittee expects honesty and transparency from the Navy. … I no longer trust that this committee is being given sufficient information required for meaningful oversight.” In the FY2021 committee report for the DoD appropriations bill, the committee wrote, “The granting of additional budget flexibility to the Department is based on the presumption that a state of trust and comity exists between the legislative and executive branches regarding the proper use of appropriated funds. This presumption presently is false” (House of Representatives, 2021).

---

[3] In another example from the Section 809 Panel, a change to terminology from “commercial item” to “commercial product and commercial service” took over 4 years to be updated in the DFARS after Congress directed the change in the FY2019 NDAA. Federal Register, (2023, January 31), “Defense Federal Acquisition Regulation Supplement: Definition of “Commercial Item” (DFARS Case 2018-D066).” https://www.federalregister.gov/documents/2023/01/31/2023-01294/defense-federal-acquisition-regulation-supplement-definition-of-commercial-item-dfars-case-2018-d066

The PPBE reform commission, in its final report to Congress, argued "by fostering transparent, consistent, and timely communication, DoD aims to keep Congress well-informed about resource needs, budget execution, and program performance" (Commission, 2024, p. iii). The commission further wrote that its "initiatives span the entire DoD, strengthen the analytic underpinning of strategic decisions, add agility and flexibility into resource management, and improve communication with Congress" (Foreword).

The commission considered communication so important that its second framework is *Modernize and Simplify Information Sharing.* Within this framework objective, two of the four reform objectives focus on communication: improved communications with Congress and establishing communication enclaves between DoD and Congress.[4]

> DoD "need[s] more trust from Congress. We will keep working to build trust with Congress, but it is a two-way street."
>
> Former Deputy Secretary of Defense Kathleen Hicks

Better communication with Congress and internal communication with the various Department stakeholders will improve analytic strategic decisions (by incorporating more data and information to support data-driven decision-making) and add agility and flexibility into resource management (by fostering trust, providing insight, and promoting collaborative policy discussions). If implemented, the commission's recommendations in this area can go a long way in improving PPBE outcomes.

Communication between DoD and Congress works both ways, and Congress bears part of the responsibility for the state of the relationship (perhaps the subject of a future paper).

How Can DoD Have More Effective Communication?

Effective communication consists of two elements: sharing/providing information to another and receiving/understanding information being transmitted by others (Radovic Markovic & Salamzadeh, 2018). This communication requires:

- That you are conveying what the other person needs to know
- Ensuring that what you mean to convey is understood by the other party (this includes the other party being comfortable to ask for clarification or propose other approaches)
- That the other party trusts the information being conveyed
- A willingness to receive new information and be open to changing one's position/ideas.

A Framework for Communication

Organizational communication exists in three spheres (Radovic Markovic & Salamzadeh, 2018):

- Outbound communication to external parties, which in the case of DoD includes contractors and Congress
- Internal communications throughout the organization, such as between the military departments, between the Office of the Secretary of Defense and the components within the Department, and the warfighters/requirements generators and the contracting workforce

[4] Communication also appears in the report's discussion on workforce, recommending that DoD develop standardized training for liaisons "incorporating best practices to ensure effective communication with Congress". See page 33.

- Inbound communications, such as DoD managing the comments from the regulatory public comment process, feedback from contractors at industry days, or implementing/internalizing legislation and Congressional communications.

Having mechanisms for communication does not mean that effective communication is taking place or that the right information is being shared. Formal processes without substance often result in communication theatre. For example, the mere existence of a debriefing process after contract award does not *ipso facto* mean that effective communication, feedback, and learning are taking place.

Bid Protests – How Better Substantive Communication Decreases Protests and Improves Future Competition

In 2018, RAND published a report analyzing GAO bid protests (Arena et al.). The report identified poor quality of post-award debriefings as one driving cause of bid protests (p. 20). According to the report, standard debriefings conducted under FAR 15.505 and 15.506 often fail to provide unsuccessful offerors with sufficient information to determine whether their proposals were properly evaluated (p. 15). Standard debriefings generally did not provide the government's underlying rationale for its evaluation conclusions. Industry characterized the debriefings as "skimpy, adversarial, evasive, or failing to provide required reasonable responses to relevant questions" (p. 20). As a result, offerors sometimes filed protests to simply gain access to award evaluation information that they could have received in a good debrief (Field, 2019).

The Office of Federal Procurement Policy (OFPP) acknowledged a common misconception among acquisition officials that limiting communication with industry—especially during debriefings—will avoid bid protests. Rather, it explained that enhancing the quality of debriefings may improve competition and help diminish bid protests (Field, 2019).

Congress initially established debriefings to encourage the free flow of information and provide offerors an indication of how they could improve their chances for success in future procurements (S. Rep. No. 103-258, at 7). Providing an opportunity for meaningful feedback helps vendors better understand the deficiencies of their proposal so they can avoid repeating the same issues and make stronger offers on future procurements (Arena et al., 2018, p. 20; Field, 2017, p. 4). Good debriefs also improve the perception of fairness and equality in the evaluation process (Field, 2017, p. 4; Schooner, 2020). OFPP noted that this communication increases the pool of competition where the government can obtain more responsive offers in the future and help mitigate the risk of protest (Field, 2017, p. 6; Arena et al., 2018, p. 65).

When contractors receive vague or insufficient information to ascertain whether the evaluation was conducted properly, contractors often adopt a "kitchen sink" approach in filing bid protests (Edwards & Schooner, 2021). An offeror can submit multiple claims in a bid protest so long as they can allege some harm or prejudice. If offerors gain access to meaningful evaluation information through a debriefing, it allows them to narrow protest claims to ones where actual prejudice might have occurred or even dissuade offerors from filing a protest altogether (Edwards & Schooner, 2021).

The U.S. Air Force's Extended Debriefing process exemplifies the success of heightened communication in debriefings and reduced bid protests (Arena et al., 2018, p. 65). Unfortunately, departing from the tradition of robust communication, the Air Force refused to discuss their approach with the authors.

Case Study: Solid Rocket Motors Call for White Papers

DoD has succeeded in effective communication with industry. In May 2024, the Defense Industrial Base Consortium (DIBC) issued a request for white papers for solid rocket motors, with the goal of increasing the number of suppliers in this critical and underdeveloped part of the defense supply chain. DoD engaged in intentional and collaborative communication with industry throughout the process to maximize participation and information sharing from all stakeholders.

The white paper request did not just reflect DoD's needs but was developed in collaboration with industry partners—both primes and sub-primes—through multiple engagements including industry days, posts on LinkedIn, and a call on the DIBC website. The contracting team talked to potential primes and subcontractors at many meetings and conferences to understand their thinking, get feedback, and further shape the request for white papers. The team learned from primes about supply chain issues that informed the acquisition strategy. The request itself was also left intentionally broad to encourage maximum participation from partners with diverse capabilities and offerings, a strategy suggested by industry members.

As a result of this intentional communication, the DIBC received over 60 white papers, well over the 10–15 papers initially expected, and got a better understanding of the supply chain landscape for solid rocket motors. Many of these responses came from sub-tier vendors, who were encouraged by primes to participate. Industry members were also motivated to participate because DoD communicated seriousness and funding certainty from Defense Production Act, Title III funding.

Why DoD Succeeded in These Cases: The Keys to Creating Effective Communication

The success stories included in this report share a few characteristics of effective communication. They reflect and reinforce the larger goals shared by DoD, Congress, and industry: to ensure national security by meeting warfighting needs with the right capabilities at the right time. To achieve those goals, these communications must

- be timely and responsive to feedback,
- allow for information to flow both ways along the communication channel,
- be meaningful, not formulaic,
- treat individual communications as part of a larger relationship,
- build new knowledge collaboratively,
- establish trust, and
- most importantly, foster an environment where communication is encouraged.

DoD is a large and complex organization, prone to creating communications that are procedural and mechanistic rather than dynamic and mission-driven. The hierarchical nature of DoD makes meaningful communication even harder. But this natural tendency leads to broken relationships, missed opportunities, and subpar outcomes. To remedy this approach, DoD should view all parts of the acquisition life cycle as opportunities for communication, including requirements, the regulatory process, debriefings, bid protests, budget requests to Congress, and requests to industry for white papers.

When executed effectively, individual communications set the stage for success over the long term. Effective debriefings not only discourage time-consuming bid protests; they also inform industry partners about how to improve their business models and approach to future contracting opportunities. In the case of the solid rocket motor request for white papers, communications provided multiple opportunities for industry to influence DoD's understanding of the marketplace. DoD was able to adjust, make necessary changes, and clarify its communication with industry. The opportunity for timely communication and responses are also one reason why using Other Transaction Authority through consortia is so popular with industry. Using communication to forge successful relationships between industry and DoD creates an active marketplace in which suppliers understand mission needs well enough to supply capabilities that may even go beyond those prescribed by DoD requirements. Similarly, successful relationships between DoD and Congress ensure that requirements, authorities, and funding are aligned to develop and deliver those capabilities effectively.

Conclusion

Private industry is one of the greatest strengths of the United States. As Jason Rathje, then director of the Office of Strategic Capital, stated in 2024, "The U.S. capital markets are the largest and the most liquid in the world. We see them as a national competitive advantage for the U.S." (Carberry, 2024).

Private industry, including such engines of innovation as private equity and venture capital, are giving DoD a new look and increasing investments in national security. In the words of one industry partner, "I don't think I've ever seen such excitement, enthusiasm for investing in defense tech across a wide variety of investment firms. I think it's something that pretty much every serious … traditional Silicon Valley investment firm has at least one partner who's focused on aerospace or defense" (Katherine Boyle, General Partner at Andreessen Horowitz, quoted in Carberry, 2024).

To maintain technological advantage on and off the battlefield, DoD needs to leverage U.S. capital markets and the full strength and innovation of domestic industry. To do this, DoD must learn how to be a better customer and to more effectively communicate. Until that happens, DoD's current state of communication—or the lack thereof—is holding it back from fully leveraging its greatest competitive advantage: America's industrial power.

References

Arena, M. V., Persons, P., Blickstein, I., Chenoweth, M., Lee, G. T., Luckey, D., & Schendt, A. (2018, January 4). *Assessing bid protests of U.S. Department of Defense procurements: Identifying issues, trends, and drivers*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR2356.html

Bresler, A., & Bresler, A. (2021, May 10). Why marketing matters: Strengthening the defense supplier base through better communication with industry. *Proceedings of the 18th Annual Acquisition Research Symposium at the Naval Postgraduate School*. https://dair.nps.edu/handle/123456789/4404

Carberry, S. (2024, May 20). Venture capital getting bullish on defense tech. *National Defense Magazine*. https://www.nationaldefensemagazine.org/articles/2024/5/20/budget-matters-venture-capital-getting-bullish-on-defense-tech

Commission on Planning, Programming, Budgeting, and Execution Reform. (2024, March). *Defense resourcing for the future*. https://ppbereform.senate.gov/wp-content/uploads/2024/03/Commission-on-PPBE-Reform_Full-Report_6-March-2024_FINAL.pdf

Defense Acquisition Regulations System. (2023, January 31). *Defense federal acquisition regulation supplement: Definition of "commercial item"* (DFARS Case 2018-D066). Federal Register. https://www.federalregister.gov/documents/2023/01/31/2023-01294/defense-federal-acquisition-regulation-supplement-definition-of-commercial-item-dfars-case-2018-d066

Defense Acquisition Regulations System. (2025, February 14). *Open FAR cases*. https://www.acq.osd.mil/dpap/dars/opencases/farcasenum/far.pdf

Defense Pricing and Contracting. (2024, January 19). *Defense federal acquisition regulation supplement (DFARS): NDAA implementation tracker (FY10–Present).* https://www.acq.osd.mil/dpap/dars/docs/DFARS_NDAA_Implementation_Tracker_2024-01-09.pdf

Department of Defense. (2022, December 12). *Federal acquisition regulation: Effective communication between government and industry*. Federal Register, 87 FR 73902. https://www.federalregister.gov/documents/2022/12/01/2022-25959/federal-acquisition-regulation-effective-communication-between-government-and-industry

Department of Defense. (2024). *DoD strategic management plan, Fiscal years 2022–2026: FY24 annual performance report*. https://dam.defense.gov/Portals/47/Documents/Publications/Annual%20Performance%20Plan/FY24%20Annual%20Performance%20Report%20FINAL%20DSD%20Approved.pdf?ver=eo4y-GqILynd4RPlaM1X8Q%3d%3d

Edwards, V. J., & Schooner, S. L. (2021, August). Postscript III: Enhanced debriefings: A simple strategy for a more manageable protest process. *The Nash & Cibnic Report*, *35*(8). https://www.arnoldporter.com/-/media/files/perspectives/publications/2021/08/postscript-iii-enhanced-debriefings-a-simple-strat.pdf?rev=e988e3723c344ec0b97e8ec15002dd30&sc_lang=en&hash=BED72047A89186F3FC4024D7A0709129

Ferrari, D. K. (2024, August 8). *CGB comments on CASB Case No. 2021-01, Regarding the application of cost accounting standards (CAS) to indefinite delivery vehicles (IDVs).* Financial Executives. https://www.financialexecutives.org/Influence/Comment-Letters/2024/CGB-CASB-Public-Comments-re-CAS-Applicability-to-I.aspx

Field, L. A. (2017, January 5). *Myth-busting memorandum #3: Further improving industry communication with effective debriefings*. Office of Federal Procurement Policy. https://business.defense.gov/Portals/57/OFPP%20Myth-busting%203%20-%20Further%20Improving%20Industry%20Communication%20with%20Effe.._.pdf?ver=2017-12-05-152530-177&timestamp=1512506152466

Field, L. A. (2019, April 30). *Myth-busting memorandum #4: Strengthening engagement with industry partners through innovative business practices*. Office of Federal Procurement Policy. https://thecgp.org/images/OFPP-Myth-Busting-4-Strengthening-Engagement-with-Industry-Partners-through-Innovative-Business-Practices.pdf

GAO. (2014, September). *Standards for internal control in the federal government* (GAO-14-704G). https://www.gao.gov/assets/gao-14-704g.pdf

Halcrow, S., & Schwartz, M. (2022, July 20). *The power of many: Leveraging consortia to promote innovation, expand the defense industrial base, and accelerate acquisition.* George Mason University Center for Government Contracting. https://business.gmu.edu/news/2022-07/power-many-leveraging-consortia-promote-innovation-expand-defense-industrial-base-and

House of Representatives. (2021). *DoD Appropriations Bill, Report to accompany FY21 appropriations*. https://democrats-appropriations.house.gov/sites/evo-subsites/democrats-appropriations.house.gov/files/FY21%20Defense%20Full%20committee%20report.pdf?_gl=1*b4forl*_ga*MTAxNDY1MTU5LjE3MzEwMDk5NjQ.*_ga_L2WB5KYYFC*MTc0MjMyNjI2NC4xLjAuMTc0MjMyNjI2NC4wLjAuMA

National Defense Industrial Association. (2025, February). *Vital signs 2025: The health and readiness of the defense industrial base*. https://www.ndia.org/-/media/sites/ndia/policy/vital-signs/2025/vitalsign_2025_final.pdf?download=1?download=1

Office of Federal Procurement Policy. (2024, June 18). *Application of cost accounting standards to indefinite delivery vehicles.* Federal Register. https://www.govinfo.gov/content/pkg/FR-2024-06-18/pdf/2024-12225.pdf

Office of Management and Budget. (2025, January 15). *Broadening public participation and community engagement with the federal government*. https://digitalgovernmenthub.org/library/m-25-07-broadening-public-participation-and-community-engagement-with-the-federal-government/

Olkkonen, R., Tikkanen, H., & Alajoutsijärvi, K. (2000). The role of communication in business relationships and networks. *Management Decision*, *38*(6), 403–409. https://doi.org/10.1108/eum0000000005365

Radovic Markovic, M., & Salamzadeh, A. (2018). The importance of communication in business management. In Radovic Markovic, M., & Salamzadeh, A., *The Importance of Communication in Business Management, The 7th International Scientific Conference on Employment, Education and Entrepreneurship, Belgrade, Serbia*.

S. Rep. No. 103-258 (1994).

Schooner, S. L. (2020, February 1). Enhanced debriefings: A toothless mandate? *34 Nash & Cibinic Report* 10 (2020). https://ssrn.com/abstract=3532671

Schwartz, M. & Johnson, M. (2023, May 1). How not to alienate business partners: A framework for addressing factors impacting retention of defense contractors. *Proceedings from the 20th Annual Acquisition Research Symposium at the Naval Postgraduate School*. https://dair.nps.edu/handle/123456789/4828

Tabansi, O. (2023, April 15). McDonald's supplier management principles. *Supply Chain Nuggets*. https://supplychainnuggets.com/the-mcdonalds-supplier-management-principles-and-why-it-works/

Vitasek, K., Downsard, A., & Kling, J. (2022, June). *The royal Australian navy FFG Enterprise: A journey from antagonism to high-performing relationship*. Vested Way. https://www.vestedway.com/wp-content/uploads/2022/06/FFG-case-study-TEACHING-case-PDF-OK-for-Distribution-May-2018-v2.pdf

Assessing the Impact of Department of Defense Weapons Systems on the Defense Industrial Base

Amanda Bresler—serves as President of PW Communications, Inc. She runs SHELDON, a subsidiary of PW Communications that provides custom analytics products and data-driven advisory services to federal and commercial clients. Prior to joining PW Communications, she worked as Chief Operating Officer for Maurice Cooper Brands. She serves on the board of directors of PW Communications; St. Dalfour SAS, a French food company; and Chatham International Inc. She graduated cum laude from Georgetown University's McDonough School of Business. [abresler@pwcommunications.com]

Alex Bresler—serves as Chief Data Officer for PW Communications Inc. He is a data-driven technologist, investor, and advisor to early stage venture funded technology companies. He is a data scientist and programmer with experience supporting clients in defense, financial services, law, real estate, and sports. He graduated from the Wharton School at the University of Pennsylvania. [alexbresler@pwcommunications.com]

Abstract

From fiscal year (FY) 2013 through FY 2023, the Department of Defense (DoD), on average, directed nearly 20% of spending into 400+ weapons systems (also referred to as equipment systems, major systems, or equipment programs). The DoD designates something as an equipment program (EP) based on factors including mission criticality; the extent to which system capabilities depend on a combination of hardware, software, and equipment elements; and the level of resourcing required.

In this paper, we leverage public data from the Federal Procurement Data System (FPDS) to analyze more than $800 billion in DoD funding allocated for EPs from FY2013–FY2023. We explore the features of these contract awards, and the pool of entities that received this funding. We examine the challenges faced by the DoD in delivering EP capabilities on time and on budget and explore possible causes for these cost and schedule overruns. We offer a series of recommendations for new policies and protocols that will enable the DoD to better manage these programs and ensure they serve as a strategic boon to the military's critical mission priorities.

Introduction

The concept of a "weapon system"—also referred to as an "equipment program" (EP) or "major system"—emerged in the Department of Defense (DoD) in the 1950s. They refer to mission critical, technically complex items that depend on a combination of hardware, software, and equipment elements (Fox, 2011). Since 1970, the Government Accountability Office (GAO) has reported on ballooning costs associated with EPs, and the share of the defense budget allocated to EPs grew significantly starting in the 1980s. From fiscal year (FY) 2013 through FY 2023, the DoD directed nearly 20% of procurement dollars into 400+ weapons systems annually.

Despite efforts to reform acquisition policy and increase program oversight, EPs remain plagued by cost overruns, scheduling delays, and routinely failing to reach technological milestones (GAO, 2024). In this paper, we leverage public data to analyze more than $880 billion in DoD funding allocated for EPs between FY2013 and FY2023, with the goal of understanding how resourcing for EPs reflects DoD mission priorities, and the extent to which this resourcing has contributed to broader trends in the defense industrial base. We also consider the role of EPs in today's world, particularly given the proliferation of low-cost weapons technologies like drone swarms and the threat of cyber-attacks. We offer recommendations for improving the management, oversight, and resourcing of EPs to ensure they meet their objectives and most effectively respond to the current, dynamic threat landscape.

Background: EP Investment

To analyze data pertaining to weapon systems, we aggregated contract award data from the Federal Procurement Data System (FPDS), the centralized, real-time database for government procurement transactions. We then filtered the data to isolate contracts funded and awarded by the DoD from FY2013 through FY2023. FPDS contains a structured field for EP, so we then isolated contracts awarded and funded by the DoD for which that field was completed. Figure 1 shows the total DoD procurement outlays annually (contracts funded and awarded by the DoD), and the share of these outlays associated with an EP.

FY	Total DoD Procurement (Funded & Awarded)	Total DoD EP Procurement	Total DoD Non-EP Procurement	% Total DoD Procurement on EPs
2013	$269,017,594,888	$78,207,054,095	$190,810,540,794	29%
2014	$256,118,674,160	$65,748,247,606	$190,370,426,554	26%
2015	$246,499,913,917	$59,781,559,745	$186,718,354,172	24%
2016	$272,614,429,221	$71,262,605,215	$201,351,824,006	26%
2017	$298,820,054,772	$76,382,601,753	$222,437,453,019	26%
2018	$351,276,245,245	$79,910,910,067	$271,365,335,177	23%
2019	$386,118,124,062	$84,853,729,518	$301,264,394,543	22%
2020	$436,878,063,385	$110,928,352,549	$325,949,710,836	25%
2021	$401,837,146,161	$71,019,308,770	$330,817,837,391	18%
2022	$377,897,271,584	$68,343,434,135	$309,553,837,450	18%
2023	$521,941,878,995	$116,384,440,405	$405,557,438,590	22%

Figure 1. DoD EP Procurement Spend as a Share of DoD Procurement, Annually

We then grouped the EP-associated contract actions by EP to isolate the unique EPs funded during our analysis period. More than 100 EPs had a negative or $0 spend associated with them, so we filtered for those with positive total obligations.[1] **From FY2013 through FY2023, the DoD funded 440 distinct EPs.** Figure 2 shows the count of unique EPs that received funding in each year.

[1] Negative or zero balances can occur for several reasons, including instances in which there were de-obligations and/or if the timing of payments required readjustment.

FY	Count EPs >$0
2013	247
2014	222
2015	205
2016	179
2017	185
2018	167
2019	148
2020	142
2021	152
2022	181
2023	242

Figure 2. Count of EPs by Year

We observed significant variability in the procurement funding received by different EPs. To explore this distribution, we grouped EPs by total procurement funding during our analysis period and split them into bins. As shown in Figure 3, while hundreds of EPs have received less than $100 million in total DoD procurement, the majority of EP spend is concentrated in the multi-billion-dollar weapons systems.

Total DoD Obligations to Individual EPs, Binned, FY2013–FY2023	Count EPs	Total Obligations Within Bin, FY2013–FY2023	%Total EP Obligation
$0 to $100,000	38	$1,244,003	0.0001%
$100,001 to $10.00M	143	$333,696,056	0.0377%
$10.01M to $15.00M	14	$164,473,492	0.0186%
$15.01M to $100.00M	71	$3,088,036,706	0.3493%
$100.01M to $500.00M	59	$14,909,087,532	1.6864%
$500.01M to $1.000B	26	$18,439,306,018	2.0857%
$1.01B to $10.00B	73	$280,189,888,103	31.6933%
> $10.00B	16	$566,940,194,467	64.1287%
	Total	$884,065,926,377	

Figure 3. Total DoD Obligations to EPs by Bin

Figure 4 provides a list of the 16 EPs that received more than $10 billion in DoD contracts during our analysis period, including a brief description of the program.

EP	Description	Total Obligations, FY2013–FY2023
F-35	Joint Strike Fighter program; largest DoD procurement initiative in history	$193,684,357,325
MISSILE DEFENSE AGENCY SUPPORT	Encompasses ballistic missile development and sustainment	$73,315,797,734
SSN 774	Virginia-class submarine	$54,081,887,858
DDG 51	Arleigh Burke-class destroyers	$26,971,995,363
TRIDENT II MISSILE	Three-stage, solid-fuel, inertially-guided submarine-launched ballistic missile	$26,690,118,950
KC-46A	Aerial refueling and strategic military transport aircraft	$25,718,471,267
C-130J	Tactical airlift	$24,798,475,323
P-8A	Multi-mission maritime patrol and reconnaissance aircraft	$22,352,861,892
V-22	Vertical takeoff aircraft	$20,261,130,937
LCS	Littoral combat ship	$16,800,964,672
CVN 78	USS Gerald R. Ford aircraft carrier	$16,085,769,254
GMLRS/GMLRS AW	Guided Multiple Launch Rocket System/Alternative Warhead	$14,213,226,852
E-2D AHE	Hawkeye early warning aircraft	$14,155,874,423
NSSL	National Security Space Launch program	$13,999,811,494
MRIC	Medium Range Intercept Capability	$13,795,895,289
EA-18G	Electronic warfare aircraft	$10,013,555,833

Figure 4. EPs in Receipt of More Than $10 Billion, FY2013–FY2023

EP Suppliers: Prime Contractors and Future Research

Next, we were interested in understanding the pool of entities that received EP contracts. We recognize that a significant share of EP contract dollars trickle down to lower-tier suppliers, and we recommend future research incorporate second- and third-tier supplier data. While we have access to subcontract award data from USASpending, it is not as comprehensively reported as prime contractor data. For the purposes of this paper, we opted to focus our analysis on the prime contract awards associated with EPs to establish higher-fidelity baseline metrics to understand this pool of funding.

To identify the companies that performed as prime contractors on EP contracts, we joined the Unique Entity Identifier (UEI) for each EP-associated contract action. Because many entities have won multiple EP contracts, the number of distinct UEIs is significantly lower than the number of contract actions. So, we grouped the contracts by their UEIs and filtered for distinct UEIs. Many large USG contractors operate with multiple UEIs, so we then manually joined UEIs for entities clearly associated with the same parent company.

For the last 30+ years, the defense industrial base has experienced significant consolidation. Many of the largest DoD suppliers have combined, with prominent acquisitions including but not limited to:

- Northrop Grumman acquiring Orbital ATK
- Lockheed Martin acquiring Sikorsky
- Raython merging with United Technologies
- L3 Harris acquiring Aerojet Rocketdyne

For the purposes of our analysis, we reviewed the parent companies with the most in EP awards and, in instances when we were aware of acquisition events, we consolidated the companies to a parent company (the acquirer). For instance, we merged Orbital ATK data with Northrop Grumman, and treated Northrop Grumman as the parent. In future research, we recommend analyzing the full list of suppliers for acquisition events and comprehensively merging them accordingly.

In instances where two primes formed a joint venture, we treated the joint venture as its own entity. Doing so established an important distinction between mergers, wherein one supplier ceases to exist upon acquisition by another, and instances in which two large corporations team together and collectively expand their market share.

As a final step, we filtered the entities to exclude those with less than $1,000 in total EP-associated contract awards during our analysis period.

Using this methodology, we identified 5,677 companies in receipt of EP-associated contract awards from FY2013 through FY2023. Of these, 31 companies received more than $1 billion in EP contract obligations during our analysis period. These 31 companies are shown in Figure 5.

Company	Total EP Obligations, FY2013–FY2023
LOCKHEED MARTIN	$325,775,571,468
BOEING	$106,539,237,903
GENERAL DYNAMICS	$76,773,885,401
RAYTHEON	$70,673,827,378
NORTHROP GRUMMAN	$45,188,718,754
HUNTINGTON INGALLS INC	$43,586,145,027
RTX CORPORATION	$34,816,197,577
BELL BOEING JOINT PROJECT OFFICE	$16,814,755,793
UNITED LAUNCH SERVICES LLC	$12,550,261,676
BAE SYSTEMS	$10,949,171,046
GENERAL ATOMICS	$10,045,561,769
OSHKOSH DEFENSE LLC	$7,509,296,463

TEXTRON	$7,422,797,922
GENERAL ELECTRIC	$6,780,981,312
ALTUS LLC	$6,175,240,351
L3HARRIS	$5,888,520,692
ROLLS-ROYCE	$4,491,107,655
THE CHARLES STARK DRAPER LABORATORY INC	$4,228,047,194
BECHTEL GROUP INC	$3,416,990,216
BECHTEL PARSONS BLUE GRASS A JOINT VENTURE	$3,279,476,001
LEIDOS	$3,161,722,506
MARINETTE MARINE CORPORATION	$2,597,768,554
DATA LINK SOLUTIONS LLC	$2,260,784,585
JACOBS SOLUTIONS	$2,120,243,764
ROCKWELL COLLINS	$2,036,122,964
PARSONS GOVERNMENT	$2,028,913,569
VIASAT INC	$1,870,841,569
HONEYWELL INTERNATIONAL INC	$1,794,148,371
AIRBUS US SPACE & DEFENSE INC	$1,725,905,650
AM GENERAL LLC	$1,527,618,805
INTREPID LLC	$1,096,217,819

Figure 5. Companies in Receipt of $1 Billion+ in EP Contract Obligations, FY2013–FY2023

These 31 companies collectively received over $824.7 billion in EP contracts during our analysis period. ***Their EP obligations represent more than 93% of all funding allocated to major weapons systems.***

Vendor Location

Next, we wanted to understand the geographic composition of the EP supplier base. For each EP company, we joined location data from FPDS and grouped the companies by state (which includes Washington D.C., Guam, Puerto Rico, and “Foreign Domicile”). For companies with associated entities in multiple locations, we counted them towards all their affiliated states. As shown in Figure 6, for more than half of all states in the United States, EP vendors based there received more than $1 billion in EP obligations from FY2013 through FY2023.

State	Total EP Obligations, FY2013–FY2023	Count of Companies
TEXAS	$247,720,372,425	310
CONNECTICUT	$102,983,367,768	103
CALIFORNIA	$68,478,507,576	756
WASHINGTON	$52,797,191,229	107
ARIZONA	$50,816,898,480	103
MASSACHUSETTS	$40,340,952,261	177
VIRGINIA	$37,854,945,792	620
FLORIDA	$33,206,498,316	380
ALABAMA	$28,021,098,009	225
GEORGIA	$25,979,526,143	162
MISSISSIPPI	$24,660,803,858	13
COLORADO	$23,098,395,659	180
MARYLAND	$20,408,024,233	287
MISSOURI	$19,292,534,913	83
PENNSYLVANIA	$17,488,340,482	191
MAINE	$13,679,585,307	7
NEW YORK	$11,899,040,803	272
WISCONSIN	$10,391,228,263	53
NEW JERSEY	$8,018,859,612	153
FOREIGN DOMICILED COMPANY	$7,443,502,691	412
UTAH	$7,169,025,909	70
INDIANA	$6,821,952,309	65
MICHIGAN	$5,861,095,984	116
IOWA	$4,344,003,861	26
KENTUCKY	$3,373,241,948	39
OKLAHOMA	$2,429,407,965	59
TENNESSEE	$2,117,022,861	70
OHIO	$1,683,778,766	171
ALASKA	$1,079,751,795	97

OREGON	$1,036,814,499	34
ILLINOIS	$1,006,151,785	123
MINNESOTA	$796,069,508	58
KANSAS	$694,035,251	45
NEW HAMPSHIRE	$675,129,427	48
DISTRICT OF COLUMBIA	$590,062,144	57
HAWAII	$402,782,320	60
RHODE ISLAND	$397,886,704	18
VERMONT	$343,351,638	12
NEBRASKA	$325,905,972	23
SOUTH CAROLINA	$291,468,733	48
NORTH CAROLINA	$261,939,746	90
NEW MEXICO	$220,870,745	60
NEVADA	$199,173,064	25
LOUISIANA	$194,657,159	32
IDAHO	$124,485,410	27
WYOMING	$86,013,002	5
UNKNOWN	$69,800,195	58
WEST VIRGINIA	$59,474,712	9
MONTANA	$48,596,200	17
SOUTH DAKOTA	$15,979,145	9
DELAWARE	$12,005,277	17
GUAM	$11,419,653	4
ARKANSAS	$9,172,269	16
NORTH DAKOTA	$2,511,921	8
PUERTO RICO	$578,815	1

Figure 6. EP Obligations by Company Location

Next, we wanted to explore the distribution of EP funding geographically, by place of performance. As shown in Figure 7, for more than half of all states, EP-associated contract actions generated more than $1 billion from FY2013 through FY2023.

State	Total EP Obligations, FY2013–FY2023	Count of Companies
TEXAS	$245,703,261,538	392
CONNECTICUT	$103,546,096,630	120
CALIFORNIA	$65,241,164,831	872
WASHINGTON	$54,379,427,806	152
ARIZONA	$51,697,040,676	161
MASSACHUSETTS	$39,431,327,923	192
VIRGINIA	$33,402,833,607	784
FLORIDA	$32,872,362,877	416
ALABAMA	$32,787,442,440	504
COLORADO	$28,249,642,443	229
GEORGIA	$25,859,021,760	568
MISSISSIPPI	$24,969,074,007	32
MISSOURI	$18,353,049,992	138
MARYLAND	$17,062,662,205	311
PENNSYLVANIA	$16,022,653,679	198
MAINE	$13,625,615,419	32
NEW YORK	$13,026,229,281	279
FOREIGN PERFORMANCE LOCATION	$11,521,691,713	690
WISCONSIN	$10,091,781,012	64
NEW JERSEY	$8,232,738,091	202
UTAH	$7,339,249,418	100
INDIANA	$6,599,433,795	98
MICHIGAN	$5,361,100,442	130
IOWA	$4,270,715,296	45
KENTUCKY	$3,466,299,929	63
OKLAHOMA	$2,164,069,019	120
OHIO	$1,672,251,952	219
DISTRICT OF	$1,562,930,809	243

COLUMBIA		
HAWAII	$1,035,298,083	130
KANSAS	$885,462,461	85
ALASKA	$874,358,958	148
ILLINOIS	$843,020,898	133
NEW HAMPSHIRE	$743,635,576	50
MINNESOTA	$704,530,944	54
SOUTH CAROLINA	$572,728,826	65
NEW MEXICO	$475,277,947	124
OREGON	$397,328,733	59
RHODE ISLAND	$395,911,130	24
VERMONT	$347,029,400	14
ARKANSAS	$279,747,777	34
NEVADA	$228,765,189	73
NEBRASKA	$223,287,526	76
NORTH CAROLINA	$195,561,899	87
GUAM	$143,087,546	39
WEST VIRGINIA	$134,465,667	14
TENNESSEE	$93,768,631	72
LOUISIANA	$87,728,270	41
MONTANA	$80,177,383	38
WYOMING	$76,220,925	30
SOUTH DAKOTA	$63,643,598	33
NORTH DAKOTA	$51,464,989	29
IDAHO	$47,994,039	40
PUERTO RICO	$32,677,550	11
DELAWARE	$23,560,059	23
NORTHERN MARIANA ISLANDS	$20,269,176	6

Figure 7. EP Obligations by EP Contract Place of Performance

High-Stakes Contracts and a Handful of Suppliers with Tremendous Influence

This analysis demonstrates that a handful of suppliers (31) are at the helm of the most significant EPs. These programs not only account for nearly 20% of DoD procurement dollars, but also serve to represent the most complex and sophisticated American defense capabilities. We recognize there are valid reasons for the DoD to rely on a small number of companies for its largest, most complex weapons systems. By design, few companies possess the combined resources, technical expertise, and experience to meet the requirements for these programs. In addition to having access to top technical talent, these firms must be extraordinarily well capitalized to manage the costs associated with designing, producing, and delivering complex systems. They must also have the ability to rapidly and securely identify and integrate thousands of lower-tier, often globally-distributed, suppliers. And they must have extensive knowledge of and experience working with the USG and DoD.

However, reliance on such a small number of suppliers poses significant security risks. To the extent a major supplier experiences production issues or otherwise cannot perform, the DoD has no alternatives. Changes to the global threat landscape mean relationships maintained by these large firms with lower-tier suppliers internationally can suddenly become problematic. Given the contract dollars at stake, these prime contractors may not be aptly incentivized to proactively elevate potential conflicts/security risks.

Thus, Congressional oversight is critical. Yet the fact that these programs drive such significant revenue into so many states arguably gives them—and their suppliers—political cover.

Consequences: Cost and Schedule Overruns

It is not surprising, then, that cost overruns, delays, and production issues have plagued many of the largest weapons systems. For instance:

- The F-35 program delivered aircraft 10 years behind schedule and 80% over budget (La Monica, 2023). As of April 2024, estimated sustainment costs for the fleet through 2088—$1.6 trillion—were 44% higher than estimates produced in 2018 (DiMascio, 2024).
- The Virginia-class submarine program has existed since FY1998. Production has never managed to reach two boats per year, as the program intended, and, since 2022, has not exceeded 1.2 boats per year (Congressional Research Service, 2025). Estimated cost overruns exceed $17 billion, and the rapid expansion of China's maritime fleet means the production shortfalls pose a significant national security risk (Suciu, 2024).
- According to a January 2025 Congressional Budget Office Report, costs for the Navy's Arleigh Burke destroyers have ballooned from $2.1 billion per hull to $2.5 billion per hull. Costs are expected to continue to rise, while production delays routinely range from six months to more than two years (Congressional Budget Office, 2025).
- A 2017 GAO report looking into the Navy's Ford-class aircraft carriers found that production costs for the initial ship were $2 billion more than estimated. They concluded that the cost estimate for the second aircraft carrier was "not reliable and does not address lessons learned from the performance of the lead ship" (GAO, 2017, p. 18).
- The Air Force's KC-46 program has resulted in $7 billion in cost overruns and multi-year delays (Losey, 2024).

Conflicting Interests: Returns vs. National Security

Another important consideration, in light of the persistent cost and performance issues associated with EPs, is the fact that the majority of the largest EP suppliers are publicly traded companies. Figure 8 denotes which of the EP suppliers with $1+ billion in EP procurement are publicly traded.

Company	Is Publicly Traded?
LOCKHEED MARTIN	Yes
BOEING	Yes
GENERAL DYNAMICS	Yes
RAYTHEON	Yes
NORTHROP GRUMMAN	Yes
HUNTINGTON INGALLS INC	Yes
RTX CORPORATION	Yes
BELL BOEING JOINT PROJECT OFFICE	Joint Venture between Boeing & Textron (two public companies)
UNITED LAUNCH SERVICES LLC	Technically private, but 50/50 JV between Boeing & Lockheed Martin (two public companies)
BAE SYSTEMS	Yes
GENERAL ATOMICS	No
OSHKOSH DEFENSE LLC	Yes
TEXTRON	Yes
GENERAL ELECTRIC	Yes
ALTUS LLC	No
L3HARRIS	Yes
ROLLS-ROYCE	Yes
THE CHARLES STARK DRAPER LABORATORY INC	No, nonprofit
BECHTEL GROUP INC	No
BECHTEL PARSONS BLUE GRASS A JOINT VENTURE	No (but Parsons is)
LEIDOS	Yes
MARINETTE MARINE CORPORATION	No
DATA LINK SOLUTIONS LLC	Technically private, but JV between Collins and RTX (two public companies)

JACOBS SOLUTIONS	Yes
ROCKWELL COLLINS	Yes
PARSONS GOVERNMENT	Yes
VIASAT INC	Yes
HONEYWELL INTERNATIONAL INC	Yes
AIRBUS US SPACE & DEFENSE INC	Yes
AM GENERAL LLC	No
INTREPID LLC	No

Figure 8. List of Publicly Traded EP Suppliers with $1+ Billion in EP Procurements

The fact that these firms serve as the backbone of the largest and most significant weapons systems in America does not absolve them of their fiduciary obligation to maximize shareholder value. How, then, do these suppliers balance the need to maximize shareholder value if and when doing so may not align with America's defense and national security needs?

For instance, if a changing threat landscape requires a supplier to abandon production in a certain part of the world and manufacture elsewhere: from the perspective of shareholder value, it could be better for the supplier to fight this change and/or delay implementing the new procedures, rather than swiftly pursue the new course of action. Likewise, EP contracts can extend years or even decades. With no explicit stipulations from the government that the prime contractor integrate innovative new technologies over the course of the contract, what incentive does the prime have to do so? In fact, if new innovations have the potential to reduce the government's dependency on legacy aspects of the prime contractor's system, they could be incentivized to thwart the diffusion of innovation, which could come at great cost to America's national security.

This misalignment also presents itself for many of the large private companies that supply to the DoD. For instance, AM General is owned by a private equity (PE) firm. Generally speaking, PE funds are incentivized to leverage balance sheets, reduce headcount, and otherwise increase profitability to generate a higher internal rate of return (IRR). These objectives may not align with the best interests of a defense end-user. Furthermore, there can be limited transparency into the investors that contribute to PE funds, known as limited partners (LPs). As such, there is a risk of nefarious LPs gaining information about critical defense technologies, and otherwise putting American security at risk via their investments.

To the extent the DoD continues to direct substantial contract dollars into major weapons systems while relying on a small number of companies for the delivery of these capabilities, these misaligned incentives must be addressed. We recommend that the DoD, the current administration, and policymakers establish a task force focused specifically on weapons systems. One major focal point for this task force should be addressing the fundamental disconnect between the needs of the military, the best interests of the external suppliers it relies on, and the taxpayers that fund this work.

The Role of EPs Today

DoD stakeholders and policymakers must address the inherent supply chain risks and performance issues that have plagued EPs. However, they must also consider a bigger-picture question: What role should large-scale weapons systems play in today's world, given how asymmetric warfare has transformed the battlefield?

For instance, low-cost drone swarms have the ability to handicap or even down multibillion-dollar assets. Based on our analysis of annual procurement allocations for EPs, this new reality does not appear to have materially impacted resourcing for EPs. In what ways, then, is the DoD responding? The largest defense contractors exert tremendous influence over the development of defense technologies. Given that they stand to lose billions in revenue if the DoD changes course on investing in large-scale weapons systems, what other voices can participate in this conversation to ensure it remains objective? For many people that devote years in military or civil service, the logical next step in their career is to work for a defense contractor—their skills are transferable, and they understand the customer. However, to the extent that people are concerned about career opportunities after service, it is critical to consider how this "revolving door" may affect their objectivity in evaluating performance and making contracting decisions.

Regardless of changes to the threat landscape, major weapons systems remain critical to American military dominance, both tactically and defensively. It is essential that stakeholders involved in resourcing and delivering these systems protect their integrity at all costs. Doing so demands difficult conversations about the relationship between the public and private sectors, and what new incentives and rules should be implemented to ensure parties' priorities align.

References

Congressional Budget Office. (2025, January). *An analysis of the Navy's 2025 shipbuilding plan* (CBO Publication No. 60732). https://www.cbo.gov/publication/60732

Congressional Research Service. (2025, February 11). *Navy Virginia-class submarine program and AUKUS submarine (Pillar 1) project: Background and issues for congress* (CRS Report No. RL32418). https://sgp.fas.org/crs/weapons/RL32418.pdf

DiMascio, J. (2024, December 11). *F-35 Lightning II: Background and issues for congress* (CRS Report No. R48304). Congressional Research Service. https://www.congress.gov/crs-product/R48304

Fox, J. R. (2011). *Defense acquisition reform, 1960–2009: An elusive goal.* Center of Military History United States Army. https://history.defense.gov/Portals/70/Documents/acquisition_pub/CMH_Pub_51-3-1.pdf

La Monica, P. R. (2023, May 11). *Lockheed Martin's $1.7 trillion F-35 fighter jet is 10 years late and 80% over budget—and it could be one of the Pentagon's biggest success stories.* Fortune. https://fortune.com/longform/lockheed-martin-f-35-fighter-jet/

Losey, S. (2024, January 9). *Cautionary tale: How Boeing won a U.S. Air Force program and lost $7B.* Defense News. https://www.defensenews.com/industry/2024/01/09/cautionary-tale-how-boeing-won-a-us-air-force-program-and-lost-7b/

Suciu, P. (2024, December 27). *Virginia-class submarine: $17,000,000,000 over budget and years behind.* The National Interest. https://nationalinterest.org/blog/buzz/virginia-class-submarine-17000000000-over-budget-and-years-behind-212894

U.S. Government Accountability Office. (2006). *Defense acquisitions: Major weapon systems continue to experience cost and schedule problems* (GAO-06-368). https://www.gao.gov/assets/gao-06-368.pdf

U.S. Government Accountability Office. (2017, June). *Ford-class aircraft carrier: Follow-on ships need more frequent and accurate cost estimates to avoid pitfalls of lead ship* (GAO-17-575). U.S. https://www.gao.gov/products/gao-17-575

U.S. Government Accountability Office. (2024). *Weapon systems annual assessment: DOD is not yet well-positioned to field systems with speed* (GAO-24-106831). Report to Congressional Committees. https://www.gao.gov/assets/gao-24-106831.pdf

Tactics to Strengthen the U.S. Defense Industrial Base with Private Capital

Sam Moyer—is a Research Fellow at the National Defense Industrial Association's Emerging Technologies Institute (ETI), where he focuses on private investment trends, supply chain analysis, and technology policy. Prior to ETI, he was a management consultant and researcher at Garter, a research and advisory firm. He has worked for more than a decade in defense technology policy, with a focus on innovative contracting, small business, and technology transition. He holds a BA in Economics from the College of William and Mary and an MBA from the University of Maryland. [smoyer@ndia.org]

Abstract

Private capital is a major source of research and development and capital investment dollars in the U.S. defense industrial base (DIB). As a strained budget environment limits the Department of Defense's (DoD) ability to capitalize defense supply chains, private capital is one resource available to help fill the gap.

This paper provides a structured assessment of both well-established and emerging tactics the DoD uses to engage private capital—ranging from multi-year procurement to demand aggregation, credit enhancements, and catalytic co-investment. Drawing on interviews with more than 30 stakeholders across government and industry and incorporating detailed case studies, the paper illustrates how these tactics can be deployed individually or in combination to mobilize private capital and strengthen the DIB.

Introduction

Key decisionmakers in the Department of Defense (DoD) have come to view the capital markets as potentially essential stakeholders in enhancing and expanding the defense industrial base (DIB). The 2023 National Defense Industrial Strategy states, "We need to build a modernized industrial ecosystem that includes . . . finance streams, especially private equity and venture capital."

The challenge is certainly urgent: in contrast to the United States, which invests on average about 20% of its gross domestic product (GDP) across its economy each year, China invests 43% of its GDP—capital investment that is used to expand China's industrial base of factories and strategic infrastructure. If the DoD is to successfully engage in great power competition it must draw upon a defense and dual-use industrial base of comparable size to China's. Increased investment by private capital in the DIB may help the United States to keep pace.

Private capital investment in the DIB carries with it a number of distinct benefits. First, private capital is used by DIB companies to invest in research and development, or R&D. While the DoD funds some R&D in federal labs, and indirectly through IR&D reimbursement, many private companies are using private investment dollars to support large amounts of critical R&D that can create new capabilities for the warfighter. Similarly, company capital expenditures ("capex") are used to build new factories, buy machinery, and scale up production. This activity can be critical when the DoD wishes for domestic supply chains and onshoring of critical production.

Private capital is also used to provide an "exit" for existing owners of defense firms who wish to retire, realize a return on their investment, or transfer management of their enterprise. This capital is typically provided during an initial public offering (IPO) or through mergers and acquisitions (M&A). Capital invested in the M&A process can also assist companies with commercializing their products and services for DoD end users.

DoD supply chains are strengthened when companies have greater access to capital. Well-capitalized companies find it easier to secure funding from reputable, regulated institutions such as commercial banks, which not only support business growth but are also a prerequisite for participation in certain DoD programs. By contrast, companies with limited access to capital are more susceptible to bankruptcy, liquidation, or acquisition by adversarial entities.

Private investors are often experts in growing companies and helping them commercialize their products and services. The DoD often struggles to encourage technology transition, and investors can be a critical part of the commercialization team to achieve this goal.

Since private capital can typically be allocated much more rapidly than government dollars, investor-backed companies can more readily accelerate products and services from research through fielding and production, allowing the government to quickly understand a commercial product's relevance to the warfighter and make necessary modifications to enable sales to DoD end users. Moreover, private capital can help diversify funding streams for companies, allowing them to carry on their work during DoD funding gaps and continuing resolutions.

Private investors can also serve as a critical alternative source of capital for firms whose leadership and business strategies demonstrate strong market potential—factors that may be underemphasized in DoD source selection processes.

As can be seen in Figure 1, the American defense sector already attracts a significant amount of private capital. By most metrics, the venture capital and private equity industries invest a similar amount of capital in the defense industrial base as the largest defense contractors and DoD itself.[1]

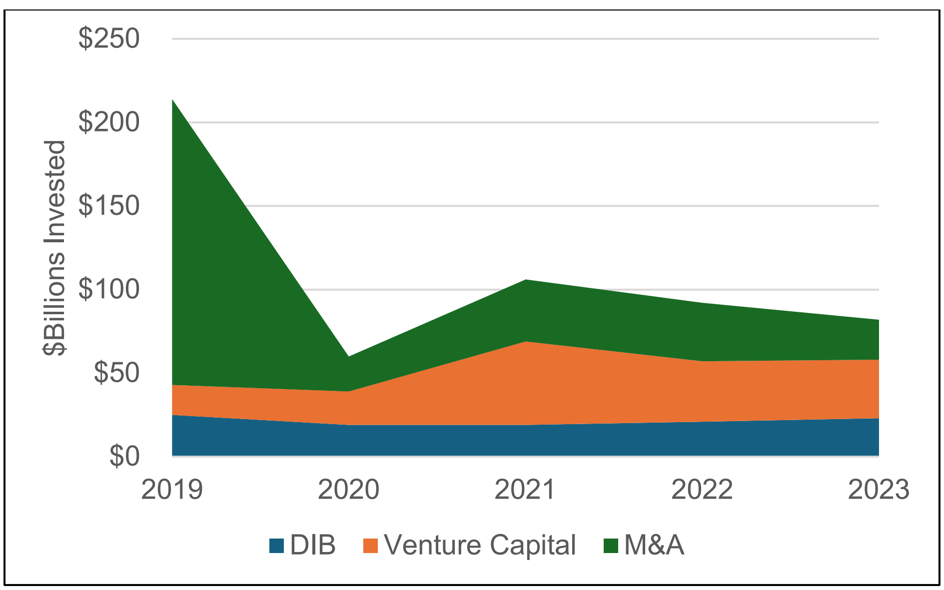


Figure 1. Investment in the U.S. Defense Sector, by Sour

[1] It is difficult to present precise, apples-to-apples comparisons for DIB investment among different sources of capital. See the appendix for an explanation of the different sources of capital and how they are measured.

Private capital is drawn into the U.S. defense sector for a variety of reasons: investment returns are relatively stable compared to the commercial sector and defense budgets are growing worldwide. In addition, the DoD uses a range of tactics to engage private capital and encourage investment in the sector.

Methodology

This paper seeks to provide a structured overview of well-established, as well as emerging tactics, that the DoD uses to engage private capital. The findings are based on interviews conducted with more than 30 individuals representing the full range of stakeholders involved in defense investment, including:

- Investors representing multiple asset classes within the capital markets, including private equity, venture capital, commercial banking, and others.
- Investor-backed companies representing multiple defense sub-sectors.
- Government personnel with experience engaging private capital.

When speaking with interviewees, it became evident that there are many different types of tactics for the DoD to engage private capital, addressing different parts of the problem space. Since no one single tactic allows the DoD to significantly increase private capital investment, it may be helpful to view the problem holistically.

This paper proposes a taxonomy that organizes the DoD's private capital engagement tactics into three categories: demand signal enhancement, catalytic capital, and dealmaking capabilities. This framework is intended as a practical tool to help researchers, policymakers and practitioners identify and apply relevant tactics.

1. Demand Signal Enhancement: Refers to tactics that help companies and investors forecast a clearer path to revenue. These include tools that improve visibility into the DoD's purchasing intent—such as memoranda of understanding (MOUs), fixed-price contracts, multi-year procurement authorities, and prize competitions. Each of these mechanisms helps reduce uncertainty about the size, timing, and likelihood of future defense sales, which in turn improves the financial case for investment.

2. Catalytic Capital: Financial tools that reduce risk for private investors by providing early government funding or favorable capital structures. These include co-investment programs, matching funds, loan guarantees, and other credit enhancements administered at the federal, state, or local level. These tactics are designed to "crowd in" private capital by acting as a signal of confidence and reducing downside risk.

3. Dealmaking Capabilities: The organizational tools, skills, and authorities that allow DoD personnel to structure investable transactions. These include the use of flexible contracting authorities like Other Transactions Agreements (OTAs), engaging in matchmaking between companies and investors, training DoD staff in commercial deal structures, and cultural or organizational changes that help the DoD act as a more predictable and responsive customer.

Each of these categories plays a distinct role in allowing the DoD to productively engage private capital. These categories are not mutually exclusive—in fact, as the case studies later in this paper demonstrate, they are most effective when used in combination.[2]

The sections that follow present the specific tactics within each category, describe how they function, where they've been applied, and how they can be combined.

Demand Signal Enhancement

The primary way that the DoD stimulates private investment is by providing a demand signal, since companies and investors will only invest capital if they can be assured that there is a reasonable chance of recouping their investment, plus a profit or "return on investment" (ROI). It's no surprise that 65% of defense industry executives agreed that it is important for the DoD to "provide a clear, consistent demand signal through contract vehicles" (National Defense Industrial Association, 2025). Demand signal is akin to a clear weather forecast for someone embarking on a risky voyage. (See box, "How Investors Evaluate Demand," for more details.) The DoD's demand signal is chiefly, though not exclusively, tied to its role as a buyer of products and services through programs of record, so demand signal enhancement tactics typically involve stakeholders in program executive offices (PEOs), program offices, and end user communities. The DoD has a number of tactics available to signal demand, ranging from highly informal to agreements that are similar to purchase guarantees.

How Investors Evaluate Demand

Professional investors, as well as companies, use a number of rigorous analytical approaches to evaluate demand. Most important to investors is the amount of dollar profits that a given demand signal represents. For example, this can be calculated as the number of anticipated unit sales multiplied by an expected profit margin. Since demand signals must be adjusted for risk and uncertainty, investors must carefully estimate the true likelihood of a future sale. In the commercial technology investment sector, the starting point for a sales forecast is typically called the total addressable market, or TAM.

Reducing the time it takes for a company to obtain sales can also have a very significant impact on the calculated demand signal through a financial mathematical principle known as the time value of money: the value of money that arrives sooner is worth more to investors than money that arrives later, because that money has an opportunity cost (Fernando, 2024). One commonly used formula known as "net-present value" allows investors to compare the financial value of different investments which return money on different timelines. As an illustration, using the net-present value formula, a project that delivers $1 million in 1 year, instead of over 5 years, is worth $290,000 more to the investor, thus justifying that much more investment.

Reports and Rhetoric: Because DoD resourcing decisions are typically made years in advance through the Planning, Programming, Budgeting, and Execution (PPBE) system, demand signals can often be inferred from internal planning documents such as the Future Years Defense Program (FYDP) and the Program Objective Memorandum (POM). However, since these documents are not publicly available, companies and investors must rely on less detailed but related sources. These include agency and service-level reports, public statements from senior defense and military leaders, and official budget documents—such as the President's Budget Request and the annual Defense Appropriations Bill—which collectively

[2] It must be noted that the most powerful factors driving investment in the DIB are outside of the DoD's control, such as the overall interest rate and regulations controlling capital markets.

provide insight into Programs of Record and other planned research, development, and procurement activities for the current and upcoming fiscal years. Some companies and investors can even gain access to controlled or classified information about defense programs and expected battlefield threats. Documents like the National Defense Strategy are typically less useful since they contain little if any budget information.

This information is extended and reinforced in rhetoric from a range of DoD stakeholders, most notably senior acquisition executives for the services, and others in the programming community, such as PEOs and program office leaders. Multiple constructs are used by the DoD to communicate pre-solicitation information to industry about future demand, such as Advanced Planning Briefings for Industry (APBI), industry days, proposer's days, requests for information (RFIs), and Technical Exchange Meetings. These each represent tactics that acquisition offices use to communicate with industry about intended future solicitations and awards. Although ethical rules about pre-solicitations apply, there is no standardization concerning how these events are carried out and how much information is conveyed to industry.

Unfortunately, it is not typically feasible to derive a useful TAM estimate based on leadership rhetoric and technical and budget information alone. Key information, such as unit sales forecasts and unit prices, is difficult to infer. In addition, due to classification challenges, the DoD is often unable to clearly explain its true demand signal beyond the small circle of individuals who possess security clearances.

DoD personnel are most effective at signaling demand to industry by communicating clear metrics or specifications tied directly to long term strategies that address DoD needs. Some PEOs share detailed strategies that provide companies with critical context for how their technology could be used. For example, the U.S. Navy PEO Digital created a highly detailed, public strategic roadmap containing organizational goals, outcomes sought, and specific metrics targeted (e.g. "reduce network downtime"), as well as an overview of the portfolio of 138 technology offerings overseen by the PEO (Navy Program Executive Office, 2024). In addition, PEO Digital supplies clear criteria that will be used to engage with any company, such as "support 10% of users uniquely in the [Department of the Navy]."

In general, demand signal communicated by DoD rhetoric suffers from a lack of credibility with industry. Especially during times of strategic turbulence, such as during a presidential transition, even seniormost DoD leaders may not have accurate perceptions of future DoD product demand. In addition, information sharing can be plagued by a "tyranny of abundance," whereby companies and investors lack an ability to track the multitude of communication channels that the DoD uses.

Test and Experimentation Events: Test, experimentation, or demonstration events provide venues for companies and investors to receive detailed feedback on their products or prototypes. Most importantly, such events provide an opportunity for stakeholders in the DoD acquisitions and end-user communities to rapidly inform their concepts of operation for using products and services and to refine the requirements that could be used to create a program of record for eventual product sales. In a recent survey of the defense industry on the value of DoD prototyping, respondents ranked "ability to communicate with government customer on requirements" as the most valuable element of a prototyping project, ahead of other elements such as "time to award" (Seraphin and Halcrow, 2025).

For an example of this tactic in action, in Spring 2023, the Joint Fires Network (JFN) initiative was launched by then–INDOPACOM commander Admiral Aquilino to create a C2 capability to coordinate joint fires. In collaboration with the Office of the Secretary of Defense (OSD) Rapid Defense Experimentation Reserve (RDER), a series of demonstration events was launched beginning in December 2023, focused first on modeling and simulation, followed by technical demonstrations, and finally full integration in-theater by April 2024—12 months after project conception (Miles et al., 2024). Rapid iteration within a series of test and experimentation events enabled the companies involved, such as SAIC and Anduril Enterprises, to deploy products and services quickly and with a high degree of information about product-to-market fit (Pomerleau, 2024).

Personalized feedback from government personnel during test and experimentation events can help companies understand the demand signal for their products and services, allowing them to forecast a more credible TAM and reduce uncertainty about their product-to-market fit.

Memoranda of Understanding: MOUs or Memoranda of Agreement (MOAs) are used by the DoD to add specificity and credibility to demand signals. In simple terms, an MOU serves as a formal handshake, outlining clear paths for future purchases.

Historically, MOUs have been used at different levels of the DoD to create or communicate demand signals for companies or other offices within the DoD (e.g., DARPA).[3] One approach is for an MOU to quantify the DoD market for a product or service in development—such as the platform, program of record, or specific program office which will acquire the product—as well as create a soft commitment for such offices to seek appropriations in the POM and FYDP. The MOU can then be signed by relevant individuals, such as service and agency acquisition leadership, science and technology executives, and other senior leaders.

The Air Force AFWERX Small Business Innovation Research (SBIR) program has required MOUs be signed to create credible demand signals for small businesses receiving SBIR Phase II awards. Such MOUs are typically signed by representatives of both the acquisition and end-user communities, who can provide commitments that a demand signal exists. Other AFWERX MOU signers can include representatives from government offices that will be essential for enabling the acquisition in practical terms, such as finance, contracting, small business offices, legal, information assurance/cybersecurity, engineering, public affairs, or security (e.g., to supply CAC cards).

MOUs cannot provide legal guarantees to companies and investors, but can instead provide descriptions of credible pathways to acquisition. By the same token, by forcing government personnel to create an acquisition plan in coordination with relevant offices, the act of creating MOUs may increase the probability of eventual transition.

Demand Aggregation: In some situations, the DoD requires products that are similar or the same as those purchased by commercial buyers. In these situations, the DoD can partner with these commercial buyers, to project a larger, aggregated demand signal.

Demand aggregation also enhances demand signals by providing diversification. With diverse paths to product sales, the risk of a total loss is lower for the company. Even when funds are in the POM and FYDP, DoD sales may be blocked by continuing resolutions, changes in priorities, unexpected cuts, or other risks. DoD personnel can therefore diversify and aggregate demand by finding potential additional government buyers, such as other program offices or joint offices. Even if the commitments from those offices are not firm, or are relatively small, the diversity of buyers itself enhances demand signal (Perley, 2024).

[3] Interview with Kathleen Harger, former Deputy Assistant Undersecretary of Defense for Innovation and Technology Transition, December 19, 2024.

Demand Aggregation Case Study: MCEIP Critical Chemicals[4]

Under the DoD's Critical Chemicals Pilot, the Office of Industrial Base Policy (IBP) Manufacturing Capability Expansion and Investment Prioritization (MCEIP) Directorate has leveraged $177 million in private capital to create domestic supply chains for 12 essential chemicals for the DoD. By aggregating its demand signal with U.S. commercial industry buyers, the DoD increased the total addressable market for chemical productors and therefore was able stimulate private capital to invest in new domestic supply facilities for critical chemicals.

This project began with a critical problem for the DoD: many critical chemical supply chains were sourced from companies in high-risk nations. MCEIP addressed this challenge using three tactics:

- Demand Aggregation—working with a commercial chemical company to adapt their chemical engineering approach to produce a critical chemical.
- Commercial Market Adaptation—supporting the certification of a lower-cost, commercially-available material for use in place of an existing, domestically unavailable, and more exacting military specification.
- Process Innovation—developing a modern production process to enable future domestic production for multiple critical chemicals (ACMI Group, 2024).

MCEIP first exhaustively catalogued DoD chemicals demand by convening the DoD Critical Energetic Materials Working Group (CEMWG), comprising experts from the defense laboratories, the acquisition community, joint warfighting, and interagency communities. The group developed an initial list of critical chemicals for which the DoD had potential supply chain vulnerabilities. MCEIP also drew from a list of chemicals restricted from foreign import in the 2023 NDAA, as well as solicitations prepared by the Defense Industrial Base Consortium (DIBC).

In parallel, MCEIP built a network in the relevant commercial chemicals production and buyer communities. At industry associations like the Society of Chemical Manufacturers and Affiliates (SOCMA), MCEIP learned that several large commercial chemical buyers had overlapping supply chain vulnerabilities with the DoD and also wished to shift purchases to U.S. domestic production companies.

Next, working with the American Center for Manufacturing and Innovation (ACMI)—the lead performer on this program—MCEIP conducted workshop events bringing together the industry and government stakeholders involved. Critically, these in-person workshops enabled the chemical production companies to understand the combined commercial and government demand signal, and justify the use of their own private capital to create domestic supply chains for the relevant chemicals.

MCEIP developed a pilot project, launched in July 2022, setting a target of mobilizing $50 million in private capital investment against $5 million in DoD funding (10:1), focused on onshoring eight critical chemicals to U.S. chemical production companies (DoD, 2022). To be selected, the chemical production companies were required to show an investment level of 10:1, private to public, before contract award, with the ability to stimulate additional capital during execution. During execution, MCEIP required monthly reporting on private capital leveraged for the project and prospects on generating new private capital. Production companies were also held to these goals during program management reviews of their work. For example, Lacamas Laboratories, a commercial contract manufacturer of high-quality

[4] Based on interviews with Christopher Zember, Senior Advisor and Portfolio Manager, Manufacturing Capability Expansion Pathfinders

pharmaceutical intermediates and fine chemicals in Portland, Oregon, used a combination of MCEIP funding and private capital to develop a fully domestic supply chain for 1,3,5-Trichlorobenzene (TCB), which is the first domestic U.S. production of this critical chemical in 15 years. Lacamas has since acquired a production facility valued at $110 million to be used to support scaling production of TCB and other critical chemicals for defense and commercial applications (Chemicals Knowledge Hub, 2025).

With successful proofs-of-concept for four chemicals, and validated demand signal from the DoD and commercial buyers, these commercial chemical production companies exceeded the $50 million goal, with $80 million in private investment secured. In September 2023, the program expanded to additional critical chemicals and by January 2025 achieved a remarkable 25:1 leverage ratio, translating to $177 million in private capital to address barriers to domestic production of 12 critical chemicals through a combination of demand aggregation, commercial product adaptation, and process innovation, demonstrating viable domestic sourcing for DoD and commercial market needs (ACMI Group, 2024).

Fixed price contracts: Fixed price contracts, even if budgeted at the same dollar level as cost-plus contracts, can sometimes create a more attractive demand signal for companies and investors. Therefore, switching from a cost-plus to a fixed price contract can sometimes expand the demand signal for a company, without requiring the DoD to obligate more money.

This results from several factors. First, fixed price contracts often raise profit margins for companies, since they can involve lower administrative costs, including tracking of costs and labor hours using a distinct accounting system. Most importantly, if a company is able to reduce its costs significantly below the price of the fixed milestone payments, it can retain any cost savings as profits. A company that is confident in its ability to continuously reduce costs through efficiency gains will prefer fixed price contracts over cost-plus contracts, for which profit potential is limited at a static percentage of costs.

Prizes: Like fixed price contracts, prizes provide a fixed payment to a company in exchange for success criteria. Prizes differ by broadcasting the opportunity to the general public or some subset of companies, such as a cohort of pre-selected companies. If a prize is of sufficient size, it will represent a demand signal that investors and companies can use to justify investing private resources into a project.

The DoD possesses several prize authorities which can be tailored to specific types of projects. Prize awards can be cash or other inducements, such as contract awards (Dunn, 2019).

Multi-Year Appropriations Authority: Funding gaps caused by a lack of appropriated funds are a common cause of reduced demand signal. Congress sometimes chooses to authorize and appropriate funds for multiple years, or until expended. This latter approach, commonly called non-expiring or “no year” funding, alleviates investor and company concerns about funding gaps (Congressional Research Service, 2024). A common source of no-year dollars is the Title III and Defense Production Act program, which can be used if expressly authorized by presidential determination.

Advance Procurement: Advance procurement contracting authority allows the DoD to provide funds to companies for major components before delivery of final products for which the appropriation exists. Advanced procurement must be authorized for a procurement program in statute. Per DoD 7000.14-R, advance procurement can be justified either for products with long lead times, or for situations where buying in bulk can bring down unit costs (referred to as Economic Order Quantity procurement). Advance procurement decreases uncertainty to companies and investors. By providing government dollars more quickly, advance procurement also enhances demand signal through the time value of money principle.

Multi-Year Procurement: Multi-year procurement is a contracting tactic provided by Congressional statute and is typically tied to specific programs of record (Multiyear Contracts: Acquisition of Property, 2020). Specifically, multi-year procurement provides a planned set of product purchases for up to five years. While each year of payments under the contract depends upon annual appropriations from Congress, a cash payment “cancellation ceiling” is provided to compensate companies in the case of government deferral from its procurement obligation. Multi-year procurement is one of the strongest demand signals that the DoD can create. Under current processes, multi-year procurement is complex to implement and must be tied to programs of record and meet specific requirements as defined in statute and regulations, including: substantial savings (the typical benchmark is at least 10%), stable end-user requirement, stable product design, and enhancement of national security (Defense Acquisition University, 2025).

Securitization of Leases: Guaranteed, or near-guaranteed, revenue streams for statutorily qualifying large and reliable projects allow companies to raise capital at very attractive terms. For example, the DoD regularly procures facilities, housing, or solar energy installations using long-term leases authorized under 10 U.S.C. § 2667 (Enhanced Use Leases), 10 U.S.C. § 2871–2885 (Military Housing Privatization Initiative), or agreements structured through Energy Savings Performance Contracts.

These lease and performance contracts provide highly secure, legally enforceable demand signal from the DoD—payment guarantees that may extend for 25 years, or even beyond 50 years in some circumstances. Companies are therefore able to finance the capital expenditure for these projects with bonds, typically the lowest cost form of financing available to companies.

In sum, demand signals are essential for private capital to invest in the DIB, and stronger signals can induce more investment. A range of tactics exist to provide demand signals of varying strength. Many of these tactics are only available in highly specified circumstances, defined by a combination of law, regulation and policy, and require express permission from a range of DoD and external stakeholders, including Congress and the president. Demand signal is often blocked by frictions such as classification challenges and an overwhelming or contradictory information contained in DoD forecasting documentation.

Catalytic Capital

In addition to stimulating investment with demand signals, the DoD can also financially incentivize companies to raise investment capital during the R&D and capital expenditure stages. Through the use of catalytic capital, the DoD acts as an initial or “anchor” investor by providing early funding that triggers private investors to contribute their own capital to support a project or company. A diverse array of government stakeholders can typically be involved in contributing catalytic capital, including RDT&E funders or dedicated entities such as the Office of Strategic Capital (OSC).

Investment Capital as Pre-award Criteria: The DoD employs a range of mechanisms during the source selection and pre-award phases that, directly or indirectly, incentivize private investment. The DoD sometimes uses investor participation as selection criteria for award: programs such as AFWERX sometimes score proposals more highly when private investors have invested in the proposing company, as indicated in letters of intent or capitalization tables[5] that can be provided as part of proposal packages.

In addition to influencing evaluation criteria, private investment and financial strength can also factor into pre-award responsibility determinations, where contracting officers assess whether offerors possess adequate financial resources to perform.

The DoD's systems for assessing the financial health of companies, such as DCMA financial capability reviews, and contracting officer required financial reviews, further incentivize companies to pursue financial stability, to include securing private capital. For example, DFARS 232.072 (2025) requires contracting officers obtain financial information from companies, such as balance sheets and income statements to "perform a financial review" of contactors.

Federal, State, and Local Credit Programs: The federal government administers a number of credit programs that provide catalytic capital to "crowd in," or catalyze the entry of private capital into targeted companies or projects. There are a range of tools, usually called "credit enhancements," that can fall under this heading, including:

- Direct loans to qualifying companies, sometimes at subsidized or reduced interest rates.
- Loan guarantees for qualifying projects or companies.
- Subsidies to investment funds which invest in qualifying domains.

Agencies such as the U.S. Development Finance Corporation, the Small Business Administration, the departments of Energy and Commerce, and others have each launched federal credit programs, which use one or more federal credit enhancements (Murphy et al., 2024).

To trigger engagement with a project or company, these government programs typically require a large majority of project funds to be derived from private capital. Letters of intent or other conditional commitments are provided by the private co-investors, which are then triggered when the government delivers its capital infusion. Proposal selection criteria may ask companies to demonstrate that every other source of private capital was exhausted before the government was approached. This tactic has been used extensively by the Department of Commerce semiconductor fabrication facility loan program and the Department of Energy Loan Program Office. State and local governments, and economic development agencies, also frequently provide similar credit enhancements, frequently projects that create jobs in a particular locale.

The DoD's OSC, established in late 2022, has authority to provide a variety of credit enhancements, such as direct loans to companies and loan guarantees for qualifying investment funds focused on qualifying technology domains (DoD, 2025). OSC began taking applications for its unique "accrual debenture" loan guarantee tool, providing subsidized debt to small business investment companies (SBICs), which in turn invest in qualifying technology companies. The first cohort is expected to catalyze investments of more than $4 billion into more than 1,700 small businesses. Most recently, OSC received "more than 200 applications totaling $8.9 billion in financing requests" for its direct loan program ("equipment financing"); the program currently has lending capacity of $984 million (DoD, 2025).

Companies planning R&D or capital investments can review the criteria associated with the various federal, state, and local credit programs to determine if they can apply for catalytic capital from one or more programs. In many instances, credit enhancements can be layered,

[5] Also known as "cap tables," these internal company documents track equity ownership and can help the government infer how much private capital a company has raised over time.

bringing together catalytic capital from the federal, state, and local level alongside private capital.

Investment Matching: A powerful way to incentivize the investment of private capital in the DIB is to directly match private investment dollars with government dollars: in other words, paying companies to raise private capital. Investment matching functions similarly to matching charity donations: the government promises to put dollars into a project or company at a set ratio against every dollar committed by private investors. In practice, this arrangement can be carried out in a variety of ways.

> Investment Matching Case Study: AFWERX STRATFI
> In this example, the DoD directly catalyzed the investment of $27 million of private capital into X-Bow Systems, a New Mexico solid rocket company, to advance "rapidly produced, low-cost solid rocket motors (SRMs) using X-Bow's proprietary advanced manufacturing technology and culminate in a flight test series" (X-Bow, 2023).
>
> This arrangement was facilitated by the now well-established Air Force Strategic Financing, or STRATFI, program and carried out by the AFWERX office overseeing the Air Force SBIR program. Under STRATFI, for approved topic areas, Air Force SBIR dollars of up to $15 million, combined with up to $15 million Air Force non-SBIR dollars (for a total government-dollar cap of $30 million), can be matched against private investment dollars flowing into private companies. This 1:1 dollar match, up to $30 million, provides a strong financial incentive to raise private capital and contribute it to a company.
>
> The Air Force limits STRATFI awards to companies which have proposed R&D projects of relevance to a list of approved Air Force SBIR topics. Moreover, awards are limited to companies which have secured an MOU from a DoD acquisition office (such as a PEO) and end user, pledging to purchase and use final products and services if the development stage of the project proves successful.
>
> Since its STRATFI award, X-Bow has been selected to deliver solid rocket engines to multiple services as well as development contracts for the Mk 72 booster and Mk 104 dual-thrust solid rocket engines. In 2024, X-Bow received contracts totaling $60 million to advance energetics readiness at Naval Surface Warfare Center Indian Head Division (NSWC IHD; PR Newswire, 2024). X-Bow has also gone on to raise another $70 million in private capital to further advance its technology. (X-Bow, 2024)

The Air Force STRATFI program, discussed in the case study above, publicizes clear criteria and thresholds for companies and investors to meet to receive a pre-specified amount of government matching dollars, up to a cap. Because the terms of the investment match are transparent and allow no negotiation, STRATFI has proven to be a highly scalable program. In just one year, AFWERX used this construct to leverage $332 million of private capital against $606 million government funds and the program has continued to grow since then (AFWERX, 2023). Similarly, the other transactions carried out for prototyping require a pre-set one third of project resources to be contributed (provided by the company) in many circumstances (10 USC 4022).

Investment matching can also be conducted on a case-by-case basis. In the case study below, DoD dollars catalyzed a very large capital investment from a wide range of non-DoD stakeholders. This arrangement involved the sequenced injection of DoD capital alongside up to nine independent stakeholders. Unlike a typical acquisition plan, where proposals differ on concretely differentiated variables such as technical capability and price, an investment matching arrangement requires complex communication and negotiation at all stages of the acquisition process.

> Investment Matching Case Study: e-VAC
> This case study concerns the creation of a $550 million advanced magnets production facility, triggered by only $94.1 million in DoD dollars, constituting 17% of the total, or a 5:1 ratio of private capital combined with DoD dollars.
>
> In September 2023, the DoD's Office of Industrial Base Policy (IBP) awarded $94.1 million to Vacuumschmelze (VAC) a leading manufacturer of rare earth permanent magnets, to partially assist with the establishment of a large American manufacturing facility ("e-VAC") to ". . . acquire and install manufacturing equipment, operationalize technical infrastructure, and engineer production lines" (DoD, 2023). The magnet production facility will produce Neodymium Iron Boron (NdFeB) rare earth permanent magnets, a critical component of many defense products, such as high-performance engines and communications equipment.
>
> The process began with an NDAA provision requiring the DoD develop a "mine-to-magnet" supply chain free from covered nations. This created a credible DoD demand signal for an onshore magnet production facility. IBP created a funding announcement for onshore magnet production, requiring at least 1:1 (50%) project cost share. VAC submitted a white paper which included contingent commitments from private investors. The award was ultimately made via an Other Transactions Agreement using no-year (non-expiring) Title III dollars.
>
> Since the facility, which is located in Sumter County, South Carolina, is expected to create 300 jobs, state and local governments provided additional financial incentives, including job development credits, plus a total of $15 million in state grants to assist the county with site preparation, road improvements, water and wastewater improvements (South Carolina Department of Commerce, 2023). In addition, the Department of Energy provided a Qualifying Advanced Energy Project Tax Credit ("Section 48c") in March 2024, of $111.9 million. By reducing e-VAC's future taxes, this credit functions similarly in financial terms to a cash grant or co-investment for the facility.
>
> Finally, a strong commercial demand signal was provided by General Motors (GM), which provided a binding MOU to e-VAC, agreeing to purchase magnets from the facility to supply GM's growing fleet of electric vehicles, such as the Chevrolet Silverado and Cadillac Lyriq, for purchases of at least 10 years (Onstad, 2023).
>
> This combination of government catalytic co-investment, and commercial demand signal, was sufficient to unlock significant private capital investment: VAC's private equity owner, Ara Partners, announced that $335 million in private capital had successfully been raised to complete the construction of the facility, which is expected to begin production in fall of 2025 (Ara Partners, 2024). Representatives at Ara Partners remarked that they "are grateful for the support from our local and state governments and the federal initiatives that have made this project possible, and we extend our sincere thanks to General Motors for being a key partner in this endeavor" (2024).

Government Equity Investing: A less common approach to injecting catalytic capital is for the government to buy company stock in the manner of a venture capitalist or other equity investor. This approach is not normally authorized and must be expressly authorized by Congress. Among the most established programs of this type are the intelligence community's

IQT and the U.S. Army's OnPoint programs (now defunct), each of which was executed by a non-profit corporation at arms-length from the government. These types of programs typically participate in investment syndicates during equity investments, meaning that the government capital is joined in the investment by additional non-government sources of capital (e.g., private venture capital firms). In these programs, if companies are successful, the government receives a return as any other equity investor would, allowing this capital to be used for other investments.

Because Other Transactions Authority (OTA) provides flexibility in resource sharing, certain Other Transactions have also been structured to return capital to the government. For example, in the DoD's very first (1990) Other Transaction, DARPA awarded $4 million to Gazelle Microcircuits to develop high-speed gallium arsenide (GaAs). DARPA's funds were used for the "development, design, production engineering, and working capital to develop and bring to market high-speed data communication GaAs components, electronic modules or subsystems, and application development tools." In return, DARPA retained "access to research and development results; certain rights in data patents; and in the case of technology developments that resulted in commercially marketable products, a fair return on its investment and discounts for government purchases of such products" (Dunn, 2018).

To summarize, DoD funding used as catalytic capital can be a powerful tactic to "crowd in" private capital to strategic companies and projects. Catalytic capital can be a challenging tactic to deploy because it depends on the DoD understanding a number of variables:

- A company's true availability of capital. Companies have an incentive to downplay the availability of capital to secure DoD catalytic capital, which may come at more favorable terms (e.g., no-cost, in the case of investment matching).
- The technical feasibility and likelihood of recouping the catalytic capital. The DoD should reserve catalytic capital for projects that are likely to succeed and, if appropriate, repay the government, especially if the catalytic capital was a loan or equity investment.
- How much capital is likely to flow into a company or project if DoD catalytic capital is provided. Typically this must be validated by follow-up milestones holding companies accountable to their commitments to provide or raise capital.
- The sequencing and coordination of capital injection from multiple public and private stakeholders. The successful deployment of catalytic capital often involves layering funding and credit enhancements across federal, state, and local levels, which can be complex to coordinate due to differing application timelines, eligibility criteria, and oversight requirements.

Catalytic capital is more likely to succeed in tandem with other tactics, such as demand signal enhancement. Investors will not contribute capital into a company or project without a credible demand signal, and stronger demand signals can induce more private investment and therefore require fewer government investment dollars to catalyze them.

Dealmaking Capabilities

For the DoD, the process of engaging private capital, while beneficial, introduces risks and complexities. There are few established processes that can be followed that will reliably lead to strong private capital investment in the DIB. Engaging capital requires creativity and sophistication on the part of government personnel. For this reason, a number of dealmaking capabilities are essential to ensuring that the DoD engages private capital effectively.

Private Capital Ecosystem Development Programs: The DoD has instituted a number of formal as well as informal matchmaking and ecosystem development programs, such as the AFWERX Project Vanguard program, the OASD IBP Office of Industry Engagement, and the National Security Innovation Network, that can facilitate interactions between government personnel, defense companies, and the investment community. Many private sector venues of this type also exist.

By reducing transaction costs, these programs make it easier for defense companies and investors to become aware of each other, and therefore catalyze an investment. In addition, DoD personnel can often provide critical subject-matter expertise to investors and companies to help justify the deployment of capital. Lacking relevant technical expertise in many defense-relevant technology domains, awareness of DoD supply chain risks, or access to controlled or classified programmatic and threats information, investors often struggle to understand investment opportunities, and are thus deterred from investing. Face-to-face communications can also help investors and companies to more deeply understand the demand signals that have already been provided by the DoD (e.g., advanced planning briefings).

M&A Regulatory Environment: One of the main ways that companies return capital to investors is by being acquired or merged into another company, at which time cash is usually paid to existing shareholders. Therefore, a healthy and predictable M&A market is a major stimulus for investment in the defense industrial base.

For an example of how this works in practice, consider Oshkosh's acquisition of Pratt Miller. After its founding in 1989, Pratt Miller grew into a world-class advanced vehicle research center. When Oshkosh paid $115 million to acquire the company in 2021, the capital was used to pay Pratt Miller's shareholders in exchange for the company, which was then folded into Oshkosh, serving as an internal "Skunk Works" for the defense company (Yu, 2024).

DoD OASD IBP is charged with collaborating with the Federal Trade Commission (FTC) and Department of Justice (DOJ) as part of the Premerger Notification and Merger Review Process under the Hart-Scott-Rodino Antitrust Improvements Act (Federal Trade Commission, 2025). The DoD's role in this process is to assess and provide feedback on potential national security and defense industrial base implications of M&A transactions in the defense sector. The FTC and DOJ use this input to determine whether enforcement actions are required, such as blocking a merger, demanding divestures, or other remedies. PEOs and program offices may provide technical consultation as part of this process.

While not a direct regulator, DoD leadership can play an important advisory role in shaping M&A activity in the defense sector by monitoring industry investment trends and clearly communicating its priorities to companies, investors, and the regulatory community.

A notable example of this influence was demonstrated by the now-famous "last supper" dinner meeting, when former Deputy Secretary of Defense William Perry announced his wish for greater defense industry consolidation through M&A to cope with declining defense budgets; industry responded with a historic uptick in defense M&A investment activity (Mintz, 1997). Conversely, then-Secretary Ash Carter and then–Assistant Secretary for Acquisition, Technology and Logistics Frank Kendall signaled caution about the increase in industry concentration due to M&A throughout the early 2010s (Clark, 2015).

Being a Good Customer: As discussed above, when the DoD can reduce uncertainty, delay, and lower transaction costs, investors and companies perceive future DoD cash flows as larger, and therefore more attractive investment targets. At the same time, the warfighter receives goods and services closer to the speed of relevance. There are several tactics the DoD can use to mitigate delay and uncertainty, making itself a more appealing customer.

Transactions carried out under OTA are not subject to the Federal Acquisition Regulations and other related and derivative DoD regulatory requirements imposed on government grants and contracts. There are several distinct OTAs, including advanced research, procurement for experimental purposes, and prototyping OTAs, as well as prize authority. OTAs result in reduced administrative cost, such as FAR-based prescribed competitive procedures, burdensome Cost Accounting Standards (CAS) accounting practices, and Bayh-Dole prescribed IP regimes.

Since approval authority under OTAs can be delegated to the most relevant government stakeholders, such as program managers or innovation offices, decision-making can be significantly faster than under FAR-based contracts, which typically require formal legal and compliance reviews. OTAs also allow for faster and more flexible payment structures, including options like advance payments, that are often restricted under the FAR. Last, prototyping OTAs allow for direct-to-production contracts, greatly reducing potential time and uncertainty associated with transitioning to a procurement stage.

Another useful approach is FAR Part 12—Acquisition of Commercial Products and Commercial Services—which is designed to provide a streamlined contracting process for commercial items. Compared to FAR Part 15, FAR Part 12 involves a simpler solicitation, pricing, and contracting process. Since this approach was designed to pay companies on near-commercial terms, such as fixed-price contracts, companies are responsible for using their own capital to support production and delivery.

The newly-implemented adaptive acquisition framework (AAF) provides a range of new acquisition pathways that allow qualifying programs to speed through traditional acquisition checkpoints companies to traverse the acquisition process much more quickly than what had been the default Acquisition Category system that was previously central to DoD 5000 (GAO, 2024). Pathways such as the Middle Tier of Acquisition and Software Acquisition allow for rapid prototyping, rapid fielding, and iterative delivery of products and services—each of which can allow companies to get products into production more quickly, unlike the traditional path to a program of record which can take seven years or longer.

The payment terms of government contracts can also have a significant effect on the ability of companies to raise capital. Many commercial lenders and investors are used to investing in companies that receive regular, subscription-based payments from commercial customers. Therefore, companies whose government contracts have lengthy or irregular milestone schedules are often ineligible to borrow from regulated lenders, such as retail and commercial banks, forcing them to turn to unregulated lending markets which may offer less favorable lending terms, such as higher interest rates or less access to financial services.

Training and Culture: When the DoD engages private capital, it relies on specific DoD personnel to assemble investable deals for companies, using a variety of appropriate tactics, which may vary depending on the nature of the project. Negotiating with investors and investor-backed companies requires a deep understanding of the typical business practices in the financial services industry, types of investment capital, legal aspects of investing, commercial accounting practices, and financial concepts like the time value of money.

Training and education programs, such as those available at the Dwight D. Eisenhower School for National Security and Resource Strategy or Defense Acquisition University, can help the DoD workforce understand commercial business practices, including private investment.

The DoD has also used rotation programs to embed government personnel into investment companies or government innovation offices such as the Defense Innovation Unit, where they can absorb knowledge about innovative tactics for engaging private capital. Many

DoD innovation offices have chosen to physically embed their entire teams into hubs of commercial innovation, such as commercial startup accelerator facilities, giving government personnel informal exposure to investors and investor-backed companies (Shah, 2024). In some offices, the DoD has adopted relevant hiring practices: offices such as the Office of Strategic Capital have deliberately sought personnel with experience working in financial services.[6]

Beyond specific knowledge and skills, successfully engaging private capital requires empowering individuals with a unique mindset—one that is focused on creatively using government resources to create win-win transactions with industry. For such individuals to flourish inside the government, they must be embedded in an organizational culture that is comfortable articulating why a particular transaction will be profitable to a company, instead of viewing company profit as something to be avoided. Creating attractive investment opportunities in the DIB can take a great amount of time and labor, making it critical for organizational incentives to be established that prioritize private capital engagement.

Data and Information Technology Tools: For DoD personnel to effectively engage private capital, they must also be equipped with the appropriate tools to navigate the financial services and commercial technology industries. A range of information technology (IT) tools can play a supporting role, including:

- Market intelligence tools to track trends in the private investment markets, such as fundraising/investment trends, mergers and acquisitions trends, and real-time financial data (e.g., interest rates, valuations data).
- Financial modeling tools to assist with valuation analysis, comparables analysis, and benchmarking of deal terms.
- Industry analysis focused on tracking trends in specific global and commercial market segments, strategic positioning and technology strength of specific companies, and innovation tracking (e.g., using patent data or company announcements).
- Regulatory tracking and analysis tools to interpret the complex and evolving regulatory environment in the investment sector.
- Supply chain risk analysis platforms to assist with due diligence and risk analysis, including assessment of adversarial capital.
- Customer relationship management (CRM) tools to assist with tracking the large volume of stakeholders that can be involved in investment dealmaking.

These tools are routinely used by companies and investors, and many DoD personnel do have access to some tools that offer similar functionality, especially industry analyst platforms, CRM tools, and supply chain risk management (SCRM) tools. In particular, SCRM has been an area of increased investment by the DoD since the 2020 COVID pandemic disrupted global and DoD supply chains. Market intelligence tools like PitchBook and CrunchBase are in widespread use in the investment industry, providing analysis on technology companies and the broader investment environment; in recent years, many of these tools have added data on the defense technology industry.

Federally-funded research and development centers and open-source intelligence offices, such as the Air Force Office of Commercial and Economic Analysis, often provide

[6] For example, in one job listing OSC sought "[e]xperience serving as a credit or risk officer in a reputed financial institution, dealing with highly complex and large-scale (multimillion or billion-dollar) transactions . . . [and] developing credit risk models, which may include corporate finance, asset-based lending, and/or project finance transactions" (DoD, 2023).

analysis to help DoD personnel understand specific technology domains. The DoD also acquires a large volume of very detailed data on private companies from a variety of sources, such as Hart-Scott-Rodino M&A reviews, and financial data sharing under DFARS 232.072, but this data is often fragmented and difficult to access.

A number of programs, such as the AFWERX Project Vanguard program, the OASD IBP Office of Industry Engagement, and the National Security Innovation Network, have directly engaged segments of the private investor community, providing channels by which the DoD can learn about trends in the investment industry and directly engage.

Certain DoD offices also make use of commercial or homegrown market intelligence systems to monitor relevant investment and technology markets. Recently, the OSC has focused on creating an analytical toolkit focused on understanding investment, corporate, and technological trends that is "panoramic in scale," yet with "pinpoint accuracy" (DoD, 2025). The approach uses a set of homegrown analytical tools, such as network analysis, corporate finance analysis, capital flow mapping, and IP licensing mapping, in combination with data collected from RFIs and structured interviews in "key global financial centers—including, New York, Silicon Valley, Boston, Dallas, London, Dubai, Tokyo, Singapore, and Sydney—and contested markets in South East Asia, South America, and Africa."

IT tools can be useful for DoD personnel charged with executing specific programs by helping them understand relevant trends and analyze specific companies and investors. There may be opportunities to widen the availability of tools that already exist, or have been procured, through a federated model. If the DoD adopts goals or metrics relating to private capital, then market intelligence or dashboarding tools may be required for leadership to track those goals.

Leadership: As is the case for many innovative acquisition practices, the actions of DoD leadership play a critical role when the DoD engages private capital. As discussed above, demand signal in the form of rhetoric from DoD senior leaders is more meaningful if it is perceived as credible and consistent.

As has been seen in the case studies, a DoD process for engaging capital often requires many years of preparation, along with consistent messaging to investors and companies. Because there is high turnover among DoD leaders, who in turn have much discretion over program design, it is easy for leaders to deter private investment by making unexpected changes.

Engaging private capital usually involves forming coalitions of stakeholders both inside and outside the government. Within government, separate institutions may be responsible for creating demand signal, providing catalytic capital, and executing contracts. These activities must be carefully orchestrated to ensure they reinforce each other, and leadership is essential to conducting this orchestration. Senior leaders provide strategic direction, secure buy-in across government, and use their convening power to maintain and grow stakeholder coalitions. At the highest level, leaders may also be called upon to clearly explain novel business processes to Congress, the president, and interagency stakeholders to gain necessary approvals or resources. Leadership can also create incentives for cultural change when necessary, deliberately rewarding employees who successfully engage private capital.

Conclusion

The DoD possesses a wide and varied array of tactics for engaging private capital that can be powerful when combined wisely. None of these tactics will, on its own, unlock enough private capital to solve the DoD's budgetary challenges, but together they can make a difference. As seen in the case studies, when appropriate tactics are used in concert, a significant amount of private capital can be brought to bear to solve the DoD's most pressing challenges.

Appendix: Calculating Private Investment in the Defense Industrial Base

DIB Contractors

One of the largest sources of private capital investment in the DIB is defense contractors, which regularly conduct R&D and capex investments to support their work with DoD. Since the majority of DoD contract dollars flow to recipients that are publicly traded companies, R&D and capex trends among those companies is available in the filings they provide to the SEC.

Year	2023	2022	2021	2020	2019
DIB	$23B	$21B	$19B	$19B	$25B

For this analysis, the methodology used in the National Defense Industrial Association's Vital Signs report (2025) was used. For each year, the 20 largest publicly traded US defense firms were identified. Companies which received a majority of revenue from non-DoD sources (e.g. healthcare companies) were excluded. The total R&D and capital expenditure of those 20 companies was summed for each year.

M&A

Flowing from both strategic acquirers and private equity funds, M&A transactions are one of the major sources of private capital investment in the DIB. However, it is challenging to calculate an exact, or even approximate value for the amount of dollars invested in these transactions in any given year. Disclosure requirements for private transactions are inconsistent, and most transactions are not required to be reported to the public or government regulators.

Market analysts use a variety of methods to track the amount of M&A activity in a given industry. In certain situations, such as when a transaction may have a material impact on a public company, public announcements are required, or made voluntarily. In other instances, transaction data can be inferred from related public filing data.

For M&A transactions above $1B, investment data tends to be easier to collect. Data on M&A activity in the US aerospace and defense sector, supplied by Capstone Partners (2024), an investment bank, is shown in the table below.

Year	2023	2022	2021	2020	2019
M&A	$24B	$35B	$37B	$21B	$171B

This data does not include transactions below $1B, which tend to constitute the large majority of M&A transactions, so it is probable that the total dollars invested in aerospace and defense are substantially higher.

It should also be noted that for a given M&A transaction the purchase price may not strongly correspond to the amount of capital that company will invest in the future. In most instances, the dollars paid in M&A are used to pay shareholders of the company as part of the purchase. M&A transaction volume does provide an indication of the amount of private capital active in the industry, and potentially available to support new capex or R&D.

Venture Capital

Venture capital is typically invested by limited partnerships in small business or startups with high growth potential. Because venture capital typically targets rapid growth, much of the money invested into companies tends to be used for activities that DoD would consider RDT&E or capital expenditure, although funds could be used by companies for any other purpose, such as sales or management costs.

Although venture investments are typically private transactions, and therefore not required to be made public, in actuality both venture firms and companies receiving venture investment tend to publicize their transactions, making the data publicly available, if not comprehensive. The data used in this paper was provided by PitchBook (2024), a market intelligence service.

Year	2023	2022	2021	2020	2019
Venture Capital	$35B	$36B	$50B	$20B	$18B

It should be noted that these dollar numbers account are inclusive of investments by venture capital in any company deemed by PitchBook to have potential defense applications, including dual-use technologies like AI.

References

ACMI Group. (2024, February 8). *ACMI awarded contract extension on successful critical chemical pilot program* [Press release]. https://acmigroup.com/2024/02/08/acmi-awarded-contract-extension-on-successful-critical-chemical-pilot-program/

AFWERX. (2023). *2023 annual report*. https://afwerx.com/wp-content/uploads/AFWERX_2023-Annual-Report-Final_CLEARED_AFRL-2024-4054_WEB.pdf

Ara Partners. (2024, September 6). *Ara Partners secures $335 million of non-recourse financing for permanent magnet manufacturing facility.* https://www.arapartners.com/news/ara-partners-secures-335-million-of-non-recourse-financing-for-permanent-magnet-manufacturing-facility/

Authority of the Department of Defense to carry out certain prototype projects, 10 U.S. Code 4022 (d) (C). (2015). https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title10-section3501&num=0&edition=prelim

Capstone Partners. (2024). *Sector fundamentals strengthen while M&A activity declines.* https://www.capstonepartners.com/insights/report-annual-aerospace-and-defense-ma-activity-and-outlook/

Chemicals Knowledge Hub. (2025, January 15). *Lacamas Laboratories purchases Albany fine chemicals facility.* https://www.chemicalsknowledgehub.com/news/lacamas-laboratories-purchases-albany-fine-chemicals-facility/

Clark C. (2015, October 2). *Whoa, Lockheed & Co.! Kendall urges congress to protect innovation*. Breaking Defense. https://breakingdefense.com/2015/10/whoa-lockheed-co-kendall-urges-congress-to-protect-innovation/

Congressional Research Service. (2024, June 7). *Appropriations duration of availability: One-year, multi-year, and no-year funds.* https://www.congress.gov/crs-product/R48087

Defense Acquisition University. (2025). *Acquipedia: Multiyear procurement.* https://www.dau.edu/acquipedia-article/multiyear-procurement

DoD. (2022, August 29). *Defense department launches innovative manufacturing pilot program* [Press release]. https://www.defense.gov/News/Releases/Release/Article/3143134/defense-department-launches-innovative-manufacturing-pilot-program/

DoD. (2023, September 19). *Department of Defense enters an agreement to expand domestic manufacturing to strengthen U.S. supply chains for rare earth magnets*. https://www.defense.gov/News/Releases/Release/Article/3529874/department-of-defense-enters-an-agreement-to-expand-domestic-manufacturing-to-s/

DoD. (2023). *Supervisory loan specialist* [Archived job posting]. https://www.usajobs.gov/job/805507900

DoD. (2024, September 27). Notice of funding availability—Covered technology categories-equipment financing. *Federal Register, 89*(188), 79271–79278. https://www.govinfo.gov/content/pkg/FR-2024-09-27/pdf/2024-22229.pdf

DoD. (2025, January 2). *Investment strategy for the Office of Strategic Capital.* https://media.defense.gov/2025/Jan/02/2003623435/-1/-1/1/FY25-INVESTMENT-STRATEGY-FOR-OFFICE-OF-STRATEGIC-CAPITAL.PDF

DoD. (2025, April 1). *Office of Strategic Capital receives $8.9 billion in financing requests for first domestic manufacturing loan program* [Press release]. https://www.defense.gov/News/Releases/Release/Article/3143134/defense-department-launches-innovative-manufacturing-pilot-program/

Dunn, R. (2019a, September 18). *Prize authority—from prize to what?*. Strategic Institute for Innovation in Government Contracting. https://strategicinstitute.org/other-transactions/prize-authority/

Dunn, R. (2019b, September 12). *Origins and evolution of other transactions—Part 3.* Strategic Institute for Innovation in Government Contracting. https://strategicinstitute.org/other-transactions/origins-evolution-transactions-part-3/

Federal Trade Commission. (2025). *Premerger notification and the merger review process.* https://www.ftc.gov/advice-guidance/competition-guidance/guide-antitrust-laws/mergers/premerger-notification-merger-review-process

Fernando, J. (2024, August 21). *Time value of money: What it is and how it works*. Investopedia. https://www.investopedia.com/terms/t/timevalueofmoney.asp

GAO. (2024, December). *Military departments should take steps to facilitate speed and innovation.* https://www.gao.gov/products/gao-25-107003

Javaheri, A. (2024). *Vertical snapshot: Defense tech update*. PitchBook. https://pitchbook.com/news/reports/2024-vertical-snapshot-defense-tech-update

Maranda, J. (2023, May 3). *X-Bow lands $60M in funding to support its rocket, aerospace technology.* New Mexico Inno. https://www.bizjournals.com/albuquerque/inno/stories/news/2023/05/03/x-bow-lands-jolt-of-funding-to-mature-tech.html

Miles, W., Seraphin, A., Clark, B. (2024, November). *The value of experimentation, not demonstration.*

Mintz, J. (1997, July 3). How a dinner led to a feeding frenzy: With cuts on the way, defense executives were told to combine, conquer at “last supper” in ’93. *Washington Post*. https://www.washingtonpost.com/archive/business/1997/07/04/how-a-dinner-led-to-a-feeding-frenzy/13961ba2-5908-4992-8335-c3c087cdebc6/

Multiyear Contracts: Acquisition of Property, 10 U.S. Code § 3501. (2021). https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title10-section3501&num=0&edition=prelim

Murphy, E. L., Cook, C., Harding, E., Sanders, G., Velazquez, N., Dickson, J., & Sockwell, M. (2024, November 21). *Alternative funding mechanisms in Review*. CSIS. https://www.csis.org/analysis/alternative-funding-mechanisms-review

National Defense Industrial Association. (2025). *Vital signs 2025*. https://www.ndia.org/-/media/sites/ndia/policy/vital-signs/2025/vitalsign_2025_final.pdf?download=1?download=1

Onstad, E. (2023, January 30). *Germany's Vacuumschmelze agrees binding deal with GM on magnet factory.* Reuters. https://www.reuters.com/markets/commodities/germanys-vacuumschmelze-agrees-binding-deal-with-gm-magnet-factory-2023-01-30/

Perley, C. (2024). *Performance report*. Army Applications Laboratory. https://aal.army/assets/files/pdf/aal-performance-report.pdf

Pomerleau, M. (2024, July 19). *Joint force, international partners, contractors test command and control capabilities in Pacific exercise.* DefenseScoop.

PR Newswire. (2024, December 3). *U.S. Navy selects X-Bow Systems to modernize and automate energetics industrial base.* https://www.prnewswire.com/news-releases/us-navy-selects-x-bow-systems-to-modernize-and-automate-energetics-industrial-base-302321127.html

PR Newswire. (2024, July 10). *X-Bow raises more than $70 million Series B to accelerate expansion of hypersonic and solid rocket motor technologies.* https://www.prnewswire.com/news-releases/x-bow-raises-more-than-70-million-series-b-to-accelerate-expansion-of-hypersonic-and-solid-rocket-motor-technologies-302193599.html

Required Financial Reviews, DFARS 232.072. (2025). https://www.acquisition.gov/dfars/232.072-financial-responsibility-contractors.

Seraphin, A., and Halcrow, S. (2025). *From breakthroughs to the battlefield: Best practices for tapping into the power of prototyping.*

Shah, R. M. (2024). *Unit X: How the Pentagon and Silicon Valley are transforming the future of war*. Scribner.

South Carolina Department of Commerce. (2023, December 13). *E-Vac Magnetics to build first U.S. facility in Sumter County*. https://www.sccommerce.com/news/e-vac-magnetics-build-first-us-facility-sumter-county

U.S. Navy Program Executive Office. (2024). *Industry engagement overview*. https://www.peodigital.navy.mil/Industry/

X-Bow Systems. (2023, April 26). *X-Bow Systems announces selection for $60 million STRATFI award* [Press release]. https://www.prnewswire.com/news-releases/x-bow-systems-announces-selection-for-60-million-stratfi-award-301808692.html

Yu, K. (2024, June 1). *Mergers, acquisitions vital for defense innovation.*

Shipbuilding and Repair: Navy Needs a Strategic Approach for Private Sector Industrial Base Investments

Lindsey Cross—is a Senior Analyst with the U.S. Government Accountability Office, specializing in Navy shipbuilding acquisitions. She has worked in this issue area for nearly 8 years, reporting and making recommendations to improve shipbuilding programs outcomes. [crosslm@gao.gov]

Abstract

The Navy relies on private companies in the defense industrial base to build—and in many cases—repair its ships. In an era of strategic competition with adversary nations, performance in both shipbuilding and ship repair is critical to achieving the Navy's desired future fleet. However, GAO's recent work has shown that the Navy continues to fall short of its goals in these areas. To achieve its goals for the future fleet, the Navy—in partnership with the ship industrial base—will need to reverse these trends. For its current report, anticipated to be published in January 2025, GAO examined the ship industrial base's ability to meet the Navy's shipbuilding and ship repair goals, including by conducting 50 interviews with government offices and private companies.

Background

The Navy plans for a larger, more capable fleet of ships to counter evolving threats. However, by fiscal year 2026, the Navy expects to have no more ships than it did when it released its first 30-year shipbuilding plan in 2003. This is due to a combination of slower than expected new ship construction and the decommissioning of older ships. Its performance in shipbuilding and ship repair is critical to achieving the desired future fleet.

The private companies that the Navy contracts with to build vessels and repair surface ships are key components of the Navy's ship industrial base. These private companies augment the repair work conducted at the Navy's public shipyards.

Ship Industrial Base Struggles to Meet the Navy's Goals

The shipbuilding industrial base has not met the Navy's goals in recent history. Our prior work has shown that Navy shipbuilding has regularly fallen short of schedule and cost goals, and current performance is consistent with these trends.

Schedule. The Navy's 45-day review of its shipbuilding programs, completed in early 2024, states that its major shipbuilding programs continue to struggle with schedule delays. Our analysis found that schedule delays continue for most ships currently under construction, in addition to the number of ship delays reported in the 45-day review

Cost. Cost increases erode the Navy's buying power to execute its shipbuilding plan, particularly because the plan assumes that ships will be delivered in alignment with cost targets. Yet we found that many shipbuilding programs face cost overruns.

The Navy would need to deliver more ships at a quicker rate to meet its goals. Yet, the Navy continues to base its goals on an assumption that the industrial base will perform better on cost and schedule than it has historically.

The shipbuilders have infrastructure and workforce challenges that have made the Navy's goals difficult to accomplish. For example, our analysis found that shipbuilders have insufficient or aging infrastructure and struggle to hire and retain an appropriately trained workforce, which will make such improvements to performance difficult to accomplish.

Similarly, the Navy has historically not met its ship repair schedule goals, though it has achieved some improvements since 2019. The industrial base has grown since then, and representatives from some companies that GAO spoke with stated they often had more capacity than the Navy used. The Navy attributes some of these improvements to a change it made to its contracting strategy in 2015, which it stated has increased competition in the ship repair industrial base. Unlike in shipbuilding, in ship repair, there are often enough companies with capacity that there may be multiple companies able to compete for repair periods.

But companies may not be able to take on unplanned work due to infrastructure or workforce limitations. We found, however, that there is not always sufficient infrastructure capacity available to manage unplanned repair work, such as growth work or emergent repairs. Growth work refers to additional tasks identified during performance that is related to a work item already specified on the original contract, some of which may be identified after a repair period has begun. For example, a dry dock of the right size may not be empty when needed.

Department of Defense Invests Billions to Support the Shipbuilding Industrial Base

The Department of Defense (DoD)—specifically the Navy and Office of the Secretary of Defense (OSD)—spent billions to support the shipbuilding industrial base. This included funding for infrastructure and workforce improvements for shipbuilders and their suppliers. But it has yet to fully determine the effectiveness of that support (i.e., its return on investment), though it has taken steps to do so. More specifically, the DoD spent more than $5.8 billion on the shipbuilding industrial base from fiscal years 2014 through 2023. It plans to spend an additional $12.6 billion through fiscal year 2028. The DoD spent this funding on contract incentives and direct investments.

However, the Navy and OSD are not fully coordinating their shipbuilding investments to prevent duplication or overlap in spending. For example, the Navy and OSD do not coordinate across all investment efforts—such as between submarines and surface ships—though they both make related investments in workforce and infrastructure for these ship categories. Further, the Navy has yet to fully establish performance metrics, such as measurable targets that link to the agency's goals that would enable it to consistently evaluate the effectiveness of its investments in building a larger fleet or achieving other intended outcomes. However, the Navy has taken recent actions to make progress in this area, such as through the development of the Maritime Industrial Base Program Office. Without better visibility across investments and established performance metrics, the Navy and OSD cannot ensure their investments in the shipbuilding industrial base are an effective use of federal funds to help build a larger fleet.

The Navy Has Not Developed a Strategy for Managing the Ship Industrial Base

The Navy's current approach for managing the ship industrial base has been largely ineffective at encouraging private industry to invest independently. The Navy has sought to spur the industrial base to invest in infrastructure and workforce through its efforts to communicate stable demand. Yet, the Navy's reported methods for doing so—long-range planning and the use of contracting strategies intended to provide stability—have not resulted in sufficient industry investments to date to meet the Navy's capacity needs.

Further, the Navy does not have an industrial base strategy and has not had coordinated leadership to guide future efforts in this area. Developing a ship industrial base strategy would help the Navy better address these challenges to improve the likelihood of achieving its shipbuilding and ship repair goals. The GAO's prior work has shown that a consolidated and comprehensive strategy enables decision-makers to better guide program efforts and assess results. The DoD issued its national industrial strategy in November 2023. However, Navy officials told the GAO that it established a new program office in September 2024 that will be positioned to develop a strategy for the ship industrial base. Officials said they plan to have additional details available in early 2025. Until the Navy implements a ship industrial base strategy, it will not be able to effectively align or assess its actions to manage the industrial base for shipbuilding and repair.

This is an excerpt from a full length report. See GAO-25-106286 for additional details, including additional report contributors: gao.gov/assets/gao-25-106286.pdf

PANEL 7. ADVANCING ACQUISITION: LARGE LANGUAGE MODELS, KNOWLEDGE METRICS, AND SMALL BUSINESS IMPACT

Wednesday, May 7, 2025	
0950 – 1105 PT 1150 – 1305 CT 1250 – 1405 ET	**Chair: Dr. Raymond D. O'Toole, Jr., Acting Director, Operational Test and Evaluation, Office of the Secretary of Defense** ***Test and Evaluation of Large Language Models to Support Informed Government Acquisition*** Erin Lanus, Research Assistant Professor, Virginia Tech ***Knowledge based Metrics for Test and Design*** Craig Arndt, Principal Research Faculty, Georgia Tech Research Institute ***The Impact of the Joint Interagency Field Experiment (JIFX) on Small Business Success*** Ashley Book, Researcher, Naval Postgraduate School

Dr. Raymond D. O'Toole, Jr.—was appointed Acting Director of DOT&E on January 10, 2025. During his time at DOT&E he has served two previous stints as Acting Deputy Director.

Dr. O'Toole joined DOT&E on March 3, 2019, when he was appointed Deputy Director of Naval Warfare. On February 16, 2020, he was promoted to Principal Deputy Director, Operational Test and Evaluation.

Prior to this promotion, Dr. O'Toole was DOT&E's Deputy Director for Naval Warfare. In this capacity, he oversaw the operational and live-fire testing of ships and submarines and their associated sensors, combat and communications systems, and weapons. Dr. O'Toole also was responsible for overseeing the adequacy of test infrastructure and resources to support operational and live-fire testing for all acquisition programs across the Defense Department.

Before joining DOT&E, he was the Deputy Group Director of Aircraft Carrier Design and Systems Engineering at Naval Sea Systems Command (NAVSEA). Prior to that, Dr. O'Toole was the Director of Systems Engineering Division (Submarines and Undersea Systems). His other NAVSEA assignments included ship design manager and Navy technical authority.

Dr. O'Toole has more than 30 years of experience as a naval officer (active and reserve), retiring at the rank of captain. His significant tours included five as commanding officer.

Dr. Raymond D. O'Toole, Jr. earned a Bachelor of Engineering in marine engineering from State University of New York - Maritime College. He also holds a Master of Engineering in systems engineering from Virginia Polytechnic Institute and State University, a Master of Science in national resource strategy from the Industrial College of the Armed Forces, and a Doctorate in Engineering in the field of engineering management from The George Washington University where he taught as a professional lecturer of Engineering Management and Systems Engineering until April 2024.

Dr. O'Toole was awarded the U.S. Presidential Rank Award of Meritorious Executive in FY2024. He also has received the Secretary of Defense Meritorious Civilian Service Award twice and the Department of the Navy Meritorious and Superior Civilian awards.

Test and Evaluation of Large Language Models to Support Informed Government Acquisition

Jaganmohan Chandrasekaran—is a Research Assistant Professor at the Sanghani Center for AI and Data Analytics, Virginia Tech. His research is at the intersection of software engineering and artificial intelligence (AI) and focuses on developing frameworks, methods, and tools to address the multi-faceted challenges in the test and evaluation of AI-enabled software systems across its life cycle. He earned an MS and a PhD in Computer Science from the University of Texas at Arlington. [jagan@vt.edu]

Brian Mayer—is a Research Scientist at the Sanghani Center. Mayer works with sponsors and collaborators to help apply Sanghani faculty research to solve government, corporate, and societal problems. Mayer has helped deploy several long-term government-funded research projects to government sponsors. Mayer's research interests are in the areas of decision support, forecasting, intelligence, and human-guided systems. Mayer received an MS and BS in Industrial and Systems Engineering from Virginia Tech. [bmayer2@vt.edu]

Heather Frase—is the head of Veraitech, a Senior Advisor for Testing & Evaluation of AI at Virginia Tech's National Security Institute, and Program Lead for the AI Risk and Reliability working group at MLCommons. Her diverse career has spanned significant roles in defense, intelligence, and policy, involving projects ranging from AI assurance and policy to missile defense, drug trafficking prevention, and financial crime analysis. She also serves as a member of the Organisation for Economic Co-operation and Development (OECD) Network of Experts on AI and on the board of the Responsible AI Collaborative, which researches and documents AI incidents. [hnfrase@vt.edu]

Erin Lanus—is a Research Assistant Professor in the Intelligent Systems Division of the Virginia Tech National Security Institute and affiliate faculty of Computer Science at Virginia Tech. Her research focus is testing and evaluation of consequential AI. She applies her background in combinatorial testing to developing metrics for measuring coverage of datasets and algorithms for constructing test sets as well as identifying novel applications of combinatorial testing to AI assurance. She earned a BA in Psychology and a PhD in Computer Science both from Arizona State University. [lanus@vt.edu]

Patrick Butler—is a Senior Research Associate at the Sanghani Center. The main thrust of his research focuses on the IARPA funded EMBERS project that uses open-source indicators, such as tweets, news, blog, weather, etc., to forecast population level events such as civil unrest, elections, and epidemics. He graduated with a PhD in Computer Science from Virginia Tech in 2014 during which time he was awarded a SMART Fellowship. His dissertation was in the area of using knowledge discovery for intelligence analysis. [pabutler@vt.edu]

Stephen Adams—is a Research Associate Professor and Assistant Director of the Intelligent Systems Division in the Virginia Tech National Security Institute. He received an MS in Statistics from the University of Virginia (UVA) in 2010 and a PhD from UVA in Systems Engineering in December of 2015. His research focuses on applications of machine learning and artificial intelligence in real-world systems. He has experience developing and implementing numerous types of machine learning and artificial intelligence algorithms. His research interests include feature selection, machine learning with cost, transfer learning, reinforcement learning, inverse reinforcement learning, anomaly detection, explainable AI, and probabilistic modeling of systems. His research has been applied to several domains including fraud detection, activity recognition, prognostics and health management, cyber-physical systems, psychology, cybersecurity, internet-of-things devices, and sports analytics. [scadams21@vt.edu]

Jared Gregersen—currently serves as Lead Research Systems Engineer at the Virginia Tech National Security Institute. His experience includes working for more than 20 years within information technology providing infrastructure, application and database, and software development and 3D visualization. His roles have included positions within higher education, the energy sector, fortune 500 companies, healthcare, and he has supported local and federal government. His comprehensive expertise provides a unique capability and perspective in solving technical challenges and designing creative solutions for complex problems. His interests include research related to AI agents, and their implementation and interaction with information technologies and systems. [jaredgregersen@vt.edu]

Naren Ramakrishnan—is the Thomas L. Phillips Professor of Engineering at Virginia Tech and leads AI and machine learning for Virginia Tech's Innovation Campus. He directs the Sanghani Center for AI and Data Analytics and the Amazon-Virginia Tech Initiative in Efficient and Robust Machine Learning. Ramakrishnan's research interests span data science, forecasting, urban analytics, recommender systems, and computational epidemiology. He is a fellow of the Association for Computing Machinery (ACM), the American Association for the Advancement of Science (AAAS), and the Institute of Electrical and Electronics Engineers (IEEE). [naren@cs.vt.edu]

Laura Freeman—is a Research Professor of Statistics and serves as the Deputy Director of the Virginia Tech National Security Institute. She is also the Assistant Dean for Research for the College of Science. Her research leverages experimental methods in the domains of cyber-physical systems, data science, artificial intelligence, and machine learning to address critical challenges in national security. Previously, Freeman was the Assistant Director of the Operational Evaluation Division at the Institute for Defense Analyses. Freeman has a BS in Aerospace Engineering, an MS in Statistics, and a PhD in Statistics, all from Virginia Tech. [lamorgan@vt.edu]

Abstract

As large language models (LLMs) continue to advance and find applications in critical decision-making systems, robust and thorough test and evaluation (T&E) of these models will be necessary to ensure we reap their promised benefits without the risks that often come with LLMs.

Most existing applications of LLMs are in specific areas like healthcare, marketing, and customer support and thus these domains have influenced their T&E processes. When investigating LLMs for government acquisition, we encounter unique challenges and opportunities. Key challenges include managing the complexity and novelty of Artificial Intelligence (AI) systems and implementing robust risk management practices that can pass muster with the stringency of government regulatory requirements. Data management and transparency are critical concerns, as is the need for ensuring accuracy (performance). Unlike traditional software systems developed for specific functionalities, LLMs are capable of performing a wide variety of functionalities (e.g., translation, generation). Furthermore, the primary mode of interaction with an LLM is through natural language. These unique characteristics necessitate a comprehensive evaluation across diverse functionalities and accounting for the variability in the natural language inputs/outputs. Thus, the T&E for LLMs must support evaluating the model's linguistic capabilities (understanding, reasoning, etc.), generation capabilities (e.g., correctness, coherence, and contextually relevant responses), and other quality attributes (fairness, security, lack of toxicity, robustness). T&E must be thorough, robust, and systematic to fully realize the capabilities and limitations (e.g., hallucinations and toxicity) of LLMs and to ensure confidence in their performance. This work aims to provide an overview of the current state of T&E methods for ascertaining the quality of LLMs and structured recommendations for testing LLMs, thus resulting in a process for assuring warfighting capability.

Keywords: Large Language Models, Test and Evaluation, Government Acquisition, Generative Artificial Intelligence, Benchmarking

Introduction

Large language models (LLMs), a subset of generative AI, have demonstrated the potential to accomplish diverse activities with minimal or no human intervention. As a result, LLMs have found utility across domains, and recent developments have indicated there is an increasing interest among people across various domains in adapting and trying to leverage LLMs in their activities. However, successful adaptation of LLMs is contingent upon the ability to thoroughly evaluate and ensure these systems perform as expected after adaptation.

LLMs, similar to AI/ML systems, are data-intensive software systems. Unlike traditional software systems where the core functionality is encoded by a human (referred to as source code), the data-intensive systems derive their decision logic from a training dataset; this decision logic is commonly referred to as a model. An LLM, a type of deep learning system, is fundamentally a language model trained on a vast amount of training data, capable of performing a variety of tasks. Furthermore, these systems exhibit non-determinism, are stochastic, and have a decision logic that is not easily understandable to humans (opaque) Moreover, both the data and the algorithm used to train the model influence its behavior. Thus, traditional T&E methods and practices, which primarily focus on assessing the functional correctness of a deterministic software system with pre-defined test inputs and outputs, might not sufficiently evaluate the LLM. Additionally, given the characteristics of the LLM—interaction via natural language, ability to perform a variety of tasks, and continual learning—necessitates extra care and additional assessments when it comes to their evaluation. Therefore, a comprehensive assessment of LLMs is essential to harness its benefits successfully.

From an acquisition perspective, numerous LLMs are currently available to practitioners. In addition to addressing warfighter requirements, acquisition specialists and T&E professionals have to make sure that all acquired and deployed LLMs are effective, safe, and reliable. The deployment of LLMs in government settings raises significant concerns regarding operational safety, data privacy, and the potential for inadvertent exposure of sensitive information, to name a few. This paper aims to present a discussion on the current practices in the T&E of LLMs to better inform acquisition professionals when seeking to acquire these tools. The ideas are presented based on the findings from the survey of academic literature and industrial best practices for T&E of LLMs.

LLMs can complete many different complex tasks, which increases the difficulty and necessary variability in testing. Due to the versatility of LLMs, T&E activities generally involve running a range of evaluations on a range of tasks (e.g., question and answer, information retrieval, text classification, and summarization) to evaluate a range of characteristics (e.g., understanding, reasoning, generation, fairness, security, and toxicity).

While LLMs can range in complexity, this paper is focused on based models but is applicable regardless of the model's size or openness. Sometimes, LLMs are single-base models (e.g., BERT, GTP 4.0, etc.). However, frequently those based models in combination with other AI or systems are also considered LLMs. This is often seen in LLMs that have added "guardrails" that provide safety and security. Models can also vary in size and whether they are open, closed, or somewhere in between. Typical LLMs can range in size from millions to even trillions of parameters. The number of parameters can have a significant impact on an LLMs capabilities and quality. Open LLMs are those whose training data, code, architecture, and model weights are fully open to the public. In closed LLMs none of those are available to the public and may not be available to the deployer. There are also partially open LLMs where only some of that content is available.

Ensuring that an LLM can be a reliable and safe solution means it must be able to provide accurate results, robust to many different scenarios, and resilient to variable and potentially hostile inputs. We categorize the LLM acquisition scenarios along two key dimensions: 1) the information about the LLM that the acquisition team has access to, such as training data, code, architecture, and model weights, and 2) how often or how many times the team can carry out T&E activities to assess the quality of the LLM (test scheduling).

Different types of information (model artifacts) access at the time of acquisition:

- White box: LLM model is developed in-house (i.e., by the government) so T&E personnel have access to all the model's artifacts
- Grey box: An off-the-shelf pre-trained LLM, so T&E personnel do not have access to training information (data, hyperparameters, code, or model weights); however, they do have access to data and other artifacts used in fine-tuning the LLM.

- Black box: An off-the-shelf pre-trained LLM without modifications (no-fine-tuning) so there is no access to training information (data, hyperparameters, code, or model weights).

Different test schedules:

- Continuous testing: The ability to perform T&E activities throughout the LLM's life cycle
- Periodic testing: No testing access during development but the ability to evaluate during fine-tuning
- One-time testing: Testing is limited to evaluating the final model output and performance

Combining these two dimensions, we identify three use case scenarios:

- Use Case 1: In-house development—White-box and the ability to perform continuous testing.
- Use Case 2: Fine-tuning an off-the-shelf LLM—Grey-box, and periodic testing.
- Use Case 3: Off-the-shelf LLM (as-is)—Black-box and one-time acceptance testing.

The acquisition team's strategy for evaluating an LLM depends on how and what access they have to the LLM and its artifacts. Next, a detailed description of three use cases is presented, which will serve as practical examples to facilitate our discussion in the subsequent sections.

Use Case 1 White box, in-house development, continuous testing, software only

A department completed an in-house effort to develop a software application for processing free-text records about financial transactions. The software application's task is to identify named entities in a user-provided collection of records, extract relationships between entities, do entity resolution, and provide network graphs of the relationships. The LLM is a key component contributor to performing the named entity recognition and relationship extraction. Separate components of the software application perform the entity resolution and the network graphs. Additionally, the application has a user interface with a quality feedback mechanism. The contract for the software application includes creating a new LLM and will provide the department with the training data and the model weights.

Use Case 2 Grey box, off-the-shelf LLM that is fine-tuned in-house, periodic testing.

A department wants to have an application to help its staff complete internal documents that traditionally require a lot of manual labor. These documents contain fixed fields, short free-text, and long blocks of free-text. The fields will be completed by extracting information from user provided records and questions responses. To build the application, the department has found a quality LLM developed by another organization in its agency. The department does not have access to the weights or training data of the LLM. They will fine-tune the LLM using their own data and development team. The department has funded retraining of the LLM every 6 months.

Use Case 3 Black box, off-the-shelf LLM, one-time acceptance testing, LLM system within hardware.

An agency is purchasing small drones for searching natural disaster sites. The drones can be commanded by text messages from the operators. While the text can be manually typed, operators are more likely to send texts created through verbal transcription. The drones have an LLM that converts text messages into commands that they can implement. The agency has no information about the specific LLM model or the data used to develop it. Because this is a Consumer-Off-The-Shelf (COTS) system, the agency cannot alter the system or its internal LLM but will test the drones before committing to a large purchase.

The remainder of the paper is structured as follows: We first present the current T&E practices, including an overview of the steps in testing LLMs. This is followed by a discussion on establishing the scope of LLM evaluation by categorizing the LLM purposes as capabilities and properties as qualities to outline "what to test" in a test plan. We then discuss the limitations of current practices and finally present our concluding remarks and directions for future research.

Current T&E Practices

An Overview of Testing of LLMs

Next, we will provide an overview of the steps required for the T&E of an LLM. Testing an LLM typically follows the procedure shown in Figure 1. For Use Case 1, this activity starts at the end of in-house model development. For the other two use cases, this series of steps begins either when the LLM is obtained as an off-the-shelf model or after fine-tuning the LLM (applicable only to Use Case 2).

Figure 1. Overview—Testing LLM

Step 1: Installing Prerequisites: The first step is to install all the required software packages and dependencies. This is useful for handling sensitive information such as Application Programming Interface (API) keys and configuration settings. The next step is to download the necessary libraries, which will be used for various activities such as data processing, API interaction, and environment management tasks.

Step 2: Loading the LLM: The procedure to load the LLM will vary depending on the specific LLM. The two common distribution modes for LLMs are 1) host the LLM locally and 2) API-based access. For a locally hosted LLM, load it using the appropriate code. For example, an LLM developed in-house or an open-source pre-trained LLM like Llama2 that can be downloaded and executed locally. If the LLM is accessed via an API, establish the connection using an API key. For example, OpenAI's GPT3.5 Turbo can be accessed via an API key.

Step 3: Loading Test Dataset: When evaluating an LLM, the test dataset will be specific to the task the LLM is asked to perform and the desired evaluation methodology (described further under Step 5: Assessment/Evaluation). LLMs can perform many different tasks (Chang et al., 2024). Some common tasks are:

- Text Classification: assigning a label or class to a given text
- Sentiment Analysis: identifying the emotional category/state of the text
- Named Entity Recognition (NER): locating and classifying named entities mentioned in text
- Multiple Choice Question (MCQ): responding to a multiple-choice question with the correct answer
- Question and Answer (Q&A): responding to an open-ended question with an appropriate answer
- Text Completion: providing words to proceed with a sequence of text
- Information Retrieval (IR): identifying relevant information to a prompt
- Summarization: summarizing, reformulating, or condensing text meaningfully based on a prompt (Allahyari et al., 2017; Nguyen et al., 2024)

Furthermore, an LLM can be evaluated using either or both of the following types of datasets:

- Established test dataset: The tester can utilize established or published datasets from the AI community.
- Custom test dataset: The tester can create a specific dataset that assesses cases or scenarios tailored to their particular use case. This custom dataset can be hosted locally as a CSV or JSON file and used to evaluate the model's capability based on specific criteria.

Step 4: Prompting: Unlike traditional software systems, user interaction with an LLM is primarily with a text input called a "prompt." A prompt is typically natural language text but can include code or pseudo code. A prompt is a set of instructions that informs the LLM about the user's request. It comprises the input, desired LLM behavior, and any other instructions that users expect the LLM to follow while processing their request. Usually, a prompt consists of three main components:

- User role: User's query
- System role: Instructions on how the model should behave or respond
- Assistant role: Provides a method for giving examples of what a response should look like.

When testing an LLM, the prompt you use and the characteristics of the prompt will be specific to the task you are asking the LLM to perform. Thus, creating effective prompts is crucial for better engagement with the LLM.

Prompting strategies are techniques used to guide language models in generating desired responses. Three common strategies are (Schulhoff et al., 2024; Wei et al., 2022):

1. Zero-Shot Prompting: involves providing no prior examples to the model.
2. Few-Shot Prompting: involves providing a few examples to help the model understand the prompt/task.
3. Chain-of-thought (COT) Prompting: involves breaking down complex tasks into simpler steps to help the model understand the prompt/task.

Overall, prompt construction plays a vital role in testing LLMs. In other words, how the prompt is constructed affects the model behavior and, thereby, model evaluation. Therefore, creating effective prompts that combine the test scenario (user input) with other contextual information relevant to the LLM is essential.

In addition to the prompting strategies, parameters significantly influence the LLM's outcome. Key parameters include:

- Temperature: Controls the randomness of the generated output. A higher temperature value increases creativity in LLM outcomes by sampling from a wider range of possible tokens, while a lower value (i.e., closer to 0) produces consistent and predictable outcomes.
- Top-p (nucleus sampling): Limits the selection of words (tokens) whose cumulative probability reaches or exceeds the specified top-p value.
- Max tokens: Sets the maximum number of tokens that can be generated in a response. In other words, the max tokens parameter allows you to limit the length of the generated response.
- Frequency penalty: To minimize the likelihood of repetitive tokens by penalizing tokens based on how frequently they have already appeared in the output text.

Step 5: Assessment and Evaluation: The prompt, which consists of the user instructions and a test case, is provided as input to an LLM. Upon receiving the prompt, the LLM processes it with any contained instructions, and produces an outcome. Next, the LLM's output is recorded and analyzed. The metric by which the LLM is assessed depends on the task it was asked to perform. Some commonly used metrics (Hu et al., 2024) are:

- **Classification-based metrics:** accuracy, precision, recall, F1-score
- **Token-similarity metrics:** Recall-Oriented Understudy for Gisting Evaluation (ROUGE), Bilingual Evaluation Study (BLEU), Metric for Evaluation of Translation with Explicit Ordering (METEOR).
- **Embedding-similarity metrics**: Bidirectional Encoder Representations from Transformer Scores (BERTScore).

Note that the assessment is specific to the task (e.g., Named Entity Recognition), and the prompt must be designed according to the task that is currently being evaluated.

Benchmarks: Evaluating an LLM using a standard test data set provides insights into the LLM's abilities on a specific task compared to other models. However, LLMs are versatile and possess the ability to perform a variety of tasks with varying degrees of complexity. Thus, evaluating an LLM on a single test set will not be sufficient. Benchmarks are tools for exploring an LLM's strengths and weaknesses over a diverse range of tasks or functions.

A benchmark is a standardized framework for the holistic evaluation of an LLM. It consists of diverse task sets (e.g., NER, MCQ) to test an LLM on its various abilities, metrics for evaluating each ability, and a systematic methodology to assess different dimensions of an LLM's abilities. Furthermore, they enable objective comparison between different LLMs. For example, Massive Multitask Language Understanding (MMLU) is a widely used benchmark that evaluates LLMs across 57 subject areas across humanities, STEM, social sciences, and others (Hendrycks et al., 2021). Overall, a diverse collection of benchmarks provides a holistic understanding of an LLM's range of abilities, offering a more comprehensive assessment than any single test set could provide.

While evaluating using benchmarks delivers valuable holistic insights into the LLM's abilities, real-world deployments of LLMs necessitate targeted evaluations that align with specific use cases and operational conditions. This is in part because of the limitations of benchmarks. By being highly structured and constrained in their implementation, benchmarks offer results that are comparable across LLMs. This means that benchmarks typically do not incorporate the context of use for a specific LLM application. Additionally, LLM benchmarks often experience benchmark saturation, where the usefulness or integrity of the benchmark reduces overtime. As a result, benchmark testing usually needs to be combined with tests for specific use cases or real-world use context. The following section outlines the key dimensions—LLM capabilities (what it can do) and qualities (how well it does it)—what specific capabilities and qualities testers should prioritize based on their intended applications and operational needs.

Key Dimensions of LLM Evaluation: Capabilities and Quality Attributes

LLMs appear to have human-like abilities, which makes us want to use them like a human, e.g., relying on general intelligence to perform a variety of tasks across a variety of domains. However, like other AI, they are developed from training data and may not generalize outside the training distribution. This makes it very important to ensure that they are tested within the operational contexts and for the specific operational purposes they will be used for. When a single LLM is expected to be used very broadly, this creates an extremely large test universe.

This expectation of broad utility and human-level performance necessitates a thorough assessment of the LLM. Therefore, a comprehensive evaluation framework for testing LLMs must include two primary dimensions: (1) evaluation of fundamental capabilities, namely understanding, reasoning, and generation, and (2) ascertaining its quality attributes, such as reliability, performance, robustness, and privacy. The evaluation of fundamental capabilities is essential due to an LLM's core function as a language model that facilitates human-like interactions. As a software component, LLMs must meet expected quality standards.

Capabilities

Our structured approach to testing LLMs is based on three aspects of LLM input processing: understanding, reasoning, and generation. These were chosen because they are core to an LLM's functionality. Upon receiving input, an LLM is expected to 1) parse and understand the input, 2) reason based on that understanding, and 3) combine both the reasoning and understanding to generate an outcome. While we will separately discuss testing these three aspects, it is possible for them to sometimes overlap.

Understanding is the capability of LLM to successfully interpret text inputs. We will test if the LLM can successfully interpret the user inputs by looking at its ability to receive, parse, and comprehend natural language. Below, we discuss specific tasks that testers can use to evaluate LLM understanding.

- **Named Entity Recognition (NER):** NER is a widely used task in the natural language processing (NLP) community to evaluate the language model's ability to parse inputs and assign appropriate entity categories to each word from the input text. NER is considered one of the fundamental evaluation tasks in NLP. It helps determine whether the model can understand each word (also referred to as a token) from the input text and classify them into entity categories. There are standard entity categories found across most NER implementations (e.g., person, organization, location). However, specific applications may develop custom entity categories. Testers should identify any relevant common and custom entity categories. Evaluating the LLM on this task helps assess its token-level understanding capabilities; however, it does not evaluate the model's ability to understand the relationship between the tokens.
- **Text classification:** In text classification, the LLM assigns a label to an input text based on the overall theme of the input. The assigned labels are predefined. They frequently have just two categories (e.g., yes/no, pass/fail, etc.), but multiple categorical labels are possible. Additionally, text classification commonly assigns one label, but some applications may be designed to assign multiple labels. Unlike the NER task, which is limited to evaluating the LLM's understanding at an individual word level, text classification helps in ascertaining whether the LLM is able to understand the overall relationship between words in the input text.
- **Sentiment analysis:** Sentiment analysis is a computational method that assesses a text's tone or sentiment. Testers can use sentiment analysis to assess an LLM's ability to understand the nuanced relationship between the words in the input text and derive the overall sentiment or emotion of the input. Sentiment analysis can be performed at the sentence level, paragraph level, or document level. Outputs from sentiment analysis can be binary (i.e., positive/negative). However, outputs that are a probability or range of scores are better at assessing an LLM's ability to infer contextual nuances and meaning across various text lengths.
- **Natural Language Inference (NLI):** NLI is also known as textual entailment (TE) or Recognizing Textual Entailment (RTE). It is the task of determining the logical relationship between two short texts which are denoted as the premise and the hypothesis. A premise and a hypothesis are provided as inputs to the LLM which analyses the relationship

between them and assigns an appropriate label. In general, NLI identifies three types of relationships: 1) Entailment if the premise logically implies the hypothesis; 2) Contradiction, when the hypothesis contradicts the premise; and 3) Neutral, when the hypothesis can neither be logically deduced as true nor false based on the premise. In other words, it might be possible but is not 100% likely (not enough information to conclude). NLI represents a more advanced level of understanding, requiring the LLM to integrate token-level understanding and contextual understanding and then reach a conclusion.

Reasoning is the capability of LLMs to process information (from input text), draw inferences from the information, and derive conclusions based on the available information (Mondorf et al., 2024). Evaluating reasoning capabilities provides insights into LLMs' logical reasoning and analytical thinking abilities. Reasoning capability evaluations in LLMs are broadly categorized into core and integrated reasoning tasks, and the current T&E practices include various tasks within these categories. Common reasoning tasks include:

- **Logic:** This is the LLM's ability to logically derive valid conclusions. Based upon the objective, it can be divided into three subtasks (Mondorf et al., 2024):
    - **Deductive** reasoning tasks aim to assess if the LLM reaches a conclusion based on its valid premise or deriving cause-and-effect relationships.
    - **Inductive** reasoning tasks help evaluate the LLM's ability to identify patterns from the input and arrive at reasonable generalizations. In other words, given a specific set of examples, the task evaluates if the LLM is capable of deriving generalizable conclusions.
    - **Abductive** reasoning tasks test the LLM's ability to use given observations to formulate plausible explanations or possible hypotheses.
- **Mathematics:** The LLM's ability to perform mathematical tasks.
- **Multi-hop:** The LLM's integrated reasoning ability, assessing if the LLM can successfully make a series of logical steps or interferences to reach a conclusion.
- **Common sense:** This reasoning task assesses the LLM's capability to apply real-world knowledge, such as everyday situations and human-like interactions, including social norms and constraints, basic laws of physics (e.g., ice melts into water and objects fall), and others.

Generation is the capability of LLMs to produce/create coherent, contextually relevant model responses. An LLM's output from a text input can range in length and complexity. The generated responses can go from a single-word or syllable to long text summaries. The below evaluation practices aim to assess the different aspects of generation evaluation through various tasks such as translation, question answering, summarization, code generation, and text generation. Note that we limit our discussion to text based LLM and do not discuss multi-model models.

- **Translation:** The LLM's ability to generate translated text that guarantees the relevance and underlying meaning of the original text, grammatical accuracy, and contextually relevant translated text.
- **Question/Answer (QA):** The LLM's ability to generate relevant and accurate responses based on provided questions. These tasks are designed to test the model's ability to respond to either an open-ended question with an appropriate answer or respond to a multiple-choice question with the correct choice.

- **Summarization:** The LLM's ability to create short content capturing the key points and concepts in a larger input text. Two types of summarizations are extractive and abstractive reasoning. In extractive reasoning the LLM is evaluated on how well it extracts key excerpts from the larger text and combines the excerpts into a coherent output. For abstractive summarization, the LLM is assessed on its ability to create concise original text that captures the meaning of the input text. Overall, the summarization task represents an advanced level in testing generation capabilities.
- **Coding**: Expanding beyond traditional text generation, the coding tasks evaluate the capabilities of LLMs in writing software code. This task primarily evaluates the LLM's capability to generate functionally correct software code based on user requirements.

Table 1 presents a list of tasks and benchmarks for evaluating the three capabilities of LLM. Testers should evaluate LLMs across all three capabilities: understanding, reasoning, and generation.

Table 1. A List of Tasks/Benchmarks Used for Evaluating Capabilities

Capability	Task Type	Benchmarks	Comments	Relative Complexity
Understanding	Named Entity Recognition	CoNLL 2003	Evaluates basic word-level understanding and categorization abilities.	Low
	Sentiment Analysis	IMDb Yelp-2 Yelp-5	The ability to grasp emotional and contextual meaning	Moderate
	Text Classification	SuperGLUE	Ability to understand and categorize text	Moderate
	Natural Language Inference	SNLI	Tests complex logical relationships between statements	High
Reasoning	Inductive reasoning	bAbI-15 EntailmentBank	Evaluate the ability to make generalizations from the observed patterns.	Moderate
	Deductive reasoning	bAbI-15	Test the ability to reach valid conclusions from premises	Moderate
	Abductive reasoning	α-NLI	Assess the ability to form plausible explanations from observations	Moderate
	Mathematical reasoning	GSM8K MATH	Tests mathematical problem-solving skills	Moderate
	Multi-hop reasoning	StrategyQA HotPotQA	Tests the ability to logically connect and reason across multiple steps.	High

	Commonsense reasoning	CommonSenseQA OpenBookQA HellaSwag	Assess the application of real-world knowledge	High
Generation	Translation	WMT IWSLT	Evaluates the translating ability	Medium
	Summarization	XSum CNN/DailyMail	Tests the summarization ability	High
	Code Generation	HumanEval MBPP	Tests the ability to generate functionally correct software code	High

Understanding: The understanding capability of an LLM influences its ability to correctly interpret inputs, including inferring complex nuances of the input, which in turn guides the reasoning and generation activities of an LLM. In other words, weaker/limited understanding capability can significantly impact the LLM's overall performance by limiting its ability to grasp the context, missing interconnected relationships in the input text, or a total misunderstanding of the user's intent, which will result in incorrect outcomes.

- For Use Case 1, testers must evaluate if the LLM can identify and accurately classify entities like person names, organizations, and transactions within user-provided records.
- Evaluation activities for fine-tuned LLM in Use Case 2 must assess the LLM's capability to understand records and identify entities that need to be mapped to specific fields in the documents.
- For Use Case 3, evaluations must ascertain if the LLM understands operator text messages by correctly identifying intended drone commands like "make a left turn" and their associated parameters "after 50 ft."

Reasoning: Shortcomings or deficiencies in reasoning capabilities can significantly impact the LLM's performance and, consequently, its adoption in operational environments. For example, a deficiency or lack of satisfactory abductive reasoning abilities can make the LLM prone to hallucinations—the tendency to generate plausible but factually incorrect information (Toroghi et al., 2024).

- For Use Case 1, logical reasoning evaluations will help determine if the LLM can identify both direct and indirect relationships among different entities across different records.
- Similarly, for Use Case 2, evaluations should be performed to determine if the fine-tuned LLM can synthesize user responses across multiple questions and extract information to update a specific field. In other words, determine if the fine-tuned LLM can perform multi-hop reasoning and extract relevant information from user responses.
- For Use Case 3, testers should consider evaluating scenarios such as whether an LLM applies common reasoning while converting text messages to commands. For instance, they should test if the LLM can recognize and avoid generating physically infeasible commands, e.g., "fly through the debris."

Generation: Evaluating generation capabilities presents its own challenges, such as measuring creativity in the generated text and the inherently subjective nature of assessment in writing tasks.

- For Use Case 2, testers should verify that the information generated by the LLM to update the fixed field, short free-text, and long blocks of free-text is accurate and matches respective user records. Note that, given the goal of Use Case 2, the automatic completion of internal documents from user records, the evaluation of the fine-tuned LLM's ability to generate coherent, concise text for updating the fields of both short free-text and long blocks of free-text, are key evaluation priorities. While Use Cases 1 and 3 might involve minor generation, it is less critical than Use Case 2's core task. Hence, we limit our discussion to Use Case 2.

Quality Attributes

Below, we describe some common quality attributes. We include some illustrative examples using the use cases described above. While quality attributes assessment is essential for all three use cases, due to space limitations we limit our discussion to one or two use cases per quality attributed.

Reliability: Reliability is "the ability of a system or component to perform its required functions under stated conditions for a specified period of time" (ISO/IEC/IEEE, 2017). Reliability assessments of an LLM evaluate its ability to produce consistent, coherent outputs under normal or expected operational conditions. The main goal is to evaluate an LLM's behavioral consistency. It includes a variety of assessments, such as testing the LLM for consistent behavior across variations in the input text and different contextual settings, factual consistency across outputs for the same or similar prompts (hallucination detection), the LLM's ability to quantify and communicate its confidence in its outcomes (uncertainty quantification) and assessing the accuracy of confidence estimates to actual performance (calibration; Sun et al., 2024; Walsh et al., 2024; Zhuang et al., 2023). A lack of comprehensive reliability assessment significantly increases the risks in operationalizing LLMs, as unreliable models can lead to serious and potentially catastrophic outcomes.

For Use Case 2: Evaluate the fine-tuned LLM's consistency; the reliability assessments should evaluate whether the fine-tuned LLM consistently extracts the same information from the user-provided records and question responses. For example, provide the employee performance report document (input) multiple times and check if the LLM consistently extracts the employee's name and the manager's feedback.

Performance: Assessing an LLM's performance is a multifaceted activity that involves evaluating both the quality of the model's outcome and its efficiency in producing the outcome. Quality assessment includes evaluating the output's coherence, ensuring logical flow and contextual consistency, determining its relevance to the given task, ensuring the outcome is factually correct, and evaluating the logical soundness of the LLM's reasoning process (Huang et al., 2024; Zhuang et al., 2023). Efficiency evaluation measures the LLM's computational performance through latency, inference speed, and throughput, which are critical for adapting an LLM across different operational environments.

Use Case 3 will be used in real-time operational conditions. Therefore, testers should evaluate LLM latency and assess if the latency level is sufficient to support operational needs.

Maintainability: Unlike traditional software systems, updating or modifying an LLM is not limited to structural changes to the software code. Given an LLM is a data-intensive system, modifications range from retraining the LLM with a revised dataset or hyperparameters to fine-tuning an LLM with a domain-specific dataset. Furthermore, in some cases, LLMs are designed to learn from their operational environment (continual learning). Thus, it is subject to frequent adjustments (or updates) upon deployment. While this is a desired behavior, ensuring that an LLM continues to work as expected is critical. Maintainability assessment must account for the characteristics of LLMs and focus on evaluating the model's ability to incorporate updates as well as their ability to adapt to different operational environments without comprising performance.

Use Case 2: Given the planned semi-annual retraining cycle, evaluations must be performed to ensure that periodic retraining does not adversely impact the LLM's performance. For instance, after each retraining with recent documents and potentially new data sources, testers should compare the LLM information extraction accuracy between previous and newly introduced data sources.

Scalability assesses the LLM's ability to deliver satisfactory performance under varying operational conditions, including fluctuating demands in user queries, input text length, and computational resource consumption. In other words, scalability evaluates the LLM's ability to handle increased demand or workloads (serve a significant number of concurrent user requests, handle larger inputs, and operate in diverse environments) without significantly comprising its performance (i.e., output quality).

Scalability assessments of an LLM help determine operational conditions suitable for optimal model performance (resource requirements), identify potential bottlenecks that may hinder the model's performance, and thereby determine the model's viability for deployment. To this end, testing scenarios will be designed to systematically evaluate LLM's performance across various operational conditions. Key testing scenarios include:

- **Size of input:** Understanding the LLM's ability to handle various types of inputs, including the complexity of input and length of the inputs. For example, can an LLM handle longer, lengthier inputs/documents (Context window limitations)? Does handling longer inputs result in a memory crash? Does handling a substantial number of longer inputs (requests) impact the LLM's processing or inferencing time? If there is a delay in inferencing time, is it within a reasonable time? Measuring the LLM's response quality with increasing complexity in prompts.
- **Number of users:** Does the increase in the number of users impact the model's performance (e.g., increase in model inference time)?
- **Frequency of input:** Can the LLM handle a higher volume of inputs without significant performance degradation?
- **Operating environments:** Can the LLM be deployed on different hardware configurations? Can it work with limited hardware resources? Can it handle new data types?
- **For Use Case 1:** Focus on whether the LLM can identify entities and extract relationships within a reasonable time frame, specifically when dealing with increasing workloads (>=1M records).
- **For Use Case 2:** Evaluate the fine-tuned LLM for varying lengths of short free-text (e.g., between 50 and 450 words) and long blocks free-text (e.g., 1000, 2000 words) and see if it impacts the LLM's extraction ability.

Robustness: "The degree to which a system or component can function correctly in the presence of invalid inputs or stressful environmental conditions" (ISO/IEC/IEEE, 2017). In the context of LLMs, robustness is defined as the ability to produce a consistent performance across different operating conditions, such as previously unseen scenarios, handling noisy input data, out-of-distribution inputs, variations or perturbations to input prompts (prompt sensitivity), and resistance to adversarial attacks (Sun et al., 2024). Robustness assessment must evaluate an LLM's resilience across various input scenarios, including its ability to handle out-of-distribution input values, resistance to prompt injection attacks (a type of adversarial attack targeted at manipulating the LLM's behavior), its performance under perturbed input values, and its ability to generate relevant, accurate outputs in the presence of misleading or irrelevant information. Failure to perform a comprehensive robustness assessment opens up the LLM for potential operational risks, such as increased vulnerability to prompt injection attacks, susceptibility to generating inconsistent model outcomes when dealing with noisy input values, and potential performance degradation when deployed in rapidly evolving operational environments (lack of generalizability).

For Use Case 3: Examine the LLM's behavior in handling informal or abbreviated language and uncommon jargon that the model might not have been exposed to at the time of its training. Furthermore, tests must be conducted to assess the LLM's sensitivity to text perturbations. For example, testers can introduce common or likely misspellings or additional white spaces to text messages and determine if the LLM continues to perform as expected or results in a misinterpretation.

Privacy: The ability to safeguard sensitive information, including training data, personally identifiable information (PII), and confidential information received through training or user interactions. LLMs are trained on large datasets, making them susceptible to privacy attacks. Furthermore, the interactive nature of LLMs, combined with its tendency to align itself with users' requests, significantly increases privacy risks. For example, malicious actors could employ sophisticated prompt injection techniques to trick the LLM into revealing sensitive information.

Privacy evaluation in LLMs focuses on ascertaining its ability to protect against unauthorized access, accidental disclosure of training data or user information, and resilience to various data extraction attacks (Sun et al., 2024). Failure to perform adequate privacy evaluation poses significant risks, as it increases the likelihood of exposing training data, revealing sensitive information, and confidential user interactions.

For Use Case 2: Evaluate the fine-tuned LLM for privacy guarantees. They should also perform privacy assessments to ensure that the LLM does not inadvertently expose sensitive information from the document.

Security: At a high-level, security refers to protecting a system from threats and risks that may lead to harm. When assessing the security for an LLM (which is likely a component of a larger system), testers should focus on assessing safeguards of the model and its related artifacts. This includes assessing protections for its training data and model weights, defenses against unauthorized entities, processes for detecting and mitigating adversarial threats and malicious manipulation of the LLM's behavior, and other processes for ensuring the LLM's integrity.

A comprehensive security assessment of LLMs includes evaluating its resistance to prompt injection attacks, its ability to defend against evasion attacks, and testing for vulnerabilities in access control. This includes assessing how resilient the model is to unauthorized access, potential modifications, or tampering with model weights and protecting against model inversion attacks. The overall goal is to assess and implement robust safeguards that prevent malicious manipulations of the model for generating harmful outcomes, thereby ensuring the integrity of the LLM across various operational contexts.

For Use Case 1: Evaluate whether it can detect malicious user requests and does not disclose privileged information (i.e., prevent unauthorized access).

Explainability: LLMs function as black boxes, with their internal decision-making processes remaining opaque to the users (Cambria et al., 2024). LLMs that are explainable facilitate user trust in outputs and effective debugging processes. Explainability in LLMs aims to provide users with insights into the LLM's reasoning as to why it produced a particular outcome, thereby facilitating a better understanding of its behavior. Explainability assessments evaluate an LLM's inherent capacity to generate human-comprehensible explanations for its outcomes (Zhao et al., 2024). For example, they may assess if the LLM provides rational support for its response.

For Use Case 2: Have the LLM provide step-by-step explanations for filling the fixed fields vs. short free-text vs. long blocks of free-text. This assessment will help stakeholders understand the LLM's information extraction and document completion process.

Fairness: LLMs are data-intensive systems that inherently reflect the distribution of the data they are trained with (Chandrasekaran et al., 2024). In other words, the behavior of the LLM is, to a large extent, a representation of its training data. Due to the practical limitations in comprehensively capturing the operational universe within the training dataset, an LLM may inadvertently reflect the inherent biases in the training data and thus exhibit discriminatory behaviors. Fairness in LLMs refers to the model's ability to generate outcomes without preference or discrimination across protected groups, ensuring no demographic is disadvantaged or misrepresented (Chu et al., 2024; Li et al.; 2023, Schwartz et al., 2022). Fairness evaluation is essential to guarantee that the LLM exhibits non-discriminatory behavior. This activity spans the LLM's life cycle, including the data collection, training, and fine-tuning phases. Unlike assessments of other quality attributes, where a single test instance (i.e., occurs only once) may suffice to identify/uncover the underlying issue, fairness evaluation may require multiple test instances (i.e., more than one occurrence) to establish patterns of discriminatory behavior (Weidinger et al., 2022). Thus, necessitating comparatively increased testing efforts.

For Use Case 3: Fairness evaluations ensure that the off-the-shelf LLM treats different dialects and linguistic styles equally when converting the operator's text messages to commands. Is the LLM prone to misinterpret certain linguistic styles while converting text messages into commands?

Safety: Evaluates the LLM's ability to avoid generating harmful, toxic, unethical, deceptive, unlawful, or otherwise undesirable content or behavior during its intended use, thereby maintaining a safe operational environment (Weidinger et al., 2022; Zhang et al., 2023). It involves systematically assessing an LLM's behavior across diverse scenarios to achieve this goal. Furthermore, it aims to validate an LLM's safety guardrails in preventing the model from generating unsafe outputs either intentionally (tricked by a malicious actor) or unintentionally (non-malicious intent yet can cause harm).

For Use Case 3: Most significant for Use Case 3 given its operational environment and application (but crucial for all use cases). Safety evaluations must be performed for this use case to guarantee that the LLM (in the drone) identifies and rejects potentially dangerous text messages or asks for further clarification (from the user) before converting to a command in case the input text is ambiguous or borderline risky.

Adaptability: "The degree to which a product or system can effectively and efficiently be adapted for different or evolving hardware, software or other operational or usage environments" (ISO/IEC/IEEE, 2017). Given the nature and characteristics of LLMs, testers can measure an LLM's adaptability by assessing its ability to adjust to new operational conditions, including new tasks and domains. They can also assess its ability to perform inferencing in resource-constrained hardware environments. Furthermore, evaluations must ascertain an LLM's ability to improve its performance through continual learning (feedback from the operational environment). Inadequate adaptability evaluation creates significant bottlenecks in operationalizing LLMs across diverse environments. For example, limited or lack of adaptability can make an LLM obsolete too soon. Since LLMs require substantial resources for training and deployment, retraining or replacing an obsolete model becomes expensive and time-consuming.

Similarly, poor adaptability to diverse operational environments significantly restricts an LLM's utility and its cross-environment applicability.

> **For Use Case 1:** Ascertain its ability to adapt to new domains or usage environments. For instance, evaluations must determine if the LLM can accurately identify and extract relationships from previously unseen entity types.
>
> **For Use Case 3:** Evaluate the LLM's inferencing performance across various resource-constrained environments. Specifically, they should assess whether the LLM maintains acceptable performance when deployed on different hardware architectures. Additionally, testers must assess whether there is a significant drop in performance between the original LLM and its optimized versions, such as a quantized LLM.

Framework Boundaries

The evaluation of LLM across capabilities and qualities reveals an interconnectedness where there is an inherent lack of rigid boundaries between assessment areas. In most cases, evaluating an LLM on a task intended for a particular capability can potentially assess other capabilities. For example, NLI evaluation assesses not only understanding but also the LLM's reasoning abilities. Likewise, evaluating an LLM for hallucinations goes beyond strictly assessing reliability, as it also reflects on the LLM's performance in generating factually correct outcomes. Although we present the framework by grouping tasks under the category they primarily assess, testers should be mindful that, in most cases, an LLM evaluation can typically provide insights into multiple capabilities, quality attributes, or a combination of these. Moreover, test plan design must be guided by the operational conditions, prioritizing specific capabilities and quality attributes based on operational requirements.

Challenges and Limitations in the Current T&E Practices

The current T&E practices for LLMs, while providing a baseline for evaluation, suffer from key limitations. Firstly, there is a disconnect between benchmark performance and real-world utility. Open-source benchmarks barely reflect the full spectrum of operations an LLM will encounter in operational scenarios. Moreover, most benchmarks remain static over time and lack domain specificity. Adding to this, the controlled nature of a test environment fails to account for variability in operational environments. Consequently, a strong benchmark performance may not necessarily translate to success in an operational environment. Secondly, the use of aggregate metrics (accuracy, F1 score) provides insights into LLM performance. However, they fail to provide a granular understanding of the LLM's behavior, thus limiting the ability to gain insights into specific strengths, weaknesses, and potential failure points in LLMs. Third, LLMs' non-deterministic and opaque nature presents unique challenges in failure analysis and debugging activities. Existing T&E approaches developed for traditional software systems with understandable decision logic and deterministic behaviors are often ineffective for LLMs, limiting the ability to systematically detect and address failure modes. Finally, a lack of standardized techniques to measure test adequacy, potentially leading to inadequate test design and incomplete evaluation.

Conclusion and Future Directions

This paper presents an overview of the current T&E practices for evaluating LLMs based on a survey of academic literature. We outline the key steps in testing LLMs and discuss how to establish an evaluation scope by categorizing LLM capabilities and properties, illustrated with three acquisition scenarios. Our findings indicate that while the T&E of LLMs is nascent and rapidly evolving, significant challenges remain. Current practices provide a foundation for evaluation but require substantial improvements to address the challenges in evaluating the multi-faceted nature of LLMs. For instance, public benchmarks offer a starting point for evaluation; however, their utility is limited as they cannot be generalized to all possible operational scenarios. Our analysis identifies several key areas for future research. First, developing new T&E approaches to comprehensively evaluate LLMs in specific operational contexts. Second, standardized approaches for measuring test adequacy should be established. Finally, we observe a significant imbalance in the T&E of an LLM across its life cycle. While a significant amount of work is reported for model-level T&E, there remains a significant gap in research regarding the system level and post-deployment (operational) evaluation.

Acknowledgment

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense (DoD) through the Office of the Under Secretary of Defense for Acquisition and Sustainment and the Office of the Under Secretary of Defense for Research and Engineering under Contract HQ003424D0023. The Acquisition Innovation Research Center is a multi-university partnership led and managed by the Stevens Institute of Technology through the Systems Engineering Research Center—a federally funded University Affiliated Research Center. Any views, opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Government (including the DoD and any government personnel).

References

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). *Text summarization techniques: A brief survey*. arXiv preprint arXiv:1707.02268.

Cambria, E., Malandri, L., Mercorio, F., Nobani, N., & Seveso, A. (2024). *Xai meets LLMs: A survey of the relation between explainable ai and large language models*. arXiv preprint arXiv:2407.15248.

Chandrasekaran, J., Cody, T., McCarthy, N., Lanus, E., Freeman, L., & Alexander, K. (2024). Testing machine learning: Best practices for the life cycle. *Naval Engineers Journal 136*(1–2), 249–263.

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., ... & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology 15*(3), 1–45.

Chu, Z., Wang, Z., & Zhang, W. (2024). Fairness in large language models: A taxonomic survey. *ACM SIGKDD Explorations Newsletter 26*(1), 34–48.

Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2021). Measuring massive multitask language understanding. In *International Conference on Learning Representations*.

Hu, T., & Zhou, X. H. (2024). *Unveiling LLM evaluation focused on metrics: Challenges and solutions*. arXiv preprint arXiv:2404.09135.

Huang, X., Ruan, W., Huang, W., Jin, G., Dong, Y., Wu, C., ... & Mustafa, M. A. (2024). A survey of safety and trustworthiness of large language models through the lens of verification and validation. *Artificial Intelligence Review*, *57*(7), 175.

ISO/IEC/IEEE**.** (2017). *ISO/IEC/IEEE international standard—systems and software engineering—vocabulary* (ISO/IEC/IEEE 24765:2017) [Standard]. ISO/IEC/IEEE. Retrieved from [IEEE]

Li, Y., Du, M., Song, R., Wang, X., & Wang, Y. (2023). *A survey on fairness in large language models*. arXiv preprint arXiv:2308.10149.

Mondorf, P., & Plank, B. (2024). *Beyond accuracy: Evaluating the reasoning behavior of large language models—a survey*. arXiv preprint arXiv:2404.01869.

Nguyen, H., Chen, H., Pobbathi, L., & Ding, J. (2024). *A comparative study of quality evaluation methods for text summarization*. arXiv preprint arXiv:2407.00747.

Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, C., ... & Resnik, P. (2024). *The prompt report: A systematic survey of prompting techniques*. arXiv preprint arXiv:2406.06608.

Schwartz, R., Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence* (Vol. 3, p. 00). US Department of Commerce, National Institute of Standards and Technology.

Sun, L., Huang, Y., Wang, H., Wu, S., Zhang, Q., Gao, C., ... & Zhao, Y. (2024). *Trustllm: Trustworthiness in large language models*. arXiv preprint arXiv:2401.05561, 3.

Toroghi, A., Guo, W., Pesaranghader, A., & Sanner, S. (2024, November). Verifiable, debuggable, and repairable commonsense logical Reasoning via LLM-based theory resolution. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 6634–6652).

Walsh, M., Schulker, D., Lau, S-H. (2024). Beyond capable: Accuracy, calibration, and robustness in large language models. *SEI Blog*. https://insights.sei.cmu.edu/blog/beyond-capable-accuracy-calibration-and-robustness-in-large-language-models/

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems 35*, 24824–24837.

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. S., Mellor, J., ... & Gabriel, I. (2022, June). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214–229).

Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., ... & Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, *15*(2), 1–38.

Zhuang, Z., Chen, Q., Ma, L., Li, M., Han, Y., Qian, Y., ... & Liu, T. (2023). *Through the lens of core competency: Survey on evaluation of large language models*. arXiv preprint arXiv:2308.07902.

Knowledge Based Metrics for Test and Design

Dr. Craig Arndt—has extensive experience as a senior executive and technology leader in research, education, engineering and defense, homeland security and intelligence technologies, with extensive experience as an innovative leader in industry, academia, and government. He is often called upon as an industry expert to chair major program reviews, lead new programs, and research efforts. Dr. Arndt currently serves as a principal research engineer on the research faculty of the Georgia Tech Research Institute (GTRI) in the System Engineering Research division of the Electronic Systems Lab. Prior to coming to GTRI, Dr. Arndt served as Professor and Chair of the engineering department at the Defense Acquisition University and as technical director of the Homeland security FFRDC at the MITRE Corporation. Dr. Arndt holds engineering degrees in electrical engineering, systems engineering, and human factors engineering and a Master of Arts in strategic studies from the U.S. Naval War College.

Dr. Valerie Sitterle—is a Principal Research Engineer and Chief Scientist for the Systems Engineering Research Division in the Electronic Systems Laboratory at the Georgia Tech Research Institute (GTRI), where her technical work synthesizes systems science, complex systems, and defense analysis. She is an expert with over 20 years of experience in engineering science, integrating engineering, natural, and physical sciences leading to the design and analysis of systems. Two primary areas of her work revolve around the design and assessment of defense systems alongside concepts of operation in theater environments, and the synthesis of architectures with decision methods, processes, and tools to meet the next generation of systems engineering challenges. She earned a BME and MS in mechanical engineering from Auburn University, an MS in engineering science from the University of Florida, and a PhD in mechanical engineering from the Georgia Institute of Technology.

Jeremy Doerr—has over 15 years of experience in software and systems engineering across several domains and is currently an MBSE researcher at Georgia Tech Research Institute (GTRI). He has a BS in software engineering and a master's in systems engineering from Georgia Tech. At GTRI, Jeremy's work ranges from basic research into technical language semantics to applied research, guiding DoD sponsors to adopt MBSE and integrated digital engineering approaches. Prior to GTRI, he successfully certified several flight deck systems using new and novel touchscreen technology while working for Gulfstream on the G500 and G600 platforms. Before Gulfstream, Jeremy worked in logistics for Walgreens, performing database and automation work in operations, as well as conducting timing optimizations studies and Lean Six Sigma projects on the industrial engineering side.

Abstract

The task of developing the best military equipment in the world has long fallen on the U.S. Department of Defense and the military industrial base that supports them. The United States made the decision years ago (and have succeeded in going to war with the best equipment since the second half of WWII (1943)) that their military would have the best equipment in world. As the 21st century continues to unfold, this commitment is becoming ever more difficult and more costly, and hard to execute in a timely manner.

Over the past few years, the leadership of the DoD acquisition community have listed the acceleration of development testing and fielding systems as their top priority. To try to make this happen, the DoD is implementing digital transformation. Another major part of accelerating the acquisition process has been a movement to integrate the design and test functions of the acquisition process. This includes moving test earlier in the development process.

When looking at the test and acquisition process, it is important to understand what the goal of test is in the development process. Traditionally, the goal of test has been to validate that a design will meet specific requirements created for the system. This traditional goal, however, is becoming less relevant, and the role of test as part of the development process is consuming much more test resources. So, what is the goal of test? If the role of test is to help ensure that we are developing the best product for our customer, then we might think of test's role being to increase knowledge about the future performance of a system still in design while there is time to improve the design. At a practical level this means two things. First, that testing should be designed specifically to support decision-making; the development of the Integrated Decision

Support key (IDSK) was intended to support this goal. Second, that we need to integrate all activities that provide additional knowledge about the future performance of the system together in meaningful ways to support decision-making.

In order to integrate and measure the amount of knowledge needed to make specific decisions (about things like requirements, risk, system design, and test resource allocation), we need to be able to measure the amount of knowledge needed for decisions and the amount of knowledge that we expect to generate in a given activity (including design, test, or history).

In this paper, we will demonstrate the development of a mathematical based knowledge metric and how it can be applied to specific DoD acquisition and test decision-making. The paper will document the development of the decision add and use it in practical programmatic decisions.

Introduction

Digital Transformation

In June 2018, the Department of Defense established its expectations for digital transformation in The DoD Digital Engineering Strategy. The strategy outlines five goals aimed at establishing a digital engineering environment for more rapid and effective development and fielding of weapon systems. The goals include the use of models to inform decision-making, establishing an infrastructure to enable the digital engineering methods, and transforming the workforce to adopt digital engineering methods over the acquisition life cycle.

Figure 1 was developed by the DoD to help communicate the different elements of the transformation effort. The development and use of standardized models is critical to the success of the transformation and the resulting advantages of digital engineering to the operations of all aspects of the department.

The DoD followed up this strategy with the release of formal guidance via DoD Instruction 5000.97, which ensures that the director of Operational Test and Evaluation (DOT&E) will utilize digital engineering methods to achieve their test objectives for operational assessment and Live Fire Testing. Also in 2023, DOT&E released their DOT&E Strategy Implementation Plan (I Plan), which includes objectives and key actions to develop digital, or model based Test and Evaluation Master Plans (TEMP) and Integrated Decision Support Keys (IDSK). As recently as December 2024, the department has released an update to DoD Instruction 5000.89 and five DoD manuals further refining the description and use of digital methods for the entire DoD test community (DoDM 5000.96, DoDM 5000.99, DoDM 5000.100, DoDM 5000.101, and DoDM 5000.102).

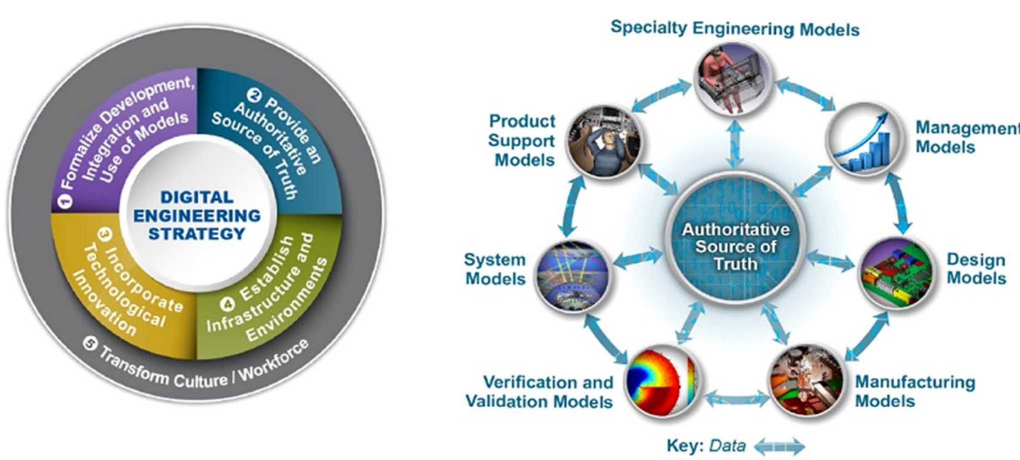


Figure 1. DoD Digital Transformation

Industry Trends / Knowledge Engineering

Knowledge engineering is a field that has grown and evolved significantly over the past few years and is now in many industries. Knowledge engineering is used to manage both knowledge in systems and development processes, and also knowledge created in manufacturing and use of systems to improve the performance and quality of a wide range of products. There are many applications of knowledge engineering in DoD applications. Some of these applications include the management and development of AI systems, and the management of data in operations, and combat systems. In industry and in the DoD, the need for and the management of knowledge is a becoming a critical concern in the development and use of systems.

Background

The Test and Evaluation Master Plan (TEMP) is one of the core artifacts in the DoD acquisition process (DoD 5000.01). However, the TEMP, and the test process as a whole, need to be understood as part of the larger research, engineering, development and acquisition process. Figure 2 is an illustration of the larger process needed to understand test and its relationship to requirements and the use of the system (mission data).

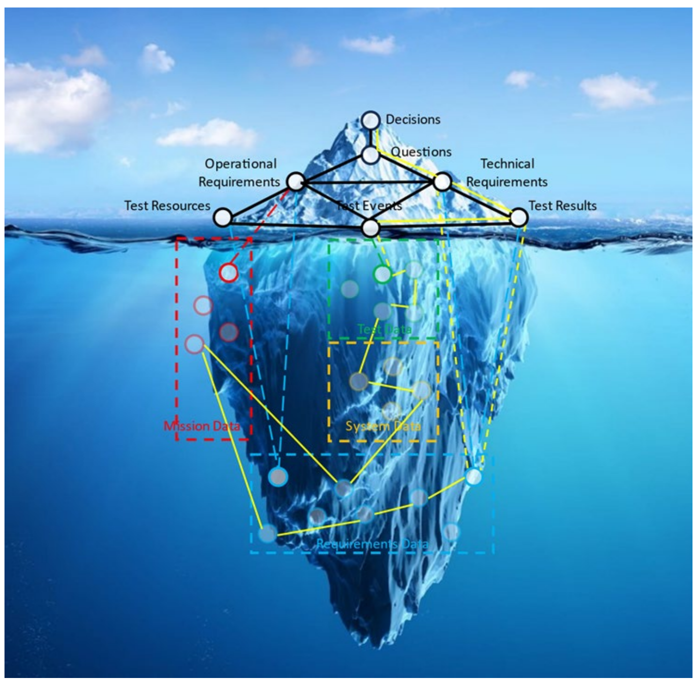


Figure 2 Development Iceberg

Prioritization of Speed in Defense Engineering

When we look at the priorities that the DoD has for process improvement to meet the significant challenges of the great power competition, it is clear that at the top of the list is the acceleration of the acquisition and development process. “As Undersecretary of Defense for Acquisition and Sustainment Dr. Bill LaPlante has emphasized, with some capabilities the department needs to be able to field several software revisions a day” (Shaffer & Whitley, 2024).

Defining the Requirement

Acquisition reform has taken many forms over the years. The defense acquisition system has been developed with many specific goals and provisions built into it. At this point, the acquisition system has a good deal of flexibility due to the adaptive acquisition pathways.

However, in the ever-changing world of the 21st century acquisition, there is a priority for the rapid development of high-performance systems. High performance systems, however, create significant risk. Moving forward, this creates a growing need to better recognize high risk in technical development, and to accelerate the development of system when possible.

Historical Issues

In order to better understand the history of acquisition risk from the perspective of, we conducted a detailed requirements analysis. The primary sources for this analysis were Government Accounting Office reports and scholarly papers on the acquisition process. Congress and DoD leadership have long used the same source to formulate acquisition policy. Table 1 summarizes issues in nine major areas of defense acquisition.

Table 1

Area	Issues	Knowledge Gap	Risk	Source
1. Capability Need	Business case development	Is there sufficient detail in the business case allowing for clearly defined requirements?	Insufficiently developed business cases leads	GAO-23-106059, GAO-21-511T
1. Capability Need	Key stakeholders' project and technology knowledge to make appropriate decisions	Does the key stakeholders, have sufficient knowledge, training	Insufficient experience, training	Defense ARJ, October 2012, Vol 19 No, 4422-443, GAO-20-439, GAO-23-106059, GAO-24-106831
Alt. Multiple areas	Incorrect inflation assumptions	Does the AoA include current approved inflation assumptions?	Not incorporating approved inflation assumptions leades to cost over/underestimation.	GAO-24-106831
2. Decisions	Programs outside acquisition pathways	Are there any programs within the Service/Department which will impact this capability?	Limited oversight of non-AAF pathway projects	GAO-24-106831
2. Decisions	Production Decisions out of sync with testing	Has the program conducted prototype testing prior to making a production decision?	Testing the prototype after making production decisions	GAO-24-106831
4. Acquisition Strategy	Acquisition pathway flexibility	Are the requirements to switch between acquisition pathways acknowledged and deliberately planned for?	Allowing contracts which plan to use multiple acquisition pathways without a deliberate plan to address known pathway	GAO-24-106831
4. Acquisition Strategy	Official cost estimates as programs transition between pathways	Are the program's official costs developed and published prior to transitioning to a new pathway?	Insufficient cost development limits informed investment decision-making by perpetuating the sunk-cost fallacy.	GAO-24-106831
5. Requirements	Cyber-security/cyber-physical interconnectivity	Are the cyber requirements for the capability full developed?	Not identifying all cyber requirements leaves the capability vulnerable to non-kinetic/EW attacks	GAO-24-106831, DoDI 5000.90

5. Requirements	All or nothing approach to requirements development	Is the program using or facilitating the use of iterative requirements development?	A monolithic approach to requirements development limits adaptability as technology matures.	GAO-24-106831
6. Source Selection	Single contract for total program	Does the source selection include modular contracting terms?	Single, large-scale contracts limit incremental capability development	GAO-24-106831
6. Source Selection	Supplier/Defense Industrial base disruptions	How stable are the logistics pipelines for suppliers and the Defense Industrial Base?	Logistical disruptions increase production timelines and overall costs.	GAO-21-511T, GAO-23-106059, GAO-24-106831
6. Source Selection	Developed a software factory	Does the Service/Department have a Software Factory to serve as Software SMEs throughout a project's lifespan?	Software factories are designed to speed up software development and acquisition by providing consistent user feedback, secure DevSecOps	GAO-23-106059, GAO-24-106831
6. Source Selection	Diminishing manufacturing sources	Does the program rely on at-risk parts?	Reduced supplier options	GAO-20-439, GAO-24-106831
6. Source Selection	Cost-reimbursement contracts on major development items	Does the program use a cost-reimbursement contract or some other type like a fixed-price incentive contract?	Cost-reimbursement contracts introduce substantial funding risks	GAO-18-238sp
7. Design	Technology maturation/readiness level	How mature is the technology for the various components?	Overestimating technology maturity leads to extended timelines and increased costs	GAO-21-511T, GAO-23-106059, GAO-24-106831
7. Design	Key decision points are not clearly defined nor are the requirements for those decisions addressed	Is there sufficient detail to the information requirements for the project to transition between acquisition phases?	Not adequately addressing decision point information requirements allows for vulnerabilities and deficiencies	Defense ARJ, October 2012, Vol 19 No, 4422-443, GAO-20-439
7. Design	Inconsistent cost data	How consistent is the data reporting?	Inconsistent cost reporting data limits oversight and potential increases risks	GAO-24-106831
7. Design	Limited use of digital engineering	Does the program use digital engineering methods to increase efficiency throughout the project's lifespan	Relying solely on static models to measure impacts of design changes limits efficiency	GAO-24-106831
8. Testing	Testing procedures for cyber/cyber-physical vulnerabilities	Does the testing plan include early and frequent testing for potential vulnerabilities?	By not testing all known or potential cyber threats, the capability becomes	GAO-24-106831, DoDI 5000.90
8. Testing	Software testing limited to "testers," not "users"	Has the program provided incremental deliveries to users for feedback?	By limiting software testing to a team of "testers" as opposed to end-state "users," programs can experience substantial risks to costs	GAO-20-439

9. Fielding	Cyber deficiencies not corrected	Were there any late-stage cyber vulnerabilities not addressed due to funding or available timeline?	Not addressing all cyber vulnerabilities prior to fielding increases risk	DoD CyberSecurity Test and Evaluation Handbook, DoDI 5000.90
9. Fielding	Production lines not achieving statistical process control	Has the production process achieved statistical process control prior to full rate production?	Achieving process control prior to full rate production limits delays in fielding and downstream user reliability concerns	GAO-20-439
9. Fielding	Accepting serious deficiencies identified during testing	Is the program accepting equipment with serious deficiencies identified during testing as defined by the respective Service?	Accepting equipment for fielding with uncorrected serious deficiencies increases risk	GAO-18-238sp
8. Testing	Production representative prototype not tested in its intended environment	Has the program completed operational environment testing with a production level prototype?		GAO 23-106059

This requirement analysis reinforces our understanding of future needs of the acquisition community.

Role of Test in the Department of Defense

Traditional test and evaluation in the DoD has been focused on two aspects of test process. Specifically, operational test, which looks at whether new system can perform functions it was intended to perform in its intended environment, and development test, which is design to add in the development of the final system design. Adding to the traditional functions of test, we need to look at test as a means to forecast the future performance of a system in the field and to assess the health and risk of the development process.

Availability of Knowledge

As we look at of knowledge of the performance of the new system of interest, the function of testing takes on a different perspective. As we have discussed, test has traditionally been about the verification of requirements. In the current digital world of system development, we can also look at test as a knowledge source contributing to our ability to forecast future performance of the system. In order to improve the ability of decision-makers to evaluate risk in system development, it is critical that we use all the available knowledge about the system as it is developed to make these decisions.

In addition to traditional developmental and operational test data as a source for knowledge about the performance of a system, a great develop of knowledge can be gain be 1. Data about legacy systems that use the same or similar subsystems, 2. design of the system itself, and 3. Modeling and simulation of the system. For the purposes of this work therefore are using five specific classes of knowledge sources for future performance of the system under development. 1. Legacy systems data, 2. Design data, 3. Modeling and Simulation, 4. Developmental test data, and 5 operational test data.

Specific Needs of Decision-Makers

As we demonstrated in our analysis of the GAO reports, the acquisition community has a number of different needs including technology development. Historical, it is critical in development to accurately asses the amount of risk there is in the baseline plan for development test and fielding of the system. By more effectively using all of the available data from all available knowledge sources, we can better asses the dynamic risk in a specific program and program development approach. Note that the test program is part of the program development plan (baseline). When we look at the specific needs of decision-makers, it is instructive to look at critical use cases. In the case of acquisition leaders, there are two important use cases that stand out.

Use Case 1: High Risk, development programs: In the past 25 years, there have been a number of high-tech defense programs that have, significantly overrun costs and schedules, including the Future Combat Systems/DDG-1000/Railgun programs. Post program analysis of these and other programs have determined that the technical risk on these programs were much higher that program managers were aware of early in the program life cycles, even though the knowledge of these risks did exist. Use case 1 is there for the need to better assemble knowledge about program risk early in a program to better understand shortfall in knowledge about the system that would indicate high development risk.

Use Case 2: In the opposite case when the DoD is developing a new system with lower technical risk, it is important to also have a clear view of the technical risk profile of the system throughout its development. On programs with lower technical risk profiles, there is also a good deal of knowledge about the system design and future performance of the system that is known. By capturing more of this knowledge, program decision-makers can structure programs to accelerate schedules based on reasonable risk profiles, given that knowledge about the system allows them to reduce the cost and time of gathering additional (and redundant) knowledge.

These specific use cases inform the functional development, design, and implementation of a metric and metric reporting system within the digital program, and Model Based Systems Engineering (MBSE) methods that are currently being integrated into Defense program management and engineering.

Development of Digital Models in the Test Process

The digital models that are being developed as a part of digital test engineering are a key part of implementing knowledge-based decision making in the DoD. Specifically, the Integrated Decision Support Key (IDSK) was developed to link decisions to specific Knowledge sources and tests. The Model Based TEMP in kind was developed to link requirements, design, and test planning though digital modeling methods. This work extends these models, by the addition or other knowledge sources and the development of specific metrics analysis.

Model Based IDSK

To realize the IDSK's potential to positively impact acquisition outcomes and program decisions, the concept of a MB-IDSK developed using model-based systems engineering will address a majority of the shortcomings of the traditional IDSK and provide great benefits to decision-makers and all stakeholders across the acquisition and T&E enterprise. These benefits include (i) its ability to support the T&E-as-a-Continuum (Collins & Senechal, 2023) framework by integrating the IDSK into a program's digital engineering ecosystem, (ii) an MB-IDSK would provide mapping of decisions to development (i.e., acquisition) risk, test risk, and test resource models, thereby allowing for more sophisticated analysis including probabilities of success analysis, (iii) an MB-IDSK will expand the ability to link different aspects of the system design, capabilities, and testing to critical program decisions.

In Anyanhun and Arndt (2024), a MB-IDSK reference architecture (MB-IDSK RA) was proposed and developed to support digital transformation efforts of DOT&E. The motivation behind defining a MB-IDSK RA was based on the premise that an architecture should reflect the organization of the owning enterprise (CAS, 2022). Specifically, the MB-IDSK RA represents an essential tool to facilitate communication and alignment efforts of current and future IDSK architectures. Figure 3 depicts the IDSK architecture strategy as adapted from the DoD Comprehensive Architecture Strategy.

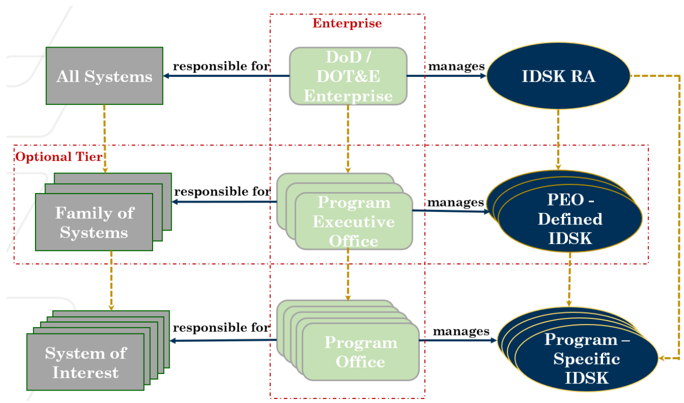


Figure 3. IDSK RA Architecture Strategy. Adapted from Figure 1 of the DoD CAS (CAS, 2022).

Equipping DoD acquisition programs with overarching guidance on how to leverage digital engineering for decision support is critical to achieving the enterprise-wide business and mission objectives of providing weapon systems at the speed of need and relevancy. As reported in CAS (2022) and Muller (2007), a Reference Architecture provides a method for focusing all architecture and design decisions.

Model Based TEMP

The MB-TEMP reference architecture captures the essence of the test planning and decision support domain relative to the needs of program offices, DOT&E, DTE&A, T&E practitioners and decision-makers. For the purpose of this article, abstractions and simplification concepts have been utilized in relation to how some diagram views appear and how they are presented in this work to enhance legibility. More importantly, the architectural strategy employed in the development of the TEMP RA results in a digital engineering artifact (tool) that, when instantiated, will seamlessly integrate into the digital engineering ecosystem of a program. Figure 4 is an example of the complete set of views that together make up the MB-TEMP RA description.

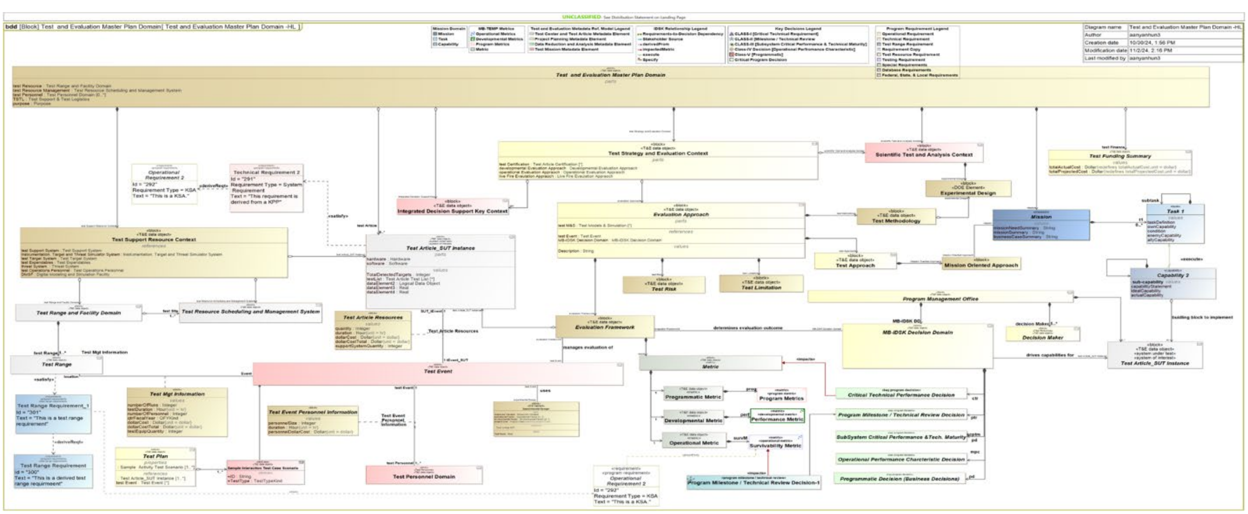


Figure 4. The TEMP Domain view of the MB-TEMP RA

The TEMP Domain view of the MP-TEMP RA provides crucial insights into the top-level composition of the TEMP domain. The RA view links together elements defined within the TEMP model and elements already defined in digital models that exist within a program's digital engineering ecosystem. These digital models include a program office model, requirements model, system model, SUT model, and test range models.

Development / Theory

There are two primary questions related to test and evaluation (T&E) relevant to effective and efficient fielding of warfighting capabilities: 1) Can we proceed to the next stage of development, and do we have an adequate understanding of the level of risk? 2) Do we have enough redundancy and reliability in our knowledge sources to accelerate the development and/or the test?

With respect to the first, the risk, we need to understand the requirements being levied, from request for proposal (RFP) to source selection, and be able to ascertain if we are proceeding at a higher risk than we realize. For the second, we need to determine if we know enough that we can truncate testing activities such that we do not undertake more testing than we need to do. For example, if we have enough redundancy and quality of data across relevant legacy performance data, design data, and modeling and simulation (M&S) data, can we reduce or otherwise compress the amount of data we need from developmental test activities? The development of a high-level characterization method to address risk and the maturation of knowledge for T&E activities is detailed in the following sections to answer these questions.

Knowledge Source Characterization Categories

There are five primary sources of data and information regarding performance of a capability under development that come together to produce knowledge about how that system will perform across the intended concepts of employment:

1. Legacy data and performance data from prior components or similar systems (K_1)
2. Design data (K_2)
3. Modeling and simulation data (K_3)
4. Developmental test data (K_4)
5. Operational test data (K_5)

These sources are typically sequential in time and in degree of "reliability."

Knowledge Based Metrics

Development of Knowledge for a Given Source

Every test decision should be linked to a specific knowledge source; this concept extends such that every key performance parameter (KPP) should also be tied to a knowledge source or sources and a test program. This concept forms the basis for how we produce an abstract representation of knowledge. While the ensuing approach is quantitative in nature for ease of propagation and understanding, there is a strong degree of qualitative expertise and assessment that underlies the numbers. The intent is that this approach will improve as use and experience mature its implementation.

Our approach develops a knowledge source characterization for a given data element that contributes to a specific K_j for a system under development and test, that is:

- $K_j(S)$ – total knowledge obtainable about a system (S) from a given source (j, 1 through 5 above).
- $K_j(s_i)$ – knowledge obtainable about a specific data element, here corresponding to specific part of the system or a subsystem relevant to the new system under development and test.

Each $K_j(s_i)$ is approximated as having three contributing dimensions: i) quality or fidelity of data from the given source, ii) similarity of data from the given source to the specified system or subsystem under current development and evaluation, and iii) completeness of data from the given source with respect to the KPPs being assessed for the specified system or subsystem under current development and evaluation.

All $K_j(s_i, t)$ will exist on the same range of [0,1], effectively representing a normalized level of representation. Similarly, each of the three contributing dimensions $K_j(s_{i\_quality})$, $K_j(s_{i\_similarity})$, and $K_j(s_{i\_completeness})$ are defined on [0,1]. In the absence of grounded data or experience to suggest otherwise, we used a simple minimum to bring dimensions of quality, similarity, and completeness together at a single point in time. The rationale for this is the maximum knowledge attainable for the specific measure or subsystem in question with respect to the relevant KPP should not exceed its minimum value across these dimensions. For example, if if $K_j(s_{i\_quality}, t) = 0.8$, $K_j(s_{i\_similarity}, t) = 0.8$, and $K_j(s_{i\_completeness}, t) = 0.1$, then $K_j(s_i, t) = 0.1$.

Next, we considered how $K_j(s_i, t)$ would come together to produce $K_j(S, t)$. A geometric mean is well suited to describing proportional and varying growth and is appropriate when the data in question may be sustainably different in either its properties (i.e., what it represents) or across its range. Intuitively, it represents the average position of the "center of mass" of a system of particles if each particle had the same weight. In this problem, with data, it represents the centroid of the finite collection of values for $K_j(s_i, t)$ across $s_i = 1, 2, 3, \ldots n$. Table 2 provides an example for these steps for four subsystems. We assume that we start with zero knowledge, treating the accumulation of knowledge from legacy information as the starting point in our process, and that legacy data varies over time only due to discovery and effective interpretation of that discovery. Knowledge values can remain constant over multiple time points if no new information is gleaned from one step to the next.

Table 2. Example Calculation of Subsystem and then System Knowledge Accumulation for a Given Source Type over Time

	t0	t1	t2	t3	t4	t5
Kj(si_quality)	0	0.3	0.3	0.3	0.4	0.4
Kj(si_similarity)	0	0.2	0.25	0.25	0.25	0.4
Kj(si_completeness)	0	0.3	0.4	0.45	0.45	0.5
K1(s1, t)	0	0.2	0.25	0.25	0.25	0.4
Kj(si_quality)	0	0.1	0.1	0.1	0.1	0.1
Kj(si_similarity)	0	0.3	0.3	0.4	0.4	0.4
Kj(si_completeness)	0	0.1	0.2	0.2	0.25	0.25
K1(s2, t)	0	0.1	0.1	0.1	0.1	0.1
Kj(si_quality)	0	0.8	0.8	0.8	0.8	0.8
Kj(si_similarity)	0	0.2	0.25	0.25	0.25	0.4
Kj(si_completeness)	0	0.33	0.35	0.375	0.45	0.5
K1(s3, t)	0	0.2	0.25	0.25	0.25	0.4
Kj(si_quality)	0	0.6	0.6	0.6	0.7	0.7
Kj(si_similarity)	0	0.9	0.9	0.9	0.9	0.9
Kj(si_completeness)	0	0.8	0.8	0.8	0.85	0.85
K1(s4, t)	0	0.6	0.6	0.6	0.7	0.7
K1(S, t)	***0.00***	***0.221***	***0.247***	***0.247***	***0.257***	***0.325***

The subsequent steps considered integrative loss and the maximum attainable knowledge possible from a given source (i.e., legacy or developmental test). For the former, all of the point information gained through the development and test process will not perfectly combine into knowledge without loss. This loss will exist because perfect integration across disparate dimensions is exceedingly difficult in practice, and all of the dimensions, their attributes, and higher-order relational effects may not be visible. To capture this effect, a parametric equation was selected from the development in (McDermott et al., 2019) and adapted in meaning for application to the knowledge problem in this paper as follows:

$x\_self(t) = x\_self(t-1) - sgn(\Delta)*\eta*|\Delta|$

where:

x_self (t)	Knowledge at the current time t
x_self (t-1)	Knowledge at the prior time (t-1)
Δ	The difference in the current node state and the target state: (x_self (t-1) - x_target(t))
sgn(Δ):	The sign, or signum, function of Δ.
η	Eta. The shaping parameter that governs how the knowledge is matured (i.e., suddenly vs exponentially). $0 \leq \eta \leq 1$
$\lvert\Delta\rvert$	The absolute value of Δ (also expressible as (Δ*sgn(Δ)).

The target, *x_target(t)* in the *Δ* above, is defined by the parents of the knowledge at time t, *x_self(t)*. Specifically, for the knowledge metric problem described here, the parents of each K(S, t) will correspond to its $K_j(s_i, t)$. Here, the sign of Δ will be negative only if some discovery invalidates or calls prior information into question.

The shaping parameter η controls how quickly *x_self(t)* approaches *x_target(t).* The closer η is to 1, the faster *x_self(t)* approaches *x_target(t).* The shaping parameter can be altered based on experience, the nature of the knowledge source, and/or the nature of the system under development and test. Moreover, there are multiple ways to define *x_target(t)* as a function of the parent contributions; a deterministic maximum or minimum, or any function of the parents are potential approaches. In this work, *x_target(t)* is defined as the geometric mean of the parent contributions as described previously. To illustrate the equation's behavior, Figure 5 illustrates the notional knowledge increase, *x_self(t)*, for a constant target *x_target(t)* = 0.5 across varying values of η.

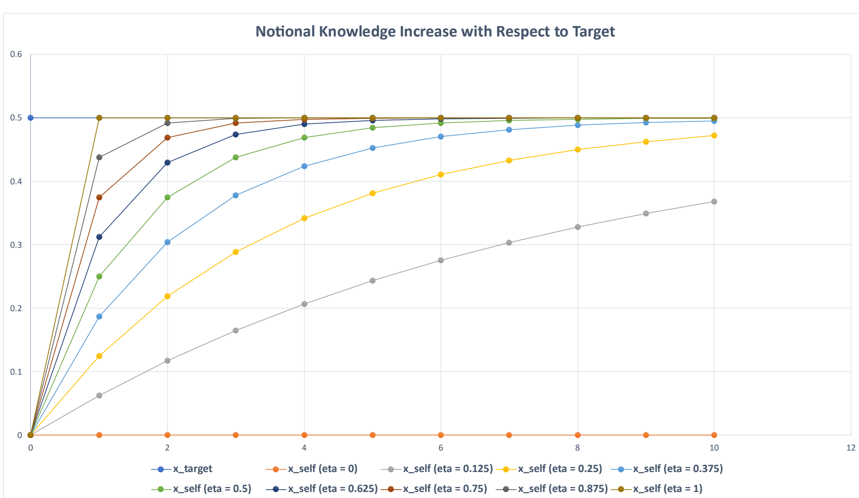


Figure 5. Illustration of Parametric Equation Behavior for Various Values of the Shaping Parameter

As per the description earlier in this section, each $K_j(S, t)$ will be defined by the geometric mean of its contributing $K_j(s_i, t)$. The parametric equation will be used to define the level of knowledge remaining after taking integrative loss into effect.

Specifically, $K_j(S_{\_parametric}, t)$ represents the actual knowledge level attained from $K_j(S, t)$ due to less than perfect knowledge integration for the whole system. Note that if the shaping parameter η is set to 1, then $K_j(S_{\_parametric}, t)$ will equal $K_j(S, t)$ (i.e., no integrative loss is considered).

Further, the maximum attainable knowledge possible from a given source needs to be defined based on the maximum amount of knowledge the development and/or test team anticipates is possible to gain from a given knowledge source (i.e., legacy data). Again, this value, $K_j(S_{\_max})$, will be defined by the team as a value in the range of [0,1], treating it as the decimal equivalent of a percent. $K_j(S_{\_max})$ is considered constant over time. $K_j(S, t)$ and $K_j(S_{\_parametric}, t)$ will be scaled against this maximum value, for example, $K_{j\_scaled}(S, t) = K_j(S_{\_max})* K_j(S, t)$. Continuing the example from 2, 3 illustrates these parts of the process.

Table 3. Continued Example Showing Integrative Loss and Scaling with Maximum Attainable Value for a Given Knowledge Source

Eta	**0.7**	(Constant over time)				
K1(S_max)	**0.2**	(Constant over time)				
	t0	**t1**	**t2**	**t3**	**t4**	**t5**
K1(S, t)	0.000	0.221	0.247	0.247	0.257	0.325
K1(S_parametric, t)	0.000	0.155	0.220	0.239	0.252	0.303
K1_scaled(S, t)	0.000	0.044	0.049	0.049	0.051	0.065
K1_scaled(S_parametric, t)	0.000	0.031	0.044	0.048	0.050	0.061

Development of the Knowledge Accumulation, Gaps, and Integration of Knowledge Sources over Time

We need to capture (i) how much total knowledge about a new system under development and test is being accumulated over time, and (ii) how far behind or ahead the process of knowledge accumulation is compared to its anticipated levels.

Total knowledge accumulated over time is represented through $K_j(S_{\_parametric}, t)$, the actual knowledge level attained from $K_j(S, t)$ due to less than perfect knowledge integration for the whole system. More specifically, this is represented by $K_{j\_scaled}(S_{\_parametric}, t)$, which places the value on the same comparative basis as the maximum, needed, and planned knowledge levels for the given source.

How far behind or ahead the knowledge development process is from what is needed is defined by evaluating the total knowledge accumulated against the amount of knowledge needed. Similar to $K_j(S_{\_max})$, the maximum level of knowledge deemed attainable for a given source, the amount of knowledge needed is also determined by the development and/or test team. These values may, however, increase in time. For example, $K_{j\_needed}(S, t)$ may be defined to increase linearly in time, reaching a maximum value equal to $K_j(S_{\_max})$. The knowledge delta, or gap, may be evaluated as $K_{j\_needed}(S, t) - K_{j\_actual}(S, t)$. Roughly, the actual knowledge corresponds to what the team has while the knowledge delta represents what the team lacks. Figure 6 illustrates these concepts as a continuation of the previous example for a single knowledge source.

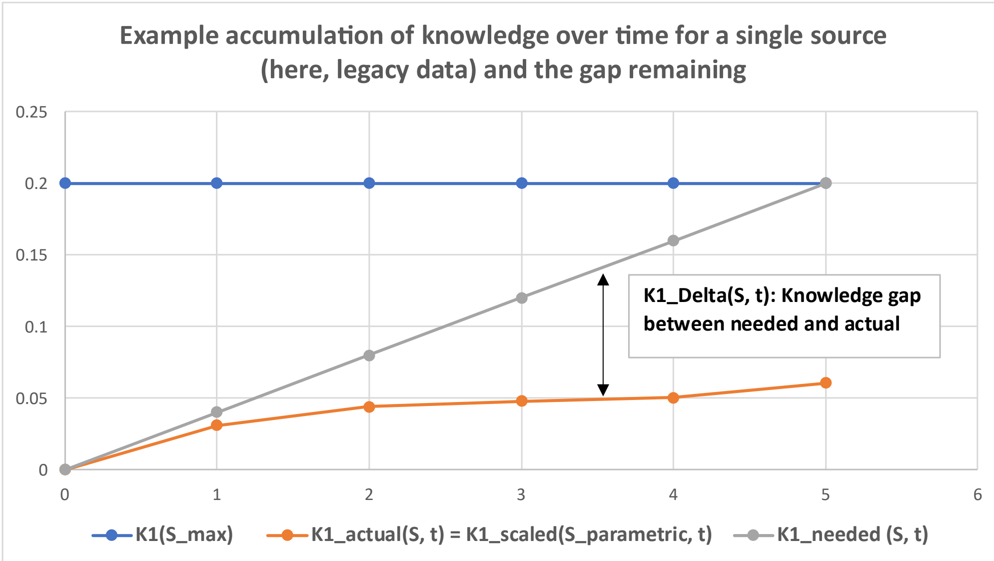


Figure 6. Example Knowledge Accumulation and Remaining Gap over Time

Figure 7 then shows how these concepts extend to combine knowledge sources over time. In the figure, the left graph shows an example where knowledge is not being gained sufficiently in the legacy data activities in comparison to what is expected. The right graph shows the potential effect of recognizing this early and shifting the collection of design data and M&S data earlier in time.

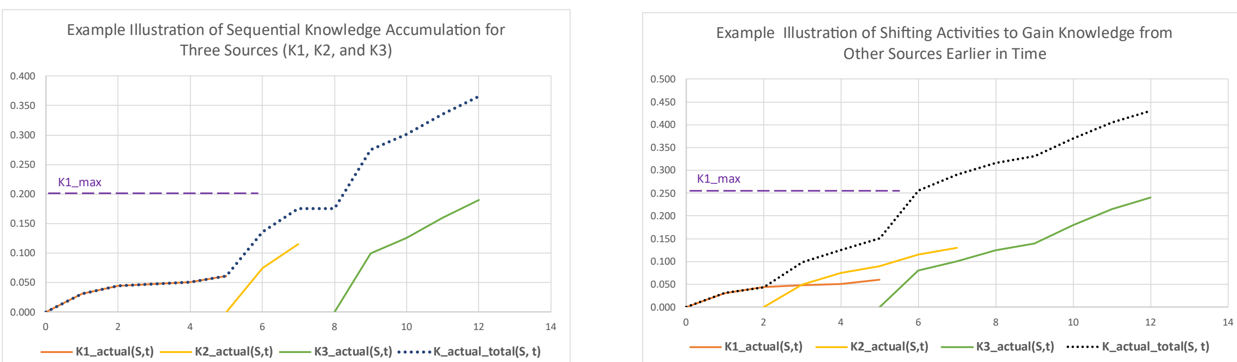


Figure 7. Example Knowledge Accumulation over Time for Three Knowledge Sources

Using the same knowledge accumulation profiles shown in Figure 7, Figure 8 illustrates how recognizing a knowledge deficit early and pushing other activities earlier can reduce the overall knowledge gap with which the development and test teams are proceeding.

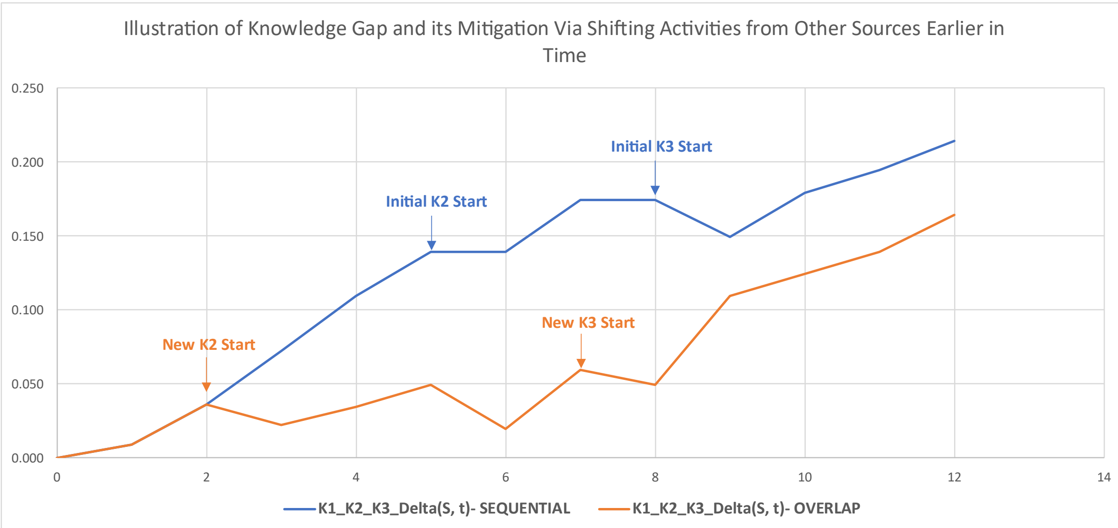


Figure 8. Illustration of Knowledge Gap Reduction via Shifting Activities Earlier in Time to Mitigate a Source Deficit

Discussion of Approach and Limitations

The math and associated methods for knowledge accumulation are simple and easily modifiable at any point in the procedure to better reflect reality as ascertained from experience. A significant challenge, however, is the quality of the values that start this knowledge metric approach. Specifically, if the evaluations of the dimensions $K_j(s_{i\_quality})$, $K_j(s_{i\_similarity})$, and $K_j(s_{i\_completeness})$ that come together to define each $K_j(s_i)$ are wild guesses, then that lack of grounded approximation will carry through the rest of the assessment. Again, each $K_j(s_i)$ represents the knowledge obtainable about a specific data element, here corresponding to specific part of the system or a subsystem relevant to the new system under development and test. This goes hand-in-hand with the quality of KPPs and the mapping of data creation activities to best evaluate the system's performance with respect to the KPPs.

Implementation and Use

In order for this process to and metric to be useful it needs to be implemented in both the digital thread of program and in a decision support tool for decision-makers. The metric will be integrated into the digital thread of the program in order to ensure that the data is automation updated as the design of the system changes, and as we gain additional knowledge sources.

Decision Support and Decision-Maker Displays

At a practical level this means the development and implementation of a set of Dashboards for program managers and other decision-making stakeholders. Figure 9 shows such a dashboard. What is displayed is a graphical representation of the current level of knowledge about the systems KPP (critical performance parameters) at specific key decisions points in the program, in this case before RFP release. On the display we see the current level of knowledge versus the expected amount knowledge and the resulting risk or opportunity. Additional dashboards will display the specific knowledge sources, and opinions to mitigate risk or exploit opportunity (accelerate schedule).

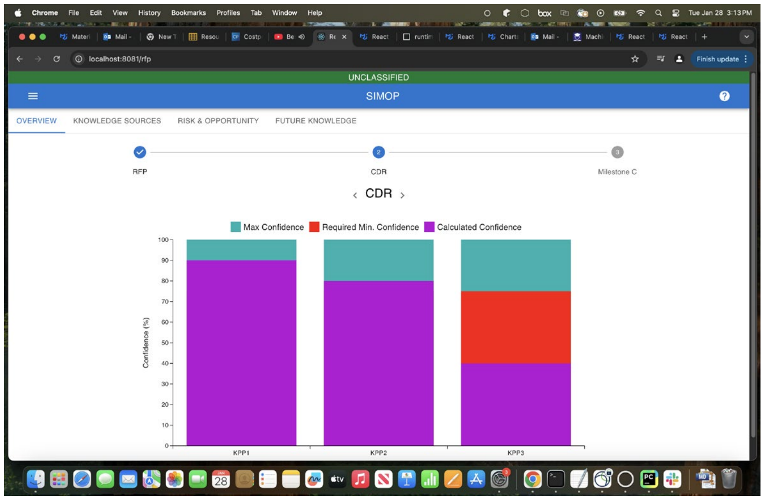


Figure 9 Program Risk by KPP by Time.

Knowledge Source Management/Test Program Trade Studies

As we talked about in the decision support section, the results of identifying knowledge shortfalls and surpluses of knowledge at specific point in time managing the development of a program will create the need and opportunity for adjustments in the baseline of the program. In the case of high risk due to knowledge shortfall, we would need to add design and development resources, and/or tests to increase knowledge. In the case of excess knowledge, the baseline for test and other knowledge sources can be modified to reduce current and future redundant sources of knowledge. However, in complex systems the relationships between the planning and execution of these knowledge sources is also complex. To facilitate better management, we as a part of this program crating standard models and characterizations of knowledge sources, so that we can manage the different knowing sources together. This allows us look at all of the knowledge sources, and make decisions about reductions or additional to knowledge sources the way you playing the video game “Tetris” (Figure 10).

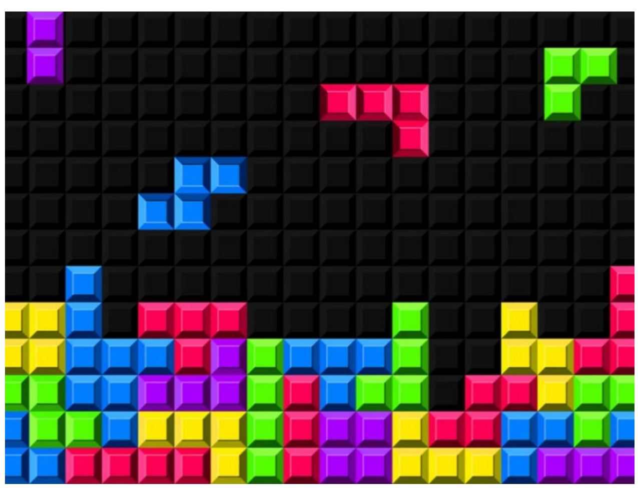


Figure 10 Test Program Dynamic Management

Knowledge Management Stack

We can better see how the knowledge management system is integrated with the rest of the models used in managing systems under the MBSE process. Figure 11 shows the bottom to top elements of the system. At the top we see the dashboards, and the knowledge source metrics that support the dashboards. Below that are the system level models that the knowledge metrics interface with, and below that are the system requirements and requirements models.

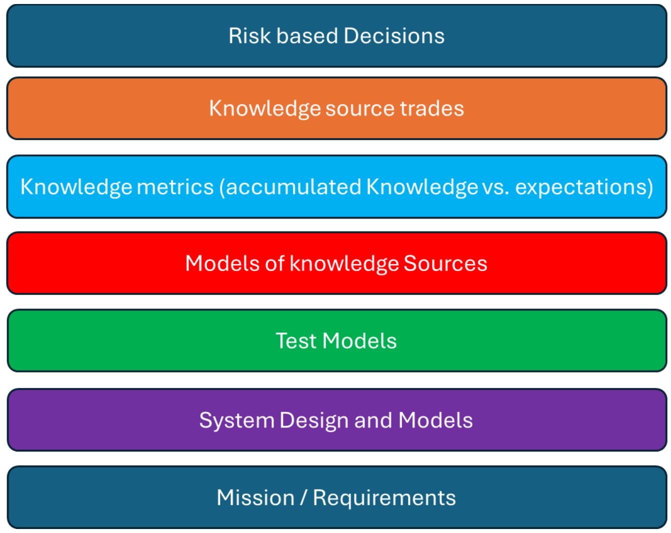


Figure 11 Knowledge Management Stack

Modeling

Central to the development of the knowledge-based metric is the modeling of the different knowledge sources. The models of the knowledge sources capture several key aspects of the knowledge source including 1. The knowledge source class, 2 the Knowledge source characteristics, and 3. The interfaces and links to the knowledge source. The classes of knowledge sources are 1. Legacy system performance, 2. System Design, 3. Modeling and simulation, 4. Developmental Test, and 5. Operational test.

Figure 12 is a SysML view of the knowledge source model that will exist for all knowledge sources. Embedded in the model for each knowledge source will be the following characteristics: 1. The knowledge source class, 2. The requirements that this knowledge source is linked to, 3. The design sub-systems related to this knowledge source, 4. The similarities of the knowledge source to the true system being designed, 5. The performance profile coverage for the specific knowledge source (including operating environment), 6. The reliability of the knowledge source, 7. The fidelity of the knowledge source, 8 the schedule associated with the knowledge source, 9. The cost associated with the knowledge source, and 10. The required inputs and predecessor event(s) to execute the knowledge source.

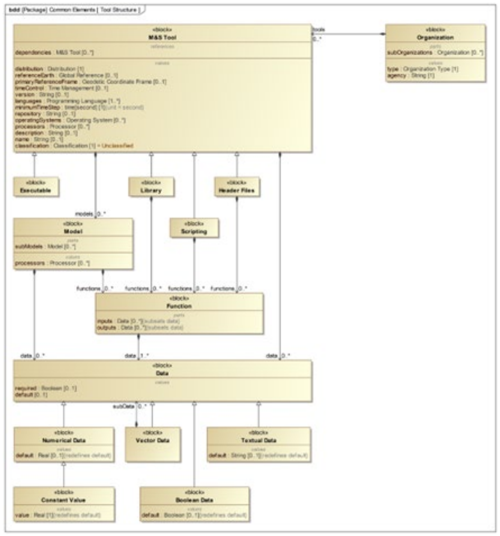


Figure 12 Knowledge Source Model

Once the different knowledge source are documented and modeled we will be able to see where there are gaps and redundancies in the knowledge we need to make decisions about the performance and development of the system of interest.

Integration with the Digital Thread

The models of the knowledge sources can then be integrated with the other models of the system including the requirements and test model described earlier in this paper. In addition, knowledge metrics can contribute significantly to informed risk management and decision-making within the program life cycle. In this way the metrics are both integrated into the digital thread of the system and are available as needed for decision-makers.

Conclusion and Recommendations

As we have seen in this paper there is an opportunity to take better advantage of available knowledge is critical, to our ability to inform and prioritize the two use cases that we highlighted earlier in this paper, first the ability to make better determinations about very high-risk technical projects, second to develop realistic approaches to accelerate development and fielding of systems when we have the knowledge needed to do so. For far too long the acquisition system's best risk management processes have been embedded in the minds and histories of its senior most program management and engineering team member.

Knowledge Source Management

The management of a program's test plan is a requirement for all programs and is captured in the Test and Evaluation Master Plan; likewise test program and test cases are developed and managed by Test and Evaluation working group and by the independent test authority. This covers operational test, and some modeling and simulation, and some developmental test. However, there is very little effort to capture the knowledge of system performance in legacy performance data and in design outside of the design process. As a result, much of the knowledge that is available about system is not captured, curated, and managed in a consistent and portable way to make it available when and where it is needed. This has led in the past to many poor assumptions and decisions being made for lack of visibility into knowledge that already exists.

Using Knowledge Metrics in Portfolio Management

The knowledge-based metrics once standardized have the potential to be used outside of individual program office. Like other digital processes in the MBSE umbrella a significant part of the metrics value is realized at a higher level than the individual program element. Because the knowledge sources and the metrics are math based, and created in model form, they can be aggregated and shared across programs to help manage portfolios of program in several ways. At the top level the metrics can be used to evaluate the status and risk of different technical development and testing efforts withing a portfolio of programs and make strategic decisions about where to spread or concentrate risk, and resources.

Using the Knowledge Metrics in Mission Engineering

Mission engineering is also an area where we can use the knowledge-based metrics. Mission engineering is the synchronization, management, and coordination of concepts, activities, technologies, requirements, programs, and budget plans to guide key decisions focused on the end-to-end mission.[20] The knowledge-based metrics will provide significant and important, information for mission-based decision making.

Using Knowledge Metrics in Service Level Budget Management

The use of this class of metrics may also have significant applications to budget management and capability versus budget trade space analysis. The knowledge-based metrics can be used to help provide data for trade studies.

Knowledge Based Interactions with Industry

Finally in order to make any of our advanced risk management methods work, we (the DoD) are going to need to work much closer with the vendor base to get better insights into the maturity of their designs, and other important knowledge sources that they rely on (legacy system data for sub-systems, design, and developmental test data) and that they often do not share. From a technical standpoint, the shift to MBSE gives the tools we need to capture and use this information, but the historical business relationships (contracting) that we have with vendors do not incentivize them to share this information. Both the technical and the business relationship with our vendors need to change significantly.

References

Anyanhun, A., & Arndt, C. (2024, September). *The use of AI and systems engineering processes in and for testing of AI-based systems*. SERC AI4SE/SE4AI Workshop,

Army Aviation and Missile Command. (2022). *Comprehensive architecture strategy (CAS)*. https://apps.dtic.mil/sti/pdfs/AD1185001.pdf

Arndt, C., Anyanhun, A., & Werner, J. S. (2023). *Shifting left: Opportunities to reduce defense acquisition cycle time by fully integrating test and evaluation in model based systems engineering*. Acquisition Research Program.

Beers, S. M. (2023). *Integrated Decision Support Key (IDSK) for capability evaluation-informed MDO decision-making*. ITEA MDO Workshop, July 2022.

Beers, S. M., Hutchison, S., & Mosser-Kerner, D. (2013). Developmental test evaluation framework: Focus on evaluation and analysis for acquisition success. *Phalanx*, *46*(3), 36–39.

[20] Mission Engineering Guild, November 2020. Office of the Deputy Director for Engineering Office of the Under Secretary of Defense for Research and Engineering.

Collins, C., & Beers, S. M. (2021). *Mission engineering, capability evaluation & digital engineering informing DoD technology, prototyping and acquisition decision*. NDIA S&ME conference, December 2021.

Department of Defense. (2020). *DoD Instruction 5000.89: Test and evaluation.*

DOT&E. (2022). *Office of the Director, Operational Test and Evaluation Strategy Update - Strategic Pillars, Final DOTE 2022 Strategy Update*. https://www.dote.osd.mil/Portals/97/pub/reports/FINAL%20DOTE%202022%20Strategy%20Update%2020220613.pdf?ver=KFakGPPKqYiEBEq3UqY9IA%3d%3d

Executive Services Directorate. (2020). DoD Instruction 5000.89. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500089p.PDF

IEEE/ISO/IEC 15288-2015. (n.d.). *IEEE SA-IEEE/ISO/IEC 15288-2015*. https://standards.ieee.org/ieee/15288/5673/

Gregory, J., & Salado, A. (2024). An ontology-based digital test and evaluation master plan (dTEMP) compliant with DoD policy. *Systems Engineering.*

Gregory, J., & Salado, A. (2024). dTEMP: From digitizing to modeling the test and evaluation master plan. *Naval Engineers Journal*, *136*(1), 134–146.

Gregory, J., & Salado, A. (2024). Towards a systems engineering ontology stack. *INCOSE International Symposium*, Dublin, Ireland.

ISO/IEC/IEEE. (2022). *Systems and software engineering—Architecture description* (document SO/IEC/IEEE 42010:2022(en)).

Muller, G., & Hole, Eirik. (2007). *Reference architectures; Why, what and how* (White paper). Architecture Forum Meeting, Embedded Systems Institute and Stevens Institute of Technology, Hoboken, NJ, March 12–13, 2007.

OpenCAESAR. (2023). Ontological Modeling Language 2.0.0 (Online). https://www.opencaesar.io/oml/

Shaffer, A., & Whitley, J. (2024, September 20). Accelerating defense innovation requires changes to the acquisition approach. *National Defense Magazine*.

Test and Evaluation Metadata Reference Model. (2011). *Range Commanders Council. Director, Operational Test and Evaluation (DOT&E) Test and Evaluation Master Plan (TEMP) Guidebook Version 3.1.*

Werner, J. S., & Arndt, C. (2023). *Development of digital engineering artifacts in support of MBSE based test planning, execution, and acquisition decision making*. Acquisition Research Program.

The Impact of the Joint Interagency Field Experimentation Program on Small Business Success

Ashley Book—is Research Faculty in the Information Sciences Department at the Naval Postgraduate School, where she also pursues a Master of Science in Applied Design for Innovation. She has almost a decade of experience supporting Department of Defense innovation through field experimentation, small business engagement, and interagency collaboration. Her professional background includes environmental planning with a focus on hazard mitigation and emergency management. Book's research explores the role of innovation ecosystems in national security, emphasizing the impact of experimentation events on technology acceleration and small business success. [Ashley.book@nps.edu]

Abstract

This paper examines the impact of the Joint Interagency Field Experimentation (JIFX) program at the Naval Postgraduate School (NPS) on the success of small businesses within the Department of Defense (DoD) innovation ecosystem. JIFX provides a collaborative, real-world experimentation environment for companies to test emerging technologies, refine capabilities, and engage with government stakeholders. Through a mixed-methods approach including quantitative data analysis and interviews with repeat and first-time JIFX participants, this study identifies how JIFX participation contributes to technological development, strategic business positioning, and long-term outcomes such as funding, acquisitions, and follow-on contracts. The research finds that JIFX serves as a strategic entry point into the defense market and fosters a resource-based competitive advantage for its participants. Recommendations are provided for enhancing the JIFX program's visibility, collaboration mechanisms, and integration with downstream acquisition pathways. These insights support policy and program development to better connect early-stage innovators with DoD operational needs.

Executive Summary

This paper examines the Joint Interagency Field Experimentation (JIFX) program and its influence on the success of small businesses that participate in its events. JIFX serves as a platform for private-sector innovators to engage with the Department of Defense (DoD) and other government agencies, providing an opportunity to test emerging technologies in an operationally relevant environment. The objective of this research is to assess the benefits of JIFX participation for small businesses. The findings of this study will be valuable to multiple stakeholders, including small businesses seeking entry into the DoD innovation ecosystem, policymakers shaping defense innovation initiatives, and JIFX organizers looking to enhance the program's impact on private-sector innovation.

Problem Statement

Despite the DoD's emphasis on fostering innovation, small businesses often struggle to access and integrate into its complex acquisition and development frameworks. JIFX presents a unique opportunity for these businesses to showcase and refine their technologies in a low-risk, collaborative setting. However, the specific impacts of JIFX participation on small business success remain largely anecdotal, with limited empirical analysis available. This research seeks to bridge that gap by systematically evaluating the role of JIFX in supporting small business growth, measuring key outcomes such as technological development, investment traction, and government contracting success. By identifying patterns and trends, this study aims to provide a data-driven understanding of how JIFX contributes to the broader DoD innovation ecosystem and how it can be leveraged to enhance small business participation and success.

Research Approach

This paper employs a comprehensive research approach to examine the perceived impact of the Joint Interagency Field Experimentation (JIFX) program on the success of participating small businesses. Specifically, the research focuses on how participation in JIFX events influences the technological development and strategic direction of these businesses. The methodology is divided into three core areas:

1. **Creation of the JIFX Participant Database.** A foundational element of this research involves compiling and structuring a comprehensive database of all JIFX participants. This database will aggregate historical data on attendees, including company information such as size, headquarter location, and founding year, and financial information such as investments and acquisitions. Future research could incorporate funding mechanisms such as the Small Business Innovation Research (SBIR) fund.
2. **Data Analysis of the JIFX Participant Database.** Once the database is established, a detailed analysis will be conducted to identify patterns and trends in repeated participation and business success following attendance at JIFX events. Key metrics such as the frequency of participation, the nature of innovations showcased, and measurable outcomes (e.g., contracts awarded, partnerships formed) will be examined. This analysis aims to quantify the impact of JIFX participation on technological development and business growth.
3. **Interviews with Historically Active Participants.** To gain qualitative insights, interviews will be conducted with select companies that have attended multiple JIFX events. These discussions will explore the motivations behind continued participation, the perceived benefits of engagement, and how JIFX has influenced their technological development and broader business strategies. The aim is to uncover longitudinal impacts and unique perspectives on how JIFX fosters innovation and collaboration for small businesses.

The combination of structured data and personal insights ensures a complete understanding of the role JIFX plays in supporting the success of small businesses through technological innovation and strategic development.

Introduction

The Department of Defense (DoD) innovation ecosystem is a dynamic network of programs, organizations, resources, and partnerships aimed at driving technological advancements to meet evolving defense and security needs. At its core, the ecosystem thrives on principles of collaboration, experimentation, and rapid innovation. By fostering relationships across government, academia, and industry, it enables the identification, development, and scaling of emerging technologies for defense applications. A widely held view within the ecosystem is that traditional development and acquisition methods are too slow to keep up with the rapid pace of technological change. To address this, many emphasize the need for the DoD to leverage commercially available technologies that are already advancing quickly to ensure the United States remains competitive and responsive to emerging challenges.

The strength of the DoD innovation ecosystem lies in its ability to leverage expertise and resources across multiple domains. By creating a pipeline for technology development, experimentation, and transition to operational use, the ecosystem accelerates the delivery of advanced capabilities to the warfighter. This collaboration supports the growth of innovative technologies that enhance national security. Additionally, it ensures the DoD can respond swiftly to emerging challenges by integrating cutting-edge solutions from both commercial and defense sectors.

NPS Innovation

The Naval Postgraduate School (NPS) plays a pivotal role in the DoD's innovation ecosystem by facilitating advanced education and research in military and defense technologies. The 2023 NPS Strategic Framework reports the establishment of an Innovation Pillar to lead naval innovation via a collaborative ecosystem connecting NPS students with academia and industry. Through programs like the Joint Interagency Field Experimentation (JIFX), the Warfare Innovation Curriculum (WIC), Naval Innovation Exchange (NIX), and the Naval Innovation Center (NIC), NPS provides a unique environment to support emerging technologies. Additionally, master's degree programs, such as the Applied Design for Innovation curricula, equip military leaders with experiential learning in design-thinking, social science methods and creative collaboration to address innovation challenges in the context of evolving technology and military challenges

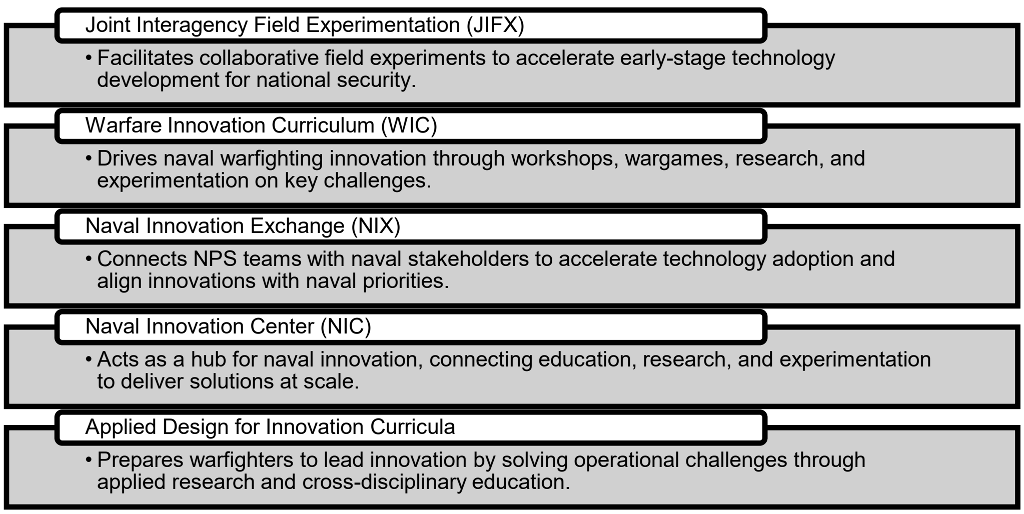


Figure 1. Naval Postgraduate School Innovation Efforts

Joint Interagency Field Experiment

JIFX is a sponsored research project at NPS, and it exists to support experimentation in alternative methods to enable rapid technological development. JIFX is a multi-faceted program focusing on: (1) a community of interest focused on supporting emerging technologies and (2) broadly scoped quarterly collaborative field events. The community of interest and quarterly field experimentation events enable DoD, U.S. government, and allied stakeholders to identify, influence, and accelerate early-stage technology development that addresses national and collective security challenges. The JIFX events are driven by five tenets, shown in Figure 2, creating a unique culture of collaboration, inclusivity, and mutual learning.

Figure 2. JIFX Tenets

JIFX operates on a quarterly cycle with events typically hosted in February, May, August, and November. The proposal submission, driven by an annual Request for Information (RFI), for each event opens at the beginning of each fiscal year and closes approximately 75 days before each event. The proposals are reviewed by JIFX stakeholders for relevancy and the NPS team for safety, legality, and applicability for the event. The NPS team manages all flight, spectrum, laser, and safety approval processes for all experiments. Once approved, experimenters are invited to participate in a series of virtual coordination calls leading up to the event. Approximately 2 weeks before the event, the coordination call is devoted to experiment introductions amongst the group to facilitate collaboration opportunities. Experimenters are encouraged to work with other technologists prior to the event to coordinate any collaborations or integration experiments (*JIFX FY24 Request for Information*, n.d.).

- Unmanned Aerial Systems
- Unmanned Systems Design, Deployment, Operation, Networking and Control
- Countering Unmanned Systems
- Communication and Networking
- Cyber, Cyber Security, and Electronic Warfare
- Intelligense, Surveillance and Reconnaissance
- Situational Awareness
- Defense Support to Civil Authorities
- Health and Safety
- Expeditionary Warfare
- Infrastructure and Power
- Mobility and Transportation
- Precision Strike, Non-Lethal Weapons, and Information Operations

Figure 3. JIFX FY24 RFI Focus Areas

The JIFX program had its creation and evolution within the NPS Field Experimentation (FX) Program, which was launched in 2002 to provide operational environments for testing and refining emerging technologies. Initially focused on enhancing capabilities like unmanned aerial vehicle (UAV) support for Naval Special Warfare, the program evolved through partnerships with U.S. Special Operations Command (USSOCOM). Early iterations, such as the Surveillance and Target Acquisition Network (STAN) and later the Tactical Network Topology (TNT), developed networks to link soldiers with tactical sensors and unmanned systems. By 2013, JIFX was established to address broader interagency needs, integrating lessons learned from the FX program's decade-long evolution (Oros, 2014). In 2025, JIFX operated under the sponsorship of the Office of the Under Secretary of Defense for Research & Engineering (OUSD[R&E]) and NavalX, focusing on technology experimentation to meet diverse operational challenges across interagency commands.

JIFX is a critical but specialized component of the DoD innovation ecosystem, offering unique opportunities for experimentation and collaboration without serving as a funding mechanism. While JIFX provides an invaluable platform for small businesses and startups to refine their technologies, its role is inherently limited to facilitating experimentation and fostering connections. This makes the integration of JIFX with other DoD organizations such as the Defense Innovation Unit (DIU), NavalX, and Small Business Innovation Research (SBIR) programs essential for ensuring the long-term success of participating businesses. These complementary organizations provide funding pathways, transition support, and additional resources that JIFX itself does not offer.

For small businesses, JIFX serves as an entry point to the broader DoD innovation ecosystem by providing a unique platform to test solutions in realistic operational scenarios, gather actionable feedback from military and government end-users, and foster connections with key stakeholders. JIFX amplifies its impact through its curated network of participants, which includes subject matter experts from across the DoD. This network enables small businesses to identify potential funding opportunities, such as SBIR or STTR programs, and gain access to mentors who can guide product refinement and commercialization efforts. Furthermore, JIFX offers visibility to military stakeholders who may champion promising technologies through procurement or development programs, thereby positioning participants for success within the DoD acquisition pipeline. Through these targeted mechanisms, JIFX transforms opportunities into tangible pathways for innovation and growth.

JIFX events stand out as uniquely collaborative and innovative within the DoD innovation ecosystem. These events are structured not as traditional trade shows or demonstrations but as dynamic, hands-on experimentation environments where companies, government stakeholders, and researchers come together to push the boundaries of emerging technologies. JIFX prioritizes collaboration and experimentation over competition, creating a supportive atmosphere that encourages participants to explore new ideas, adapt their technologies in real-time, and learn from failure. As one participant described, “The true bleeding edge can be seen live in the making at JIFX because the lack of fear of failure allows experimenters to truly push the bounds of what is possible.” This mindset fosters an environment where early-stage prototypes—even duct-taped and minimally functional technologies—are not only welcomed but celebrated for their potential and growth.

JIFX’s format is intentionally open and fluid. Held at the NPS Field Laboratory at Camp Roberts, participants are free to set up experiments wherever they see fit across the austere environment, which includes runways, open terrain, and specialized facilities. A highlight of the facilities includes access to restricted airspace allowing experimental aircraft to fly without adhering to typical Federal Aviation Administration (FAA) Class C Airspace rules. This setting allows participants to test their technologies in real-world conditions, often adapting or improving them on the fly based on feedback from government science and technology experts or insights gained during the week. For instance, technologies that may not meet specific requirements on the first day can evolve by the end of the week to address those needs, demonstrating the program's emphasis on rapid prototyping and development. The combination of JIFX's unique facilities and its diverse community of attendees creates a unique environment for innovation and collaboration.

Another hallmark of JIFX events is the ad hoc collaboration that naturally occurs among participants. Approximately 50% of attendees represent private industry—startups and established companies—while the other 50% are government stakeholders, including scientists, engineers, and operational experts. This mix creates a productive environment for cross-pollination of ideas and partnerships, where companies often integrate their experiments with one another to explore new applications. For example, one participant might fly a payload for another, while a third integrates the resulting data into a sensor system, forming a chain of innovation that would be difficult to replicate in a more rigid environment. Additionally, pre-event coordination calls and networking opportunities allow participants to connect with potential collaborators before the event even begins.

JIFX events also emphasize networking and mentorship, with many participants leaving with not only valuable feedback but also new relationships and follow-on collaboration opportunities. Government attendees often provide insights into operational needs and potential use cases for the technologies, helping companies refine their approaches. Participants are encouraged to invite additional government stakeholders or contractors to the event, further enriching the ecosystem of expertise and opportunities. This mix of experimentation, networking, and collaboration ensures that JIFX provides significant value for early-stage and cutting-edge technologies, offering a pathway for iterative development and a platform for showcasing potential game-changing innovations.

Looking forward, JIFX is poised to play an increasingly significant role in the evolving DoD innovation ecosystem. As the defense landscape continues to change with new emerging threats, JIFX is structured to incorporate new focus areas into the RFI through the annual update process. Partnerships with entities like the Defense Innovation Unit (DIU) and NavalX are expected to grow, creating a more cohesive pipeline for transitioning experimental technologies into operational capabilities. By continuing to prioritize collaboration, inclusivity, and adaptability, JIFX will remain a foundation of the DoD's efforts to harness innovation from industry, academia, and government stakeholders. Furthermore, there is growing potential for JIFX to incorporate international partnerships, enabling allied nations to collaborate on joint experimentation initiatives.

JIFX exemplifies the value of collaborative experimentation within the DoD innovation ecosystem. By providing an open, hands-on environment for testing and iterating technologies, JIFX bridges the gap between early-stage innovation and real-world application. Its unique format, emphasis on partnerships, and ability to adapt to emerging priorities ensure it remains a vital platform for advancing cutting-edge capabilities. While JIFX is not a standalone solution, its role as an entry point into the broader innovation ecosystem highlights its importance in enabling small businesses, researchers, and technologists to contribute meaningfully to national security. As the DoD continues to pursue rapid innovation and maintain technological superiority, JIFX will undoubtedly play a key role in shaping the future of defense experimentation and collaboration.

JIFX Participants Analysis

Since its inception in 2013, the JIFX program has served as a dynamic platform for fostering innovation and collaboration between government entities, private industry, and academia. Over the years, hundreds of companies have participated in JIFX, representing a wide array of industries, specialties, and organizational sizes. Understanding the composition of these participants and identifying trends across their characteristics is essential to evaluate the program's role in promoting innovation and supporting industry growth. This thesis seeks to analyze a comprehensive dataset of all companies that have participated in JIFX since its inception.

The dataset was constructed using publicly available information from LinkedIn, Crunchbase, and company websites. LinkedIn was utilized to gather detailed information about company size, industry, specialties, founding years, and headquarters location, while Crunchbase provided insights into funding types and growth trajectories. Individual company websites were used to fill in information missing from LinkedIn and Crunchbase. The integration allowed for the creation of a robust database that paints a detailed picture of the JIFX participant landscape.

Crunchbase	Headquarters Location	The primary location of the company’s headquarters. (City, State, County)
	Number of Employees	Categorized company size: 1-10, 11-50, 51-100, 101-250, 251-500, 501-1,000, 1,001-5,000, 5,001-10,000, 10,000+
	Funding Round	The type of funding the company as received: Convertible Note, Debt Financing, Equity Crowdfunding, Grant, Non-equity Assistance, Post-IPO Debt, Post-IPO Equity, Pre-Seed, Private, Private Equity, Public, Seed, Series A, Series B, Series C, Series E, Series F, Undisclosed, Venture - Series Unknown
	Crunchbase Ranking	A dynamic score that measures the prominence of entities (companies, people, investors, etc.) based on factors like connections, funding events, news coverage, and acquisitions, influencing how they appear in search results
	Total Funding Amount	Total amount raised across all funding rounds ($)
	Number of Investors	Total amount raised across all funding rounds
	Investor Names	Name of the investor who participated in the Investment
	Stock Symbol	Stock ticker symbol e.g. AAPL, FB, TWTR
	Valuation at IPO	Value of the Company at IPO ($)
	Money Raised at IPO	Amount the Organization raised at IPO ($)
	IPO Date	The date when the Organization went public (MM/DD/YYYY)
JIFX Shared Drive	# of JIFX's	The total number of JIFX events attended by the Company.
LinkedIn	Industry	The industry category self-selected by the company (over 200 options available).[1]
	Company Size (2025)	Categorized company size: 1, 2-10, 11-50, 51-200, 201-500, 501-1,000, 1,001-5,000, 5,001-10,000, 10,001+
	Founded	Year of company founding, self-selected by the company page creator.
	Specialties	A company can select up to 20 specialties for their profile.[1]
	Headquarters Location (Company Website)	The primary location of the company’s headquarters, as listed on its website. (City, State, County)

Figure 4. JIFX Database Categories

By exploring these dimensions, this paper aims to provide a comprehensive understanding of the companies that have engaged with JIFX, their characteristics, and their growth trajectories. This analysis will offer valuable insights into the program’s impact on fostering technological innovation and highlight potential opportunities for enhancing its effectiveness in the future.

Company Demographics

Understanding the demographics of companies that participate in JIFX is essential to evaluating the program’s broader impact on innovation and small business success. By examining who participates—whether startups or established firms, private sector or government-affiliated, and which industries and regions they represent—we gain insight into the types of organizations that JIFX attracts and supports. This information helps identify patterns in participation that may correlate with successful outcomes, such as commercialization,

[1] LinkedIn does not publish a complete list of industry or specialty types.

investment, or acquisition. In particular, a strong presence of small and medium-sized enterprises (SMEs) can signal that JIFX provides a uniquely accessible and valuable platform for early-stage companies seeking to test, iterate, and demonstrate their technologies in realistic field conditions.

Figure 5 illustrates the organizational diversity of companies participating in JIFX, showing that a vast majority—87%—are privately held firms, while government, academic, and other entities make up the remaining 13%. This breakdown reinforces JIFX's role as a venue primarily serving innovation-driven private companies while maintaining an inclusive environment for public and academic collaboration. The presence of government and FFRDC participants highlights opportunities for dual-use technology exploration and transition pathways between commercial and public-sector applications. This diversity of participants sets the stage for analyzing how JIFX fosters technological innovation across a wide ecosystem of actors.

The data presented in Figure 6 through Figure 9 offer a comprehensive snapshot of the organizational, geographic, structural, and industrial diversity of past JIFX participants, illustrating the program's unique position as a catalyst for small business innovation. A dominant presence of private-sector entities, particularly small and medium-sized enterprises, underscores JIFX's appeal to agile, innovation-driven companies seeking opportunities to test and refine technologies in an applied field environment. Although large corporations and government-affiliated entities are also represented, the strong SME turnout signals that JIFX plays a vital role in supporting early-stage growth and market access. The geographic concentration of participants in U.S. innovation hubs—especially near defense and research institutions—suggests that proximity to government and military stakeholders enhances opportunities for collaboration and transition. Industry-wise, the program attracts companies working at the forefront of technological development, particularly in aerospace, cybersecurity, and AI, reflecting both JIFX's emphasis on dual-use technologies and the evolving needs of national security. Taken together, these trends provide essential context for evaluating JIFX's impact on small business success.

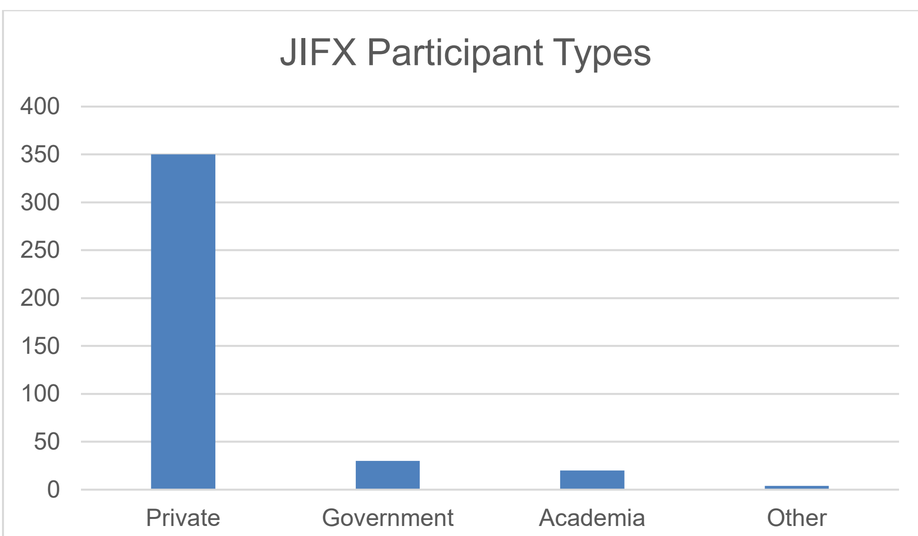


Figure 5. JIFX Participant Types.

(Note: The companies participating in JIFX range from early-stage startups to well-established corporations. The dataset includes information on company type (e.g., private, public, government-affiliated), providing insights into the diversity of organizational types engaged in the program. The data shows that most participants represent private industry (87%), with much smaller representation from government agencies (7%) and academia (5%). Representing the other category, which accounts for 1% of experiments, are Federally Funded Research and Development Centers (FFRDCs) and nonprofits.)

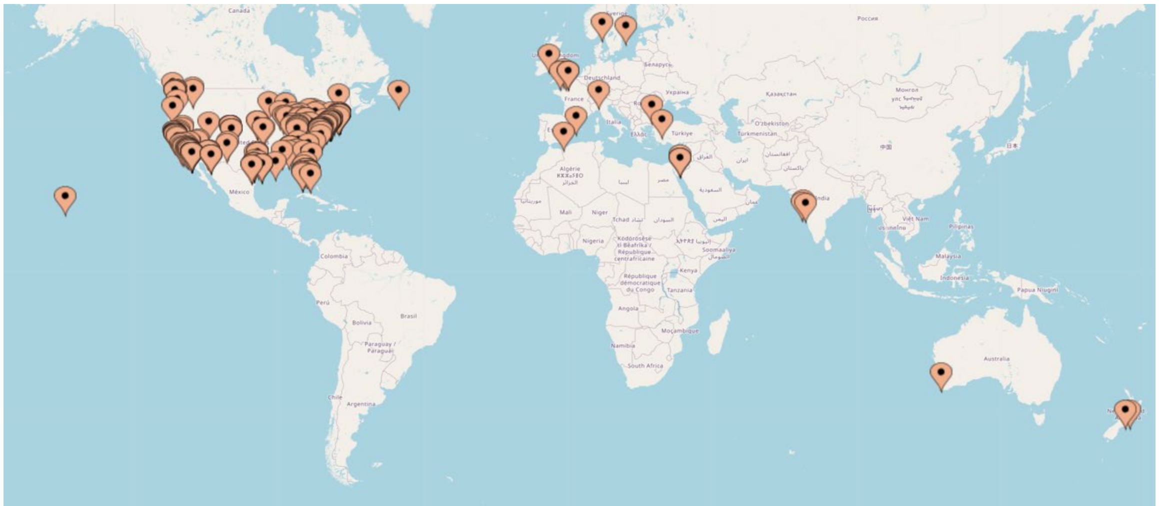


Figure 6. JIFX Participant Headquarters Location.

(Note: This map visualizes the headquarters locations of participating companies. The data indicates that a majority of participants are based in the United States, with notable clusters in California, Texas, and Virginia. International participation is observed but remains limited. The concentration of companies in key regions suggests that proximity to defense and government research institutions plays a role in fostering innovation and engagement with JIFX.)

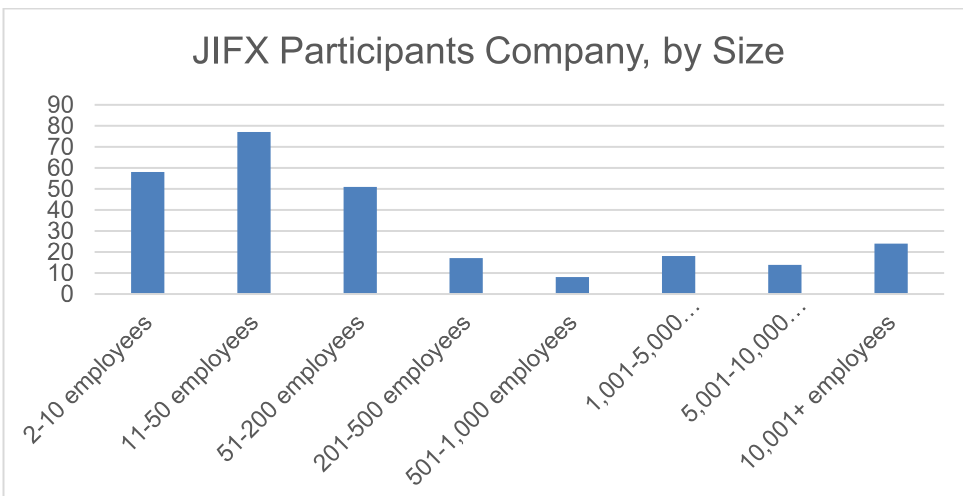


Figure 7. JIFX Participants Company, by Size.

(Note: This graph presents the breakdown of company sizes, ranging from small businesses with fewer than 10 employees to large corporations with over 10,000 employees. The analysis reveals that a significant proportion of JIFX participants are small and medium-sized enterprises (SMEs), indicating strong engagement from emerging innovators. Large corporations also participate, reflecting JIFX's ability to attract a broad spectrum of companies, from nimble startups to industry leaders.)

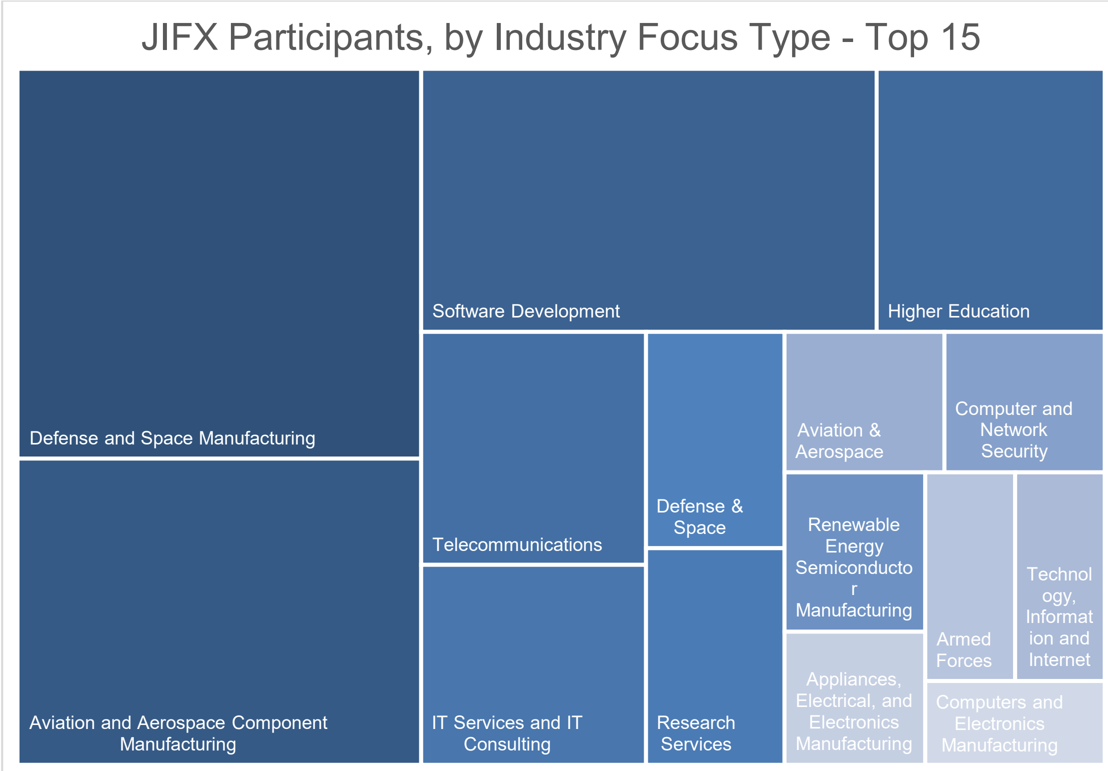


Figure 8. JIFX Participation by Industry Type.

(Note: The industries represented in JIFX participation are diverse, spanning aerospace, cybersecurity, artificial intelligence, and more. Analyzing these sectors helps to identify which technological fields are most actively involved in JIFX and which areas show the highest potential for innovation. This figure showcases the distribution of the top 15 industries among JIFX participants. The aerospace and defense sectors hold the largest share, followed closely by cybersecurity and AI-driven companies. This distribution highlights the program's focus on cutting-edge technological advancements. The presence of a strong cybersecurity segment suggests the increasing relevance of secure communications and AI-driven analytics in national security applications.)

Company Financials

Understanding acquisitions among JIFX participants provides valuable insight into industry consolidation and the commercial success of emerging technologies showcased at the event. Mergers and acquisitions (M&A) activity highlights the market relevance and long-term viability of these companies as they grow and integrate into larger corporate structures. Notable examples include Splunk's acquisition by Cisco for $28 billion and Inmarsat's acquisition by Viasat for $7.3 billion, both of which underscore the high value placed on innovations demonstrated through JIFX. Figure 9 presents a timeline of years between JIFX participation and subsequent acquisition, with most companies acquired after their involvement—sometimes many years later, as seen with Inmarsat and SystromDirect. These trends suggest that JIFX serves as a meaningful launchpad, increasing visibility and credibility among government and commercial stakeholders. By offering an environment to test and demonstrate capabilities in realistic settings, JIFX can catalyze strategic partnerships, investment, and eventual acquisition. When compared with the entire participant dataset, 3.6% of companies in the JIFX ecosystem have been acquired following their participation—demonstrating a tangible indicator of the program's role in accelerating commercial success.

Table 1. This table lists companies that participated in the JIFX program and were subsequently acquired. It includes each company's industry, acquiring entity, dates of JIFX participation and acquisition, and acquisition price when available on Crunchbase.

Company Name	Industry	Acquired By	JIFX Participation	Date of Acquisition	Price
Saab	Aerospace, Automotive, Infotech	Quantal	14-3, 20-4	March 3, 2021	
Ghost Robotics	Industrial, Military, Mining	Lig Nex1	21-1	July 28, 2024	$240M
Perforce Software	Analytics, DevOps, Infotech	Clearlake Capital Group	16-3, 16-4	January 10, 2018	
Mashable	Content, Digital Media, Infotech	Ziff Davis	15-4	Dec. 5, 2017	$50M
Orion Labs	Artificial Intelligence, Communications	Vontas	15-3, 16-2, 17-3, 18-3, 20-1, 21-4	Dec. 11, 2023	
Inmarsat	Communications Infrastructure	Viasat	15-3	May 9, 2023	$7.3B
Splunk	Analytics, Infotech	Cisco	21-4	Sept. 21, 2023	$28B
Tomahawk Robotics	Military, Product Design, Robotics	AeroVironment	22-2	August 22, 2023	$120M
Planck Aerosystems	Aerospace, Artificial Intelligence	AeroVironment	19-2, 22-2	August 17, 2022	
Haystax Technology	Analytics, Cyber Security, Public Safety	Fishtech Labs	14-2	May 21, 2018	
Javelin Networks	Computer, Cyber Security, Infotech	NortonLifeLock	17-1	Nov 5, 2018	
Asymmetric Technologies	Electronics, Infotech	Chesapeake Technology International	14-2, 23-3, 24-3	Mar 28, 2024	
Aurora Flight Sciences	Aerospace, Air Transportation	The Boeing Company	21-2	Oct 5, 2017	
Satcom Direct	Aerospace, Air Transportation	Gogo	14-2	Sept. 30, 2024	$375M
Tableau	Analytics, Big Data, Consulting	Salesforce	14-2	Jun 10, 2019	$15.7B
Leica Geosystems	CRM, Infotech	ABTECH Services Polytechniques Inc.	20-1	Feb 3, 2020	
HDT Global	Commercial, Government, Industrial	Behrman Capital	14-2	Jan 7, 2004	
Prox Dynamics	Aerospace, Air Transportation	Teledyne FLIR	16-1	Nov 30, 2016	$134M
Star Solutions	Infrastructure	OpenGate Capital	21-2	June 10, 2008	
Intelligent Automation	Innovation Management	BlueHalo	15-2	Aug 19, 2021	
OnLive	Cloud Computing, Gaming	Sony	15-1	Apr 4, 2015	
QRC Technologies	Manufacturing, Telecommunications	Parsons Corporation	22-2	Jul 22, 2019	$215M
Arcturus UAV	Aerospace, Manufacturing	AeroVironment	18-3	Jan 13, 2021	$405M
Protonex Technology Corporation	Electronics, Energy, Manufacturing	Ballard Power Systems	17-3	Jun 29, 2015	$30M

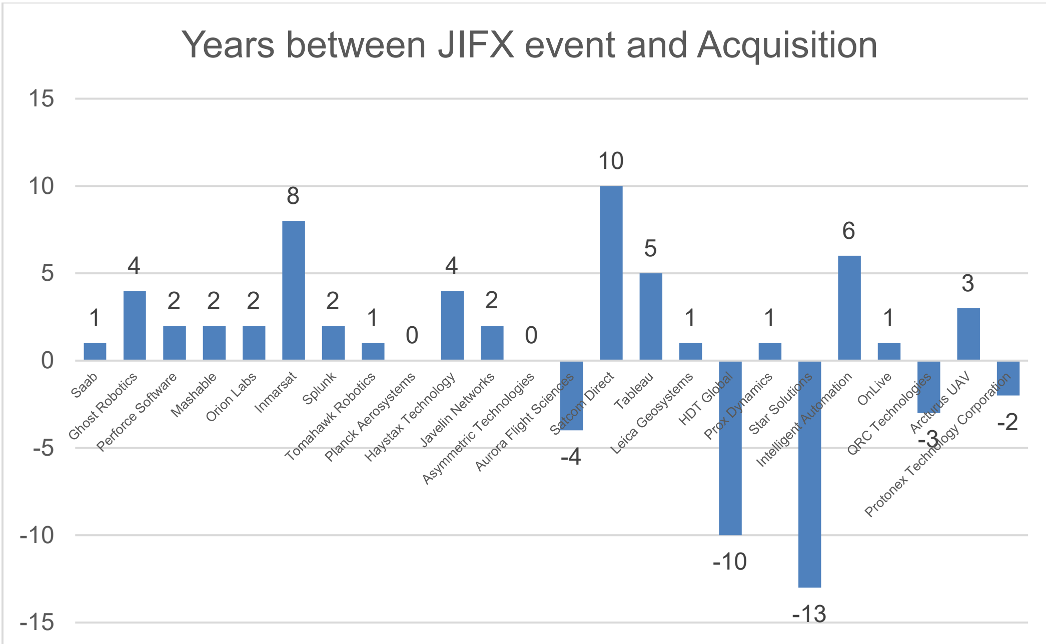


Figure 9. Time Between JIFX Participation and Acquisition for JIFX Participants.

(Note: Positive values indicate years after the JIFX event that a company was acquired, while negative values indicate acquisition occurred before the JIFX event. This timeline helps illustrate the relationship between field experimentation exposure and subsequent commercial or strategic acquisition.)

Repeat JIFX Participation

Repeat participation in JIFX is a strong indicator of the program's value to small and emerging technology companies. Many companies have returned to JIFX multiple times, with some attending as many as 14 events. To explore the motivations behind this continued engagement, interviews were conducted with representatives from seven companies that have participated in multiple JIFX events. These firms, ranging in size and sector, consistently cited the opportunity to test technologies in a realistic field environment as a key driver. The ability to conduct operationally relevant experiments—particularly for UAVs, sensors, and autonomous systems—was viewed as a major benefit, especially when compared to more traditional industry events.

Company	JIFX Participation	Motivation to Attend JIFX	How They Found Out About JIFX
Airrow	22-3, 23-2, 23-4	Wanted to explore DoD opportunities	Recommended by Roman Aerospace (another JIFX Participant)
Bluespace.ai	23-4, 24-1, 24-2, 24-3, 24-4	Seeking a military-recognized test environment for dual-use development and stakeholder engagement.	Through a contact at NAMC (National Advanced Mobility Consortium).
Gantz-Mountain	20-4, 21-3, 21-4, 22-1, 22-2, 22-3, 22-4, 23-1, 23-2, 23-3, 23-4, 24-1, 24-2, 24-4	Interested in stakeholder feedback and stakeholder engagement	Founders participated in JIFX while prior employees of NPS
Microwave Monolithics	15-3, 15-4, 19-2, 19-4, 21-1	Seeking a test environment to fly their unit on UAVs	Possibly saw something online or received outreach
Odys Aviation	23-4, 24-1, 24-3, 24-4, 24-4	Interested in stakeholder feedback and stakeholder engagement	Recommended by business consultant
Premise Data	23-2, 23-3, 23-4	Diversity of government stakeholders and networking with other companies	Likely through Vulcan SOF

Yotta Navigation	21-3, 21-4, 22-2, 22-3, 23-1, 23-2, 23-3, 23-4	Needed a location to test GPS tracking for mass fatality events, but cellular network limitations prevented attendance.	Discovered through online searches for government test events.

Figure 10. JIFX Interviews With Repeat Participants

Participants also praised JIFX's accessible application process, which lowers barriers to entry for startups unfamiliar with DoD procedures. Many discovered JIFX through word of mouth, recommendations from other participants, or defense-focused networks like NAMC or Vulcan SOF. While JIFX is primarily designed for experimentation, interviewees found significant value in the informal business development opportunities it offered. They highlighted the importance of interacting with government stakeholders and other innovators in a setting that allowed for deeper technical engagement than typical conferences. In many cases, JIFX opened doors to further funding opportunities and participation in high-visibility DoD demonstrations such as Project Convergence and USSOCOM TE.

Another prominent theme was the iterative nature of JIFX participation. Companies returned to refine their technologies across different environmental conditions and use cases. The open and collaborative atmosphere at Camp Roberts enabled spontaneous testing, peer-to-peer learning, and even unexpected collaborations. Participants valued the chance to observe and integrate with other systems on-site, which accelerated development cycles and revealed new operational applications for their technologies. For many, JIFX served as an entry point into the broader defense innovation ecosystem, providing early-stage validation before progressing to larger procurement or testing venues.

Despite its strengths, the interviews also identified areas for improvement. Notably, none of the participants mentioned the Joint Vulnerability Assessment Branch (JVAB), which suggests that either its role is not well understood or its integration into the experimentation process is lacking. Participants recommended increasing marketing outreach to startups, enabling structured collaboration with government representatives, and enhancing feedback mechanisms. They also suggested expanding the diversity of attending stakeholders and creating more direct links between JIFX and follow-on opportunities. These insights reinforce JIFX's critical role in accelerating defense-relevant innovation and highlight ways it can further support small business success.

Conclusion

The JIFX program represents a vital, though often underrecognized, mechanism for catalyzing small business success within the DoD innovation ecosystem. Through a uniquely accessible and collaborative structure, JIFX lowers the barriers that often prevent emerging companies from entering the defense market. Its open experimentation environment provides a rare opportunity for early-stage technologies to be tested, iterated, and validated under operationally relevant conditions, while simultaneously connecting small businesses to government stakeholders, potential partners, and a broader network of innovation champions. The findings of this research, derived from both quantitative data and qualitative interviews, affirm that JIFX plays an indispensable role in helping small businesses navigate the complex path from concept to capability.

Participation in JIFX has led to meaningful outcomes for companies, including iterative product development, connections to future funding opportunities, and increased visibility within the DoD. While direct contracting may not always be immediate, the cumulative impact of networking, real-time feedback, and government engagement helps position companies for downstream success. Repeat participants, in particular, view JIFX as a cornerstone of their defense innovation strategies. For many, it serves as a launchpad into larger demonstrations and funding mechanisms such as SBIR, Project Convergence, and SOCOM Technical Experimentation. These longitudinal benefits highlight JIFX's importance not only as a venue for experimentation but as a critical on-ramp to the broader acquisition and innovation ecosystem.

At the same time, this research identified several opportunities to further enhance JIFX's effectiveness and strategic alignment. Greater awareness of cybersecurity requirements and integration of the Joint Vulnerability Assessment Branch (JVAB) could strengthen the program's alignment with modern defense priorities. More deliberate pathways from JIFX to acquisition programs and structured co-development with government users would also amplify its impact. In particular, expanding outreach to underrepresented startups and venture-backed firms would diversify the portfolio of participating technologies and maximize innovation potential. Clearer mechanisms for data sharing, post-event feedback, and metrics-driven outcomes would further support participating businesses in demonstrating their progress and value.

In summary, JIFX exemplifies how a thoughtfully designed, field-based experimentation program can unlock innovation from the private sector and channel it toward national security objectives. By providing a hands-on, low-risk environment for testing and collaboration, JIFX bridges the persistent gap between emerging technologies and end-user needs. Its continued evolution—through partnerships, increased visibility, and deeper integration with DoD transition pathways—will be essential in sustaining the pace of innovation and maintaining U.S. technological advantage. For small businesses, JIFX offers not just a proving ground but a launchpad—transforming potential into progress and ideas into impact.

References

America's seed fund: SBIR and STTR. (n.d.).

Bordetsky, A., & Netzer, D. (2009). TNT testbed for self-organizing tactical networking and collaboration. In *Proceedings of the 14th International Command and Control Research and Technology Symposium* (p. 63).

Chestnut, J. (2024, November 6). *Innovative Algorithms* [In person].

Crunchbase. (n.d.). https://www.crunchbase.com

Dagan, S. (2024, November 8). *Anello Photonics* [Microsoft Teams].

Defense Analysis (DA) curricula. (n.d.). https://nps.edu/web/da/curricula

The Defense Innovation Unit FY 2023 annual report. (n.d.). https://downloads.ctfassets.net/3nanhbfkr0pc/57VfnQbajgWdONRicxv6nG/5987e9a02c4d24f16b6d5f356c55fa63/DIU_Annual_Report_FY2023.pdf

Furey, R. (2025, February 6). *Premise Data* [Microsoft Teams].

Goshorn, D. & Goshorn, R. (2025, February 13). *Gantz-Mountain Intelligence Automation Systems* [Microsoft Teams].

Hazlett, A. (2025, February 5). *Yotta Navigation* [Microsoft Teams].

Kaye, D. (2025, February 6). *Airrow* [Microsoft Teams].

Kotila, B., Drezner, J. A., Bartels, E. M., Hill, D., Hodgson, Q. E., Huilgol, S. S., Manuel, S., & Simpson, M. (n.d.). *Strengthening the defense innovation ecosystem*. RAND.

JIFX FY24 request for information. (n.d.). https://nps.edu/documents/104517539/104522581/JIFX+FY24+RFI.pdf/df2c4fa5-1e5d-8c58-3fbd-63bb7b99256b?t=1692735046133

Joint Interagency Field Experimentation. (n.d.). https://www.nps.edu/web/fx/

LinkedIn. (n.d.). https://www.linkedin.com

Livingston, A. (2025, February 14). *Thermo AI* [Microsoft Teams].

Manuel, C. E., Murphy, H. R., Jr., & Paxton, K. A. (2004). *The Surveillance and Target Acquisition Network (STAN)* [Master's thesis]. Naval Postgraduate School.

Mena-Werth, R. (n.d.). *Empowering student ideas: NPS introduces the Naval Innovation Exchange*. Naval Postgraduate School. https://nps.edu/-/empowering-student-ideas-nps-introduces-the-naval-innovation-exchange

Moffitt, A. (2024, November 12). *Clara CoPilot* [Microsoft Teams].

Naval Innovation Center. (n.d.). https://nps.edu/web/naval-innovation-center

Naval Postgraduate School academic catalog. (2022). https://nps.edu/documents/104111578/117929723/Fall+Quater_NPS-Academic-Catalog_Sept+2022.pdf/5d801f39-d6b3-8bdd-42f6-6cb4574b8b9e?t=1664562884492

Naval Postgraduate School 2023 annual report & mission measures. (n.d.). https://nps.edu/documents/10180/149156382/NPS+Annual+Report+FY2023-WEB-single%5B85%5D.pdf/e3e9009b-c980-1b63-82fb-3c41168d2463?t=1727998157146

NPS Joint Interagency Field Experiment. (n.d.). www.nps.edu/fx

Oros, C. L. (2014). A brief history of the NPS field experimentation program: Spanning STAN, TNT, and JIFX.

Peterson, W. (2025, February 7). *Microwave Monolithics* [Microsoft Teams].

Pipenberg, B. (2024, November 7). *Knightwerx* [Microsoft Teams].

Radermacher, A. (2025, February 14). *Odys Aviation* [Google Meet].

Templeton, J. (2025, February 10). *Bluespace.ai* [Microsoft Teams].

Vision & strategic framework. (2025). Naval Postgraduate School.

Warfare innovation workshops. (n.d.). Naval Warfare Studies Institute. https://nps.edu/web/nwsi/warfare-innovation-workshops

Warmoth, A. (2025, February 14). *Chariot Defense* [Microsoft Teams].

Wells, B. (2025, February 19). *uAvionix* [Microsoft Teams].

Youmans, T. (2024, November 7). *Rhoman Aerospace* [Microsoft Teams].

PANEL 8. STRATEGIC INNOVATIONS IN DEFENSE CONTRACTING: AI, TECHNOLOGY SUPPLIERS, AND COMPETENCY

Wednesday, May 7, 2025	
1115 – 1230 PT 1315 – 1430 CT 1415 – 1530 ET	**Chair: John G. (Jerry) McGinn, Ph.D. Executive Director Greg and Camille Baroni Center for Government Contracting George Mason University** ***Leveraging Generative AI for Validating the Quality of DoD Acquisition Packages and Contract Documents*** Samantha Nangia, Senior eBusiness Analyst, Deputy Assistant Secretary of the Navy (Procurement) ***High-Technology and Cosmopolitan Companies in the Federal Supply Chain*** Edward Hyatt, Senior Research Fellow, George Mason University ***Contract Management Competency in the United States Marine Corps: An Assessment of MCSC, ECPs, and LOGCOM*** Rene G. Rendon, Associate Professor, Naval Postgraduate School

John G. (Jerry) McGinn, Ph.D.—is the Executive Director of the Greg and Camille Baroni Center for Government Contracting in George Mason University's Costello College of Business. In this role, he has established and is leading the first-of-its-kind university center to address the business, policy, and regulatory issues facing the $700B+ government contracting community. He is one of the nation's leading experts on industrial base and acquisition policy issues and the Center has published over 80 influential reports and white papers on issues such as defense innovation, government contracting, intellectual property, budget reform, industrial resilience, collaboration with allies and partners, and COVID-19.

Jerry is also a trusted strategic advisor and board member sought after for his expertise in U.S. industrial policy, security cooperation, supply chain, industrial security, export control, foreign military sales, and industrial base policies. Prior to joining GMU, Jerry served as the senior career official in the Office of Manufacturing and Industrial Base Policy in the Department of Defense, leading efforts to analyze the capabilities and overall health of the defense industrial base, including the 2017-2018 interagency review of the manufacturing and defense industrial base. He also directed hundreds of reviews of high-profile mergers and acquisitions as well as transactions before the Committee on Foreign Investment in the United States. He additionally led eight manufacturing innovation institutes focused on building capacity in advanced manufacturing areas such as 3D printing, microelectronics, and robotics.

Previous to DoD, Dr. McGinn spent a decade in senior defense industry roles at Deloitte Consulting LLP, QinetiQ North America, and Northrop Grumman. Before industry, Dr. McGinn served in DoD as Special Assistant to the Principal Deputy Undersecretary (Policy) and as a political scientist at RAND.

Dr. McGinn is also a widely acclaimed thought leader and sought-after speaker. He has published influential George Mason reports and white papers, RAND monographs, and articles in The Hill, Business Insider, Defense News, Defense One, and other outlets. He has also testified before the U.S. Congress and the UK House of Commons and he has participated in advisory studies for the Homeland Security Advisory Council, the NATO Industrial Advisory Group, and the Defense Science Board.

Dr. McGinn was commissioned into the U.S. Army and served with distinction as an infantry officer and is a graduate of Ranger and Airborne Schools. He has received numerous civilian and military awards and has earned a Ph.D., M.S., and M.A. from Georgetown University as well as a B.S. from the United States Military Academy.

Leveraging Generative AI for Validating the Quality of DoD Acquisition Packages and Contract Documents

Samantha Nangia—is a Senior Program Analyst and the Enterprise Procurement System (ePS) Solution Manager for the Department of the Navy, overseeing the transformation of all Navy and Marine Corps procurement capabilities. She leads efforts to enhance procurement data quality, visibility, and integration, supporting enterprise analytics and auditability objectives. As the Lead Product Owner for the Navy's Electronic Procurement System (ePS), she ensures the delivery of key capabilities to the 1102 community. Nangia specializes in Procure-to-Pay (P2P) systems, data architecture, and automation, driving the integration of procurement and financial data. She has successfully automated the Procurement Performance Management Assessment Program (PPMAP) and developed solutions to improve procurement data quality across the Department of the Navy. She holds a bachelor's in computer science from Rollins College and an MBA from Georgetown University. Nangia's expertise in procurement systems and data solutions has contributed to improved efficiency and performance across the DON.

Tom Wardwell—is the Deputy Director of eBusiness Policy and Oversight serving as the Acquisition IT Portfolio Manager and eBusiness Transformation lead for the Department of the Navy. In this role, he oversees the strategic integration and management of a $200 million/year IT acquisition business systems portfolio, ensuring sound investment in capabilities supporting the acquisition mission and alignment with Navy-wide strategic objectives. He is critical in driving innovation, enhancing procurement efficiency, and ensuring compliance with evolving regulatory frameworks. He holds a bachelor's in physical science from the U.S. Naval Academy and master's degrees in acquisition management from the Naval Postgraduate School, information systems technology from The George Washington University, and national strategic resourcing from the Industrial College of the Armed Forces.

Randall Mora—is President and CEO of Avum, Inc., providing Avum's strategic and management direction and actively participates in multiple U.S. Navy, U.S. Army, U.S. Air Force, and the Fourth Estate (Department of Defense) Procure-To-Pay initiatives. His passion is leading Avum's research and development efforts to rapidly build solutions that Avum's customers can leverage for their success.

Carlos Parada Jr.—is a data scientist and statistician working with Avum, Inc. to provide high-quality products using the latest insights in machine learning and AI. Parada completed his degree in mathematics at Carleton College. He started in Bayesian inference and econometrics as a research developer with the Cambridge Machine Learning Group on the Turing.jl project. Since then, Parada has worked on medical trials testing cancer drugs and mass healthcare interventions in Cameroon. Parada's free time is often occupied by exploring emerging research within machine learning, artificial intelligence, and mathematics, coupled with his work on Wikipedia's mathematics articles.

Abstract

The Department of Defense (DoD) contracting process requires rigorous validation to ensure regulatory compliance, accuracy, and completeness. This paper explores the integration of NIPR GPT, a secure generative artificial intelligence (AI) model, to enhance the efficiency and reliability of the Acquisition and Contracting package validation. Deployed in a DoD-approved environment, NIPR GPT is a Government R&D Platform for GenAI models and applications serving as a comprehensive AI research and development platform featuring retrieval augmented generation. NIPRGPT enables model evaluation, shared workspaces, and secure document processing workflows, in our use case, we used it to automate key tasks such as compliance checks against FAR/DFARS/NMCARS/Local Policy Language, clause verification, contract risk identification, and data consistency validation. The proposed framework enables contracting officers to upload documents, select validation tasks, and receive detailed, actionable reports. NIPR GPT is able to leverage fine-tuned training on DoD-specific datasets to identify missing clauses, resolve ambiguities, and flag high-risk elements. By automating labor-intensive tasks, the system is able to reduce human error, accelerate processing, and ensure compliance with regulatory and policy requirements. The model is implemented within an IL-4 environment to address security concerns, with robust encryption protocols and access controls to safeguard sensitive data. Audit logging provides transparency, ensuring outputs can be reviewed and verified. A case study using a significant Aircraft procurement demonstrates the practical application of this framework. NIPR GPT identified missing compliance language and clauses, flagged ambiguous deliverable descriptions, and recommended corrective actions, streamlining the package approval process. This integration of AI into DoD workflows illustrates its potential to modernize procurement practices, improve accuracy, and maintain compliance in a highly regulated environment. This abstract highlights the transformative role of generative AI in supporting DoD contracting officers by providing reliable, secure, and efficient tools for package validation.

Keywords: Defense acquisition, Generative AI, LLM, NIPR GPT, FAR compliance, DFARS compliance, Contract validation, NMCARS, Procurement Performance Management Assessment Program, PPMAP, Regulatory automation, Compliance assurance, AI accuracy, Clause verification, Contracting workflows, FAR/DFARS interpretation, Human-in-the-loop AI, Contract analysis, Procurement automation, DoD contracting packages, DOD acquisition packages, Compliance errors, Automated legal review, Fine-tuned AI models, Regulatory compliance systems, AI benchmarking, Retrieval-augmented generation, RAG

Introduction

Background

The Department of Defense (DoD) contracting process is governed by a complex framework of regulations, including the Federal Acquisition Regulation (FAR), the Defense Federal Acquisition Regulation Supplement (DFARS), and the Navy and Marine Corps Acquisition Regulation Supplement (NMCARS). These regulations ensure that all contracts meet strict standards for compliance, accuracy, and completeness. For example, under DFARS 252.204-7012, contractors must implement cybersecurity measures to protect controlled unclassified information. Similarly, FAR Part 15 outlines detailed procedures for contract negotiations, ensuring fairness and transparency in source selection. NMCARS supplements these regulations by providing specific guidance for Navy and Marine Corps acquisitions, such as stricter validation of cost estimates and contract requirements. Rigorous validation processes, including proposal audits, compliance reviews, and independent cost estimates, help mitigate risks and ensure compliance with regulatory requirements while maintaining procurement integrity.

Problem Statement

Due to the complexity and evolving nature of federal acquisition regulations, contracting officers must navigate an intricate compliance landscape that includes the FAR, the DFARS, the NMCARS, and Department of the Navy (DON) policy directives. Each of these frameworks imposes stringent requirements on procurement processes, ranging from cost estimation and cybersecurity compliance to small business set-asides and contract auditing. The frequent updates and nuanced interpretations of these regulations add another layer of difficulty, increasing the risk of non-compliance, bid protests, and potential contract delays.

In response to these challenges, emerging generative AI models like NIPR GPT promise to automate regulatory analysis, reduce administrative burdens, and improve contract review efficiency. However, despite their potential, these models have not undergone rigorous validation to ensure their accuracy, reliability, and ability to identify critical compliance risks effectively. Errors in AI-generated recommendations could lead to overlooked compliance issues, misinterpretations of regulatory language, or unintended contract violations, ultimately jeopardizing mission readiness and procurement integrity. As such, there is a critical need to assess the efficacy of AI-driven solutions in real-world DoD contracting environments to determine their feasibility, limitations, and potential role in enhancing regulatory compliance.

Research Question/Objective

1. How accurately can NIPR GPT detect missing or misapplied language in DoD acquisition/contracting packages compared to traditional manual review processes?
2. What are AI-generated outputs' most common errors and limitations when validating regulatory and policy compliance?
3. To what extent does domain-specific fine-tuning improve the accuracy and reliability of NIPR GPT in assessing compliance with FAR, DFARS, NMCARS, and DON policy directives?
4. How can retrieval-augmented generation (RAG) improve the accuracy and completeness of AI-driven compliance validation in DoD acquisition/contracting packages?

This research aims to evaluate the effectiveness of AI-driven solutions, particularly NIPR GPT and RAG, in enhancing compliance validation for DoD contracting. The study focuses on assessing accuracy, identifying limitations, measuring the impact of fine-tuning, and exploring the potential benefits of integrating AI into existing procurement workflows. The key objectives are as follows:

1. **Evaluate AI Accuracy**: Assess NIPR GPT's effectiveness in identifying compliance issues compared to manual review.
2. **Identify AI Limitations**: Analyze common errors and gaps in AI-generated compliance assessments.
3. **Measure Fine-Tuning Impact**: Determine how domain-specific fine-tuning improves AI performance in regulatory compliance tasks.
4. **Compare Review Efficiency**: Investigate whether AI-assisted reviews can reduce the time and effort required for compliance validation while maintaining accuracy.
5. **Assess RAG Effectiveness**: Evaluate how RAG enhances AI-generated outputs by integrating real-time regulatory references.
6. **Analyze Error Reduction**: Examine whether RAG reduces common AI errors, such as hallucinations, misinterpretations, or outdated regulatory references.
7. **Optimize AI Integration**: Identify best practices for implementing AI-driven compliance validation in DoD procurement workflows.

Scope and Limitations

Scope: This research examines the potential of AI-driven compliance validation in DoD procurement, focusing on NIPR GPT and RAG to enhance accuracy, efficiency, and regulatory adherence. The study aims to:

1. Assess the accuracy of AI in detecting missing or misapplied regulatory language compared to traditional manual review processes.
2. Identify common errors and limitations in AI-generated compliance assessments.
3. Evaluate the impact of domain-specific fine-tuning on the AI model's ability to interpret and apply FAR, DFARS, NMCARS, and DON policy directives.
4. Investigate how RAG improves the accuracy and completeness of AI-generated outputs by integrating external regulatory sources.

5. Explore the feasibility of implementing AI-assisted compliance validation in real-world DoD procurement workflows.

The research will utilize a dataset of DoD acquisition packages and contract documents to test AI performance. It will also include qualitative insights from Acquisition professionals to understand AI's practical applications and limitations.

Limitations: While this study provides valuable insights into AI-driven compliance validation, certain limitations exist:

1. *Data Availability*: This research relies on contract data accessible within the DON enclaves, which are not publicly available. While this controlled environment ensures data security and regulatory compliance, it may limit the diversity of acquisition packages used to evaluate AI performance. As a result, findings may not fully account for the variability in contract structures across different DoD agencies or broader procurement scenarios.
2. *Regulatory Updates*: AI models may not instantly adapt to evolving regulatory changes, impacting the accuracy of compliance validation over time.
3. *AI Interpretability*: NIPR GPT's decision-making process may lack transparency, making it challenging to fully understand how compliance determinations are made.
4. *Scope of Fine-Tuning*: The study focuses on domain-specific fine-tuning but does not explore real-time learning or continuous retraining of AI models.
5. *Human Oversight*: AI is not intended to replace human acquisition professionals but to assist them. The research does not propose a fully automated compliance validation system but rather a hybrid approach where AI enhances manual reviews.
6. *Comparative Baseline*: Manual review processes vary across contracting offices, which may introduce inconsistencies when comparing AI performance against human evaluations.

By defining these scope and limitations, this research ensures a focused and realistic assessment of AI's potential in DoD contracting compliance validation.

Methodology

This research employs a multi-faceted methodology to evaluate AI-driven compliance validation in DoD contracting. By leveraging structured compliance checks, RAG, and expert feedback, we aim to assess the accuracy, reliability, and practical applicability of AI models in regulatory reviews. The key components of our methodology include:

1. **Structural Compliance Checks:**
    - *Annex 18 ISTRAPS, Annex 19 PSTRAP-M, and Annex 20 ISTRAP-M:* Evaluate structural compliance of acquisition packages to ensure adherence to regulatory frameworks.
    - *Annex 1 Review of Justifications & Approvals (J&As)*: Assesses the logic, strength of argument, and flow of J&As. AI-generated prompts guide reviewers in refining responses, ultimately producing revised J&As.

2. **Contract Award Document Analysis:**
    - Compares final contract award documents against FAR, DFARS, and NMCARS to identify structural compliance gaps, risks, clause compliance analysis, and regulatory adherence.
    - Utilizes PDS XML data to standardize and enhance acquisition package analysis.
3. **RAG Model for Compliance Checks:**
    - Uses "clean" acquisition packages—deemed high-quality by the team—as reference data to improve AI-generated compliance recommendations.
    - Assess the impact of RAG-enhanced AI outputs on the accuracy and completeness of contract reviews.
4. **Expert Review and Model Refinement:**
    - Incorporates DASN (P) and legal review comments to fine-tune AI-generated outputs through prompt engineering.
    - Evaluates the fidelity of AI-generated compliance assessments against expert feedback.
5. **Batch Summarization and Ranking:**
    - Processes multiple acquisition packages simultaneously to generate structured summaries.
    - Ranks packages based on quality and provides recommendations for strengthening weaker submissions.

This methodology ensures a rigorous and iterative approach to assessing AI's potential in DoD contracting, balancing automation with expert validation to enhance regulatory compliance and efficiency.

NMCARS Analysis

To assess structural compliance in DoD acquisition, we collaborated with the Program Analytics Business Transformation (PABT) team at DASN (P) to obtain contract review samples and results. This dataset includes both annotated assessments and high-quality "clean" contract versions, providing a foundation for AI model evaluation. The specific annexes reviewed include:

- Annex 1—Review of J&As: Assesses logic, argument strength, and flow; AI-generated prompts guide reviewers in refining responses, leading to revised J&As.
- Annex 18—Individual Streamlined Acquisition Plan (ISTRAP): Evaluates structural compliance in acquisition planning.
- Annex 19—Program Streamlined Acquisition Plan With Services (PSTRAP-M): Focuses on acquisition planning for service-based contracts.
- Annex 20—Individual Streamlined Acquisition Plan With Services (ISTRAP-M): Examines structural compliance for individual acquisition plans involving services.

These structured compliance checks provide a benchmark for AI-driven analysis, enabling comparison between AI-generated outputs and expert-validated contract reviews. The results inform subsequent model fine-tuning and RAG-based improvements for enhanced regulatory compliance validation.

This ensures clarity and emphasizes how the collaboration with the PABT team strengthens the research methodology.

RAG-Based Improvements for Enhanced Compliance Analysis

We leveraged RAG within the NIPR GPT large language model (LLM) to improve the accuracy and completeness of AI-driven compliance validation (Google Cloud, n.d.). By incorporating high-quality contract documents and structured compliance assessments into the retrieval system, we aim to enhance AI-generated compliance recommendations.

This process involves:

- **Feeding validated contract reviews and "clean" documents** into the RAG model as reference materials.
- **Enhancing AI outputs** by dynamically retrieving relevant regulatory information from FAR, DFARS, NMCARS, and DON policy directives.
- **Comparing AI-generated compliance assessments** before and after RAG integration to measure accuracy, completeness, and risk identification improvements.

As illustrated in Figure 1, RAG-based retrieval enables the AI model to access real-time reference data, reducing errors such as misinterpretations and hallucinations while improving the overall quality of compliance validation (Zvornicanin, 2024).

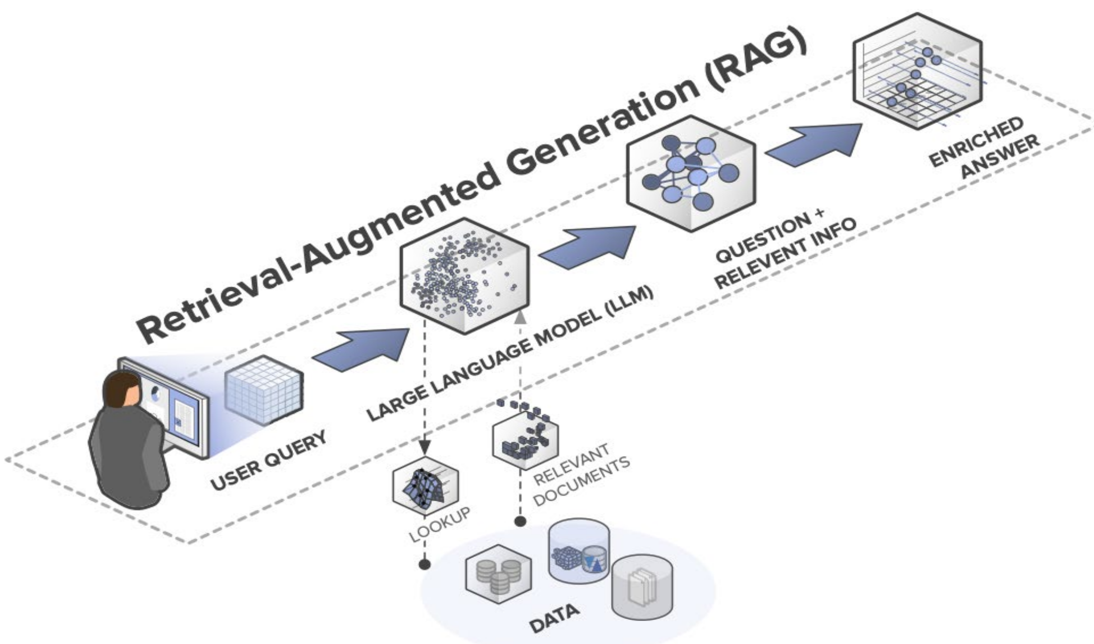


Figure 1. RAG-Enhanced Compliance Validation Workflow

FAR/DFARS/NMCARS Clause Compliance Analysis

This research utilizes the Procurement Data Standard (PDS) XML as the primary data source to systematically assess regulatory compliance, risk factors, and structural integrity in

DoD contracting (OUSD[A&S], n.d.-a). PDS is a standardized XML format that captures structured procurement information (OUSD[A&S], n.d.-b), enabling automated validation against **FAR, DFARS, NMCARS**, and DON policy directives. By leveraging PDS, we aim to:

- Standardize contract data ingestion for AI-driven compliance analysis.
- Identify structural inconsistencies, regulatory gaps, and missing clauses in contract award documents.
- Develop intelligence models that enhance AI's ability to detect compliance risks, misapplied, and missing clauses.
- Improve automation and accuracy in contract validation by integrating structured PDS data into AI and RAG models (Atamel, 2025).

Our goal is to ensure a consistent, scalable, and **data-driven approach** to evaluating compliance across diverse procurement scenarios by aligning contract award document analysis with PDS XML.

While PDS XML provides a standardized structure for analyzing contract award data, it has inherent limitations when used for comprehensive clause compliance validation. The XML primarily reflects explicitly listed clauses and those tagged for inclusion, but it does not account for clauses incorporated by reference in attachments or supplemental documents.

As such, during our analysis, clause presence or absence is inferred based on what is available in the XML structure. The system may flag potential omissions, but it is important to note:

> "Based on the provided XML and a general understanding of government contracting, certain clauses may appear to be missing. However, we cannot definitively determine omission without access to the full contract file, including all attachments and incorporated references. Some clauses may exist elsewhere in the contract documentation but are not surfaced in the PDS XML."

To mitigate this limitation, our methodology incorporates:

- Subject matter expert (SME) validation of flagged discrepancies,
- Pattern recognition from previously reviewed complete packages,
- And continuous refinement of the model prompts to account for common clause placement practices.

This ensures that while the system offers valuable insights into potential compliance issues, final validation still benefits from expert oversight and full contract context.

Key components of this approach include:

- **Comparative Clause Analysis**—Evaluating Clause Logic Service (CLS) recommendations against actual contract clauses to identify discrepancies, misapplications, and gaps.
- **Benchmarking "Clean" Contracts**—Developing a validated reference set of high-quality contract awards, serving as a compliance standard for future assessments.
- **AI-Enhanced Compliance Checks**—Integrating RAG within NIPR GPT to:

    - Analyze new contract awards against the clean contract benchmark.
    - Improve clause validation accuracy by refining AI-driven assessments (Zvornicanin, 2024).
- **Expert Review and Model Refinement**—Leveraging the expertise of experienced contracting officers, DASN (P) analysts, and legal professionals to:
    - Incorporate DASN (P) and legal review comments to fine-tune AI-generated outputs through prompt engineering.
    - Evaluate the fidelity of AI-generated compliance assessments against expert feedback to improve model accuracy and reliability (Atamel, 2025).

By combining automation with expert judgment, this methodology ensures efficient, accurate compliance validation while reinforcing the essential role of contracting officers and subject matter experts. The goal is to streamline regulatory adherence in DON contracting by reducing manual review efforts, improving compliance precision, and continuously refining AI-based assessments through SME-driven oversight.

Batch Summarization and Ranking

To enhance structural analysis and compliance validation across FAR, DFARS, and NMCARS, this research develops a Python-based process for batch summarization, ranking, and accuracy assessment of contract documents. A critical aspect of this methodology is detecting AI hallucinations, which occur when the model generates incorrect or misleading information. These errors can be identified using entropy, a measure of the model's uncertainty in its predictions. High-entropy responses indicate that the model is uncertain about the correct answer, signaling a higher likelihood of inaccuracies or hallucinations (Entropy [information theory], n.d.).

The methodology consists of the following key components:

1. **Structural Analysis at Scale**—Processing multiple contract documents to assess completeness, clause structure, and alignment with regulatory requirements.
2. **Entropy-Based Accuracy Evaluation**—Assigning entropy scores to AI-generated outputs, where:
    - Low entropy suggests a high-confidence prediction.
    - High entropy signals uncertainty and a potential AI hallucination.
3. **Flagging and Visualizing High-Entropy Responses**—High-entropy answers are flagged for additional review and highlighted in the interface, enabling SMEs to focus on ambiguous or unreliable AI outputs.
4. **Ranking and Prioritization**—Summarized contract analyses are ranked based on compliance confidence, directing attention to high-risk discrepancies for further evaluation.

To substantially improve AI model performance, we integrate advanced reasoning techniques such as:

- Prompt engineering—Combining human-generated instructions with AI enhancements to provide additional clarity and encourage the model to “think things through” (OpenAI, 2024).
- Chain-of-Thought (CoT; Gadesha & Kavlakoglu, 2024)⁻Requiring the model to provide structured, step-by-step logical reasoning (Founding Minds, 2024; Villani, 2024).
- Beam search—Replacing the traditional “greedy sampling” with a more-advanced beam search optimization algorithm (Leblond et al., 2021).
- Structured generation (OpenAI, 2024)⁻Enforcing a prespecified output formatting (given by a context-free grammar such as JSON) to provide a “scaffolding” and restrict the AI to generating valid outputs.

By combining structured automation (Shorten et al., 2024), entropy-based validation (Entropy [information theory, n.d.), and expert oversight, this approach enhances compliance accuracy, reduces AI hallucinations, and improves regulatory adherence.

Results and Analysis

The results of our acquisition document/ package structural review stem from the integration of annotated acquisition documents provided by the DASN (P) PABT team with our RAG model analysis in NIPR GPT. These annotated documents, sourced from previous acquisition package reviews, contained NMCARS sections mapped directly to contract content, serving as a ground truth dataset for evaluating and improving structural compliance validation.

The annotated acquisition documents were ingested into our RAG pipeline, enabling the model to retrieve contextually relevant regulatory references when analyzing contract clauses.

Key steps and observations are included below.

Basic Use Case—Annex 18—ISTRAP Structural Review

Doc 1: TAB A - NMCARS 18-25 Annex 18 - ISTRAP. This is the template for an ISTRAP and the rules to be followed. In this document, we provided the model for the NMCARS structure to analyze the remaining against.

Doc 2: ISTRAP CNO Avails Fast Attacks Sub. This is the first document we assess against the template for structural/content compliance. It contains three tabs—each tab is a different version of the same document—(B1) initial submission; (B2) DASN Edits; (B3) Final Clean Copy.

Doc 2-B1: TAB B1 - ISTRAP CNO Avails Fast Attacks Sub

Doc 2-B2: TAB B2 - ASN(RDA) CR_TC_ISTRAP CNO Avails

Doc 2-B3: TAB B3 - ISTRAP CNO Avails Fast Attacks Sub_20241213 CLEAN

Doc 3: DDG-FFG PY Acquisition Plan v9. This is the second document we use to assess against the template for structural/ content compliance. It contains three tabs—each tab is a different version of the same document—(C1) initial submission; (C2) DASN Edits; (C3) Final Clean Copy.

Doc 3-C1: TAB C1 - ISTRAP CNO Avails Fast Attacks Sub

Doc 3-C2: TAB C2 - ASN(RDA) CR_TC_ISTRAP CNO Avails

Doc 3-C3: TAB C3 - ISTRAP CNO Avails Fast Attacks Sub_20241213 CLEAN

Doc 4: DDG-FFG PY Acquisition Plan v9. This is the third document we use to assess against the template for structural/content compliance. It contains three tabs—each tab is a different version of the same document—(C1) initial submission; (C2) DASN Edits; (C3) Final Clean Copy.

Doc 4-D1: TAB D1 - ISTRAP CNO Avails Fast Attacks Sub

Doc 4-D2: TAB D2 - ASN(RDA) CR_TC_ISTRAP CNO Avails

Doc 4-D3: TAB D3 - ISTRAP CNO Avails Fast Attacks Sub_20241213 CLEAN

Summary Result of ISTRAP Structural Review Use-Case

Across three iterations using NIPR GPT, the model's assessment of the ISTRAP showed alignment with formal reviewer comments, though weaknesses and areas for refinement were identified. Its capacity for rapid initial assessments promises to enhance both the efficiency of the contract writing process and the accuracy of acquisition package compliance.

Enhanced Use Case—Annex 1—J&A Structural and Logical Flow Review

Doc 1: TAB A - NMCARS 18-25 Annex 1 - J&A. This is the template for a J&A and rules to be followed. We feed the model this document to assess the rest against.

Doc 2: J&A CJA No. CR-24219. This is the first document we assess against the template for structural/content compliance. It contains three tabs—each tab is a different version of the same document—(B1) initial submission; (B2) DASN Edits; (B3) Final Clean Copy.

Doc 2-B1: TAB B1 - CJA No. CR-24219 - Body - MS Word

Doc 2-B2: TAB B2 - CJA No. CR-24219 - Track Changes - DASN

Doc 2-B3: TAB B3 - CJA No. CR-24219_Final Clean 22JAN25

Doc 3: Draper CPS JA23-51. This is the second document we use to assess against the template for structural/content compliance. It contains three tabs—each tab is a different version of the same document—(C1) initial submission; (C2) DASN Edits; (C3) Final Clean Copy.

Doc 3-C1: TAB C1 - Draper CPS JA23-51 Final CLEAN

Doc 3-C2: TAB C2 - 20240806_ASN(RD&A) CR_TC_Draper CPS JA23

Doc 3-C3: TAB C3 - Draper CPS JA23-51 Final CLEAN

Doc 4: NAVSEA JA_DDG91. This is the second document we use to assess against the template for structural/content compliance. It contains three tabs—each tab is a different version of the same document—(C1) initial submission; (C2) DASN Edits; (C3) Final Clean Copy.

Doc 4-D1: TAB D1 - NAVSEA JA_DDG91_DMP2_to_NASSCO

Doc 4-D2: TAB D2 - ASN(RDA) CR_TC_NAVSEA_DDG91 DMP J&A - NA

Doc 4-D3: TAB D3 - Final USS PINCKNEY (DDG 91) DMP JA 42,9

Draper J&A (SPJA23-51):
Total Comments in Initial Review: 5
Comments with Direct Alignment: 4
Proportion of Alignment: 4/5 = 80%

Analysis: There was a strong degree of alignment for the Draper J&A, indicating that NIPR GPT was generally on the right track in identifying the key areas of concern.

USS PINCKNEY (DDG 91) Modernization J&A (J&A Number: 42,916):
Total Comments in Initial Review: 7
Comments with Direct Alignment: 6
Proportion of Alignment: 6/7 = 86%

Analysis: The alignment was even stronger for the USS PINCKNEY J&A, suggesting that NIPR GPT's understanding of the review team's priorities had improved.

Structural Review—J&A Enhanced Use Case

To assess the models' capabilities in both structurally reviewing Navy acquisition packages and assessing the strength of the document's arguments through logical flow assessments, testing was conducted across Annex 1 (J&As). NIPR GPT, across multiple iterations, showed increasing degrees of alignment with formal reviews including legal reviews, reinforcing its potential as a supplementary tool that can significantly enhance acquisition efficiency and compliance accuracy.

Structural Review—Multiple Use Cases

To assess the models' capabilities in structurally reviewing Navy acquisition packages, testing was conducted across Annex 1 (J&As), Annex 18 (ISTRAP), Annex 19 (PSTRAP-M), and Annex 20 (ISTRAP-M). NIPR GPT, across multiple iterations, showed varying degrees of alignment with formal reviews, highlighting the need for refinement but reinforcing its potential as a supplementary tool that can significantly enhance acquisition efficiency and compliance accuracy. The following is an analysis of the alignment between NIPR GPT's comments and the formal review team's comments for each of the three J&A packages, expressed as a proportion of comments that aligned. As an example, this was the Annex 1 Quantitative Result.

Overall Trend

The proportion of alignment increased over the three J&A packages, demonstrating that the review process became more closely aligned with the review team's perspective. This suggests that NIPR GPT effectively learned from the previous reviews and incorporated that knowledge into the subsequent assessments.

Key Takeaways

- **Effective Learning:** The increasing proportion of alignment indicates that NIPR GPT was able to effectively learn from the review team's feedback and incorporate their priorities into its own review process.
- **Areas for Improvement:** Even with the high levels of alignment, there were still some comments that NIPR GPT missed. This highlights the importance of continuous learning and refinement of the review process.
- **Value of Different Perspectives**: The combination of the model's general assessment and the review team's specific comments resulted in a more comprehensive and robust evaluation of the J&As.

Clause Compliance PDS-Based Analysis Approach

The clause compliance portion of this research began by analyzing the PDS XML contract award data to identify potentially missing clauses, which were evaluated based on known regulatory requirements and contract characteristics (e.g., contract type, dollar value, and acquisition strategy).

Recognizing the limitations of PDS data—which may not include clauses incorporated by reference or detailed in attachments—our initial approach focused on detecting likely omissions based on what was explicitly represented in the XML.

To facilitate structured analysis, potentially missing clauses were categorized by key areas of regulatory concern, including but not limited to:

Cost and Pricing:

- **FAR 52.216-7 Allowable Cost and Payment:** While mentioned within another clause, it should likely be a standalone clause, especially in a CPFF contract.
- **FAR 52.216-8 Fixed Fee:** Similar to the above, while referenced, it's best practice to include it directly.
- Clauses related to cost accounting standards (CAS), if applicable to the contractor.

Changes and Terminations:

- **FAR 52.243-1 Changes—Fixed-Price:** Or the appropriate Changes clause for a CPFF contract if modifications are anticipated beyond the issuance of task orders.
- **FAR 52.249-2 Termination for Convenience of the Government (Fixed-Price):** Or the appropriate termination clause for a CPFF contract.

Data Rights and Intellectual Property:

- Specific data rights clauses (e.g., **DFARS 252.227-7013 Rights in Technical Data—Noncommercial Items**) define ownership and usage of technical data. The XML mentions a "data rights strategy," making these clauses highly likely to be needed but potentially located elsewhere in the full contract.
- Clauses related to patents and copyrights, if applicable.

Subcontracts:

- **FAR 52.244-2 Subcontracts (Cost-Reimbursement and Letter Contracts):** Or the appropriate subcontracts clause for a CPFF contract.
- **FAR 52.219-9 Small Business Subcontracting Plan**, if applicable, based on the dollar value and nature of the work. The PWS mentions small business contracting, suggesting this clause might be necessary.

Other Important Areas (clauses may be needed depending on specific circumstances):

- **Inspection of Services:** A clause defining acceptance criteria and inspection procedures.
- **Insurance:** Clauses requiring specific types and levels of insurance.
- **Disputes:** A clause outlining the dispute resolution process.
- **Equal Opportunity:** Clauses related to equal employment opportunity and affirmative action.
- **Labor Standards:** Clauses related to labor laws (e.g., Service Contract Act, if applicable).

Key Takeaway

This list is not exhaustive and serves as a starting point. **We then consulted the full contract, all incorporated documents, and any applicable regulations (FAR, DFARS, NMCARS) to determine the complete set of required clauses to supplement the initial**

results, consulting with contract specialists for a thorough review. The model identified many applicable clauses and terms; however, ***not all*** left erroneous references and missing language.

Contract Risk and Compliance Assessment Using PDS XML

Next, the team provided PDS XML files representing multiple DON contract awards as part of the analysis. These structured data files were ingested by NIPR GPT, which performed a multi-layered review of each contract, focusing on both structural risks and regulatory compliance.

The system produced a detailed breakdown consisting of the following core elements: an overview of the contract's key terms, type, scope, and funding structure derived from the PDS metadata.

1. **Summary of the Contract**
2. **Risks Associated with the Contract Structure**
    - Cost-Plus-Fixed-Fee (CPFF) Risk
    - Indefinite Delivery Indefinite Quantity (IDIQ) Risk
    - Incrementally Funded Contract Risk
    - Level of Effort (LOE) Risk
    - Personnel Risks
    - Proprietary Information Agreements (PIAs) and Technical Assistance Agreements (TAAs) Risk
    - Travel Costs (Cost No Fee)
3. **Compliance with FAR, DFARS, and NMCARS Stipulations**

Contract Risk and Compliance Assessment NIPR GPT Output (*summary only*):

Based on the structured **PDS XML data** provided for each contract, **NIPR GPT** generated a detailed compliance assessment across multiple regulatory frameworks. The output was structured into the following categories:

1. **FAR Compliance**
 – Evaluation of required FAR clauses based on contract type, dollar value, and other contextual metadata.

2. **DFARS Compliance**
 – Assessment of DoD-specific regulatory provisions and how well they were represented in the contract data.

3. **NMCARS Compliance**
 – Review of DON-specific clauses and guidance under the NMCARS.

4. **Potential Compliance Concerns and Areas for Further Review**
 – Identification of clauses or contract features that may warrant additional scrutiny, including possible omissions or inconsistencies.

5. **Recommendations**
 – Actionable suggestions provided for both the **government** (e.g., clause corrections, structural risks) and the **contractor** (e.g., documentation improvements or clarifications).

A disclaimer accompanied each assessment to clarify the scope and limitations of the AI-generated analysis:

> "This analysis is based solely on the provided XML file. A complete assessment would require a review of the entire contract document, including all attachments and incorporated clauses. I am not a legal professional, and this is not legal advice."

Summary: Clause Set Completeness and Future Research Direction

During the course of our research, it became clear that **complete clause visibility**—including clauses incorporated by reference or detailed in attachments—is **essential for accurate contract compliance analysis**. Relying solely on the PDS XML representation proved insufficient for a comprehensive assessment, as it omits many contextually critical clauses.

As a result, we identified a **promising direction for further research**: leveraging the Clause Logic Service (CLS) to **associate its recommended clause set with the actual clauses present in the final award PDS**. This involves:

- Pulling both the **CLS recommendation outputs** and the **final contract award PDS** data via interface integration
- Creating **mappings between recommended and actual clauses**
- Analyzing the differences to identify **compliance gaps and best practices**

Through this process, we aim to **identify high-quality, "clean" contract examples**—those that demonstrate strong alignment between CLS guidance and final execution. These exemplar contracts can then be used to **fine-tune our RAG model**, supporting improved clause prediction and validation in future awards.

Programmatic Interface for Batch Summarization and Ranking

We developed a programmatic interface to interact with an LLM to efficiently conduct automated document analysis, requirement-specific grading, and confidence scoring. This interface enabled batch processing of contract documents, allowing for structured assessments across multiple dimensions, including:

1. **Automated Document Analysis**—Parsing and analyzing contract documents at scale to identify structural integrity, clause completeness, and regulatory compliance.
2. **Requirement-Specific Grading**—Evaluating each package against predefined compliance criteria, assessing adherence to FAR, DFARS, and NMCARS.
3. **Confidence Scoring**—Assigning an entropy-based confidence score to each assessment, flagging high-uncertainty outputs for further expert review.
4. **Requirement Grading—**Ranking packages based on compliance strength, highlighting areas that require revision or additional scrutiny.

The programmatic interface facilitated seamless interaction with the LLM, ensuring a repeatable and efficient workflow for large-scale contract evaluations. We utilized obfuscated data to test and refine the automated interface, allowing for code validation and efficiency testing while preserving data integrity and security.

This interface streamlined contract analysis workflows by integrating batch processing, ranking mechanisms, and AI-driven confidence scoring, enhancing review accuracy, efficiency, and scalability. Future iterations will focus on further optimizing response accuracy, refining ranking methodologies, and enhancing real-time feedback mechanisms for contract evaluators. Key steps and observations from the development team include:

This framework uses ISTRAP documents against a predefined set of requirements to automate the analysis. Using the Gemini 2.0 Flash Lite AI model, each document is evaluated and graded (A–F) based on how well it meets each individual requirement. A certainty score, derived from the model's perplexity (exponentiated entropy), indicates the confidence level of each assigned grade (Tornetta, 2021):

$$C(X) := 1 - PP(x) = exp \sum_{i \in A:F} p_i \, ln \, p_i$$

The final output is a table showing the grade and certainty for each requirement in the analyzed document.

1. **Automated Document Analysis:** The software takes a set of ISTRAP documents (PDFs found in `data/contracts/`) and systematically analyzes each one against a predefined list of requirements derived from the `reqs/ISTRAP.pdf` and structured in `reqs/response-schema.txt`.
2. **Requirement-Specific Grading:** For *each* document, the AI model (Gemini 2.0 Flash Lite) is prompted to evaluate how well that document addresses *each individual requirement* listed in the schema (Liu et al., 2021).
3. **Grading Scale:** The AI assigns a grade (A, B, C, D, or F) for every requirement within each document, based on the instructions provided (`system_instruction`) and the specific requirement's description. Table 1, Grading Scale for ISTRAP Reviews, shows prioritized results for ISTRAP reviews, including model certainties in outcomes (measured using model perplexity over tokens).
4. **Confidence Score (Certainty):** The software calculates a "Certainty" score for each grade assigned from the model's perplexity (Morgan, 2024). Perplexity is a measure of how surprised or uncertain the model was when generating the response (the grade). Lower perplexity (closer to 1) means higher certainty/confidence in the assigned grade. Figure 2, Histogram Count of ISTRAP Requirements Colored by Grade, illustrates the model's confidence score for each ISTRAP requirement.
5. **Tabular Output (`results` DataFrame):** The final output is a table (currently showing results for *one* processed document). This table has three columns:
    - `Requirement`: The specific ISTRAP requirement text being evaluated (e.g., "1.1: Statement of Need").
    - `Grade`: The A–F grade assigned by the AI for that requirement in the analyzed document.
    - `Certainty`: The calculated confidence score for that specific grade.
6. **Overall Confidence Assessment:** An overall perplexity score is calculated across all the grades for the analyzed document, providing a general sense of the AI's confidence in its assessment of that entire document.

Table 1. Grading Scale for ISTRAP Reviews

Requirement str unique: 61	Grade str unique: 4	Certainty f64
1.1: Statement of Need	A	42.16%
1.2: Historical Summary	A	46.88%
1.3.1: Program Objectives by Phase	A	72.34%
1.3.2: Contractor Data Requirements and Rights	B	46.65%
1.3.3: Cost Effectiveness of Buying Data	C	42.47%
1.3.4: Technical Data Package Validation	B	85.67%
1.3.5: Patents and Copyrights	B	76.33%
1.4: Funding Identification	A	14.3%
2.1.1: Product/Service Code Choice and Rationale	A	70.43%
2.1.2: Required Capabilities/Performance Standards	A	82.91%

61 rows, 3 columns   Page 1 of 7   Download

```
1 results
```

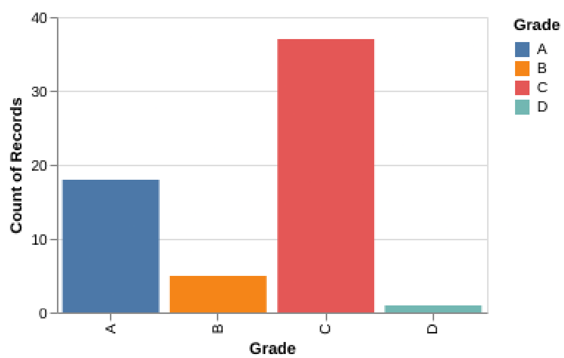


Figure 2. Histogram Count of ISTRAP Requirements Colored by Grade

Figure 3, Histogram Count of the Model Certainty Colored by Grade, illustrates ISTRAP graded point-by-point for compliance; we generate summary charts to help users understand both the evaluated quality of the ISTRAP and the model's confidence in its outputs.

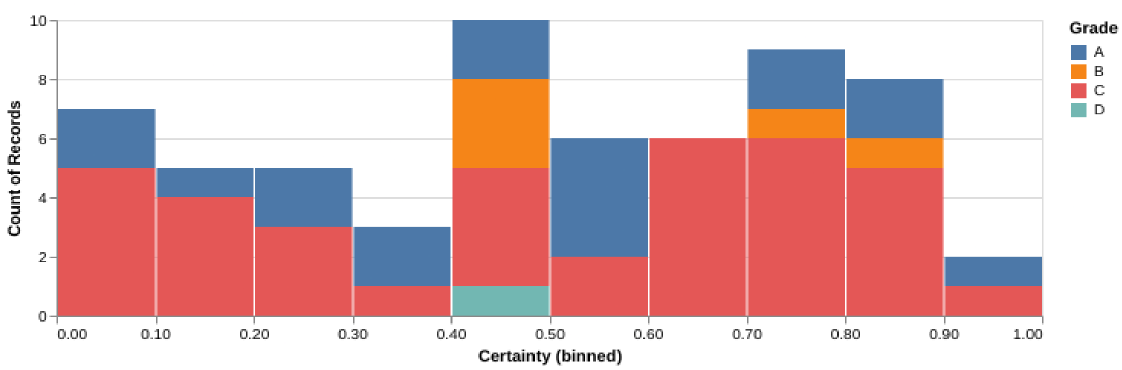


Figure 3: Histogram Count of the Model Certainty Colored by Grade

Our research identifies several significant potential improvements from incorporating our framework into the ISTRAP review process.

1. **Efficiency and Speed:** Manually checking every ISTRAP against dozens of detailed requirements is time-consuming. This software automates the initial review, drastically reducing the time needed per document. Procurement personnel can process more ISTRAPs faster.
2. **Granular Feedback:** Instead of a simple pass/fail, the system provides a grade for each *specific requirement*. This immediately highlights the exact sections or requirements within an ISTRAP that are weak (e.g., grades C, D, F) and need attention or revision.
3. **Focused Review:** Reviewers can use the results table to prioritize their efforts. They can quickly scan for low grades and focus detailed human reviews on those specific areas of non-compliance or weakness rather than re-reading sections where the AI is confidently graded as compliant (A or B; Reddy, 2024).
4. **Consistency:** The AI applies the same criteria (based on the system instruction and requirement definitions) to every document, reducing the variability and potential subjectivity that can occur with different human reviewers. This leads to more standardized initial compliance checks.
5. **Training and Template Improvement:** Common patterns of low grades can emerge by analyzing results across multiple ISTRAPs (if the code is extended to aggregate results from all documents). This data can identify systemic weaknesses in how ISTRAPs are being written, informing targeted training for personnel or improvements to ISTRAP templates and guidance.
6. **Risk Indication:** The “Certainty” score adds another layer of insight. A low grade with high certainty strongly indicates a problem. A low grade with low certainty might suggest the document is ambiguous or the AI struggled, warranting closer human inspection.
7. **Audit Trail:** The software execution and its results provide a documented record of the initial compliance check performed on each ISTRAP.

This framework provides an automated, granular, and consistent first-pass compliance check for **FAR, DFARS, and NMCARS**. It allows contracting officers to quickly identify potential issues, focus their review efforts efficiently, and visualize data in a way that leads to systemic improvements in **FAR, DFARS, and NMCARS** quality and compliance. The framework testing, while conducted specifically on the NMCARS Annex 18—ISTRAP requirements, establishes a foundational methodology applicable to structural compliance checks across all FAR, DFARS, and NMCARS regulations.

Conclusions

This research demonstrates the potential of leveraging generative AI, specifically NIPR GPT and RAG, to enhance the efficiency and accuracy of DoD acquisition package validation. The application of AI-driven tools shows promise in automating key tasks such as compliance checks, clause verification, and risk identification, ultimately reducing the administrative burden on contracting officers and minimizing the risk of errors.

Our findings indicate that AI models can effectively identify structural compliance gaps and potential omissions in contract documents, as evidenced by the ISTRAP and J&A analysis and the clause compliance assessment using PDS XML data. The integration of RAG further improves the accuracy and reliability of AI-generated outputs by providing access to real-time regulatory references, addressing limitations such as hallucinations and misinterpretations.

The development of a programmatic interface for batch summarization and ranking streamlines the analysis of multiple contract documents, enabling efficient identification of high-risk discrepancies and prioritization of review efforts. This automation not only accelerates the review process but also enhances consistency and objectivity in compliance assessments.

However, it is crucial to acknowledge the limitations of AI-driven solutions. The accuracy of AI models depends on the quality and completeness of the data they are trained on, and they may not always capture the nuances of regulatory language or adapt instantly to evolving regulations. Therefore, human oversight remains essential to validate AI-generated outputs and ensure comprehensive compliance.

Future research should focus on refining AI models through continuous learning and fine-tuning, integrating them with existing procurement systems, and developing best practices for human-AI collaboration. By addressing these challenges, the DoD can fully leverage the transformative potential of generative AI to modernize procurement practices, improve contract quality, and ultimately enhance mission readiness.

References

Atamel, M. (2025, January). *Evaluating RAG pipelines.* Medium. https://medium.com/google-cloud/evaluating-rag-pipelines-d99e007e625f

Entropy (information theory). (n.d.). In *Wikipedia*. Retrieved March 26, 2025, from https://en.wikipedia.org/wiki/Entropy_(information_theory)

Founding Minds. (2024, September 24). Chain-of-thought reasoning: The magic behind the o1 model. https://www.foundingminds.com/chain-of-thought-reasoning-the-magic-behind-the-o1-model/

Gadesha, V. & Kavlakoglu, E. (2024, August 12). *What is chain of thoughts (CoT)?*. IBM. https://www.ibm.com/think/topics/chain-of-thoughts

Google Cloud. (n.d.). *What is retrieval-augmented generation (RAG)?*. Retrieved March 26, 2025, from https://cloud.google.com/use-cases/retrieval-augmented-generation

Leblond, R., Alayrac, J.-B., Sifre, L., Pislar, M., Lespiau, J.-B., Antonoglou, I., Simonyan, K., Vinyals, O. (2021). *Machine translation decoding beyond beam search.* ArXiv. https://arxiv.org/pdf/2104.05336

Liu, L., Pan, Y., Li, X., & Chen, G. (2024). *Uncertainty estimation and quantification for LLMs: A simple supervised approach*. ArXiv. https://arxiv.org/html/2404.15993v1

Martyr, R. (2024, December 22). *Chain of thought prompting in AI: A comprehensive guide [2025]*. Orq. https://orq.ai/blog/what-is-chain-of-thought-prompting

Morgan, A. (2024, November 21). *Perplexity for LLM evaluation.* Comet. https://www.comet.com/site/blog/perplexity-for-llm-evaluation/

OpenAI. (2024, August 6). *Introducing structured outputs in the API*. https://openai.com/index/introducing-structured-outputs-in-the-api/

OUSD(A&S). (n.d.-a). *Data standards: Procurement data standard and other enterprise initiatives.* Retrieved April 2, 2025, from https://www.acq.osd.mil/asda/dpc/ce/ds/procurement-data-standard.html#:~:text=The%20Procurement%20Data%20Standard%20(PDS,including%20grants%20and%20other%20transactions.

OUSD(A&S). (n.d.-b). *Procurement data Standard and other enterprise initiatives.* Retrieved

April 2, 2025, from https://www.acq.osd.mil/asda/dpc/ce/ds/procurement-data-standard.html

Reddy, Y. (2024, October). *Avoiding LLM hallucinations and building LLM confidence scores*. Nanonets. https://nanonets.com/blog/how-to-tell-if-your-llm-is-hallucinating/

Shorten, C., Pierse, C., Smith, T. B., Cardenas, E., Sharma, A., Tengrove, J., & van Luijt B. (2024). *StructuredRAG: JSON response formatting with large language models.* ArXiv. https://arxiv.org/html/2408.11061v1

Tornetta, G. N. (2021). View of entropy methods for the confidence assessment of probabilistic classification models. *Statistica*. https://rivista-statistica.unibo.it/article/view/11479/13978

Villani, T. (2024, November 19). *How to implement chain-of-thought prompting for better AI reasoning.* New Jersey Innovation Institute. https://www.njii.com/2024/11/how-to-implement-chain-of-thought-prompting-for-better-ai-reasoning/

Zvornicanin, E. (2024, July, 16). *What are the evaluation metrics for RAGs?.* Baeldung on Computer Science. https://www.baeldung.com/cs/retrieval-augmented-generation-evaluate-metrics-performance

High-Technology and Cosmopolitan Companies in the Federal Supply Chain

Edward (Ed) Hyatt, PhD—is a senior research fellow at the Baroni Center for Government Contracting in the Costello College of Business at George Mason University. He has a decade of research experience and another decade of managerial experience in the public procurement and contracts profession. He holds several advanced degrees, including a PhD in Business Management from The University of Melbourne, and has participated in the research and publication process on dozens of projects. He has two streams of research, one centered on organizational behavior topics like personnel selection and managerial decision-making, and the other involving government acquisition matters. [ehyatt4@gmu.edu]

Abstract

Using a combination of USASpending.gov, SAM.gov, and BLS.gov data, the following research investigates two aspects of the defense industrial base. The first, a recognizable yet little operationalized concept, innovation, is often associated with new technologies and has become an important focus of defense acquisition in recent years. Unfortunately, research and policy efforts have been hindered by a lack of rigorous definition and application to present-day industry classifications. To address this shortcoming the author applied a STEM-occupation methodology developed by the U.S. Bureau of Labor Statistics to identify core and peripheral high-technology industries (as a surrogate for innovation). The second concept, newly introduced in this study, describes the existence of companies that perform work for both defense and non-defense agencies. Borrowing a term from biogeology, these contractors are dubbed "cosmopolitan," as in having a broad operating range as opposed to those contractors that are endemic to only one contracting environment. Key characteristics of these cosmopolitan contractors are presented, and their potential importance to future research is highlighted. Finally, the results of both research efforts are used to produce a report to show how these concepts can serve a practical purpose and potential follow-on empirical research is discussed.

Keywords: Innovation, High-technology, STEM, Cosmopolitan, Defense Industrial Base

Background

The composition of the U.S. defense industrial base (DIB) has recently attracted much scholarly and policy interest. The DIB is defined by the Congressional Research Service as "all organizations and facilities that provide [Department of Defense] DOD with materials, products, and services" ((Congressional Research Service, 2023, p. 1). There has been a particular focus on the concept of innovation within the DIB, most obviously embodied in the proliferation of rapid acquisition programs and policies to adopt commercial technologies. The DoD has also pursued efforts to increase the participation of small businesses and non-traditional contractors because they are often viewed as a source of fresh thinking and innovation. This view is broadly supported by the federal government that believes stronger supply chains can be built through a "greater focus on new and recent contractors that—along with established contractors—can regularly provide fresh innovative thinking and seasoned expertise to support agencies in addressing national priorities" (U.S. Office of Management and Budget, 2023, p. 1).

Two observations, at least germane to this paper, can be made about the present situation. First, innovation and technology are often used interchangeably or in combination in many DoD departments, policies, and resources. For example, the DoD's "defense innovation policy aims to improve warfighting capabilities through adopting technologies and processes" (DoD, n.d.), and the Defense Innovation Board is focused on "emerging technologies and innovative approaches" (Defense Innovation Board, n.d.). Despite the obvious emphasis on technology, there are limitations for research and policy development because at an industry level the concept has not been well defined or operationalized. Second, the clear focus of these programs is on commercial companies that exist outside the federal contracting sphere; little attention has been paid to existing non-defense federal contractors that have at least already learned how to operate in the world of federal contracting. This represents both a gap in conceptual understanding and a potential missed resource.

The following paper details two research efforts that seek to address these issues. The first is focused on the idea of "high-technology" as a surrogate for innovation, and the other introduces the concept of "cosmopolitan" companies that do both defense and non-defense work.[1] The aim of the first research effort involving "high-technology" is to provide a framework for more rigorous research, policy development, and acquisition efforts. The aim of the second research effort involving "cosmopolitan" is to introduce a new concept that appears to offer promising avenues of future research. Each research effort is first independently described in full (background, data source, method, and results). The two efforts then converge with an example report that exemplifies their potential benefit, and the paper concludes with a series of recommendations for how the concepts can be applied in future research.

Research Effort #1: High-Technology Industries

Innovation: Concept and Practice

Innovation as a concept is a pervasive yet ubiquitous term, often used to describe everything ranging from outcomes, products, processes, business models, organizational structures, or even mindsets (Kahn, 2018). To that end, there is no single, unitary theory of innovation but rather different theories that seek to explain different aspects of innovation (Downs Jr. & Mohr, 1976). A large number of typologies for innovation have been developed: radical vs. incremental, original vs. borrowed, expansionary vs. evolutionary vs. developmental, administrative, product, process, and technological (Jaskyte, 2011). Innovation also features prominently in large scales that rely on measures like number of patents; for example, the Global Innovation Index (World Intellectual Property Organization, 2024) and the National Innovative Capacity Index (Porter & Stern, 2001). At the scholarly level, there are often strict parameters for how innovation is defined and measured, but because there are hundreds of potential conceptual distinctions it can be difficult to apply to daily decision-making.

In the defense world, the concept of innovation is most strongly associated with the adoption of new technologies, with a recent emphasis on commercial and dual-use technologies. For example, the Congressional Research Service has defined the defense innovation ecosystem as "the set of organizations, activities, functions, and processes that develop, produce, and field new or improved *technologies* [emphasis added] and capabilities for military use" (Congressional Research Service, 2025, p. 1). To that end, dozens of defense-sponsored innovation initiatives such as the Small Business Innovation Research/Small Business Technology Transfer (SBIR/STTR) program have sought to increase the presence of innovative companies in the military's industrial base. Indeed, so many separate organizations have been established that the Defense Innovation Unit has recently become the focal point for commercial technologies, responsible for coordinating efforts in this area across more than 100 other military service organizations (United States General Accounting Office, 2025).

While the concept of innovation is clear, and its association with technology made evident, it is often not rigorously defined or adequately operationalized in practice, especially at the industry level. This shortcoming means that practical-minded research that touches on this topic often must take methodological shortcuts. For example, Bresker and Bresler (2020) noted

[1] According to the Congressional Budget Office, DoD spending accounts for nearly all the nation's defense budget (Congressional Budget Office, n.d.). Therefore, the term "defense" in this paper is synonymous with DoD and "non-defense" refers to all other federal agencies combined.

that associating product and services codes (PSCs) with innovation is subjective, and only briefly considered the “most obvious ‘non-innovative’ PSCs” to get a sense that the majority of new DIB vendors were not innovative commercial technology companies. This cursory approach, almost certainly born out of necessity, highlights the need to more rigorously distinguish “innovation” in defense acquisition research. Thankfully, the close association between innovation and technology offers a way forward. Although it is dated, some work has been done in this realm by the U.S. Bureau of Labor Statistics (BLS) relying on employment statistics. This paper therefore adopts “high-technology” as a surrogate for innovation and uses the BLS method to identify those industries that would be of keenest interest to DoD practitioners and researchers.

Technology: A Surrogate for Innovation

Previous research has made efforts to define and measure high-technology industries; one such approach relies on measures of the proportion of technical jobs in science, technology, engineering, and mathematics (STEM) within an industry. STEM occupations heavily involve the fields of computers, mathematics, architecture, engineering, and life and physical sciences, as well as managerial and postsecondary teaching occupations related to these functional areas and sales occupations requiring scientific or technical knowledge at the postsecondary level (U.S. Bureau of Labor Statistics, 2024b).. Computer-related occupations such as computer support specialists, systems analysts, and software engineers have historically made up roughly half of all STEM employment ((Cover et al., 2011; Fayer et al., 2017). The Occupational Employment and Wage Statistics (OEWS) program at BLS annually produces employment and wage estimates, including for STEM and non-STEM occupations, for approximately 830 occupations. See *Appendix 1* for the list of STEM occupations according to the 2018 Standard Occupational Classification (SOC) system.

The North American Industry Classification System (NAICS), which is the business classification standard used by federal agencies since 1997 to conduct statistical analyses related to the U.S. business economy (U.S. Census Bureau, 2022), does not itself distinguish high-technology industries. However, relying on BLS STEM data, high-technology industries can be identified as industries having a higher-than-average concentration of workers in STEM occupations. The philosophy underpinning this approach is that “a country’s competitive position will be largely determined by the quality of its investment in human and capital resources dedicated to science and technology” (National Science Foundation, 1989, p. vii). Indeed the outsized importance of high-technology industries for the U.S. economy has been noted in various studies; for example, in 2014 while they accounted for about 12% of total national employment those industries contributed almost 23% to output (Wolf & Terrell, 2016). Based on this, a National Science Foundation report on science and technology resources considered the employment and utilization of scientists, engineers, and technicians as one of the most important parameters of innovation and used it, along with measures of R&D activity, as surrogates for the broader concept of innovation (National Science Foundation, 1989).

This framework has been used by BLS researchers and others in a series of publications on the topic (Economic and Labor Market Information Bureau, New Hampshire Employment Security, n.d.; Hecker, 1999, 2005; Workforce Information Council, 2015). Hecker (2025) defined an industry as high-technology if the proportion of employment in technology-oriented occupations (i.e., STEM occupations) within that industry accounted for at least twice the 4.9% average for all industries at that time. He then established three levels of high-technology based on 2.0 to 2.9 times the average, 3.0 to 4.9 times the average, and at least 5 times the average. The result was 46 total industries identified as high-technology. The Workforce Information Council (2015) similarly calculated two levels of high-technology: core concentration defined as industries with at least 5 times average concentration in STEM occupations and peripheral concentration as industries with at least 2.5 times the national average. The result was 33 industries identified as high-technology, which was later replicated by Wolf and Terrell (2016) when they used the same threshold of two and a half times the national average. High-technology industries are therefore similarly defined in this study as those industries with a high proportion of STEM occupations compared to the national average.

There are two issues worth noting about the usefulness of prior research in this area, both dealing with datedness. One issue is that the nationwide average STEM employment has changed over time, as well as the proportion of STEM employment within individual industries. For instance, Hecker (2005) reported a 4.9% average for all industries, Cover et al. (2011) reported about 6% (nearly 8 million jobs) in 2009, and Fayer et al. (2017) reported 6.2% (nearly 8.6 million jobs) in 2015. This means that high-technology industries have likely shifted to some extent over time as STEM employment naturally fluctuates. The other issue is that NAICS codes undergo updates every five years; the most recent update was conducted in 2022. Hecker (2005) used the 2002 NAICS list, a list of high-technology titles built by the Economic and Labor Market Information Bureau, New Hampshire Employment Security (n.d.) adjusted the industries to accommodate the 2017 NAICS revision but did not re-run Hecker's underlying analysis, and the Workforce Information Council (2015) presumably used the 2012 list (although it was not specified) and reran the analysis using different levels than Hecker (2005). Therefore, to the author's best knowledge, no recent analysis has applied the BLS methodology to current employment data, taking into account up-to-date NAICS, which means that the prior, limited work done to identify high-technology industries is likely outdated.

Data and Method

First, the author downloaded the most recent set of BLS OEWS occupation profiles and tables, released on April 3, 2024, including an aggregated STEM data set containing national and industry-level STEM employment data (U.S. Bureau of Labor Statistics, 2024a). These files incorporate current NAICS, thereby ensuring the raw data reflect present day industry classifications. The author also downloaded files related to the SOC Policy Committee recommendations to the Office of Management and Budget (OMB) that provided a framework for defining STEM occupations under the 2018 SOC system (U.S. Bureau of Labor Statistics, 2019a, 2019b). The author then applied the STEM-occupation methodology used in previous research to identify peripheral and core high-technology industries using two thresholds of STEM concentrations: two and a half times and five times the national average of STEM employment, respectively.

Results

According to BLS OEWS data, the national average for STEM occupations was 6.7% of total employment, representing 10,165,900 jobs. Two levels of high-technology industries were calculated, (a) peripheral: two and a half to five times the national average, which constitutes 16.74% to 33.47% of industry employment, and (b) core: greater than 5 times the national average, which represents more than 33.47% STEM employment in the industry. Twenty-six high-technology industries (15 peripheral and 11 core) were identified (see Table 1). For interested readers, a second table displaying the same core industries but with 26 peripheral industries defined at two times the national average (constituting 13.39% to 33.47% of industry employment) is provided in Appendix 2.

Table 1. High-Technology Industries (Core=5x, Peripheral=2.5x)

NAICS code	Industry	STEM % of employment	High-Tech Level
541500	Computer Systems Design and Related Services	59.7	Core
541300	Architectural, Engineering, and Related Services	59.5	Core
541700	Scientific Research and Development Services	55.9	Core
513200	Software Publishers	55.8	Core
334100	Computer and Peripheral Equipment Manufacturing	49.8	Core
518200	Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services	47.2	Core
519200	Web Search Portals, Libraries, Archives, and Other Information Services	44.7	Core
334500	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	38.1	Core
334200	Communications Equipment Manufacturing	37.8	Core
334400	Semiconductor and Other Electronic Component Manufacturing	34.9	Core
334600	Manufacturing and Reproducing Magnetic and Optical Media	33.9	Core
336400	Aerospace Product and Parts Manufacturing	32.2	Peripheral
423400	Professional and Commercial Equipment and Supplies Merchant Wholesalers	29.5	Peripheral
334300	Audio and Video Equipment Manufacturing	27.1	Peripheral
521100	Monetary Authorities-Central Bank	27.0	Peripheral
325400	Pharmaceutical and Medicine Manufacturing	26.6	Peripheral
516200	Media Streaming Distribution Services, Social Networks, and Other Media Networks and Content Provid	24.7	Peripheral
517000	Telecommunications	22.5	Peripheral
333600	Engine, Turbine, and Power Transmission Equipment Manufacturing	22.0	Peripheral
333300	Commercial and Service Industry Machinery Manufacturing	20.3	Peripheral
551100	Management of Companies and Enterprises	18.9	Peripheral
211100	Oil and Gas Extraction	17.9	Peripheral
541600	Management, Scientific, and Technical Consulting Services	17.8	Peripheral
4240A2	Merchant Wholesalers, Nondurable Goods (4242 and 4246 only)	17.4	Peripheral
999100	Federal Executive Branch (OEWS Designation)	17.1	Peripheral
221100	Electric Power Generation, Transmission and Distribution	16.8	Peripheral

Note. Core = Industries with at least five times the national average concentration in STEM occupations; Peripheral = Industries with at least two and a half times the national average concentration in STEM occupations.

In short, this industry list provides a way to systematically identify suppliers that operate in high-technology, or innovative, fields that are deemed critical to the health of the DIB.

Research Effort #2: "Cosmopolitan" Companies

Background

Various aspects of the depth and extent of company participation in the federal workplace have been investigated to-date. Josephson et al. (2019) examined the performance implications of companies with varying levels of government customer breadth and depth across federal agencies. In their study, Carril and Duggan (2020) concluded that industry concentration resulting from company mergers has caused the DoD procurement process to become less competitive and to induce a shift from the use of fixed-price towards cost-plus contracts, although these impacts did not appear to produce a significant increase in acquisition costs. A study by the Baroni Center for Government Contracting relied on a large-scale survey of presumably exited contractors to investigate reasons for declining contractor participation in the DIB (Hyatt & Everhart, in press). The following research effort described here extends this limited work but is exploratory in nature. It introduces a new concept, a "cosmopolitan" contractor, and explores some of the basic characteristics of this type of contractor to justify its potential importance for future research.

This effort leans on concepts from biogeology, an integrative field that studies the distribution of species across space and time. A cosmopolitan species is one that can survive in a range of climates or environments. As examples, pigeons can be found in most urban areas around the world, and migratory animals such as orcas, blue whales, and great white sharks range across every major oceanic body on Earth (Wikipedia, 2024). Alternatively, an endemic species, like the snow leopard of Central Asian mountain ranges, is found in a single environment and is usually specifically adapted to exist in only that environment. Borrowing these terms, the author defines a “cosmopolitan” company as one that does both defense and non-defense work (i.e., performs on contracts for both DoD and non-DoD agencies). Endemic companies are those that work exclusively on either defense or non-defense contracts. The intimation here is that cosmopolitan companies can work in a broader range of environments, as there are unique requirements, policies, and standards for being a defense contractor above and beyond being a federal contractor.

Data and Method

The author relied on several data sources for the raw data to conduct the analysis. USASpending.gov contains contract transaction data for nearly all federal government prime contract awards since Fiscal Year 2001. These data for all fiscal years were downloaded in early 2024 using the Award Data Archive. SAM.gov provides data on all active contractors as well as contractors that have become inactive during the previous six months. These data were downloaded in early 2024 to obtain the primary NAICS reported by companies. All computations were run using IBM SPSS Statistics (Version: 29.0.0.0), and all figures and tables were created using Microsoft Excel 360. None of the dollar figures in this section have been adjusted for inflation since the important calculations are proportions. A complete list of unique companies and their defense and non-defense obligations in every fiscal year from FY2001 to FY2023 was generated, thereby identifying cosmopolitan and endemic companies over the time period.

Results

First, the number of unique endemic companies performing either only defense work or only non-defense work, and the number of cosmopolitan companies working in both realms, from FY2001–2023 were calculated (see Figure 1). Then, the number of endemic and cosmopolitan contractors and the amount of dollars obligated to each group for each Fiscal Year was assessed (see Table 2). Finally, the author identified key characteristics of all three types of companies, although the analysis was limited to only FY2023 for ease of computation (see Table 3). Characteristics that were identified include average and median amount of contract obligations, company size (small vs. other-than-small), organization type, commodity (based on PSC), and industry (based on NAICS).

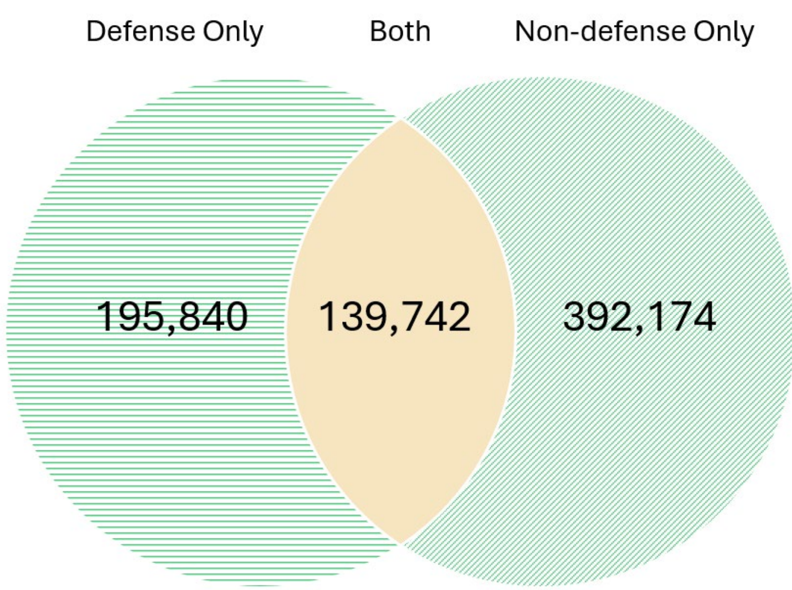


Figure 1. Unique Contractors: Defense Only, Non-Defense Only, and Both (FY2001–2023)

One of the most important conclusions from these numbers is that cosmopolitan contractors that provide products or services to both the DoD and other federal agencies (rather than exclusively to one or the other) have constituted approximately 16% of the total federal contractor base since FY2012, but they have accounted for roughly 60% of all federal obligations during that time. Their outsized presence in the federal supply chain indicates the importance of investigating cosmopolitan contractors and their unique experiences of contracting for both civil and defense agencies. The cohort of cosmopolitan companies is also clearly different from the other two endemic groups of contractors in some respects. For example, they appear more likely to not only provide services but also provide a higher proportion of R&D work, and they are more likely to be a not-tax-exempt corporation. Observations like these should be investigated further to see if there are other effects that result from these differences.

Table 2: Endemic and Cosmopolitan Companies per Year, FY2001–FY2023

	Endemic								Cosmopolitan			
	Defense Only				Non-defense Only				Both			
Year	Contractors (#)	Contractors (% of total)	Obligations ($ in millions)	Obligations (% of total)	Contractors (#)	Contractors (% of total)	Obligations ($ in millions)	Obligations (% of total)	Contractors (#)	Contractors (% of total)	Obligations ($ in millions)	Obligations (% of total)
FY2001	29,623	39%	72,548	33%	38,198	50%	46,518	21%	8,089	11%	103,892	47%
FY2002	40,483	45%	81,920	31%	39,253	44%	52,873	20%	9,814	11%	129,383	49%
FY2003	50,987	45%	105,318	32%	49,754	44%	60,964	19%	12,576	11%	158,635	49%
FY2004	57,353	39%	106,579	31%	70,694	48%	55,557	16%	18,451	13%	177,064	52%
FY2005	68,376	38%	116,305	30%	85,867	48%	59,109	15%	24,483	14%	215,722	55%
FY2006	62,121	33%	131,518	31%	97,383	52%	66,874	16%	27,417	15%	232,612	54%
FY2007	62,796	32%	140,127	30%	101,667	53%	65,605	14%	29,135	15%	263,887	56%
FY2008	61,619	31%	158,166	29%	105,991	54%	63,544	12%	30,035	15%	320,154	59%
FY2009	62,613	33%	161,262	30%	100,196	52%	77,771	14%	28,690	15%	301,746	56%
FY2010	60,316	31%	148,431	26%	103,371	54%	97,033	17%	28,924	15%	315,885	56%
FY2011	57,312	30%	139,479	26%	102,889	55%	76,462	14%	28,183	15%	323,907	60%
FY2012	52,914	31%	146,847	28%	91,888	53%	69,137	13%	27,160	16%	304,718	59%
FY2013	46,697	30%	109,770	24%	86,187	55%	68,232	15%	25,110	16%	285,547	62%
FY2014	44,207	29%	109,927	25%	84,355	55%	71,886	16%	24,703	16%	264,437	59%
FY2015	43,270	29%	103,346	24%	82,903	55%	73,050	17%	24,498	16%	263,297	60%
FY2016	41,784	28%	108,247	23%	82,082	56%	78,484	17%	23,886	16%	288,403	61%
FY2017	39,531	26%	117,622	23%	88,282	58%	87,712	17%	24,016	16%	305,148	60%
FY2018	38,398	28%	135,160	24%	75,610	56%	83,406	15%	22,104	16%	337,250	61%
FY2019	36,777	29%	144,122	24%	68,002	54%	84,184	14%	20,378	16%	361,871	61%
FY2020	34,702	28%	149,178	22%	68,104	56%	120,660	18%	19,180	16%	401,330	60%
FY2021	33,539	29%	139,539	22%	64,956	56%	127,010	20%	18,501	16%	378,820	59%
FY2022	30,787	28%	167,193	24%	62,493	56%	131,589	19%	17,745	16%	395,500	57%
FY2023	30,466	28%	182,101	24%	61,231	56%	145,225	19%	17,699	16%	431,768	57%

Table 3: Endemic and Cosmopolitan Companies, Characteristics, FY2023

Attribute	Characteristic	Defense Only	Non-defense Only	Both
Obligation	Dollars (average)	$101,856	$315,958	$97,649
	Median (median)	$494	$969	$315
Size	Small Business	21,817 (67%)	42,719 (65%)	14,548 (63%)
	Other than Small Business	10,715 (33%)	22,581 (35%)	8,551 (37%)
Organization Type	Corporate (not tax exempt)	18,782 (61%)	33,311 (54%)	13,370 (70%)
	Corporate (tax exempt)	1,475 (5%)	3,308 (5%)	831 (4%)
	Foreign government	45 (0%)	49 (0%)	5 (0%)
	International organization	662 (2%)	645 (1%)	103 (1%)
	Partnership	4,090 (13%)	7,811 (13%)	2,063 (11%)
	Sole proprietorship	3,011 (10%)	10,343 (17%)	749 (4%)
	US government entity	660 (2%)	1,752 (3%)	292 (2%)
	Other	2,166 (7%)	4,259 (7%)	1,748 (9%)
Commodity	Product	33,188 (59%)	65,181 (56%)	10,207 (39%)
	Service	20,507 (36%)	49,546 (43%)	13,994 (53%)
	Research and Development	2,783 (5%)	1,827 (2%)	2,280 (9%)
Industry	Agriculture, Forestry, Fishing and Hunting	113 (0%)	3,569 (4%)	330 (0%)
	Mining, Quarrying, and Oil and Gas Extraction	217 (0%)	190 (0%)	206 (0%)
	Utilities	517 (1%)	954 (1%)	554 (1%)
	Construction	3,995 (6%)	6,698 (8%)	5,792 (7%)
	Manufacturing	43,282 (62%)	15,602 (19%)	37,946 (43%)
	Wholesale Trade	407 (1%)	1,504 (2%)	1,956 (2%)
	Retail Trade	381 (1%)	518 (1%)	852 (1%)
	Transportation and Warehousing	1,527 (2%)	2,518 (3%)	1,661 (2%)
	Information	1,203 (2%)	4,004 (5%)	4,929 (6%)
	Finance and Insurance	44 (0%)	534 (1%)	118 (0%)
	Real Estate and Rental and Leasing	694 (1%)	1,627 (2%)	949 (1%)
	Professional, Scientific, and Technical Services	7,482 (11%)	21,734 (26%)	18,782 (21%)
	Management of Companies and Enterprises	2 (0%)	5 (0%)	3 (0%)
	Administrative & Support and Waste Management & Remediation Services	3,625 (5%)	8,241 (10%)	5,865 (7%)
	Educational Services	1,125 (2%)	2,244 (3%)	2,003 (2%)
	Health Care and Social Assistance	560 (1%)	4,813 (6%)	1,189 (1%)
	Arts, Entertainment, and Recreation	318 (0%)	647 (1%)	189 (0%)
	Accommodation and Food Services	1,728 (2%)	879 (1%)	370 (0%)
	Other Services (except Public Administration)	2,576 (4%)	4,843 (6%)	4,109 (5%)
	Public Administration	268 (0%)	1,054 (1%)	377 (0%)

Note. The numbers within each attribute are inflated because companies may exhibit more than one characteristic depending on their contracts. For example, some companies have contracts designating them alternatively as a "Small Business" and an "Other than Small Business"; those companies are therefore counted in both categories. Totals may not add to 100% due to rounding.

Output Combining Both Research Efforts

Non-Defense Companies in High-Technology Industries: An Untapped Resource

In lieu of further empirical exploration, the following is offered as an example of what can be done practically with the concepts of high-technology industries and cosmopolitan companies. It has been noted that 75% of SAM.gov registered entities do not receive an award in any given fiscal year (U.S. Office of Management and Budget, 2023). Given the importance of experience for companies to be competitive for RFPs, even more potentially valuable for DoD officials would be those companies that have at least received non-defense awards in their past. If it were possible to identify these companies, they would represent a more viable, immediate source of "new blood" than purely commercial companies. To that end, the author generated a report of companies that are solely in non-defense work and thereby represent potential future suppliers for DoD. This kind of report can facilitate both outreach and acquisition efforts; it is probably most useful for defense policy makers interested in targeted industry engagement and for contract officers who conduct market research for new contractors to notify of bidding opportunities. It should be noted that the following is a static report, and thereby best treated as a proof of concept because its usefulness will diminish with time. If it proves useful, a dynamic version could be built using real-time SAM.gov data to accompany other market research tools available for contract officers (e.g., Government-wide Procurement Equity Tool and the Periodic Table of Acquisition Innovations (PTAI)).

Two excerpts of the report are included in Appendix 3 and Appendix 4, and the full list is available upon request from the author (length and size limitations prohibit including it in this paper). The excerpts use the four six-digit NAICS (541511, 541512, 541513, 541519) that embody the most highly concentrated high-technology NAICS: 541500 (Computer Systems Design and Related Services) and only include companies that had FY2023 contract actions to reference the most recently existing companies. The initial list produced 5,643 total companies; these would probably be the companies of greatest interest for a defense contract officer conducting market research. This list was then compared against the list of companies that have only done non-defense work for their entire contractor life to remove any companies that might have done "Both" work at some point in their past. This further narrowed the list to 3,716 companies, of which 15 were pulled at random for presentation-sake in the excerpts. In short, the excerpts include examples of companies that (a) perform work in a high-technology industry, (b) have only done non-defense work their entire federal contractor life, and (c) have operated as recently as FY2023.

The first excerpt shows the simple presence of companies, with an "X" in a given Fiscal Year indicating that the company was present with contract action(s) in USASpending.gov data. What might be of more use to contract analysts, however, is the amount of money these companies have earned each year. This is because in any given year a company may not have earned any money even though they technically exist as a contractor in the dataset. For example, if an analyst were searching for contractors in 541512 (Computer Systems Design Services), one of the three companies in the sample (Z8W6PBA8MKY4) looks more promising than the other two (HP47D5ZH4U85 and CDKHHKUY4KK6) based on the total dollars earned, even though all three companies have some experience as federal contractors.

Limitations, Contributions, and Future Research

There are at least two study limitations worth highlighting. First, for ease of data exploration only the self-reported primary NAICS from SAM.gov was used to identify the industry in which a company was doing work, and therefore to determine if it operated in a high-technology industry. This information is certainly not as comprehensive as relying on contract-level data since companies may have completed actual work in a different industry or industries. Therefore, a more thorough examination would utilize all NAICS associated with companies based on USASpending.gov contract transaction data. While this is possible given the state of development of the files used for this study, it would require an extensive amount of additional work to complete. Second, the five largest DoD companies have accounted for roughly 30% of defense contractor obligations since 2013 (Semler, 2023). These companies almost certainly do non-defense federal work as well and would therefore be represented in the cosmopolitan pool of contractors. The weighty presence of these companies in defense work might sway the results; how much so remains a matter of conjecture at this point, but it should be investigated. All the exploratory tests above should be rerun without those five companies (and their child companies) to ensure the results still hold.

The first research effort makes an important contribution by applying an established method to present day employment statistics for defining high-technology industries, which is itself an acknowledged surrogate for innovative industries. This provides a rigorous foundation for future empirical research involving the concepts of technology and, importantly, innovation. The second research effort, given its relative nascency, is more limited in its immediate practical contribution. The usefulness of the concept of cosmopolitan companies to researchers and policymakers has yet to be fully explored through empirical research. The groundwork has been laid, but evidence of its full value remains in the unknown future with potential work described below. However, as an example of practical output of the research, a report was produced of companies in high-technology industries that are exclusively doing work for non-defense agencies. These high-technology endemic companies are ideal candidates to become future DIB suppliers through DoD outreach and notification of bidding opportunities.

The most intuitive recommendation for future research is to extend the existing study by fully capitalizing on the concepts of cosmopolitan and high-technology companies. This could be done in several ways, and some of this groundwork has already been developed by the author along with a colleague at the GMU Baroni Center for Government Contracting, Olivia Letts.

In one case, we have started work to investigate the cost differential for being a defense contractor above and beyond being a contractor for other federal agencies. There is an ongoing concern that DoD regulations represent a cost premium that limits the attractiveness of being a defense contractor, but there appears to have been little empirical research on the subject since the Coopers and Lybrand (1994) study. Cosmopolitan companies that straddle the defense and non-defense government contracting worlds are uniquely positioned to provide insight into the challenges and benefits of being a defense contractor. As a head start on this work, the author has identified those companies that shifted to both defense and non-defense work since FY2019 after only previously performing endemically on non-defense contracts. This initial result of 3,867 companies that became cosmopolitan during the last five years was then further narrowed to distinguish those companies working in high-technology industries, which has resulted in a final dataset of 1,358 recently transitioned cosmopolitan high-technology companies. Future research can involve surveys and/or interviews to investigate questions like what barriers to entry were uniquely difficult to overcome, what are the perceived and real costs of becoming a defense contractor, and which DoD-specific issues exist beyond baseline federal contracting challenges.

Another extension of the present research would be to conduct additional quantitative analyses on the existing cohorts of endemic and cosmopolitan companies to address questions like whether the longevity, depth of involvement (based on obligations), or mere presence as a government supplier in prior years is predictive of becoming a DoD supplier. Of specific interest for DoD officials would be whether and how cosmopolitan high-technology companies differ from non-defense endemic ones. This knowledge would help to predict which endemic companies might be most apt to become future DIB contractors, and would facilitate targeted outreach campaigns of innovative, non-defense contractors. The list of high-technology NAICS from research effort #1 can also be leveraged for future research, examining questions like how the DoD compares to the rest of the federal government and how the ebb and flow of contractors in these critical, innovative industries has evolved in recent years. In short, all the statistics presented in this paper are descriptive; they can be readily augmented with arguably more useful predictive statistics in the future.

In conclusion, further research in this vein would be significant to the acquisition community because any differences in government-wide contracting versus defense-only contracting are likely contributors to the broader issue of attracting and retaining critical suppliers in the defense industrial base.

Appendix 1: STEM Occupations

Type of Occupation	Occupation
Research, Development, Design, or Practitioner Occupations	15-1200 Computer Occupations, except 15-1230 Computer Support Specialists and *15-1299 Computer Occupations, All Other*
	15-2000 Mathematical Science Occupations, except *15-2099 Mathematical Science Occupations, All Other*
	17-2000 Engineers
	19-1000 Life Scientists
	19-2000 Physical Scientists
	19-3000 Social Scientists and Related Workers, except 19-3093 Historians
	17-1010 Architects, Except Naval
	29-1000 Health Diagnosing and Treating Practitioners
	29-9000 Other Healthcare Practitioners and Technical Occupations, except *29-9020 Health Information Technologists and Medical Registrars* and *29-9099 Healthcare Practitioners and Technical Workers, All Other*
Technologist and Technician Occupations	15-1230 Computer Support Specialists
	15-1299 Computer Occupations, All Other
	15-2099 Mathematical Science Occupations, All Other
	17-1020 Surveyors, Cartographers, and Photogrammetrists
	17-3000 Drafters, Engineering Technicians, and Mapping Technicians
	19-4000 Life, Physical and Social Science Technicians, except 19-4060 Social Science Research Assistants
	19-4060 Social Science Research Assistants
	29-2000 Health Technologists and Technicians
	29-9020 Health Information Technologists and Medical Registrars
	29-9099 Healthcare Practitioners and Technical Workers, All Other
Postsecondary Teaching Occupations	25-1020 Math and Computer Teachers, Postsecondary
	25-1032 Engineering Teachers, Postsecondary
	25-1040 Life Sciences Teachers, Postsecondary
	25-1050 Physical Sciences Teachers, Postsecondary
	25-1060 Social Sciences Teachers, Postsecondary
	25-1031 Architecture Teachers, Postsecondary
	25-1070 Health Teachers, Postsecondary
Managerial Occupations	11-3020 Computer and Information Systems Managers
	11-9040 Architectural and Engineering Managers
	11-9120 Natural Sciences Managers
	11-9110 Medical and Health Services Managers
Sales Occupations	41-4011 Sales Representatives, Wholesale and Manufacturing, Technical and Scientific Products
	41-9030 Sales Engineers

Note. Table adapted from U.S. Bureau of Labor Statistics (2019b). SOC occupations based on 2018 SOC system. Occupations in italics are split between two occupation types.

Appendix 2: High-Technology Industries (Core=5x, Peripheral=2x)

NAICS code	Industry	STEM % of employment	High-Tech Level
541500	Computer Systems Design and Related Services	59.7	Core
541300	Architectural, Engineering, and Related Services	59.5	Core
541700	Scientific Research and Development Services	55.9	Core
513200	Software Publishers	55.8	Core
334100	Computer and Peripheral Equipment Manufacturing	49.8	Core
518200	Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services	47.2	Core
519200	Web Search Portals, Libraries, Archives, and Other Information Services	44.7	Core
334500	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	38.1	Core
334200	Communications Equipment Manufacturing	37.8	Core
334400	Semiconductor and Other Electronic Component Manufacturing	34.9	Core
334600	Manufacturing and Reproducing Magnetic and Optical Media	33.9	Core
336400	Aerospace Product and Parts Manufacturing	32.2	Peripheral
423400	Professional and Commercial Equipment and Supplies Merchant Wholesalers	29.5	Peripheral
334300	Audio and Video Equipment Manufacturing	27.1	Peripheral
521100	Monetary Authorities-Central Bank	27.0	Peripheral
325400	Pharmaceutical and Medicine Manufacturing	26.6	Peripheral
516200	Media Streaming Distribution Services, Social Networks, and Other Media Networks and Content Provid	24.7	Peripheral
517000	Telecommunications	22.5	Peripheral
333600	Engine, Turbine, and Power Transmission Equipment Manufacturing	22.0	Peripheral
333300	Commercial and Service Industry Machinery Manufacturing	20.3	Peripheral
551100	Management of Companies and Enterprises	18.9	Peripheral
211100	Oil and Gas Extraction	17.9	Peripheral
541600	Management, Scientific, and Technical Consulting Services	17.8	Peripheral
4240A2	Merchant Wholesalers, Nondurable Goods (4242 and 4246 only)	17.4	Peripheral
999100	Federal Executive Branch (OEWS Designation)	17.1	Peripheral
221100	Electric Power Generation, Transmission and Distribution	16.8	Peripheral
335300	Electrical Equipment Manufacturing	16.6	Peripheral
336500	Railroad Rolling Stock Manufacturing	16.2	Peripheral
3250A1	Chemical Manufacturing (3251, 3252, 3253, and 3259 only)	15.6	Peripheral
486100	Pipeline Transportation of Crude Oil	15.2	Peripheral
486200	Pipeline Transportation of Natural Gas	15.2	Peripheral
335900	Other Electrical Equipment and Component Manufacturing	14.6	Peripheral
524100	Insurance Carriers	14.5	Peripheral
3330A1	Machinery Manufacturing (3331, 3332, 3334, and 3339 only)	14.3	Peripheral
611300	Colleges, Universities, and Professional Schools	14.2	Peripheral
339100	Medical Equipment and Supplies Manufacturing	14.2	Peripheral
811200	Electronic and Precision Equipment Repair and Maintenance	13.5	Peripheral

Note. Core = Industries with at least five times the national average concentration in STEM occupations; Peripheral = Industries with at least two times the national average concentration in STEM occupations.

Appendix 3: Non-defense Only Companies in High-technology Industries (Work Completed)

	Industry		Company	Work Completed*				
NAICS	Name	UEI	Name	FY2019	FY2020	FY2021	FY2022	FY2023
541511	Custom Computer Programming Services	JV2VNJQP6J89	TRUESCAPE LIMITED				X	X
541511	Custom Computer Programming Services	ZMJFM8EAJ2H8	STONEMILL CONSULTING LLC			X	X	X
541511	Custom Computer Programming Services	KGHLJM1178S7	DSG SYSTEMS, INC	X	X	X	X	X
541511	Custom Computer Programming Services	PE7HCR4CV495	BIODESIGN COMPANY LIMITED		X	X		X
541511	Custom Computer Programming Services	Q8TGJ4DS3CB5	THE RIGGS GROUP, P.C.		X	X	X	X
541512	Computer Systems Design Services	Z8W6PBA8MKY4	C&T TECHNOLOGIES LLC	X	X	X	X	X
541512	Computer Systems Design Services	HP47D5ZH4U85	DATA DYNAMICS, INC.	X	X	X		X
541512	Computer Systems Design Services	CDKHHKUY4KK6	GEN3 TECHNOLOGY CONSULTING LLC			X	X	X
541513	Computer Facilities Management Services	JX95KFNVMU35	WCJ CONSULTANTS, L.L.C.	X	X	X	X	X
541513	Computer Facilities Management Services	UNXHJZDJB315	CORDYACK BRIAN					X
541513	Computer Facilities Management Services	LGNZKY4RM3U3	WAVEMARK, INC	X	X	X	X	X
541519	Other Computer Related Services	GUCBYHJYTCR5	ALEUTIANSTAR JV, LLC					X
541519	Other Computer Related Services	L5C9GFPC8LY4	PARYMON CORP			X	X	X
541519	Other Computer Related Services	GAM2KTWKMFP8	GEE WHIZ SOFTWARE, LLC			X	X	X
541519	Other Computer Related Services	JL5MFBETQJ68	THE EARNEST ANALYTICS COMPANY, INC.			X	X	X

Note. NAICS = North American Industry Classification System. UEI = Unique Entity Identifier.
* An 'X' indicates a contract action record in that Fiscal Year.

Appendix 4: Non-defense Only Companies in High-technology Industries (Dollars Obligated)

	Industry		Company	Dollars Obligated*				
NAICS	Name	UEI	Name	FY2019	FY2020	FY2021	FY2022	FY2023
541511	Custom Computer Programming Services	JV2VNJQP6J89	TRUESCAPE LIMITED	--	--	--	$827,472	$0
541511	Custom Computer Programming Services	ZMJFM8EAJ2H8	STONEMILL CONSULTING LLC	--	--	$0	$0	$0
541511	Custom Computer Programming Services	KGHLJM1178S7	DSG SYSTEMS, INC	$1,869,017	$2,955,683	$1,351,098	$705,357	$726,903
541511	Custom Computer Programming Services	PE7HCR4CV495	BIODESIGN COMPANY LIMITED	--	$47,892	-$4,512	--	$299,900
541511	Custom Computer Programming Services	Q8TGJ4DS3CB5	THE RIGGS GROUP, P.C.	--	$0	$0	$0	$0
541512	Computer Systems Design Services	Z8W6PBA8MKY4	C&T TECHNOLOGIES LLC	$250	$2,381,639	$1,434,321	$782,047	$0
541512	Computer Systems Design Services	HP47D5ZH4U85	DATA DYNAMICS, INC.	$0	$0	$0	--	$0
541512	Computer Systems Design Services	CDKHHKUY4KK6	GEN3 TECHNOLOGY CONSULTING LLC	--	--	$0	$0	$0
541513	Computer Facilities Management Services	JX95KFNVMU35	WCJ CONSULTANTS, L.L.C.	$1,496,732	$0	$1,257,600	$962,495	$2,710,269
541513	Computer Facilities Management Services	UNXHJZDJB315	CORDYACK BRIAN	--	--	--	--	$52,700
541513	Computer Facilities Management Services	LGNZKY4RM3U3	WAVEMARK, INC	$7,223,695	$4,231,608	$4,174,374	$4,704,730	$3,873,354
541519	Other Computer Related Services	GUCBYHJYTCR5	ALEUTIANSTAR JV, LLC	--	--	--	--	$1,079,058
541519	Other Computer Related Services	L5C9GFPC8LY4	PARYMON CORP	--	--	$0	$0	$0
541519	Other Computer Related Services	GAM2KTWKMFP8	GEE WHIZ SOFTWARE, LLC	--	--	$205,702	$56,090	$105,316
541519	Other Computer Related Services	JL5MFBETQJ68	THE EARNEST ANALYTICS COMPANY, INC.	--	--	$50,000	$47,500	$217,750

Note. NAICS = North American Industry Classification System. UEI = Unique Entity Identifier.

* A '$0' indicates a contract action, but $0 obligated dollars. An '--' indicates no contract action record in that Fiscal Year.

References

Bresler, A., & Bresler, A. (2020, May 13). The effect of defense-sponsored innovation programs on the military's industrial base. Acquisition Research: Creating Synergy for Informed Change. Seventeenth Annual Acquisition Research Symposium. https://dair.nps.edu/handle/123456789/4208

Carril, R., & Duggan, M. (2020). The impact of industry consolidation on government procurement: Evidence from Department of Defense contracting. Journal of Public Economics, 184, 104141. https://doi.org/10.1016/j.jpubeco.2020.104141

Congressional Budget Office. (n.d.). *Defense and national security: Defense budget*. Retrieved March 26, 2025, from https://www.cbo.gov/topics/defense-and-national-security/defense-budget

Congressional Budget Office. (n.d.). Defense and national security: Defense budget. Retrieved March 26, 2025, from https://www.cbo.gov/topics/defense-and-national-security/defense-budget

Congressional Research Service. (2023). Defense primer: U.S. defense industrial base (IF10548). https://crsreports.congress.gov/product/pdf/IF/IF10548

Congressional Research Service. (2025). The defense innovation ecosystem (IF12869). https://crsreports.congress.gov/product/pdf/IF/IF12869

Coopers & Lybrand, & TASC Project Team. (1994). The Department of Defense (DoD) regulatory cost premium: A quantitative assessment. https://dair.nps.edu/handle/123456789/3000

Cover, B., Jones, J. I., & Watson, A. (2011). Science, technology, engineering, and mathematics (STEM) occupations: A visual essay. Monthly Labor Review, 3–15.

Defense Innovation Board. (n.d.). About: Our story. Retrieved March 26, 2025, from https://innovation.defense.gov/About1/

DoD. (n.d.). Spotlight: DOD innovates. Retrieved March 26, 2025, from https://www.defense.gov/Spotlights/DOD-Innovates/

Downs Jr., G. W., & Mohr, L. B. (1976). Conceptual issues in the study of innovation. Administrative Science Quarterly, 21(4), 700–714. JSTOR. https://doi.org/10.2307/2391725

Economic and Labor Market Information Bureau, New Hampshire Employment Security. (n.d.). High-tech industries by NAICS. Retrieved October 2, 2024, from https://www.nhes.nh.gov/elmi/statistics/documents/high-tech-titles.pdf

Fayer, S., Lacey, A., & Watson, A. (2017). STEM occupations: Past, present, and future. Spotlight on Statistics, 1–35.

Hecker, D. E. (1999). High-technology employment: A broader view. Monthly Labor Review, 18–28.

Hecker, D. E. (2005). High-technology employment: A NAICS-based update. Monthly Labor Review, 57–72.

Hyatt, E., & Everhart, L. E. (in press). The "shrinking" defense industrial base: A survey of former DoD prime contractors. Defense Acquisition Research Journal.

Jaskyte, K. (2011). Predictors of administrative and technological innovations in nonprofit organizations. Public Administration Review, 71(1), 77–86. https://doi.org/10.1111/j.1540-6210.2010.02308.x

Josephson, B. W., Lee, J.-Y., Mariadoss, B. J., & Johnson, J. L. (2019). Uncle Sam rising: Performance implications of business-to-government relationships. Journal of Marketing, 83(1), Article 1. https://doi.org/10.1177/0022242918814254

Kahn, K. B. (2018). Understanding innovation. Business Horizons, 61(3), 453–460. https://doi.org/10.1016/j.bushor.2018.01.011

National Science Foundation. (1989). Science and technology resources in U.S. industry (Special report NSF 88–321). https://catalog.hathitrust.org/Record/103161139

Porter, M. E., & Stern, S. (2001). National innovative capacity. In The global competitiveness report 2001– 2002 (pp. 2–18). Oxford University Press.

Semler, S. (2023, September 22). Top five weapons companies set to get $140 billion from pentagon next year. https://readsludge.com/2023/09/22/big-five-contractors-get-one-sixth-of-pentagon-budget-analysis/

United States General Accounting Office. (2025). Defense innovation unit: Actions needed to assess progress and further enhance collaboration (GAO-25-106856). https://www.gao.gov/assets/gao-25-106856.pdf

U.S. Bureau of Labor Statistics. (2019a). Options for defining STEM (Science, Technology, Engineering, and Mathematics) occupations under the 2018 Standard Occupational Classification (SOC) system—Attachment A. https://www.bls.gov/soc/attachment_a_stem_2018.pdf

U.S. Bureau of Labor Statistics. (2019b). Options for defining STEM (Science, Technology, Engineering, and Mathematics) occupations under the 2018 Standard Occupational Classification (SOC) system—Attachment B. https://www.bls.gov/soc/attachment_b_stem_2018.pdf

U.S. Bureau of Labor Statistics. (2024a). STEM data, May 2023 [Dataset]. https://www.bls.gov/oes/2023/may/stem_2023.xlsx

U.S. Bureau of Labor Statistics. (2024b, August 29). Employment in STEM occupations. https://www.bls.gov/emp/tables/stem-employment.htm

U.S. Census Bureau. (2022). NAICS, 2022 [Dataset]. https://www.census.gov/naics/

U.S. Office of Management and Budget. (2023). Memorandum for the heads of executive departments and agencies (M-23–11). Executive Office of the President. https://www.whitehouse.gov/wp-content/uploads/2023/02/M-23-11-Creating-a-More-Diverse-and-Resilient-Federal-Marketplace.pdf

Wikipedia. (2024, December 22). Cosmopolitan distribution. https://en.wikipedia.org/wiki/Cosmopolitan_distribution

Wolf, M., & Terrell, D. (2016). The high-tech industry, what is it and why it matters to our economic future (Vol. 5, No. 8; Beyond the Numbers: Employment & Unemployment). U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-5/the-high-tech-industry-what-is-it-and-why-it-matters-to-our-economic-future.htm

Workforce Information Council. (2015). High-tech industries in the U.S. economy (pp. 1–27). https://www.bls.gov/advisory/bloc/high-tech-industries.pdf

World Intellectual Property Organization. (2024). Global innovation index [Dataset]. https://www.wipo.int/en/web/global-innovation-index

Contract Management Competency in the USMC: An Assessment of MCSC, ECP/RCOs, and LOGCOM

Dr. Rene G. Rendon— is an associate professor and acquisition management area chair in the Department of Defense Management at the Naval Postgraduate School (NPS). A retired Air Force contracting officer, he served as a warranted contracting officer for major weapons system programs such as the Peacekeeper ICBM, F-22 Advanced Tactical Fighter, Space-Based Infrared Satellite program, and the Evolved Expendable Launch Vehicle program. He also served as a contracting squadron commander for an Air Force pilot training base. He was presented with the Air Force Outstanding Officer in Contracting Award, and he has received the NPS Hamming Teaching Excellence Award and the Graduate School of Business and Public Policy Research Excellence Award. [rgrendon@nps.edu]

Abstract

In April 2020, the DoD senior procurement executives established a new contracting competency model and a single, entry-level certification program for the DoD contracting workforce. The new competency model is based on the National Contract Management Association (NCMA) Contract Management Standard (CMS). This new DoD contracting competency model serves as the basis for training the contracting workforce in the new DAWIA Back-to-Basics certification. The competency model can also be used as the basis for assessing the contracting workforce's contract management competency. The purpose of this research is to present the findings of three competency assessments using the new DoD contracting competency framework. The competency assessments were conducted on the contracting workforce at the Marine Corps Systems Command (MCSC), the Marine Corps Expeditionary Contracting Platoons and Regional Contracting Offices (ECP/RCO), and the Marine Corps Logistics Command (LOGCOM). The research seeks to identify any gaps in contract management proficiency and knowledge, and to provide the USMC contracting leadership with recommendations to fill these gaps. This research will answer the question: How do the competency assessment results compare across the MCSC, the ECP/RCOs, and the LOGCOM? Based on the competency assessment results, recommendations for competency development are provided to the assessed organization.

Introduction

In the 2020 National Defense Authorization Act (NDAA; 2019), Congress directed the secretary of defense to implement a professional certification program for all members of the acquisition workforce that is based on standards developed by a third-party accredited program based on nationally or internationally recognized standards (NDAA, 2019). In September 2020, the undersecretary of defense for acquisition and sustainment (USD A&S) implemented the Back-to-Basics (BtB) talent management program to be fully deployed by October 1, 2021 (OUSD[A&S], 2020). This would be a major change to the acquisition certification program established by the Defense Acquisition Workforce Improvement Act (DAWIA) and enacted by Congress in 1990.

In February 2021, the office of the undersecretary of defense (OUSD) principal director for defense pricing and contracting (DPC) published a memorandum restructuring the DoD Contracting Professional Certification Program and Contracting Competency Model. The new contracting competency model would be based on the American National Standards Institute (ANSI)/National Contract Management Association (NCMA) accredited Contract Management Standard (CMS; OUSD[A&S], 2021). This new contracting workforce competency model complies with the 2020 NDAA (2019) requirement to base a professional certification on standards developed by a third-party accredited program (OUSD[A&S], 2021).

Purpose of Research

Given the backdrop of the congressional legislation and the establishment of the new contracting workforce competency model, the purpose of this research is to present the findings of three competency assessments using the new DoD contracting competency framework. Competency assessments were conducted on the contracting workforce at the Marine Corps Systems Command (MCSC), the Marine Corps Expeditionary Contracting Platoons and Regional Contracting Offices (ECP/RCO), and the Marine Corps Logistics Command (LOGCOM). The research seeks to compare the results of the competency assessments and identify any consistencies and patterns in contract management competency across the three organizations. This research will also identify gaps in contract management proficiency and knowledge across these organizations, and provide the USMC contracting leadership with recommendations to fill these proficiency and knowledge gaps. This research will answer the following question: How do the competency assessment results compare across the MCSC, the ECP/RCOs, and the LOGCOM? Across these three organizations, in which contract management competencies are the assessed workforce less proficient and less knowledgeable? Based on the competency assessment results, recommendations for competency development are provided to the USMC contracting leadership.

Methodology

The methodology for this research consists of comparing the results of three previous competency assessments. The three competency assessments were conducted by Hayashi and Pfannenstiel on the MCSC (2021), Hoover on the ECP/RCOs (2021), and Bute on the LOGCOM (2024). These assessments were conducted using the Contract Management Competency Assessment instrument developed by Rendon (2021). This assessment instrument has also been used on Army and Air Force contracting organizations (for example, see Davies et al., 2021; Moyer et al., 2020; Powell, 2021).

DoD Contract Management Workforce Competency Model

The new DoD contracting workforce competency model, based on the NCMA CMS, is significantly different from the legacy DoD contracting competency model in both structure and scope, and thus provides an innovative approach for talent and competency management (Rendon, 2019; Rendon & Winn, 2017). The top-level structure of the NCMA CMS is reflected in Figure 1 (NCMA, 2019b). The CMS domains (e.g., Develop Solicitation, Develop Offer, …) and processes (e.g., Plan Solicitation, Plan Sales, …) are the foundation for the competency assessment instrument.

The CMS's concise and detailed contract life cycle and greater emphasis and granularity in each of the life-cycle phases and job tasks may help develop and fortify the DoD's contracting processes and practices, as well as the training of its contracting workforce on these competencies. Providing greater emphasis on each of the contract life-cycle phases and also structuring the competencies using a hierarchical approach that aligns each competency with processes, tasks, and subtasks will support the development of a professional contracting career path that associates contracting technical competencies and key work experiences (Rendon, 2019). The CMS also has an overarching narrative of guiding principles aligned with professional competencies that apply across all phases of the contracting life cycle.

Additionally, the CMS uses contract management terms that are relevant and applicable across the DoD, federal agencies, and industry.

In terms of scope, the CMS differs from the legacy DoD contracting competency model in that the CMS also includes the industry (seller) competencies, processes, and job tasks. Expanding the DoD's contracting workforce knowledge to include industry's side of contracting (e.g., industry operations and processes) as reflected in the CMS will help in developing technical and professional skills that can transfer across government and industry, as well as improve communication and collaboration between government and industry. Including the industry side of contracting would also result in strengthening systems thinking within the DoD contracting workforce (Carlson, 2017). Contracting officers applying systems thinking to contract management will know that "problems can have hidden, indirect causes" and it is the "relationships among the parts that matter the most" (Carlson, 2017). Using systems thinking, contracting officers will be able to "see the gaps where complications or opportunities can arise" within the acquisition process and understand how their contracting decisions may impact contractors and subcontractors (Carlson, 2017).

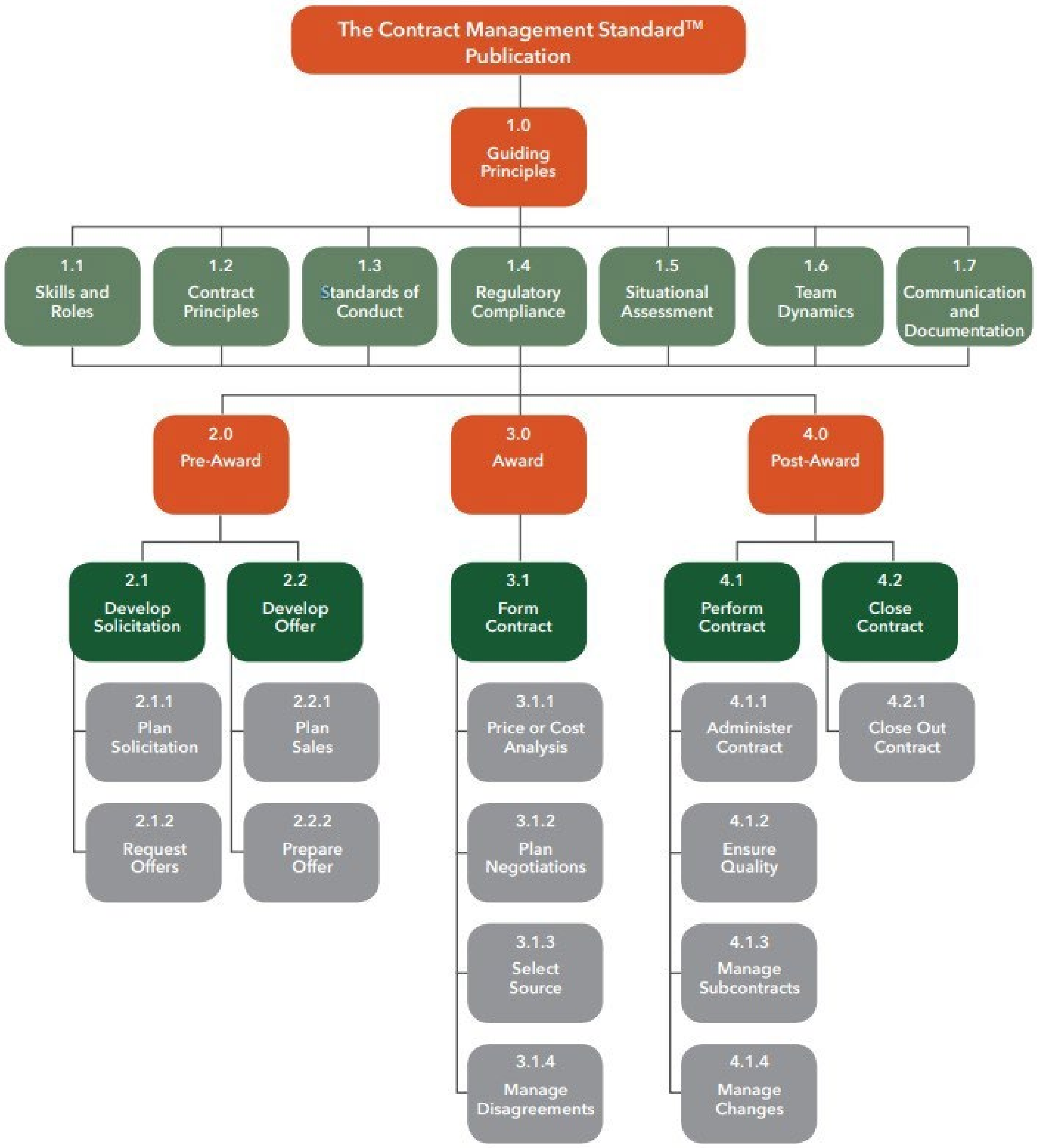


Figure 1. NCMA Contract Management Standard. NCMA (2019b).

Additionally, adopting the CMS competency framework may provide the DoD contracting workforce with a stronger foundational understanding of not only the complete contract life cycle, but also the different perspectives in contract relationships (e.g., buyer, seller, subcontractor, supplier, end users, etc.). This understanding of different perspectives may enable DoD contracting officers to introduce innovation and process change into the DoD contracting processes.

Finally, providing training on the seller-side competencies to the DoD contracting workforce may also strengthen "communication, collaboration, problem-solving, and adaptability" skills (Carlson, 2017). A recent RAND study found that within the defense acquisition workforce, knowledge gaps in business acumen, industry operations, and industry motivation exist (Werber et al., 2019). The RAND report also found that the lack of standardized definitions and competency model formats obscures the need for knowledge related to business acumen, industry operations, and industry motivation (Werber et al., 2019).

Structure of Competency Assessment Instrument

The structure of the contracting competency assessment instrument consists of contracting competency statements for each of the contract management phases (pre-award, award, post-award), as well as from both buyer and seller contracting perspectives. More specifically, the contracting competency statements reflect the contracting competencies and the specific job tasks for each contract management phase and for each perspective as reflected in the CMS. The competency statements will be rated by the contracting workforce members using a Likert scale reflecting different levels of proficiency for performing the buyer job tasks and a different Likert scale reflecting the different levels of knowledge of the seller job tasks. The proficiency rating levels for performing buyer job tasks are identified and defined as follows:

1. Aware: Applies the competency in the simplest situations and requires close and extensive guidance.
2. Basic: Applies the competency in somewhat difficult situations and requires frequent guidance.
3. Intermediate: Applies the competency in difficult situations and requires little or no guidance.
4. Advanced: Applies the competency in considerably difficult situations and generally requires no guidance.
5. Expert: Applies the competency in exceptionally difficult situations, serves as a key resource, and advises others.
6. N/A: Not applicable/not needed in my job.

The knowledge rating levels for understanding seller job tasks are identified and defined as follows:

1. None: I am not aware of this Contractor competency.
2. Aware: I am aware but have no knowledge of this Contractor competency.
3. Basic: I have some basic-level knowledge of this Contractor competency.
4. Intermediate: I have intermediate-level knowledge of this Contractor competency.
5. Advanced: I have advanced-level knowledge of this Contractor competency.

Deployment of Competency Assessment Instrument

The competency assessment instrument link was deployed to the Marine Corps Systems Command in 2020, the Marine Corps Expeditionary Contracting Platoons (ECPs) and the co-located Regional Contracting Offices (RCOs) in 2021, and the Marine Corps Logistics Command in 2024. For additional and specific information about those assessments, please refer to Hayashi and Pfannenstiel (2020), Hoover (2021), and Bute (2024).

The competency assessment instrument was deployed using the Naval Postgraduate School (NPS) Qualtrics survey tool. The web-based survey tool allows participants to respond anonymously to the self-assessment items.

Findings

Our assessment findings are presented in terms of demographics, proficiency in performing buyer tasks, and knowledge of seller tasks. Tables 1, 2, and 3 provide the demographic data for each of the assessed organizations.

Table 1. MCSC Workforce Competency Assessment Demographics

DAWIA Level		CM Years of Experience	
None	1	3 or Less	5
Level I	3	4 to 8	5
Level II	5	9 to 13	21
Level III	41	14 to 18	4
		19 or more	17
PCO	21		
		Years in Organization	
Professional Certifications		3 or Less	20
CFCM	2	4 to 8	10
CCCM	0	9 to 13	13
CPCM	1	14 to 18	4
Other	7	19 or more	4

Table 2. ECP/RCO Workforce Competency Assessment Demographics

CM Years of Experience		Years in Organization	
3 or Less	15	1 or Less	15
4 to 8	19	1 to 2	6
9 to 13	5	2 to 3	6
14 to 18	2	3 or More	12
19 or More	0		
		PCO	22
DAWIA Level Certification		**Other Professional Certifications**	
None	8	CFCM	0
Level I	11	CCCM	0
Level II	16	CPCM	0
Level III	6	Other	0

Table 3. LOGCOM Contracting Workforce Competency Assessment Demographics

CM Years of Experience		Years in Organization	
3 or less	5	3 or less	7
4 to 8	1	4 to 8	3
9 to 13	1	9 to 13	1
14 to 18	3	14 to 18	1
19 or more	5	19 or more	3
		PCO	6
DAWIA BtB Certification			
None	5	Professional Certifications	
Contracting Professional	10	None	13
		CFCM	0
		CCCM	0
		CCPM	0
		Other	2

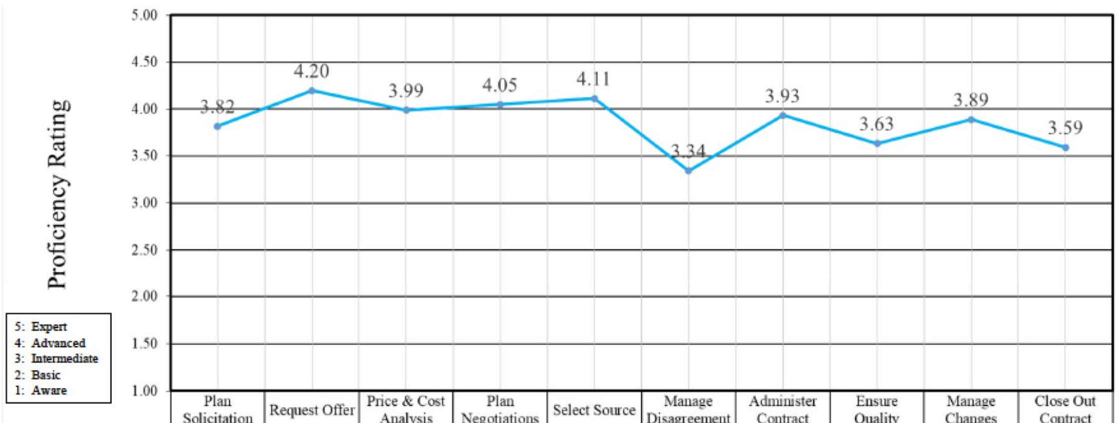
As reflected in Table 1, approximately 51 of the 220 MCSC potential participants responded to the demographic questions, resulting in a response rate of 23%. The demographics for the MCSC contracting workforce indicate a highly educated, trained and experienced workforce with 41 respondents reported being DAWIA Level III Contracting. Additionally, 21 respondents indicated that they are Procuring Contracting Officers (PCO), meaning that they hold warrants from MCSC to award contracts on behalf of the United States government.

As reflected in Table 2, approximately 41 of the 100 ECP/RCO potential participants responded to the demographic questions, resulting in a response rate of 41%. The demographics for the ECP/RCO contracting workforce indicate a less educated, trained and experienced workforce. Almost half of the respondents either had no DAWIA certification or were certified at Level 1. Additionally, the majority of the respondents (83%) had between 0 and 8 years of contracting experience.

As reflected in Table 3, approximately 15 of the 28 LOGCOM potential participants responded to the demographic questions, resulting in a response rate of 54%. The demographics for the LOGCOM contracting workforce indicate a mid-level educated, trained and experienced workforce with 10 respondents reported being DAWIA Back-to Basics certified as a Contracting Professional, and the remainder of the respondents were not certified. Additionally, about half of the respondents (60%) had 9 years or more contracting experience, with 6 respondents indicated that they are Procuring Contracting Officers (PCO).

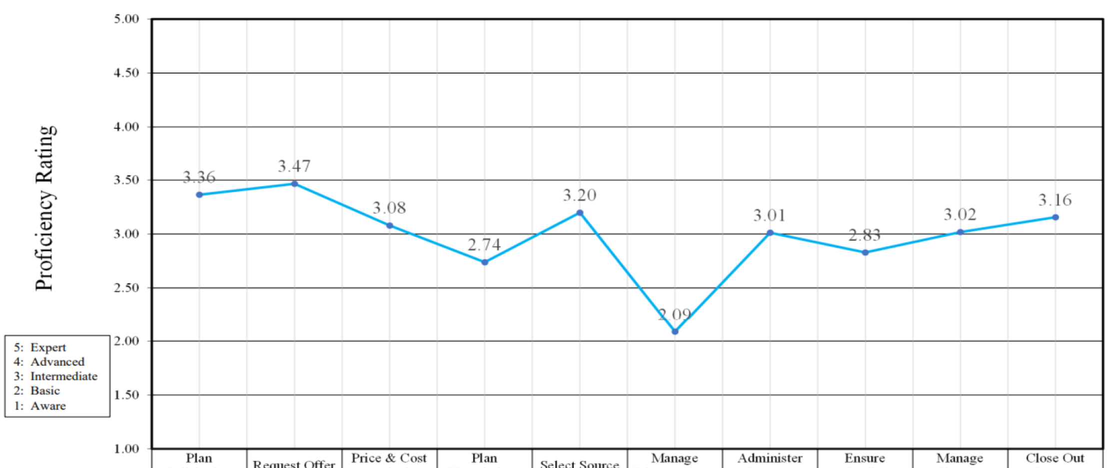
Buyer Proficiency Levels

Figures 1, 2, and 3 reflect the assessment results of the Buyer Proficiency component of the competency assessment for each organization. The figures reflect the buyer competencies (e.g., Plan Solicitation, Request Offer, etc.) that include buyer associated job tasks, as reflected in the NCMA CMS. Also reflected in these figures are the average proficiency ratings for each competency, based on the buyer proficiency rating scales discussed earlier.



	Plan Solicitation	Request Offer	Price & Cost Analysis	Plan Negotiations	Select Source	Manage Disagreement	Administer Contract	Ensure Quality	Manage Changes	Close Out Contract
Proficiency Rating	3.82	4.20	3.99	4.05	4.11	3.34	3.93	3.63	3.89	3.59

Figure 1. MCSC Contracting Workforce Competency Assessment: Buyer Proficiency



	Plan Solicitation	Request Offer	Price & Cost Analysis	Plan Negotiations	Select Source	Manage Disagreement	Administer Contract	Ensure Quality	Manage Changes	Close Out Contract
Proficiency Rating	3.36	3.47	3.08	2.74	3.20	2.09	3.01	2.83	3.02	3.16

Figure 2. ECP/RCO Contracting Workforce Competency Assessment: Buyer Proficiency

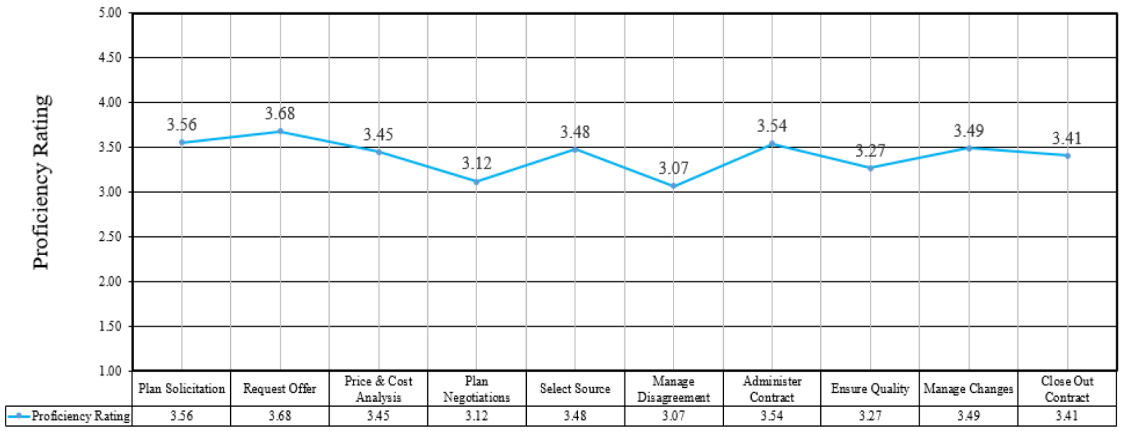


Figure 3. LOGCOM Contracting Workforce Competency Assessment: Buyer Proficiency

As can be seen in Figure 1, for MCSC, the average buyer proficiency ratings ranged from Intermediate to Advanced proficiency levels. Specifically, the lowest average proficiency rating was 3.34 (Intermediate) for Manage Disagreement, and the highest average proficiency rating was 4.2 (Advanced) for Request Offer.

As can be seen in Figure 2, for ECP/RCO, the average buyer proficiency ratings ranged from Basic to Intermediate proficiency levels. Specifically, the lowest average proficiency rating was 2.09 (Basic) for Manage Disagreement, and the highest average proficiency rating was 3.47 (Intermediate) for Request Offer.

As can be seen in Figure 3, for LOGCOM, the average buyer proficiency ratings ranged within the Intermediate proficiency level. Specifically, the lowest average proficiency rating was 3.07 (Intermediate) for Manage Disagreement, and the highest average proficiency rating was 3.68 (Intermediate) for Request Offer.

Seller Knowledge Levels

Figures 4, 5, and 6 reflect the assessment results of the Seller Knowledge component of the competency assessment for each organization. The figures reflect the seller competencies (e.g., Plan Sales, Prepare Offer, etc.) that include seller associated job tasks, as reflected in the NCMA CMS. Also reflected in these figures are the average knowledge ratings for each competency, based on the seller knowledge rating scales discussed earlier.

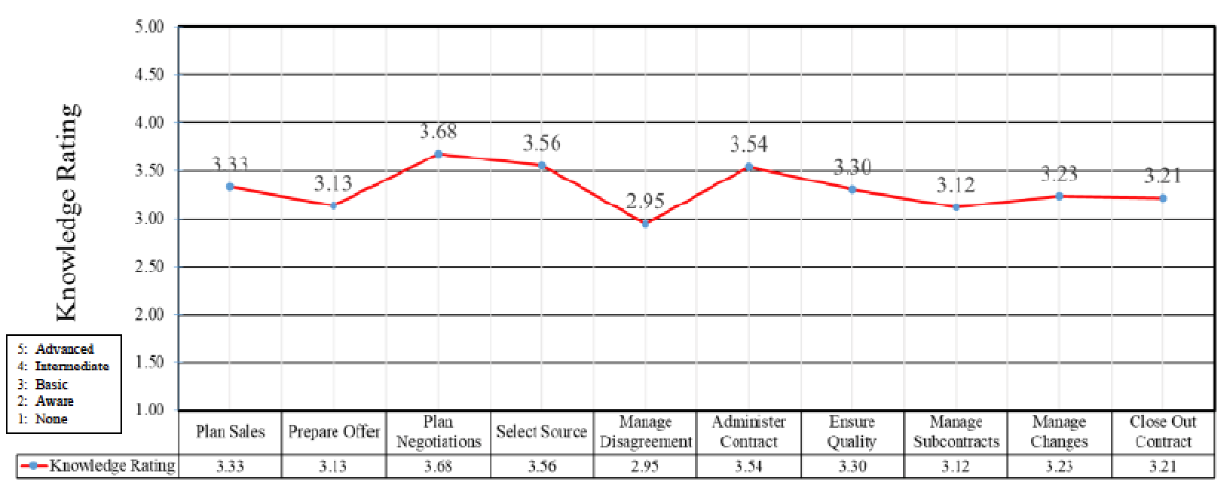


Figure 4. MCSC Contracting Workforce Competency Assessment: Seller Knowledge

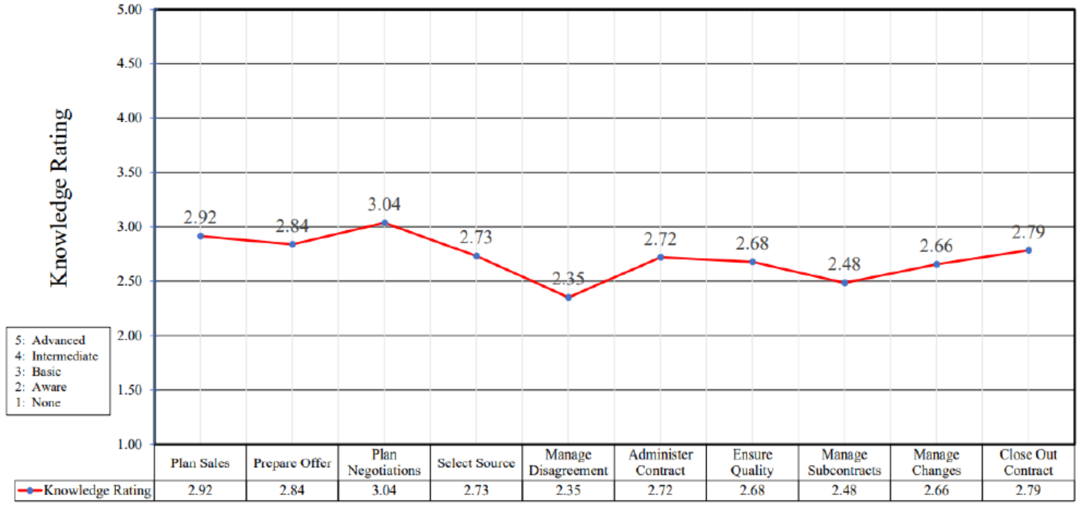


Figure 5. ECP/RCO Contracting Workforce Competency Assessment: Seller Knowledge

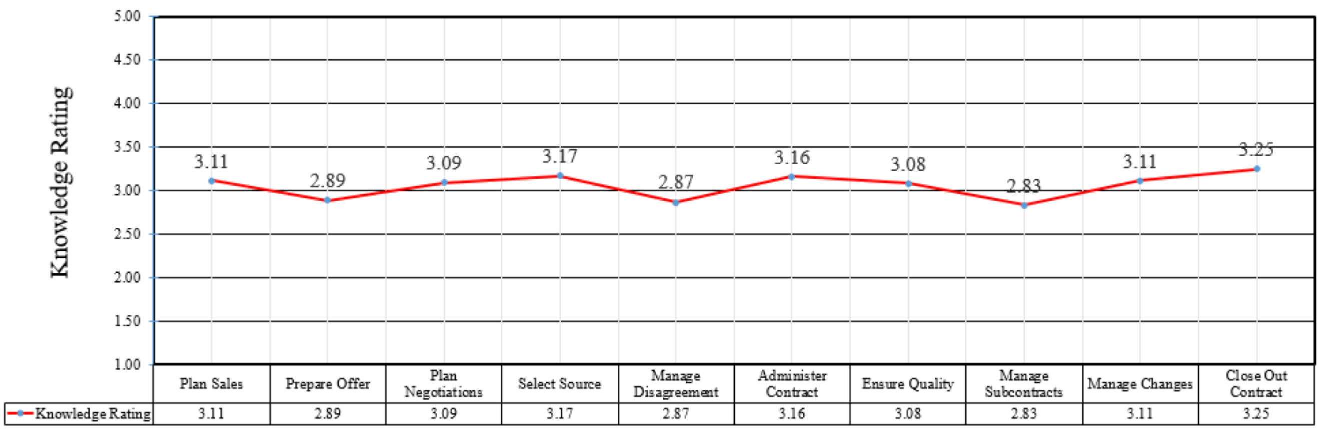


Figure 6. LOGCOM Contracting Workforce Competency Assessment: Seller Knowledge

As can be seen in Figure 4, for MCSC, the average seller knowledge ratings ranged from Aware to Basic knowledge levels. Specifically, the lowest average knowledge rating was 2.95 (Aware) for Manage Disagreement, and the highest average knowledge rating was 3.68 (Basic) for Plan Negotiations.

As can be seen in Figure 5, for ECP/RCO, the average seller knowledge ratings ranged from Aware to Basic knowledge levels. Specifically, the lowest average knowledge rating was 2.35 (Aware) for Manage Disagreement, and the highest average knowledge rating was 3.04 (Basic) for Plan Negotiations.

As can be seen in Figure 6, for LOGCOM, the average seller knowledge ratings ranged from Aware to Basic knowledge levels. Specifically, the lowest average knowledge rating was 2.83 (Aware) for Manage Subcontracts, and the highest average knowledge rating was 3.25 (Basic) for Close Out Contract.

Discussion of Findings

Buyer Proficiency Competencies

The findings from the comparison of the three Marine Corps contracting workforce competency assessments indicate that the organizations' average competency levels for the buyer proficiency tasks are rated higher than the organizations' average knowledge levels of seller tasks.

Specifically, based on the competency assessments across all of the organizations, the majority of the buyer proficiency competency ratings are at the Intermediate level. Only three competencies were rated at the Advanced level (Request Offer, Plan Negotiations and Select Source, all at MCSC) and only three competencies were rated at the Basic level (Plan Negotiations, Manage Disagreement and Ensure Quality, all at ECP/RCO). All other buyer proficiency competencies were rated at the Intermediate level.

Additionally, across all three organizations, the Request Offer competency was consistently the highest rated buyer proficiency, whereas the Manage Disagreement competency was consistently the lowest rated buyer proficiency competency.

Finally, in terms of the contracting life cycle, for MCSC and ECP/RCO, the pre-award buyer proficiency competency ratings are higher than the award and post award buyer proficiency competency ratings. The buyer proficiency competency ratings for LOGCOM seem to be consistent throughout the contract life cycle.

Seller Knowledge Competencies

As previously stated, the findings from the comparison of the three Marine Corps contracting workforce competency assessments indicate that the organizations' levels for the seller knowledge competencies are rated lower than the organizations' buyer proficiency tasks.

Specifically, based on the competency assessments across all of the organizations, all of the seller knowledge competency ratings are at the Aware or Basic levels. At ECP/RCO, all of the seller knowledge ratings are at the Aware level except for one, Plan Negotiations, which is rated at the Basic level. At MCSC, all of the seller knowledge competencies are at the Basic level, except for one, Manage Disagreement, which is rated at the Aware level. At LOGCOM, three of the seller knowledge competencies (Prepare Offer, Manage Disagreement, and Manage Subcontracts) were rated at the Aware level, with the remainder competencies rated at the Basic level.

Additionally, for MCSC and ECP/RCO, the Plan Negotiations competency was the highest rated seller knowledge competency, with Close Out Contract the highest for LOGCOM. The Manage Disagreement competency was the lowest rated seller knowledge competency for MCSC and ECP/RCO, with Manage Subcontracts the lowest seller knowledge competency for LOGCOM.

Finally, in terms of the contracting life cycle, across the organizations a distinct pattern did not appear in terms of which life-cycle phases were the highest or lowest in terms of average seller knowledge ratings. For MCSC and ECP/RCO, the findings generally indicate that within each contract life-cycle phase, the seller knowledge ratings start high for the first competency in that phase (e.g., Plan Sales, Plan Negotiations, Administer Contract) and then decrease in the later competencies within that phase. The findings indicate that the seller knowledge competency ratings for LOGCOM are generally consistent throughout the contract life cycle.

The higher buyer proficiency competency ratings for MCSC may be related to the demographics of the MCSC contracting workforce. As reflected in Table 1, the demographics for the MCSC contracting workforce indicate a highly educated, trained, and experienced workforce with 41 respondents reported being DAWIA Level III Contracting and 21 respondents indicated that they are Procuring Contracting Officers (PCOs).

The lower buyer proficiency competency ratings for ECP/PCO may also be related to the demographics of this contracting workforce. As reflected in Table 2, the demographics for the ECP/RCO contracting workforce indicate a less educated, trained, and experienced workforce. Almost half of the respondents either had no DAWIA certification or were certified at Level 1. Additionally, the majority of the respondents (83%) had between 0 and 8 years of contracting experience. This level of training and experience may indicate a lower competency level in performing the buyer tasks reflected in the CMS.

The higher buyer proficiency competency ratings compared to the lower seller knowledge ratings may reflect the scope and focus of the contracts training received by the DoD acquisition workforce. The contracts training provided by the Defense Acquisition University (DAU) and based on the previous DoD contracting competency framework reflects only the buyer processes and related tasks, as reflected in the Federal Acquisition Regulation (FAR). The legacy DAU contracts training courses do not cover the seller (industry) processes and related tasks. (See Rendon and Winn [2017] for a comparison of the previous DoD contracting competency model and the NCMA Contract Management Standard).

Finally, the general consistency in the lower buyer proficiency and seller knowledge ratings for the Manage Disagreement competency is indeed an interesting finding. This CMS competency specifically deals with the seller tasks of submitting protests and appeals and the buyer tasks of responding to protests and appeals. The low buyer proficiency and knowledge ratings from the assessed contracting workforce in this competency area may reflect a deficiency in the knowledge, skills, and abilities related to these contract management tasks.

Recommendations for Competency Development

Based on the comparison of these competency assessments across the MCSC, ECP/RCO, and LOGCOM, the following recommendations to the USMC for competency development are provided. These recommendations can be used by the USMC for developing a training roadmap for targeting buyer task proficiency and seller knowledge areas needed for improvement within the contracting workforce.

The first recommendation is to incorporate training to increase knowledge of the CMS seller competencies and related job tasks (NCMA, 2019b). The assessment results reflect that the knowledge ratings of the seller competencies are lower than the buyer proficiency ratings. The recommendation is to incorporate the seller competencies and job tasks from the CMS for all the contract life-cycle competencies (NCMA, 2019a) into the required training courses. Development of this training module could start by incorporating information from the Contract Management Body of Knowledge (CMBOK) sections 4.0, 5.0, and 6.0 (NCMA, 2019a).

The second recommendation is to emphasize training on the CMS buyer competencies and related job tasks that were rated at the Aware and Basic levels. The job tasks that were rated at the Aware level indicate that the workforce can apply the competency in the simplest situations and requires close and extensive guidance. Competencies rated at the Basic level indicate that the workforce can apply the competency in somewhat difficult situations and requires frequent guidance. Thus, the recommendation is for this training to emphasize buyer competencies and job tasks from the CMS for all of the contract life-cycle competencies that were rated at the Aware and Basic level (NCMA, 2019a).

The third recommendation for the assessed organizations is to develop and/or improve the contracting workforce training on the Manage Disagreement competency. The assessment results reflect that the Manage Disagreement competency and related job tasks within the Award phase had the lowest scores for the buyer proficiency competencies and generally the seller knowledge competency. Development of this training module could start by providing training to improve skills such as critical thinking, problem solving, and decision-making related to managing contract disagreements, as well as resolving protests and appeals.

Conclusion

The purpose of this research was to present the findings of three competency assessments using the new DoD contracting competency framework. Competency assessments were conducted on the contracting workforce at the Marine Corps Systems Command (MCSC), the Marine Corps Expeditionary Contracting Platoons and Regional Contracting Offices (ECP/RCO), and the Marine Corps Logistics Command (LOGCOM). The research compared the results of the competency assessments and identified any consistencies and patterns in contract management competency across the three organizations. This research also identified gaps in contract management buyer proficiency and seller knowledge across these organizations and provided the USMC contracting leadership with recommendations to fill these proficiency and knowledge gaps. Based on the competency assessment results, recommendations for competency development were provided to the USMC contracting leadership. Based on the research findings, the USMC can develop a training roadmap for targeting competencies and knowledge areas needed for improvement within the contracting workforce.

Areas for Further Research

The primary area for further research is to conduct a follow-on competency assessment of the three Marine Corps organizations after the contracting workforce has received the recommended training based on the initial assessment. This follow-on assessment would measure any increased learning, in terms of buyer proficiency and seller knowledge of both the buyer and seller competencies as reflected in the CMS.

A second area for further research is to conduct workforce competency assessments on additional contracting organizations throughout the DoD. This would enable benchmarking workforce competency assessment data from DoD activities with diverse contracting mission sets. Conducting competency assessments at Air Force and Army contracting organizations may provide insight and patterns on buyer proficiency and seller knowledge levels that could further inform the DoD contract training programs for these organizations.

References

Bute, M. (2024). *Analysis of the Marine Corps Logistics Command contracting workforce competency assessment* (Master's thesis, Naval Postgraduate School).

Carlson, S. (2017). A new liberal art: How systems thinking prepares students for a complex world. *The Chronicle of Higher Education*, *64*(5). https://www.chronicle.com/article/a-new-liberal-art/

Davies, J. N., Markelz, D., & Rostermundt, S. A. (2021). *Analysis of Army contracting workforce competency assessment* (Master's thesis, Naval Postgraduate School).

Hayashi, S., & Pfannenstiel, A. (2021). *Analysis of Marine Corps Systems Command contracting workforce competency assessment* (Master's thesis, Naval Postgraduate School).

Hoover, B. (2021). *Analysis of the Marine Corps expeditionary contracting workforce competency assessment* (Master's thesis, Naval Postgraduate School).

Moyer, D., Walls, K. E., & Phillips, C. V. (2020). *An analysis of Air Force contract management personnel competency and internal processes using the National Contract Management Association's third-party accredited competency standard* (Master's thesis, Naval Postgraduate School).

National Contract Management Association. (2019a). *Contract management body of knowledge* (6th ed.).

National Contract Management Association. (2019b). *Contract management standard* (2nd ed.).

https://www.ncmahq.org/Web/Standards---Practices/ContractManagement-Standard-Publication.aspx

National Defense Authorization Act for Fiscal Year 2020, Pub. L. No. 116–92, 133 Stat. 1198

(2019). https://www.govinfo.gov/content/pkg/PLAW116publ92/html/PLAW-116publ92.htm

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020, September 2). *"Back-to-basics" for the defense acquisition workforce* [Memorandum]. https://4edacm.dau.edu/assets/Back_to_Basics_Memo_2_Sep_2020.pdf

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2021, February 17). *Restructuring of the certification program for the contracting functional area* [Memorandum]. https://dair.nps.edu/handle/123456789/4487.

Powell, Richard W. (2021). *Analysis of NGB enterprise contract management competencies*. (Master's thesis, Naval Postgraduate School).

Rendon, R. G. (2019). Enhancing professional and technical excellence: Analysis of contract management competency models (SYM-AM-19-036). *Proceedings of the 16th Annual Acquisition Research Symposium*, 86–112. https://calhoun.nps.edu/handle/10945/63019

Rendon, R. G., & Winn, T. (2017). Competency in contract management: A comparison of DoD and CMBOK competency models. *Contract Management*, *57*(12), 66–81.

Werber, L., Ausink, J., Daugherty, L., Phillips, B., Knutson, F., & Haberman, R. (2019). *An assessment of gaps in business acumen and knowledge of industry within the defense acquisition workforce: A report prepared for the U.S. Department of Defense in compliance with Section 843(c) of the Fiscal Year 2018 National Defense Authorization Act* (RR-2825-OSD). RAND.

PANEL 9. ADVANCING NAVAL CAPABILITIES: PARTNERSHIPS, DESIGN EXCELLENCE, AND CASE STUDIES

Wednesday, May 7, 2025	
1115 – 1230 PT 1315 – 1430 CT 1415 – 1530 ET	**Chair: Rear Admiral Thomas J. Anderson, USN, Program Executive Officer, Ships** ***Identifying Pathways for U.S. Shipbuilding Cooperation with Northeast Asian Allies*** Henry H. Carroll, Research Assistant, Defense-Industrial Initiatives Group, Center for Strategic & International Studies ***Navy Shipbuilding: Increased Use of Leading Design Practices Could Improve Timeliness of Deliveries*** Sean Merrill, Senior Analyst, U.S. Government Accountability Office ***USMC Landing Craft Case Study*** Robert Mortlock, Professor, Naval Postgraduate School

Rear Admiral Thomas J. Anderson, USN—is a native of North Brunswick, New Jersey. He was commissioned in 1991 through the Naval Reserve Officer Training Corps (NROTC) Program at Boston University where he received a Bachelor of Science in Mechanical Engineering.

Anderson's tours as a surface warfare officer included USS Capodanno (FF 1093) and USS Arleigh Burke (DDG 51), where he coordinated the first two Chief of Naval Operations availabilities of the DDG 51 Class.

Upon selection to the engineering duty community in 1996, he attended the Naval Postgraduate School where he earned a Master of Science in Mechanical Engineering. He also completed the Total Ship Systems Engineering Curriculum and became a California State Licensed Professional Engineer.

Ashore, he has served in a variety of industrial, fleet, program office and headquarters assignments in ship design and construction, maintenance, budgeting, and requirements. His ashore assignments include: Naval Sea Systems Command (NAVSEA) executive assistant; Littoral Combat Ship Shipbuilding program manager (PMS 501); Office of the Chief of Naval Operations requirements officer (N86); chief engineer and post-delivery branch head for the DDG 51 Class (PMS 400D); and Commander, Naval Surface Forces, Atlantic, mine warfare type desk officer (N43).

As a flag officer, Anderson served as commander Naval Surface Warfare Center (CNSWC), Department of Defense executive manager for Military Explosive Ordnance Disposal (EOD) Technology and Training, commander Navy Regional Maintenance Center (CNRMC) and Naval Sea Systems Command director Surface Ship Maintenance and Modernization (SEA 21).

Since May 2020, Anderson has served as Program Executive Officer, Ships, where he is responsible for Navy shipbuilding for surface combatants, amphibious ships, logistics support ships, support craft, and related foreign military sales.

Anderson served as the 15th Commander Naval Sea Systems Command (NAVSEA) from August 2023 through January 2024. In that role, he oversaw a global team responsible for the development, construction, delivery and maintenance of the Navy's ships, submarines and systems.

Identifying Pathways for U.S. Shipbuilding Cooperation with Northeast Asian Allies

Henry H. Carroll—is a research associate with the Defense-Industrial Initiatives Group at the Center for Strategic and International Studies (CSIS). His analytical focuses include the defense industrial base, naval policy, and assessing foreign military industrial capacity. Prior to joining CSIS, he worked as an intern in the defense and international practices of Brownstein and as a defense-focused legislative intern for Senate Majority Leader Chuck Schumer. Henry holds a BA in history, with a concentration in war and society, from Yale University. His undergraduate senior thesis examined the politics of U.S. naval shipbuilding in the interwar period. [hcarroll@csis.org]

Cynthia Cook—is director of the Defense-Industrial Initiatives Group and a senior fellow in the Defense and Security Department at the Center for Strategic and International Studies. Dr. Cook is a member of the editorial board for the Defense Acquisition Research Journal and is an adjunct professor at the Pardee RAND Graduate School. From 1997 to 2021, Dr. Cook worked as a senior management scientist at RAND, where she served as the director of the Acquisition and Technology Policy Center. She holds a PhD in sociology from Harvard University and a BS in management from the Wharton School of the University of Pennsylvania. [ccook@csis.org]

Abstract

Military conflict in the Indo-Pacific will demand overwhelming American naval power. The challenges in U.S. shipbuilding, including capacity shortages, industrial base constraints, cost overruns, and delayed delivery, suggest that the United States should explore alternative pathways for delivering capability to the Navy. One option is enhancing cooperation with close allies, such as Japan and the Republic of Korea. The paper identifies options for cooperation such as allied participation in maintenance, repair, and overhaul (MRO), allied purchase and revitalization of U.S. shipyards, various methods of co-production including modular construction, and purchase of allied-built ships.

Introduction

The 2022 National Security Strategy identifies China as the United States' "pacing challenge." (Biden, 2022, p. 20) Given the vastness of the Pacific theatre, its vital shipping lanes, and the many regional allies and partners depending on a persistent U.S. security presence, sea power is critical to the U.S. strategy for promoting a free and open Indo-Pacific and competing with China. However, the United States has long struggled with shipyard capacity as well as timely and cost-effective construction of naval vessels (O'Rourke, 2025). Policies aimed at maintaining its shipbuilding industrial base have failed to yield a sector that keeps pace with those of potential adversaries (Evans, 2023; Jones & Palmer, 2024, p. 15). In April 2024, a 45-day review of the U.S. shipbuilding industrial base by the secretary of the Navy found that many of the Navy's major shipbuilding programs were one to three years behind schedule (O'Rourke, 2025, p. 55).

The lack of adequate naval shipbuilding capacity, as well as the moribund state of the U.S. commercial shipbuilding industry, significantly hinders the country's ability to increase production of ships in the event of a conflict. U.S. workforce constraints, facilities limitations, and supply chain challenges have contributed to an inability to deliver necessary capabilities on schedule and at scale. At the same time, China's share of global commercial and military shipbuilding continues to grow rapidly, accounting for 51 percent of global ship deliveries in 2023, with current trends pointing toward an eventual shift in the maritime balance of power (Mandhana, 2024).

In response to this shipbuilding crisis, new modalities are needed, particularly those that lean on the United States' unique strength: its network of allies and partners. The U.S. Navy could turn to Japan and South Korea for industrial cooperation to scale warship production, which would represent an important shift in its naval acquisition policy and broader industrial partnerships with allies. South Korea and Japan are the world's second and third largest producers of ships, respectively, and could contribute significantly to U.S. warship production, whether overseas or at U.S. shipyards (Mandhana, 2024). Carlos del Toro, the secretary of the Navy under the Biden administration, was remarkably forward-leaning in considering the possibility of looking abroad to reinvigorate the U.S. shipbuilding industry and maritime production, and recent comments from the incoming administration appear to indicate it is also favorable to a rethinking of approaches to increase capacity ("Trump's call for 'K-shipbuilding'", 2024; United States Navy, 2024). However, there is currently a lack of rigorous and public analyses on the potential advantages and challenges of the several ways in which U.S. industry could cooperate with allies on shipbuilding.

Each approach to international shipbuilding cooperation comes with its own benefits and drawbacks. These are not mutually exclusive, and the U.S. government may also choose to pursue a combination of pathways; if so, it should also consider whether and how cooperation methods affect each other when pursued simultaneously or sequentially. Such an analysis should take into account the possibility that pathways impact the same underlying factors, such as labor availability or overall demand, as well as potential path dependencies in industrial planning, wherein funding one shipyard approach may require the conversion or use of limited yard space.

Moreover, while these approaches could each serve as useful measures to enhance naval capability, they would each affect the long-term health of the U.S. shipbuilding industry. Determining the nature of this impact—¬¬¬¬¬be it positive, negative, or a combination of the two, with variation across different subsectors—is critical for policymakers as they balance meeting the imminent threat from China with the strategic need to ensure the long-term strength of the U.S. shipbuilding industrial base.

Background and Analysis

- **The Strategic Situation**

Winning any conflict with China will inevitably require maritime dominance given the vastness of the Pacific Ocean and the location of flashpoint areas such as Taiwan, the South China Sea, the Korean Peninsula, and Guam. A series of wargames conducted at CSIS in 2024 found that while China would likely lose any conflict over Taiwan, the United States would still incur significant losses in terms of ships, submarines, and planes, including naval aviation assets (Cancian et al., 2023). While the U.S. Navy retains a qualitative and tonnage edge over China's People's Liberation Army Navy (PLAN), the pace of PLAN construction greatly exceeds that of the U.S. Navy (USN, Palmer et al., 2024). Chinese shipbuilding, both naval and commercial, has been thriving. Due to vast industrial subsidies, China's shipyards have gone from producing 5 percent of the world's ships in 1999 to over 50 percent in 2024, and many of these shipyards embrace the Chinese Communist Party's military-civil fusion strategy and produce warships for the Navy (Funaiole, 2024).

In a protracted great power conflict, the United States would likely struggle to repair and replace its ships fast enough to keep up with China, let alone construct sufficient new vessels to establish and retain control of sea lines of communication. Given the well-documented struggles of U.S. shipbuilding, the United States should explore supplementing domestic production with other options for sustaining and growing its naval might. Cooperation with shipbuilding heavyweights South Korea and Japan offers one possible approach to solving this issue.

- **The United States' Shipbuilding Challenge**

The U.S. naval shipbuilding sector faces critical hurdles. The most commonly cited are skilled-workforce constraints, antiquated shipyard infrastructure and equipment, insufficient use of new technology including digital tools and modular construction techniques, and legacy organizational structures (Weddle et al., 2024). Other analysts point to issues in U.S. design capacity, hyper-specialization of military shipyards that holds back scalability, and a 20-year backlog of maintenance and repairs that constrains the Navy's ability to practice and train with its existing ships (Seavy, 2024). Some of these issues stem from the closure of U.S. shipyards during the 1990s, which hindered the domestic production of ships and left significant gaps that in industrial readiness (Di Mascio, 2024). The reasons behind the U.S. shipbuilding challenge are the subject of a substantial and growing body of literature, including recent analyses from policymakers such as Senator Jack Reed (D-RI) and former Senator Jim Inhofe (R-OK); public institutions such as the Government Accountability Office (GAO) and the Congressional Research Service; and various think tanks (Dallas et al., 1994; Oakley, 2024a; O'Rourke, 2025; Reed & Imhofe, 2021).

A near universally agreed-upon challenge for the U.S. shipbuilding industry is the dearth of a skilled workforce. Coupled with the retirement of workers with shipbuilding skillsets, a demographic shift away from manufacturing careers has created recruitment and retention challenges, resulting in an inexperienced workforce that lacks proficiency in skilled trades and requires increased supervision to avoid quality problems (Oakley, 2025b, pp. 27–28). A March 2025 Congressional Research Service report found that part of the difficulty in recruiting and retaining new workers lies in the relatively low wages and benefits of shipbuilding jobs. While such jobs still pay better than service and retail, the differential in wages has narrowed, and the latter are less likely to involve risk of serious injury, are often located in areas with easier commutes, and are generally done in cleaner indoor settings. And while increasing total wages for shipbuilding workers could reestablish a large differential in wages and benefits, it would also substantially increase ship procurement costs (O'Rourke, 2025, p. 23).

Simultaneously, shipbuilders also struggle to acquire land for expanding existing shipyard facilities, building new shipyards, or providing housing for workers near shipyards. In 1988, the first Base Realignment and Closure (BRAC) process prompted the selling of land access along coastlines, resulting in the closure of four naval shipyards in the 1990s (Di Mascio, 2024). In some places, land could theoretically be repurchased, but at a steep price, given that it has since been put to new productive uses. Greenfield development along coastlines is generally difficult, as there are few useful places that are both unused and affordable to buy and build on (Hooper, 2023). Even when land itself is cheap, shipyards can face high costs when developing it—whether for industrial use or worker housing—due to the potential for industrial contamination, which requires expensive remediation (Waxmann, 2024).

Civilian shipbuilding is a critical supporter of naval construction due to the scaling benefits of sharing skillsets and material inputs, as well as the smoothing of demand over time, as many shipyards around the world build both military and commercial vessels (Schank et al., 2005). Yet the United States has rarely been a world-leading constructor of civilian vessels, except during and immediately after World Wars I and II (Colton & Huntzinger, 2002). Consequently, maritime historians argue that the United States' strategic culture is split between alignment with territorial land empires such as Germany and true sea power states such as the United Kingdom—and therefore is less likely to maintain a consistent engagement with the sea via commercial shipbuilding the way a fully maritime state would (Lambert, 2019). This lack of persistent cultural and strategic interest is reflected in the poor state of U.S. shipbuilding today relative to the rest of the world (Frittelli, 2023).

- **U.S. Policy Options**

In response to these clear challenges in the face of growing strategic demand, the United States has several possible options, some of which have engaged leaders' interest and support. Senior political figures such as Senator Mark Kelly (D-NJ) and former National Security Advisor (and former Congressman) Mike Waltz (R-FL) have been at the forefront of efforts to revitalize shipbuilding via domestic investments (Center for Strategic and International Studies, 2024). Their "Congressional Guidance for a National Maritime Strategy," published alongside other members of Congress, proposes both incentives for U.S. shipbuilders and carrying cargo on U.S.-flagged commercial vessels (Waltz et al., 2024). This congressional effort, in the form of the Shipbuilding and Harbor Infrastructure for Prosperity and Security (SHIPS) for America Act of 2024, advances a domestic approach focused on revitalizing shipbuilding within the United States (H.R. 10493, 2024).

The SHIPS for America Act, like other current efforts, focuses on generous subsidies and legal privileges for the U.S. shipbuilding industry. This domestic approach draws upon a wide and deep body of literature that aims to diagnose the issues facing domestic shipbuilding; it has many political and intellectual champions within the United States, including from powerful industry groups that have been active since the 1930s (Paxton & Schonhaut, 2024; Shipbuilders Council of America, 1937).

Some of the most promising—yet less comprehensively studied or advocated—policy options involve the United States partnering with Asian allies. Statements from officials in the Trump administration, such as Secretary of Defense Pete Hegseth, as well as from the Biden administration, including former Secretary of the Navy Carlos Del Toro and former Ambassador to Japan Rahm Emanuel, suggest that U.S. policymakers are interested in exploring cooperating with South Korea and Japan to overcome challenges to the naval shipbuilding industry (Politico, 2025; Lagrone, 2024). Even the SHIPS for America Act's explicitly domestic strategy includes references to assistance from international actors, especially treaty allies such as South Korea and Japan. The associated congressional guidance notes that the United States should "seek mutually beneficial relationships with treaty allies, exploring comparative advantages to lower cost, time, and the complexity of rebuilding America's domestic shipping and shipbuilding industry" (Waltz et al., 2024, p. 6). Indeed, analysts have proposed a variety of international cooperation options, from realistic ones grounded in statements by Navy and political leaders to more theoretical and creative options (Seavy, 2024).

- **Why Cooperate Internationally? Examining Allied Strength in Shipbuilding**

Unlike the United States, South Korea and Japan have impressive shipbuilding industries, making them valuable potential partners. While China produces more ships overall due to its greater number of shipyards, South Korean and Japanese shipyards continue to lead the world in contemporary productivity due to technical advancements. However, China is rapidly closing the productivity gap (Chao & Yeh, 2020, p. 193-210).

South Korea rose as a commercial shipbuilding power between 1970 and 1990 as significant government subsidies, technological advancements, and favorable economic conditions such as low labor costs enabled it to outpace U.S. and European industry during a challenging period for the global shipbuilding market (Bruno & Tenold, 2011, p. 201-217). South Korea has retained its cost-competitive edge even as its labor has grown more expensive alongside the development of its economy. Advancements in automation and control systems within its shipbuilding industry have shifted the sector from labor-intensive to technology-driven (Min, 2008, p. 7185-7190).

Japanese shipbuilders currently maintain a strong market presence, building on their long period of dominance that began after World War II (OECD, 2016). The country nevertheless faces challenges in competing with South Korean and Chinese shipyards that excel in increasingly high-value and rapid ship construction (OECD, 2016). Japan's focus on high quality standards and gradual adoption of automation could be enhanced to meet military demands, especially with increased collaboration across maritime technology sectors. Much like South Korea, Japan has a strong focus on automation in its shipbuilding sector for both simple vessels and highly complex naval ships (Koenig et al., 2003, p. 131-140).

However, U.S. policymakers should also consider the impact that international approaches to shipbuilding may have on the health of the U.S. shipbuilding industry. Some pathways explored below, such as international companies purchasing U.S. facilities and incorporating their advanced production techniques, have the potential to boost domestic productivity and competitiveness. Other approaches could undermine the industry's long-term health, particularly if they direct production to foreign shipyards at the cost of domestic order books—which could have downstream impacts on labor force retention and the capacity of U.S. shipbuilding in the long run.

South Korea and Japan are not the only U.S.-aligned countries with innovative and effective shipbuilding industries. Other nations may have much to offer the United States, particularly regarding cost-effective warship construction. Given the scope of the work, however, this project is focused on South Korea and Japan as possible cooperation partners due not only to their dominance of the commercial shipbuilding market, which gives them significant scale advantage on cost, but also their history of close industrial cooperation with the United States on military production and sustainment.

- **Possible Pathways for International Cooperation with U.S. Allies**

This study identifies several approaches for international cooperation on naval shipbuilding with South Korea and Japan, including:

- allied maintenance, repair, and overhaul of U.S. ships to free up U.S. shipyard capacity;
- allied acquisition of U.S. shipyards to revitalize their production capability;
- joint distributed production of warships via modular construction methods; and
- U.S. purchase of existing allied warship designs from allied shipyards.

Although there are additional avenues for cooperation, these four pathways emerged as the most actionable and reasonable based on a survey of public discourse, existing U.S. government policies such as the Regional Sustainment Framework, and CSIS interviews with U.S. and allied industry, as well as with government officials, over the past year (DoD, 2024a). Other policy options outside the scope of these pathways and this report have been floated, including various combinations or divisions of the above ideas, as well as proposals that depend on outside parties such as the global naval export market.

The following sections discuss these possible pathways for cooperation in greater detail. They review the extant literature relevant to each pathway and describe the most viable forms of cooperation within it, as there is often more than one form of activity that the pathway could take.

- **International Cooperation on U.S. Maintenance, Repair, and Overhaul**

MRO activities are essential for ensuring a fleet's operational readiness and long-term availability for action. MRO activities range from routine inspections and maintenance actions (for example, applying surface coatings) to major service life extensions or refits of weapons systems (Marsh, 2024; Office of the Chief of Naval Operations, 2019). Robust MRO capacity enables a nation to maintain combat power during prolonged conflicts and ensure cost-effective and timely servicing of ships during peacetime.

However, MRO operations in the United States are facing significant challenges. Due to shipyard capacity, the U.S. Navy is estimated to be 20 years behind on maintenance work, leading to the decommissioning of viable ships as a result of its inability to conduct core MRO, modernization, and service-life extensions (Seavy, 2024). A February 2025 GAO report identified a lack of infrastructure and workforce capacity as the main obstacle to ship repairs, resulting in an inability to perform unplanned work such as emergency repairs. Even if hiring and retention efforts for skilled labor are successful in ameliorating the widespread workforce shortages, new employees will still be inexperienced, which will likely result in reduced efficiency in the short term (Oakley, 2025a, p. 28).

The report also notes that workforce and infrastructure capacity is dependent on "fleet concentration areas," or places where ships are homeported and undergo repair at domestic facilities. The GAO identified the five such areas that conduct major repair: one each in California, Florida, Hawaii, Virginia, and Washington. If slated repair work exceeds the capacity of one fleet concentration area, it may have to be done at another location (Oakley, 2025a, p. 31). To address these physical capacity constraints, two out of seven shipbuilders interviewed for the GAO study have outsourced work to their suppliers, with "plans to expand the volume of material they are outsourcing." Another shipbuilder has plans to use outsourcing, and an additional one is considering outsourcing if it is awarded a new contract by the Navy. Yet while outsourcing can reduce physical constraints at shipyards, suppliers often "have their own workforce and infrastructure problems" (Oakley, 2025a, p. 25).

Because of these challenges, enabling greater use of allied MRO in the Indo-Pacific region is critical for U.S. strategic goals and those of Japan and South Korea (Tanaka, 2024). In strengthening supply chains, leveraging the strategic positioning of ports, and expanding MRO capacity, some scholars believe the United States could solve its shipyard dilemma by empowering domestic yards to focus on facilities and process modernization (Kim, 2023).

The United States has already begun laying the groundwork for greater MRO cooperation across the entire Indo-Pacific through the Regional Sustainment Framework (DoD, 2024a). One of the core goals of the framework is to leverage existing regional MRO capacity within partner nations, particularly for shared weapons systems operated by allies and partners. (DoD, 2024b) Close MRO cooperation with Indo-Pacific treaty allies has already been pursued in Australia, where the United States has begun an initiative to advance combined regional MRO solutions in support of the framework (Defense MRO Playbook, 2024, p. 10). Another venue for cooperation is the U.S.-Japan Defense Industrial Cooperation, Acquisition, and Sustainment (DICAS) Forum, which aims to accelerate joint development and sustainment of defense equipment. DICAS oversees multiple working groups, including the Ship Repair Working Group, which seeks to identify opportunities and challenges for U.S. naval ships to be maintained by Japanese shipyards (DoD, 2024c).

MRO cooperation with South Korea and Japan is hardly new. The U.S. Navy has been collaborating with Japanese industry on MRO since the end of World War II, with NIPPI beginning to service assets in the 1950s (Wilson, 2021). In 2019, Mitsubishi Heavy Industries (MHI) conducted maintenance on the USS *Milius*, an *Arleigh Burke*–class guided missile destroyer, and signaled their desire for more contracts with the U.S. Navy (Wilson, 2024). In 2024, U.S. Ambassador to Japan Rahm Emanuel announced plans to build on the 2019 maintenance collaboration with MHI, saying that the U.S. Navy would send some of its vessels to Japanese shipyards for MRO (Wilson, 2024). Japanese companies have also performed

MRO activities on some U.S. auxiliary vessels (Tanaka, 2024). Similarly, in August 2024 South Korean naval shipbuilder Hanwha Ocean received their first MRO contract with the U.S. Navy to provide services to one of their cargo and ammunition ships, the USNS *Wally Schirra*, which has since been completed ("USNS Wally Schirra Completes Major Maintenance", 2025). In November 2024, Hanwha Ocean received another contract to perform MRO services on the USNS *Yukon*, a replenishment oiler (Boram, 2024). Another South Korean shipbuilder, HD Hyundai, signed a Maintenance and Ship Repair Agreement (MSRA) with the U.S. Navy, which qualifies them to bid for maintenance projects for U.S. combat and support ships ("South Korea's HD HHI Inks MRO Agreement", 2024).

However, the existing literature highlights that shifting MRO work to foreign yards could have negative economic consequences for the U.S. ship-repair and shipbuilding sector (Kim, 2023). Other sources indicate that certain parts of MRO operations, such as routine maintenance, constitute only a limited part of shipyard economies; as such, shifting some MRO operations abroad is unlikely to damage the U.S. shipbuilding industry and broader economy (Maritime Administration, 2021). Nevertheless, naval MRO activities between World Wars I and II were a source of stability for cash-strapped shipyards that were otherwise out of work—even if these activities only constituted 0.3 percent of the value of U.S. private shipyards' total commercial and naval work from 1920–39 (Smith & Brown, 1948, p. 105).

This pathway requires close study to ascertain its potential value to the U.S. Navy, especially in terms of how much it may free up new production capacity in the United States and how offshoring these small but routine (and therefore valuable for long-term financial stability) contracts to U.S. allies may affect the domestic shipbuilding industry. A critical factor in determining the viability of this approach will be its anticipated ability to create new shipbuilding capacity—as the facilities, machinery, and skilled workforce used in MRO operations are not necessarily the same as for shipbuilding and may require substantial time and money to switch. For U.S. allies and partners, the economic appeal of this pathway will depend on the potential magnitude of the market as foreign shipbuilders consider whether to dedicate existing facilities to U.S. naval MRO, expand capacity to support U.S. Navy ships in their home country, or expand capacity in third countries such as the Philippines (CSIS interviews with an international shipbuilder, November 18, 2024).[1]

Allied Acquisition of U.S. Yards—Tech Transfer and Productivity Improvements

Allied companies' acquisition of U.S. shipyards offers another approach for Japanese and South Korean shipbuilders to support U.S. shipbuilding via entering the U.S. market. The goal would be for the purchasers to impart the home nation's shipbuilding expertise and efficiencies. The partner company would need to set up a U.S. subsidiary that can prove it is not under foreign ownership, control, or influence—meaning it qualifies as a U.S. company for both Jones Act considerations and "Buy America" clauses in military contracting.[2] In a relevant example, Hanwha Ocean recently purchased Philly Shipyard, having received the necessary regulatory approval ("Hanwha Closes $100 Million Philly Shipyard Acquisition", 2024).

There is not just one way for international shipbuilders to become involved in the U.S. domestic shipbuilding market. Through reviews of the literature, qualitative research, and

[1] Allied shipbuilders are opening new facilities across the Indo-Pacific to grow their addressable market ("HHI Opens New Philippine Office", 2024).
[2] Any foreign-acquired or -built shipyard would need a Facility Clearance to be eligible to access classified information. Facilities deemed under foreign control or ownership cannot qualify (*Entity Vetting, Facility Clearances & FOCI*, n.d.).

interviews with industry over the past nine months, the following have emerged as possible sub-pathways:

- international purchase of existing operational U.S. military shipyards;
- international purchase and renovation of defunct or nonmilitary U.S. shipyards;
- creation of new government-owned, commercially operated (GOCO) U.S. shipyards with foreign shipbuilders considered in the operator bidding pool; and creation of a joint venture or consortium between U.S. and international industry to produce ships within an existing U.S. shipyard.

The key question regarding international acquisition of U.S. shipyards in any of these sub-pathways is whether new ownership can improve shipyard performance. Given that the U.S. workforce and material input costs will largely remain unchanged, the key theorized drivers of improvement would be altered management practices, possible cross-training of shipyard workforces, and technology transfer of more advanced foreign shipbuilding techniques to the United States, as well as the accompanying capital infusion required to implement those new techniques.[3] Some of these methods have been publicly discussed by officials from Hanwha as ways to improve their newly acquired Philadelphia yard (Korea Economic Institute of America, 2025). These sub-pathways are not mutually exclusive, and selecting one for a given situation would depend on both local conditions and an assessment of how its particulars would facilitate—or not—productivity gains in general.

Technology transfer is difficult to catalyze and manage properly (Andrenelli et al., 2019). The United States' experience managing military technology transfer—especially through the Department of Defense, where national competitiveness and security are paramount concerns—has overwhelmingly been as the provider, rather than the recipient, of technology transfers (Defense Security Cooperation University, 2024). This lack of DOD experience may be a complication for this pathway, necessitating further close study so policymakers and implementers are fully aware of potential hurdles and best practices.

A limited literature supports the possible returns of technology transfer from advanced shipbuilding nations like Japan to companies in the United States. One 1987 study found that Japan's Ishikawajima-Harima Heavy Industries (IHI)'s technology transfer efforts to U.S. shipbuilders—building on IHI's advanced techniques, such as block construction, process lane systems, and a strong emphasis on material management and design standardization—were able to improve productivity, but not catch the U.S. shipyards up to Japanese standards (Sasaki, 1987). More recent comprehensive studies are lacking, however, providing an opening for scholars to contribute to the extant literature on the possible returns and tradeoffs of technology transfer to U.S. shipbuilders.

This pathway has been pursued in recent history. Italian shipbuilder Fincantieri purchased the Wisconsin-based Marinette Marine shipyard in 2009, with Lockheed Martin as a minority owner (Fincantieri, 2008). The new company won the competition to build the *Constellation*-class guided missile frigate in 2020 (although construction challenges, including workforce limitations, have contributed to late delivery of the first-in-class ship) (Oakley, 2024a, p. 7). Austal, an Australian shipbuilder, started operations in Alabama in 1999 and began to expand rapidly in 2005 after winning a contract to design the *Independence*-variant Littoral Combat Ship for the U.S. Navy (Austal, 2024). Foreign shipbuilder acquisition in these cases has brought in new investment and modernization efforts. Fincantieri, since buying its yards in Wisconsin, has invested over $300 million in them (Fincantieri Marine Group, 2021). It also has

[3] CSIS interviews with international shipbuilders on, October 23, 2024, and November 18, 2024, raised the examples of investing in the capital equipment necessary to enable greater automation as well as bringing over experienced or retiring workers who could train U.S. personnel.

leased new yard space in an existing Florida shipyard near Commodores Point in Jacksonville, announcing plans to invest $30 million into improvements and modernization there to support its sustainment and repair work (Mathis, 2022).

These past examples of international acquisition offer some early insight into the challenges and opportunities of foreign ownership that could be applied to future cooperation with South Korea and Japan. The benefits of foreign ownership are clear. Foreign parent firms with commercial enterprises—a rarity in the United States—can bring the energy, expertise, and innovative capabilities of the advanced commercial market to U.S. shipyards (Oakley, 2025b, p. 15). For example, these shipbuilders often use more robotics and automation in processes such as panel making than U.S. defense shipyards do, which can reduce strains on a depleted workforce and improve efficiency (Lo, 2013). Foreign ownership by large shipbuilders can also enable buying certain inputs in bulk, especially if they are not exclusively military in nature. At a minimum, parent firms can help provide information and negotiation power to their U.S. subsidiaries as they buy components for ships, contributing to lower costs (CSIS interview with U.S. shipbuilder, April 1, 2025).

Regulatory barriers such as complying with Committee on Foreign Investment in the United States (CFIUS) and Foreign Ownership, Control, or Influence (FOCI) mitigation are unlikely to pose major barriers given government support. However, International Traffic in Arms Regulations (ITAR) are likely to pose major challenges. Ship designs are controlled by ITAR, down to the nonmilitary design elements that could benefit from foreign owners' commercial expertise, including galley and berth plans (Code of Federal Regulations, 2024). Even visits by experts from potential parent companies can involve ITAR, and long-term residency permits to allow foreign expertise to benefit U.S. yards can be difficult to obtain. A further complication is that U.S. Navy standards and procedures are unique, and communicating these requirements to foreign parent firms can require an additional ITAR waiver, making it more cumbersome for the U.S. shipyard to benefit from foreign expertise (CSIS interview with U.S. shipbuilder, April 1, 2025).

Modularity in Shipbuilding via Distributed Construction

Modularity is part of the advanced shipbuilding approaches employed by South Korean and Japanese shipbuilders. For commercial shipbuilding of massive cargo and tanker ships, it is analogous to taking 250–300 modules manufactured in workshops and assembling them like bricks in a drydock (CSIS interview with international shipbuilder, November 18, 2024). For this report, "modularity" in shipbuilding refers to two separate but related methods, both of which have the potential to improve U.S. shipbuilding capacity.[4] The first is advanced outfitting, or the construction of a ship by assembling together pre-furnished modules, such as horizontally and vertically joining sections of a ship. The other method is the use of modular systems, or the integration of various components—from weapons systems to power plants—onto a hull using common standards for key interfaces, enabling a loose coupling between the manufacture of the ship and the system.[5]

[4] As defined by a Naval Sea Systems Command (NAVSEA) document cited in a leading report on the topic, "modularity" is "a design approach in which a system has the following characteristics: functionally partitioned into discrete, scalable, and reusable modules consisting of isolated, self-contained elements; a systems engineering process that emphasizes functional analysis and the identification of key interfaces; common industry standards for key interfaces to the largest extent possible (John F. Schank et al., 2016).
[5] These definitions of "modularity" link to production. This study does not cover "mission modularity," which is the idea that ships can use rapidly interchangeable mission modules to swap in different capabilities so they can serve as multi-purpose vessels, as was the concept behind the Littoral Combat Ship (Salisbury, 2023).

Modularity is a key enabler of the distributed approaches that could bring in both international shipyards and U.S. subsidiaries of international firms. In the international context, therefore, possible complementary pathways for the United States to work with its allies include:

- a business-to business distribution system in which U.S. shipyards subcontract or enter joint ventures to assemble ships from complete, pre-furnished hull modules and systems built by the United States and its allies;
- a government-furnished equipment system in which The U.S. Navy procures U.S.- and allied-built hull modules and systems directly and provides them to U.S shipyards; and
- a system assembly system in which U.S. shipyards integrate modular systems (such as weapons and propulsion systems) onto pre-built, complete ship hulls from South Korea and Japan.

Each form of modular cooperation involves its own advantages and challenges. In general, the literature on the role of modularity in shipbuilding is overwhelmingly positive, with many studies highlighting how flexible and modular designs could reduce costs, enhance international cooperation, and support modernization and adaptability. Rains, for example, highlights the benefits of the potential reduction in ship size, while Malone and Rubeša, Fafandjel, and Kolić emphasize pre-outfitting in modules in workshops to minimize the work that must be done in dry docks (Malone, 2019; Rains & Johnson, 1993; Rubeša et al., 2011).

In particular, three types each of modularity and flexibility are identified as having potential for the modernization and adaptability of the U.S. Navy. For modularity, these include common modules, self-contained modules, and modular installations; ship infrastructure, additional space, and additional ship services are listed for flexibility (Schank et al., 2016). The U.S. Navy could additionally benefit from a more optimized and comprehensive approach to modularity, for example by integrating standardized components and standardized weapons systems into a collection of ready hull designs (O'Rourke, 2025, p. 26).

Studies of recent shared-build warship programs in the United States, France, and the United Kingdom identify risk reduction areas, key costs, and potential benefits of international modular shared-build programs and highlight the conditions and circumstances under which multi-shipyard, modular-build strategies can be adopted (Smallman et al., 2011). In addition, the works of Friedman, Lombardi, and Rudd, who outline recent challenges faced by the United Kingdom with joint shipbuilding, become particularly useful for understanding how the United States might leverage international partnerships to fill in its aforementioned gaps in production (Lombardi & Rudd, 2013, p. 1-17; Friedman, 1999).

One international shipbuilder thought it could potentially subcontract to U.S. shipbuilders to provide either modules produced in a U.S. subsidiary yard or generators built in other inland facilities (CSIS interviews with an international shipbuilder, October 23, 2024). The business-to-business path was seen as a low-margin but nevertheless appealing opportunity to generate early revenue and build trust with other shipbuilders who might be otherwise inclined to see new entrants primarily as competition (CSIS interviews with an international shipbuilder, November 18, 2024). Modularity, referred to as "federated shipbuilding" or "nation as a shipyard," has also been proposed as a solution to domestic U.S. shipbuilding constraints (O'Rourke, 2025, p. 25). If these approaches are adopted, allied firms could plug into these domestic modular approaches as suppliers, leveraging workforces and materials not just in the inland United States but across allied nations as well.

Skeptics of modularity, however, point to inefficiencies in the field that can counterintuitively lead to higher procurement costs and delayed timelines. This is due to the high degree of skill that complete modularity requires and is evidenced by the costly case of the

U.S. Littoral Combat Ship (Axe, 2009). As an additional challenge, some of the efficiency of high-productivity shipyards comes from the equipment they employ to move the largest of modules, so additional capital investments may be required for some U.S. producers to take advantage of offsite module production, particularly across the long distances involved.

U.S. Purchase of Ships from Allied Yards

The final pathway identified by this work is the U.S. Navy purchasing ships that are produced in allied yards. There are numerous sub-pathways for this form of cooperation, including:

- allied yards building licensed U.S. designs,
- allied yards building a new co-developed design, and
- the United States buying allied-built and allied-designed ships.

However, the existing literature on such approaches is limited due to the novel nature of this idea in U.S. shipbuilding history.

This is perhaps the most difficult and unlikely pathway explored in this paper. For instance, although reusing the existing designs for foreign ships would likely offer the most cost effective and rapid solution to at-sea capacity gaps, these ships might not meet the U.S. Navy's specific operational requirements, including full interoperability with U.S. systems (Oakley, 2024a, p. 39). Past experience in trying to adapt foreign designs shows that the U.S. Navy's tendency to "gold plate" design requirements can cause scope creep, raising costs and time frames. Moreover, its standards and procedures are not shared by other navies, requiring a major rework of allied designs to be acceptable for domestic use. The Navy's attempt to have the existing Italian design of the European multipurpose frigate (FREMM) quickly converted into the U.S. Navy's Constellation class, for example, has resulted in a final U.S. design that reportedly bears less than a 15 percent similarity to the FREMM, down from a planned 85 percent, at great cost of time and money to the United States government (Shelbourne and LaGrone, 2024; Oakley, 2024b, p. 24). Some allied designs are similar to existing U.S. designs, such as South Korea's KDX-III Batch I Aegis destroyers, which are said to be based on the DDG-51 Arleigh Burke–class of the U.S. Navy (Vavasseur, 2021). Use of these mostly shared designs could potentially ease the compatibility issue.

A new codeveloped design could take advantage of partner shipbuilding expertise to incorporate manufacturability in the design phase. However, designing new ships is a notoriously hard and slow process. Modern warships are incredibly complex machines; even a single amphibious assault ship contains 4.7 million parts from over 700 companies (Thompson, 2022). The lengthy process of designing and building new destroyers means this approach would have a long time horizon, making it susceptible to changing political winds. While some experts float this option, linking it to potential export sales as a way to spread out production costs across more customers, its technical and political challenges are daunting.

The allied build of U.S. ships alleviates many of these considerations but raises new challenges of its own. Designs will need to be licensed to allied yards, which will take time to negotiate, as will securing funding for intellectual property rights. It can take two years for even a comparatively expedited technical-assistance agreement to address export controls and the release of closely held U.S. weapons-system designs (Interview with international shipbuilder, November 18, 2024). However, the United States has managed to share its advanced capabilities with Korean shipbuilders before. For instance, HD Hyundai recently delivered the Republica of Korea's Jeongjo the Great to South Korean Navy; it is the first of the new KDX III Batch II Aegis destroyers and the fourth domestically designed and built South Korean Navy ship to incorporate the U.S. Aegis system ("HD HHI Delivers First Jeongjo the Great-class Destroyer", 2024).

These alternative approaches can all be used to build warfighting capability through the delivery of additional ships. It remains unclear, however, whether and how they might address core concerns about the capability and capacity of the U.S. shipbuilding industrial base—a broader strategic issue that needs to form part of any consideration of an alternative approach. While the United States has some partnerships with other nations on specific programs, relying on allies as a complete solution to capability gaps would be unprecedented. Moreover, given South Korea's and Japan's location close to China, U.S. policymakers should also consider the possibility of damage to their shipyards during any active conflict. Finally, the health of the U.S. shipbuilding industry is not only a national security concern but also a political one given the well-documented impact of shipyards on their local economies and the interest of Congress in ensuring domestic capability (Keating et al., 2015; Maritime Administration, 2021).

Implication for Policy:

It is no secret that the United States has a shipbuilding problem. The U.S. policy community has long studied this issue and has produced a strong body of work exploring a wide range of possible domestic solutions. However, a continual lack of progress within the United States, the increasingly pressing threat of a fraught naval war with China, and recent shifts in political support for more creative solutions means a window of opportunity is opening for the Navy to consider adopting novel strategies that leverage the United States' strong and unique network of allies and partners. Yet policymakers lack a clear and comprehensive analysis of the options for industrial maritime cooperation.

Security cooperation policy is difficult to get right, and industrial cooperation policy can be even harder. For the United States to strike the right balance between leaning on its allies and partners to alleviate its shipbuilding problems and investing in its own capabilities at home, it will need to properly understand the advantages and challenges inherent to each kind of international cooperative activity in isolation and, critically, as they relate to one another. Recent history is littered with attempts at international cooperation that were partially or completely stymied by their starting conditions. For example, Constellation-class ships face three years of delays due to alterations to meet U.S. Navy requirements that lowered commonality with FREMM designs from 85 to 15 percent; both Australia and the United Kingdom have ratified technology-control treaties that go largely unused because industry is not confident in the regulatory implementation; and the F-22 fighter jet ended production rather than being exported because there were no initial investments in addressing technology-release concerns and the systems proved too expensive to retrofit (Greenwalt & Corben, 2023, p. 20-21; Shelbourne and LaGrone, 2024; Trevithick, 2021). Policymakers pursuing any of the pathways would greatly benefit from understanding the prerequisites for success as early in the process as possible.

As the Department of Defense will likely need to come to a decision on this key issue in the coming years, this project aims to support policymakers as they grapple with these difficult but critical decisions. Future work will evaluate these pathways using various assessments of interest to policymakers, including time to implement, cost to government, economic viability for industry, political and regulatory viability, and creation of new U.S. shipbuilding capacity.

References

Alex Wilson. (2021, July 28). Japanese Firms Sign $225 Million Deals to Maintain Ospreys for Navy, Marine Corps. Stars and Stripes. https://www.stripes.com/branches/navy/2021-07-28/osprey-tiltrotor-repairs-us-military-japan-2334831.html

Alex Wilson. (2024, January 19). Navy warships to be repaired at Japanese shipyards, ambassador says. Stars and Stripes. https://www.stripes.com/branches/navy/2024-01-19/navy-ship-repair-japanese-shipyards-12724715.html

An Analysis of the Navy's Fiscal Year 2024 Shipbuilding Plan | Congressional Budget Office. (2023, October 26). https://www.cbo.gov/publication/59508

Andrea Andrenelli, Julien Gourdon, & Evdokia Moisé. (2019). International Technology Transfer Policies (OECD Trade Policy Papers No. 222; OECD Trade Policy Papers, Vol. 222). OECD. https://doi.org/10.1787/7103eabf-en

Axe, D. (2009, August 26). Debating the Navy's "Plug-and-Play" Warship. Wired. https://www.wired.com/2009/08/debating-the-navys-modular-warship/

Boram, K. (2024, November 12). Hanwha Ocean wins 2nd maintenance deal from U.S. Navy. Yonhap News Agency. https://en.yna.co.kr/view/AEN20241112008100320

Brian T. Di Mascio. (2024, October 22). Foreign Shipyards Can Help the U.S. Navy Build Its Fleet. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2024/october/foreign-shipyards-can-help-us-navy-build-its-fleet

Brooke Weddle, Nick Mellors, & Ryan Brukardt. (2024, June 5). Shipbuilding in America: Charting a New Course. McKinsey. Retrieved November 26, 2024, from https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/charting-a-new-course-the-untapped-potential-of-american-shipyards#/

Bruno, L., & Tenold, S. (2011). The Basis for South Korea's Ascent in the Shipbuilding Industry, 1970–1990. The Mariner's Mirror, 97(3), 201–217. https://doi.org/10.1080/00253359.2011.10708948

Chao, Shih-Liang, & Yeh, Yi-Hung. (2020). Comparing the productivity of major shipyards in China, South Korea, and Japan – an application of a metafrontier framework. Maritime Business Review, 5(2), 193–210. https://doi.org/10.1108/MABR-12-2019-0060

Chris Lo. (2013, August 26). The digital shipyard: Robotics in shipbuilding. Ship Technology. https://www.ship-technology.com/features/feature-the-digital-shipyard-robotics-shipbuilding/?cf-view

Code of Federal Regulations. (2024). PART 121—THE UNITED STATES MUNITIONS LIST. https://www.ecfr.gov/current/title-22/chapter-I/subchapter-M/part-121

Colton, T., & Huntzinger, L. (2002). A Brief History of Shipbuilding in Recent Times: Defense Technical Information Center. https://doi.org/10.21236/ADA409101

Craig Hooper. (2023, October 10). Desperate For New West Coast Shipyards, Navy Eyes San Francisco Bay National Park. Forbes. https://www.forbes.com/sites/craighooper/2023/10/10/desperate-for-new-west-coast-shipyards-navy-eyes-san-francisco-bay-national-park/

Dallas, Andrew, McGrady, Edsel D., Perla, Peter P., & Kathleen J. Robertson. (1994). The Shipbuilding Game: Summary Report. The CNA Corporation. https://wargaming.hoover.org/view/ark:/54723/h3cr5nn6k/the-shipbuilding-game-summary-report

Defense Regional Maintenance, Repair, and Overhaul (MRO) Playbook. (2024). Defense Acquisition University. https://www.dau.edu/tools/defense-regional-maintenance-repair-and-overhaul-mro-playbook.

Emma Salisbury. (2023). Beware the Allure of Mission Modularity. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2023/may/beware-allure-mission-modularity

Entity Vetting, Facility Clearances & FOCI. (n.d.). Retrieved December 1, 2024, from https://www.dcsa.mil/Industrial-Security/Entity-Vetting-Facility-Clearances-FOCI/

Evans, M. (2023, July 12). China leaves US trailing in race to build warships. https://www.thetimes.com/world/asia/article/china-leaves-us-trailing-in-race-to-build-warships-hd0p6lszw

Fincantieri to Acquire the Manitowoc Marine Group. (2008, August 4). Fincantieri. https://www.fincantieri.com/en/media/press-releases/2008/000240/

Frittelli, J. (2023). U.S. Commercial Shipbuilding in a Global Context. Congressional Research Service.

Funaiole, M. P. (2024). The Threat of China's Shipbuilding Empire. https://www.csis.org/analysis/threat-chinas-shipbuilding-empire

Greenwalt, D. W., & Corben, T. (2023). Breaking the Barriers: Reforming US Export Controls to Realise the Potential of AUKUS. United States Studies Center. https://www.aei.org/wp-content/uploads/2023/05/Breaking-the-barriers-Reforming-US-export-controls-to-realise-the-potential-of-AUKUS.pdf

H. Gerrish Smith & L. C. Brown. (1948). Shipyard Statistics. In F. G. Fassett (Ed.), The Shipbuilding Business in the United States of America (pp. 61–200). The Society of Naval Architects and Marine Engineers.

Hanwha Closes $100 Million Philly Shipyard Acquisition – Hanwha Philly Shipyard. (2024, December 19). https://hanwhaphillyshipyard.com/hanwha-closes-philly-shipyard-acquisition/

HD HHI Delivers First Jeongjo the Great-class Destroyer to ROK Navy. (2024, November 28). Naval News. https://www.navalnews.com/naval-news/2024/11/hd-hhi-delivers-first-jeongjo-the-great-class-destroyer-to-rok-navy/

HHI Opens New Philippine Office to Target Southeast Asian Naval Ship Market. (2024, March 14). Naval News. https://www.navalnews.com/naval-news/2024/03/hhi-opens-new-philippine-office-to-target-southeast-asian-naval-ship-market/

Hiroshi Sasaki. (1987, August 26). IHI's experience of Technical Transfer & Some Considerations on Further Productivity Improvement in U.S. Shipyards. https://apps.dtic.mil/sti/tr/pdf/ADA454096.pdf

Jack Reed & Jim Imhofe. (2021, July 7). To Provide and Maintain a Navy: Understanding the Business of Navy Shipbuilding. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2021/july/provide-and-maintain-navy-understanding-business-navy-shipbuilding

Jeffrey L. Seavy. (2024, February 8). The United States Must Improve Its Shipbuilding Capacity. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2024/february/united-states-must-improve-its-shipbuilding-capacity

Jeong Soo Kim. (2023, May 1). Use Allies in Shipyard Modernization. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2023/may/use-allies-shipyard-modernization

John F. Schank, Scott Savitz, Ken Munson, Brian Perkinson, James McGee, & Jerry M. Sollinger. (2016). Designing Adaptable Ships. RAND Corporation.

Jones, S. G., & Palmer, A. (2024). Rebuilding the Arsenal of Democracy. CSIS. https://www.csis.org/analysis/china-outpacing-us-defense-industrial-base

Karen Brune Mathis. (n.d.). Fincantieri plans more than $30 million investment. Fincantieri Marine Repair. Retrieved December 1, 2024, from https://fincantierimarinerepair.com/fincantieri-marine-systems-teams-with-motor-services-hugo-stamp/

Keating, E. G., Danescu, I. E., Jenkins, D., Black, J., Murphy, R., Peetz, D., & Bana, S. H. (2015). The Economic Consequences of Investing in Shipbuilding: Case Studies in the United States and Sweden. RAND Corporation. https://www.rand.org/pubs/research_reports/RR1036.html

Koenig, P., Narita, H., & Baba, K. (2003). Shop-Floor Automation and Market Strategy in Japanese Shipbuilding. Journal of Ship Production. https://www.semanticscholar.org/paper/SHOP-FLOOR-AUTOMATION-AND-MARKET-STRATEGY-IN-Koenig-Narita/fd89c7c73d3e382d21aa154cc9dfb22703f9f488

Korea Economic Institute of America. (2025, March 26). All Hands on Deck: Korea's Strategic Role in Revitalizing U.S. Shipbuilding. https://www.youtube.com/watch?v=-ypEWmKupfw

LaGrone, M. S. and S. (2024). Constellation Frigate Delivery Delayed 3 Years, Says Navy. USNI News. https://news.usni.org/2024/04/02/constellation-frigate-delivery-delayed-3-years-says-navy

Lambert, A. D. (2019). Seapower States: Maritime Culture, Continental Empires and the Conflict that Made the Modern World. Yale University Press.

Laura Waxmann. (2024, October 12). Prolonged S.F. shipyard cleanup moves into final phase focused on offshore 'toxic hot spot' – the San Francisco Chronicle. San Francisco Chronicle. https://www.sfchronicle.com/sf/article/sf-hunters-point-shipyard-cleanup-19792980.php?

Laurence Smallman, Hanlin Tang, Schank, J. F., & Stephanie Pezard. (2011). Shared Modular Build of Warships: How a Shared Build Can Support Future Shipbuilding. RAND Corporation. https://www.rand.org/pubs/technical_reports/TR852.html

Lombardi, B., & Rudd, D. (2013). The Type 45 Daring-Class Destroyer. U.S. Naval War College Review, 66(3).

Marinette Marine | Shipyards | Fincantieri. (2021, September 30). https://fincantierimarinegroup.com/about-us/us-shipyards/marinette-marine/

Maritime Administration. (2021, March 30). Economic Contributions of U.S. Shipbuilding and Repairing Industry. https://www.maritime.dot.gov/sites/marad.dot.gov/files/2021-06/Economic%20Contributions%20of%20U.S.%20Shipbuilding%20and%20Repairing%20Industry.pdf

Mark F. Cancian, Matthew Cancian, & Eric Heginbotham. (2023, January). The First Battle of the Next War: Wargaming a Chinese Invasion of Taiwan. Center For Strategic and International Studies. https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/230109_Cancian_FirstBattle_NextWar.pdf?VersionId=WdEUwJYWIySMPIr3ivhFolxC_gZQuSOQ

Marsh, K. (2024, September 25). Maritime MRO for the Naval Industry. Naval Technology. https://www.naval-technology.com/buyers-guide/maritime-mro/

Mike Waltz, Mark Kelly, Marco Rubio, & John Garamendi. (2024, April 30). Congressional Guidance for a National Maritime Strategy. https://www.kelly.senate.gov/wp-content/uploads/2024/05/Congressional-Guidance-for-a-National-Maritime-Strategy.pdf

Min, K.-S. (2008). Automation and Control Systems Technology in Korean Shipbuilding Industry: The State of the Art and the Future Perspectives. IFAC Proceedings Volumes, 41(2), 7185–7190. https://doi.org/10.3182/20080706-5-KR-1001.01216

Moyuru Tanaka. (2024). Navigating the Competitive Seas. Center For Strategic and International Studies. https://www.csis.org/analysis/navigating-competitive-seas

Niharika Mandhana. (2024, February 13). China's Shipyards Are Ready for a Protracted War. America's Aren't. Wall Street Journal. https://www.wsj.com/world/china/chinas-shipyards-are-ready-for-a-protracted-war-americas-arent-d6f004dd

Norman Friedman. (1999, June 1). World Naval Developments: British Deep-Six Project Horizon. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/1999/june/world-naval-developments-british-deep-six-project-horizon

OECD. (2016). Peer Review of the Japanese Shipbuilding Industry. OECD. https://doi.org/10.1787/3f30a5a1-en

Office of the Chief of Naval Operations. (2019, May 8). OPNAV INSTRUCTION 4700.7M: Maintenance Policy for Navy Ships. Department of the Navy. https://www.secnav.navy.mil/doni/Directives/04000%20Logistical%20Support%20and%20Services/04-700%20General%20Maintenance%20and%20Construction%20Support/4700.7M.pdf

Our History. (n.d.). Austal USA. Retrieved December 1, 2024, from https://usa.austal.com/our-history

Palmer, A., Carroll, H. H., & Velazquez, N. (2024). Unpacking China's Naval Buildup. https://www.csis.org/analysis/unpacking-chinas-naval-buildup

Paxton, M., & Schonhaut, A. (2024, August 5). Outsourcing the US shipyard industrial base will outsource American sovereignty. Breaking Defense. http://breakingdefense.com/2024/08/outsourcing-the-us-shipyard-industrial-base-will-outsource-american-sovereignty/

Philip Malone. (2019, May 6). JohnF. Kennedy (CVN 79) Enterprise (CVN 80) & Unnamed (CVN 81)—Two Ship Buy. Sea Air Space Exposition. https://www.navsea.navy.mil/Portals/103/Documents/Exhibits/SAS2019/Capt%20Malone-SAS-05062019.pdf?ver=2019-05-06-200404-223

Politico. (2025, March 30). Hegseth calls Japan indispensable in the face of Chinese aggression. https://www.politico.com/news/2025/03/30/hegseth-calls-japan-indispensable-in-the-face-of-chinese-aggression-00259311

Rains, D. A., & Johnson, J. A. (1993). Naval Ship Affordability Through Machinery Modularity. U.S. Department of the Navy.

Rebuilding America's Maritime Strength with Senator Kelly and Congressman Waltz. (2024, September 25). Center For Strategic and International Studies.

https://www.csis.org/analysis/rebuilding-americas-maritime-strength-senator-kelly-and-congressman-waltz

Ronald O'Rourke. (2025, March 31). Navy Force Structure and Shipbuilding Plans: Background and Issues for Congress. Congress.gov. https://www.congress.gov/crs-product/RL32665

Rubeša, R., Fafandjel, N., & Koli, D. (2011). PROCEDURE FOR ESTIMATING THE EFFECTIVENESS OF SHIP MODULAR OUTFITTING.

Sam Lagrone. (2024, March 7). SECNAV Del Toro Tells U.S. Shipyards 'Invest More', Encourages Foreign Investment. U.S. Naval Institute. https://news.usni.org/2024/03/07/secnav-del-toro-tells-u-s-shipyards-invest-more-encourages-foreign-investment

Schank, J. F., Pung, H., Lee, G. T., Arena, M. V., & Birkler, J. (2005). Commercial Shipbuilding Techniques: Can They Be Applied to Warship Production in the United Kingdom? RAND Corporation. https://www.rand.org/pubs/research_briefs/RB9085.html

SECNAV Del Toro Travels to the Indo-Pacific to Further Maritime Cooperation, Explore Opportunities to Enhance Naval Capabilities. (2024, February 24). United States Navy. https://www.navy.mil/Press-Office/Press-Releases/display-pressreleases/Article/3686156/secnav-del-toro-travels-to-the-indo-pacific-to-further-maritime-cooperation-exp/

Shawn Harrison. (2024, August 2). Defense Regional Maintenance, Repair, and Overhaul (MRO) Playbook. Defense Acquisition University. https://www.dau.edu/tools/defense-regional-maintenance-repair-and-overhaul-mro-playbook

Shelby S. Oakley. (2024a, May 2). NAVY SHIPBUILDING: Increased Use of Leading Design Practices Could Improve Timeliness of Deliveries. (2024, May 2). U. S. Government Accountability Office. https://www.gao.gov/products/gao-24-105503

Shelby S. Oakley. (2024b, May 29). NAVY FRIGATE: Unstable Design Has Stalled Construction and Compromised Delivery Schedules. U.S. Government Accountability Office. https://www.gao.gov/products/gao-24-106546

Shelby S. Oakley. (2025a, February 27). SHIPBUILDING AND REPAIR: Navy Needs a Strategic Approach for Private Sector Industrial Base Investments. Government Accountability Office. https://www.gao.gov/assets/gao-25-106286.pdf

Shelby S. Oakley. (2025b, March 11). NAVY SHIPBUILDING: A Generational Imperative for Systemic Change. Government Accountability Office. https://www.gao.gov/assets/gao-25-108136.pdf

Shipbuilders Council of America (Ed.). (1937). Commercial Shipyards and the Navy. National Council of American Shipbuilders.

SHIPS for America Act of 2024, 118th Congress, H.R. 10493 (2024). https://www.congress.gov/bill/118th-congress/house-bill/10493

South Korea's HD HHI Inks MRO Agreement with the US Navy. (2024, July 11). Naval News. https://www.navalnews.com/naval-news/2024/07/south-koreas-hd-hhi-inks-mro-agreement-with-the-us-navy/

Technology Transfer, Disclosure, Export Controls, and International Programs Security. (2024, August 7). Defense Security Cooperation University. https://dscu.edu/sites/default/files/2024-08/07-chapter.pdf

The Biden-Harris Administration's National Security Strategy. (2022, October 12). The White House. https://www.whitehouse.gov/briefing-room/statements-releases/2022/10/12/fact-sheet-the-biden-harris-administrations-national-security-strategy/

Thompson, L. (2022, July 19). How To Slash The Time And Money Needed To Build Warships—Without Cutting Capabilities. Forbes. https://www.forbes.com/sites/lorenthompson/2022/07/19/how-to-slash-the-time-and-money-needed-to-build-warships-without-cutting-capabilities/

Trevithick, J. (2021, September 16). F-22 Export Briefing Shows What It Would Have Taken To Sell The Raptor Abroad. The War Zone. https://www.twz.com/42318/f-22-export-briefing-shows-what-it-would-have-taken-to-sell-the-raptor-abroad

Trump's call for 'K-shipbuilding' signals new opportunities in his second term. (2024, November 7). The Chosun Daily. https://www.chosun.com/english/opinion-en/2024/11/08/EGVWKW5OYRETZIUEFLH6LWDZFU/

U.S. Department of Defense. (2024a, May 15). Regional Sustainment Framework. https://www.acq.osd.mil/asds/docs/RSF-9MAY24.pdf

U.S. Department of Defense. (2024b, February 22). Regional Sustainment Framework (RSF) Fact Sheet. https://www.dau.edu/sites/default/files/2024-02/RSF%20INFO.pdf.

U.S. Department of Defense. (2024c, October 7). Second U.S. - Japan DICAS Meeting Conducted. https://www.acq.osd.mil/news/archive/2024/Second-US-Japan-DICAS-Meeting-Conducted.html

USNS Wally Schirra completes major maintenance at South Korean shipyard. (2025, March 13). U.S. Pacific Fleet. https://www.cpf.navy.mil/Newsroom/News/Article/4119656/usns-wally-schirra-completes-major-maintenance-at-south-korean-shipyard/

Vavasseur, X. (2021, October 12). South Korea's HHI Lays Keel of New KDX III Batch II Destroyer For ROK Navy. Naval News. https://www.navalnews.com/naval-news/2021/10/south-korea-new-kdx-iii-batch-ii-destroyer/

Navy Shipbuilding: Increased Use of Leading Design Practices Could Improve Timeliness of Deliveries

Sean Merrill—is a Senior Analyst with the U.S. Government Accountability Office, specializing in U.S. weapon system acquisitions. He has worked in this issue area for 20 years, with a primary focus on Navy shipbuilding issues for the past decade. This work has included reports and recommendations related to the Navy's destroyer and frigate programs as well as broader reviews of Navy ship design, testing, and shipbuilding oversight. [merrills@gao.gov]

Abstract

Changing maritime threats are pushing the U.S. Navy to increase its pace for designing and delivering new ships. Since 2009, GAO has used leading practices in commercial shipbuilding to evaluate the plans and execution of Navy shipbuilding programs. GAO's numerous recommendations have spurred Navy action to improve acquisition practices and the use of taxpayer dollars. Yet, the Navy has continued to face persistent challenges in its ability to design and deliver timely, affordable new ships that perform as expected. In response to the Navy's shipbuilding issues and interest in identifying how modern design practices support timely delivery of new ships, GAO completed a review to assess (1) the leading design practices used by commercial ship buyers and builders to inform their understanding of design maturity and readiness for construction, and (2) how the Navy's ship design practices compare to the leading practices in commercial ship design.

Leading Companies' Design Practices Support Timely and Predictable Ship Delivery

Commercial ship buyers and builders use four primary leading practices to enable shorter, predictable cycles for designing and delivering new ships (see Figure 1).

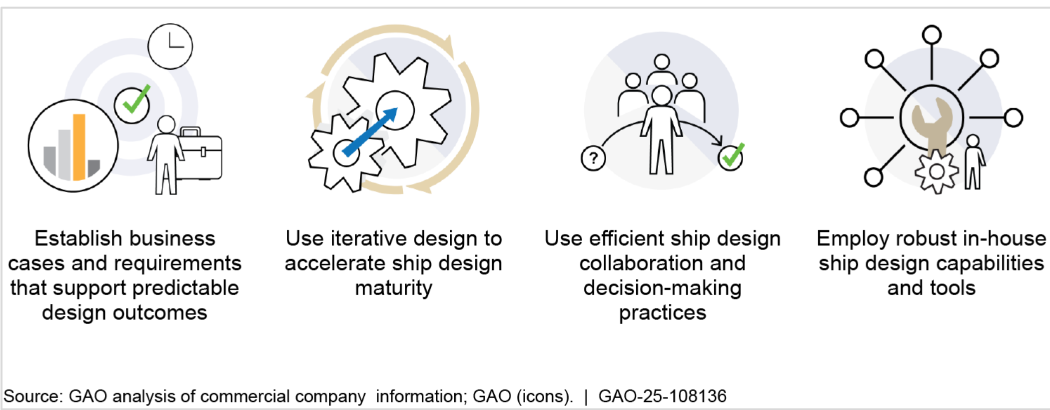


Figure 1. Primary Leading Practices GAO Found in Commercial Ship Design

Companies' Business Cases and Requirements Support Predictable Ship Design Outcomes

Prioritize Timeliness of Ship Design and Delivery

Leading commercial companies in ship buying and building have strong business cases that prioritize cycle time for ship design and construction over additional capability. These companies prioritize schedule because shorter periods for design and delivery help them preserve their business case and meet strategic business interests. Specifically, ship buyers and builders have an interest in compressing their design and build cycle time to avoid delivering ships with design features that are obsolete or no longer in demand by their customers. Predictability is also a fundamental element of their schedule prioritization. For both parties, delays to designing and delivering a ship as contractually agreed to pose unacceptable financial consequences.

For buyers, delays can prevent them from fulfilling obligations to their customers. Depending on the type of ship, these obligations can include honoring thousands of passenger reservations for a cruise ship vacation. They can also involve transport across oceans of hundreds of cargo containers full of consumer goods or hundreds of thousands of cubic meters of liquid natural gas. Commercial shipbuilders noted that the firm-fixed-price design and construction contracts that they agree to generally include significant financial penalties, such as liquidated damages, for late ship delivery.[1] Such penalties for delayed ship delivery could involve, for example, liquidated damages to the buyer that exceed $500,000 per day of delay.

For buyers, shorter design and construction cycles also support their interests in being the first to provide the latest innovative technologies and design features at sea for their customers. Further, shorter cycles hasten the start of buyers receiving a return on investment through the revenue received from customers once the ships begin operating. These financial considerations provide incentive for timeliness when considering large, complex ships can cost hundreds of millions of dollars and reach into the billions in some cases, such as with Royal Caribbean Group's recently delivered *Icon of the Seas,* with a reported cost of $2 billion.

GAO also found that short, predictable design and build cycles support commercial shipbuilders' interest in optimizing shipyard workflow and maintaining a steady design and construction workforce. In general, leading commercial shipyards have multiple ships under design and construction at any given time. The shipyards also typically have a backlog of new ship builds—for the same or different buyers—waiting to start design and construction. Under these conditions, a delivery delay for one ship can create a cascading negative effect on other ongoing and future builds at the shipyard and the builder's financial bottom line. As a result, builders' design decisions reflect the circumstances of their respective shipyards and their interest in upholding the schedule for designing and delivering new ships.

Avoid Overly Prescriptive Requirements

The practices commercial ship buyers use to establish requirements help preserve the builders' autonomy for decisions on how to efficiently design and construct ships that meet schedule, cost, and capability requirements. The requirements can include functional specifications, preliminary general arrangements, and ship renderings. Collectively, these requirements serve as the foundation for buyer and builder collaboration. This helps them to reach early agreement on key attributes of the ship design concept and to progressively define the final ship design. Buyers typically share requirements that capture high-level operational needs with prospective shipbuilders and collaboratively develop detailed requirements during iterative planning.

Buyers and builders use feedback from ship engineers and operators—as well as passengers in the case of cruise ships—to inform ship requirements for new designs. Before contract award, they also ensure both parties have a clear understanding of the relationship between requirements, cost, and schedule for each new ship design. This ship design practice is consistent with what GAO previously found leading companies across different industries do to successfully develop and deliver products to users with speed.

[1]For nongovernment contracts, a fixed-price contract is a type of contract in which the buyer agrees to pay the seller a definite, predetermined price, regardless of costs.

Maintain a Sound Business Case

As the pursuit of new ship designs and builds progresses, commercial ship buyers and builders regularly reassess their respective business cases. For example, a cruise ship buyer may determine that feedback collected from cruise ship passengers warrants a change in design to either add high-demand design features or remove less-valued features. Further, a cargo ship buyer may identify a changing business case based on feedback from ship operators, indicating opportunities to gain efficiencies in operations or maintenance from incorporating different equipment into ship designs. For any design decisions that may affect the delivery date, buyers and builders reach agreement on a way forward that aligns with their respective interests.

Prior to contract award, if a builder believes that a ship cannot be designed and constructed to meet the buyer's operational requirements and schedule and cost objectives, trade-offs must be made for the project to proceed. Such trade-offs can include removing or revising ship capability requirements, including innovative features that may carry outsized schedule or cost risk. They can also involve the buyer agreeing to take responsibility for all or portions of the development, testing, procurement, and installation of a ship's design features. In such cases, the buyer may also accept responsibility for any financial consequences or delays to the ship's delivery associated with those buyer-supplied design features.

Companies Use Iterative Design to Accelerate Ship Design Maturity

Leading commercial ship buyers and builders use iterative processes to efficiently establish requirements and designs focused on timely delivery of ships with capabilities desired by customers. Knowledge about the ship's design is progressively refined and documented through ship specifications, contract requirements, and design products supporting construction. As they proceed, the buyer and builder make design trade-offs as needed to support timely delivery of affordable ships that commonly operate at sea for decades delivering required capabilities. This approach incentivizes buyers to identify the capabilities needed for customers to recognize value in a ship's design and avoid chasing immature or expansive innovations to the detriment of timely ship delivery. These commercial ship design practices are consistent with broader leading practices for product development across different commercial industries. Specifically, these practices being used for commercial ship design reflect a cyclical process to determine what capabilities are achievable within a fixed period, design and deliver one or more ships with those capabilities, and repeat this process for successive ship designs.

Prioritize User Involvement in Design Process

Commercial ship buyers and builders prioritize user involvement in iterative design processes by obtaining and applying design input from ship operators and the broader user community. This includes direct ship operators' and engineers' involvement in the review of design models and drawings during design maturation. Additionally, commercial buyers and builders receive feedback post-ship delivery to inform designs for subsequent ships and modifications to operational ships. For cruise ships, buyers told GAO that they use their extensive market research—including passenger feedback from operational ships—to inform ship design decisions from the concept stage of the design process through to relatively late-cycle construction. This market research helps them make design decisions that align with user needs and expectations and helps ensure that cruise operators receive a return on their investment.

Chevron and Maersk provided other examples of how ship operators and engineers contribute to design reviews and decisions. Chevron uses its officer development program to involve first mates and engineers directly in the review of ship designs. The company sometimes also includes ex-chief engineers in its design teams to ensure operational perspectives are accounted for in designs. The operators and engineers review design drawings and contribute to the overall comments that Chevron provides to the shipbuilder. Chevron also performs “lookback” reviews, through which comments can be added to and preserved for design drawings as a form of lessons learned for use in future designs. Once ships are delivered, Chevron uses operational feedback, which includes lessons learned from incidents or near misses, to inform future designs. Maersk has “sea-to-shore” contracts with its captains and chief engineers, who are experienced ship operators, and assigns one of each position to the design review and approval team for new ship designs. These individuals will typically move with the approved design to the shipyard to serve as oversight during construction and then sail on the lead ship (or a retrofitted ship for smaller-scale design efforts) when it is delivered. This approach enables the personnel to experience the ship from the design stage to operations.

Leverage Existing Ship Designs and Systems

GAO found that commercial shipbuilders draw heavily from their respective libraries of existing ship designs and ship systems to speed design maturity and reduce risk. Use of proven ship designs and makers lists—which identify buyer-approved vendors for major equipment such as main engines and propellers—minimizes design, cost, and schedule uncertainties for buyers and builders. Use of existing ship designs and systems also supports earlier technical maturity for new designs and reduces the need to validate that designs or equipment meet vessel standards.[2] Further, use of existing ship design information helps companies incorporate maintenance and operations considerations in their new designs. Maintenance and operations contribute significantly to a ship’s total cost for its buyer, with much of the associated cost fixed at the time when requirements are set and the ship is designed. As a result, efforts to account for these factors in new ship designs support improvements to life-cycle costs for the ships.

Leveraging existing designs and mature equipment also creates opportunities for shipyards to use their prior experiences building to those designs and incorporating that equipment to create efficiencies in new ship construction. For example, Meyer Werft used its library of design data to create a high number of design iterations to determine how to optimize a new design for a recent Carnival cruise ship from a vast array of options. The company’s use of design iterations created flexibility that better enabled it to adapt the design if Carnival Corporation wanted to make changes during the design and construction cycle.

Commercial shipbuilders told GAO that using existing design and system knowledge enables them to start new ship designs with greater baseline design maturity. As an example, Samsung Heavy Industries uses its existing ship design library to identify a baseline design, or “mother ship.” This practice provides an optimal design with significant design maturity from the outset. Samsung Heavy Industries then works with the buyer to incorporate new design features that address the buyer’s specific needs not already addressed by the mother ship design. For Damen Shipyards Group, the company uses a stable, “Damen Standard” design to build some of its most highly in-demand ship classes without having a specific buyer. Damen stated that the company understands how to efficiently build a baseline ship and will tailor it to meet specific capability interests once the buyer is confirmed.

Prioritize Timely Vendor Decisions and Information

Commercial builders facilitate a shorter design and construction cycle by rapidly selecting vendors (i.e., equipment suppliers) and managing the timely receipt of associated

[2]The International Maritime Organization requires a ship’s design and construction to be approved by ship classification societies, such as the American Bureau of Shipping, Det Norske Veritas, or Lloyd’s Register. These societies (1) establish and maintain standards for the construction and classification of ships and offshore structures, (2) supervise construction in accordance with these standards, and (3) carry out regular surveys of ships in service to ensure the compliance with these standards.

vendor-furnished information (VFI). Builders noted that rapid selection can include reaching vendor agreements before contract awards or shortly thereafter, such as within 2 months. Commercial builders are incentivized to finalize agreements with vendors for equipment as early as possible to avoid design uncertainty or instability from having incomplete or unreliable VFI in ship designs. For example, Seatrium commonly identifies and selects equipment and vendors before the shipbuilding contract is finalized, noting this practice is especially important for complex ship designs that include unique mission equipment—such as pedestal cranes for heavy lift vessels—for which vendor options are limited.

Prompt vendor selection also helps commercial ship buyers or builders expedite any additional development and testing equipment vendors need to complete to meet the needs of the new ship design and establish reliable VFI. An example of reliable VFI would be having finalized specifications for a piece of equipment but awaiting the results of factory acceptance testing to validate those specifications through manufacturing. Shipbuilders told GAO that, until vendor agreements are reached, the best available VFI could involve basic specification sheets that provide limited details on the characteristics for previous models of equipment. Builders noted that delays in obtaining reliable VFI constrain ship design progress and can negatively affect the builder's readiness for construction and ship delivery schedule.

Make Risk-Based Decisions to Off-Ramp Design Features

Commercial ship buyers and builders told GAO they use off-ramping practices to support decisions that remove or amend design features or specifications from new ship designs. This includes decisions to exclude design features through collaborative efforts between ship buyers and builders prior to contract awards as well as changes after contract awards. Use of off-ramping can occur when the design feature presents significant risk to achieving the ship delivery date. It can also occur when risk identified from a business case change supports removing design features from the ship's design, such as with the previously discussed cruise ship restaurant example.

In cases where a design feature is removed or significantly changed, that feature can be deferred to future commercial ship designs. Companies perform risk assessments in these instances and may decide to defer the feature because they determine that including it in the design poses an unacceptable risk to meeting the objectives of the existing build. For example, cruise ship buyers and builders noted cases where the buyer may desire an innovative design feature not explicitly defined in contractual requirements that cannot be achieved within the agreed to ship delivery schedule. In such cases, the builder typically works with the buyer to find a solution that aligns with the existing schedule. The builder and buyer will also discuss using the desired design feature in future ships when the longer lead time required to incorporate that feature can be accounted for in up-front decision-making.

Minimize and Isolate Changes to Existing Designs

Commercial shipbuilders isolate changes within the total ship design to maximize the value of using an existing design as their foundation for new ship designs. This approach helps preserve design maturity and reduces total work required for new ship designs. For example, Fincantieri officials told GAO that the company reduces design time and design labor hours by 90% or more for "sister" ships—a second ship on the same contract—by carrying over most of the previously validated design of the first ship to the sister ship design. By managing design changes in a manner that minimizes the amount of ship spaces affected, commercial builders and buyers limit total risk to the ship design and maximize the shipyard's experience in building to the prior ship design. This practice supports shorter design and construction cycles as well as more predictable cost and construction performance.

For example, as part of the company's efforts to become carbon neutral, Maersk explored existing green technology options for its shipping vessels. As part of these efforts, the company identified an opportunity to use methanol-based technology to power a new class of ships. To develop a ship design that included methanol-fueled technology, Maersk worked with Hyundai Heavy Industries—which had used similar technology in tanker vessels—to use an existing container ship design already operating in Maersk's fleet. The resulting design—which includes dual-fuel methanol- and conventional-fueled systems—limited total ship design changes to those areas of the ship where the new methanol-fuel system is integrated. The lead ship, Laura Maersk, was delivered roughly 2 years after contract award and began operations in 2023.

Carefully Manage Design Innovation

In general, significant innovation—which can include novel design features and advanced technologies—must be technically mature for a commercial shipbuilder to agree to include it in the design. This means that the innovation must be well understood and proven—which can be accomplished through its use on other ships or formal testing, such as physical or digital prototyping.

Commercial buyers and builders also told GAO that they limit the amount or scale of novel design features they are willing to include in a ship design as part of their risk management. Royal Caribbean noted that financial factors play a role in bounding the number of new features that can go into a ship, with a finite amount of money available for such features given all the baseline costs involved with any new cruise ship. Two other buyers noted a clear link between introducing innovations and maintaining shorter cycles for design and construction. One of those companies added that its responsibility as the buyer is to ensure the timing of its orders support delivery of the ships by a certain date, so if the company wants ships sooner, it can consider a more standard ship design. One company also noted that too many innovations in a ship design can undermine the builder's ability to maximize its business model and more rapidly design and build ships.

GAO found that buyers—particularly of cruise ships—will sometimes pursue design innovations through an iterative design process that informs final requirements for reserved areas, or "white spaces," in designs. For these undefined design elements, determined prior to contract award, the buyer works with the builder and vendors, as well as a classification society when needed, to validate compliance with technical standards and finalize detailed design requirements.

Companies Use Efficient Ship Design Collaboration and Decision-Making Practices

Use Processes That Support Timely Design Decisions

GAO found that commercial ship buyers and builders use consistent, effective collaboration to support timely decision-making practices from design concept to ship delivery. Their use of extensive up-front communication establishes a common understanding of ship requirements, schedule, and cost before contract award, which hastens design maturity. This collaboration includes candid conversations between ship buyers and builders at the concept stage regarding what can and cannot be reasonably incorporated into a design based on technical, cost, and schedule parameters. Seatrium stated that, as the ship designer and builder, it uses early engagement with buyers to ensure the company's understanding of the buyer's requirements. Seatrium also uses this early engagement with buyers to identify key factors that will affect the ship's design, such as requirements for a vessel to achieve a certain speed, as early as possible, which minimizes potential issues in later stages of the design and construction cycle.

The decision-making processes employed by commercial ship buyers and builders are also designed for efficiency. For example, Royal Caribbean told GAO that it uses measurements of risk to determine responsibility for decision-making. For higher risk design elements, the program manager for the new ship is the primary decision-maker. For lower-risk design decisions, the company supports timeliness by delegating authority to lower working levels, such as an assistant project manager for a specific design element of the ship.

Commercial ship buyers and builders also told GAO that their design and construction contracts—which include firm-fixed prices and fixed ship delivery schedules—include a period typically ranging from 10 to 21 days for buyers to review and comment on design products. They added that design products requiring buyer approval, such as drawings or other design deliverables, may be considered approved by default if the ship buyer does not respond within the period agreed to in the contract. These typical expectations for design review support a timely process for maturing designs to support construction. As ship design updates are requested and accepted, commercial buyers and builders maintain steady communication with each other, enabled by access to a shared electronic communication platform. The platform provides a real-time means for conveying design decisions among stakeholders and access to information related to the ship design. The overall collaboration and decision-making practices used by these companies allow them to efficiently decide how, if at all, to incorporate design updates without significantly disrupting the overall design and ship delivery schedule.

Align Decision-Making with Design Maturity Measures

Commercial ship buyers and builders ensure key decisions are closely linked to consistent measures of design maturity and associated effects on construction readiness. Although GAO found some variation among companies in how much of the total ship design must be completed before they will begin construction, they consistently expect a high degree of design maturity to proceed with construction. For example, Damen told GAO the company completes the full detail design before starting construction for the first ship in a new class. Samsung Heavy Industries expects at least 90% of production design drawings to be completed at the time of its ship model gate review that supports a decision to begin construction—only smaller design elements can remain unfinished.

Overall, GAO found that commercial ship buyers and builders only begin construction when design maturity and related measures demonstrate their readiness to do so. To ensure such readiness, companies set and uphold expectations that (1) basic and functional design will be fully 3D modeled with reliable VFI included to achieve design stability before construction begins; and (2) at a minimum, detail design for any given block of the ship will be completed prior to beginning construction of that block.

Table 1 provides more details on key tasks in different design phases that support the leading ship design practices GAO found being used by commercial ship buyers and builders.

Table 1. Leading Practices for Commercial Ship Design

Design phase	Key tasks involved
Basic and functional design	• Fix ship steel structure and set hydrodynamics • Design safety systems and get approvals from applicable authorities • Route all major distributive systems, including electricity, water, and other utilities • Provide information on position of piping, ventilation, equipment, and other outfitting in each block • 3D model the ship structure and major systems, with reliable vendor-furnished information (VFI) incorporated to support understanding of final system design. Reliable VFI reflects a firm understanding of the characteristics for ship equipment and components, including requirements for space, weight, power, water, and other utilities. An example of reliable VFI is a piece of equipment with finalized specifications but awaiting the results of factory acceptance testing to validate those specifications through manufacturing.
Design stability achieved upon completion of basic and functional design	
Detail design	• Use 3D modeling information to generate work instructions for each block—basic unit of ship construction—that show detailed system information and support construction; includes guidance for subcontractors and suppliers, installation drawings, schedules, material lists, and lists of prefabricated materials and parts • At a minimum, complete detail design for any given block of the ship prior to beginning construction of that block

(Source: GAO analysis of commercial ship design information.)

Companies Employ Robust In-House Ship Design Capabilities and Tools

Maintain Strong In-House Design Workforce Capabilities

Commercial ship buyers and builders maintain strong in-house ship design capabilities. Doing so ensures both sides have a firm and common understanding of the ship design concept and required performance before agreeing to contracts that lock in ship prices and delivery dates. In general, commercial shipbuilders in GAO's review employ an extensive amount of personnel to support ship design efforts. For example, Damen has the equivalent of over 1,100 personnel involved in its design and engineering for first-in-class and single-ship designs. Commercial shipbuilders use their own personnel to perform most of the design work for the ships they build. For detail design, builders noted that their in-house expertise supports decisions that align the ship's design with the shipyard's characteristics to create an efficient build strategy.

Commercial ship buyers use in-house resources to develop design concepts and evaluate the builders' design proposals, development, and execution during construction. For example, Royal Caribbean personnel complete engineering feasibility and packaging assessments and architectural design work—including for buyer-supplied equipment—before finalizing contract awards. Royal Caribbean's department for new ship builds creates a specific team for each project that typically includes a project manager, outfitting manger, technical engineering manager, and area managers for different portions of the ship designer. One buyer noted that having robust in-house resources to advance a design through functional design provides the company with a firm understanding of how design affects cost, which helps set achievable expectations and supports better decisions. As another example, Maersk has a team of about 100 engineers to support its ship design activities at its offices and on-site at builder shipyards. Within this engineering team, 10% of personnel specifically focus on new concept development for ship design and innovation. These personnel regularly leverage subject matter expertise within Maersk's overall engineering team for specific functional design aspects to support design development and oversight.

Use Ship Design Tools to Shorten Cycle Time

GAO found that commercial ship buyers and builders use advanced 3D modeling and—to varying degrees—other modern ship design tools to accelerate design maturity and support efficiencies in design and construction. Overall, they noted that their use of modern digital design tools creates efficiencies for design validation, optimization, and completion, among other benefits. For example, Samsung Heavy Industries uses a paperless system to manage ship design and construction. The system combines 3D modeling and scheduling information to produce what Samsung refers to as "4D" modeling. The system is available on mobile devices throughout the shipyard to enable digital access to design drawings and models for use in construction. Samsung also uses augmented reality tools that enable personnel to overlay 3D modeling on actual construction work to evaluate results against design. Damen uses its Triton "internet of things" platform to enable access by the company and others, such as suppliers or ship owners, to specific data on system performance. The Triton platform provides a dashboard where data from onboard ship sensors can be leveraged for real-time or point-in-time data extraction and analysis. This information can be used to optimize ship designs.

Commercial companies have used advances in 3D modeling capabilities since GAO's 2009 work on shipbuilding practices to increase the amount of design knowledge in modeling and its availability to stakeholders. The 3D modeling systems can increase design efficiency by, for example, customizing the systems to automatically route pipes and electrical cable trays in accordance with preconfigured rules for the ship design. Modern digital engineering, product life-cycle management, and enterprise resource planning systems have also contributed to improved design processes. For example, Fincantieri's engineering tools perform automatic checks between technical specifications and materials used for modeling. The checks identify any inconsistences and focus on data and 3D model updates to support design changes as opposed to updating 2D drawings. Collectively, these systems enable commercial builders and buyers to refine, store, and communicate design and requirements information that helps stakeholders make decisions throughout the life cycle for a ship's design and construction.

The advances in tools supporting commercial ship design enable builders to mature basic, functional, and detail design earlier in the overall project cycle than previously achieved with less capable tools. These advances help builders achieve the leading ship design practice of complete 3D modeling of all basic and functional design before starting ship construction. When combined with reliable VFI, the 3D modeling capabilities that commercial builders employ help reduce design uncertainty prior to construction and improve cost and schedule predictability.

Commercial ship buyers and builders varied in their use of other modern design tools that provide virtual representations of physical products—referred to as digital twins—and virtual or augmented reality that immerses users in a virtual environment using head-mounted displays or other technology, to support ship design and construction. Some builders were using virtual or augmented reality tools for activities like testing ship design ideas and virtual walk-throughs of the ship design. For example, one builder tests the company's design ideas in a virtual environment—using virtual reality in certain cases—from the initial ship design to the production of the final vessel. The company noted that this approach saves time and money as well as enables constant delivery of new innovations to the ships it designs and builds.

GAO found commercial ship buyers and builders view digital twinning as an area of opportunity for future ship design, with present use limited to twinning of ship systems or shipyards rather than entire ships. GAO's work on leading practices in product development highlights the use of digital twins as a tool to support testing and validation of a product's integrated functionality in its operating environment. For example, Chevron is using digital twinning models to analyze the effects of different loading and damage scenarios and the impact of grounding, flooding, and collision on the ship. One builder has also used digital twinning for virtual commissioning, verification, and validation for new designs.

Cumbersome Practices and Ship Design Capability Limitations Challenge the Navy's Ability to Improve Timeliness

Navy shipbuilding programs often take significantly longer to design and deliver new ships compared to the typical timelines for commercial ships. GAO found that several factors contribute to the differences in the pace of ship design and delivery, as shown in Figure 2.

Establish business cases and requirements that support predictable design outcomes

Commercial	Navy
• Prioritizes timeliness of ship design and delivery • Avoids overly prescriptive requirements • Maintains a sound business case through continued reevaluation	• Progresses through an extensive requirements process, with significant time elapsing before detail design and construction contracts • No regularly required reevaluation of approved requirements to confirm their continued relevance

Use iterative design to accelerate design maturity

Commercial	Navy
• Ensures schedule, cost, and requirements expectations are informed by sufficient design knowledge • Prioritizes user involvement in the ship design process • Leverages existing ship designs and systems in digital libraries • Prioritizes timely vendor decisions and information	• Sets expectations for schedule, cost, and operational requirements when design is unstable, resulting in less design knowledge available to inform key decisions and increased program risk • Generally uses a longer, more linear approach—with less consistent user involvement—focusing on new designs with extensive and novel capability rather than speed to delivery • Makes some use of existing ship designs, but lacks a robust design library to support iterative design and shorten time needed to mature new designs • Generally takes extended time to finalize vendor decisions for ship systems and receive vendor-furnished information needed to mature ship designs

Use efficient ship design collaboration and decision-making practices

Commercial	Navy
• Uses processes that support timely design decisions • Aligns decision-making with design maturity measures	• Lacks streamlined, more time-constrained processes, with numerous stakeholders having decision-making authority and contributing to extended cycle times to finalize designs • Lacks consistent design maturity measures and a clear connection between those measures and decision-making

Employ robust in-house ship design capabilities and tools

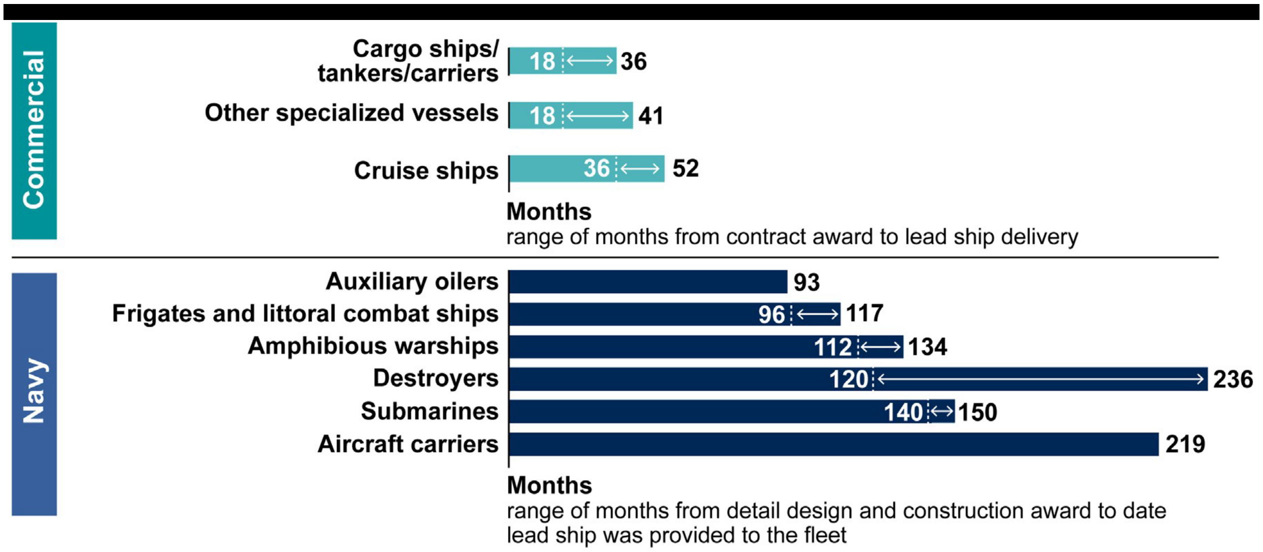
Commercial	Navy
• Maintains strong in-house design workforce capabilities • Uses ship design tools to shorten cycle time	• Evaluating ways to address acknowledged shortfalls in its in-house design workforce and tools • Adopting modern design tools to varying degrees, with the potential for expanded, more consistent use to provide efficiencies that support shorter, more predictable cycle times for ship design

Source: GAO analysis of commercial company and Navy information; GAO (icons). | GAO-25-108136

Figure 2. GAO Comparison of Leading Ship Design Practices for Commercial Companies and U.S. Navy

Long Cycle Times Increase Program Risks for New Ship Designs

GAO found notable contrast in the design and construction cycle times that is typical for selected types of commercial ships compared to the lead ships for Navy shipbuilding programs, as shown in Figure 3.



Source: GAO analysis of commercial company and Navy information. | GAO-24-105503

Figure 3. Comparison of Design and Construction Cycle Times for Selected Commercial and Navy Ships

Notes: "Commercial other specialized vessels" includes ship types such as offshore support vessels, ferries, icebreakers, tugboats, and research and science vessels. For Navy ships, the number of months indicate the shortest and longest periods for the Navy to provide selected lead ships to the fleet since 2007. GAO measured Navy cycle times based on the actual obligation work limiting date (OWLD), or planned date for lead ships that had yet to reach OWLD. OWLD generally coincides with when a Navy ship is provided to the operational fleet. Since GAO found that commercial ships typically enter operation soon after delivery, Navy OWLD provides the best proxy for comparison to commercial delivery dates. For Navy programs that had a contract prior to the detail design and construction award, GAO used that contract award date as the start of the cycle.

A lengthy cycle time creates business case challenges as threats and mission needs can change. For example, 11 years elapsed between the start of the DDG 1000 program and construction beginning on the lead ship. During that time, the Navy shifted from a focus on capability needs for operations in nearshore waters to deeper water operations. With this shift, the Navy determined that the DDG 51 class of destroyers would be a more effective option to meet operational needs and reduced the total DDG 1000 class from 32 to three ships.

Requirements Practices Hinder Business Cases and Ship Design Maturity

The extensive process used by the Navy to establish capability requirements for new ships contrasts significantly with the typical commercial process used to efficiently move from basic requirements to specifications that support a contract award for ship design and construction. Specifically, Navy shipbuilding programs progress through a protracted process to solidify requirements in the capability development document (CDD) prior to contract award for detail design and lead ship construction. The CDD outlines the operational requirements that will deliver the capability to meet operational performance expectations for the ship. The Navy's acquisition guidance also includes gated reviews intended to ensure that requirements align with acquisition plans. These reviews support the Navy's efforts to develop and endorse capability requirements before submitting them for Joint Staff review.

The overall requirements setting process leads to significant time elapsing before Navy shipbuilding programs can move forward with contract awards for detail design and construction. For example, it took over 4 years from when the Navy initiated its pursuit of DDG 51 Flight III to validate its CDD. This included 2 years between the Navy's CDD approval at the program's third gate review and the Joint Requirements Oversight Council's CDD validation. DoD's guidance for the Joint Capabilities Integration and Development System portion of the CDD review and validation process indicates that it should be accomplished in no more than

103 calendar days. However, GAO's prior work reviewing this process found that none of the DoD programs GAO reviewed completed the process within this time. That work also found a variety of issues that could affect the length of elapsed time, with the comment adjudication period cited by Joint Staff officials as the biggest contributor to the length of reviews.

In addition to timeliness issues, GAO found that the Navy's processes do not require confirmation of the continued relevance of its business case—a leading practice—through formal reevaluation of CDDs during ship construction or prior to the start of construction for each ship. Specifically, the Navy's acquisition guidance includes a gate review after detail design and construction contract award to endorse any CDD updates. However, the guidance does not require that the Navy proactively continue to assess its business case supporting approved capability requirements as a shipbuilding program progresses. The lack of such a requirement limits formal opportunities to identify changes that could improve the capability delivered to the fleet. It also increases the risk of the Navy investing resources in ship designs with capabilities that are no longer needed.

A recent law requires DoD to develop and implement a streamlined requirements development process.[3] However, GAO identified some steps that the Navy has already taken for its recent shipbuilding programs to improve the requirements process, which are also consistent with leading practices. Specifically, Navy officials said that they have focused on increasing communication with prospective shipbuilders during requirements setting and conceptual design activities. They have also held requirements open later into the acquisition cycle for more recent shipbuilding programs. This helps the Navy and builder increase their understanding of the requirements' effect on design, schedule, cost, or other factors before finalizing the CDD. Navy officials told GAO that communication with shipbuilders can help shape requirements and design to get a ship with desired capability at a reduced cost by leveraging the builders' knowledge of available innovations and current shipyard capabilities. These efforts support improvements to requirements setting and early design that could contribute to more predictable program outcomes for future ship classes.

Linear Acquisition Approach Increases Cycle Times for New Ships

The Navy generally uses a longer, more linear approach to design and deliver new ships that contrasts to the iterative design practices that GAO found in use for commercial ship designs. This linear approach defines and locks down requirements relatively early, and development focuses on compliance with original requirements. The Navy's approach also focuses on designing and delivering extensive, and often novel, capability with the lead ship, with reduced emphasis on the length of time needed to deliver the ship compared to commercial practices.

For instances of major design changes to existing ship classes—such as those included in DDG 51 Flight III and LPD 17 Flight II—the Navy treats them much like new shipbuilding programs, with linear requirements setting and design maturation processes. This leads to a considerable amount of time elapsing before a lead ship is delivered to the fleet. For example, about 14 years elapsed between the Navy's decision to pursue DDG 51 Flight III and its June 2023 acceptance of lead ship delivery.

As part of the linear approach used for its shipbuilding programs, the Navy measures results against an acquisition cost, schedule, and performance baseline. GAO found challenges with the Navy setting these baselines for programs before achieving a stable design for the new

[3]Section 811 of the National Defense Authorization Act for Fiscal Year 2024 requires that, by October 1, 2025, the secretary of defense develop and implement a streamlined requirements development process for DoD to improve alignment between modern warfare concepts, technologies, and system development and reduce the time to deliver needed capabilities to warfighters.

ships. Specifically, DoD policy requires that Navy shipbuilding programs receive approval for their acquisition program baseline—which outlines capability, cost, and schedule requirements—before awarding a detail design and construction contract for the lead ship. However, the Navy generally does not work with builders to achieve design stability before setting these baseline requirements and awarding these contracts. Instead, the Navy commonly defers significant amounts of basic and functional design work—which provides such stability—until after the detail design and construction contract awards. For example, shortly after the detail design and construction contract award for FFG 62, the program office stated that most of the ship's design drawings for basic and functional design remained incomplete.

As a result of setting baseline requirements without a stable ship design, key decisions for Navy shipbuilding programs are informed by less design knowledge than what commercial ship buyers and builders expect to have before entering into contracts. Further, the Navy's approach poses greater risk that the business case for its new ships will erode because cost, schedule, and capability requirements are set before the design has sufficiently matured to support more predictable outcomes.

Limited User Involvement

GAO found less consistent and direct involvement of ship operators and engineers in the Navy's ship design activities compared to commercial practices. The Navy has extensive guidance to support its ship design management and ensure the human component—operators, maintainers, and support personnel—is reflected in design. This guidance supports the Navy's establishment of ship design teams with extensive subject matter expertise in the design and engineering of ships. However, GAO found that this guidance does not explicitly include the type of consistent user involvement employed in commercial ship design—such as the inclusion of ship operators on design teams and in direct design reviews—to incorporate user input in design decisions.

Further, Navy shipbuilders indicated direct user involvement in the design process varied. For example, one builder stated that the Navy's end users for new ships have little or no involvement in the design process unless such involvement is explicitly included in the contract requirements. In contrast, another Navy shipbuilder told GAO that ship operators and maintainers are consistently involved in the 3D model review process for ship designs, providing lessons learned for consideration. Without consistent practices to ensure direct user involvement in design efforts across Navy shipbuilding programs, the Navy falls short of leading practices and increases its risk of design decisions that do not fully account for the needs of its sailors.

Inconsistent Off-Ramping Practices

In another contrast to commercial practices, the Navy has a history of remaining committed to its pursuit of originally approved capability requirements on the lead ship when technical, cost, schedule, or other business case issues arise, rather than deferring desired capability to future designs. As GAO previously found, the Navy's lack of adaptability has proven particularly challenging when pursuing ambitious requirements for ships that require innovations that have yet to be proven out.

Further, GAO found that, when the Navy has decided to off-ramp design innovations, it has been after it made significant investments. For example, the Navy invested hundreds of millions of dollars to develop the remote multi-mission vehicle systems for the Littoral Combat Ship before replacing them with a different system due to performance shortfalls.

Limited Design Library

The Navy makes some use of existing designs but lacks a digital design library like those used by commercial industry to support iterative design and shorten the time needed to mature new designs. The limitations of the Navy's library reduce the range of existing ship designs that the Navy can leverage to evaluate and optimize baseline designs for its new ships. It also hampers the Navy's ability to expedite design and construction by increasing initial design maturity for new ships. A senior Navy official noted that, while the Navy has a solid digital library for ship systems and components, its library is more limited for ship designs. The official also said the Navy would benefit from a more expansive library of ship designs but noted that developing one would likely require a collaborative effort with Navy shipbuilders. He cited builders' intellectual property interests for their respective ship designs as a reason for needing collaboration.

Challenges with Timely Vendor-Furnished Information

In addition to design library limitations, GAO found that the Navy generally incorporates reliable VFI in its ship designs later than commercial ship buyers and builders. The companies' speed compared to the Navy stems from efficient processes for finalizing vendor agreements, regular adoption of equipment in use on existing commercial ships, and intolerance for including immature technologies in commercial ship designs. Navy shipbuilders commonly make vendor decisions after the award of detail design and lead ship construction contracts, with extended time elapsing in some cases before vendor finalization. Causes of delay include the lack of an existing relationship between the shipbuilder and vendors requiring more time to reach agreement. Navy practices add time to the design cycle by delaying the start of any development efforts needed for equipment to meet Navy requirements. They also delay the receipt of reliable VFI needed to mature the ship design. Without timely receipt of reliable VFI, design maturity is limited by inaccurate or incomplete design information, which could result in design and construction rework if the actual specifications vary significantly from estimates.

Decision-Making Practices and Inconsistent Design Maturity Measures Affect Timeliness and Risk

GAO found that the Navy and its shipbuilders generally have less direct communication prior to contract award than commercial ship buyers and builders. GAO's prior work found that shipbuilders may communicate less openly when the request for proposals process is the primary means for communication with the Navy in order to preserve their competitive interests. Reduced early communication increases the risk of shipbuilders and the Navy experiencing challenges post-award due to a lack of common understanding about requirements. The Navy has worked to increase early communication in recent programs, such as with the FFG 62 frigate and DDG(X) destroyer. This includes awarding multiple contracts to prospective builders for the early design phase. This approach is intended to enable greater communication and collaboration before decisions are made on contract awards for detail design and lead ship construction.

Extended Stakeholder Involvement in Decision-Making

The Navy's decision processes for new ship designs lack the streamlined and more time-constrained processes GAO found commercial ship buyers and builders use to reduce cycle times for ship design. Instead, Navy shipbuilding programs have many stakeholders with the authority to affect design decisions. This can prolong timelines for design decisions. Interoperability requirements for ships across the Navy's fleet can create design demands not present for commercial fleets that necessitate additional stakeholder involvement in design decisions. Still, timely decision-making for commercial ship design is supported by empowering project leaders to make most decisions without layers of stakeholders needing to weigh in. This approach is consistent with leading ship design practices as well as broader leading practices for product development identified in prior work.

As an illustration of the extended Navy timelines, GAO found through an assessment of selected Navy ship design and construction contracts that they generally allotted anywhere from 21 to 60 days for the Navy to review and respond to ship design documentation submitted by shipbuilders. In contrast, the longest typical timeline any commercial ship buyer and builder in GAO's review identified for these activities was 21 days. Additionally, Navy officials noted instances where the Navy and builder agreed to extend design review periods when additional time is needed. With these review timelines potentially applying to hundreds of contractually defined design products for a shipbuilding program, timeliness of design approval can weigh on the pace of design progress and contribute to a longer design cycle for Navy programs.

Navy officials noted that design decision-making is challenging because the Navy often manages key technologies as unique programs. As a result, shipbuilding programs do not have control of all the systems on the ships. Coordinating with these different programs to reach a decision for a ship's design can be time-consuming. Navy officials also told GAO that the number of stakeholders has grown over time due to risk aversion—principally the risk of overlooking key factors when making program decisions—and challenges with ensuring a single stakeholder has sufficient knowledge of all systems to support decision-making and accountability. Navy shipbuilders agreed that many design decisions require layers of Navy review or consensus of many stakeholders for approval, which results in an administratively burdensome and time-consuming process. For example, one shipbuilder noted that design changes can sometimes take weeks or months to finalize because of the Navy's layers of technical review that support decision-making, and the associated internal coordination required to make such decisions.

The Navy's recently acknowledged shortfalls with its in-house ship design capability further contribute to its timeliness challenges for design decision-making. Specifically, in May 2023, the acting assistant secretary of the Navy for research, development, and acquisition stated that the department did not have the ability to fully execute a Navy-led ship design due to, among other factors, workforce deficiencies. Navy officials told GAO that significant reductions to their design-related workforce over time affected the Navy's timelines for evaluating design products and resolving design issues. For example, a senior Navy official told GAO that, instead of the 10 technical experts and 10 supporting staff that the Navy had in the past to review hydrodynamics for all surface ship designs, the Navy currently relies on one technical expert for these reviews. The official stated that similar circumstances exist for reviewing general arrangements for ship designs. Beyond the workforce capacity considerations, Navy officials noted that a significant loss of experience and institutional knowledge within the Naval Sea Systems Command negatively affects the command's in-house ship design capability.

Inconsistent Connection between Design Maturity Measures and Decisions

The Navy's ship design practices have a less consistent and clear connection between design maturity data and decision-making compared with commercial practices. When evaluating design maturity and making decisions on construction readiness, commercial companies generally focus on key ship design knowledge attained—including design product approvals, VFI completeness, and material availability for construction—rather than calculations of design completion. Use of this information at key decision points in the design cycle helps the buyer and builder ensure a clear understanding of existing maturity and remaining risks.

The Navy's design maturity expectations and results vary across shipbuilding programs. For example, GAO found that programs were mixed as to whether they set an expectation that basic and functional design would be completed before starting ship construction. GAO similarly found variation in whether the programs achieved 100% completion for basic and functional design before beginning ship construction. GAO also found that Navy shipbuilding programs generally do not expect complete 3D modeling of basic and functional design before ship construction begins, which is inconsistent with leading ship design practices.

The Navy has taken some actions in recent shipbuilding programs to formalize design maturity measures. For example, GAO found that several Navy shipbuilding programs set thresholds for the degree of design maturity they require before deciding to begin ship construction. How programs measured their achievement of these thresholds varied but typically reflected percentages of design drawings or design-specific contract deliverables expected to be submitted at key milestones. Navy shipbuilders noted that using this type of metric does not necessarily provide a clear understanding of overall design maturity. For example, the metrics may overstate design completeness by giving builders credit for submitting design-related documentation without fully accounting for the quality or completeness of associated design. Drawings that appear complete could include design placeholders that lack necessary VFI for key equipment and, consequently, mask design uncertainties and remaining design work. Further, Navy officials noted cases where builders submitted blank design products, which met the submittal deadline to the Navy but did not contribute to advancing design maturity.

A recent law emphasizes the role design maturity should play in Navy decision-making and could help better align its shipbuilding program activities with the leading ship design practices. Specifically, the National Defense Authorization Act for Fiscal Year 2022 required the secretary of the Navy to certify to congressional defense committees the completion of basic and functional ship design before approving the start of construction for the first ship. The Act also required the secretary of the Navy to provide these committees certain design maturity information as part of its production readiness review reporting and certification.[4]

The Navy stated that it has not issued any guidance on its approach to evaluating design maturity for programs to support these statutory design certification and reporting requirements. Navy officials also told GAO that they have no plans to issue such guidance. Instead, they said that they use engineering judgment to establish working definitions for what a major shipbuilding program must achieve to meet the statutory requirement to certify completion of a ship's basic and functional design. They added that shipbuilding programs can choose how to define detail design.

The lack of Navy guidance to support the statutorily required certification and production readiness review reporting on design maturity increases the potential for confusion and inconsistencies in the Navy's approach to fulfilling these requirements across its shipbuilding programs. For example, the secretary of the Navy certified in August 2022 that the FFG 62 frigate program had completed basic and functional design, as defined by the statute. The certification included technical data showing 90% of the frigate's functional design was completed before beginning construction, which is counter to leading ship design practices. Navy officials told GAO that the statutory definition of basic and functional design includes a subset of the overall design characteristics that the Navy reviews and considers when determining readiness for ship construction. They also stated that the Navy requires a more

[4] National Defense Authorization Act for Fiscal Year 2022, Pub. L. No. 117–81 (2021), § 1013 (codified at 10 U.S.C. § 8669c). Section 8669c(a) of title 10, United States Code requires the Secretary of the Navy to submit a report to the congressional defense committees on the results of any production readiness review before approving the start of construction for the first ship for any major shipbuilding program.

rigorous level of design maturity than what is required by the statute's basic and functional design definition. Navy officials said that these factors and other metrics tracked by the FFG 62 program supported certification that basic and functional design—as defined by statute—was complete.

While the Navy's approach meets the statutory requirement to certify completion of basic and functional design, the FFG 62 certification and production readiness review reporting did not demonstrate the type of clear connection between design maturity data and decision-making expected by leading practices to support construction readiness. Further, subsequent functional design problems encountered by the FFG 62 program, which have contributed to cost and schedule issues for the lead ship, raise concerns about the Navy's approach to measuring functional design maturity.

Limitations in In-House Ship Design Capabilities and Tools Hinder Timeliness

As previously discussed, the Navy has acknowledged shortfalls in its design workforce, which contrasts to the significant in-house design capabilities that GAO found typical of commercial ship buyers and builders. The Navy's workforce shortfalls present challenges to minimizing the overall cycle times for ship design and effectively managing design risk for design and construction. In recognition of the challenges, the acting assistant secretary of the Navy for research, development, and acquisition initiated activities in May 2023 to improve the Navy's in-house ship design capabilities and enable the Navy to effectively lead ship design efforts.

In December 2023, the Navy confirmed that it had developed a draft strategic plan focused on reinvigorating the Navy's in-house ship design capabilities. The draft plan's high-level objectives include strengthening the Navy's technical community to support in-house design capabilities; better aligning Naval Sea Systems Command and other Navy organizations to support efficient and effective design efforts; and establishing new ship design team facilities at certain Navy locations. Navy leadership stated that, without a reinvigorated Navy ship design capability, the department risks overreliance on shipbuilders for design work. Further, the Navy will remain challenged in its ability to reduce the cycle time for design and construction and effectively manage design risk. GAO plans to monitor the Navy's progress in finalizing a strategic plan to address the identified design shortfalls and the Navy works toward implementing that plan.

For design tools, GAO found commercial ship buyers and builders and the Navy and its builders using a range of digital 3D modeling applications to mature ship designs. Similar to commercial companies, Navy shipbuilders GAO spoke with noted significant advancements in recent years with 3D modeling capabilities and the integration of design data from other systems in the models. However, Navy shipbuilding programs generally encounter more challenges in integrating 3D modeling with other information systems to enhance the depth of knowledge available to stakeholders. The challenges include incompatible systems and continuing use of 2D design information for legacy ship classes, such as *Arleigh Burke* destroyers and *Virginia* class submarines. These programs used less sophisticated digital design technologies or methods to document their ship design before the rise of 3D modeling capabilities.

By using 2D design information instead of 3D information, Navy shipbuilding programs face increased risk that 2D designs obscure issues—such as multiple design components occupying the same space. Such issues are more easily identifiable when visualizing a space using 3D modeling. Further, shipbuilders noted that 2D design is limited, compared to 3D design capabilities, in its ability to provide for simultaneous access of designs by multiple users, rapid assessment of many design options, and effective modeling of designs earlier in the design cycle to inform decision-making.

Additional challenges cited by Navy shipbuilders include the Navy's continued use of 2D design products for reviews and the timeliness of VFI receipt. For example, one Navy shipbuilder noted that programs continue to rely more on 2D drawings to support design review despite the availability of 3D design products to support these reviews. Further, Navy shipbuilders told GAO that their ability to capitalize on the opportunities that design tools offer to expedite ship design maturity is predicated on the timely receipt of reliable VFI. Without it, the 3D modeling is held back by the risk of design changes from unstable information on ship equipment.

Beyond 3D modeling, the previously discussed May 2023 design shortfalls acknowledged by Navy leadership also included capability gaps with in-house design tools. As with the design workforce issues, the Navy expects its ongoing work related to a strategic plan for design capabilities to set a course to replenish its in-house design tool set. In addition, Navy shipbuilders told GAO they are adopting other modern design tools to varying degrees, noting limited use of digital twinning and early-stage employment of virtual or augmented reality to support ship design and construction. For example, one Navy shipbuilder told GAO that its increased capability in 3D modeling and recently introduced virtual reality allow for design testing using the ship model as a digital prototype. The company is also creating a digital twin of its shipyard to support production efficiencies. However, Navy shipbuilders' use of these tools remains more limited overall than what GAO found for commercial builders.

Navy shipbuilders told GAO that use of modern design tools can advance design maturity and inform design decision-making. Specifically, the tools can help validate the physical integration of the ship, which ensures that multiple systems or features are designed into the ship without creating design conflicts, such as two systems occupying the same space. In the absence of a Navy requirement to use these design tools, Navy shipbuilders indicated that one challenge to expanding their design tools is building the business case to support the investment required to acquire and implement them. Still, without assessing potential opportunities to expand the use of modern design tools—within the Navy and across its shipbuilders—the Navy will not have a solid understanding of the types of investments required to ensure modern design tools are consistently used across its shipbuilding programs. The Navy could miss opportunities to gain efficiencies that support shorter, more predictable cycle times for ship design.

Conclusions

The demands pushing the Navy to increase the pace of design and construction for new ships will likely go unfulfilled without reforms to its ship design approach that provide greater flexibility and enhanced timeliness. Since GAO's initial shipbuilding leading practices work in 2009, the Navy and its shipbuilders have taken steps to improve design practices, which include implementing many of GAO's recommendations directed at increasing design maturity before the start of construction. GAO's analysis of the practices used by commercial ship buyers and builders indicates that the Navy has additional opportunities to embrace leading ship design practices to support timely, predictable outcomes for its shipbuilding programs. These opportunities involve:

- Improving consistency and communication of ship design maturity measures that support decisions to begin construction.
- Ensuring validated requirements continue to reflect operational needs before making decisions to proceed with the construction of each ship.

- Increasing the level of design maturity achieved before making decisions on detail design and construction contract awards and cost and schedule expectations for shipbuilding programs.
- Ensuring consistent, direct user involvement throughout the ship design process to inform decision-making.
- Improving processes and resources to streamline decision-making by ensuring that the amount of stakeholder involvement matches the significance of decisions, and decision-makers have the support needed to efficiently make them.
- Improving the Navy's digital ship design resources to increase its inventory of existing design knowledge and its efficiency in maturing and validating new ship designs.

Without additional action to better align its ship design efforts with leading practices, the Navy will be significantly challenged in its ability to rapidly confront evolving maritime threats with new ships that have the capabilities to combat those threats. These challenges affect current programs' timelines for delivery of new ships. They also create headwinds from the outset for the Navy's major future programs planned for the coming decades to deliver the next generation of destroyers, attack submarines, and amphibious assault ships, among other new additions to its fleet. In addition, without increased use of leading ship design practices, Navy shipbuilding programs will likely continue to regularly take a decade or more to move from concept to ship delivery. This increases the risk that capabilities approved in the earlier stages of a program lose their relevance and puts the Navy perpetually on the defensive because it cannot deliver timely, new capability to match the pace of new threats.

Excerpt of U.S. Government Accountability Office (GAO) report GAO-24-105503 (May 2, 2024)

USMC Landing Craft Case Study

Dr. Robert F. Mortlock, COL, USA (Ret.)—is a professor of the practice and associate chair for acquisition sciences for the Department of Defense Management at the Naval Postgraduate School in Monterey, CA. He holds a PhD in chemical engineering from the University of California, Berkeley, an MBA from Webster University, an MS in national resource strategy from ICAF, and a BS in chemical engineering from Lehigh University. [rfmortlo@nps.edu]

Samuel T. Irvine, Major, USMC—is currently serving as operational contract support advisor for II Marine Expeditionary Force stationed in Camp Lejeune, NC. He holds an MBA from the Naval Postgraduate School and a BS in business administration from Drake University. [samuel.t.irvine.mil@usmc.mil]

Abstract

This case study is written to produce an active learning environment to increase the capability of acquisition/program management professionals and senior leaders regarding program planning, decision-making, and affordability. The U.S. Marine Corps (USMC) Landing Ship Medium (LSM) program is a USMC priority acquisition program originating from USMC Force Design 2030 organizational changes and managed within the Naval Sea Systems Command (NAVSEA) Program Executive Office Ships acquisition portfolio. The USMC LSM procurement objective is 35 ships, and the initial cost estimate for each ship was between $100 million and $150 million. The U.S. Navy (USN) has expressed concern over the LSM's limited survivability requirements. To meet the USN's more stringent survivability requirements, the LSM cost would increase to more than $350 million per ship and threaten the program's affordability within the USN's shipbuilding budget. Moving forward, the USMC faces challenges addressing the best option to solve the medium-size amphibious ship capability gap as well as determining the optimal acquisition pathway and contracting strategy. The program must balance the following in determining the path forward: performance and security requirements; affordability/cost constraints; schedule need dates; program, technical, and manufacturing risks; and industrial base challenges.

Keywords: ship building, affordability, decision-making, critical thinking, project management

Introduction

The Landing Ship Medium (LSM) program is a U.S. Marine Corps (USMC) priority acquisition program with an acquisition objective of 35 ships originating from USMC Commandant General David Berger's (2023a) Force Design 2030 organizational and equipment changes. U.S. Navy (USN) leadership has expressed concern over initial LSM survivability requirements and potential increased cost estimates to over $350 million per ship to meet additional survivability requirements (O'Rourke, 2023b). Differences in ship capability requirements and Naval Sea Systems Command (NAVSEA) concerns with a limited shipbuilding budget have delayed procurement contract award to fiscal year (FY) 2025 (O'Rourke, 2023b).

The USMC's Force Design 2030 requirement identified a need for 35 additional amphibious connectors larger than a Landing Craft Air Cushion (LCAC) or Landing Craft Utility (LCU) and smaller than a Landing Platform Dock (LPD; Berger, 2023a). Figures 1, 2, and 3 depict an LCAC, LCU, and LPD, respectively, to show the vessel size differences and capability limitations between ship-to-shore LCU/LCAC connectors and larger amphibious LPD warships.

Figure 1. An LCAC Moving USMC Vehicles to Shore (Eckstein, 2023).

Figure 2. An LCU Transporting Marines to Shore (USN, 2019).

Figure 3. USS *New Orleans* (LPD 18) with an LCAC in the Background (Eckstein, 2022).

These additional vessels are needed because Marine Littoral Regiments (MLRs) operating as stand-in forces in the Pacific lack tactical mobility and maneuverability to move company-sized forces and equipment between Pacific islands (Berger, 2023a). The LSM will provide the USMC with a low-signature ship attached to the MLR that can deliver a Marine company to shore; it will also be larger and more effective than current smaller LCU and LCAC connectors assigned to Marine Expeditionary Units (MEUs; Oakley et al., 2023, p. 171). The LSMs will augment larger amphibious vessels assigned to support MEUs in the Pacific theater, such as the LPD and Landing Helicopter Assault LHA; Berger, 2023b).

The initial cost estimate for each LSM was between $100 million and $150 million with an acquisition program and desired procurement contract award in FY2023 Quarter 1 (O'Rourke, 2023b). Currently, the program is behind the USMC's desired schedule, and the program could slip further due to issues solidifying the acquisition quantity and requirements (Oakley et al., 2023, p. 171). The USMC requirement outlines the need for additional medium amphibious connectors as a priority to meet increasing operational demand in the Indo-Pacific Command (INDOPACOM) and expects the first LSMs in the fleet by 2028 to meet Force Design 2030 implementation timelines (Feichart, 2023, p. 1). Given budgetary constraints, shipbuilding backlogs, limited industry participation, and other issues, the risk of further schedule slip is high (O'Rourke, 2023a).

Background

> I woke up this morning, checked what's the readiness rate. It's 32 [percent]. We can't live with a 32 percent readiness rate. And over the last decade it's below 50 percent.
> —38th Marine Corps Commandant General David Berger (Kenney, 2023, p. 1)

During World War II, the rapid production and availability of Landing Ship Tanks (LSTs) played a pivotal role in transporting troops, equipment, and supplies in the European and Pacific theaters. These vessels were designed to carry heavy cargo, up to 431 troops, and 510 tons of vehicles, and conduct amphibious beach landings. After the war, the Department of Defense (DoD) recognized the naval utility value of these ships and kept the LST in service until 2002. However, after the Newport-class LST was decommissioned that year, a logistical void surfaced that could not be filled by smaller or less capable connectors or medium-sized vessels. Despite evolving warfare dynamics and technologies, there is still a need for modern multi-functional LSTs that provide the naval services with the ability to conduct amphibious operations, humanitarian missions, and evacuation operations. The LST's historical significance is amplified by its World War II production efficiency, which was due in part to its modular assembly and design. This unique feature enabled large-scale production at 18 shipyards that produced over 1,000 LSTs in only 3 years. Surprisingly, many of these vessels originated from inland shipyards located in Illinois, Indiana, and Pennsylvania because of the ship's smaller size, modular design, and ability to navigate inland rivers to reach the oceans (Phillips, 2023). The LST's historical significance is amplified by its World War II production efficiency, which was due in part to its modular assembly and design. This unique feature enabled large-scale production at 18 shipyards that produced over 1,000 LSTs in only 3 years.

Figure 4. World War II LSTs Onloading Equipment and Supplies in England in Preparation for Operation Overload (Ussery, 2008).

Amphibious Capability Gap and Requirements

USN amphibious L-class ships (e.g., LPD, LHA) are crewed by Navy sailors and used to transport Marines, weapons, equipment, and limited supplies to expeditionary operations in littoral areas (O'Rourke, 2023b, p. 5). Figure 5 depicts an LHA, which is the largest type of USN L-class ship and, unlike other amphibious vessels, does not possess a well deck.

Figure 5. USS *America* (LHA 6) Conducting a Replenishment-at-Sea (Defense Visual Information Distribution Service, 2020).

The FY2023 National Defense Authorization Act (NDAA) directs that the minimum necessary amphibious fleet shall consist of 10 amphibious assault ships (LHA/Landing Helicopter Dock [LHD]) and 21 LPDs (Berger, 2023b). The L-class ships are organized into Amphibious Readiness Groups and combine with MEUs to provide overseas naval deterrence and response capability to support combatant commanders. Kenney (2023) reported that the deployable USN amphibious fleet averaged 46% readiness over the past decade. In 2023, deployable L-class amphibious ship readiness reached its lowest recorded point, at 32%. The USN attributes these operational availability issues to a ship maintenance backlog, which is a fleet-wide problem. The lack of availability impacted the USMC's ability to respond quickly with an MEU in 2022 to the Russian invasion of Ukraine and provide humanitarian aid to Turkey and Syria earthquake victims (Kenney, 2023). Currently, the USN amphibious fleet is unable to meet the National Defense Strategy requirement to consistently provide 31 amphibious ships to ensure MEU forces for combatant commanders (Berger, 2023b, pp. 16–17).

Force Design 2030 introduced a new force structure by transforming two infantry and one artillery regiments into three MLRs possessing balanced infantry, fire support, low-altitude air defense, and logistics battalions organic to the new formation. These forces are designed to operate dispersed within the first island chain of the Pacific Islands, including Senkaku (Japan), Ryukyu (Okinawa), and the Philippines, providing land-based sea lane control and sea denial capabilities. The MLR structure promotes decentralized company-level operations within the area of operations to reduce detectability. USMC wargames identified that the MLR requires organic sea mobility to enable small company-size movements between the numerous Pacific first island chain nations. Sea mobility provides the MLR with the ability to blend into dense commercial shipping routes using comparably sized vessels, thus limiting detectability among similar commercial vessels, which increases the MLR's survivability during conflict. The LSM is envisioned to fulfill tactical sea mobility in politically and militarily contested Pacific environments while complementing L-class amphibious ships by offering a new remote island connector capability (Berger, 2023b, p. 13). This platform offers a lower risk of escalation when maneuvering in gray zone areas to facilitate security cooperation, humanitarian assistance, and

MLR logistics support mobility (Berger, 2023b, p. 13). According to General Berger (2023c),

> After extensive research and wargaming, we calculated a need for nine LSMs to support a single regimental sized unit. The DON's Amphibious Force Requirements Study over the last two years validated this number, articulating a requirement of no fewer than 18 LSMs to support littoral maneuver. Given that current force structure plans call for three MLRs, we require 35 LSMs to account for operational availability and mobility for those units. We anticipate an initial request for 18 of the 35 LSMs we seek will be a step toward enabling us to more effectively counter adversaries' strategies, support and reinforce alliances and partnerships, and do so at a relatively low cost. (pp. 13–14)

In 2020, the LSM ship requirements were simple and inexpensive, and could be based on commercial ship design (O'Rourke, 2023b). Figure 6 depicts an LSM concept design based on the following vessel requirements and specifications outlined in the System for Award Management's (2020a, 2020b) LSM Circular of Requirements and Industry Day brief, which were consolidated by O'Rourke (2023b):

- length of 200–400 feet
- maximum draft of 12 feet
- displacement of up to 4,000 tons
- ship's crew of no more than 40 USN sailors
- ability to embark at least 75 Marines
- 4,000–8,000 square feet of cargo area for the Marines' weapons, equipment, and supplies
- stern or bow landing ramp for moving the Marines and their weapons, equipment, and supplies from the ship to shore (and vice versa) across a beach
- modest suite of C4I equipment
- 30mm gun system and .50 caliber machine guns for self-defense
- transit speed of at least 14 knots, and preferably 15 knots
- minimum unrefueled transit range of 3,500 nautical miles
- tier 2+ level of survivability (i.e., ruggedness for withstanding battle damage), a level broadly comparable to that of a smaller USN surface combatant (e.g., a corvette or frigate), that would permit the ship to absorb a hit from an enemy weapon and keep the crew safe until they and their equipment and supplies can be transferred to another LSM
- ability to operate within fleet groups or deploy independently
- 10-year minimum and 20-year expected service life

Figure 6. An LSM Concept Design (Grady, 2023).

Key to the LSM design and survivability is mobility to hide within commercial shipping lanes and surrounding Pacific Islands. The capability to move forces, equipment, and supplies between small commercial ports and remote island beaches is crucial to fill the MLR amphibious vessel gap. The LSM is a fraction of the size of L-class ships, and initial requirements described a desire for it to resemble commercial shipping vessels navigating the same maritime arena. Hubbard (2023) described the LSM as a "transport vessel in the tradition of vessels like the Landing Ship, Tank (LST) of World War II [WWII] vintage. LSTs were designed to bring materiel from American factories at home across oceans and deposit this equipment on a foreign and often hostile shore" (p. 68). The LSM, like the LST, was initially envisioned as an inexpensive vessel able to deploy dispersed surface forces across the INDOPACOM theater. Like the LST, the LSM provides intra-theater tactical lift able to fulfill multiple transportation requirements in conjunction with larger L-class ships. The LSM is required to be less detectable than L-class amphibious ships and able to operate in a channel distribution system to move people and things between vessel platforms to dispersed remote island end points (Hubbard, 2023).

The LSM capability forecasts a vessel able to support a "dispersed, agile, constantly relocating force" (Apte et al., 2021, p. 305) operating in accordance with the Expeditionary Advanced Base Operations concept. As a medium-sized ship, the LSM is required to conduct amphibious landings on beaches to offload Marines, equipment, and supplies while also possessing greater carrying capacity, range, and survivability in comparison to LCUs and LCACs. Apte et al. (2021) described the LSM requirement as a "risk-worthy vessel (defensible enough that risks are not excessive or cheap enough that we can afford to lose it) with priority for personnel survivability" (p. 306), which is a different employment concept from L-class ships.

The Deputy Commandant of Marine Corps Combat Development and Integration (CD&I), Lieutenant General Karsten Heckl, described the LSM as a shore-to-shore connector not requiring a pier or another ship (Easley, 2022). CD&I is the USMC's requirements generation, experimentation, and wargaming command responsible for defining what the USMC needs from the LSM to be effective in the INDOPACOM region. LtGen Heckl described the LSM as a priority for modernization efforts despite budget constraints delaying production and USN leadership concerns about survivability in a conflict. In 2022, CD&I leased a commercial stern vessel to deploy with 3rd MLR for experimentation in the INDOPACOM area of operations to reaffirm minimum viable product LSM requirements and demonstrate urgency of need (Easley, 2022).

The U.S. Army possesses a large fleet of aging watercraft capable of transporting soldiers and equipment short distances and conducting beach landings. Under the U.S. Army's Maneuver Support Vessel initiative, two new watercraft variants are being developed for operations in the Indo-Pacific region. The Army Program Executive Office for Combat Support and Combat Service Support (PEO CS&CSS) launched the Maneuver Support Vessel-Light (MSV-L) prototype at Vigor LLC's Vancouver, WA, facility, which marked the introduction of a new and improved class of Army watercraft (Higgins, 2022). Vigor was awarded a 10-year contract in 2017 to produce up to 36 of these MSV-L craft that are intended to replace the

Vietnam-era Landing Craft Mechanized-8, which is like the USN LCU vessel. The MSV-L is 117 feet long, is crewed by eight soldiers, has a top speed of 21 knots fully loaded with soldiers and equipment, and has a maximum range of 360 nautical miles (Higgins, 2022). Further, the MSV-L is designed to transport either an M1 Abrams tank, two Stryker combat vehicles, or four Joint Light Tactical Vehicles (Luckenbaugh, 2023). After initial testing, the Army determined the MSV-L baseline requirements necessitated modification to address design changes and cost increases, with projections for initial operational capability in 2028 (Roque, 2023). Notably, the MSV-L design lacks the defensive systems and survivability features the USN desires to incorporate in the LSM design, which increase the LSM's cost per ship (The Maritime Executive, 2023). Figure 7 shows the MSV-L concept design and resemblance to USN LCUs in service.

Figure 7. U.S. Army MSV-L Concept Design (Vigor, n.d.).

Brigadier General Samuel Peterson, U.S. Army PEO CS&CSS, highlighted collaboration with the USN and USMC in defining the larger Maneuver Support Vessel (Heavy; MSV-H) requirements (Roque, 2023). The MSV-H is planned to be up to 400 feet in length, have a top speed of 18 knots, carry as many as 175 soldiers and their equipment, possess a crew of approximately 30, and be capable of beach landings (Luckenbaugh, 2023). The Army plans to select multiple shipyards to develop virtual prototypes with a planned low-rate initial production (LRIP) decision in 2028 and the first delivery in 2030. The MSV-H design specifications resemble the USN LSM vessel requirements; however, the MSV-H provides slightly greater speed and carrying capacity. The similarities between the two programs in meeting INDOPACOM warfighter requirements create the possibility for a joint solution that would provide reduced life-cycle operations and sustainment operation and costs as well as Army/USN/USMC collaboration opportunities for budgetary resources allocation.

Program Development

The LSM program, previously named the Light Amphibious Warship (LAW) program, received a Material Development Decision and entered the Materiel Solution Analysis phase of the major capability acquisition (MCA) process with a procurement goal of 18–35 LSMs and the awarding of initial production contracts in FY2025 (O'Rourke, 2023b). The initial capabilities document outlined the validated threshold requirements for the ships (System for Award Management, 2020a, 2020b), which supported the completion of a draft Analysis of Alternatives (AoA; Oakley et al., 2023, p. 171). As of 2023, the DoD had not approved the AoA (Oakley et al., 2023, p. 171). According to DoD Instruction 5000.85, without AoA approval, the acquisition program is unable to proceed to the MCA Milestone A decision to develop the system further in the Technological Maturation and Risk Reduction (TMRR) phase (Office of the Under Secretary of Defense for Acquisition and Sustainment, 2020). Figure 8 displays the LSM program schedule (as of 2023) from concept to system development and through production.

Figure 8. LSM Acquisition Timeline as of June 2023 (Oakley et al., 2023, p. 171).

The LSM AoA studies the necessity to proceed in developing and producing a new amphibious ship design over repurposing existing USN, Maritime Sealift Command, or U.S. Army watercraft to meet the sea transportation requirement. According to O'Rourke (2023b), the DoD has not yet approved the AoA because the "key requirements of the new vessels are very similar to the capabilities of vessels operated by U.S. Army Transportation Command" (p. 22). Further, O'Rourke (2023b) recommended that "the Navy and Marine Corps should delay any new construction and immediately acquire some of these existing vessels to drive experimentation and better inform their requirements for the LAW program" (p. 22). O'Rourke's (2023b) recommendation to delay production and further explore opportunities to leverage existing Army Transportation Command watercraft systems could benefit the USN and USMC to reduce their operational capability gap risk.

Though the AoA study plan is still pending approval, the LSM program office awarded concept design contracts to five production-capable shipbuilders with the option to award a follow-on Preliminary Design Review (PDR) contract (Shelbourne, 2021). These five finalist shipbuilders, tasked with creating digital prototypes, could be viable manufacturers during the production phase even though they are not all traditional Navy amphibious shipbuilders (Quigley, 2022). These shipbuilders and engineering design firms included Fincantieri, Austal USA, VT Halter Marine, Bollinger, and TAI Engineers. In total, 11 industry teams worked with NAVSEA to understand the vessel requirements and competed for the design contract award (Eckstein, 2021). One of the 11 firms was SeaTransport; Figure 9 displays its LSM concept design.

Figure 9. SeaTransport's Proposed LSM Concept Design (Shelbourne, 2021).

The contract winners will use the requirements to produce ship designs, which will include engineering analyses and trade-off studies to assist in the TMRR phase (Royal Institution of Naval Architects, 2021). The winning concept will receive a follow-on preliminary design contract to refine technology maturation in preparation to enter the Engineering and Manufacturing Development (EMD) phase post–Milestone B. The five concept design awards amounted to less than $7.5 million (Shelbourne, 2021). Additionally, in the FY2024 budget, the

USN programmed $14.7 million for research and development to refine the five awarded design review contracts through prototyping.

Originally, the USN and USMC requirements and acquisition team projected enthusiasm and willingness to begin initial production as early as FY2022 (Eckstein, 2021). However, capability requirements differences delayed initial production. Shelbourne (2021) described LSM planning, programming, budgeting, and execution funding as an issue, for the "Navy only sought the research and development funding in the recent FY2022 request" (p. 1). The USMC's aggressive acquisition requirement timeline did not match the USN's desire to refine the concept studies and did not program procurement appropriation funding to meet the expected FY2022 initial production goal.

Rear Admiral John Gumbleton, deputy assistant secretary of the USN for budget, commented on the LSM development as part of the USN's FY2023 budget by stating, "The Marine Corps and the Department are getting the requirements tight on that ship before we choose to put it in our [shipbuilding appropriations account]. So, there is funding in R&D for LAW" (O'Rourke, 2023b, p. 17). While RAML Gumbleton argued that USN shipbuilding leadership preferred to reduce the risk through research and development funding, Major General Tracy King, former director of expeditionary warfare for the Office of the Chief of Naval Operations (OPNAV 95), proclaimed that the LSM acquisition schedule was "aiming at lead ship construction in FY '22, it's going to be late in FY '22, but I still consider that pretty fast" (Eckstein, 2021, p. 1). O'Rourke (2023b) outlined the developing program schedule risk, stating that "another issue for Congress concerns the date for procuring the first LAW. As noted earlier, previous USN plans envisioned starting procurement of LAWs in FY2023. Compared to this, the USN's FY2023 five-year shipbuilding plan in effect defers the start of LAW procurement two years, to FY2025" (pp. 16–17). O'Rourke (2023b) highlighted the LSM program schedule delays and increased per-ship procurement costs, opining the need for further cost–benefit analysis and enhanced congressional oversight.

Currently, the USN is planning for LRIP beginning with procurement contract award in 2025, with the first LSM estimated to cost $187.9 million (O'Rourke, 2023b). Using a single shipbuilder, the follow-on manufacturing contract award for the second LSM would occur in FY2026 and cost $149.2 million, while the third and fourth ships would be procured in FY2027 and cost a combined $297 million, or $148.5 million per ship. The LRIP fifth and sixth LSM procurement contract awards are scheduled for FY2028, costing an estimated combined total of $296.2 million, or around $148.1 million per ship. Included in the cost estimate for the lead ship are the detailed design and nonrecurring engineering costs, which are traditionally how the USN generates ship cost estimates for the first procurement (O'Rourke, 2023b).

Compared to larger LPD and LHA amphibious ships, the LSM's reduced size enables a greater number of shipyards and shipbuilders to manufacture it. O'Rourke (2023b) stated, "The Navy's baseline preference is to have a single shipyard build all the ships, but the Navy is open to having them built in multiple yards to the same design if doing so could permit the program to be implemented more quickly and/or less expensively" (p. 2). The LSM concept is a modified commercially produced stern landing vessel design that can be built at many U.S. shipyards, creating greater production capacity beyond the limited larger L-class shipyard producers (Royal Institution of Naval Architects, 2021). With the USN's proposed LRIP acquisition strategy, the time between procurement contract award and delivery is estimated at 3.5 years for the first ship, so a FY2025 contract award will deliver the lead ship to the fleet in FY2028. Former Commandant of the Marine Corps Gen Berger (2023b) described the current problem set in congressional testimony by stating,

> We have adapted to this challenge and are developing

> bridging solutions to experiment with LCU-1700s and leased Expeditionary Fast Transports (T-EPF) and Stern Landing Vessels. While these platforms will inform the eventual employment of the LSM, they will fall short of desired capabilities if called upon in an operational setting. Our modernized expeditionary forces need a comparably modern mobility platform to bring the full weight of their capability to bear on competitors or adversaries, particularly in littoral regions. (p. 14)

Optimistically, the first LSM will complete production in 2028, and the fleet will not be fully operational and capable of effectively supporting MLRs until at least a decade later. In the interim, pressure to achieve the USMC's high priority need for additional amphibious ships can only be fulfilled by commercial vessel leasing options and existing alternative legacy Army Transportation Command watercraft. These will be the only solutions available in the near term to meet an increasing need for light sea transportation in INDOPACOM.

Program Challenges

With the LSM, the USN aims to provide a modern adaptation of the World War II–era LST for transporting Marines and equipment throughout INDOPACOM. In a major war, LSMs would be susceptible and slow targets, just like World War II LSTs were, though the LST's versatility outweighed its vulnerability (Hooper, 2023). Additionally, the modest 40-person LSM crews led by junior officers conflict with current naval personnel shortfalls. A 35-LSM fleet would require 280 junior naval officers, further challenging recruitment, and would deviate those officers from traditional surface warfare officer career pathways (Hooper, 2023). Contrary to common sense, commanding an LSM as a USN lieutenant (O-3) could put junior officers at a disadvantage in terms of remaining competitive for promotion due to their peer group gaining greater warship systems experience while serving aboard actual warships (e.g., destroyers; Hooper, 2023).

O'Rourke (2023b) described that the LSM program experienced significant delays, with the detail design and construction contract award pushed from FY2023 to FY2025. O'Rourke (2023b) opined that the 19-month slippage stems from ongoing engagement with industry to refine requirements and delays approving the program's AoA. O'Rourke (2023b) detailed that the LSM program continues working toward a contract award in 2025 and aims to shorten development time by modifying an existing commercial ship design rather than creating a new design. The LSM program seeks to streamline the schedule by eliminating certain oversight reviews, which risks senior leaders lacking information necessary for making sound decisions (O'Rourke, 2023b). The USN has engaged industry on LSM concepts since 2020 through multiple rounds of studies with numerous participating designers and shipbuilders. The USN aims to rapidly iterate designs to meet evolving requirements and provide feedback on requirement impacts.

Key LSM program elements, including survivability requirements and procurement quantity, remain undefined. The USMC proposed acquiring 35 LSMs, but the USN supports only 18. Without a clearly defined acquisition objective and concurrence on commercial ship design modification requirements, the LSM vessel procurement cost ranges from $150 million per ship to produce the minimum viable product the Marines desire to $350 million per ship to add the Navy's desired survivability requirements comparable to L-class amphibious ship survivability and systems technology (O'Rourke, 2023b). At its core, the disagreement over LSM capability systems and survivability reflects the USN and USMC's differing attitudes toward risk tolerance. The USN is extremely reluctant for its vessels to suffer catastrophic battle damage, whereas the

USMC acknowledges that losses of Marines and equipment, while regrettable, are an unavoidable hazard during combat operations (Larson, 2022).

Critics of the LSM program stress that the USMC values ship procurement and delivery speed by requesting appropriation funding before the final requirement is determined, which is reminiscent of the flawed LCS program (Baird et al., 2022). Deviating from major capability acquisition processes and milestones increases program risk and can lead to requirements creep. LCS construction began before prototype testing did, which led to cost overruns and unmet operational needs after 20 years of design and program management failures, resulting in terminating future production and retiring ships early (Baird et al., 2022). Currently, the FY2024 shipbuilding budget supports the first LSM construction contract being awarded in 2025.

Also, the USN prefers a single shipyard that manufactures all LSMs but would allow a multi-yard approach if it accelerated schedule or reduced costs (O'Rourke, 2023b). Key design considerations reflect these trade-offs, including a maximum 12-foot draft, which facilitates transit in shallow waters and beach landings, and ample cargo space, as open deck storage differs from most current amphibious ships. The modest speed of about 15 knots, compared to 22 knots for larger amphibious ships, allows for a less expensive and more fuel-efficient propulsion system (O'Rourke, 2023b). The 20-year service life is less than the 30–45 years that is typical for bigger amphibious ships but enables a lower cost for this smaller ship class. The services are working to strike the right balance between affordability gained through simplified designs and survivability requirements aimed at enhancing fleet capabilities.

The LSM survivability is questionable due to its slow speed and limited maneuverability, which makes it susceptible to enemy detection when transiting contested seas and vulnerable to missile strikes (Jenkins, 2022). Further, any direct hit on the lightly defended ship would result in unrecoverable catastrophic damage. Adding enhanced survivability features increases the per-unit procurement cost and the operations and maintenance cost, resulting in the necessity to trade off other features or reduce the number of ships procured. It is inevitable that the final cost of building the new ship will be far higher than initial estimates, as more unforeseen expenses and requirements will emerge during the long construction process. Additionally, given the new naval ship class's record of cost overruns and delays, there is considerable uncertainty about when this capability will be delivered to the fleet (Jenkins, 2022).

In April 2023, the USN and USMC communicated that they were close to reaching agreement on the requirements and costs for the LSM program (O'Rourke, 2023b). BGen Marcus Annibale, the director of expeditionary warfare on the chief of naval operations staff, indicated there was progress in drafting the capability development document (O'Rourke, 2023b). The author further reported Vice Admiral Scott Conn, the deputy chief of naval operations for warfighting requirements and capabilities, recognized the importance of procuring these smaller ships. Additionally, LtGen Heckl, deputy commandant of CD&I, explained that he, VADM Conn, and BGen Annibale were able to work together to find common ground on survivability and vulnerability features to incorporate into the LSM design (O'Rourke, 2023b). LtGen Heckl also noted that the original concept emphasized low cost, larger quantities, and a commercial-style design (O'Rourke, 2023b). However, discussions between the USN and USMC led to the USN and the Office of the Secretary of Defense demanding greater capability and survivability requirements—and, therefore, greater costs—and now the program is returning to its initial size and cost (O'Rourke, 2023b). On May 17, 2023, the USN issued a request for information to shipbuilders about the LSM program and asked interested firms to provide responses on several production capacity and investment topics. According to O'Rourke (2023a), those questions included the following:

- Do you have the resources and production capacity available to be awarded 4 LSM ships per fiscal year?
- If so, how can your shipyard support production of 4 LSM hulls per year?
- If not, what is the maximum number of LSM ships that can begin production each year?
- If not, are there investment or shipyard improvements that can be done to enable increasing production capacity to 4 LSM hulls per year? (p. 5)

This request for information showed the USN's interest in manufacturing multiple LSMs per year, and, given the USN's previously stated acquisition strategy to produce 18 LSMs, this four-ships-per-year rate would complete production within 5 years of accelerated production. This is a key insight into the USN's goals and willingness to accept increased risk to achieve greater production speed for the warfighter.

In the Government Accountability Office's 2023 annual weapons system report, Oakley et al. (2023) described the current LSM (referred to as the LAW) program status by stating,

> Since our last review, the Navy delayed the detail design and construction contract award for LAW from fiscal year 2023 to fiscal year 2025. According to Navy officials, this change was due to ongoing efforts to engage with industry and refine program requirements, as well as delays in gaining approval of the program's analysis of alternatives (AOA)—a key document to help DOD and the Navy decide if a new ship class is needed. As of January 2023, the Office of the Secretary of Defense had yet to approve the AOA, which is at least a 19-month delay in the planned approval since our last review.
>
> Although an approved AOA has yet to confirm the need for LAW, the program continues to work toward a detail design and construction contract award and is looking for opportunities to shorten LAW's development time. For example, the program plans to modify an existing parent ship design, instead of creating a new one, and has been assessing potential designs with five companies since 2021. The program also plans to seek approval to streamline its schedule by eliminating certain early acquisition oversight reviews. We previously found that eliminating such reviews can increase the risk that senior acquisition and warfighting leaders lack information needed for sound investment decisions.
>
> Currently, several key program elements remain undefined. In particular, the Navy is still determining LAW's requirements. In alignment with leading principles for iterative development, the Navy is making changes to draft requirements based on industry feedback and ongoing AOA efforts. DOD has also yet to determine LAW's total procurement quantities. The Marine Corps suggested 35 ships, but the Navy proposed acquiring only 18. The Navy cannot estimate LAW's costs until it defines requirements and quantities. (p. 171)

Oakley et al. (2023) received the following summarized comments from the LSM/LAW Program Office:

> It stated that the Navy is following a deliberate requirements process to determine its needs for the LAW program. It noted that the Navy endorsed the AOA in March 2022 and is awaiting the sufficiency review by the Office of the Secretary of Defense. It added that it is incorporating the analysis results and feedback from the five industry preliminary designs into the upcoming Capabilities Development Document. (p. 171)

Path Forward

The LSM program faces acquisition options and decision points that include finalizing the vessel requirements and procurement quantity and maturing the commercially modified design (Oakley et al., 2023, p. 171). The program must also determine whether the design and construction contract will be awarded to a sole shipbuilder or multiple concurrent shipbuilders. The shipbuilding industrial base's capability and capacity to produce four LSMs per year to meet the USMC's operational need dates remain key constraints. Finally, the program must determine the best acquisition path forward to manage cost, schedule, performance, and manufacturing risk. In an attempt to shorten development timelines and streamline oversight reviews, the program plans to modify an existing commercial ship design rather than develop a completely new design.

In summary, the LSM program faces decisions on balancing performance capability, schedule, costs, and manufacturing risks as it proceeds toward a production contract award. Careful oversight is necessary to avoid past shipbuilding program pitfalls. The acquisition team's challenge is to tailor, combine, and transition between acquisition pathways to deliver the LSM to the warfighter before 2030 while also reducing per-unit costs through capability trade-offs to meet shipbuilding budget constraints. The team must maximize value for the warfighter by creating a realistic program baseline despite cost overruns, budget limits, and a need for faster shipbuilding.

Recommendations for the path forward must address the following questions and key decisions:

- What is the best option to solve the warfighter's medium-size amphibious ship capability gap?
- Assuming the doctrine, organization, training, materiel, leadership and education, personnel, and facilities assessment justifies a materiel solution and the LSM AoA is approved, what is the best acquisition pathway to follow?
- What is the best LSM contract award strategy?

Options to address the warfighting capability gap include using current amphibious ships, pursuing a joint acquisition program with the Army's MSV-H program, acquiring commercially available vessels (commercial off-the-shelf), or pursuing an LSM development program. If the USMC decides that the LSM program is best path forward, then the appropriate acquisition approach can leverage multiple Adaptive Acquisition pathways based on the urgency of need, available resources, and technical/manufacturing readiness levels. Acquisition approaches to consider include continuing in the major capability acquisition (MCA) pathway toward an MS B, using the MCA pathway but going directly to MS C, using the middle tier acquisition (MTA) pathway with both rapid prototyping and rapid fielding, using the MTA pathway with rapid prototyping followed by entry to MCA at MS B, and using the MTA pathway with rapid prototyping followed by entry to MCA at MS C. Finally, the LSM contracting strategy to engage with shipbuilders can include contracts with a single domestic shipbuilder, multiple domestic shipbuilders, or multiple domestic shipbuilders and international shipbuilders. Decision criteria used to compare these options could include performance (meeting more USMC requirements is better), cost (lowering total life-cycle costs is better), schedule (meeting the USMC operational need dates is better), technical and manufacturing risk (leveraging high TRLs and MRLs is better), defense industrial base considerations (supporting the capacity and capability of shipbuilding industrial base is better), and security considerations (lowering the risk with use of international shipyards is better).

References

Apte, A., Apte, U., & Doerr, K. (2021). *Framework for augmenting current fleet with commercially available assets for logistics support in contested environment* (NPS-LM-20-155). Naval Postgraduate School.

Baird, S., Schuster, C., & Wells, T. (2022). The light amphibious warship is a bad sequel to the littoral combat ship. *The National Interest*. https://nationalinterest.org/feature/light-amphibious-warship-bad-sequel-littoral-combat-ship-205887

Berger, D. H. (2023a). *Force Design 2030: Annual update June 2023*. U.S. Marine Corps.

Berger, D. H. (2023b). *Statement of General David H. Berger, commandant of the Marine Corps, on the posture of the United States Marine Corps before the Senate Appropriations Committee*. Senate Committee on Armed Services. https://www.armed-services.senate.gov/imo/media/doc/230414_CMC%20Written%20Testimony%20(SASC).pdf

Berger, D. H. (2023c). *Statement of General David H. Berger, commandant of the Marine Corps, on the posture of the United States Marine Corps before the Senate Appropriations Committee*. U.S. Marine Corps. https://www.cmc.marines.mil/Speeches-and-Transcripts/Transcripts/Article/3360019/statement-of-general-david-h-berger-commandand-of-the-marine-corps-on-the-postu/

Birdon. (n.d.). *Our story*. https://birdon.us/our-story/

Defense Visual Information Distribution Service. (2020). *USS* America *(LHA 6) and USS* New Orleans *(LPD 18) conduct a replenishment-at-sea with USNS* Charles *crew (T-AKE 10) [image 3 of 7].* https://www.dvidshub.net/image/6339882/uss-america-lha-6-and-uss-new-orleans-lpd-18-conduct-replenishment-sea-with-usns-charles-drew-t

Easley, M. (2022). Improved sensors, loitering munitions on Marine Corps wishlist.*National Defense*, *107*(824), 28–29.

Eckstein, M. (2021, January 11). Marines, Navy moving quickly on light amphib, anti-ship missiles to create more warfighting options. *USNI News*. https://news.usni.org/2021/01/11/marines-navy-moving-quickly-on-light-amphib-anti-ship-missiles-to-create-more-warfighting-options#:~:text=Maj.,recent%20long%2Drange%20shipbuilding%20plan .

Eckstein, M. (2022, March 28). U.S. Navy seeks to end San Antonio-class ship production, reducing fleet by 8 amphibious hulls*. Defense News.* https://www.defensenews.com/naval/2022/03/28/us-navy-seeks-to-end-san-antonio-class-ship-production-reducing-fleet-by-8-amphibious-hulls/

Eckstein, M. (2023, October 4). RAND calls for reexamination of surface connectors' survivability. *Defense News.* https://www.defensenews.com/naval/2023/10/04/rand-calls-for-reexamination-of-surface-connectors-survivability/

Feichart, A. (2023). *U.S. Marine Corps Force Design 2030 initiative: Background and issues for Congress* (CRS Report No. R47614). Congressional Research Service.

Grady, J. (2023). SECNAV Del Toro "excited" about new landing ship mediums. *USNI News*. https://news.usni.org/2023/02/22/secnav-del-toro-excited-about-new-landing-ship-mediums

Higgins, R. (2022). *New vessel class enters Army watercraft fleet with prototype launch*. U.S. Army. https://www.army.mil/article/260993/new_vessel_class_enters_army_watercraft_fleet_with_prototype_launch

Hooper, C. (2023, February 6). Lay down the LAW: Stop fighting and build the light amphibious warship. *Forbes*. https://www.forbes.com/sites/craighooper/2023/02/06/lay-down-the-law-stop-fighting-and-build-the-light-amphibious-warship/

Hubbard, J. (2023). The purpose of the light amphibious warship. *Marine Corps Gazette*, *107*(4), 68–69.

Jenkins, H. (2022). The ugly truth: Can the light amphibious warship survive war with China? *The Hill*. https://thehill.com/opinion/national-security/3708190-the-ugly-truth-can-the-light-amphibious-warship-survive-war-with-china/

Kenney, C. M. (2023, March 11). Fewer than 1/3 of Navy's amphibious ships are ready to deploy. *Defense One*. https://www.defenseone.com/threats/2023/03/fewer-13-navys-amphibious-ships-are-ready-deploy/383874/

Larson, C. (2022). Kill it or build it? Navy and Marines don't agree on the light amphibious warship. *The National Interest*. https://nationalinterest.org/blog/buzz/kill-it-or-build-it-navy-and-marines-don%E2%80%99t-agree-light-amphibious-warship-205207

Luckenbaugh, J. (2023). Army finalizing requirements for "heavy" watercraft. *National Defense*. https://www.nationaldefensemagazine.org/articles/2023/3/27/army-finalizing-requirements-for-heavy-watercraft

The Maritime Executive. (2023). *U.S. Army's next heavy landing ship may outclass Marine Corps' LSM.* https://maritime-executive.com/article/u-s-army-s-next-heavy-landing-craft-may-outclass-marine-corps-lsm#:~:text=To%20complement%20its%20existing%20acquisition,of%20the%20%20Marine%20Corps%E2%80%99%20LSM

Oakley, S., et al. (2023). *Weapon systems annual assessment: Programs are not consistently implementing practices that can help accelerate acquisitions* (GAO-23-106059). Government Accountability Office.

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020). *Major capability acquisition* (DoD Instruction 5000.85). Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500085p.pdf

O'Rourke, R. (2023a). *Navy force structure and shipbuilding plans: Background and issues for Congress* (CRS Report No. RL32665). Congressional Research Service.

O'Rourke, R. (2023b). *Navy medium landing ship (LSM) (previously light amphibious warship [LAW]) program: Background and issues for Congress* (CRS Report No. R46374). Congressional Research Service.

Phillips, E. (2023). *LST redux: Adapting to the future of maritime warfare by understanding the past*. Marine Corps University Press. https://www.usmcu.edu/Portals/218/EXP_LST%20Redux_Phillips_PDF.pdf

Quigley, A. (2022). Five shipbuilders emerge as leading light amphibious warship contenders. *Inside Defense*. https://insidedefense.com/daily-news/five-shipbuilders-emerge-leading-light-amphibious-warship-contenders

Roque, A. (2023). *With Indo-Pacific in mind, Army eyeing heavy support ship prototyping competition*. Breaking Defense. https://breakingdefense.com/2023/03/with-indo-pacific-in-mind-army-eyeing-heavy-support-ship-prototyping-competition/

Royal Institution of Naval Architects. (2021). U.S. progresses to next phase of light amphibious warship project. *Warship Technology*, 25–26.

Shelbourne, M. (2021). Navy awards 5 companies light amphibious warship “concept design” contracts. USNI News. https://news.usni.org/2021/06/17/navy-awards-5-companies-light-amphibious-warship-concept-design-contracts

System for Award Management. (2020a). *RFI: US Navy light amphibious warship (LAW) (Notice ID N0002420SN018): Industry day brief.* https://sam.gov/opp/90a9ece86ade48089e9f6d57d2969d23/view

System for Award Management. (2020b). *RFI: DRAFT US Navy light amphibious warship preliminary design/contract design statement of work (Notice ID N0002421SN01): Circular of requirements.* https://sam.gov/opp/3d313f4c926746028630a5bd70204b3b/view

U.S. Navy. (2019). *Landing craft, mechanized and utility – LCM/LCU*. https://www.navy.mil/Resources/Fact-Files/Display-FactFiles/Article/2171588/landing-craft-mechanized-and-utility-lcmlcu/

Ussery, P. (2008). *LST in WWII*. U.S. Army. https://www.army.mil/article/10067/lst_in_wwii

Vigor. (n.d.). *Maneuver support vessel (light)*. https://vigor.net/projects/msvl

PANEL 10. EMPOWERING ACQUISITION THROUGH EDUCATION AND INNOVATION

Wednesday, May 7, 2025	
1115 – 1230 PT 1315 – 1430 CT 1415 – 1530 ET	**Chair: Frank Kelley, BGen, USMC (Ret.), Vice President of Defense Acquisition University (DAU)** ***TradeSpace: Designing A Wargame for USSF Acquisition Program Managers*** Lt Col Alan Lin, Chief, Strategic Development, Space Systems Command Chief Learning Officer ***DAU Innovate to Win: DOD Upskilling Innovation Readiness*** Jonathan Barkand, Learning Systems Program Manager, Defense Acquisition University (DAU) ***Tailored Interventions to Foster Acquisition Innovation: Piloting the Innovation Alliance Program*** Amanda M. Girth, Associate Professor, Ohio State University

Frank Kelley, BGen, USMC (Ret.)—is the Vice President of Defense Acquisition University (DAU). In this position, he is responsible for aligning DAU strategic plans to the goals of both the Secretary of Defense and Under Secretary of Defense for Acquisition and Sustainment, while continuing to build the outstanding reputation of DAU as the Department of Defense primary learning institution for acquisition. He oversees the development and expansion of acquisition curriculum and learning opportunities and the delivery of those learning assets throughout the five DAU regional campuses, the Defense Systems Management College, and the College of Contract Management.

Prior to this assignment, Mr. Kelley served as the Deputy Assistant Secretary of the Navy Unmanned Systems from October 2015 – June 2018. In this capacity, he was the principal advisor to the Assistant Secretary of the Navy for Research, Development and Acquisition on matters relating to unmanned systems across all domains—land, sea, and air.

Mr. Kelley joined the civil service in 2015 following a 32-year career as a United States Marine.

In 1983, Mr. Kelley graduated from the University of Notre Dame with a degree in Aerospace Engineering and was the recipient of the Naval ROTC Donald R. Bertling Award. He was commissioned as a Second Lieutenant after completing Officer Candidate School.

He attended the Marine Corps War College and taught at the Command and Staff College. Mr. Kelley transferred to Marine Corps Systems Command in Quantico, VA, where he was the Program Manager for Unmanned Systems. His next assignment was Military Assistant to Dr. Delores Etter, then the Assistant Secretary of the Navy for Research, Engineering, and Acquisition.

In August 2007, Mr. Kelley was assigned to the position of Marine Corps Systems Command Program Manager for Training Systems in Orlando, FL. In August 2009, he was reassigned as the command's Chief of Staff before being promoted to the rank of Brigadier General and assuming command from July 2010 to July 2014. He then served in the position of the Vice Commander, Naval Air Systems Command, preceding his last military assignment as Director for Prototyping, Experimentation and Transition in the Office of the Deputy Assistant Secretary of the Navy.

TradeSpace: Designing A Wargame for USSF Acquisition Program Managers

Lt Col Alan Lin, USSF, PhD—is the Chief of Strategic Development, Space Systems Command and in the developmental engineering (62E) career field. He has 21 active duty service years across USAF and USSF. He was previously an assistant professor of computer science at the Air Force Institute of Technology (AFIT) where he conducted and applied game-based education research. [alan.lin@spaceforce.mil]

Chad Millette—is the former Chief Learning Officer for the United States Space Force's Space Systems Command (SSC). His responsibilities included identifying and providing education and training for the SSC acquisition workforce. Chad is a retired U.S. Air Force Lieutenant Colonel, having served 21 years first as an intelligence officer and then as an acquisition program manager. Chad ended his active-duty career teaching at the Air Force Institute of Technology's (AFIT) School of Systems and Logistics where he taught project management to Air Force personnel. [primetimex2@gmail.com]

John Walsh—is a manager with Toffler Associates. He has worked in a variety of roles across the defense industry and consulted on emerging technologies, supply chain risk, and wargame design and application. He received his graduate degree in Strategic Studies and Economics from the Johns Hopkins School of Advanced International Studies. [jwalsh@tofflerassociates.com]

Cherie Chauvin—is a strategic advisor supporting Toffler Associates' defense and national security customers. Her areas of focus include industrial-organizational psychology, human-systems integration, systems dynamics, behavioral and social sciences in military contexts, ethical and moral implications of military technology, and decision making under uncertainty. She holds a BS in Cognitive Science from the University of California at San Diego, an MA in International Relations from the Maxwell School at Syracuse University, and an MS in Strategic Intelligence from the National Defense Intelligence College. [cchauvin@tofflerassociates.com]

Abstract

The United States Space Force (USSF) Space Systems Command (SSC) worked with Toffler Associates, a future-focused strategic advisory firm based in Arlington, VA, to design, develop, and test an acquisition education wargame. This game fills a known gap in educating and training mid-level space acquisition program managers who have not had an opportunity to experience and practice acquisition-specific critical thinking skills in a realistic scenario. Toffler Associates' facilitators ran two instances of the game with teams of five and six participants from SSC's acquisition workforce. The game used a space acquisition scenario for which participants developed an acquisition strategy within the game mechanics. After a rapid briefing on their approach for approval to proceed, they played the game, creatively developing courses of action to mitigate risks and exploit opportunities. Feedback from participants indicated the game was logical and engaging, and met learning objectives. This initial playtest points to a future for this type of experiential environment in acquisition education as well as other acquisition strategy situations, including team building and real-world acquisition strategy testing.

Introduction

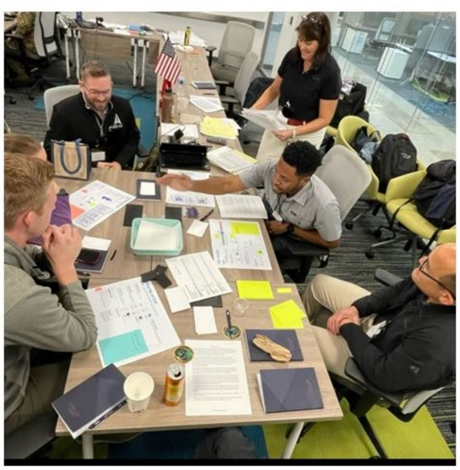
A USSF program management office is in trouble. Their plan: to regain space superiority over a peer competitor by delivering a new generation of space-based sensor. To mature the next-generation technology, they are working with a cohort of small startups, owners of the relevant intellectual property and expertise, guiding them to scale with generous cost-plus contracts. Unwilling to compromise on the capability they deliver to the warfighter, design complexity escalates, and the team finds themselves 8 months behind schedule and 10% over budget.

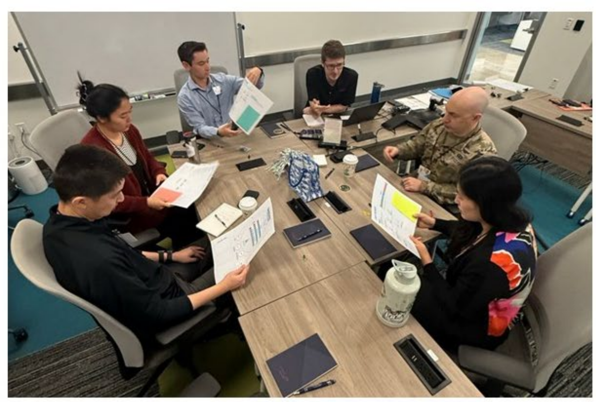
The near-peer competitor demonstrates their version of this capability in orbit. Under pressure to accelerate, the team decides to transition the technology to production, despite outstanding technical risk and incomplete development. The manufacturers take on ambitious contracts but find themselves in a perfect storm of high interest rates and supply chain disruptions. Hat in hand, they inform the program team that it will take significantly more investment to get the capability's first tranche to the launchpad.

The capability is late and approaching obsolescence. Industry partners are threatening to pull out. The Service Acquisition Executive asks difficult questions about risk management and professional judgment. This is the kind of disaster that draws news headlines and Congressional testimonies.

Except, in this case, the scenario is unfolding in a tabletop game. The participants are sitting around a table, drawing cards, throwing dice, and deliberating on the risks and the opportunities of different courses of action. This affords the program office the opportunity to *experiment*. They could reset the clock: if they return to technology maturation and risk reduction, are there steps they could take to begin manufacturing with lower risk? Would a different contracting approach have provided more options? Or perhaps they play out the scenario as it unfolds. If this is an acquisition disaster, what steps can they take to mitigate the impact? What risks can they take to try to steer the program toward a successful outcome?

USSF SSC worked with Toffler Associates' team of futurists and strategic planners to develop the game—TradeSpace℠—as a tool for program manager professional development. The game provides a unique chance to experience acquisition failure in a safe environment, affording an opportunity to build crucial skills. As a prototype, the game remains under development, and its initial playtest with space acquisition professionals demonstrated TradeSpace℠ as a powerful tool to improve acquisition workforce competencies.





USSF SSC Program Managers playing TradeSpace℠ with Toffler Associates facilitators at the SpaceDen, Los Angeles, CA (Nov 13-14, 2024).

FIGURE 1 Images from Initial Playtest of TradeSpace℠

Objectives

Training for most military operational specialties involves a progression that is not mirrored in DAF acquisition training. Using pilot training as an example, USAF pilots first learn how to fly an airplane at initial flight training (IFT). Upon graduation, they transition to learning how to fly the specific airplane with which they've been assigned. Over several years, they spend countless hours in the classroom, in simulators, and in training aircraft. And their training continues when they arrive at their unit to learn how that unit employs the aircraft weapon system—*how do we fly the airplane*. At some point, many pilots get the opportunity to participate in a large-scale realistic exercise like the Air Force's Red Flag. Red Flag organizers often describe their objective as giving pilots the opportunity to fly their first five combat missions before they fly in combat. Only after this training and preparation are pilots ready to execute their mission in the real operational environment.

In contrast, DAF acquisition training starts with 100-level courses from the Defense Acquisition University (DAU). Then, the Air Force Institute of Technology's School of Systems and Logistics (AFIT/LS) offers the first two levels of operational training. Together, they provide the fundamentals of acquisition rules and regulations and DAF-specific lessons. However, unlike pilot training, acquisition officers are then thrown into the fire to execute the mission with real acquisition programs. Often, these officers do not receive any deliberate unit-specific training or the opportunity to practice their craft in a realistic environment where the only consequence for failure is learning.

Literature Review and Background of Past Efforts

Following Department of Defense (DoD) projections that more than 31% of the acquisition workforce would be eligible for retirement by 2026, Defense Civilian and Training Corps (DCTC) created an acquisition game to rapidly train early career acquisition civilians (MacGregor & Cuff, 2024). The creators of this game were motivated to develop additional training and experience methods to supplement traditional acquisition coursework, such as that provided by DAU, because studies suggest that building expert proficiency in acquisitions requires a minimum of 10 years of steady practice (Ericsson et al., 2007; Murphy & Bouffard, 2017).

Similarly, SSC developed the tabletop wargame, Operation Kodiak Dawn, for early career acquisition professionals in the science, engineering, and cyber career fields (Lin et al., 2023). Players were split into red and blue teams, with asymmetric starting conditions, in a cold war space race. The objective of the game is to help the players understand how successes and failures at early stages of technology and system development may have lasting impacts to meeting national-level security goals over time. Through team play, the players also experienced challenges reconciling varying levels of individual team members' risk tolerance when deciding on courses of action for technical maturation and system development strategies. Operation Kodiak Dawn adopted a rule-based approach for mechanically straightforward adjudications, although experienced facilitators were required for the game debrief to ensure lesson objectives were met.

With traction in using wargaming as a tool for defense acquisition training, Georgetown University's Wargaming Course designed a game for SSC around the acquisition process, focused on the defense industrial base and the challenges of developing dual-use technology. Titled "Acquisition Wars," the game's target audience is government acquisition professionals (Shala et al., 2024). Unlike Operation Kodiak Dawn and other DoD wargames, the players play as commercial and private industry partners. The game was designed to be stand-alone and playable in a box, without the need for white cell adjudication, though the ideal use-case setting would also include skilled facilitators for effective debriefing.

In non-education and training settings, DAF Global Futures employs foresight methods to build scenarios and exercises to understand how highly volatile, uncertain, complex, and ambiguous (VUCA) environments impact strategies, opportunities, and challenges. Predictive analytical tools capable of creating accurate and detailed predictions remain beyond state-of-the-art, and the report establishes that this analytical process does not provide predictions.

Instead, it offers future insights to improve decision making (HAF A5/7 Air Force Futures, 2023). Similarly, USSF holds Title 10 wargames for senior level decision-makers that focus on policy and operations, not force modernization at the program office level.

TradeSpace^SM builds upon the existing body of work in wargaming and more specifically, wargaming the acquisition process. It expands on the concepts of previous games to allow players greater opportunity to exercise critical thinking and decision-making about defense program management in a no-risk environment, over and over. The DAF needs new ways to create an experienced acquisition workforce, because it cannot afford to wait for experience through time alone. This urgency is even greater with the implementation of USSF's Officer Training Course, where future force modernization officers will only be identified after their initial 4-year operational tour, reducing the amount defense acquisition experience in comparison to their predecessors at the same time point in their careers. Thus, TradeSpace^SM provides the experiential learning environment for the next generation of acquisition officers to do more "sets and reps" in less time.

Designing an Experiential Learning Environment

The design for TradeSpace^SM sought to simultaneously reflect complex real-world acquisition challenges and appropriately abstract those challenges into an executable game.

For the game to be accepted as a learning tool for program managers, it had to challenge players with the experience of managing major acquisition programs. The defense acquisition system is notoriously complex, with a huge variety of processes, stakeholders, and variables. That complexity is an important feature of the system. The struggle to identify and pursue clear strategic priorities despite that complexity is an important learning experience for players.

Simultaneously, for the game to be engaging and playable, it had to be as straightforward as possible to play from both the perspective of the players and the facilitator. The game is intended to be played in iterative loops: teams or players repeatedly experience failure, observe the consequences, and carry lessons over into their next play session. To achieve that goal, the game was designed to be playable with no more than one support staff and rules that players can learn with no more than 30 minutes introduction.

To meet these competing goals, the game simulates the defense acquisition system at a level of abstraction that focuses players on strategic choices and trade-offs rather than processes. It does not teach acquisition law, policy, or regulations, and it assumes players begin the game with a basic understanding of acquisition authorities.

Game Phases

With some exceptions, defense acquisition follows a natural, repeatable phase structure: program managers define requirements, mature technology, transition it to production, and deploy it into the world. The Adaptive Acquisition Framework allows for flexibility, but this fundamental structure serves as a useful abstraction and facilitates game design as well: games benefit from phases because they allow a varied gameplay loop and encourage players to observe the consequences of their actions and adapt. Figure 2 illustrates how the phases of TradeSpace^SM (bottom) map onto the phases of the Adaptive Acquisition Framework. TradeSpace^SM thereby provides a general-purpose framework that can reflect any acquisition pathway while focusing players on the important differences between phases and how choices in each set up programs for success and failure in subsequent phases.

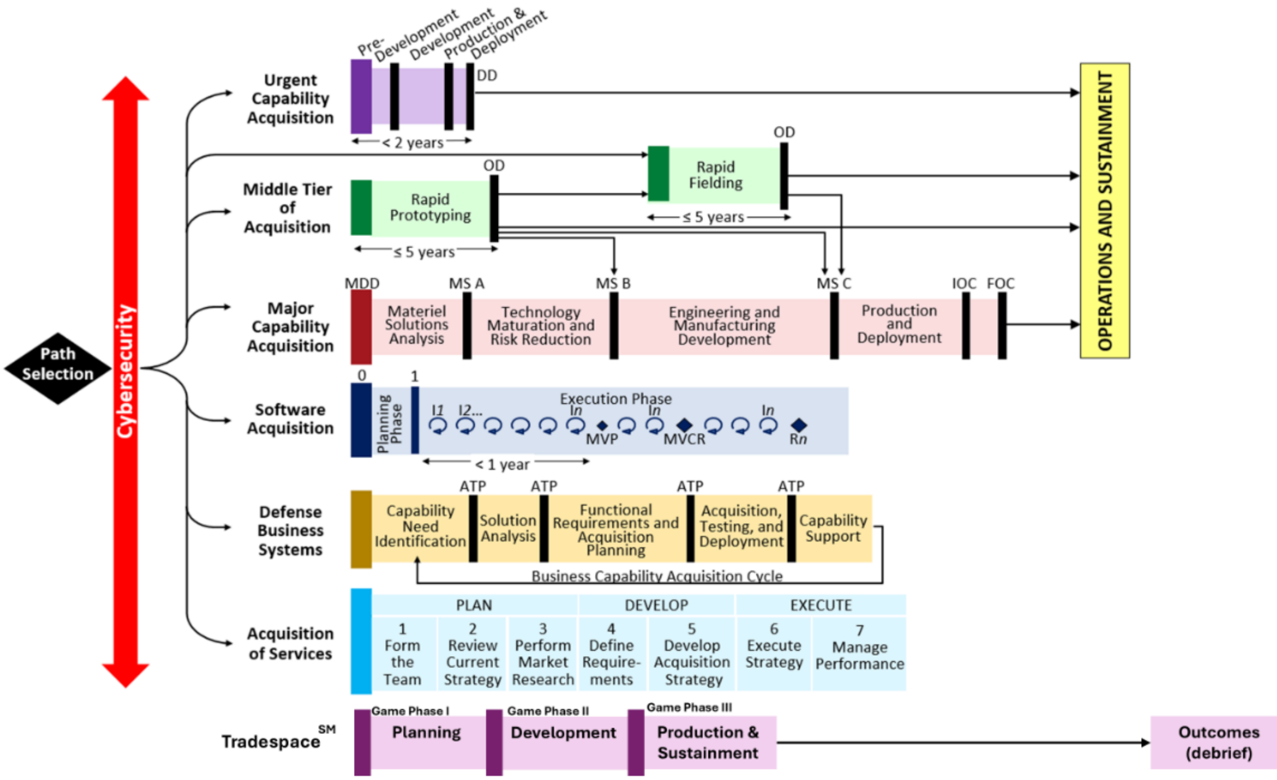


FIGURE 2 Alignment of TradeSpace SM to the Adaptive Acquisition Framework

- **Modeling the Defense Acquisition System**

TradeSpace SM presents players with a discrete set of variables to represent an abstraction of the defense acquisition system. These variables are the "tradespace," in which players must respond to challenges by making strategic trades to protect some priorities at the expense of others. Players define their game objectives in terms of these variables (e.g., deliver ahead of schedule and support the growth of industry partners) and then objectively assess their progress through reference to them. The variables include:

Program Scores. The classic "iron triangle" against which program performance is conventionally measured:

- Cost
- Schedule
- Performance

Contextual Scores. Scores that represent the program's relationship with its environment. High scores allow the program flexibility, and low scores can pose crises:

- Favorability
- Industry Health
- Mission Alignment & Interoperability

A final variable, **Complexity**, captures the relative technical challenge of the program. Complexity is determined by planning choices, presenting players with the trade-offs of designing exquisite versus minimally viable capabilities, and increases the risk that vendors will fail to deliver and engineering challenges will incur cost and schedule penalties.

Other variables matter in acquisition, but we found that in practice these seven provide tools to enforce consequences for all player actions. And important for the game's usability, they are few enough to form a comprehensible tradespace for players to visualize their decision-making. We also found that the mere existence of any variables at all beyond the "iron triangle" was a powerful revelation for program managers, who reported that it opened their eyes to the true scope of tools at their disposal to manage their program.

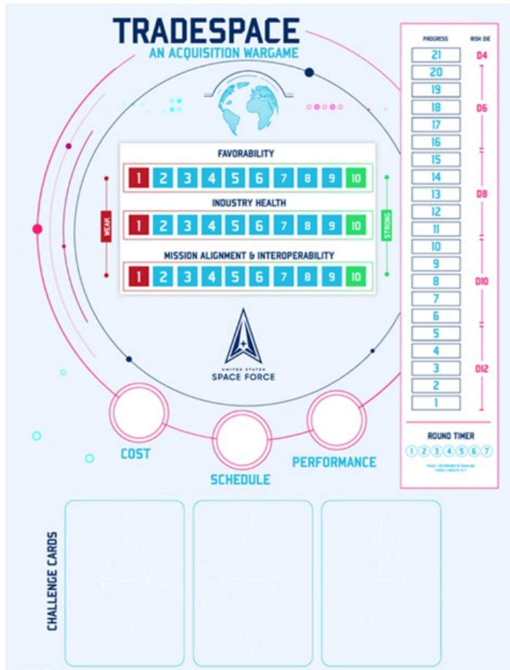


FIGURE 3 TradeSpace^{SM} Scorecard

Challenge & Response

The core gameplay loop of TradeSpace^{SM} challenges teams with a series of interruptions to their acquisition process: new threats or opportunities that impact the program negatively or positively. The game represents the team's time and energy as an Attention resource, the scarcity of which forces an immediate and highly consequential choice: what crises *must* be dealt with to deliver the acquisition strategy and which do the players believe can be ignored? To experience that not all crises are worth the cost of distraction is itself a valuable learning objective. We frequently observed teams struggling to prioritize the *important* over the *urgent* in accordance with their acquisition strategy. In these cases, the facilitator can play a valuable role in shepherding players through analysis paralysis by challenging them to recall their priorities.

Players resolve challenges by describing a course of action to the facilitator and allocating the requisite Attention resources. TradeSpace^{SM} uses a semi-structured adjudication framework. The facilitator interprets the team's course of action in terms of either a *trade-off*, one variable traded for another, or an *uncertain outcome*, in which they must roll a twelve-sided die (adding modifiers representing their investments in program office capabilities) to determine if their course of action succeeds or fails. This approach increases the burden on the facilitator, as they must fluently translate player intention to game mechanics but allows the players complete freedom to creatively define their courses of action. We observed significant learning occurring in this process as players shared lessons-learned from their professional experience and experimented with novel, and in many cases entirely unexpected, courses of action.

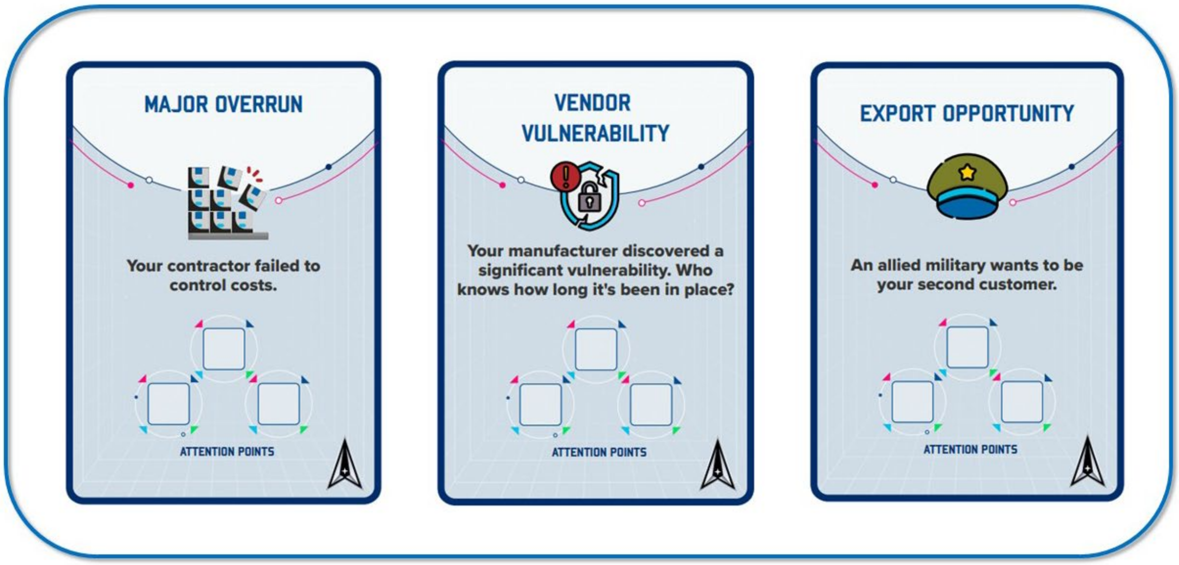


FIGURE 4 TradeSpace^{SM} Challenge Cards

Gameplay Experience

TradeSpace^{SM} was designed to be played by groups of three to six over six to 12 hours. The game is flexible, allowing it to be played in a variety of different formats:

- Small teams testing a real-world acquisition strategy or using the game as a team-building activity.
- Larger professional education cohorts playing multiple games in parallel to test different approaches and observe different outcomes.
- Simple games with one facilitator, or self-adjudication, and expedited discussion.
- Complex games with dedicated facilitator roles (such as a dedicated "Service Acquisition Executive" to provide realistic input) and discussions structured to address specific learning objectives.

The most crucial component of the experience of playing TradeSpace^{SM} is discussion. Experiential learning is effective when the experience allows participants to independently arrive at learning objectives. In TradeSpace^{SM}, players learn by talking about how the experience represented in the game reflects their actual experiences and their formal acquisition education. As such, the game prompts discussion at multiple points.

In the Planning Phase, the facilitator prompts players to discuss the scenario and their strategy to respond to it, with an envisioned end state and a theory for how to deliver it. The game places players (or groups of players) in the roles of the Program Manager, Contracting Officer, and Chief Engineer. This encourages players to share information and approach the problem from different objectives.

In the Development and Production & Sustainment phases, the cycle of each round promotes discussion at two points. First, players discuss *which* challenges and opportunities need a response. The challenge for them is to identify how each will impact their program and then prioritize them according to the elements of their strategy. Second, players discuss *how* they will respond to those challenges and opportunities. This provides an opportunity for them to draw on their professional experience and weigh the pros and cons of different courses of action.

Lastly, the Outcomes phase provides a forum for open-ended, forward-looking discussions of the consequences of the acquisition for all stakeholders.

TradeSpace[SM] functions as a structured framework for facilitating scenario-based discussion. It has significant promise for professional development applications where participants already possess subject matter knowledge but have not had the opportunity to develop expertise through practical applications.

Findings

Toffler Associates and SSC playtested TradeSpace[SM] with a group of acquisition professionals over 2 days in November 2024. The playtest group primarily consisted of early-career acquisition professionals who have received practitioner-level textbook and classroom education.

Two groups played through the game in parallel using a common scenario. The scenario challenged teams to develop a counter to a revolutionary new near-peer adversary communications technology, grappling with significant schedule urgency and an underdeveloped domestic vendor base. Although each group made different strategic choices in response to injects, both successfully realized a strategy of prioritizing rapid delivery in exchange for significant cost overruns and mild underperformance. The game concluded with a facilitated discussion of the consequence of those trade-offs.

Players provided feedback on their experience and learning progress against objectives through surveys. These revealed two primary themes.

First, players found the game to successfully create a safe environment for experiencing risk and the consequences of failure. Players commented that:

- “It was valuable to experience the consequences of choices made in planning.”
- “I liked the dilemma of having many decisions and consequences.”
- “The game encouraged big picture decision making: look at a problem, critically think about possible courses of action, learn from team member experiences throughout.”
- “It was important to maintain margins for dealing with unexpected contingencies.”

This feedback corroborates behaviors observed during the course of the game. Players routinely proposed unconventional solutions to challenges and robustly debated the potential consequences. These included strategies to leverage international collaboration, to sustain the business operations of failing vendors, and to use competition to promote better, cheaper solutions. Not all of these strategies succeeded, but by taking the risk players improved their ability to weigh and manage the consequences.

Second, players found that the game generally improved their critical and creative thinking skills in defense acquisition:

- “The variety of difficult challenges led to valuable discussion.”
- “The hard dilemmas and no-win scenarios were fun and informative to think through.”
- “I learned more about options for handling program risk.”
- “I enjoyed open-ended exploring of options for dealing with problems.”

- “The game gave a broader perspective than I get in my office.”
- “I can imagine a wider variety of challenges in the future.”

Players generally reported that the game presented a unique and thought-provoking challenge and noted that it opened their eyes to risks and opportunities associated with a wider variety of forces, including labor, technology, finance, and regulation.

The players were also evaluated in terms of their self-reported improvement in understanding of the Nine Tenets of Space Acquisition, a set of directives issued in a 2022 memo by Assistant Secretary of the Air Force Frank Calvelli. The overwhelming majority of players reported an improved understanding of seven of the nine directives.

Broadly, TradeSpace^{SM} succeeded in promoting adoption of an adaptive planning approach. Players managed their programs without strict adherence to baselines and instead engaged in nuanced discussions of the relative value of cost, schedule, and performance benchmarks. By presenting players with clear variables and constant demonstrations of the trade-offs between them, the game helped players determine the priorities for their notional program and challenged them to adhere to or evolve that value determination throughout the course of the game.

Future Development

The playtest feedback suggests several areas for refinement. One of the first considerations is applying the existing mechanics to cover trade-offs around cybersecurity, Special Access Programs, and overclassification. This mechanic could interact with challenge cards related to security leaks and opportunities for collaboration. Lower classification could potentially increase the risk of security leaks but enable international cooperation, and the decision to pursue higher classification incurs significantly greater program costs and fewer opportunities for industry collaboration.

On the event cards themselves, the randomness and distribution of the events sometimes created conflicting narratives that disturbed the suspension of disbelief, damaging the immersion and subsequently, the learning objectives. For instance, during the playtest, some teams drew “retiring SMEs” and “labor strike/mass layoffs” cards on the same turn, significantly derailing program schedule. While real-world 2025 events show that these events can certainly coincide, at the time of the playtest, it was considered unlikely. As the game continues to be developed, there are several potential mediations. One design consideration is to modify the distribution of event cards where there are more minor issue event cards than major issue cards, and even fewer catastrophic events cards. In addition, the next iteration may have event cards that more uniquely specify the type of attention, such as requiring “Engineering” attention or “Contracting” attention, rather than treating all attention costs equally. The intent is to have more constraints when dealing with the event cards per round, such that even a confluence of minor issues can present a challenge if all hardships fall on a single function to address. Additionally, opportunities to leverage a digital database of event injects creates opportunities for more deliberate inject timing and combinations to improve the flow of events from round-to-round.

Finally, regarding the functional roles, the playtest revealed that the team played fully cooperatively and collaboratively. While ideal, this is slightly negative learning, as in practice, PM, Engineering, and Contracting teams have their own internal objectives. Future iterations of the game may implement a system where each functional team must trade between meeting their internal goals and the goal for the entire program, that puts their function at odds with the rest of the team.

Bibliography

Caffrey, M. (2019). *On wargaming.* Naval War College Press. https://digital-commons.usnwc.edu/newport-papers/43/

Defense Acquisition University. (n.d.). *Adaptive acquisition framework pathways*. Retrieved March 31, 2025, from https://aaf.dau.edu/aaf/aaf-pathways/

Department of the Air Force. (2022, October 31). *Memorandum for the Department of Air Force space acquisition workforce*. https://www.safsq.hq.af.mil/Portals/76/documents/ASAF%20-%20Space%20Acquisition%20Tenets%20(31%20Oct%2022)%20FINAL.pdf

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). *The making of an expert.* Harvard Business Review. https://hbr.org/2007/07/the-making-of-an-expert

HAF A5/7 Air Force Futures. (2023). *Air Force global futures report: Joint functions in 2040.* HAF. https://www.af.mil/Portals/1/documents/2023SAF/Air_Force_Global_Futures_Report.pdf

Lin, A., Bull, D., DiLuca, E., & Kennedy, B. (2023). *Operation Kodiak Dawn*.

MacGregor, M., & Cuff, T. (2024, May 8–9). The acquisition game: Learning through play. *In Proceedings of the 21st Annual Acquisition Research Symposium.* Naval Postgraduate School, Monterey, CA. https://dair.nps.edu/bitstream/123456789/5168/1/SYM-AM-24-114.pdf

Murphy, C., & Bouffard, A. (2017). *Understanding defense acquisition.* MITRE Corporation. https://www.mitre.org/sites/default/files/publications/17-1585-understanding-defense-acquisition-workforce-challenges.pdf

Shala, R., Beto, J., Spreenberg, O., Bae, S., & Lin, A. (2024). *Acquisition wars*. https://www.linkedin.com/posts/activity-7194350773424590848-N0fR

DAU Innovate to Win: DoD Upskilling Innovation Readiness

Jonathan M. Barkand—Learning Systems Program Manager**,** Defense Acquisition University

Lolita M. Horne—Learning Systems Program Manager, Defense Acquisition University

Julie A. Knechtel—Professor of Program Management, Defense Acquisition University

Abstract

In the context of divisive geopolitics, the Department of Defense (DoD) faces challenges in assessing innovation readiness due to the lack of a unified language and metric. Despite over 200 active innovation cells, efforts are fragmented and underfunded (Theodotou, 2023). In response to the National Defense Authorization Act of 2023 and Deputy Secretary of Defense Dr. Kathleen H. Hick's commitment to innovation, the Defense Acquisition University (DAU) launched the "Innovate to Win" pilot program in 2022.

This program aims to codify the innovation skillset required by the DoD workforce by integrating academic research and industry practices into a DoD context. Key elements include an Innovation Competencies and Skills Model, Innovation Readiness Self-Assessment resulting in Curated Learning Pathways, and an Innovate to Win Playbook for all levels of leadership. Over 3,700 personnel have completed the self-assessment, receiving tailored learning pathways to enhance their innovation skills. This research examines the correlation between self-assessed innovation readiness and organizational support, aiming to refine the self-assessment tool and provide resources to foster innovation across the DoD workforce.

Keywords: Innovate to Win, Innovation Readiness, Innovation Ecosystem, DoD Workforce, Defense Acquisition University (DAU)

Background

The Department of Defense (DoD) operates in an era of rapid technological change, among shifting geopolitical forces and nascent threats that demand relentless and swift adaptation. To gain and maintain the operational advantage, the DoD must scale innovation at an unprecedented pace, ensuring that both technological advancements and operational concepts evolve to meet the challenges of modern warfare (Defense Innovation Board, 2023). Despite efforts to drive innovation, challenges persist. Bureaucratic resistance, risk aversion, and rigid hierarchical structures often hinder the ability to rapidly develop, test, and implement new solutions (Mahnken et al., 2023). An institutionalized culture of innovation can serve as an antidote to those challenges (Bowdren, 2024). A critical component of innovation at scale is fostering an innovation mindset among the DoD workforce (Theodotou, 2023a).

Research highlights that innovation is not solely driven by advanced technology, funding, or watershed moments, but by the cognitive and behavioral traits of individuals supported within an enabling system wide approach of an organization (Miller & Brankovic, 2011). Organizations play an essential role in accelerating the speed of innovation within the DoD. Studies show that "effective management of [organizational] culture lies at the heart of organizational innovation" (Tushman & O'Reilly, 2002, Chapter 5). Organizations that promote both creativity and speed to implementation report more success than organizations that focus on only one; both are needed (Tushman & O'Reilly, 2002).

Within the DoD, innovation-enabling entities such as AFWERX, Army Futures Command, Defense Innovation Unit (DIU), and NavalX have demonstrated that providing dedicated resources, streamlined acquisition pathways, and partnerships with private sector innovators can significantly enhance innovation outcomes (Defense Innovation Board, 2025). However, the DoD's innovation efforts are fragmented, with over 200 active innovation organizations operating independently. These organizations employ diverse definitions, languages, skill sets, and metrics. The absence of a unified language and metric for innovation has led to inefficiencies, fragmentation, and the underfunding of innovation efforts within the DoD (Theodotou, 2023b). Results garnered by varied innovation organizations validate the success of organizations that foster both innovative mindsets and rapid implementation. They also highlight the need for a unified approach to streamline these efforts and enhance overall innovation readiness across the DoD workforce (Theodotou, 2023b). In response, DAU developed the Innovate to Win initiative and deployed it in 2023 (https://www.dau.edu/innovatetowin). It is designed to:

- Establish the baseline innovation competencies and skills needed to cultivate an innovation mindset.
- Define innovation readiness to standardized lexicon and optimize innovation metrics across the DoD workforce.
- Provide tools and resources that enhance innovation behaviors for individuals and assist leaders in motivating and cultivating a culture of innovation.

An innovation mindset is characterized by curiosity, adaptability, and a willingness to take calculated risks, essential for developing novel solutions to complex national security challenges (Theodotou, 2023b). This paper analyzes 3,700 self-assessment responses. Specifically, it examines how employees rate their innovation mindset and what organizational barriers and incentives are prevalent.

Methodology

Inspired by language in the National Defense Authorization Act of 2023 and the commitment of the former Deputy Secretary of Defense Dr. Kathleen H. Hicks, DAU launched the “Innovate to Win” pilot program in 2022. The program aims to codify the innovation skillset required by the DoD workforce by leveraging successful tactics from academic research, industry practices, and existing innovation cells. The three key elements of the program are: 1. Innovation Competencies and Skills Model, 2. Self-Assessment and Curated Learning Pathways, and 3. Innovate to Win Playbook.

Innovation Competencies and Skills Model

The Innovation Competencies and Skills Model is a standardized framework for assessing innovation readiness across the DoD, shown in Figure 1. Data triangulated from academic research, workforce insights, and industry best practices are represented in one of three domains of practice:

- Thinking - Thinking innovatively hinges on embracing a growth mindset, using critical thinking approaches, cultivating creativity, taking calculated risks, and practicing futures thinking.
- Collaborating - Collaborating to foster innovative behaviors includes networking and communicating effectively, including asking probing questions and listening actively.
- Cultivating - Cultivating innovative behaviors includes experimenting, observing, embracing a comprehensive approach, driving change, integrating people, ideas, and learning, and embracing lifelong learning.

Each of the three domains is characterized by supporting competencies and skills. The Innovation Competency Model introduced innovation competencies, enables self-evaluation of innovation skills and reports an innovation readiness metric. It then provides a curated pathway of targeted learning assets. The Innovation Readiness Metric calculates the median of each response aligned to each competency skill and rolls it up to the domain level. A workforce member can gauge their own innovation readiness by completing the self-assessment. Team leaders can use the Innovation Readiness Dashboard to review and report the aggregate innovation readiness of the team based on the individual self-assessments of each team member.

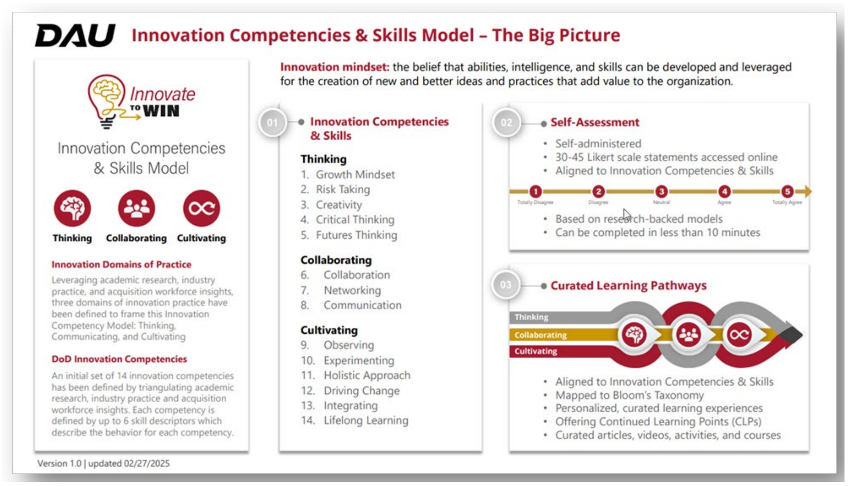


Figure 1. DAU Innovation Competencies & Skills Model: A Three Step Approach

Self-Assessment

The self-assessment instrument provides a gauge of self-reported innovation readiness shown in Table 1. It is central to the model and is crafted to pinpoint strengths and areas for growth. Access is available at https://www.dau.edu/innovatetowin/self-assess.

Table 1. DAU Innovate to Win: Self-Assessment

THINKING	RATING				
GROWTH MINDSET	1	2	3	4	5
My basic intelligence is something that is not static and can be expanded over time.					
As I learn new things I am improving my overall intelligence.					
When I face a challenge, I find a way to persevere.					
I frequently iterate on a idea and work on it in small doses.					
I prefer to dig deep into the details of a process to insure I fully learn and undestand it.					
RISK TAKING					
I am willing to try a new way of accomplishing my work.					
When faced with a challenge, I'm willing to embrace the possibility of failure as a way to learn.					
I regularly ask questions that challenge the status quo.					
I am comfortable making calculated decisions in the absence of complete information.					
CREATIVITY					
I frequently connect seemingly unreleated concepts and ideas from diverse disciplines.					
I frequently tinker with problems and seek new ways to tackle them.					
I am familiar with Design Thinking process models or other tools/models to create innovative solutions.					
I am comfortable using my imagination to come up with new ideas.					
CRITICAL THINKING					
I gather relevant information and think through multiple solutions before making a decision.					
When tackling a problem, I listen to other people's opinions, points of view and their perspective.					
I am willing to make decisions without all the data I would prefer to have.					
I understand how unconscious bias, assumptions, and empathy can impact decision making.					
FUTURES THINKING					
To envision future scenarios, I leverage data, trends, and technology.					
I use analysis to determine plausible future scenarios.					
COLLABORATING	RATING				
COLLABORATION	1	2	3	4	5
I spotlight my team members that do great work.					
When I have a new idea, I try to involve people who are able to help improve and adopt it.					
I recognize when stakeholders are not aligned and that it impacts my work products and I reach out to facilitate					
I seek to understand opposing views to my work efforts, and facilitate discussion to reach desired outcomes.					
NETWORKING					
I deliberately seek to connect with other people to learn and seek advice.					
I seek to connect people I know that may enjoy meeting each other.					
I am active on Professional social media sites and participate frequently in professional organizations in my					
ALLYSHIP					
I recognize that the more diverse my team is, the more innovate we are.					
I seek out learning opportunities beyond just mandatory courses to learn about diversity, equity, inclusion and					
I seek to amplify diverse and innovative ideeas from people who don't always have a voice of their own.					
COMMUNICATION					
I use storytelling to influence others.					
When I have a new idea, I am comfortable and have had success reaching out to involve people who are able to					
I often ask open ended questions to expand on a converstion.					
CULTIVATING	RATING				
OBSERVING	1	2	3	4	5
I often explore a topic from different angles and uncover new ideas.					
I seek opportunities to learn how others in other industries and sectors solved their problems.					
I watch and listen intently to understand how things or processes work.					
EXPERIMENTING					
During a project, I perfer to try multiple solutions and options rather than being told what or how to do it.					
When I am working on a project, I usually try new ways of doing things.					
HOLISTIC APPROACH					
I often zoom into the details of a situation and then zoom out to the big picture.					
I maintain a high degree of situational awareness.					
DRIVING CHANGE					
During a crisis, I manage to reign in my emotions and focus.					
When I have a new idea, I champion my cause with leadership.					
I engage multiple stakeholdes to build coalitions to drive results.					
INTEGRATING					
I seek resources (people, dollars, tools, etc.) needed to implement new ideas.					
I thrive in connecting seemingly unrelated people, ideas, or topics.					
I develop suitable plans and schedules for the implementation of new ideas.					
LIFELONG LEARNING					
When I am faced with a new challenge or unexpected problem, I spend time seeking out people or information					
I voluntarily take training courses and feel motivated from learning new skills.					
QUALITATIVE QUESTIONS:					
What barriers have you encountered when trying to be innovative (new processes, new idea) in your organization?					
What incentives or other resources does your organizatoin devote to innovation?					

RATING SCALE	
SCORE RATING	**DESCRIPTION**
1	Strongly Disagree
2	Disagree
3	Neither Agree nor Disagree
4	Agree
5	Strongly Agree

Curated Learning Pathways

The curated learning pathways are tailored to align learning assets to the innovation competencies and skills, cultivating the engagement of learners with the content and ensuring fit with the practical and actionable learning based on individual self-assessment results. DAU identifies a list of learning experiences and generates a learning asset bank from which the online learning management system (Cornerstone on Demand [CSOD]) algorithm will curate specific, personalized learning pathways for each member of the workforce based on the self-assessment results. Providing learning resources enables benefits for the individual learners and organizations at large (Figure 2).

Figure 2. DAU Innovate to Win: Curated Learning Pathways Benefits for the Learner and Organization

Alignment to Adult Learning Theories

While there are more than 30 research-based learning theories and models, Innovate to Win aligns itself with four foundational schools of thought: cognitivism, behaviorism, constructivism, and connectivism (Drew, 2024). The Innovation Competency Skills Model Curated Learning Pathways hinge on cognitivism by aligning with the updated Bloom's taxonomy (Wilson, 2016). In 2001, Anderson and Krathwohl revised Bloom's taxonomy to focus on how the learners remember, understand, apply, analyze, evaluate, and create meaning. Additionally, behaviorism is reflected in the structured and measurable approach to learning, where specific behaviors are reinforced through guided practice and feedback. The curated pathways and Innovate to Win Playbook focus on how learners understand and apply the innovation competency skills (Bloom's Levels 2, 3, and 4) which center on understanding, applying, and analyzing innovative skills and behaviors. For example:

> **Bloom's Level 2 -** Understanding: Compare and contrast fixed and growth mindset behaviors, relate with other colleagues, rephrase assertions into great questions.
>
> **Bloom's Level 3** - Applying: Experiment with new ways of doing things, observe and interview users to build new innovative solutions, model lifelong learning by reading one page of a book per day.
>
> **Bloom's Level 4** - Analyzing: Take part in experiments, test for user experience, build strategic network relationships.

The Curated Pathways incorporate elements of Constructivism as some of the learning assets are scaffolded. The learning experiences build on each other, and learning happens when the learner interacts with tools, language, and organizational structures. Most importantly, the pathways leverage elements of Connectivism whereby learning results from a variety of input. Connectivism has been hailed as the learning theory of the digital age because it acknowledges the power of technology as an enabler in learning whereby learners can quickly select content from their Curated Learning Pathway, dig deeper by searching for an instructional video, connect with others through a community of learning, chat live, and together find solutions to support the warfighter (Duke et al., 2013).

Learning Asset Modalities

The learning modalities used within the Innovation Pathways are visual (watch), auditory (listen), and kinesthetic (read). By including an array of instruction methods, all preferred learning styles are enabled to support preferred learning style. Offering multimodal learning options also creates an exciting learning environment and leads to increased learner engagement and retention.

Learning Asset Quality

All recommended learning assets are first benchmarked against DAU's quality rubric and value metrics, demonstrated by superior feedback on Net Promoter Scores, learning hours, and learner qualitative and quantitative survey results. Quality Assurance for each proposed asset required more than perusing content and visuals. It also included validating elements of navigation, interactivities, 508 compliance, alignment of knowledge checks to the asset objectives, and ease of user access. DAU ensured alignment, engagement, and fit using rubric as a guide to result in the standardized selection of productive, engaging, and effective assets in support of an innovative mindset.

Content Rubric

All assets were evaluated using the content rubric criterion in Table 2. Microlearning assets, videos, articles, job tools, and infographics required a minimum of 12 points to be considered for inclusion in the learning asset bank.

Table 2. Learning Content Rubric

RUBRIC CRITERION/RATING		0 POINTS	1 POINT	3 POINTS
1	Practice: Includes opportunities to practice	Zero opportunities	One opportunity	More than one opportunity
2	Relevance: Demonstrates relevancy to the Workforce	Never	Once	More than once
3	Engagement: Offers the learner engagement opportunities (click a button, solve a puzzle, write down thoughts, answer a question, etc.)	Offers zero engagement opportunity	Offers up to two engagement opportunities	Offers three or more engagement opportunities
4	Cross-Referencing: Connects to other learning / Facilitates self-directed learning	Never	Passively by mentioning other learning	Actively by offering hyperlinks, book recommendations, etc.
5	Assessment	None	Assessment present, but offers no feedback to user	Assessment present and offers feedback to user
6	Access: Ease of Access	Not easy to access	Somewhat easy to access	Easy to access
7	Currency: Age of Content	>2 years old	Between one and two years old	Less than a year old
8	Time to Complete	>6 hours	Between three and six hours	Less than three hours

Innovate to Win Playbook

The Innovate to Win Playbook serves as a practical tool for leaders and supervisors to create a culture that motivates and fosters innovation within their teams (https://www.dau.edu/innovatetowin/perform). It includes a step-by-step guide with seven plays that team leaders and supervisors can use to motivate and cultivate innovation within their teams. Each play focuses on a different aspect of fostering innovation and is designed to be actionable, accessible, and user centered. The seven plays are:

1. Define a Compelling Vision and Goals
2. Provide Top Cover
3. Collaborate and Communicate
4. Embrace Risk
5. Foster Curiosity
6. Cultivate a Learning Culture
7. Recognize and Reward Innovators

The playbook is part of a larger effort to increase the culture of innovation across the DoD at large. By utilizing the content included within each play, leaders gain measurable, valuable insights and practical knowledge that can be applied to real-world challenges. It encourages teaming to enhance problem-solving skills, explore new technologies, and collaborate with other professionals. To this end, each play includes four key elements:

- What the play is about
- Why it is useful
- How to use the play
- How to measure the play's success

The playbook can be used independent of, or in conjunction with, the Innovate to Win Self-Assessment. The aggregated results of each team member's self-assessment are provided on a custom dashboard report, which helps leaders select and apply the plays that are most relevant to the needs identified by team results in the dashboard. Alternatively, leaders can use their knowledge of the team, or any other barometer of choice to define utilization of the playbook. Regardless, the iterative process of assessing, applying plays, and recalibrating the innovation readiness baseline helps grow innovation and support an innovative culture.

Results

The findings suggest that a supportive organizational environment is crucial for fostering innovation. The self-assessment tool provides valuable insights into individual and organizational innovation readiness, informing the development of resources to encourage psychologically safe work environments. Data from 3,719 individual competency self-assessments were collected. Within this population, 72% (2,676 total) of the respondents were identified as Defense Acquisition Workforce (DAW) members. The remaining responders were DoD military/civilian employees (1,011), federal government employees (141), and industry (9). Component representation of responders is identified in Table 3.

Table 3. Component Representation of Responders

Component	Total Number of Responders	Count of DAW Members	% DAW
4th Estate	19707	14415	73%
Air Force	11389	8318	73%
Army	13247	8891	67%
Navy	11997	8965	75%

After removing all other federal users and industry responders, the average score for each competency area was calculated. Respondents are not required to answer every question and can leave questions unanswered. A non-response was not considered a zero, but a null. Each competency had more than one question; the score for each question (ranging from 1 to 5) was averaged to give an overall competency score. On a scale of 1 to 5, all component categories reported a score of "Agree" to the competencies across each of the three domains in Table 4.

Table 4. Component Domain Scores

Component	Collaborating	Cultivating	Thinking
4th Estate	4.0	4.1	4.2
Air Force	4.0	4.1	4.1
Army	3.9	4.0	4.1
Navy	3.9	4.0	4.1

The average score of all 3,719 respondents was calculated across all domains and competencies in Table 5.

Table 5. Average Score by Domain and Competency of 3,719 Respondents

Domain and Competency	Average Score
Collaborating	**3.96**
Allyship	4.04
Collaboration	4.30
Communication	3.97
Networking	3.55
Cultivating	**4.04**
Driving Change	3.98
Experimenting	3.86
Holistic Approach	4.04
Integrating	3.85
Lifelong Learning	4.32
Observing	4.19
Thinking	**4.11**
Creativity	3.88
Critical Thinking	4.21
Futures Thinking	4.06
Growth Mindset	4.34
Risk Taking	4.08
Grand Total	**4.05**

While the quantitative scores on average showed that in general employees agree that they have innovation competencies, the qualitative questions provided a deeper insight into local barriers and incentives at their organization. Two qualitative questions were asked in the competency self-assessment, and a response was not required.

For the question **"What barriers have you encountered when trying to be innovative (new process, new idea) in your organization?"** responses are consolidated into four emergent themes:

1. **Resistance to Change**
   - Many respondents highlighted a status quo mentality and a fear of failure as significant barriers. This included a reluctance to support new ideas and a preference for entrenched old-school methods.
   - There is also a notable leadership reluctance to change, with some leaders being complacent or procrastinating.

2. **Resource Constraints**
   - A common theme is the lack of funding, equipment, and other resources. Respondents referenced limited manpower and time, as well as overly burdensome processes that focus more on compliance than mission accomplishment.
   - Information technology infrastructure issues and difficulty in finding historical information and understanding current systems were also noted.
3. **Organizational Culture**
   - Hierarchical barriers and a lack of procedural knowledge among leadership were frequently stated.
   - Groupthink, process paralysis, and biases related to gender and rank also hinder innovation.
   - There is a lack of learning culture and a fear of the unknown, which makes employees unwilling to take risks.
4. **Communication and Support**
   - Poor communication and collaboration, along with a lack of buy-in from leadership and team members, were significant barriers.
   - Delays in review and approval processes and non-responsiveness from principals were also highlighted.

For the question "**What incentives or other resources does your organization devote to innovation?"** responses are consolidated into five emergent themes:

1. **Recognition and Awards**
   - Respondents shared various forms of recognition, including cash awards, time-off awards, and certificates.
   - Public and private acknowledgment of achievements and on-the-spot awards for innovative efforts were noted.
2. **Training and Development**
   - Opportunities for training and professional development were highlighted, including Digital Development Fridays and Lean Six Sigma belts for process improvement projects.
3. **Supportive Programs**
   - Some organizations have innovation programs that allow employees to work on new ideas, internal investment projects, and innovation cells.
   - Results Accelerators and working groups were also mentioned.
4. **Leadership and Organizational Support**
   - Encouragement from senior leaders and platforms for sharing ideas and feedback were noted as important resources.
   - Flexibility and a focus on end results rather than strict processes were also highlighted.
5. **Limited or No Incentives**
   - Some respondents reported no incentives or resources devoted to innovation, indicating a lack of clear incentives and support from the organization.

Results Summary

The competency assessment was unchanged from FY2023 to FY2024. In FY2025 the allyship section was removed from the competency assessment in accordance with the Presidential Executive Order on diversity, equity, and inclusion (Executive Order No. 14151, 2025). The instrument has remained consistent other than this change. Overall, the three highest scoring competencies were Growth Mindset (4.34), Lifelong Learning (4.32), and Collaboration (4.30). The statements from these competencies focused on improving as an individual through training and learning through collaboration with team members. The three lowest scoring competencies were Networking (3.55), Integrating (3.85), and Experimenting (3.86). While these scores are still within neutral to agree, the statements focused on seeking new ways of doing things, connecting unrelated people and topics, and actively seeking connections and connecting others. The highest and lowest scoring competencies were across each domain, which shows overlapping domains on specific competencies.

Conclusions

The DAU "Innovate to Win" program represents a comprehensive approach to embedding innovation into the foundation of DoD organizations. By standardizing the assessment of innovation readiness and providing curated learning pathways, the program aims to upskill the DoD workforce at scale. Based on the success of the DoD Innovate to Win program, it was extended to the federal workforce in 2024 and is also accessible via the Federal Acquisition Institute (FAI, 2024). FAI used the same competency model but adjusted the suggested training courses to align to the broader federal workforce. This allows future research to compare self-assessed competencies from DoD and federal employees. Future research should focus on identifying the most effective training methodologies for enhancing innovation competencies, exploring the correlation between innovation readiness and job performance, and analyzing the long-term impact of the program.

References

Bowdren, S. J. (2024, April 15). *A culture of innovation drives acceleration!* Marine Corps Association. https://www.mca-marines.org/gazette/a-culture-of-innovation/

Defense Acquisition University. (2022a, October). *Innovate to win*. https://www.dau.edu/innovatetowin

Defense Acquisition University. (2022b, October). *Innovate to win: Self-assessment*. https://www.dau.edu/innovatetowin/self-assess

Defense Acquisition University. (2024, September 30). *Innovate to win playbook*. https://www.dau.edu/tools/innovate-win-playbook

Defense Innovation Board. (2023, July 17). *An innovation strategy of the decisive decade: National Defense Science & Technology Strategy Review Task Force*. https://innovation.defense.gov/Portals/63/documents/Recommendations/DIB_An%20Innovation%20Strategy%20for%20the%20Decisive%20Decade_230717.pdf?ver=9b09Qgn2SKAvOL7i1sCgBA%3d%3d

Defense Innovation Board. (2025, January 8). *Scaling nontraditional defense innovation*. https://innovation.defense.gov/Portals/63/DIB%20Scaling%20Nontraditional%20Defense%20Innovation%20250113%20PUBLISHED.pdf

Drew, C. (2024, May 30). *31 major learning theories in education, explained!* https://helpfulprofessor.com/learning-theories/

Duke, B., Harper, G., & Johnston, M. (2013). *Connectivism as a digital age learning theory*. HETLReview2013SpecialIssueArticle1.pdf

Exec. Order No. 14151, 90 F.R. 8339 (2025). https://www.federalregister.gov/documents/2025/01/29/2025-01953/ending-radical-and-wasteful-government-dei-programs-and-preferencing

Federal Acquisition Institute. (2024). *Innovate to win*. https://www.fai.gov/training/innovate-to-win

James M. Inhofe National Defense Authorization Act for Fiscal Year 2023, Pub. L. No. 117-263, § 236. 2490. (2022). https://www.congress.gov/117/plaws/publ263/PLAW-117publ263.pdf

Mahnken, T. G., Montgomery, E. B., & Hacker, T. (2023). *Innovating for great power competition: An examination of service and joint innovation efforts*. Center for Strategic and Budgetary Assessments. https://csbaonline.org/uploads/documents/CSBA8337_(Innovating_For_Great_Power_Competition_Report)_web2.pdf

Miller, P., & Brankovic, A. (2011). Building a creative culture for innovation. *IESE Insight*, *11*, 51–58. https://doi.org/10.15581/002.ART-2049

Theodotou, M. (2023a, July/August). Innovate to win – Overcoming DoD's innovation scaling problem. *Defense Acquisition Magazine*, *LII*(4, DAU 293), 20–23.

Theodotou, M. (2023b, July/August). Workforce innovation readiness: Vital in this decisive decade. *Defense Acquisition Magazine*, *LII*(4, DAU 293), 48–51.

Tushman, M. L., & O'Reilly, C. A. (2002). *Winning through innovation: A practical guide to leading organizational change and renewal*. Harvard Business Review Press.

Wilson, L. O. (2016). *Anderson and Krathwohl Bloom's taxonomy revised*. Anderson-and-Krathwohl_Revised-Blooms-Taxonomy.pdf

Tailored Interventions to Foster Acquisition Innovation: Piloting the Innovation Alliance Program

Amanda M. Girth, PhD—Associate Professor at The Ohio State University's John Glenn College of Public Affairs, specializes in acquisition policy and practice. Based in Washington, DC, her research focuses on contract management, performance incentives, and acquisition innovation. She has published in top journals such as the Journal of Public Administration Research and Theory and Journal of Supply Chain Management. Girth has testified before Congress and has led research sponsored by the Department of Defense. She is an Adjunct Fellow in the Defense-Industrial Initiatives Group at CSIS, and a member of the Research Council for the SERC/AIRC. [girth.1@osu.edu]

Michael F. Rayo, PhD—is an Associate Professor in Integrated Systems Engineering at The Ohio State University, where he also directs the Cognitive Systems Engineering Lab and is a core faculty member of the Translational Data Analytics Institute. His research focuses on resilience engineering, system signals, and human-machine teamwork across healthcare, military, and other high-stakes domains. He has authored numerous peer-reviewed publications and has led funded work from agencies including the Air Force and AHRQ. Rayo holds degrees in cognitive systems engineering, chemical engineering, and music performance, and received the 2019 IBM Faculty Award for his work on organizational resilience. [rayo.3@osu.edu]

Laura Maguire, PhD—is a Research Scientist with the Cognitive Systems Engineering Lab at The Ohio State University. She specializes in resilience engineering, cognitive systems engineering, and organizational process improvement, working with fast-paced, high-stakes environments. Maguire has led research and advisory efforts for Fortune 500 companies, start-ups, government agencies, and professional associations. She previously served as Head of Research at a Series A software start-up based on her doctoral work, acquired in 2023. Known for her practical approach, she is also a skilled facilitator and international keynote speaker on human performance in complex systems. [maguire.81@osu.edu]

E. Miriam Balkin, PhD Candidate—became interested in the science of human and organizational performance during her career as a Registered Nurse, and she is currently pursuing a PhD in Cognitive Systems Engineering at The Ohio State University. Balkin has conducted human factors research in various industries including healthcare, commercial aviation, and defense infrastructure. She is currently developing proactive work system monitoring strategies that integrate cognitive interviewing and data analytics, and her doctoral research is exploring how program managers mitigate barriers and seize opportunities when implementing system-level innovation. [balkin.7@osu.edu]

Morgan E. Reynolds, PhD—is an engineer trained in both cognitive systems and resilience engineering. Her work focuses on how practitioners adjust their coordination, monitoring, and authority relationships to accomplish fast-paced, high-risk work. Her experience studying coordination spans healthcare, air traffic control, software engineering, and defense domains. When not explicitly focused on coordination, Morgan can be found researching human–machine teaming, representation design, building resilience, and supporting system safety. [reynolds.818@osu.edu]

Abstract

A key challenge in the U.S. Department of Defense (DoD) is aligning, prioritizing, and adopting acquisition innovation. This research explores the underlying systemic pressures that both impede and enable innovation in the acquisition workforce and proposes a structured intervention program, the Innovation Alliance Program (IAP), aimed at promoting a healthy innovation culture. Piloted at the Department of the U.S. Air Force, the IAP develops capacity through training and strengthening the collaborative networks within the organization to enhance the scalability and adoption of innovative practices, which addresses a critical issue of promulgating micro-innovations for larger-scale impact across the DoD.

Keywords: Innovation in Acquisition; Acquisition Workforce

Introduction

The 2022 National Defense Strategy (NDS) underscores the need for the Department of Defense (DoD) to modernize its acquisition processes, emphasizing innovation and adaptability to meet evolving strategic demands. Facing rapidly evolving threats from near-peer adversaries such as China and Russia, the DoD must continuously adapt its acquisition system to remain operationally effective and mission-ready. Innovation enables acquisition professionals to respond with agility, ensuring that the DoD can deliver timely and effective solutions to support the warfighter.

Innovation is perhaps best understood as a set of behaviors that introduce novel tools, modify or reengineer existing processes, or hybrid efforts that integrate both (Girth et al., 2022). Innovation encompasses not only technological advancements, but also cultural and organizational changes aimed at creating public value and improving service delivery (Osborne & Brown, 2011). Public sector innovation often involves collaboration, co-creation, and experimentation within complex organizational settings to achieve increased administrative efficiency and effectiveness (Demircioglu & Audretsch, 2024). Yet, the pursuit of novel solutions in the public sector frequently collides with institutional norms that prioritize procedural compliance and status quo. This dynamic creates the "innovation–risk paradox" whereby public organizations are increasingly called upon to innovate to address complex social, economic, and national security challenges, yet the institutional environment often penalizes the very behaviors—experimentation, iteration, and failure tolerance—that innovation requires (Moynihan & Landuyt, 2009; Osborne & Brown, 2011).

Risk aversion in public agencies manifests through procedural rigidity, short-term performance pressures, and reluctance to deviate from established norms or practices, all of which can inhibit organizational learning and adaptive change (Kettl, 2015; Walker, 2006). However, the development of psychological safety, supportive leadership, and cross-functional collaboration can mitigate perceived risks (De Vries et al., 2016; Demircioglu & Audretsch, 2017). The challenge lies in constructing systems and cultures where risk-taking is encouraged and supported, rather than reflexively avoided.

The objective of this study was to uncover and mediate underlying systemic pressures on the acquisition workforce in the DoD that impede agility and innovative behaviors. A program of cohesive interventions is presented—the Innovation Alliance Program (IAP)—to address systemic factors to incentivize lasting behavioral and cultural changes to meet the NDS to block Russia and China and restore America's competitive edge.

Originally developed for safety-critical domains, components of the IAP were adapted and integrated for DoD acquisition. It provides a lightweight but analytically robust process that supports early identification of innovation, structured stakeholder dialogue, and data-driven refinement of implementation strategies. By embedding systems thinking into both assessment and design phases, the IAP enhances an organization's ability to scale locally developed solutions and overcome systemic barriers to innovation.

The IAP was piloted in collaboration with the Air Force Installation Contracting Center (AFICC) in two implementation phases. Employing the IAP's novel Systemic Contributors and Adaptations Diagramming (SCAD) interview technique (Walker et al., 2016; Jefferies et al., 2022), a current state analysis of innovation culture was conducted with AFICC. SCAD analysis revealed critical patterns of systemic pressures—including regulatory constraints, policy demands, and organizational norms, which impact the translation of DoD leadership intentions into actionable behaviors. Findings from this analysis suggest that these pressures, more than individual differences between analysts, contribute to inhibiting or facilitating innovative behaviors. While certain system attributes, such as a willingness to embrace failure and support for organizational learning, can enhance innovative efforts, other pressures, such as stringent adherence to procedures and time constraints, inhibit innovation. The insights gained also highlight the importance of leadership support, aligning team goals, fostering internal and external collaboration, and granting autonomy to empower the acquisition workforce in overcoming barriers to innovation.

Another foundational element of the IAP is stress-testing innovations for scalability using the IMPActS framework (Fitzgerald, 2019) to design and revise interventions and address system attributes found critical to enabling successful innovative behaviors. The IMPActS framework captures the interdependencies of the Ideas behind the innovation (including its relationship to the systemic pressures noted above), Mental Model alignment, Pragmatics, availability of Actors, and resources to Sustain it. The framework was developed into a workshop to generate a lightweight, actionable method to deploy in practice (Balkin et al., 2024). Accelerating IMPActS workshops (AIWs) created a collaborative space for AFICC personnel to iterate on innovation interventions emerging from the SCAD interviews. AIWs are designed to increase stakeholder motivation to scale innovations, reduce the cost of risk-taking behaviors, while at the same time ensuring solutions are implementable and sustainable in the organization.

This research advances a practical approach for embedding innovation within the DoD by targeting the underlying systemic factors that shape individual behavior. Using novel tools (e.g., SCAD, AIW), the IAP creates a program for continuous monitoring, cultivating the conditions that enable innovation to take root and grow by supporting cultural transformation through sustained capacity building and organizational learning. The results presented provide a deeper understanding of persistent structural barriers in defense acquisition, as well as a model for strengthening and validating grassroots innovations for broader deployment. In doing so, it addresses a critical challenge in the DoD: How to ensure promising local innovations are not lost but instead matured and mobilized to create lasting, enterprise-wide impact.

Innovation Alliance Program

Developed by researchers at the Cognitive Systems Engineering Lab at The Ohio State University (OSU), the IAP is a tailored suite of techniques designed to elicit grounded insights into how individuals navigate complex, high-pressure work environments and create sustainable solutions. Originally applied in high-stakes safety domains such as commercial aviation and healthcare, the IAP has been adapted for use in DoD acquisition, where complexity and operational risk similarly demand a nuanced, systems-informed approaches to innovation. There are three core functions of the program, as illustrated in Figure 1:

1. A method for continuous monitoring to **identify** signals of barriers and facilitators to a healthy innovation culture using the lightweight SCAD interview technique.
2. A model and a tool to aid in the **interpretation** of the signals collected in the identification activities and target improvement efforts.
3. AIWs **implement** a codesign process to stress test high-potential innovations for supporting the transition of high potential ideas to improve their implementability and sustainability at increasing scale.

Figure 1. Innovation Alliance Program

SCAD

The SCAD interview technique is a novel approach to systems analysis, designed to reveal the often-hidden pressures and adaptive behaviors within complex organizational settings. Developed in response to the limitations of traditional event-driven methodologies like Root Cause Analysis (RCA) and Failure Modes and Effects Analysis (FMEA), SCAD focuses not on singular adverse events, but on the systemic pressures that shape day-to-day behavior across the organization (Jefferies et al., 2022). Rather than attributing deviation to individual factors or prescribing linear causal chains, SCAD collects and synthesizes frontline narratives that reveal how and why adaptations emerge in response to conflicting demands, resource constraints, and systemic tensions. This reorientation allows for generating a current state of an organization that is both diagnostic and prognostic—highlighting not only existing points of friction, but where future adaptations and breakdowns are likely to occur.

SCAD interviews are conducted by external researchers (e.g., OSU researchers) alongside internal domain experts (e.g., AFICC project leads). Interviews begin with a simple prompt: *Describe a time when work was done differently from the "textbook" approach.* Follow-up questions probe the systemic conditions that give rise to the adaptation. By treating each adaptation as a data point reflecting broader systemic dynamics, SCAD provides a high-resolution view into how people navigate their operational environment (Walker, 2021). The technique enables rapid insight generation at minimal cost and with limited training, proving especially effective in uncovering patterns of pressure and adaptation that would have remained opaque using conventional methods (Jefferies et al., 2022). This practical accessibility, paired with its systemic orientation, positions SCAD as a valuable tool for organizations seeking to move beyond a compliance mindset and toward transformation.

AIW

The AIW is a systems-based co-design method developed to improve the implementation and sustainment of organizational interventions in complex operational environments. It is grounded in the IMPActS Framework, which represents five key components essential for sustained implementation success: *Ideas*, *Mental Model Alignment*, *Pragmatics*, *Actors*, and *Sustainment* (Balkin et al., 2024; Fitzgerald, 2019). This framework is explicitly designed to address common pitfalls in intervention efforts, where promising initiatives experience early enthusiasm but ultimately fail due to misalignment across stakeholders, insufficient system capacity, or unaddressed contextual pressures (Fitzgerald, 2019; Nilsen & Bernhardsson, 2019). AIW provides a structured setting for stakeholders—including frontline personnel, managers, and researchers—to collaboratively assess candidate interventions against the five IMPActS dimensions, using both qualitative ratings and facilitated discussion to identify system constraints, clarify assumptions, and redesign interventions for scalability (Balkin et al., 2024).

A central goal of the workshop is to proactively uncover barriers to implementation before they lead to intervention failure. Participants are trained in systems thinking and guided through structured discussions that generate novel insights, align mental models, and sharpen intervention designs using shared institutional knowledge and adaptive systems research (Girth et al., 2022; Woods, 2018). In doing so, the workshop moves beyond theoretical frameworks to serve as a practical decision-support tool, bridging the gap between high-level design and ground-level execution (Balkin et al., 2024). AIWs increase the likelihood of sustained success by anticipating misalignments and incorporating implementation requirements into the intervention itself, offering a lightweight yet high-value process.

See–Do–Teach Model

Critical to the success of the IAP is the development of internal capacity by employing the See–Do–Teach approach to organizational learning. The See–Do–Teach model illustrated in Figure 2 is a progressive framework designed to build internal capability for executing and sustaining the IAP (i.e., conducting and analyzing SCAD interviews and facilitating AIWs). The model, which is applied across the three core competencies of the IAP—identifying (e.g., conducting interviews), interpreting (e.g., analyzing data), and implementing (e.g., leading workshops)—guides participants through three stages of skill acquisition: **See**, where individuals observe and receive foundational training from OSU researchers; **Do**, where individuals begin actively conducting and interpreting research activities with support and coaching from OSU researchers; and **Teach**, where individuals train others and facilitate activities independently. The scaffolding approach ensures a structured transition from novice to expert and anchors knowledge transfer by building internal capacity to execute the program.

Figure 2. Innovation Alliance Program: See–Do–Teach Model

Methodology

The IAP was implemented in two phases at AFICC. The initial pilot of the SCAD and AIW components included 15 SCAD interviews and one AIW. The second phase implemented the IAP and included 10 interviews and two AIWs; in total, 25 SCAD interviews and three AIWs were conducted with AFICC. Critical to the implementation of the IAP is collaboration between AFICC and OSU researchers. Embedded project leads at AFICC participated in weekly project meetings, identified participants for interviews and workshops, and participated in and ultimately conducted said interviews and workshops.

The OSU–AFICC team conducted semi-structured SCAD interviews with personnel at various levels, including leadership and the frontline workforce, illustrated in Table 1. Interviews were typically 1 hour in duration and were recorded and transcribed. Participant characteristics were de-identified to preserve confidentiality.

Table 1. SCAD Interview Participant Characteristics

Positions	Phase I	Phase II	Functions
Leadership (6)	Military (1) Civilian (4)	Military (0) Civilian (1)	Contracting (3) Program Management (3)
Frontline (19)	Military (5) Civilian (5)	Military (9) Civilian (0)	Contracting (19)

SCAD interviews began with a prompt to describe situations where the participant deviated from standard or “textbook” procedures, a hallmark of adaptive behavior. Probing questions then explore the contextual pressures, conflicts, and reasoning behind the deviation, enabling researchers to surface systemic dynamics that impede and facilitate innovative behavior. Employing the See–Do–Teach model, OSU researchers initially led interviews, training AFICC program leads as they observed the interview process, began participating in the interviews, and then led interviews at the end of the pilot period. Interviews were analyzed through iterative coding by trained researchers to identify patterns of pressure and adaptation.

Candidate innovations surfaced through the SCAD interviews, and promising innovations were identified by the AFICC project leads for AIWs. AIWs function as structured, participatory sessions designed to assess and refine candidate interventions prior to implementation. OSU researchers facilitated the workshops with participants drawn from various organizational levels at AFICC, including frontline personnel and supervisors. AIWs were 3 to 4 hours. The workshop opened with a brief training on the IMPActS framework and systems thinking. Participants then independently assessed proposed innovations using rating scales for each of the five IMPActS dimensions. These ratings are designed to elicit thoughtful judgments about an intervention’s feasibility, alignment, and sustainability within the organizational system. Following the individual assessments, participants engaged in facilitated group discussions to explore divergent perspectives, clarify assumptions, and collaboratively refine the intervention design. Two AIWs were conducted online and one in-person; thus, they were designed to be pragmatic, utilizing either physical materials or digital collaboration platforms (e.g, Miro).

Results

Innovation: Initiating, Sustaining, Spreading

The initial objective of the research team was to examine how acquisition innovation manifests within the organization and to trace how it evolves and spreads. Drawing from the AFICC SCAD interviews, acquisition innovation included adaptive responses to emerging constraints, creative problem-solving, and context-sensitive decision-making within the bounds of complex acquisition systems. Interview data showed that innovation is primarily sustained through the voluntary, discretionary efforts of individuals, often undertaken in parallel with their regular duties. Instances of innovative behavior were observed across various hierarchical levels and geographic locations, yet there was minimal dissemination of these efforts beyond the unit level, limiting opportunities for shared learning or scaling. Moreover, innovation tended to surface in response to acute, time-bound challenges, such as urgent mission demands, high-visibility projects, or situations requiring rapid remediation of issues related to schedule, cost, or quality.

Innovation was observed to progress through stages of *initiating, sustaining,* and *spreading* novel practices. As depicted in Figure 3, this trajectory begins with a state of “No Innovation,” wherein individuals operate strictly within the parameters of standard procedures. The earliest departure from this baseline is the “One Time” phase, in which a novel idea, process, or technology is introduced and used a single time but is not retained or institutionalized. This form of innovation is typically reactive and constrained by systemic barriers such as procedural rigidity, time scarcity, or lack of leadership support.

Advancing beyond these ad hoc efforts, the “Locally Adopted” phase represents the sustained integration of innovative practices within a specific unit or organizational context. In this stage, personnel consistently apply new tools, processes, or methods to enhance acquisition effectiveness. The most advanced stage, “Globally Adapted,” is marked by the successful identification, understanding, and transfer of these locally validated innovations to broader contexts across the organization. This stage requires mechanisms for recognizing emergent innovation, facilitating cross-unit learning, and reducing organizational friction. Each stage in this continuum is shaped by dynamic interactions between individual and institutional constraints, underscoring that innovation in DoD acquisition is less about isolated technological breakthroughs and more about the systemic conditions that enable novel practices to emerge, take root, and scale across the organization (Girth et al., 2022).

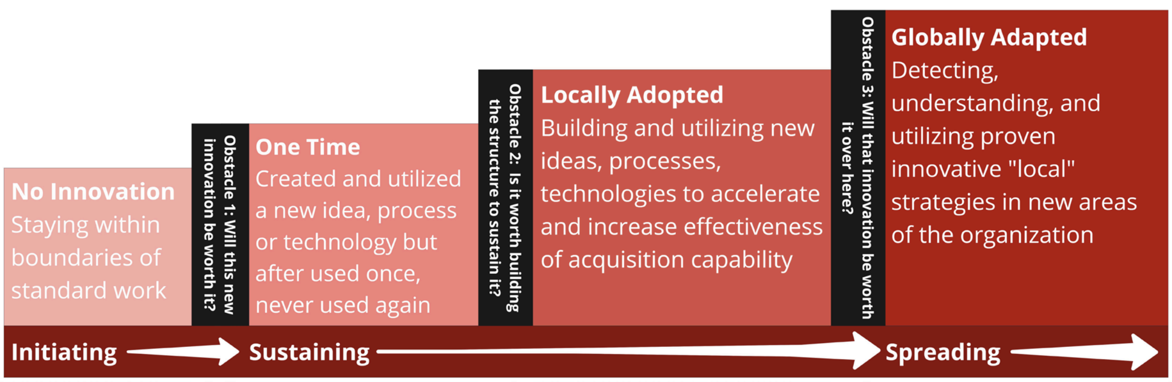


Figure 3. Innovation Continuum

The full value of innovation is only realized when its benefits are more broadly scaled. While localized or one-off innovations may offer immediate operational improvements, they fail to produce systemic change or address enterprise-level challenges if they remain siloed. This limitation reduces the return on investment in terms of time, labor, and institutional learning. When innovations are confined to specific units, other parts of the organization continue to experience inefficiencies or capability gaps that have already been solved elsewhere, resulting in duplication of effort, wasted resources, and lost opportunities for enhanced mission outcomes.

From an organizational perspective, the spread of innovation is essential to fostering institutional adaptability and resilience. When supported by leadership and institutional mechanisms, these grassroots innovations can be assessed for scalability and integrated more broadly across the enterprise. Programs such as SPARK[1] illustrate how the right structures can translate isolated successes into systemic improvements.

Without mechanisms for identifying, evaluating, and transferring effective innovations, the DoD risks perpetuating fragmented improvements rather than building coherent, enterprise-wide advancements. Spreading innovation also enables the standardization of best practices, creates shared language and understanding across teams, and builds a culture that normalizes and rewards adaptive behavior. In a strategic environment where speed, integration, and agility are imperative for outpacing near-peer adversaries like China and Russia, diffusion of innovation is not merely a desirable outcome—it is an operational necessity (Girth et al., 2022).

SCAD

Acquisition innovation is shaped by the interaction between **system attributes**—organizational conditions that enable innovative behavior—and **system pressures** that constrain it. Key system attributes that emerged from the AFICC pilot include *organizational learning, internal and external collaboration, goal alignment, autonomy,* and *making room for failure and risk-taking.* These elements create fertile ground for experimentation, reflection, and sustained change. However, these conditions are frequently moderated or suppressed by system pressures such as *excessive workload, time scarcity, procedural rigidity, low prioritization of innovation,* and *limited resources*. Among these, *leadership support* emerges as a particularly potent **compound system pressure**, in that it amplifies or mitigates the effects of other pressures depending on how it is exercised.

System attributes were frequently moderated by a range of system pressures, which could either enable or erode the organizational conditions necessary for innovation. The most frequently reported pressures included *procedure, time, innovation prioritization, reputation, workload*, and *organizational relationships*. For example, procedural constraints were found to have dual effects: in some contexts, flexible interpretation of rules encouraged creative approaches, while in others, rigid adherence stifled initiative and reinforced rote behavior. Similarly, high workloads and time scarcity constrained personnel's willingness to engage in experimentation, as the perceived cost of failure translated directly into additional burdens, effectively disincentivizing risk-taking behaviors.

Table 2. System Attributes That Support Innovation

Attribute name	Definition	Participant count (total per phase)	
		Phase I (15)	Phase II (10)
Creating room for failure and risk	Organization encourages risks and creative solutions without fear of punishment for trying something new	7	4
Organizational learning	Supports institutional learning, keeps people up-to-date on new tools and methods, and uses past situations as a source of information	5	5
Collaboration	Organization facilitates collaboration internally and externally with other departments and industry partners throughout a project lifespan	5	7
Goal alignment	People and groups (moving horizontally and vertically through the organization) share the same goal and understand their role in reaching the goal	5	4
Autonomy	Organization allows people to have flexibility and freedom to complete work through their own means, less leadership involvement, and more personal authority over projects	3	5

[1] For more information on SPARK, see https://afwerx.com/divisions/spark/overview/

The SCAD analysis further elaborated on how certain pressures interact with one another to shape innovation outcomes. One particularly illustrative case is leadership turnover, which surfaced as a compound pressure influencing both goal alignment and momentum. While turnover often disrupted ongoing initiatives and introduced misalignment between incoming leaders and existing innovation efforts, some cases revealed a more positive dynamic: when leadership transitions were managed through explicit bridging mechanisms—such as outgoing leaders designating continuity agents or codifying innovation as a team priority—the disruptive effects were significantly mitigated. These bridging practices functioned as organizational "throughlines," allowing promising innovations to persist beyond the tenure of their original champions.

Table 3. System Pressures That Impede Innovation

Pressure name	Definition	How strengthens (+)/weakens (-) system attributes	Participant count (total per phase)	
			Phase I (15)	Phase II (10)
Procedure	Policies, processes, and regulations can both enable change (if not explicitly prohibited) or stifle it through rigid adherence	Organizational learning (+/-): (+) Reducing the number of rules encouraged critical thinking and development of new skills (-) Following protocol, everything is a checklist rather than an evaluation of foundational skills and education Autonomy (+): Procedures that allow flexibility of execution encourages individualized solutions to problems Room for failure (-): Protocol provides a comfort zone that people fall back onto rather than attempting something risky	7	7
Time	The urgency to complete tasks quickly often reinforces the status quo, but crises or complex problems can accelerate creative problem-solving	Organizational learning (-): Desire to go fast leads to reliance on current/old procedures Collaboration (+): Need for results in a strict timeframe encourages collaboration and communication	6	8
Innovation prioritization	Reflects how an organization signals its commitment to innovation through resource allocation, messaging, policies, and support structures	Organizational learning (+/-): (+) Leads to developing critical thinking skills and seeking new information on improving current practices (-) Prioritizing innovation increases options, which can lead to an overwhelming amount of new information Goal alignment (-): The people working have a primary goal of getting work done, and if innovation is overly prioritized it gets in the way of that goal Room for failure (+): The desire to innovate allows more risks to be taken and boundaries to be pushed	4	5
Workload	When there is a mismatch between work demands and available resources, making it difficult for employees to support one another	Organizational learning (-): With high workload, additional dissemination and educational tasks are a burden and take a lower priority Room for failure(-): High workload decreases desire to take risks because a failed risk adds more work	3	5
Budget constraint	Lack of financial resources to execute	Goal alignment (-): Unknown budgetary restrictions disrupt ability to align intentions	3	4

	and impede ability to attract vendors			
Turnover	Particularly among enlisted personnel, disrupts momentum and can lead to employees delaying adoption of innovations in anticipation of leadership changes	Organizational learning (-): Rotating individuals through does not develop experts with a deep understanding of foundational skills Organizational learning (+): However, rotational programs increase the variety of perspectives and experiences the rotating individual gets to learn from and then take back to their team. Some practitioners also suggest this rotation and variety of perspectives (+) increases the willingness to try new ideas and take risks. Collaboration (-): Constant rotation of people does not support consistent collaboration Goal alignment (-): When people leave the project, it's hard to get a replacement with similar goals and enthusiasm about the project Goal alignment (+) When the originator of an innovation leaves the team, leadership or another team member acts as a throughline for an innovation, orchestrating the handoff and providing ongoing momentum.	3	4
Reliance on routines	Reinforces existing work habits, often making newer employees more open to change while longer-tenured members of the workforce may resist adopting new practices	Organizational learning (-): Becoming reliant on routine decreases the ability to embrace new information and processes Room for failure (-): People get attached to their way of doing things and create an environment that devalues trying new ideas	3	4
Political exposure	Increased scrutiny in public sector leads to risk aversion	Room for failure (-): Backlash and public scrutiny make people wary of attempting new ideas in the future	2	4
Reputation	Worrying about personal reputation if take risk and fail	Room for failure (-): Fear of damaging their reputation and hurting their career makes people less inclined to take risks and try new things	2	7
External events	Exogenous factors that force change	Organizational learning (+): External events push people to learn new ways of dealing with situations and can be applied to future scenarios	2	4
Organizational relationships	Quality of the relationship—at odds or strong working relationship	Collaboration (+/-): (+) Good relationships increase the likelihood for future collaboration (-) Strained relationships and lack of desire for communication decreases ability to collaborate	2	6

Leaders who provide "top cover" for experimentation, model openness to uncertainty, and align authority with responsibility foster autonomy and resilience in their teams. These leaders were consistently cited as enablers of innovation. Conversely, leaders who were perceived to lack alignment with team goals, failed to delegate authority, or signaled low tolerance for deviation from established norms (whether overtly or in their aggregate response to previous events) were seen as significant barriers to sustained innovation. The SCAD findings illustrate that leadership behavior is not merely a contextual variable but a central force in sustaining or suppressing innovation within the acquisition workforce.

Table 4. Leadership Support Pressures That Influence Innovation Behaviors

Leadership support	Definition	Incidence count (Phase II only)
Availability	Leaders are available/accessible to their team, encouraging them to find solutions but providing support when needed	3
Feedback	Getting more frequent feedback from leadership and customers creates opportunities to (a) realign goals across levels, (b) address and learn from issues, and (c) generate new insights and innovations	2
Openness	Leadership makes it "okay" not to know everything. They encourage people to ask questions and share knowledge to enable a culture of openness to learning. Leaders provide "top cover" for teams and individuals experimenting with innovative solutions	3
Bridging	When the originator of an innovation leaves the team, leadership or another team member acts as a throughline for an innovation, orchestrating the handoff and providing the ongoing momentum	4
Accounting for trade-offs	Goal alignment specifically on the risk versus reward trade-off is important to getting an innovation off the ground	3
Authority–responsibility alignment	Allowing people to have flexibility and freedom to complete work they are responsible for through their own means (i.e., more personal authority over work)	6
Goal alignment	One person in the right position of authority who does not share common goals can stop an innovation in its track	4
Incoming orientation toward innovation	A change in leadership greatly impacts the goals and innovation capability of the team: (+) New leaders who have a desire to innovate can create an environment that allows more risks to be taken and boundaries to be pushed (-) New leaders who prioritize status quo can halt previously developed innovations as new ideas	3

Taken together, these findings suggest that fostering innovation in DoD acquisition is not merely a matter of individual initiative or isolated process change. The SCAD methodology proves especially effective in uncovering these interactions, offering a diagnostic lens that provides a more nuanced understanding of how innovation behaviors are supported—or undermined—by the organizational environment.

AIW

Drawing on qualitative data from SCAD interviews and other innovation signals within AFICC, the research team employed the IMPActS framework—Ideas, Mental Model Alignment, Pragmatics, Actors, and Sustainment—to assess whether interventions are likely to be successfully implemented and sustained. Three AIWs were conducted across both pilot phases, demonstrating a structured and replicable process for evaluating and refining candidate interventions by integrating systems thinking at the design stage. Participants consistently emphasized the framework's utility in guiding design-phase discussions, promoting reflective dialogue, and surfacing latent implementation challenges. Participants reported that the IMPActS process enabled convergence of diverse perspectives, identification of potential barriers, and refinement of interventions in ways that would have been difficult to achieve through less structured methods.

One of the most salient themes to emerge from the AIWs was the emphasis on *aligning stakeholder mental models*. Participants described the structured discussion format as instrumental in clarifying assumptions and revealing misalignments in how interventions were understood across roles. This alignment process was viewed as foundational not only to the design of the intervention but to its long-term viability in a dynamic operational environment. Another key insight was the distinction between initiating and sustaining innovation. Participants found the *sustainment* dimension particularly valuable, as it highlighted anticipating downstream resource demands, institutional support structures, and long-term stakeholder engagement. This concern aligns with broader implementation science findings emphasizing that the mechanisms enabling adoption are often distinct from those required for ongoing execution (Lewis et al., 2022; Proctor et al., 2011).

The first AIW revealed the intrinsic value of the framework itself. Participants described it as “lightweight” yet “high value,” capable of catalyzing strategic dialogue in environments with limited bandwidth and high operational tempo. Feedback emphasized that the process would not be difficult to justify to leadership—an essential factor for adoption within hierarchical organizations like the DoD. The second two workshops added a complementary component focused on identifying practical mitigations and mobilizing resources in response to assessment results. This evolution reflects the recognition that assessment alone is insufficient; successful implementation also requires structured planning for action and coordination. Importantly, the two later workshops explored sourcing interventions not only from SCAD interviews but also from emergent opportunities within innovation communities, such as the Innovation Rodeos and the internal AFICC Crosstalk meetings for innovation sharing.[2]

Ultimately, participants reported that the AIW gave them a sense of ownership and agency over the redesign process and greater confidence in the proposed interventions. The AIW structure successfully supported cross-role collaboration, clarified implementation trade-offs, and enhanced foresight around systems-level risks and requirements. These findings underscore the framework’s value as both a diagnostic and generative tool, capable of improving intervention quality and increasing the likelihood of scalable, enterprise-wide adoption.

Discussion

This study was designed to pilot the IAP, enhancing AFICC’s innovation culture with practical, scalable tools that allow for monitoring, supporting, and propagating innovative practices throughout its workforce. Rather than viewing innovation as a sporadic or personality-driven endeavor, this research used a systems approach to identify and respond to emerging barriers or enablers in real time. The methodology supports both top-down strategy and bottom-up experimentation, fostering an environment where promising ideas can be adapted and scaled across varying operational levels.

AFICC has a collaborative, capable, and motivated workforce in the innovation ecosystem. However, the analysis reveals several persistent challenges that could limit the sustainment and spread of innovative behaviors. Chief among these is the issue of time and workload. Although participants express strong interest in contributing to innovation initiatives, they consistently report a lack of capacity to support implementation efforts beyond initial engagement. Initiatives demanding sustained participation or project ownership often falter due to competing responsibilities and limited bandwidth. The lack of clearly allocated time and responsibility for innovation advocates constrains the effectiveness of the organization, whereas institutionalization of the IAP can create innovation leaders throughout the organization to redistribute workload and explicitly assign authority and responsibility. Ironically, innovative

[2] For more information on Innovation Rodeo, see https://www.afimsc.af.mil/innovationrodeo/

activities that could result in better long-term time and resource management are deprioritized or abandoned because of short-term time and resource demands.

In addition to time constraints, innovation at scale requires adequate financial and structural support. Participants highlight the difficulty of preparing grassroots innovations—often developed with minimal resources—for wider adoption. Transitioning from a local pilot to an enterprise-level solution frequently necessitates new materials, expanded technical infrastructure, or additional staffing, few of which are readily available under current resourcing models. The cross-functional nature of many high-value innovations introduces ambiguity about who holds the authority and the budget to support broader implementation. Without a formalized mechanism for cross-cutting resource allocation, these innovations risk stagnation at the local level, unable to realize their full potential within the enterprise.

Fostering a culture of innovation—one that encourages calculated risk-taking, supports organizational learning, promotes alignment across teams, and grants autonomy—is essential for sustainable transformation. Leaders play a pivotal role in cultivating this environment by providing top cover for experimentation, aligning incentives, and acting as champions for promising innovations. The data from this study underscores that innovation is not simply a function of individual creativity or motivation, but a product of deliberate organizational design and leadership commitment.

Conclusion

This study demonstrates that in everyday work there are high-potential grassroots innovation interventions being developed by the acquisition workforce at a local level to quickly and effectively solve problems encountered. However, to benefit from and amplify this innovative behavior, the organization needs to (1) assess innovations as they emerge to determine their scalability and overcome or mitigate systemic barriers that hinder broader adoption and (2) dynamically prioritize those that it will foster and proliferate, and (3) provide resources to support the transition from local to global interventions.

By addressing the identified systemic contributors, the IAP aims to promote lasting behavior changes that support a culture of innovation within the DoD through continuous monitoring and capacity building using our novel SCAD and IMPActS tools. This research not only contributes to the understanding of the systemic barriers to innovation in DoD acquisition but also offers actionable recommendations for developing an adaptive, resilient workforce. It provides a lightweight, adaptable model for stress testing and strengthening innovations for larger-scale deployment across the DoD, addressing the issue of micro-innovation that is high-potential but fails to launch beyond an individual or localized unit.

In sum, acquisition innovation is not optional; rather, it is foundational to modernizing defense acquisition. By using the IAP to identify and enable the conditions that support innovative behaviors, leadership can ensure that the acquisition enterprise is capable of accelerating change when it matters most.

Disclaimers

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD[A&S]) and the Office of the Under Secretary of Defense for Research and Engineering (OUSD[R&E]) under Contract HQ0034-19-D-0003, TO#0285, and TO#0309.

The views, findings, conclusions, and recommendations expressed in this material are solely those of the authors and do not necessarily reflect the views or positions of the U.S. government (including the Department of Defense [DoD] and any government personnel), the Stevens Institute of Technology, or The Ohio State University.

References

Balkin, E. M., Maguire, L. M. D., Girth, A. M., & Rayo, M. F. (2024). Leveraging systemic contributors and adaptations to sharpen interventions: The IMPActS Workshop. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *68*(1), 627–633. https://doi.org/10.1177/10711813241269240

De Vries, H., Bekkers, V., & Tummers, L. (2016). Innovation in the public sector: A systematic review and future research agenda. *Public Administration*, *94*(1), 146–166. https://doi.org/10.1111/padm.12209

Demircioglu, M. A., & Audretsch, D. B. (2017). Conditions for innovation in public sector organizations. *Research Policy*, *46*(9), 1681–1691. https://doi.org/10.1016/j.respol.2017.08.004

Demircioglu, M. A., & Audretsch, D. B. (2024). What is public sector innovation? In *Public Sector Innovation* (pp. 10–22). Cambridge University Press. https://doi.org/10.1017/9781009279277.002

DoD. (2022). *2022 National Defense Strategy of the United States.* https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.pdf

Fitzgerald, M. C. (2019). *The IMPActS Framework: The necessary requirements for making science-based organizational impact* [Doctoral dissertation, The Ohio State University]. OhioLINK Electronic Theses and Dissertations Center.

Girth, A., Rayo, M., Newton, E., Balkin, E., Jefferies, C., & Tewani, P. (2022). *Systemic factors influencing risk aversion: Diagnosing behaviors and tailoring interventions for lasting transformation* (Final Technical Report SERC-2022-TR-007). Acquisition Innovation Research Center. https://acqirc.org/publications/research/systemic-factors-influencing-risk-aversion-diagnosing-behaviors-and-tailoring-interventions-for-lasting-transformation/

Jefferies, C. M., Balkin, E. A., Groom, L., & Rayo, M. F. (2022). Developing Systemic Contributors and Adaptations Diagramming (SCAD): Systemic insights, multiple pragmatic implementations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *66*(1), 75–79. https://doi.org/10.1177/1071181322661334

Kettl, D. F. (2015). *The transformation of governance: Public administration for the twenty-first century*. Johns Hopkins University Press.

Lewis, C. C., Boyd, M. R., Walsh-Bailey, C., Lyon, A. R., Beidas, R. S., Mittman, B., Aarons, G. A., Weiner, B. J., & Chambers, D. A. (2022). A systematic review of empirical studies examining mechanisms of implementation in health. *Implementation Science*, *17*(1), 1–20. https://doi.org/10.1186/s13012-021-01181-3

Moynihan, D. P., & Landuyt, N. (2009). Explaining turnover intention in state government: Examining the roles of gender, life cycle, and job characteristics. *Review of Public Personnel Administration*, *29*(2), 111–134.

Nilsen, P., & Bernhardsson, S. (2019). Context matters in implementation science: A scoping review of determinant frameworks that describe contextual determinants for implementation outcomes. *BMC Health Services Research*, *19*(1), 189. https://doi.org/10.1186/s12913-019-4015-3

Osborne, S. P., & Brown, L. (2011). Innovation, public policy and public services delivery in the UK: The word that would be king? *Public Administration*, 89(4), 1335–1350. https://doi.org/10.1111/j.1467-9299.2011.01932.x

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Rayo, M. F., Girth, A. M., & Maguire, L. (2024). *Systemic factors influencing risk aversion: Diagnosing behaviors and tailoring interventions for lasting transformation* (WRT-1081.8.4-v1.2). Acquisition Innovation Research Center. https://acqirc.org/publications/research/building-acquisition-innovation-capacity-phase-ii-final-report/

Walker, K. E. (2021). *Using a corpus of accidents to reveal adaptive patterns that threaten safety* [Doctoral dissertation, The Ohio State University]. OhioLINK Electronic Theses and Dissertations Center.

Walker, R. M. (2006). Innovation type and diffusion: An empirical analysis of local government. *Public Administration*, *84*(2), 311–335. https://doi.org/10.1111/j.1467-9299.2006.00004.x

Woods, D. D. (2018). The theory of graceful extensibility: Basic rules that govern adaptive systems. *Environment Systems and Decisions*, *38*(4), 433–457. https://doi.org/10.1007/s10669-018-9708-3

PANEL 11. MODERNIZING DEFENSE ACQUISITION: MODULAR SYSTEMS, AI, AND SOFTWARE COMPLIANCE

Wednesday, May 7, 2025	
1240 – 1355 PT 1440 – 1555 CT 1540 – 1655 ET	**Chair: Van Hendrey, Executive Director Program Executive Office, Integrated Warfare Systems Naval Sea Systems Command** ***DoD Needs Better Planning to Attain Benefits of Modular Open Systems*** Nate Vaught, Senior Defense Analyst, U.S. Government Accountability Office ***Accelerating Software Acquisition Using Generative AI for Regulatory Compliance*** John Robert, Deputy Director, Software Solutions Division, Carnegie Mellon University Software Engineering Institute ***Accelerating the Future Leveraging AI for the Transformative Federal Acquisition*** Adam Bouffard, Principal Contract Analyst, MITRE Corporation

Ms. Van Hendrey—was appointed to the Senior Executive Service and began serving as the Executive Director, Program Executive Office for Integrated Warfare Systems (PEO IWS) in September 2023. In this role, Ms. Hendrey directs the acquisition and Fleet support of the Surface Navy's combat systems, weapons, radars, and related international and foreign military sales programs. She is responsible for an organization of over 400 civilian and military personnel and 170 programs and projects with an annual acquisition budget of over $6 billion.

Prior to her current role, Ms. Hendrey was appointed to the Defense Intelligence Senior Executive Service in August 2020 and served as the Program Manager for the Defense Intelligence Agency's flagship Machine-Assisted Analytic Rapid-Repository System (MARS) program. As the MARS Program Manager, she led the comprehensive digital modernization of the Defense Intelligence Enterprise's system for foundational military intelligence.

Prior to her selection to SES, Ms. Hendrey served in a number of senior leadership positions to include: Deputy Program Manager, Mine Warfare (PMS 495) in PEO Unmanned and Small Combatants; Director of Undersea Warfare Systems and Littoral Combat Ship for the Assistant Secretary of the Navy, Research, Development and Acquisition (ASN RD&A); Principal Assistant Program Manager (PAPM) for Advanced Radars (IWS 2A) in PEO IWS; and PAPM and Chief Engineer for the Cobra Judy Replacement program in PEO IWS. She spent 11 years in industry working on PEO IWS, PEO C4I and Army programs where she gained invaluable insights into the perspectives of small and medium sized businesses working on Government contracts. Ms. Hendrey's Navy career began as a high school intern in the Magnetic Materials Group, Naval Surface Warfare Center White Oak.

Ms. Hendrey is a Vietnamese refugee whose family immigrated to the United States after being sponsored by an Air Force family. She holds a Bachelor of Science degree in Physics and a Master of Science degree in Electrical Engineering, both from the University of Maryland, College Park. Ms. Hendrey is a member of the Acquisition Professional Community and has a Level III Certification in Program Management. She was the winner of the USD(AT&L) 2014 Workforce Achievement Award for Program Management.

Ms. Hendrey's awards also include two Navy Meritorious Civilian Service Awards, a Navy Meritorious Unit Commendation, the Office of the Director for National Intelligence Exceptional Service Award, the Defense Intelligence Agency Director's Award, and the Under Secretary of Defense (Intelligence and Security) Visionary Award. She resides with her husband and two children in Arlington, Va.

DoD Needs Better Planning to Attain Benefits of Modular Open Systems

(Executive Summary)

Nathaniel Vaught—is a Senior Defense Analyst with the U.S. Government Accountability Office, specializing in weapon system acquisitions. He has worked in this issue area for more than 10 years, reporting and making recommendations on a variety of systems including the KC-46A tanker aircraft, F-35 Joint Strike Fighter, Columbia class submarine, and Virginia class submarine. [VaughtN@gao.gov]

Abstract

A modular open systems approach (MOSA) is a business and technical strategy that can help the Department of Defense (DOD) design weapon systems that take less time and money to maintain and upgrade. Recent legislation requires acquisition programs to implement a MOSA to the maximum extent practicable. The benefits of designing weapon systems with a MOSA have been long established. However, GAO reported as recently as 2023 that implementation of a MOSA in acquisition programs was inconsistent. For its current report, anticipated to be published in January 2025, GAO reviewed the use of a MOSA in 20 selected acquisition programs and assessed policies and processes at the Office of the Secretary of Defense and military departments.

Background

Legislation enacted over the past several years required DoD to change the way it buys and designs weapon systems by implementing a modular open systems approach (MOSA) to the maximum extent practicable. A MOSA, which includes a modular design and standard interfaces, allows programs to easily replace components of a product. This approach allows the product to be competitively upgraded with new, improved components that can be made by a greater variety of suppliers. It may also help address concerns we have previously reported on about rising sustainment costs by increasing competition for sustainment among potential vendors. Otherwise, these costs may limit DoD's ability to afford the force structure it expects to need in future conflicts.

Objectives, Scope and Methodology

This report assesses the extent to which (1) programs implemented MOSAs and why; (2) programs and portfolios planned for MOSAs; (3) the military departments invested in necessary MOSA resources; and (4) DoD developed MOSA policy, regulations, and guidance. GAO reviewed planning documents for 20 acquisition programs that started after relevant laws were passed in 2016. GAO selected the programs based on their acquisition approach and military service. GAO also reviewed policy and guidance documents and interviewed DOD officials

Summary

GAO found that 14 of the 20 programs reviewed reported implementing a MOSA to at least some extent. Other programs cited barriers to doing so, such as added cost and time to conduct related design work. While a MOSA has potential benefits, it may also require programs to conduct additional planning, such as to ensure they address cybersecurity concerns. However, none of these 20 programs conducted a formal analysis of costs and benefits for a MOSA because DoD's policy does not explicitly require one. As GAO reported in March 2020, program officials often focus on reducing acquisition time and costs. Unless required to consider the costs and benefits of a MOSA, officials may overlook long-term MOSA benefits.

Most programs did not address all key MOSA planning elements in acquisition documents, in part, because the military departments did not take effective steps to ensure they did so. As a result, programs may not be well-positioned to integrate a MOSA into early key investment decisions. Also, DoD's process for coordinating MOSAs across portfolios does not ensure the level of collaboration needed to achieve potential benefits such as lower costs from using common components across programs.

The military departments are statutorily required to ensure availability of certain resources and expertise related to MOSA implementation. However, they have yet to assess their departments' MOSA needs or determine how resources should be aligned across their respective departments. Until they do this, programs risk having insufficient resources and expertise to achieve the potential benefits of a MOSA.

DoD has updated some acquisition and engineering policies and is drafting regulations and guidance to address MOSAs. But gaps remain that could hinder MOSA implementation. For example, DoD policy does not address how MOSA requirements apply to programs using the middle tier of acquisition pathway those intending to complete rapid prototyping or fielding in 5 years or less.

Full Report: https://www.gao.gov/products/gao-25-106931

Accelerating Software Acquisition Using Generative AI for Regulatory Compliance

John Robert—is a Principal Engineer at the Carnegie Mellon University Software Engineering Institute (SEI) and currently serves as the Deputy Director of the Software Solutions Division. In this role, Robert leads software engineering research and development efforts in partnership with programs in the Department of Defense and industry. Robert is a co-author of Architecting the Future of Software Engineering: A National Agenda for Software Engineering Research & Development. Robert holds a Master of Software Engineering from CMU and a BS in Electrical Engineering from West Virginia University.

Carlos Olea—is a Graduate Research Assistant at the Magnum Research lab, Vanderbilt University. His work focuses on interdisciplinary AI evaluation and utilization, with applications in fields from cyber-physical security to aerospace design to sports analytics.

Yash Hindka—graduated with a BS in Computer Science from the University of Wisconsin-Madison in December 2021 with Highest Distinction. After graduating, he joined Raytheon as a software engineer, driven by the warfighters he grew up admiring from a young age. Yash currently works at the Software Engineering Institute (SEI) at Carnegie Mellon University (CMU). He joined the SEI to learn from and be a part of the strong academic tradition at CMU, all while continuing to support the warfighter. At the SEI, he analyzes embedded weapon system software for vulnerabilities and participates in research efforts to increase automation in the analysis process.

Nanette Brown—is a Senior Member of the Technical Staff at the Software Engineering Institute. In that capacity she has worked as a researcher in the areas of agile and architecture and technical debt management. She has consulted with government agencies on Agile and Lean Practices, and Operational Test practices and policy development. Prior to joining the SEI, Brown spent more than 20 years in product development, starting as a developer and continuing on to lead development organizations and change initiatives as the executive director of Architecture and Quality Management of a Fortune 500 company.

Douglas C. Schmidt—is the Dean of Computing, Data Sciences & Physics at William & Mary. Schmidt has served the president-appointed and Senate-confirmed Director of Operational Test and Evaluation, where he was responsible for overseeing the evaluation of the operational effectiveness, suitability, survivability, and (when necessary) lethality of United States defense systems to defend the homeland and prevail in conflict. He also served as the Chief Technology Officer at the Carnegie Mellon University Software Engineering Institute (SEI).

Abstract

Detecting document incompleteness, inconsistencies, and discrepancies between regulatory documents and software artifacts is a common and people-intensive task for acquisition teams. Department of Defense (DoD) Acquisition environments have extensive documentation describing policies, guidance, and standards that must be repeatedly compared to delivered software artifacts for a DoD program to ensure regulatory conformance throughout a project's life cycle. Acquisition professionals in these environments must learn the extensive and complex regulatory information, apply the knowledge to multiple projects, and identify document incompleteness, inconsistencies, and discrepancies (DIID) that could indicate non-compliance or high-risk areas. Currently, teams of people review multitudes of documents and data, reading and using general search on keywords to find relevant text to review and compare to regulatory documents. As the DoD continues moves toward DevSecOps with continuous integration and rapid capability deployment approaches, people-intensive approaches to ensure regulatory compliance are slow, do not scale, and delay mission capability.

This paper investigates the use of large language models (LLMs) to improve the efficiency and accuracy of DIID detection while enabling customization through prompt engineering. The proposed approach leverages LLMs to augment acquisition professionals by providing semi-automated and meaningful connections of software artifacts to regulatory documents. Testing approaches are proposed to assess the effectiveness of LLMs for DIID detection, and preliminary results are provided for detecting DIID with augmented LLMs. This paper also proposes prompt engineering approaches for DIID detection and suggests benefits for DIID detection in software acquisition activities.

Introduction

Department of Defense (DoD) software acquisition programs must adhere to a complex web of regulations, standards, and contractual requirements. Ensuring every requirement is addressed involves producing and reviewing extensive documentation, such as requirements, design specifications, and test plans. Currently, teams of analysts manually cross-check these artifacts against DoD policy checklists, derived from DoD directives or standards, to identify any section that is incomplete, inconsistent, or discrepant with respect to the official guidance. This manual process is tedious, error-prone, and people-intensive, so important omissions or contradictions can be overlooked due to sheer volume or reviewer fatigue. Moreover, as the DoD increasingly adopts DevSecOps practices to enable continuous integration and rapid deployment, human-centric compliance reviews struggle to keep pace, introducing delays in delivering mission capabilities.

Automating regulatory compliance checks has become an area of great interest to improve efficiency and reduce human error. A common regulatory compliance task is identifying where software acquisition or engineering artifacts are inconsistent, incomplete, or discrepant from regulatory or standards documents. Such document incompleteness, inconsistencies, and discrepancies (DIID) could indicate a regulatory requirement is not addressed or does not apply to a system, or it could be a simple omission. Software acquisition and engineering teams often begin regulatory or milestone reviews searching for DIID because they can be an indicator of potential risks or deviations from expected system performance. Currently, acquisition and engineering teams are using blunt word searches or lengthy reading and review to find DIID, which is both slow and error-prone because documents can use different words for similar topics.

Our goal is to augment acquisition professionals by automating the tedious detection of DIID issues, while human experts remain central to interpreting and resolving these issues. Determining the applicability generative AI to specific use cases, including software acquisition, requires assessing the opportunities and risks among multiple considerations for their contexts (Bellomo et al., 2023). The risks and opportunities of applying large language models (LLMs) to DIID detection are discussed throughout this paper, along with some of the strengths of LLMs and understand the risks. This paper does not explore the question of AI-based recommended resolutions to detected DIID because automating DIID resolutions has higher risks given current AI capabilities.

Not every document discrepancy or omission is critical because some may be contextually justified. Identifying all potential DIID issues is a necessary first step before human judgment can determine their severity. Automated support to augment humans in this task could significantly improve efficiency. However, any automation to detect DIID, particularly in mission- or safety-critical environments, must keep humans as an integral part of the process and as the ultimate decision makers.

There is a pressing need for tools that can rapidly and reliably detect DIID across large collections of acquisition documents. This paper describes how we are assessing the use of LLMs to find DIID in software engineering artifacts efficiently by cross-referencing them with regulatory documents. We also outline prompt engineering techniques to tailor LLMs for this task and ensure they work effectively alongside humans. Finally, we describe gaps in testing and evaluating the LLM's performance on DIID detection and discuss preliminary findings. Our approach illustrates initial observations and proposes approaches to improve scalability and accuracy in compliance checking, ultimately accelerating software acquisition while maintaining confidence in regulatory conformance.

New Opportunities in AI-Augmented Regulatory Compliance

Recent advances in generative AI, including LLMs, offer a unique opportunity to transform software engineering and software acquisition (Robert et al., 2024). This transformation extends to automating and accelerating software regulatory compliance workflows. Unlike keyword-based search tools, LLMs can consider natural language context and make judgements about document content based on training and other data. The ability of LLMs to interpret semantics makes them suitable for text comparisons (Brown et al., 2020). Using LLMs for document analysis is a common topic of interest across many domains, including healthcare (Moilanen et al., 2022; Shokrollahi et al., 2023), financial or business applications (Cao et al., 2024; Shukla et al., 2023), or for analyzing legal documents (Prasad et al., 2024). In most of these domains, LLMs are used for document summaries and to provide insights into data for that domain.

LLMs like GPT-4, Gemini, and Ollama can analyze documents, respond to prompts, or generate natural language from unstructured text. LLM abilities to consider natural language context has generated interest in automation of document heavy tasks. For example, an LLM can interpret a requirement like "The system shall implement encryption in accordance with FIPS 140-2" and check that the design document specifies FIPS 140-2 compliant encryption beyond a simple word search of the document to include some semantic understanding.

By embedding the relevant regulations and policy documents into a vector database and using it to provide context, an LLM can be prompted to cross-reference an acquisition document against applicable rules. This approach, known as retrieval augmented generation (RAG), allows an LLM to find explicit regulatory clauses or past examples from the vector store to confirm its compliance checks (Nextra, n.d.). Using RAG to augment LLMs is currently part of many online services offered by many LLMs and can be implemented on stand-alone LLM solutions.

LLMs also offer the opportunity for people to interact with the documents and information in new ways. Prompt engineering (Liu et al., 2023) and prompt patterns (White et al., 2023) are techniques humans use to create and refine inputs to optimize responses from a generative AI model, but they also represent new opportunities for humans to interact with AI systems (Bozkurt, 2024). For software acquisition and engineering teams, prompt engineering and prompt patterns provide a flexible and natural way to customize and refine DIID detection.

Role of DIID in Regulatory Risk and Non-Compliance

DIID issues in DoD program documents are not just theoretical inconveniences. In practice they have been linked to project risks and delays. For example, (Brownsword, 2012) identified DoD program anti-patterns that contribute to cost overruns and schedule delays. Several of these patterns are examples of discrepancies or incompleteness in a DoD program's acquisition artifacts. A more recent GAO report identified that several programs failed to conduct or report cybersecurity testing phases despite a DoD policy requiring the tests (GAO, 2022). These findings explain why the DoD mandates rigorous document reviews and checklist-driven inspections because even minor omissions can indicate compliance violations or additional risk if left undetected.

Despite these precautions, human reviewers struggle with information overload. Critical DIID issues can be missed simply because they are buried in hundreds of pages of specifications and plans. The more complex the program, the more likely an overwhelmed analyst might overlook a subtle inconsistency, which is why automated DIID detection assistants can prove invaluable. Such AI-powered assistants (e.g., an LLM scrutinizing the documents in parallel) act as a second pair of eyes, scanning at scale and flagging potential issues that warrant human attention. By catching contradictions and gaps that a single reviewer might miss, these tools can enhance overall compliance assurance.

Figure 1. Role of DIID in Software Acquisition

Figure 1 visualizes the role of DIID in software acquisition. Multiple documents are created and connected as part of a program life cycle, and compliance with DoD regulations is both required and verified. However, DIID examples exist in many documents and remain unidentified due to the limitations in human reviewers and result in delays and other consequences. An AI-powered DIID detection assistant helps identify DIID and prevent delays or other issues.

It is important to note that the findings of AI tools would feed into the human review process, not replace it. The goal is to improve review effectiveness and reduce risk while allowing human experts to focus on judgment calls and solutions. In short, improving DIID detection through automation reduces risk in acquisition programs by increasing the likelihood that potential issues are found, and humans can review and analyze these issues.

DIID issues can surface at many stages of the software acquisition or software development life cycle. DoD policies drive some requirements but also contribute to acquisition strategy, architecture, and testing. When DoD policies recommended or require processes or standards, these must be reflected in the relevant program documents. Some examples are provided in Table 1.

Table 1. DIID Examples by Life Cycle

Lifecycle Stage	DIID Examples
Acquisition Strategy	Goals conflicting with
Requirements	Missing safety constraints or security clauses from DoD standards or policies
Design Documents	Conflicting interface definitions
Test Plans/Reports	Incomplete traceability to requirements
Certification Artifacts	Discrepant terminology from regulatory standards

Defining DIID in the Acquisition Life Cycle

DIID all indicate potential problems in software acquisition artifacts, but each term has a distinct meaning:

- *Incompleteness.* Important content is missing either within a document or across a set of documents. For example, if a safety policy mandates a "detailed analysis of safety-critical signals" in the architecture description, but the project's Software Architecture Document only provides an analysis of "signals" without the safety-critical qualifier, this omission constitutes an incompleteness. Incompleteness typically means some requirement or context that should be present is absent, potentially necessitating additional review or rework to address the gap.
- *Inconsistency.* There are contradictions or a lack of uniformity in terminology or content either within one document or between documents. For instance, using "safety" and "security" interchangeably in a system description (when in some domains they imply different requirements) can confuse readers about the true requirements. Inconsistencies might not always signify an error because they can be acceptable in the context, but they often create confusion and may hint at deeper incompleteness or misunderstandings.
- *Discrepancy.* There is a direct conflict or divergence between facts or statements in one place versus another. Discrepancies can be factual (e.g., one document states a response time requirement as 0.01s while another cites 0.1s for the same event, a tenfold difference), policy-related (a process described in the project plan deviates from what a governing policy mandates), narrative (two documents describing the system's behavior tell conflicting stories), or theoretical (project results or assumptions conflict with established theory or prior data). Discrepancies often indicate more serious issues because one of the conflicting statements could be wrong. If critical parameters differ, the impact can be severe.

These definitions are typical of what is found in dictionaries and literature on human interpretation or document analysis. Although there is no comprehensive list of different types of DIID, common types and examples are summarized in Table 2.

Table 2. Types of Incompleteness, Inconsistencies, and Discrepancies

Incompleteness	Inconsistency	Discrepancy
• **Incomplete**: Important context or terms are missing.	• **Terminology**: Using different terms interchangeably without clear definitions or consistency. • **Structural**: Lack of uniform structure in presenting information.	• **Factual discrepancies**: Conflicting factual information. • **Policy or procedural discrepancies**: Deviations from established protocols. • **Narrative discrepancies**: Different user stories fail to align. • **Theoretical discrepancies**: Actual results conflict with theoretical predictions.

Exploring DIID for Software Safety in DoD Regulatory Compliance

Exploring specific examples provides important insights into the opportunities and current limitations of automating DIID detection in DoD regulatory use cases. We have performed initial tests using currently available online LLMs, such as OpenAI ChatGPT 4o and

Google Gemini 2.0 as well as locally hosted LLMs, such as Ollama, to perform simple DIID comparisons. Although testing is in progress, our initial results provide important insights into opportunities and risks to enable generative AI for regulatory compliance. Some of the testing insights are shared in blogs or webinars, to provide general observations while in progress (Robert & Schmidt, 2024; Schmidt & Robert, 2024).

The initial testing has been performed on public (Distro-A) data comparing DoD and NASA software safety standards to text statements that represent different types of DIID. There are currently more than 100 statements to help identify false positives, the majority of which do not contain DIID. There *are* examples of DIID, although not all sub types of discrepancies are represented at this time.

A common use case is when DoD programs must follow system or software safety standards to identify safety requirements or risks. For example, MIL-STD-882E is used for hazards analysis, and Safety Assurance Cases (*MIL-STD-882E*, 2023), which must align with and program-specific safety policies. The initial testing of DIID detecting used DoD standards, including or similar to MIL-STE-882E, to compare with different statements that do or do not contain DIID when compared to the standard. Inconsistencies could include mismatch in system functions between FHA and test procedures or missing causal chains in hazard analyses.

Using this DoD standard, we illustrate DIID examples in Table 3 that might be found in requirements documents, design documents, or testing documents. These are just a few of the many examples of possible DIID, which can be in any of the DIID types, across many documents, at different parts of the software development lifecycle, and for difference software releases over time.

Table 3. DIID Examples

DIID Type	DIID Example	DIID Justification
Incompleteness	The software test report includes results from functional and integration testing.	The DoD standard may require explicit evidence of software safety testing, including corner cases or hazardous conditions. The absence of test cases for safety-critical scenarios indicates incompleteness in test coverage.
Inconsistency	“Cyber resiliency is ensured through encryption and secure boot mechanisms,” (in a System Description) vs. “Safety-critical software must handle fail-safe transitions during fault events for resiliency”	If the term “resiliency” is used in both contexts (cybersecurity and safety) but without consistent definitions, and without clarifying boundaries, it introduces inconsistency.
Discrepancy	“The system shall complete data processing within 5 milliseconds,” (in Requirements Specification) vs. “The implemented architecture meets the 50 millisecond timing constraint for processing.” (in Design Document)	There is a factual discrepancy in performance in timing requirements with a 10x difference between what is required and what is implemented or claimed, possibly leading to a violation of real-time constraints mandated in mission-critical DoD systems.

The examples are a small sample of the many types of DIID tests. The statements are simple, but each statement must be reviewed by people, sometimes acquisition experts, to confirm that the tests are relevant DIID that human reviewers would find interesting or consider investigating in more detail. In addition, examples of statements without DIID, as perceived by human reviewers, are also needed to consider false positives. The number of tests needed to really exercise LLM solutions for DIID detection is extensive, particularly to consider all types of regulatory documents and software artifacts. Work is currently underway to further characterize DIID testing and gaps.

An important comment from software acquisition professionals and software engineers is that DIID could appear to be limited when compared to software artifacts, such as design or testing documents, but DIID could be extensive when compared to software code for that same system. It is common to have software plans or software artifacts claim performance or processes, only to review code or detailed testing to discover that the implementation is not consistent with the documentation. This known issue requires extension of DIID detection of regulatory documents beyond software artifacts to code. Commercial software development and analysis tools are already integrating LLMs into their workflows, and these solutions can be tested for DIID detection. In addition, the DIID detection for software artifacts remains important for software acquisition teams and this may be the earliest opportunities to automate some of the regulatory burden on acquisition teams. We strongly agree that DIID solutions must extend to code, but in practice today we frequently observe programs only reviewing software artifacts.

Figure 2 illustrates a notional DIID detection testing pipeline. One or more regulatory standards, DoD or other government standards, are used (top left) to create public DIID detection test data sets. The test data is created and/or reviewed by humans for accuracy and used to test multiple DIID detection architectures, including online LLMs, stand-alone (locally hosted) LLMs, or LLMs combined with RAG solutions. The test results are summarized to understand which solutions provide the best DIID detection, categorized by DIID type and other groupings.

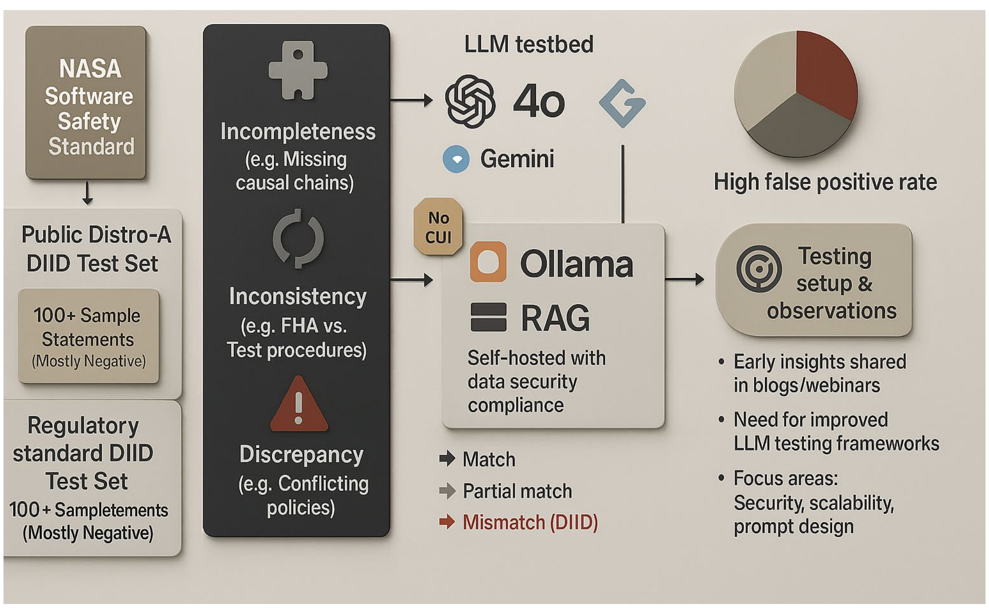


Figure 2. AI-Augmented DIID Detection Pipeline in DoD Safety Standards

Our initial testing is simplified from the actual needs for an operational system. For example, the following aspects are, for now, not the focus of our testing:

- *Data security*—Sensitive information must be protected. Some software acquisition activities may involve sensitive or proprietary information to LLMs, which raises concerns about data security and privacy. We never use CUI or secured information for these tests because the online LLM services do not support data security for CUI or proprietary data, which is a known concern. We therefore include self-hosted LLMs as part of our testing to observe similarities or differences in the responses. Organizations should always be aware of how to mitigate the data disclosure risks of LLMs and prevent access to private or protected data.
- *Scalability*—We initially uses LLMs with RAG for testing, but more online services have increased their support for scaling to support multiple documents. RAG is still needed for the self-hosted solutions.

These initial tests highlight a few important areas of discussion to achieve AI-augmented DIID detection for software acquisition use cases:

- *Prompt engineering*, which is important to ensure clear definition of DIID types and provide guidance on what documents to compare.
- *LLM testing frameworks.* As we assess current LLM testing frameworks and data sets, we found gaps for the DIID use case that point to the need for additional testing resources.

The Role of Prompt Engineering in DIID Detection

The flexibility of LLMs to support a wide range of uses effectively is supported in part by prompt engineering (Liu et al., 2023), which involves structured interactions with LLMs to optimize outputs via natural language interfaces. Prompt patterns (White et al., 2023) codify best practices for phrasing prompts to maximize extraction accuracy and provide knowledge transfer mechanisms to problem-solve with LLMs more effectively and accurately. Prompt patterns also enable more effective and repeatable performance of LLMs, and many patterns have been identified for a range of task objectives.

Prompt engineering is a key capability that provides new opportunities for DIID detection in software acquisition, e.g., the *Context Manager* pattern (White et al., 2023) directs an LLM to set the context. Prompt engineering, including identification and refinement of relevant prompt patterns, is thus central to our research. When DIID types are narrowly defined and examples are provided to LLMs through fine tuning or prompt engineering, DIID detection can be improved for specific use cases. In our initial experiments, failure to include DIID definitions in the prompt led to a high variance in the output and the increase confusion for the user. For instance, a user would have to read the output several times to understand is the output was consistent with the DIID types.

To detect DIID in software artifact analysis, an LLM must act like a reviewer, i.e., it should analyze the content, the context, and then apply reasoning to identify gaps. For incompleteness, this analysis means knowing what should be in the document (perhaps via a template or a checklist derived from policy) and checking if it is there. For inconsistencies, it entails comparing statements for logical alignment. For discrepancies, it requires cross-referencing external sources.

Figure 3. The role of Prompt Engineering in DIID Detection is a summary of prompt engineering for DIID detection. A DIID taxonomy must be provided to ensure consistency in the types of DIID. Specific context is sometimes needed to define how different documents are related.

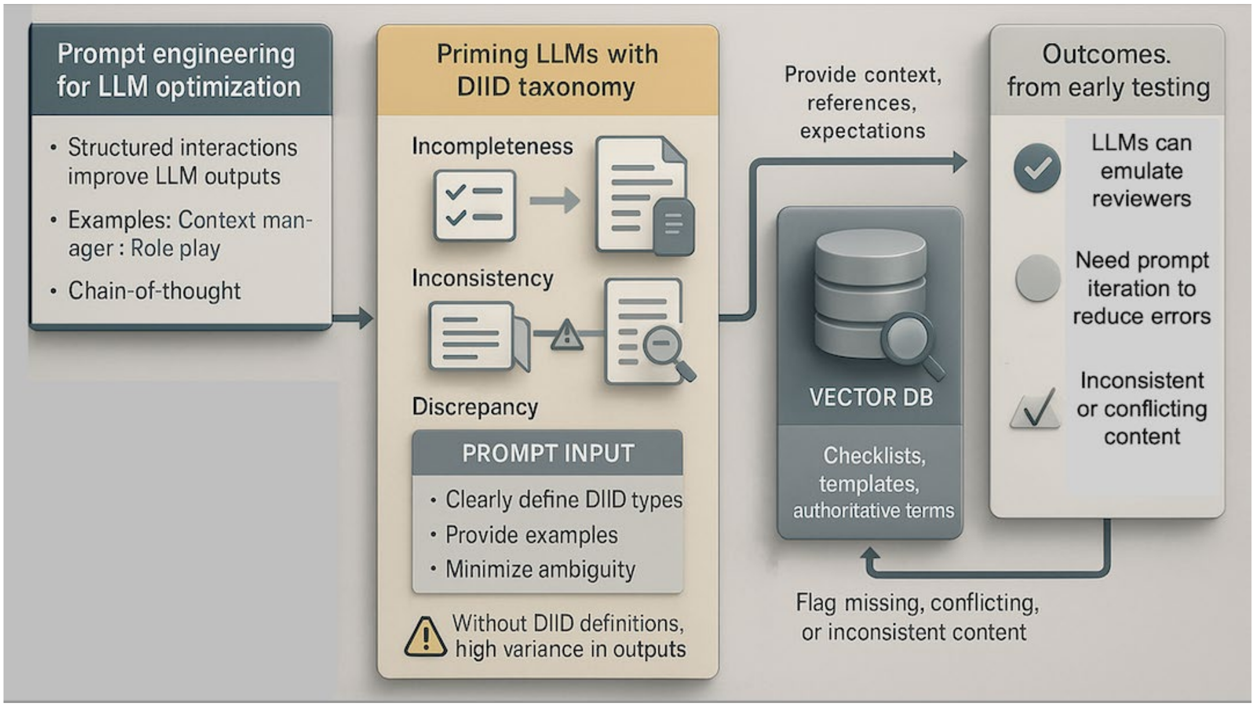


Figure 3. The Role of Prompt Engineering in DIID Detection

These complex tasks are where the combination of LLM + vector database shines. A vector database can supply the LLM with a checklist of expected content (for completeness checks) or the authoritative values/terms from reference documents (for discrepancy checks). An LLM, in turn, can combine this information and flag where the document deviates. By training or priming the LLM with plenty of examples of DIID issues (possibly from past project documentation annotated with findings), we can further improve its detection abilities. Early experimental setups indicate that LLMs can replicate the judgment of reviewers in identifying such issues, although careful evaluation and iteration on the prompts are required to minimize false positives and false negatives.

LLM Testing Frameworks for DIID Use Cases

Evaluating the effectiveness of LLMs on DIID detection requires specialized testing frameworks and data. Common LLM testing frameworks and benchmarks provide a starting point but often do not directly assess the kind of document analysis needed for regulatory compliance. For example, OpenAI's Evals toolkit allows users to create custom evaluations and has been used to measure GPT-4's performance on tasks ranging from math problems to code generation. For DIID detection on software acquisition, however, we need to design evaluations that mirror the task, e.g., given a set of documents and related regulations, does the LLM correctly identify all instances of DIID?

LLM testing protocols for detecting inconsistencies in document summaries could be extended because it provides tests specific to DIID and benchmarks for some of the latest LLMs, as described in Laban et al. (2023). This work provides a relevant DIID detection test framework and dataset that considers multiple discrepancy examples, with multiple LLMs, and compares results of a few benchmarks. This current solution is limited to only a subset of the DIID detection needed for software acquisition, however, and without representation about levels of abstraction.

Figure 4 further describes the LLM testing framework for DIID. Using the test data set previously discussed, multiple LLMs are tested for DIID detection. Additional DIID tests with real systems are performed to ensure consistency of the DIID detection.

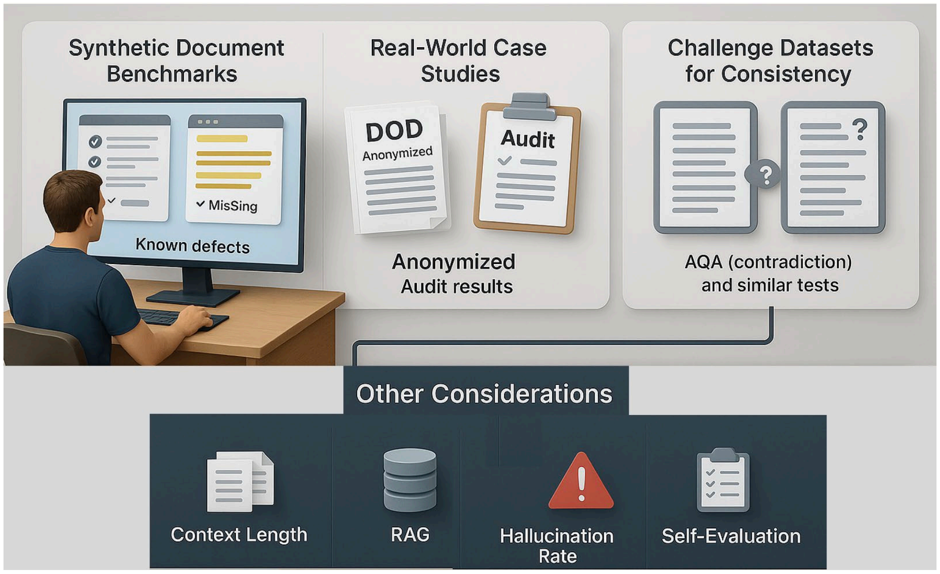


Figure 4. LLM Testing Frameworks for DIID Use Cases

Traditional natural language processing (NLP) benchmarks, such as MMLU (Massive Multitask Language Understanding) or BIG-bench, include reading comprehension and logical reasoning tasks, but they don't capture the multi-document, domain-specific nature of DIID detection. Similarly, academic benchmarks like Holistic Evaluation of Language Models (HELM; Liang et al., 2023) provide broad assessments on accuracy, robustness, and fairness across tasks, but lack metrics for "document consistency checking" or "completeness verification" which are critical in our context. These gaps indicate that testing benchmarks should be extended to better consider DIID detection. Potential approaches to test LLMs for DIID include the following:

- *Synthetic document benchmarks*—Construct pairs of documents and regulatory checklists where we intentionally introduce known incompleteness or inconsistencies. For instance, create a requirement spec with one requirement missing a subsection, or two documents with slightly differing data for the same item. These synthetic cases (with ground-truth labels of where the issues are) can be used to evaluate LLMs quantitatively. Metrics like precision, recall, and F1-score for issue detection can be computed to indicate how many of the known missing sections did the model correctly flag (recall) versus how many flags were incorrect (precision).
- *Real-world case studies*—Use actual past project documents from DoD or industry, with appropriate anonymization, where known issues were found through audits. These documents serve as a gold standard to see if the LLM finds the same issues. This approach is more challenging because the model might find different issues than the human auditors. In these cases, additional analysis is needed to determine if the issues are valid or hallucinations. An evaluation framework must account for these differences by having experts review the AI-identified DIID issues.
- *Challenge datasets for consistency*—Leverage existing datasets focusing on contradiction detection or question answering consistency. For example, some academic tasks require a model to identify contradictions in text (similar to our inconsistency detection). Adapting such tasks (e.g., the BoolQ or contradiction questions from NLI benchmarks) could provide some

insight into how well the LLM maintains consistency understanding. If an LLM struggles with basic contradiction identification in short texts, it will likely struggle with large documents too.

In testing current LLMs for DIID, limitations and challenges have been identified. One consideration is how to scale LLMs to support large documents or many documents, which is very common in DoD software acquisition and software engineering. RAG is a popular approach to scale LLMs for multiple documents and has proven both effective and relatively fast response but effectively implementing RAG systems requires some decision about how to chunk and store data (Godfrey, 2024). An alternate approach is to increase the context length and the trend has been for LLMs to increase context length to ~ 1 million for current models. This larger context length simplifies application of LLMs to larger document sets because it does not require RAG. We are considering architectures with and without RAG, which complicates evaluation.

Another challenge is ensuring the model doesn't hallucinate. A powerful generative model might "imagine" a discrepancy that isn't there, especially if prompts are not tight. This type of hallucination is analogous to a false positive in testing. For reliable adoption, we prefer an LLM to miss a minor issue (false negative) rather than invent a problem (false positive) that sends teams on a wild goose chase. Testing frameworks must measure hallucination rates. One way is to include control documents that are error-free (or have no DIIDs) and then verify that the LLM mostly outputs "No issues found" for those. Any issues it does claim in a known-good document count as false alarms.

We also can examine the use of LLM self-evaluation and iterative refinement in testing. Some frameworks allow an LLM to evaluate its own answers or engage in a back-and-forth (e.g., asking the LLM to provide rationale and then verifying the rationale). For DIID, this self-evaluation could mean the LLM first lists what it thinks should be in the document, then checks off what's present. By evaluating how well this self-check correlates with actual document quality, we gauge the LLM's thoroughness.

Citing testing benchmarks and prior work, some LLMs have demonstrated human-level performance on various professional and academic benchmarks, suggesting its reasoning abilities are strong (Minaee et al., 2025). This finding gives optimism that LLMs can handle complex compliance tasks with the right prompts. DoD domain–specific testing with software artifacts is limited, however, and does not include specific testing of DIID. Early internal tests we constructed indicate that LLMs can catch subtle requirement inconsistencies (e.g., mismatched units or thresholds) that earlier models or simple scripts would miss. Conversely, some LLMs struggled unless the inconsistency was blatant, highlighting a performance gap. Such observations underscore the need for continued testing and prompt refinement, as well as possibly fine-tuning models on compliance data to improve their capabilities in DIID detection.

In summary, adapting LLM testing frameworks to DIID use cases involves creating realistic test scenarios and measuring detection performance carefully. These activities not only help quantify the effectiveness of current LLMs but also guide future improvements. For example, if we find that LLMs often miss certain types of discrepancies (e.g., terminology mismatches), we can then focus on training/prompt strategies to address those problems. In addition, if we find LLMs are flagging too many non-issues, we can adjust prompts so they are more conservative. Systematic testing and evaluation is essential for robust AI-augmented compliance in practice.

Quantifying the Benefits of LLM-Augmented DIID Detection

Beyond qualitative improvements, it is important to measure how LLM-augmented DIID detection impacts the software acquisition process. Prior research on humans detecting discrepancies in text provides some clues. For example, Schoor et al. (2023) found that when readers encountered conflicting information, it increased their "attention to sources," i.e., they spent extra effort cross-checking where the information came from. In the software acquisition DIID context, this suggests that an AI tool must not only flag an issue but also point analysts to the source evidence (e.g., the specific sections of the documents in conflict) to effectively aid and not hinder the human's workflow.

Figure 5 illustrates some of the potential benefits of DIID detection. Discussions with acquisition professionals and software engineers that review program artifacts for regulatory compliance have indicated that manual review is both exhausting and error prone, although quantifying the frequency of these issues is difficult. The benefits represented are based on discussions and will be used to define future testing.

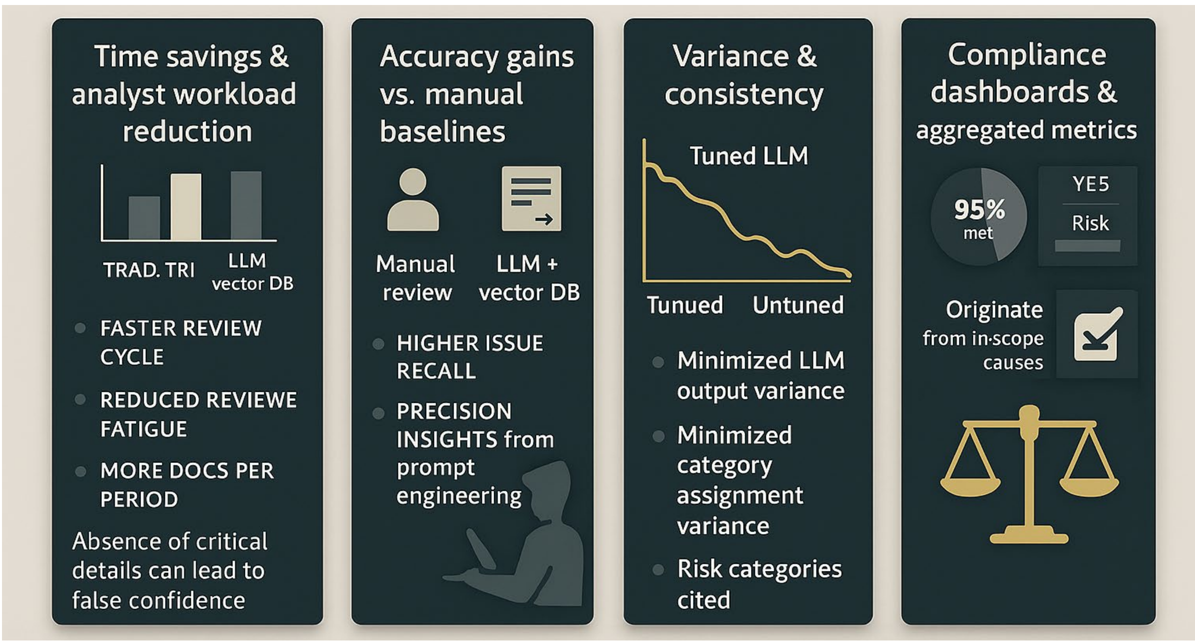


Figure 5. Measuring the Impact of AI-Augmented DIID Detection

We are incorporating the insights from prior work by designing our LLM's output to include references or quotes from the documents, so analysts can quickly validate the AI's findings. To evaluate the overall value proposition of LLM augmentation, we consider the following key questions and metrics:

- *Time savings and analyst workload reduction*—Does an LLM-assisted review significantly reduce the time and effort required by human analysts? We hypothesize that automating the tedious parts of the comparison will allow analysts to review more documents within the same time frame or to focus on deeper analysis of each document. Our hypothesis can be evaluated by controlled experiments where one group uses the LLM assistant and another follows traditional methods, measuring total person-hours spent. We are also exploring cognitive load indicators (perhaps via surveys) to see if LLM-based tools alleviate reviewer fatigue.
- *Accuracy gains compared to manual baselines*—Can the LLM + vector database approach match or exceed human performance in finding DIIDs, and with more consistency across different document types? We hypothesize that LLM-based tool will have higher recall (catch more true issues) than a typical manual review, especially for cross-document inconsistencies that humans might overlook. Precision (avoiding false flags) might initially be lower, but we expect the AI's precision to improve with effective prompt engineering and

prompt patterns. We will measure this improvement by comparing the set of issues found by humans versus the AI in test scenarios.

- *Variance*—An interesting aspect is the variance of DIID detection in both human reviewers and LLMs. Human reviewers' thoroughness can vary widely and is not well quantified, but this issue is one of the reasons why regulatory compliance reviews include multiple reviewers. Given the same prompt and data, LLM output can vary, which is a common issue for generative AI solutions. We aim to demonstrate that the LLM output variance can be minimized, ensuring even challenging documents get a thorough check every time.
- *Enhanced traceability and new metrics*—A potential ancillary benefit of using an LLM for DIID detection is the traceability data it can generate. As the LLM links sections of policy to sections of project artifacts, it could produce a traceability matrix or graph as a by-product. For instance, if requirement *A* in a standard is fulfilled by section *X* of a design document, the LLM's analysis can record that link. Over time, this approach generates a graph of which policies map to which project artifacts. Such an output could be used to answer questions like "Have all policy clauses been addressed by the project's documents?"—providing an assurance metric for compliance coverage. We are exploring the extent to which an LLM's outputs can be aggregated into useful metrics, such as "percentage of requirements with at least one corresponding implementation evidence." This analysis goes beyond what manual reviews typically produce and could transform how compliance is reported by generating dynamic dashboards of compliance status.

In summary, our research will quantify the LLM's impact on efficiency (time/workload), effectiveness (issues detected vs. missed), and augmented capabilities (like automatic traceability). Demonstrating concrete improvements in these areas is essential for gaining stakeholder buy-in, especially in DoD programs where any new tool must justify its adoption in terms of saved resources or improved outcomes.

Future Research

Our initial results are promising, but we are investigating the following avenues to enhance and validate this approach further, as shown in Figure 6 and described below.

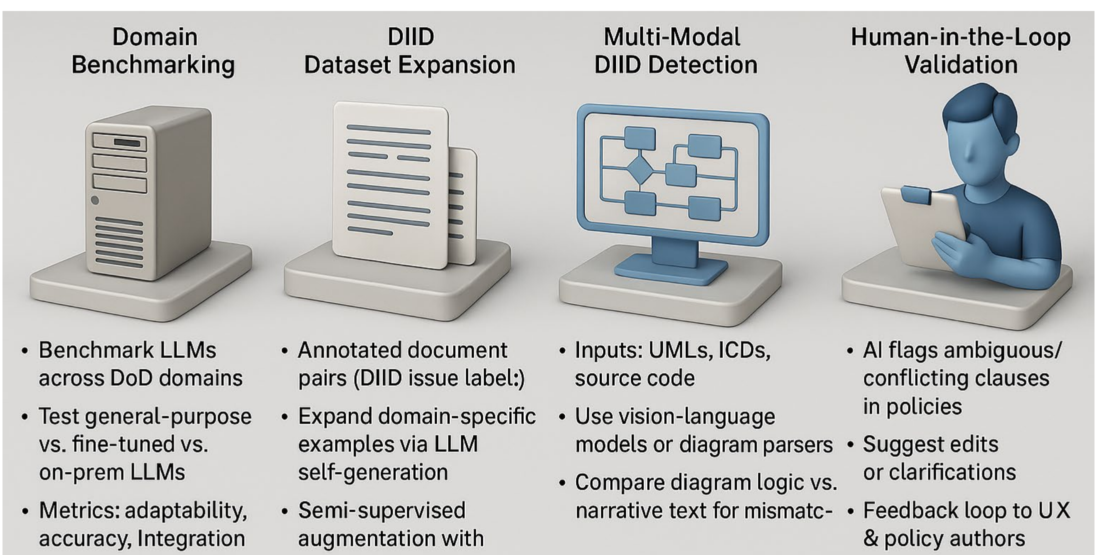


Figure 6. Future Research Directions for AI-Augmented DIID Detection

1. *Benchmark LLMs across different software acquisition domains.* The DoD acquisition landscape is diverse, i.e., what works for a software-intensive weapons system might need adaptation for an IT business system or a healthcare system. We will test the DIID approach on documents from multiple domains (with their own terminologies and compliance standards) to evaluate the model's adaptability and identify if certain domains require specialized tuning or domain-specific training data. Likewise, we are experimenting with various LLMs (including newer models as they become available, or smaller fine-tuned models that could be deployed on-premises for security reasons) to see which offer the best balance of accuracy, context length, and ease of integration.
2. *Refine our DIID datasets and training resources.* A current limitation in our work is the lack of large, labeled datasets of DIID occurrences. We will therefore curate and release a dataset of annotated document pairs with known incompleteness, inconsistencies, and discrepancies. A DIID specific dataset will not only help us fine-tune models for better performance, but it can also benefit the research community by providing a benchmark for this problem. In addition, we might explore semi-supervised approaches: using the LLM itself to generate plausible DIID examples to augment training data (with careful validation to avoid reinforcing errors). We are currently extending existing software life cycle data sets to include considerations of DIID, such as Moreno et al. (2024).
3. *Incorporate multi-modal inputs like diagrams and code into the DIID detection process.* Many regulatory and technical documents include information that isn't pure text, e.g., architecture diagrams, interface control drawings, or source code snippets for critical algorithms. Our future work will examine multi-modal LLMs or pipelines to detect DIID issues across different media. For example, a UML diagram inconsistent with the written design description. Early exploration might involve feeding textual descriptions of diagrams into the LLM to see if it can relate them to the text.
4. *Rigorously quantify the benefits of LLM-augmented DIID detection in realistic settings.* We are conducting user studies where experienced acquisition professionals use the AI assistant on real tasks and provide feedback on workload reduction and confidence in the results. We will measure time saved, error rates reduced, and collect qualitative feedback on how the tool affects their workflow. Such studies will be vital for refining the tool's usability and demonstrating return on investment (ROI) to decision-makers in the DoD.
5. *Drive changes in the regulatory documents themselves.* If AI tools struggle due to ambiguous or conflicting policy language, that signals an opportunity to improve the source materials. We foresee a feedback loop where the AI not only checks compliance but also suggests improvements to the policies and standards. For example, the LLM might frequently flag a particular paragraph in a policy as confusing or internally inconsistent; this could be reported to the policy authors, who can then clarify the text in the next revision. In fact, we have observed that with simple prompting, LLMs can recommend rewordings of requirements for clarity.

By adopting the approaches described above, organizations can create AI-friendly regulations that are easier for both machines and humans to parse. This co-evolution of policy and AI tooling—essentially writing requirements with automated checking in mind—is an intriguing area for future exploration. It aligns with the DoD's interest in more modular and unambiguous requirements and could further accelerate compliance activities. We will pursue research into how best to formulate guidance for policy authors, potentially in collaboration with DoD standards bodies.

In summary, our future work will scale up our current DIID research to cover more domains and data types, systematically evaluate its impact, and explore the bidirectional relationship between AI and policy (e.g., to determine how each can inform improvements in the other). We believe these steps are crucial to transition our research into a practical capability for the DoD and beyond.

Concluding Remarks

This paper discussed accelerating regulatory compliance checks in software acquisition by leveraging LLMs. We defined document incompleteness, inconsistencies, and discrepancies, and demonstrated how an LLM, guided by prompt engineering, can effectively assist in detecting DIID. Our approach has the potential to reduce the manual burden on acquisition professionals significantly, increase the consistency and thoroughness of compliance reviews, and ultimately help DoD programs deliver capability faster without sacrificing rigor.

Our preliminary exploration indicates that even with current LLM technology, there is improvement in DIID detection accuracy and efficiency when compared to traditional methods. There remain challenges to address, including creating test frameworks and data sets to ensure the LLM's outputs are trustworthy and consistent, to tailoring the solution for different domains and document types. Ongoing work is tackling these through targeted evaluation and iterative refinement.

Overall, the integration of human expertise with LLM capabilities is a powerful new paradigm. This human-AI collaboration can not only speed up the acquisition process but also improve its outcomes by catching issues early and iteratively with different software releases. As generative AI tools mature, and as organizations adapt their practices (and possibly their documentation standards) to better accommodate AI assistance, we anticipate that LLM-augmented compliance checking will become an invaluable part of software engineering researchers and practitioners. By continuing to evaluate and improve our approach in realistic settings, and by sharing insights with the community, we aim to move one step closer to a future where compliance is maintained more proactively and efficiently.

Acknowledgements

Copyright 2025 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. Requests for permission for non-licensed uses should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

DM25-0544

References

Bellomo, S., Zhang, S., Ivers, J., Cohen, J. B., & Ozkaya, I. (2023). *Assessing opportunities for LLMs in software engineering and acquisition*. Carnegie Mellon University Software Engineering Institute. https://doi.org/10.58012/m3hj-6w28

Bozkurt, A. (2024). Tell me your prompts and I will make them true: The alchemy of prompt engineering and generative AI. Open Praxis, 16(2). https://doi.org/10.55982/openpraxis.16.2.661

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. Proceedings of the 34th International Conference on Neural Information Processing Systems, 1877–1901.

Cao, T., Raman, N., Dervovic, D., & Tan, C. (2024). Characterizing multimodal long-form summarization: A case study on financial reports (No. arXiv:2404.06162). ArXiv. https://doi.org/10.48550/arXiv.2404.06162

GAO. (2022). Weapon systems annual assessment: Challenges to fielding capabilities faster persist. https://www.gao.gov/assets/gao-22-105230.pdf

Godfrey, C. (2024, November 22). RAG in the era of long-context LLMs. Medium. https://medium.com/tr-labs-ml-engineering-blog/rag-in-the-era-of-long-context-llms-b8ecda2d5693

Laban, P., Kryściński, W., Agarwal, D., Fabbri, A. R., Xiong, C., Joty, S., & Wu, C.-S. (2023). LLMs as factual reasoners: Insights from existing benchmarks and beyond. https://doi.org/10.48550/ARXIV.2305.14540

Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Yuan, B., Yan, B., Zhang, C., Cosgrove, C. A., Manning, C. D., Re, C., Acosta-Navas, D., Hudson, D. A., … Koreeda, Y. (2023). Holistic evaluation of language models. Transactions on Machine Learning Research. https://openreview.net/forum?id=iO4LZibEqW

Liu, P., Zhang, L., & Gulla, J. A. (2023). Pre-train, prompt and recommendation: A comprehensive survey of language modelling paradigm adaptations in recommender systems (No. arXiv:2302.03735). ArXiv. https://doi.org/10.48550/arXiv.2302.03735

MIL-STD-882E. (2023). https://quicksearch.dla.mil/qsDocDetails.aspx?ident_number=36027

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2025). Large language models: A survey (No. arXiv:2402.06196). ArXiv. https://doi.org/10.48550/arXiv.2402.06196

Moilanen, T., Sivonen, M., Hipp, K., Kallio, H., Papinaho, O., Stolt, M., Turjamaa, R., Häggman-Laitila, A., & Kangasniemi, M. (2022). Developing a feasible and credible method for analyzing healthcare documents as written data. Global Qualitative Nursing Research, 9, 23333936221108706. https://doi.org/10.1177/23333936221108706

Moreno, G., Hristozov, A., Robert, J., & Klein, M. (2024, July 23). A model problem for assurance research: An autonomous humanitarian mission scenario. https://insights.sei.cmu.edu/library/a-model-problem-for-assurance-research-an-autonomous-humanitarian-mission-scenario/

Nextra. (n.d.). RAG for LLMs. https://www.promptingguide.ai/research/rag#rag-evaluation

Prasad, N., Boughanem, M., & Dkaki, T. (2024). Exploring large language models and hierarchical frameworks for classification of large unstructured legal documents (No. arXiv:2403.06872). ArXiv. https://doi.org/10.48550/arXiv.2403.06872

Robert, J. & Schmidt, D. (2024, January 22). 10 benefits and 10 challenges of applying large language models to DoD software acquisition. https://insights.sei.cmu.edu/blog/10-benefits-and-10-challenges-of-applying-large-language-models-to-dod-software-acquisition/

Robert, J., Ivers, J., Schmidt, D. C., Ozkaya, I., & Zhang, S. (2024). The future of software engineering and acquisition with generative AI. Crosstalk, AI Taming the Beast, 26–43. https://community.apan.org/wg/crosstalk/m/documents/464157

Schmidt, D. C., & Robert, J. (2024, April 1). Applying large language models to DoD software acquisition: An initial experiment. Carnegie Mellon University, Software Engineering Institute's Insights (Blog). https://doi.org/10.58012/s32x-gc46

Schoor, C., Rouet, J.-F., & Britt, M. A. (2023). Effects of context and discrepancy when reading multiple documents. Reading and Writing, 36(5), 1111–1143. https://doi.org/10.1007/s11145-022-10321-2

Shokrollahi, Y., Yarmohammadtoosky, S., Nikahd, M. M., Dong, P., Li, X., & Gu, L. (2023). A comprehensive review of generative AI in healthcare. https://doi.org/10.48550/ARXIV.2310.00795

Shukla, N. K., Katikeri, R., Raja, M., Sivam, G., Yadav, S., Vaid, A., & Prabhakararao, S. (2023). Investigating large language models for financial causality detection in multilingual setup. 2023 IEEE International Conference on Big Data (BigData), 2866–2871. https://doi.org/10.1109/BigData59044.2023.10386558

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT (No. arXiv:2302.11382). ArXiv. https://doi.org/10.48550/arXiv.2302.11382

Accelerating the Future
Leveraging AI for the Transformative Federal Acquisition

Ryan Novak—Acquisition Outcome Lead at MITRE, brings 30 years of experience as a USAF Contracting Officer and acquisition innovator. He holds multiple advanced degrees, including two MBAs, and is DAWIA III Contracting certified. A published author and speaker, Ryan created the Challenge-Based Acquisition Handbook and taught innovative acquisition approaches to 800+ professionals across government, industry, and MITRE. He has led strategic programs like iCAMS, held an unlimited CO warrant, and completed executive programs at Oxford and Yale. As MITRE's Acquisition Innovation Center Lead for AI in Acquisition, he drives AI strategies, tools, and prototypes to transform federal procurement under the ACQAIRE Portfolio. [rnovak@mitre.org]

Adam Bouffard—is a Group Leader and Principal Decision Analyst for MITRE. He has over 20 years of experience in acquisition and contracts supporting multiple DoD and Civilian Government Agencies. Prior to joining MITRE in 2015 he worked as a civilian supporting Naval Sea Systems Command and Office of Naval Research. He has a master's degree in Systems Engineering from George Washington University and has multiple contracts certifications. [abouffard@mitre.org]

Wilson Miles—is an Associate Research Fellow at ETI. Miles's research and analysis portfolio focuses on emerging technology supply chains, acquisition policy, artificial intelligence, workforce issues, and other modernization technology policy issues. Miles previously held internships at multiple nonprofit organizations including CRDF Global, the Hudson Institute, and the Foundation for Defense of Democracies. He received his master's in International Affairs: U.S. Foreign Policy and National Security from American University's School of International Service and his bachelor's in International Relations from Linfield University. [wmiles@ndia.org]

Christopher Barlow—is a Sr. Acquisition Analyst for MITRE with over 8 years of experience and has demonstrated expertise in audit risk mitigation, acquisition data analytics, process development and optimization, as well as innovative acquisition approaches including challenge-based acquisitions (ChBA), and artificial intelligence applications. His diverse experience spans multiple federal agencies. Chris holds a bachelor's in business administration from the University of Pikeville and is a Certified Business Intelligence Professional from Villanova University. [cbarlow@mitre.org]

Greg Forman—is a business and economic analyst at MITRE with over 15 years of experience supporting U.S. government agencies in cost estimation, financial analysis, and acquisition strategy. Since 2015, he has advised Department of Defense and civilian agencies on large-scale IT system procurements, with a particular focus on Health IT. Greg holds a master's degree in Systems Engineering Management and is currently focused on modernizing cost estimation practices through artificial intelligence and automation. He draws on his experience sharing this work at internal briefings and industry forums—including the 2023 ICEAA Workshop—and is currently authoring several papers on the application of AI in acquisition. [gdforman@mitre.org]

Zineb Kharbouch—is a Decisions Lead at the MITRE Corporation, with over 10 years of experience in project management, acquisitions, and contracts. Her current work focuses on developing AI acquisition guidance for government use, including contributions to the Acquisition Integration and Interoperability (AI2) project under the Assistant Secretary of Defense for Acquisition (ASD[A]) and the Army's Project Linchpin AI Acquisition Playbook. Zineb holds a bachelor's degree in physiology and neurobiology from University of Maryland. [zkharbouch@mitre.org]

Kevin Forbes—is a Principal Software Engineer at the MITRE Corporation, with over 25 years of software development experience. His current work centers on the application of Generative AI to tackle challenges in government acquisition and software development life cycle. Kevin earned his bachelor's degree in computer science from Georgetown University and a master's degree in Systems Engineering from Johns Hopkins University. [forbes@mitre.org]

Abstract

The U.S. Department of Defense (DoD) has requested $310.7 billion in funding for Fiscal Year 2025 dedicated to Procurement, Research, Development, Test, and Evaluation (RDT&E). This includes an allocation of $167.5 billion for Procurement and $143.2 billion for RDT&E. As the largest federal contracting entity, the DoD awards contracts across a wide spectrum, from major weapon systems to software development to food services.

To support these vast acquisition programs, the DoD relies on a skilled acquisition workforce equipped with the management, technical, and business capabilities needed to oversee these programs and activities from conception to completion. However, the rise of technologically advanced strategic adversaries highlights the need for the DoD to streamline its acquisition practices. Modernizing these traditionally time-consuming processes to include technologies like artificial intelligence (AI) is critical to bolstering acquisition professionals' efficiency and effectiveness.

To address these challenges, The MITRE Corporation and the National Defense Industrial Association's (NDIA) Emerging Technologies Institute conducted a joint research initiative. In October 2024, MITRE and NDIA ETI co-hosted a symposium on "Leveraging AI in Acquisition," which explored potential AI applications to enhance the defense acquisition workforce and their processes. Drawing on insights from the symposium and interviews with AI and acquisition experts, the research team identified key acquisition areas where DoD-developed and commercial AI could be applied, as well as the barriers to integrating AI within various federal acquisition phases. These targeted areas for AI application potentially include:

- Market Research
- Request for Proposal/Quote Creation
- Evaluation
- Contract Management

While the integration of AI holds significant promise for enhancing defense acquisition, certain critical functions must remain human-led. This research offers a range of recommendations to support the transition to a modern, AI-enabled acquisition system and offers strategies to address structural and cultural barriers.

Introduction

The rapid pace of technological change, evolving threats, and the complexity of global supply chains has created a sense of urgency to modernize the federal acquisition process. Traditional procurement methods, characterized by lengthy manual reviews, duplicative paperwork, multiple layers of approval, and inconsistent data management too often result in extended acquisition lead times. The Government Accountability Office found that the Department of Defense (DoD) took an average of 309 days to award complex service contracts due to administrative bottlenecks and fragmented information systems (U.S. Government Accountability Office, 2018). Such bureaucratic delays can lead to missed opportunities, increased costs, and diminished readiness in an environment where agility and responsiveness are paramount. Moreover, the DoD acquisition workforce is regularly described as overworked, with their workload having "doubled in the last couple of years" (Obis, 2024).

At the end of FY 2022, the DoD acquisition workforce consisted of 157,594 members, which includes DoD civilians and military personnel. The RAND Corporation states that the size of the DoD uniformed acquisition workforce is consistent with previous years; however the DoD civilian acquisition workforce decreased by roughly 28,000 from FY 2021 to FY 2022 with most of the cuts coming from the Army and Navy. The figures below outline the career fields—categorized by functional area—for both DoD civilian and military personnel acquisition professionals (Gates, 2024).

Conventional wisdom holds that contracting officers have little capacity to be creative and innovative due to a demanding workload that requires focus on executing daily tasks using traditional processes.

"PROGRESS WILL DEPEND ON THE RIGHT MIX OF EXPERTISE AND COMMITMENT ACROSS GOVERNMENT AND INDUSTRY"

Artificial Intelligence (AI) offers sets of tools to shift mundane work away from acquisition professionals to IT systems. For the purposes of this paper, AI includes several types of technologies, including machine learning, generative AI, retrieval augmented generation, multi-modal, and robotic process automation. AI-driven analytics can automate parts of initial market research, generate draft solicitations, and identify performance trends across suppliers.

These technologies can also help acquisition professionals detect potential risks earlier in the process and provide data-driven recommendations that inform strategic human decision-making. A 2023 study conducted by Massachusetts Institute of Technology researchers found that using generative AI tools, such as ChatGPT, substantially raised productivity: the average time to complete controlled writing tasks decreased by 40% and output quality rose by 18% (Noy & Zhang, 2023). Not all agencies will benefit equally from AI's capabilities given differences in mission and because effective implementation depends on factors such as data quality, workforce readiness, organizational culture, and authority to operate these advanced tools within DoD systems.

The DoD has explored efforts to integrate AI into various acquisition-related workflows. Some agencies, such as the Defense Logistics Agency, have employed AI to optimize inventory management vis-à-vis supply chain forecasting and demand planning, so that future acquisition decisions are informed by accurate, data-backed predictions (DLA, 2020). The Air Force has also explored AI tools for personnel and resource management (Bistarkey, 2024). In late 2024, the Army announced a pilot program, where they will be experimenting with a generative AI tool that was created to assist with multiple acquisition activities (U.S. Army Public Affairs, 2024). Recent legislation is also supportive of efforts to make use of AI tools to support the acquisition workforce, including:

- The 2021 National Defense Authorization Act (NDAA) encouraged DoD to leverage AI and machine learning in acquisition programs to increase speed, reduce costs, and enhance decision-making.
- Section 237 of the 2025 NDAA directs the secretary of defense to establish a pilot program to evaluate the utility of using AI-enabled software to optimize depots, shipyards, or other manufacturing facilities run by DoD as well as contract administration for DoD, "including the adjudication and review of contracts managed by the Defense Contract Management Agency."

The ability of DoD to incorporate new AI tools in its acquisition workflows depends on dedicated resources and leadership commitment. But it also will be shaped by its people, whose collaboration with industry and leadership across hierarchies will be the foundation of success. There is widespread recognition that DoD's acquisition processes take too long and contracting officers have a heavy workload. There is an inherent risk to technology development and deployment in how the Department approaches acquisition. Progress will depend on the right mix of expertise and commitment across government and industry, who are all working to advance the digital transformation of how the acquisition workforce buys hardware, software, and services.

In October 2024, the National Defense Industrial Association Emerging Technologies Institute (NDIA ETI) and the MITRE Corporation partnered to host a one-day symposium, where representatives from industry, academia, and government gathered to examine the opportunities and challenges of integrating AI in acquisition. To supplement the findings of the symposium, the authors conducted a series of interviews with AI and acquisition experts. Both the event and interviews informed the key takeaways of this paper.

Use Cases in Acquisition

AI will have the biggest impact by assisting humans in sifting through large amounts of data to reduce administrative burdens and enable more strategic decision-making. AI-enabled tools are regularly developed in the commercial sector to streamline mundane tasks associated with procurement, contract management, and data analysis.

AI appears likely to enable humans to delegate more and more routine tasks. When considering where to apply AI, DoD leadership and staff should target areas with the lowest cognitive load and mature AI technology, leaving tasks demanding the highest attention span for human experts to maximize return on investment. There is an abundance of use cases for both pre- and post-award contract management.

Although widespread adoption is still pending, early movers in federal acquisition have seen tangible benefits based on initial applications. By relying on AI to handle routine tasks, such as pre-screening vendors, checking contract compliance, and generating initial market intelligence reports, professionals can focus on complex negotiations, supplier relationships, and strategic planning. As AI tools become more robust and interoperable with existing systems, and as agencies continue to invest in data governance and workforce training, the federal acquisition environment will evolve into a more agile, data-driven ecosystem, delivering better value, reduced timelines, increased efficiency and improved mission outcomes.

Points of AI Integration

There are numerous examples of capabilities from companies serving all segments of the federal acquisition life cycle. The following are some illustrative examples of these capabilities and where they can be applied based on company-provided information and market research (Figure 1).

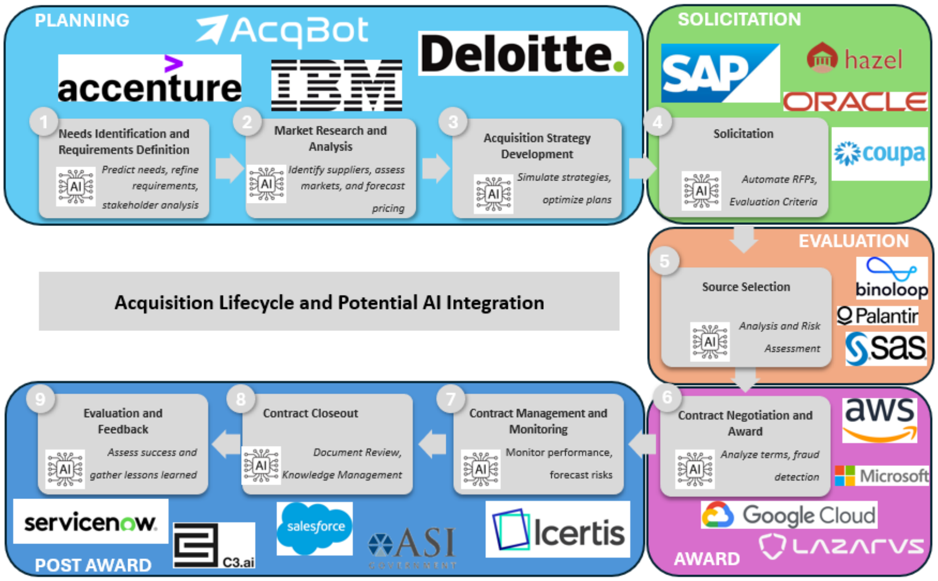


Figure 1 - Acquisition Life Cycle's AI Integration Points

Planning

The planning phase encompasses several areas where AI can be leveraged to increase efficiency, transparency, and effectiveness. NetImpact Strategies, renowned for its innovative solutions, states that in using AI's "newfound ability to harness data-driven insights revolutionizes the requirements development process, enabling agencies to make informed decision and set the stage for successful procurements" (NetImpact Strategies, 2024). Some examples of companies that are providing AI services and tools with potential application in this phase include:

- *ACQBOT*: Offers an array of acquisition-related AI tools including requirements articulation, market research reports, statements of objectives, performance work statements, and several other features related to the planning phase and other acquisition processes and phases.
- *IBM*: Offers AI-driven analytics to help agencies define requirements and plan acquisitions.
- *Deloitte*: Provides AI tools for market research and needs assessment.
- *Accenture:* Uses AI to assist in strategic planning and decision-making processes.

Solicitation

Development of the solicitation package is one of the most important phases of the acquisition life cycle. AI can be leveraged to help generate many different facets of a solicitation. Much of solicitation generation is a repetitive templatized process. There are several different AI tools that can help to supplement this process and reduce the burden of the more tedious and repetitive tasks, such as the development of Sections I of a solicitation.

- *Hazel*: Offers a suite of tools including the capability of writing solicitations with an AI copilot.
- *SAP*: Offers AI-powered solutions for automating solicitation documentation.
- *Oracle*: Provides AI tools for streamlining the creation and distribution of solicitation documents.
- *Coupa*: Delivers AI-driven procurement solutions to optimize the solicitation process.

Evaluation

In this phase AI can be used to help supplement the evaluative analysis based on the evaluation criteria and solicitation developed in the previous stages. Leveraging AI in this phase must prioritize impartial analyses conducted in a way that removes any bias a human would provide. Some of the companies providing tools for this phase include:

- *Binoloop*: Provides Tally, an AI-powered tool that provides intelligent, impartial, and regulation-compliant analysis powered by AI.
- *Palantir*: Uses AI to analyze and evaluate proposals, offering insights and data-driven decision support.
- *SAS*: Provides advanced analytics and AI tools for evaluating bids and assessing vendor capabilities.
- *Booz Allen Hamilton*: Offers AI solutions for risk assessment and proposal evaluation.

Source Selection and Award

In this phase it is important to distill all the applicable information, analysis, and subject matter expert inputs to support fair decision-making processes and then make an award. Examples of tools available to support this phase include:

- *Lazarus AI*: Provides a tailored integration of AI capabilities to support award decisions.
- *Microsoft*: Provides AI tools for contract management and award decision support.
- *Amazon Web Services (AWS)*: Offers AI services for automating contract award processes and compliance checks.
- *Google Cloud*: Delivers AI solutions for optimizing award decisions and ensuring transparency.

Post-Award

This phase has several repetitive and tedious tasks for which AI could be leveraged. AI can help in the analysis of contract milestones, analysis of work completed and adherence to the contract, and to help gain insight into where program improvements can be made. Several companies offer AI tools to help with this phase, including:

- *ASI*: Provides multiple AI tools and solutions to support processing modification back logs.
- *Icertis*: Offers many AI tools, including contract maintenance and contract analytics capabilities.
- *Salesforce*: Uses AI to manage contract performance and vendor relationships.
- *ServiceNow*: Offers AI-driven solutions for post-award contract management and performance monitoring.
- *C3.ai*: Provides AI applications for monitoring and optimizing contract execution and compliance.

These companies are just a sample of the AI service offerings that can be integrated into various stages of the federal acquisition life cycle to enhance efficiency, accuracy, and decision-making capabilities.

Next, we look at how AI can be embedded in each phase of the acquisition life cycle in more detail, along with initial estimates of the level of human oversight needed. Further analysis is needed to validate these estimates.

Detailed AI Integration Analysis

Applications within Acquisition Life Cycle

The following tables are designed to illustrate functions within the acquisition life cycle where AI is well suited, and descriptions of the expected benefits and potential oversight anticipated.

Needs Identification and Requirements Definition

Example Acquisition Function	Benefit(s)	Potential Oversight
Automated Requirements Generation	Assist in drafting clear and precise requirements by analyzing previous contracts and identifying common patterns and language	High
Stakeholder Analysis	Map and analyze stakeholder interests and influence, ensuring that all relevant parties are considered in the requirements definition process	Medium

Acquisition Planning & Strategy

Example Acquisition Function	Benefit(s)	Potential Oversight
Market Research	Identify potential vendors, assess market trends, and predict future needs. Natural language processing can be used to extract relevant information from industry reports and databases	Low
Request for Information (RFI)	Automate data extraction and evaluation, retrieving relevant information based on key criteria. Create more efficient workflow management by automating mundane tasks (e.g., scheduling, document management, and reminders)	Low

Solicitation

Example Acquisition Function	Benefit(s)	Potential Oversight
Demand Forecasting	Predict future demand for goods and services based on historical data, helping agencies plan more accurately	Low
Bid Matching	Match agency needs with potential suppliers by analyzing supplier capabilities and past performance data	Medium

Fair and Reasonable Price Determination	Assist in evaluating proposals by scoring them against predefined criteria, reducing human bias and increasing consistency	High

Evaluation and Source Selection

Example Acquisition Function	Benefit(s)	Potential Oversight
Past Performance Analysis	Predict the likelihood of a vendor's success based on historical performance data, helping acquisition teams make more informed decisions	Low
Risk Assessment	Assess the risk associated with different vendors by analyzing financial health, past performance, and other criteria	Medium
Matching FAR Clauses to Contracts	Suggest the appropriate FAR clause(s) for a given contract	Low
Cost and Schedule Risk	Assist with forecasting the financial and other resources needed in the acquisition of an end item within defined parameters	Low

Contract Award

Example Acquisition Function	Benefit(s)	Potential Oversight
Contract Optimization & Negotiation	Suggest optimal contract terms and conditions by analyzing similar contracts and outcomes. Support the generation of fair and reasonable price determinations	Low
Fraud Detection	Detect anomalies and potential fraud in contract awards by analyzing patterns and flagging suspicious activities	Medium
Contract Writing	Reduce wait times for contract awards, allowing organizations to adapt quickly to new opportunities and requirements	High

Contract Management

Example Acquisition Function	Benefit(s)	Potential Oversight
Performance Monitoring	Monitor contract performance data analytics, alerting managers to potential issues	Low

Automated Reporting	Generate reports on contract performance, compliance, and financials, reducing administrative burden on contract managers	Low
Determination and Findings (D&F)	Assist with generating the necessary written approval documents needed for an authorized official to take certain contract actions	Low

Contract Closeout

Example Acquisition Function	Benefit(s)	Potential Oversight
Document Review	Automate the review of contract documents to assess whether all obligations have been met and identify any outstanding issues	Low
Knowledge Management	Capture lessons learned and best practices from closed contracts	Low

Cautions and Limitations of AI Integration

While AI offers numerous benefits, organizations must also consider challenges such as data privacy, the need for high-quality data, integration with existing systems, and the potential for bias in AI algorithms and training data. Reviews by human experts is still critical because many AI techniques can hallucinate, make errors, or reach false conclusions, especially when data are conflicting, missing, or containing errors. Other techniques are not perfect or can reach false conclusions as is possible with statistics. Ensuring that AI systems are transparent, ethical, and aligned with organizational goals is crucial for successful implementation.

In summary, AI can transform the acquisition life cycle by automating tasks, analyzing data, and providing actionable insights, leading to more efficient, accurate, and speedy processes. Organizations that effectively leverage AI in their acquisition strategies are likely to gain a competitive edge in the marketplace. However, human experts must still be involved to ensure quality, objectivity, and accuracy while taking the outputs from AI tools into human decision-making.

Embedding AI in Contracting Phases

The federal contracting and procurement life cycle consists of several phases, each of which can benefit from the integration of AI to enhance efficiency and effectiveness. Figure 2 illustrates contracting elements in the acquisition life cycle tagged with ways that AI can be injected to each phase to provide support.

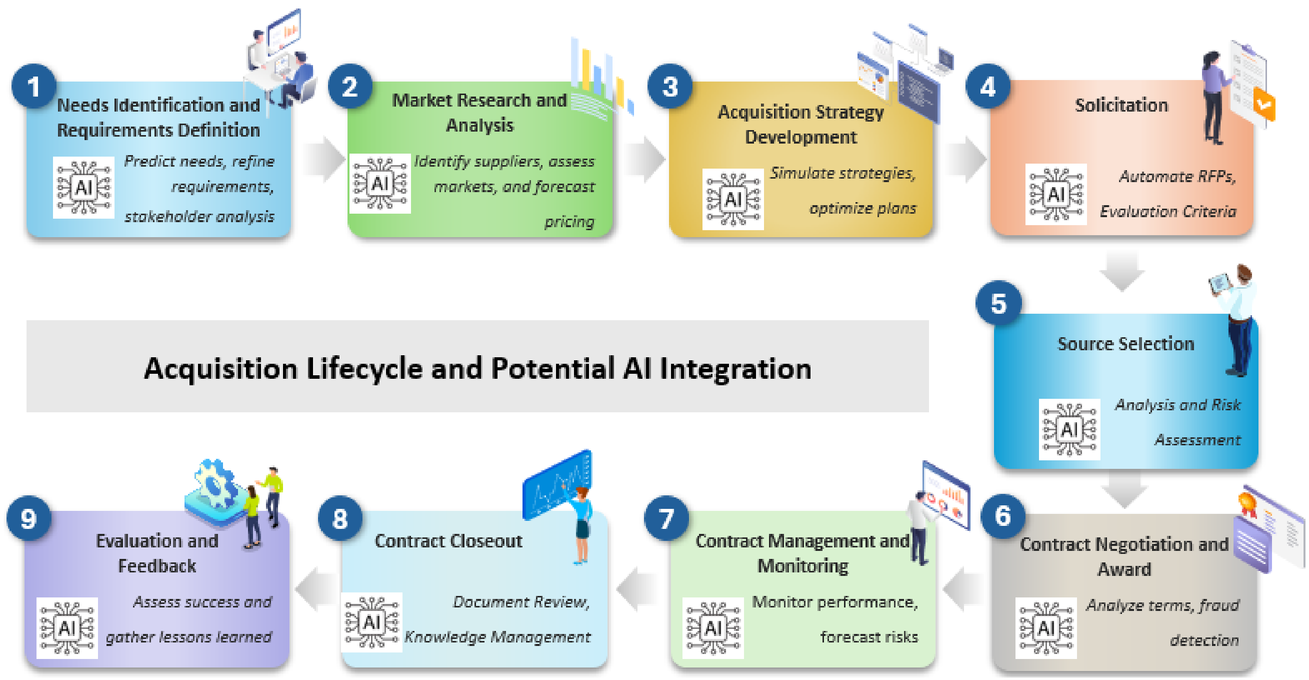


Figure 2 - Leveraging AI in the Acquisition Lifecycle

Step 1: Needs Identification and Requirements Definition

A foundational step of any acquisition which should be a collaboration between the program leadership, stakeholders and contracting officer and supporting functions. Soliciting detailed requirements from program staff can often be challenging, but AI tools can be used to assist in drafting clear and precise requirements by analyzing previous contracts, results of previous programs, and identifying common patterns and language. This could be organization-specific contracts or across a wide data set of contracts from across the federal landscape. In this first step AI can also be used to map and analyze stakeholders' interests and influence, ensuring that all relevant parties are considered in the requirements definition process. AI can help to supplement this process and provide a more technically and operationally feasible starting point to be validated by further market research and analysis.

Step 2: Market Research and Analysis

A common complaint is that this step is cumbersome and overwhelming, and consumes a significant portion of time and energy, particularly in data rich environments. AI could make market research more insightful by analyzing vast amounts of data to identify potential vendors and gauge capabilities offered, assess market trends, and predict future needs. Natural language processing can be used to extract relevant information from industry reports and databases quickly. Machine learning algorithms can predict future demand for goods and services based on historical data, helping agencies plan more accurately. AI could also be used to identify trends in vendors' responses to RFIs or review transcripts from industry days or vendor one-on-ones. By using AI tools to supplement and assist in market research, this step becomes more efficient and provides the information needed to help acquisition professionals execute the next step. By finding more potential vendors, DoD can reduce its use of sole source contracts and task orders.

Step 3: The Acquisition Strategy

Contracting professionals sometimes struggle to develop acquisition strategies due to cultural, bureaucratic, policy, time, and budget constraints. AI could be leveraged to assist in analyzing the optimal path for acquiring goods and services, benefiting from the data captured from past program strategies with successful outcomes. More importantly AI can be used to help educate program and project managers on why particular strategies were chosen and what is required to execute such a strategy. AI could summarize the policies, procedures, and other factors that are driving the chosen strategy and help to inform all the documentation and reviews required. This will help to optimize the execution of the chosen strategy and ensure the contracts and contracting professionals take the best approach and award the most capable vendors.

Step 4: Solicitation

Tools can be used to assist in the development of the evaluation criteria that will be associated and how the award will be made to winning proposal. AI tools could be used to automate the development of the request for proposal (RFP) into a standardized format, resulting in a quicker time to solicitation. AI tools could look across the federal landscape and develop evaluation criteria based off large data sets of previous solicitations and the responses to them. AI could be used to ingest the finalized requirements or objectives, so that selection criteria can be used to properly differentiate between vendors. The use of these AI tools would result in more effective solicitations that have been vetted by an unbiased tool and finalized by the human in the loop. Better solicitations lead to better evaluations, source selections, and awards.

Step 5: Evaluation and Source Selection

Well suited to be optimized with AI tools by helping to evaluate proposals by analyzing them against the predefined criteria such as analyzing supplier capabilities and past performance data. Imagine using AI to assist in a review of supplier capabilities and past performance across all the most recent data across all their contracts with the federal government. AI can then take this information and predict the likelihood of a vendor's success based on this historical data helping to provide more informed decisions. For example, the Contractor Performance Assessment Report Reporting System (CPARS) provides source selection officials with information on contractor past performance. Acquisition professionals have anecdotally expressed the strong need for support in this area and conclude that LLMs could improve how this information is generated and used.

Solicitations produce large volumes of information that must be reviewed in a proposal. AI can help distill this into usable information quickly and help source selection teams to more quickly assess the viability of a vendor against their solicitation requirements. On top of these particular use cases, AI could be leveraged to assess the risk associated with different vendors by analyzing financial health, past performance, and other risk indicators that programs or projects identify.

Step 6: Contract Award and Negotiation

AI tools have many potential applications in this space, such as recommendation engines for optimal contract terms and conditions by analyzing similar contracts and outcomes, better balancing the needs of both the customer and vendor. Additionally, AI can detect anomalies and potential fraud in contract awards by analyzing patterns and flagging suspicious activities. This could help to identify contractors looking to enter the market at all costs, which can be bad for the government if the vendor can't deliver. It also allows for an unbiased review and analysis that helps to inform the contracts and acquisition professionals around things like terms and conditions, clauses, and even data and intellectual property rights. By building the use of AI tools into the negotiation process it gives the contracting officer/specialist the power to quickly produce useful information resulting in better value for the government.

Step 7: Contract Management

This step can be enhanced by using AI-enabled tools by easing the workload for contracting officer representative or contract officer technical representatives. AI can continuously monitor contract performance using data analytics, alerting managers to potential issues before they become significant problems. Along with this AI can generate reports on contract performance, compliance, and financials, reducing the administrative burden. With many CORs and COTRs often having this role as an additional duty, leveraging AI could help to reduce the burden created by these activities, while allowing personnel to focus on higher priority issues.

For example, AI tools could help turn requirements into solicitation packages or more efficiently streamline justification and approval (J&A) paperwork by leveraging past examples of J&A documents. Within DoD's business and financial management enterprises, there exist many opportunities to modernize legacy systems that are responsible for contract, cost, and pricing audits.

Step 8: Contract Closeout

This step could be streamlined using AI-enabled tools by automating the review of contract documents to assess whether all policy, regulatory, and contractual obligations have been met and to identify any outstanding issues. This can help all stakeholders to be able to close out complex and longstanding confidently and to meet all documentation requirements more easily.

Step 9: Evaluation and Feedback

Responsibilities can be supported with AI-enabled tools by capturing lessons learned and best practices from closed contracts or past programs, which would provide valuable insights for future acquisition strategies. AI-supported analysis can help program officials to understand what happened during execution, summarize what occurred, and develop the best program strategies based on all available sources of information. By embedding AI in these acquisition life cycle steps, federal agencies can streamline processes, reduce costs, improve decision-making, and enhance overall acquisition outcomes.

Policy, Cultural, and Technical Challenges

AI—especially Generative AI—shows immediate promise for increasing efficiency and productivity. There is a recognition that AI is changing how the Department views enterprise-level management and some individuals are advocating for the use of AI tools to support acquisition professionals. Its utility for different areas of enterprise management is clear and raises questions related to the efficacy and efficiency of current practices in activities such as legal analysis, compliance, and procurement. However, the application of AI in acquisition processes does pose certain risks such as a lack of transparency in decision-making and potential cybersecurity vulnerabilities. Additionally, much of the current discourse on the use of these tools is focused on how to educate acquisition professionals on AI and how to create incentives for the use of new capabilities. Ultimately, cultural barriers among other considerations constrain the government's ability to collaborate with industry and adopt new technologies.

Policy Considerations

Inherently Governmental Functions (IGFs)

Within acquisition, certain functions are inherently governmental, which statutes and regulations generally define as a particular task or function that must be performed by a government official (Defense Acquisition University, n.d.). IGFs require officials to exert discretion over governance areas such as policy decision-making, performance/mission accountability, and execution of monetary transactions and entitlements. FAR 7.503(c) provides a list of 20 examples of IGF, which includes functions such as awarding and terminating contracts, among others. FAR 7.503(d) describes 19 functions which are closely associated with IGF. If a task or function is not determined to be inherently governmental, DAU states that it may be eligible for performance by private sector contractors through a contract or other service arrangement. The type of performance that private industry can provide to the government is limited to services that gather information on the government's behalf for the purpose of advising, offering opinions, and providing recommendations or ideas. The criteria are that contractors cannot establish government policies, provide actionable organizational decisions, or spend taxpayer dollars. Similarly, AI tools must be treated as advisory rather than authoritative. While they can enhance decision-making, they cannot independently dictate policy, make binding organizational choices, or allocate government resources. Just as government functions require human oversight when working with contractors, the same principle applies to AI: it should support, but not replace, human judgment and accountability.

Classification & Large Language Models (LLMs)

While not a significant barrier, it is important to note that existing DoD policy restricts national security information and controlled unclassified information (CUI) or use in a publicly accessible tool, including personally identifiable or protected health information. DoD personnel are prohibited from entering such information into common commercial generative AI tools, such as ChatGPT or any other LLM that is connected outside protected firewalls and thus divulges CUI. In instances where government-generated data, code, text, or media does not fall into classified or CUI categories, DoD personnel may only input such information into publicly accessible generative AI tools if that content has been approved for public release. This is intended to protect sensitive but unclassified information while allowing for the responsible use of generative AI tools to enhance efficiency and innovation within federal government operations. However, there are LLMs that are authorized to host CUI material; these LLMs must be DoD Impact Level (IL5) compliant, which falls under the responsibility of the Defense Information Systems Agency (Nicewick, 2024). DoD is currently experimenting with LLMs that work inside Non-Secure Internet Protocol Router (NIPR). In June 2024, the Department of the Air Force launched NIPRGPT, which is intended to assist users with a wide range of tasks, such as correspondence, background papers, and code (Secretary of the Air Force Public Affairs, 2024). CamoGPT, which was developed by the Army AI Integration Center, is currently available on NIPR and is built to optimize equipment maintenance, logistics, and supply chain management using data analytics and algorithms (Pharathikoune, 2025).

Barriers

Data Quality & Ownership

The quality and completeness of data available for analysis, refining algorithms, and for use to support machine learning profoundly impacts the effectiveness of AI applications. DoD and defense and commercial industry have struggled to collect and retain data about acquisition processes and program execution. They have also struggled to share data across government and the private sector; between industry partners; and even between government organizations. These issues are generally attributed to business models and incentives, the lack of effective technical infrastructure to support sharing and collaboration, the desire to avoid intrusive management and oversight, classification issues, and other bureaucratic dynamics.

One frequently asked question is: What rights does the federal government have over the data? Intellectual property (IP) rights have grown in importance to DoD as U.S. defense research and development (R&D) spending as a share of global R&D spending has declined. IP rights are also increasingly important to industry partners who rely on their portfolios of developed IP to generate profits from their R&D investments.

Another specific worry regarding technical data bias is potential cases of corruption and compliance issues. This introduces, for example, possible instances of organizational conflicts of interest, where companies have built a product—which is used by DoD—that could help benefit the same company later in a future procurement. Trust in new AI tools and the organization, cleanliness, and management of data are intertwined. Ultimately, the Department understands that trust in data underpins its adoption strategies and has attempted to craft policies that are specifically designed to empower leaders at all levels. However, the issue of explainability of AI-generated results will continue to present challenges for all professionals who are expected to justify decisions or outcomes when AI tools are used.

Security Compliance & Authority to Operate (ATO)

Due to legitimate security and reliability concerns, the ATO process is used to determine when new software can be installed on and used in most of these systems. Waiting for an ATO and working through assessments is often the longest step in deploying software. This process requires a government sponsor or Authorizing Official (AO), who will work hand-in-hand with the industry partner. The pace of technological development, however, requires agile development practices that continuously integrate and deliver software while still maintaining security. Small and non-traditional businesses are playing a key role in developing new AI technologies and often state that obtaining an ATO is slow and expensive, which limits the ability of these companies to work with DoD. By limiting the software available to the government to only products from companies that can accept high initial costs and long-term horizons, the government misses out on innovative solutions from non-traditional players. This bottleneck not only slows adoption but also hinders the government’s ability to rapidly leverage cutting-edge AI solutions, micro tools, and emerging innovations. By streamlining ATO pathways, the DoD can inject these advancements more quickly—where they can make an immediate impact on a timeline more consistent with the development of new AI capabilities and with changing operational needs.

Continuous ATO (cATO) is an emerging approach to deploying secure software faster. DoD’s Chief Information Officer (CIO) defines cATO as “a modernized authorization process designed to work with software delivery organizations that want to move faster and are willing to adopt the necessary culture change” (U.S. Department of Defense, 2024). cATO eliminates the delays that come with leaving testing and certification to the end. By doing those things in parallel, time is saved and mission effectiveness improved; however this process still requires adequate resourcing, which may not be realistic for some government offices. The “Assess Only” process is an alternative option, which allows organizations to incorporate and use products and services that fall below the system level (e.g., system components, hardware, software, IT services) without going through the full ATO process (U.S. DOD Chief Information Officer, 2024). By using fast-track ATOs or testbed environments, vendors could deploy and iterate AI solutions immediately ensuring the government can harness innovation at the speed of relevance rather than being bogged down by lengthy certification timelines. Sec. 1522 of the FY2025 NDAA took a step to help with reciprocity of ATOs. Specifically, it states DoD must develop a policy requiring DoD officials to accept security analysis and artifacts of a cloud capability that has already been authorized by another DoD official or component (U.S. Congress, 2025). The policy must also provide for standardization of accreditation documentation and other measures to enhance reciprocity between DoD components’ respective ATO processes, including Federal Risk and Authorization Management (FedRAMP) ATOs. This legislation marks a step forward in reducing bureaucratic hurdles for both technology companies and DoD.

Acquisition Barriers

The integration of AI-enabled tools into contracting and the acquisition life cycle writ-large suffers from the same challenges that the entire DoD acquisition ecosystem is plagued with, including slow budgets, unclear requirements, and difficulty procuring commercial technologies and services. The Planning, Programming, Budgeting, and Execution (PPBE) Commission found that "the opaque and unresponsive nature of the PPBE process is antithetical to the kind of market signaling, commitment, and certainty that they need to attract private investment supporting development of emerging technologies" (Commission on PPBE, 2024). This is compounded by the fact that there is no coordinated procurement effort to purchase AI-enabled tools, consolidate requirements, and distribute these tools across DoD's acquisition program offices. Moreover, while AI is a highly technical and fast-moving technology, there is also no supportive science and technology enterprise to help DoD understand, leverage, and modify these technologies so that they can be used within DoD's acquisition processes.

Education & Training

According to research at RAND, approximately more than 50% of the DoD civilian acquisition workforce consists of individuals aged 40 and above (Gates et al., 2024). The age distribution—due to the pace of technology development—demands attention toward the need for retraining and upskilling. The education or formal training that a 20–29 year old received is much different than their peers who graduated decades ago. It is imperative for contracting organizations to gain a deep understanding of the workforce skills and processes needed to ensure all contracting officers are able to consistently use and scale new AI tools.

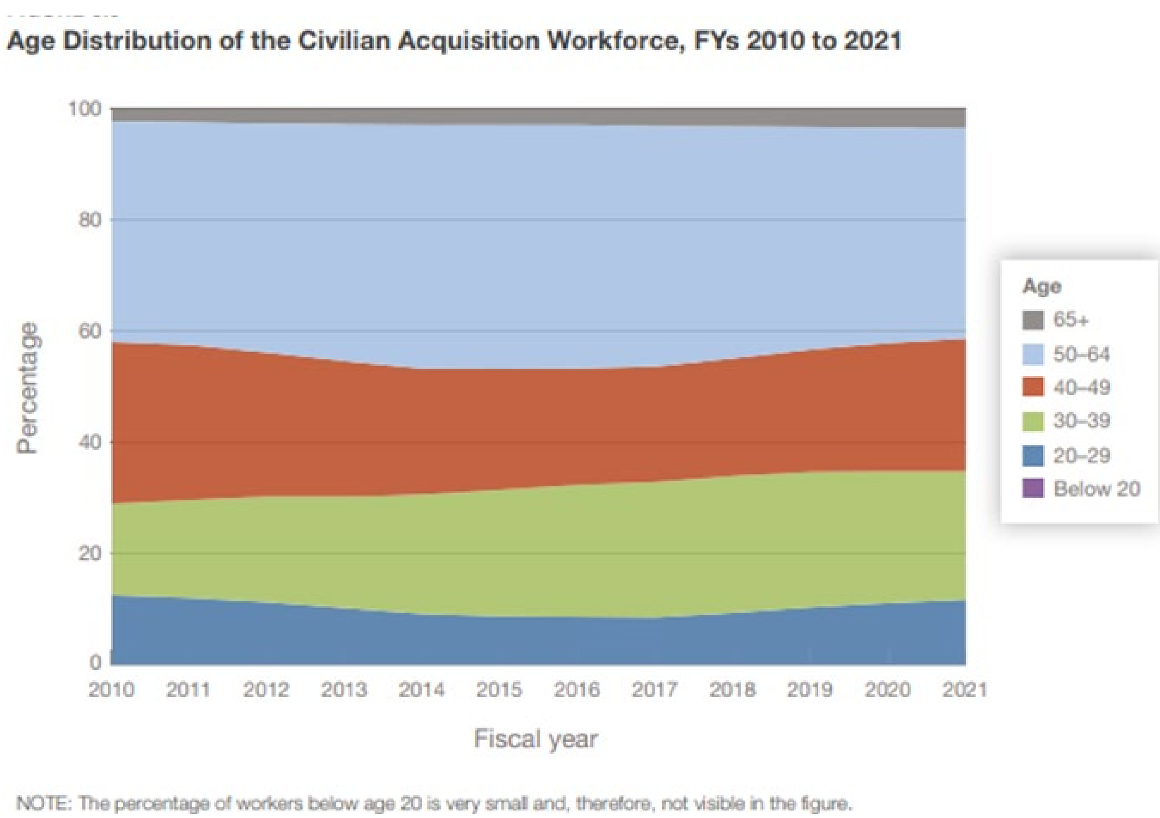


Figure 3 - RAND DoD Civilian Acquisition Workforce

Contracting professionals rarely have the time to enroll in a course to learn a tool; therefore, it is critical that their cognitive load is also well-understood by industry service providers. As such, potential solutions aimed at improving efficiency should be easily integrated into existing systems and processes. New tools should be built with the intention that integration should require minimal training.

Cultural Resistance & Need for Clearer Incentives

Cultural barriers in the workforce also limit the ability to make full use of AI in acquisition, specifically the fear of job loss and risk-aversion. Data from a 2023 Gallup poll shows that just over a fifth of U.S. workers are worried technology will make their jobs obsolete; this certainly applies to the contracting and overall acquisition workforce. The fear of job loss is a real concern. Education and training will need to demystify the realities of AI integration. Moreover, the acquisition community works in such a highly regulatory field, so there is also an intrinsic risk aversion among contracting officers, which leads to a reluctance to explore new tools and processes. Other factors can disincentivize the use of AI tools. For example, it can be difficult for people to use AI-empowered tools when they have to justify the results, yet the AI tools are a black box to them. These factors must force organizations to determine how to incentivize their workforce to use new AI tools.

Considerations for Successful AI Integration

Laying the groundwork for long-term AI adoption in acquisition processes requires a strong emphasis on technical infrastructure, interdepartmental coordination, and clearly articulated standards for data management, training, and ethical governance. Robust infrastructure includes high-performance computing environments and secure, cloud-based architecture equipped to process large datasets, as noted by the National Security Commission on Artificial Intelligence (2021). Ensuring interoperability and integration across existing contract writing, financial management, and logistics platforms is equally essential to minimize data silos and maintain a coherent flow of information throughout the acquisition life cycle (GAO, 2021).

Equally critical is the need for interdepartmental alignment. Establishing clear data-sharing policies, while respecting privacy, classification, and compliance requirements creates a uniform foundation for building and refining AI models. Organizational readiness also entails upskilling the workforce, recruiting and retaining acquisition professionals to operate and interpret AI tools, and enabling them to leverage the insight generated by the AI tool effectively and make informed decisions. Cultivating leadership support, open communication channels, and cultural receptiveness helps overcome resistance to change and encourages a collaborative atmosphere that values innovation over routine practices. These efforts, combined with adherence to legal and ethical guidelines, such as transparent AI outputs, robust cybersecurity protocols, and strong conflict-of-interest policies, will support the development and deployment of systems that develop AI-based actionable recommendations that are both trustworthy and accountable.

Acquisition and Innovation Activities

To promote the use of AI tools by the acquisition workforce, acquisition programs must be established and resourced with the mission to develop and deploy such tools, in the same way traditionally done for weapons systems and other IT capabilities. This includes processes to acquire the AI-tools and transition them into use, some of which will be based on existing commercial capabilities and some of which may be unique to the defense enterprise. There are also very limited science, technology, innovation, and testing activities executed by the military service and defense agencies to support the development and delivery of tools to meet the needs of the acquisition workforce and contractors conducting the "business" of acquisition.

Without the establishment of these kinds of activities, the adoption of AI tools will be sporadic and anecdotal.

User & Leadership Buy-In

Ultimately, widespread adoption of AI-enabled tools will be intrinsically linked to the openness of the acquisition workforce and insistence of their use by DoD senior leadership. It is often colloquially known that "Cultural barriers tend to manifest themselves as policies and budgets." With limited funding, decision-makers prefer to spend appropriated dollars on more urgent needs. To continuously gain buy-in from both users and senior government leadership, embedding AI into the acquisition life cycle must also include a clear understanding of investment returns. The individual, or group of individuals, who are acting as the change agent, or "champion," should be able describe metrics of effectiveness and performance, linked to the program characteristics variable being addressed (i.e., schedule, cost savings, retention, and competition). The metrics must be quantifiable and should be able to be clearly understood by all stakeholders, including government appropriators, users, and agency leadership—all of which will help the AI-enabled tools scale into broader use.

Procedural Reviews

The embedding of AI in acquisition on both the commercial and government side is leading to new workflows, where AI tools are compared to highly skilled assistants and therefore create possibilities for improved decision-making, streamline processes, and ultimately drive toward better acquisition outcomes. On the industry side, the barrier to entry for contract writing, for example, is much lower; it is becoming easier to quickly generate documents when responding to RFPs using AI tools. However, this presents a challenge to government officials. AI-generated responses are becoming easier, making it difficult for contracting officers to weigh bids. This example illuminates the need to adopt appropriate review processes to ensure that all stakeholders' needs are met in accordance with existing laws and regulations.

Coordination and Information-Sharing Between Pilot Programs

As pilot programs continue to develop and successful efforts transition, the lessons learned need to be shared and implemented across DoD. However, robust information-sharing is historically a challenge within the Pentagon, sometimes leading to unwanted duplicative efforts and heterogeneous practices leading to confusion for both government and industry. Current practices for developing AI-enabled tools include having a clear understanding of the problem and workflows, not dictating the solution, and requiring demonstrations. Coordination is necessary so that resources are used more efficiently and will help inform DoD efforts that look to build on past successes and failures. The insights gained should also be used to inform future policies, training, and budgetary requests, which would be based on clear and documented data.

Insufficient data collection, labeling, and storage is a deterrent to adopting or even experimenting with AI capabilities. While there have been attempts at tracking the acquisition workforce's time on a given task, these efforts have failed. Most professionals do not have these steps documented and therefore it is difficult to assess how effective an AI-enabled tool is without a baseline understanding of knowing what is consuming an individual's time. As these steps or as more data is collected, AI-enabled tools will become more available as well as accurate.

Addressing Risk of Skill Decay

One of the concerns with using an AI-enabled tool to perform an acquisition task is the risk of losing competency in the given acquisition function. Simply, the function is being outsourced, which could lead to a loss of skills. The downside to automation is that the workforce is not sharpening their critical thinking skills and may lose the ability to determine the quality of the AI-delivered product. Senior government leaders will also need to consider how to approach ensuring that a workforce—augmented with AI-enabled tools—can still be effective if the tools experience any technical failures. While discussed within a medical context, the National Institute of Health concluded in a 2024 study that multidisciplinary research is needed to 1) understand the potential cognitive consequences of leveraging AI, 2) design artificial intelligence systems to mitigate skill decay, and 3) develop training and use protocols to prevent negative impacts on users' cognitive skills (Macnamara et al., 2024). Building a AI-literate workforce is contingent upon finding the right balance between embedding AI and preserving human judgment and safeguarding the foundational capabilities that humans are well-suited for.

Education

Education is and will continue to play an instrumental role in convincing the acquisition workforce that new AI tools can ease their workload, minimizing the cultural resistance to change. Emphasizing AI literacy among the DoD civilian and military acquisition workforce will help build trust between users and AI systems. Partnerships between government and industry, Federally Funded Research and Development Centers (FFRDCs), nonprofit organizations, or academia can help introduce AI tools, deliver training, and overcome cultural barriers. Low-risk opportunities for individuals to engage with AI tools will help them learn how to interact with new AI tools and build trust in the technology. The goal should be to understand and determine how to increase user trust in AI systems in the same way they would trust other pieces of software commonly used for acquisition functions.

Build Modern Data Infrastructure

To bridge the gap from what is technologically possible today to future applications, DoD will need to improve its data infrastructure and overall data collection efforts. The National Security Commission on Artificial Intelligence's final report noted that DoD currently lacks the digital infrastructure in place to support the development of AI tools (2021). Moreover, data are critical to the development and operation of AI-enabled capabilities, but according to various DoD officials, the department often lacks data that are usable for AI. High performing AI typically requires accurately labeled historical data to train the system. While DoD recognizes this need is incongruent with historical data collection practices, it is beginning to adopt collection, storage, and cleaning processes needed to train models.

To accelerate the adoption of advanced AI-enabled capabilities, DoD will also need to prioritize more computing power to support the processing of AI's algorithms and data. Part of the feedback received after the launch of NIPRGPT focused on the shortfalls in compute in DoD's high-performance computers and commercial vendors (Harper, 2024).

Recommendations for Harnessing AI's Transformative Potential

As AI continues to reshape industries globally, its transformative potential within the federal acquisition system has never been more significant. Drawing on recent insights from the GSA, DoD, National Science Foundation, emerging best practices from AI research and policy experts, and insights from the joint NDIA ETI-MITRE symposium in October 2024, this paper provides actionable guidance for modernizing federal acquisition systems by leveraging AI. This paper outlines several strategic recommendations aimed at ensuring the successful integration of AI technologies in defense and acquisition practices. The recommendations focus on a holistic approach to AI adoption that includes early engagement, iterative development of AI tools, comprehensive training, policy reform, infrastructure investments, and fostering cross-sector collaboration.

Cultivate Collaboration and Stakeholder Engagement

Successful integration of AI-enabled tools in the federal acquisition space requires strong, coordinated leadership and collaboration across multiple sectors, including government, industry, and academia. Agency leadership should prioritize creating cross-sector partnerships that bring together diverse expertise, from technological development to ethical considerations. These partnerships should have a mandate to accelerate the development of AI solutions for acquisition process so that they address practical needs. The National AI Initiative Office is a good example of supporting this approach, advocating cross-sector partnerships to advance AI technologies in the defense sector, based on identified needs.

The successful integration of AI technologies within the federal acquisition system hinges on early engagement with key stakeholders, including both the user and policy-making communities, in government agencies, defense contractors, policy-makers, and technology developers. Engaging stakeholders early in the research, development, prototyping, experimentation, and testing phases is critical to building a comprehensive understanding of the benefits and risks of using AI in acquisition as well as breaking down the cultural resistance to AI adoption. Early engagement can identify potential concerns, such as suitability for use, ethical considerations, security risks, and workforce impacts, while also creating opportunities for collaboration in designing AI solutions that meet the specific needs of the acquisition workforce.

Analyze the Acquisition Life Cycle

All federal agencies must truly understand workflows, organizational incentives, and workforce issues to understand where to prioritize development efforts and technology insertions. This information will be presented to senior leaders on where it makes sense to prioritize the use of AI, the best way to go about educating the workforce to use AI and considering AI as an assistive tool rather than a replacement for the acquisition professional.

Previous sections of this paper broke down at a very high level the contracting life cycle and gave examples of where AI could be inserted. It is recommended that federal agencies break this down even further and look at where it would make sense to potentially pilot the use of AI and where it would make sense to educate and test in potential agency labs. This should be done side-by-side with AI subject matter experts who understand the intricacies of the use of AI and where it would make sense to insert new tools.

An example of this that is being done by several agencies currently was the identification of how AI could be leveraged to complete more impactful market research. These agencies identified that if the use of AI were to fail or not be that effective in piloting it in the market research phase that this failure could be learned from and have less impact on the acquisition. For example, the Department of Homeland Security use the market research phase to gain buy-in from the workforce while also educating the workforce on the use of different AI tools and how to properly develop things like prompts that are important for the successful use of these tools (DHS, n.d.). By starting small and inserting AI-enabled tools in lower risk areas, agencies help cultivate user buy-in and mitigate potential risks to workflows.

Pilot Programs and Innovation Activities

The federal government struggles with change if senior leaders at the implementing agencies don't allow acquisition professionals to learn, for example through pilot initiatives, and most importantly tolerate failure in some efforts. Pilots should be designed with end uses in mind, a strategy to transition them into broader use across the enterprise, appropriate resources for executing the pilot, and a comprehensive set of metrics, data collection activities, and independent evaluation processes to judge the value of the pilot to federal goals and missions.

Leadership plays a crucial role in driving the adoption of AI technologies. Support from top leadership increases the likelihood that AI initiatives receive the necessary resources, attention, and momentum to succeed. This includes advocating for AI within the broader organizational strategy, fostering an environment that encourages experimentation, and providing adequate funding and resources for AI-related initiatives and their transition across the enterprise.

One of the crucial ways senior leaders can show full buy-in from the top down is the establishment of dedicated innovation activities which would perform research and prototyping activities on technologies, practices, and new fundamental concepts that would enhance the capabilities of the federal acquisition workforce, including through the use of AI. Congress should draft a provision requiring DoD to show it a plan to use commercially available tools to address targeted areas for improvement in the contracting life cycle (e.g., market research, J&A paperwork, Contractor Performance Assessment Reporting (CPAR, or RFIs). In DoD's plan, there should be strict criteria to evaluate the success of these pilots and a plan to scale the tools, if appropriate. Such activities, if they incorporate both intramural and extramural efforts, would allow different users to test AI technologies in practical settings before full-scale implementation. The DoD stresses the value of pilot programs in refining AI technologies before large-scale deployment but sometimes fails to capture lessons learned from pilots or have strategies to transition successful activities.

Innovation labs serve as dedicated spaces for testing and prototyping new AI-driven ideas. These labs allow teams to experiment with AI applications in a controlled environment, providing valuable insights that can inform future development. Leadership support is critical for creating an atmosphere of innovation, where failure is seen as a learning opportunity rather than a setback.

Pilot programs should be designed to demonstrate quick wins and generate confidence in AI's value. Regular evaluation of these pilots will allow for iterative improvements, ensuring that AI tools are continually refined and adapted to meet the unique needs of the federal acquisition process. Pilot programs can provide vital feedback and improve the alignment of AI tools with real-world operational requirements.

Transition strategy

Institute Training Programs and Foster AI Literacy

For AI to be fully harnessed within federal acquisition, comprehensive training programs must be developed to ensure AI literacy and technical competence across the workforce. In some cases, this means understanding the strengths and limitations of what AI can do while being an informed manager of AI inputs and consumers of AI outputs. Not everyone needs to know how to develop and use AI, but everyone needs to be literate enough to ensure proper usage. This is essential for both acquisition professionals and technical personnel. For example, the Defense Acquisition University could consider developing a standardized AI in Acquisition certification for contracting officers/specialists, program managers, and acquisition executives. This could be mandated for all acquisition professionals, ensuring AI becomes an integral part of the contracting life cycle and would allow acquisition professionals to begin their training in simulation-like environments. In other cases, selected people need skills in building or using AI, particularly in data analysis and AI tool usage.

Training programs should be designed with varying levels of expertise in mind, from foundational AI literacy for general employees to specialized courses for data scientists, analysts, and acquisition experts. These initiatives should emphasize practical knowledge, focusing on the application of AI tools within the acquisition life cycle. In addition to formal training, mentoring programs and internal workshops could foster a culture of continuous learning and skill development.

The aim is to create a workforce that is not only proficient in using AI tools but also capable of adapting to the rapidly evolving landscape of AI technologies. Building a culture of AI literacy across the federal workforce is fundamental to ensuring the effective adoption and deployment of AI tools. AI literacy training initiatives should be designed not just for technical teams but for all employees involved in decision-making processes. These programs should educate staff about the fundamentals of AI, its potential applications, and its ethical implications.

AI literacy initiatives will empower workers to understand the capabilities and limitations of AI, enabling them to make more informed decisions when interacting with AI-powered systems. The goal is to create a workforce that is not only capable of using AI tools but also prepared to make critical decisions regarding AI's implementation, security, and ethical considerations. The Artificial Intelligence Workforce Act supports the expansion of AI literacy programs across government sectors (Peters & Schmitt, 2024).

Invest in Infrastructure and Data Governance

One of the biggest takeaways from the symposium was that for AI adoption to succeed, significant investments in infrastructure and robust data governance policies are necessary, including data sharing and access across the DoD. AI tools require advanced computational resources, cloud services, and secure networks to function effectively. The federal government must invest in AI-supportive infrastructure, including high-performance computing systems and data storage solutions that meet the demands of AI applications. The National Institute of Standards and Technology (NIST) provides guidelines for developing secure and ethical data governance structures necessary to support AI (NIST, 2024). Along with this NIST is creating AI Risk Management Frameworks, ARM enables organizations to tailor risk management practices to their unique needs to coincide with the governance necessary (NIST, 2024).

Additionally, effective data governance is crucial to ensure that AI tools are trained and operated using high-quality, secure, and compliant datasets. Establishing standards for data collection, usage, and sharing will help mitigate risks related to data privacy and security while enabling AI to deliver actionable insights. Furthermore, a commitment to data interoperability across agencies will promote more cohesive AI-driven solutions. It cannot be overstated that the key to successful implementation of AI tools is the data that the AI will be ingesting to produce its outputs. To support this effort, Congress could request that DoD create a S&T portfolio to support the development of acquisition tools intended to improve its acquisition and IT performance.

A great example of this type of investment is the DoD's Joint All-Domain Command and Control (JADC2) initiative, which made key investments in AI-supportive infrastructure. It aims to modernize IT systems and improve data accessibility, enabling seamless data sharing across different military branches. JADC2 relies on cloud computing to provide the infrastructure needed for AI tools to analyze vast amounts of data in real-time. Data governance policies under JADC2 ensure that data across the DoD is consistent, secure, and usable by AI systems, facilitating better decision-making and operational efficiency (Obis, 2023). These are the type of investments and implementations that the federal government needs to continue to make if it wants to be successful with AI, and that was evident throughout the symposium and in the multiple interviews that were conducted.

Establishing Standardized Evaluation Metrics

Establishing a standardized set of AI evaluation metrics is crucial for fostering trust, resiliency, accuracy, and effective monitoring of AI systems. These metrics provide a consistent framework for assessing AI performance, ensuring that systems operate reliably and transparently across diverse applications. By standardizing evaluation criteria, stakeholders can more easily compare and validate AI models, promoting accountability and confidence in AI technologies. Furthermore, standardized metrics facilitate the identification and mitigation of biases and errors, enhancing the resilience and accuracy of AI systems. This structured approach to evaluation also supports ongoing monitoring and improvement, enabling AI systems to adapt to evolving challenges and maintain their integrity over time.

Final Encouragement: Building an AI-Driven Acquisition System

It is essential to recognize that the successful integration of AI within federal acquisition is not just about adopting new technologies; it is about building a resilient, future-proof system that can adapt to evolving challenges. AI has the potential to radically improve the way the government acquires goods and services, from streamlining acquisition processes to enhancing decision-making capabilities. To achieve this, we must take decisive action now, embracing innovation, collaboration, and continuous learning. This includes implementing the many published recommendations that have been provided in a strategic and logical fashion that will allow for the true adoption of the use of AI in the acquisition life cycle. The call to action for all federal agencies is now as AI tools and capabilities continue to evolve daily, and the more collaborative learning and implementation that can occur around these tools and capabilities, the more successful the federal government will be in harnessing its potential.

Bibliography

Bistarkey, D. (2024, July 26). *AI-powered agile talent identification systems support joint force*. Air Force. https://www.af.mil/News/Article-Display/Article/3851781/ai-powered-agile-talent-identification-systems-support-joint-force/

Commission on Planning, Programming, Budgeting, and Execution Reform. (2024). *Defense resourcing for the future*. https://ppbereform.senate.gov/wp-content/uploads/2024/03/Commission-on-PPBE-Reform_Full-Report_6-March-2024_FINAL.pdf

Defense Acquisition University. (n.d.). *Inherently governmental functions*. DAU. https://www.dau.edu/acquipedia-article/inherently-governmental-functions

Defense Logistics Agency. (2020). *DLA Fiscal Year 2020 historical report*. https://www.dla.mil/Portals/104/Documents/Headquarters/History/Fiscal2020AnnualHistory.pdf

Department of Defense Chief Information Officer. (2024). *ATO 101 for small businesses*. https://dodcio.defense.gov/Portals/0/Documents/Library/ATO-101SmBusinessInfo.pdf

Department of Defense. (2024). *Continuous authorization to operate (cATO) evaluation criteria: DevSecOps use case*. https://dodcio.defense.gov/Portals/0/Documents/Library/cATO-EvaluationCriteria.pdf

Gates, S. M., Esteves, F., Roth, E., & Kempf, J. (2024). *Implementation of the new Defense Acquisition Workforce Improvement Act framework: End of Fiscal Year 2022 update*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA758-3.html

Harper, J. (2024, November 7). *How things are going with the Air Force's experimental NIPRGPT chatbot*. Defensescoop. https://defensescoop.com/2024/11/07/air-force-niprgpt-experimental-chatbot-how-things-are-going/

Macnamara, B. N., Berber, I., Çavuşoğlu, M. C., Krupinski, E. A., Nallapareddy, N., Nelson, N. E., Smith, P. J., Wilson-Delfosse, A. L., & Ray, S. (2024). Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? *Cognitive Research: Principles and Implications*, *9*(46). https://doi.org/10.1186/s41235-024-00572-8

National Institute for Standards and Technology. (2024). *Secure software development practice for generative AI and dual-use foundation models*. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-218A.pdf

National Institute for Standards and Technology. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile*. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

National Security Commission on Artificial Intelligence. (2021). *Final report*. https://assets.foleon.com/eu-central-1/de-uploads7e3kk3/48187/nscai_full_report_digital.04d6b124173c.pdf

NetImpact Strategies. (2024, June 25). *The role of AI in transforming federal acquisition*. https://netimpactstrategies.com/insights/point-of-views/the-role-of-ai-in-transforming-federal-acquisition/

Nicewick, A. (2024, June 10). *Generative AI and the DOD*. Defense Acquisition University. https://www.dau.edu/sites/default/files/2024-06/June102024-AI%20presentation%20.pdf

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. https://www.science.org/doi/10.1126/science.adh2586

Obis, A. (2023, May 15). *Pentagon developing data integration layer to enable JADC2*. GOVCIO Media & Research. https://govciomedia.com/pentagon-developing-data-integration-layer-to-enable-jadc2/

Obis, A. (2024, May 16). *DoD's acquisition workforce is stretched thin*. Federal News Network. https://federalnewsnetwork.com/defense-main/2024/05/dods-acquisition-workforce-is-stretched-thin/

Pharathikoune, S. (2025, March 7). *Enhancing military operational effectiveness through the integration of CAMO and NIPR GPT*. U.S. Army. https://www.army.mil/article/283601/enhancing_military_operational_effectiveness_through_the_integration_of_camo_and_nipr_gpt

Secretary of the Air Force Public Affairs. (2024, June 10). *Department of the Air Force launches NIPRGPT*. Air Force. https://www.af.mil/News/Article-Display/Article/3800809/department-of-the-air-force-launches-niprgpt/

Technology Workforce Framework Act of 2024, S. 3792, 118th Cong. (2024). https://www.congress.gov/118/bills/s3792/BILLS-118s3792is.pdf

U.S. Army Public Affairs. (2024, October 18). *Army launches pilot to explore generative AI for acquisition activities*. U.S. Army. https://www.army.mil/article/280500/army_launches_pilot_to_explore_generative_ai_for_acquisition_activities

U.S. Congress. (2021, January 1). FY21 National Defense Authorization Act (Public Law 116–283). https://www.congress.gov/116/plaws/publ283/PLAW-116publ283.pdf

U.S. Congress. (2025, December 11). FY25 National Defense Authorization Act (Public Law 118-31). https://www.congress.gov/bill/118th-congress/senate-bill/4638/text

U.S. Department of Homeland Security. (n.d.). *Artificial intelligence for market research contract*. https://www.dhs.gov/sites/default/files/2022-12/AI%20for%20MArket%20Research%20DHS-gov.pdf

U.S. Government Accountability Office. (2021). *Artificial intelligence: An accountability framework for federal agencies and other entities*. https://www.gao.gov/assets/gao-21-519sp.pdf

U.S. Government Accountability Office. (2018). *Defense contracting: Enhanced information needed on interagency contracting* (GAO-18-578).

PANEL 12. STRENGTHENING DEFENSE TIES: COOPERATION, INNOVATION, AND PARTNERSHIP

Wednesday, May 7, 2025	
1240 – 1355 PT 1440 – 1555 CT 1540 – 1655 ET	**Chair: Raymond D. Jones, Chair and Professor of Practice, Department of Defense Management, Naval Postgraduate School** ***Enhancing Defense Industrial Cooperation Between Australia and the United States*** Katy Buda, Associate Director, Center for Strategic and International Studies ***Competing at the Upstream of Innovation: The US-China Balance in Critical Minerals*** Nathan Picarsic, Chief Operating Officer, Horizon Advisory ***Improving Arms Sales, Technology Transfer, and Defense Industrial Cooperation with Allies and Partners*** Audrey Aldisert, Research Associate, Center for Strategic and International Studies

Raymond D. Jones—retired as a Colonel from the U.S. Army in 2012 and is a Chair and Professor of Practice with the Department of Defense Management at the Naval Postgraduate School. His last assignment in the Army was as the Deputy Program Executive Officer for the Joint Tactical Radio System (JTRS). Additionally, he served as the Military Deputy for the Director of Acquisition Resources and Analysis in the Office of the Under Secretary of Defense for Acquisition Technology and Logistics (USD(AT&L)), managed three Major Defense programs for the DoD in addition to his many operational and research and development assignments. He graduated from the U.S. Naval Test Pilot School in 1995 and is 1983 graduate of the United States Military Academy. He has a Bachelor of Science degree in Aerospace Engineering, a Master of Science Degree in Aeronautical Engineering from the Naval Postgraduate School, a Master's in Business Administration from Regis University, a Master's Degree in National Resource Strategy from the Industrial College of the Armed Forces and is currently a PhD candidate with the Graduate School of Information Sciences at the Naval Postgraduate School in Monterey California.

Enhancing Defense Industrial Cooperation Between Australia and the United States

Cynthia R. Cook—is director of the Defense-Industrial Initiatives Group and a senior fellow in the Defense and Security Department at the Center for Strategic and International Studies. Cook is a member of the editorial board for the Defense Acquisition Research Journal and is an adjunct professor at the Pardee RAND Graduate School. From 1997 to 2021, Cook worked as a senior management scientist at RAND, where she served as the director of the Acquisition and Technology Policy Center. She holds a PhD in Sociology from Harvard University and a BS in Management from the Wharton School of the University of Pennsylvania. [ccook@csis.org]

Henry H. Carroll—is a research associate with the Defense-Industrial Initiatives Group at the Center for Strategic and International Studies (CSIS). His analytical focuses include the defense industrial base, naval policy, and assessing foreign military industrial capacity. Prior to joining CSIS, he worked as an intern in the defense and international practices of Brownstein and as a defense-focused legislative intern for Senate Majority Leader Chuck Schumer. Henry holds a BA in History, with a concentration in war and society, from Yale University. His undergraduate senior thesis examined the politics of U.S. naval shipbuilding in the interwar period. [hcarroll@csis.org]

Katy Buda—is an associate director with the Defense-Industrial Initiatives Group at CSIS. She previously interned at the U.S. Army War College. She holds an MA in International Studies with a focus in security from the Korbel School at the University of Denver. Katy completed an undergraduate degree at Temple University, where she studied political science and French. [kbuda@csis.org]

Abstract

The alliance between Australia and the United States is becoming increasingly salient as threats have evolved and challenges have multiplied. In the Indo-Pacific region, the primary threat to regional stability is China, but Australia and the United States are well postured to work in partnership against global threats. Defense industrial cooperation is a distinct interest of both partners, with public statements from leaders culminating in the AUKUS agreement. Drawing on desk research, interviews with approximately 30 business organizations in the both nations and with government personnel, and outputs of two discussion workshops, this paper gives an overview of the goals of the bilateral Australia-U.S. partnership with a focus on strengthening the defense industrial ecosystem of emerging technology acquisition between Australia and the United States. The strategic imperative of enhancing defense industrial cooperation between Australia and the United States requires a response rooted in a clear understanding of the specific challenges of this bilateral relationship combined with a broader mastery of the strategic imperatives of both countries, the acquisition process, and the numerous obstacles to any form of defense industrial cooperation. Simple solutions and single policy changes (i.e., “fix ITAR”) are not going to yield the desired results. A longer-term plan for change management—with a focus on sharing the strategic vision, providing resources and training, continually looking to identify and address barriers, and highlighting wins, can enhance cooperation outcomes. A plan to measure and track cooperative activities will provide a useful metric that can be used to assess whether policy changes are having an effect.

Introduction

Australia and the United States have had a close partnership for over a century, fighting side by side in every major war since World War I. The next potential conflict that could involve both countries working together is in the Indo-Pacific. Over the past few years, China has become increasingly aggressive with its stated intentions and its actions, and concern over China’s posturing is laid out in each nation’s strategic planning documents. In the United States, the National Security Strategy characterizes China as the U.S. military’s pacing challenge, with the People’s Republic of China (PRC) as “the only competitor with both the intent to reshape the international order and, increasingly, the economic, diplomatic, military, and technological power to advance that objective" (Biden, 2022). Australia's 2024 National Defense Strategy identified Chinese actions in the Indo-Pacific—and especially in the South China Sea—as directly endangering regional stability and Australia's national security interests, while also emphasizing Australia's role as an active shaper of the Indo-Pacific strategic environment. Australia and the United States are well postured to work in partnership against global threats.

Both countries' visions recognize the importance of a strong industrial base and how working with allies and partners can contribute to this goal. Turning these ideas into action will require time and attention from senior policymakers to translate the strategic vision into tactical-level steps to identify and fund requirements and overcome existing barriers. Necessary actions include facilitating critical technology transfers, allowing for the sharing of classified and sensitive information, and providing opportunities for defense firms to enter each other's markets. There have been important recent steps to address some of the policy challenges, such as changes to the International Traffic in Arms Regulations (ITAR) regime, which is designed to limit the proliferation of advanced technology to problematic actors. On September 1, 2024, the AUKUS exemptions to ITAR went into effect, granting Australia and the United Kingdom the same privileged status within the U.S. defense industrial base (DIB) as Canada, after the State Department determined that the export control systems of both countries are "comparable" to those of the United States (Cooper, 2024). The reforms add an expedited licensing process for exporting some defense articles to AUKUS partners and increase the scope of exemptions for transferring defense articles to some dual nationals who have security clearances from Australia and the United Kingdom (State Department, 2024).

However, there has been less research on the experiences that private enterprises face when trying to do business in the partner nation. If the U.S. and Australian governments truly want to encourage defense industrial cooperation, they should address the full spectrum of challenges and adjust policies accordingly. The research analyzed in this report aims to support this goal by providing a framework to better understand barriers and catalysts to cooperation. This framework builds off existing research by diving deeper into the defense industrial cooperation component of security cooperation, assessing the challenges faced by industry, and providing recommendations to both government and industry.

Methodology and Report Outline

Drawing on research, interviews, and workshops, this paper gives an overview of the goals of the U.S.-Australia alliance, with a focus on strengthening the entire defense industrial ecosystem. The project team reviewed policy documents focused on the strategic goals of the partnership, the breadth and depth of defense industrial ties between Australia and the United States, and the laws, policies, and cultural barriers that limit these ties, including the export control regimes of both countries.

Interviews included discussions with government personnel on both sides of the Pacific. The team gathered data from engagements with government representatives at the working and senior levels. These engagements included a conference in Canberra, Australia, and a follow-on conference in Washington, D.C (Cook et al., 2024a; Cook et al., 2024b). Each conference included private track 1.5 dialogues to establish space for forthright discussion, and the twin conferences allowed both Australian and U.S. stakeholders to have the opportunity to take center stage. Some of the industry executives interviewed for the project had previously worked in government, allowing them to offer perspectives from both viewpoints.

The team also conducted interviews with business leaders from Australian and U.S. companies on their experiences doing business in the partner nation. The interviewees included representatives from nine large U.S. defense contractors, eight of which are doing business with Australia. The team also spoke with representatives from 19 Australian companies. Only one of these had decided not to pursue business with the United States because of perceived challenges. The other 18 are either doing business with the United States or are trying to break into the market. These companies were identified through two means. The Austrade representative in the Australian embassy in Washington offered a list of companies that they worked with on marketing to U.S. and global industry (Austrade, n.d.). To avoid any bias by using a sample defined by the Australian government, the study team also worked with a Washington-based consultant who advised Australian companies more generally, including those trying to break into the U.S. market.

Current Strategic Challenge and Relationship

While the United States and Australia share a long history of deep cooperation, the contemporary strategic environment is propelling both countries to pursue a significantly expanded and deepening set of initiatives. The rapidly deteriorating strategic environment has been a key driver in strengthening the U.S.-Australia alliance in recent years, with an increasingly assertive China, and its ever-more-aligned partner Russia, standing at the forefront of this challenge. This is a fight the United States and its like-minded partners and allies can win—or, better yet, can avoid through deterrence. There is no place more promising to start than by revamping the U.S.-Australia alliance to ensure that it can both align national strategic visions and also produce the vital defense equipment needed.

Fortunately, there is already a strong history of alignment upon which to build. Australia's close strategic relationship with the United States has translated into support in the U.S. Congress and from the president for closer industrial relationships. This was shown in 2017 with the expansion of the U.S. National Technological and Industrial Base (NTIB) to include Australia and the United Kingdom. All three countries also hold a bilateral Reciprocal Defense Procurement Agreement (RDP-A) with the United States, which means that "Buy American" provisions do not apply to U.S. Department of Defense (DoD) purchases from those nations. Along with New Zealand, these countries also enjoy a close intelligence-sharing relationship with the United States through the Five Eyes arrangement.

The arrangement that has received the most attention recently is the AUKUS partnership, announced in 2021 (Vaughn, 2023). AUKUS, including its two pillars, is designed to counter twenty-first-century threats through enhanced technology partnerships. Pillar I is designed to enhance Australia's military capability with a new fleet of conventionally armed nuclear-powered submarines. The Pillar II technology partnership focuses on new technologies, including artificial intelligence and autonomy; quantum technologies; hypersonic, undersea, and advanced cyber capabilities; electronic warfare; and innovation (The White House, 2022).

Understanding Defense Cooperation

A variety of public statements and formal policies affirm the strength and endurance of the defense relationship between Australia and the United States. This relationship fits into a broader frame of U.S. defense cooperation, which the DoD defines as: "a generic term for the range of activity undertaken by DoD with its allies and other friendly nations to promote international security. Such activity includes . . . security assistance, industrial cooperation, armaments cooperation, Foreign Military Sales (FMS), training, logistics cooperation, cooperative research and development (R&D), Foreign Comparative Testing (FCT), and Host-Nation Support (HNS)" (Defense Acquisition University, n.d.).

One highlighted approach is industrial cooperation, which can take a variety of forms. McGinn's 2023 study highlights five pathways for a "Build Allied" approach to defense industrial cooperation: 1) an increase in the number of U.S. subsidiaries of foreign defense companies, 2) co-development of systems or subsystems across two or more countries, 3) co-production of defense systems across two or more countries, 4) second-sourcing or licensed production to qualify multiple producers for the same part or system, and 5) foreign sustainment (maintenance, repair, or overhaul) of existing systems (McGinn, 2023, p. 5).

Defense industrial cooperation across national borders requires catalysts—reasons for action—as well as deliberate efforts to overcome barriers. One approach to specify barriers is offered by Jennifer D. P. Moroney et al. in a 2023 report on U.S. security cooperation, which provides a typology of barriers that impede U.S. security cooperation with highly capable allies and partners in the air, space, and cyber domains (Moroney et al., 2023). The authors examine defense industrial cooperation as part of the broader framework and note that it is subject to numerous barriers. The author's research and framework focus on security cooperation, which is led by government actors. Adding information derived from CSIS interviews with industry conducted as part of this research, this report expands the framework to include a broader range of issues which government policies may help resolve. Table 1 includes an overview of barriers and is an extension of the framework offered in the work of Moroney et al. The figure includes the addition of the economic barriers experienced by industry, as derived from the literature and informed by this project's interviews.

Table 1. Barriers to Security Cooperation
(Adapted from Moroney et al., 2023)

Type of Barrier	Representative Examples
Budgetary	Differences in funding priorities or availability of resources Inability to determine or agree to fair share (costing requirements)
Bureaucratic	Sheer number of stakeholders and organizations Over-classification of communications (default to NOFORN) Conflicting priorities and incentives within U.S. and partner organizations
Cultural	Differing approaches or expectations regarding military cooperation Reluctance or inability to share sensitive or classified data Historical experience in bilateral or multilateral engagements/relationships
Political	Government restrictions or limitations external to a nation's defense department Domestic pressures or influences from industry, legislatures, or popular opinion
Regulatory	Written prohibitions or limitations to collaboration in U.S. legal code, congressional legislation, or departmental instructions Ally/partner legal or executive-level restrictions on collaborations with foreign partners
Strategic	Diverging national interests and threat perceptions Differences in priorities concerning collaboration with the United States and other allies and partners
Technical	Lack of compatible systems or procedures to share information Imbalances in scientific or domain experience Lack of confidence in ally/partner's ability to effectively protect classified or sensitive information
Economic	*Insufficient business case to incentivize cooperation for industry *Cost of learning new, foreign acquisition system or setting up a subsidiary and office in the partner nation *Misaligned business strategies as companies prioritize different end markets and products

*Note: *New elements added by CSIS to the Moroney et al. framework.*

CSIS expanded Moroney's framework by including an "economic" barrier category with three representative examples. The first is an insufficient business case to incentivize industry to conduct international cooperation activities. The second identifies the costs—both financial and the opportunity cost of personnel time—associated with learning a new acquisition system in the partner country or the cost of acquiring or establishing a subsidiary company and offices in the partner nation. The third is the potential misalignment of corporate strategies between defense firms which prioritize different end markets and products, thereby hampering their ability to effectively cooperate with each other.

The Industry Perspective: Barriers to Cooperation

The policy benefits of defense industrial cooperation between Australia and the United States are well understood and have led to the many cooperation agreements. Despite those agreements, however, the interviews conducted for this project revealed numerous tactical-level barriers to cooperation, with roots both in the government and within industry itself.

Budgetary: Getting the Money, from Budgets to Contracts

Business organizations must continually consider their financial picture, and interviewees frequently mentioned budget concerns when working with both the United States and Australia. These concerns related both to budgets on the national level (making this a political issue as well) along with the timely and long-term funding of projects (overlapping with bureaucratic acquisition concerns).

For Australia, businesses' budgetary concerns centered on how government funding levels impacted the extent of what the government of Australia was able to invest in. For example, one U.S. company speculated that a large project was terminated because the Australian government wanted to put those resources toward AUKUS Pillar I nuclear-powered submarines instead. Another issue raised was a lack of consistent funding across governments, without which industry is less able to access the resources necessary to invest in production to scale and sustain production over time. One interviewee summarized this as "lots of promises are being made and not a lot of money is flowing." This impacts both Australian and U.S. companies working to do business in Australia. For the most part, the U.S. government does not face the same overall funding issues. Funding stability is mostly ensured once new requirements become programs of record, with the caveat that frequent continual resolutions limit new program starts. However, the delays caused by the budget process are seen as a barrier by both Australian and U.S. industry when doing business with the U.S. government, indicating that some of the challenges identified by industry when considering international cooperation projects are in fact generic challenges endemic to working in the defense industrial base.

It can also be difficult for companies to get longer-term, multiyear funding. More complex projects benefit from firm commitments for funding across multiple years. Companies may be eager to work together across national lines, but as the U.S. subsidiary of an Australian firm noted, "both governments struggle to give long-term contracts" which would enable this cooperation. Ramping up a supply chain and the workforce requires a multiyear commitment to be economically worthwhile, but those longer-term commitments are difficult to extract and run into political barriers within both nations.

There are solutions to all these issues, but they are challenging to enact. The Australia budget challenge could be addressed by the government consistently funding defense over time, but this is a policy that will depend on national government decisions. Another approach for the Australian government is to deliberately help Australian industry strengthen its export market to ensure a more consistent customer base over time. In fact, this is the approach being taken in the Guided Weapons Explosive Ammunition (GWEO) enterprise, which is focused on munitions production. An export-focused approach could build upon Australia's 2018 Defence Export Strategy, which advocated for Australia to become a top 10 defense exporter (Defence Australia, 2018).

Another challenge links budgeting and contracting processes. For small businesses, delays in finalizing a contract, which can relate to the government's availability of funds, can mean that they do not have the resources to pay their employees and, moreover, will otherwise interfere with the longer-term viability of the enterprise itself. Small enterprises frequently noted that "the government does not understand cash flow." Payment delays are disruptive to any business, but small and medium-sized enterprises (SMEs) may be particularly vulnerable. Furthermore, Australian firms often complain that banks and venture capital groups in Australia lack sufficient capital supply because of reliance on foreign sources, exacerbating critical balance-of-payment issues for Australian SMEs (Connolly & Jackman, 2017, p. 59). Finally, one DoD interviewee suggested that the United States has more funding streams available than Australia does to bring projects from early-stage science and technology into more advanced stages of development, using vehicles such as the Defense Innovation Unit or the Defense Advanced Research Projects Agency (DARPA). The individual recommended creating a dedicated pot of funds in Australia for investing in these early-stage efforts.

Other discussions raised the specifics of getting paid. One Australian company highlighted that a U.S. defense innovation organization's default payment method was via check, and that it cost more to get them to wire funds. Another complained that the aforementioned organization did not know what a SWIFT key was, delaying payment by eight months. This was reported to create months-long delays before the DoD could find the correct form to process the payments.

The challenge of industry being able to access stable funding with enough certainty to grow production to scale and scope is by no means unique to the international cooperation context, and it represents a challenge for business enterprises doing business for their own national governments as well as for those of partner nations. While strengthening relations with partner nations via defense industrial ties is not the main goal of any nation's budget process, the negative impacts on partners of budget perturbations are real and should be a consideration in deciding if industrial cooperation with allies is a true priority.

Bureaucratic: Navigating Complex Acquisition Systems

The U.S. defense acquisition system can be difficult to navigate, even for U.S. firms. This difficulty is magnified for small Australian firms and amplified by the lack of personal connections with decisionmakers, societal and cultural differences, and the tyranny of distance and tight travel budgets. Australian firms report finding it hard to understand who U.S. decisionmakers are and how to connect with them, a challenge also shared by small U.S. businesses. Interviews with Australian industry often noted that "if Australian companies don't understand the U.S. procurement system, they can't sell." Many of these Australian firms added that hiring U.S. advisers as guides to understanding the system is expensive, which serves as a deterrent for trying to make the jump into the U.S. defense ecosystem.

U.S. government organizations need a contract vehicle as a pathway to get funds to performers. Many Australian SMEs reported that U.S. contracting vehicles are difficult for them to use, requiring them to go through larger U.S. companies as resellers, which reduces the SMEs' profit margins and, therefore, their incentive to cooperate across the Pacific. The issue of the reseller dynamic arose several times as a complicating factor in Australian firms closing deals. An Australian firm going through a U.S.-based reseller is going to have longer lead times for contracts and may miss out on business opportunities due to speed (or lack thereof) rather than capability—an issue complicated and magnified by the already long lead times for export control licenses and security clearance processes. Another complicating factor is the U.S. acquisition workforce's unfamiliarity with acquisition processes for foreign firms, which leads some DoD acquisition staff to encourage Australian firms to work with resellers. This well-meaning advice may be given without fully considering the impact it could have on the speed of the Australian firm's subsequent contract and therefore the firm's business case for working with the DoD.

Australia does have a dedicated agency called Austrade whose mission includes helping Australian exporting companies grow their business. Austrade has representatives stationed in the United States, and according to this project's interviewees, the agency has provided useful guidance in navigating U.S. processes and providing information about tenders. One company noted that "We would not have been able to do what we did without Austrade." Another offered, "Austrade has been incredibly helpful—[they help companies] plug into shows and be part of delegations and make it easier to go to [conferences] like SeaAirSpace." On the other hand, other companies noted that Austrade has provided briefings on a less detailed (and therefore less useful) level and has not helped them navigate the U.S. market, largely due to Austrade lacking the necessary contacts with customers. Smaller businesses interested in exporting did not always know how to access Austrade's tools. Policymakers should consider providing additional support for the agency so that it could more effectively reach out to new defense companies interested in exporting. A throughline in Australian interviews was the necessity of education both for industry and the government in terms of the opportunities for partnership and the specific bureaucratic challenges that need to be overcome.

Cultural: Two Nations Divided by a Common Language and an Ocean

Cultural barriers were among the most common types of impediments mentioned by the interviewees and spanned across most of the pathways of connection in the defense industrial relationship. One challenge centers on the differences between corporate culture in Australia and the United States regarding considerations like self-promotion and seeking legal advice from counsel. According to both Australian and U.S. interviewees, Australian firms tend to be much less self-promoting when discussing their products with U.S. officials or businesspeople than Americans are used to, leading to occasional moments of mismatched expectations and underappreciated offerings from Australia. One interviewee noted that "Australians are not pushy. . . . They don't puff themselves up." Furthermore, Australian firms are much less likely than their U.S. counterparts to seek legal recourse or the advice of counsel when encountering regulatory difficulties in U.S. acquisition efforts or business-to-business (B2B) engagements, such as with export controls restrictions. This can hamper the ability of Australian firms to get contracts compared to U.S. entities. As one U.S. firm explained, "Cultural differences get overlooked between the U.S. and Australia. . . . We like to believe that they are very similar, but the cultures are very dissimilar. The similarities outweigh the dissimilarities, but they are very different cultures."

Another cultural barrier identified by interviewees relates to bureaucratic and regulatory issues. Both nations prefer local suppliers, which affects the relationship on both sides of the Pacific and limits cooperation. Even as the countries align closer on defense cooperation, and senior officials make statements and policies about the essential nature of AUKUS and U.S.-Australia cooperation, interviewees noted persistent parochial favoritism from mid-level officials in acquisition and business decision-making processes (DoD, n.d.; Garamone, 2023). This was pervasive in the United States, including with the DoD acquisition workforce and when partnering with U.S. companies. One Australian company noted, "Americans like things made in America, a culture which flows through to procurement people from corporate leadership." Government and industry preference for local suppliers is a known factor across global procurement, with the literature highlighting an even greater local preference by government than industry actors (Mulabdic & Rotunno, 2022).

Risk aversion is a persistent cultural challenge across the defense cooperation space, one often linked to other barriers, including regulatory ones. Australian interviewees noted that U.S. government officials are hesitant to green-light cooperation with Australia, even if they have the authority to do so. Working-level officials are perceived as adhering to the status quo, even as senior leaders advance new visions of cooperation (Henneke & Stephens, 2024). This is mirrored in industry, where companies are sometimes overly cautious despite encouragement from government officials with regulatory authority. This occurs even when considering sharing information already publicly available on corporate websites. One Australian company offered an example where a U.S. government agency approved the use of their products, but then lawyers at the prime contractor directed an additional review out of an excess of caution. One interviewee highlighted risk aversion as a serious problem when thinking about the competitive global landscape, which links cultural factors to political ones: “If China is our pacing threat, [we should] find things we can go jointly after. . . . Time is our enemy; we’re squandering opportunity. Only so long we can say PRC won’t catch up with us; they have smart people too. . . . Accept some amount of failure.”

An increased recognition of the strategic challenge may help the U.S. government embrace a global supply chain. Policies like the National Defense Industrial Strategy highlight the importance of allies and partners as part of an economic deterrence strategy (*National Defense Industrial Strategy*, 2024). Consistent messaging across administrations in both nations (which demands strong leadership support) is required to evolve government and industry cultures over time—though this alone may not be sufficient.

Political: Moving from Policy Announcements to Tactical Support

Industry in both nations observed that even with the necessary political support, translating policymaker intent into action has proven difficult. One interviewee observed that all politicians “love announceables” but that moving from the policy level into tactical execution was more of a problem. In spite of pronouncements, there was some cynicism as to whether national governments were truly behind cooperation. In general, businesses did not always see the next steps necessary to move the vision into action even though they thought the governments believed what they were saying.

One U.S. prime interviewee argued that the most likely business successes came from U.S. companies selling systems to Australia but getting beyond that was difficult, observing that the goals of highly publicized agreements like “AUKUS . . . [don’t] trickle down to the small companies.” An Australian firm echoed this same sentiment, stating, “The ambiguity around how these high-level strategic agreements translate into business opportunities is harmful to the Australian business community because they are making assumptions on how to export and if they lose money they may exit the business.” Several Australian interviewees expressed frustration with inconsistent signals from the United States, where government policies seemed to support cooperation but failed to result in business opportunities for Australian companies. One employee stated that “If the U.S. is just saying no to Australian products, that’s fine. . . . The problem is unclear messaging. [When] he talks to other defense companies . . . they all don’t have clarity.” The offered solution was that the ADOD “needs to be more direct about the need for the United States to understand and support Australia business,” which would depend on the DoD listening to and valuing this message from the ADOD over its other priorities.

There are government efforts in support of industrial cooperation, such as Australia's investment in the Global Supply Chain (GSC) Program and Austrade (covered above).[1] The GSC Program is an export initiative aiming to give Australian companies better opportunities to enter the supply chains of global defense primes by providing funding to "establish a team within their company dedicated to identifying export opportunities leading to contract award for Australian suppliers in both Civil and Defence businesses of the Prime."(Department of Defence, 2024a) Not every interviewee at the U.S. primes mentioned the GSC Program, but at least one found it to be a very useful support for bringing Australian firms into their supply chain. That said, in the interviews, many Australian SMEs expressed distrust of the commitment of the major U.S. primes—or their Australian subsidiaries—to incorporating Australian companies into their supply chains. The SMEs were concerned that U.S. primes use Australian SMEs as "window dressing" for their Australian government bids to comply with local preference regulations, only to squeeze the Australian SMEs out of these contracts later. One colorful Australian interviewee, describing the defense system in Australia, said that the "organ grinder is the [U.S.] primes, and the monkey varies between the government and the Australian industry."

One U.S.-based Australian interviewee was a former employee at a U.S. prime contractor and had also spent time working in the U.S. government. They suggested that that while the United States says it wants cooperation, what it really wants is "U.S. companies to split off units and do business for Australia. We are much less interested in building up native Australia companies—don't want them to be too competitive." Nevertheless, other U.S. defense prime figures repeatedly stressed their commitment to the GSC Program and noted the successes they had seen under it. U.S. industry individuals did note that the GSC Program was more effective when the Australian government had previously provided funding for Australian SMEs to qualify as subcontractors under the program, an initiative that could be revitalized.

Regulatory: Policy Underpins Many Cooperation Challenges

Government regulatory policy, including national export control regimes, shapes and limits defense industrial cooperation. Export controls have multiple goals, including limiting the export of sensitive military technologies that could find their way into the hands of adversaries. The challenge that export controls create is extensively highlighted in the literature on cooperation, but changes are being instituted to support AUKUS . In the U.S. system, there are at least 37 departments, agencies, and commissions with export control authority, including the Departments of State, Commerce, and the Treasury (U.S. Customs and Border Protection, 2024). The State Department handles ITAR compliance and the U.S. Munitions List for traditional military capabilities, and the Commerce Department enforces the Export Administration Regulations (EAR) and the Commerce Control List for dual-use technology (Kerr & Casey, 2021). The Defence Export Controls (DEC), part of the ADOD, oversees military and dual-use export controls through the Defence Trade Controls Act 2012, Defence Trade Controls Regulations 2013, and the Defence and Strategic Goods List (Australian Government, Department of Defence, n.d.). In March 2024, Australia changed legislation to place controls on the re-export of articles originally from Australia, information sharing on controlled technology areas to certain foreign persons in Australia, and various defense services, which went into effect on September 1, 2024 (Department of Defence, 2024b; Industry and Security Bureau, 2024). Following changes approved to Australian export controls, the U.S. State Department amended ITAR to provide licensing exemptions for Australia in technology areas not included

[1] One reviewer highlighted the existence of additional government programs designed to support cooperation, which were not raised during the interviews. More details can be found here: "Policy and Engagement," ADOD, accessed December 11, 2024, https://www.defence.gov.au/business-industry/industry-governance/industry-regulators/australian-defence-export-office/policy-engagement.

on the Excluded Technology List, which also went into effect on September 1, 2024 (Department of State, 2024). The U.S. Commerce Department implemented EAR changes as well, proposed and in effect by April 2024, enabling Australia to be treated in the same manner as Canada (Industry and Security Bureau, 2024).

The U.S. State Department's goal for export controls is to "mitigate diversion and proliferation risks, which both bolsters U.S. national security and contributes to regional and international security and stability"(Department of State, 2023). Interviewees universally understood and supported the goals for ITAR and other methods of information and export control regulation, but they highlighted that these regulations also create delays and other challenges for industrial cooperation. One U.S. prime specifically noted that the pace of the regulatory review did not match the pace of the acquisition cycle. Australian export controls are simpler, but export controls from each country are not the only regulatory barrier.

For Australian companies, fear of U.S. export control penalties can affect their business dealings with the United States. One U.S. company stated that "if they violate ITAR . . . [Australian firms] are worried about getting put out of business by a foreign regulator." Smaller vendors fear that they will be put out of business if they receive a penalty for violating ITAR. On the other hand, one larger Australian firm noted that the concern is often misplaced: "People think 'I am at the risk of going to jail,' but if they follow the process then that's just not going to happen." The small scale of many Australian firms amplifies the stifling effects of information security regulations, as their compliance teams and their financial margins for error are much smaller.

Protracted wait times can also be barriers to business. A U.S. prime interviewee indicated that if they wanted to work with an Australian SME, they could face production delays of up to 90 days while waiting for a license. They noted that "oftentimes those opportunities come and go within 90 days." There can also be holdups if a firm changes suppliers, and delays can be detrimental to smaller companies. One U.S. firm argued that "regulation shouldn't be easier [to navigate] for those with resources." Australian regulators have their own resource limitations and likely will need additional resources and funding following the enactment of new export controls to successfully implement these changes (Greenwalt & Corben, 2024).

Information sharing is another barrier for foreign suppliers interacting with U.S. primes, as foreign suppliers can face hurdles that in-country business dealings will never encounter. From the very start, conversations between a U.S. company and a foreign partner on sensitive topics can require ITAR approval—and the line between the two can be unclear, causing delays while this is determined. If the Department of State issues approvals only for a portion of the conversations needed, that is insufficient for building international cooperation. While some State Department personnel interviewed noted their progress in approving thousands of these approvals, a U.S. industry figure argued that the department needs "to approve millions to cover all the potential conversations—or else to change policy." The September 1, 2024, ITAR reforms between the United States and Australia may be able to abate this problem once companies become part of the "authorized user list," although there is still a range of excluded technologies that remain a concern.

The U.S. export and information control apparatus has led participants in the U.S. defense sector to proceed with caution when dealing with foreign actors, even if allies. For the United States, "they're built to never engage with a foreigner." While export control reforms have taken place in both Australia and the United States, there will likely be questions over who can operate license-free. Moreover, if businesses expect to need licenses and face delays, they could avoid certain suppliers or partners. Stakeholders need to allow time for industry to understand these changes and feel safe operating under the greater flexibility of a new export control policy. The regulatory barrier posed by export controls is therefore intertwined with the cultural barriers of risk aversion and resistance to change within the U.S. acquisition workforce and compliance departments in U.S. defense primes. One U.S. firm mentioned “ITAR/export controls focus on a thing or a defense article. Even if all of these are removed [after reforms], we need to accept that it will take time for folks to understand what it means and to execute it. The real barrier, when that’s gone, there will be a hesitation/blockage of the systems. It will be OK if it takes three to six months. But if it takes years…”

Another regulatory barrier relates to limitations on sharing information, including classification and controlled unclassified information (CUI).[2] Classification of information is meant to prevent damage to national security by controlling information release. The challenge this presents is not always visible to Australian companies, but their employees with military experience sometimes remarked that it is much easier to share information between the partner militaries than it is to access classified information on the industry side. This creates challenges for partnering on some national security projects and also sometimes limits Australian access to marketing opportunities in cases where tenders are classified. Even though it notionally presents less of a danger to national security, CUI was described as sometimes more difficult to handle than classified information because, while there are carefully established channels for classified information, the way to properly handle CUI is not as clear-cut. Interviewees also raised concerns about NOFORN markings prohibiting access to non-U.S. persons, which creates similar difficulties for access and also lacks a clear-cut and expeditious process for removal. One industry representative suggested that to enhance cooperation “YESFORN is the objective, NOFORN is the barrier.” Interviewees suspected that NOFORN labels are sometimes simply the default habit of an overly cautious acquisition and industrial workforce, rather than reflective of the contents.

Strategic: The Need for a Common View of the Challenge

The U.S. and Australian governments are closely aligned on their strategic outlooks, with both seeing China as the main strategic concern. However, one area where the United States and Australia have been reported not to see eye-to-eye on strategic issues is the ability of Australia to acquire certain U.S. systems. Interviewees noted that Australia often wants to purchase advanced U.S. technology that was still in the early R&D phase using the Foreign Military Sales (FMS) approach. The U.S. government, however, does not like to sell equipment that is still in early-stage R&D because, as one U.S. government official noted, “we don’t sell systems that won’t meet the requirements. . . . This is the problem with FMS, which is transactional, not meant to be flexible and have vision and work with different strategies and designs.” U.S. caution on FMS sales processes has occasionally clashed with Australian eagerness to acquire cutting-edge U.S. tech, a downstream problem from slight misalignments on strategy and timing of acquisitions.

Technical: Aligning Engineering Details

An ongoing challenge to cooperation is the existence of different technical standards and varying technical standards regimes. These differences may be physical incompatibilities, or there may be regulations that impose specific policies depending on the source. Differing standards can impede the ability of companies to work together and limit their ability to sell to partner governments or participate in partner supply chains. This issue was raised in several of the interviews as one of the tactical challenges that the interviewees did not see being

---

[2] Though not covered in detail here, CMMC cyber mandates for working with the DoD could also be a problem based on different standards than the Australian “Essential Eight” for cybersecurity. The issue of diverging standards is covered in the Technical Barriers section.

addressed in policy statements: "People want to get technology as fast as possible, [but] a company will have to produce a completely different variant for different [customers]." This challenge demonstrates that strategic vision in support of partnership cannot drive industrial cooperation without the identification and solving of specific challenges.

One of the examples raised during the interviews is the fact that the two nations have different regimes for the non-destructive testing of defense articles. Australia has the National Australian Testing Authority (NATA), but this authority is not recognized by the United States. Meanwhile, the United States has the National Aerospace and Defense Contractors Accreditation Program (Nadcap), an industry-led cooperative accreditation program for aerospace and defense industries. These two programs have different training of test operators and different standards for test success. Companies can put their articles through the other nation's test regimes, but this duplication of testing adds time and additional costs. One interviewee argued that standards could be assessed to see if they are close to those of the other nation; if they were, companies could then be permitted to test in only one nation while being certified in both. One Australian firm stated, "Standards is a small thing but could be impactful moving forward. Under AUKUS, there are working groups at technology levels, [but they have] not extended into standards . . . those are the big barriers."

Any deviation in standards also means that when technical details or parts need to be changed, even slightly, the part may need to be recertified. This increases the costs of non-recurring engineering, which can then increase the average cost of a part. One company raised the issue of different voltage standards as a source of recertification requirements (standard voltage in the United States is 120 V, Australia's is 240 V): "We meet a higher standard, but we don't meet the U.S. standard. Because the standard is slightly different, we have to recertify."

One interviewee noted that there is an AUKUS working group to address technical standards, but this challenge is complex and cross-cutting from the defense industrial base to the broader national manufacturing industrial base. To fully tackle this long-term challenge, the interviewee recommended that the working group should have the authority to create an action plan to address different standards, with the goal of continually identifying and addressing differences that create challenges for cooperation and cross-border sales. Other frameworks and organizations have found ways to establish common standards. One interviewee pointed out that NATO has a common standard, demonstrating that a solution across national borders is possible.

Finally, one interviewee raised a complication to overcoming this barrier: there may be benefits for incumbents in maintaining different standards, as it limits competition: "The primes developed this and don't allow outsiders in."

Economic: Making It Worthwhile to Invest

Economic challenges for industry are intertwined with their access to capital, the scale of their firm, and their corporate aspirations. Companies mentioned the challenge of finding funding, which plays out differently in the two economies. It is perceived to be easier to raise funds in the United States because of the size and strength of the venture capital community. In Australia, there is "skepticism in their own market," and the "investment community [is] skeptical about the U.S. market." Better funding from government could bridge this funding gap. One company argued that we "need to unlock funding . . . quickly to win opportunity—look at process, what Australia does well and what [the] U.S. does well, look at convergence of how we do funding and financing for small business." In the United States, the DoD's Office of Strategic Capital has been set up to try to "crowd in" private capital for defense and dual-use technologies, a model that could be adopted in Australia. One Australian firm mentioned, "We haven't seen a lot of actionable things come to market about workforce problems, and how we will share knowledge and skills across the two continents. . . . The broader population is completely unaware of what all the jobs and job opportunities will be. Joe plumber doesn't know about this. There is available talent that is disconnected from the demand."

Another issue highlighted was the challenge of scale, especially when partnering with, selling to, or competing with U.S. firms. Australian companies are often much smaller than their U.S. partners, and a deal that is a matter of survival for an Australian company may be relatively unimportant for their U.S. partner. One Australian firm noted that they "have to partner with U.S. companies and then [work] at the whim of their creative ambitions. [U.S.] companies were making huge gains in passing through the work." For a small company, the costs of the procurement process can be overwhelming: "How do I compete with a 10-person team at IBM focused on this Request for Proposal (RFP) when I'm three people and a cat." Industry as a whole faces strong government pressure to manage costs, but the customer can be oblivious to the economic repercussions of delays and other bureaucracy. One U.S. firm noted that "government regulators on both sides don't seem to be concerned about the economic impacts. They are focused on national security. . . . The Hill underestimates the scale of these transaction costs. Industry doesn't make the case effectively about the impact of delays." Small companies feel these burdens acutely.

Some Australian SMEs noted that small firms occasionally hold unrealistic expectations about what their role in U.S.-Australian B2B cooperation could look like given their size and production capacity. As a recent Australian Strategic Policy Institute report on AUKUS Pillar II cooperation notes, "size matters" in defense cooperation (Stephens, n.d., p. 18). Medium or large enterprises have greater scale than small ones, which allows them to more easily upscale their production when needed and to navigate complex bureaucratic tasks like export controls or acquisition. An Australian firm noted that being a "perennial smaller company working with larger companies," their firm faces the "challenge of being treated as a peer or equal. Larger or established companies, [find it hard] to take the reputational risk of partnering with a start up." Multiple Australian SMEs noted that they had more success working with small U.S. firms than with the primes. One ADOD official similarly commented that companies needed to find firms of similar sizes with which to match up. U.S. firms stated that they believe Australian companies are less willing than U.S. ones to partner up in order to conquer new, non-U.S. markets and are naive about what technology is required to compete globally.

Another barrier is a lack of industry alignment between Australia and the United States regarding the form and focus of possible future industrial collaboration. Indeed, some of the aforementioned Australian SMEs' lack of trust in U.S. primes and their frustration with the GSC Program may stem from misaligned expectations of what Australian firms and U.S. firms can productively partner on. The GSC Program identified second sourcing and exploiting innovative technology as potential inroads for Australian firms into prime supply chains, with success seeming to have been found more in the latter than the former (Department of Defence, 2024a). Interviewed companies reported that it is very difficult to bring in Australian firms as new second-source suppliers due to non-recurring engineering costs, the difficulties of technology transfer, and a high learning curve leading to a production cost well above target price. In short, Australian firms struggle to compete against entrenched, typically U.S., firms. In DoD contracting, even when second sources of products or systems are identified, there is no guarantee they will be cost competitive, and those which are competitive are often U.S. firms that do not face the same challenges as Australian SMEs (Adjei & Hendricks, 2022).

In connection with the cultural barrier of parochialism, more than one Australian company highlighted the necessity of being both aggressive and sensitive to local issues to successfully enter the U.S. market. Simple steps such as registering a website with a ".com" address instead of a ".com.au" address may overcome initial U.S. suspicion about working with foreign suppliers and lead to companies getting past an initial screen. Australian businesspeople will need to travel to the United States and do their own marketing. Another Australian firm stated that "they [Austrade] won't do business for you—you have to knock on doors, build [your] own pipeline [and sell] into [your] own market. . . . I look for what events are happening and go to as many as possible to meet in person." Oftentimes, success relies on the personal determination of entrepreneurs.

One last barrier, mentioned in almost every interview, is the "tyranny of distance." Australia and the United States are half a world away from each other. Vastly different time zones mean that connecting on the phone during business hours is a challenge, and the flights take almost a day in both directions and cross the International Date Line. There is no easy solution for this. The fact that industry continues to press to cooperate despite this challenge is a signal of a broad commitment to cooperation.

Recommendations

The strategic significance of enhancing defense industrial cooperation between Australia and the United States requires a response grounded in a clear understanding of the specific challenges of this bilateral relationship combined with a broader mastery of the acquisition process and the numerous obstacles to any form of defense industrial cooperation. These barriers are often mutually reinforcing. For these reasons, some topics, such as standards, appear under multiple categories of barriers, and many recommendations have implications that go beyond the barrier that they primarily address.

Budgetary and Technical Recommendations

For the United States and Australia, increasing the speed and quantity of defense production is crucial to addressing the threats that have been identified by their respective national strategies. Achieving these goals will be expensive, even with the opportunity provided by rising defense budgets. Australia recognizes that its desired sovereign defense industrial capabilities cannot be sustained without integration into other defense ecosystems and funding streams. The U.S. National Defense Industrial Strategy correctly identifies greater commonality with partners as an imperative. In short, commonality is an area where industrial integration can and should have a return on investment that offsets fiscal barriers to cooperation. The recommendations below take aim at the technical obstacles to commonality, which in turn will aid in justifying the budgetary investments in cooperation.

1. The United States and Australia should endeavor to align their requirements for new weapons systems or produce shared requirements, where possible, especially in the context of AUKUS.

2. Groupings such as AUKUS or the overlapping members of Five Eyes and the NTIB should be used as venues for implementing shared standards working groups.

Regulatory and Bureaucratic Recommendations

Regulatory and bureaucratic barriers inevitably add friction to international cooperation. To alleviate some of these obstacles, the U.S. Congress legislated a partial ITAR exemption for AUKUS countries, only the second such exemption in existence. This AUKUS ITAR exemption is a generational shift in U.S. export and technology control policies, moving from mandating licensing to requiring tracking for a range of technology in the territories of Australia, the United Kingdom, and the United States. Some analysts, such as William Greenwalt and Tom Corben, call for further liberalization in the next steps of implementation (Greenwalt & Corben, 2024). The interviews for this project took place before the implementation language was released;

therefore, the recommendations below do not seek to evaluate the implementation of the AUKUS ITAR exemption itself but instead focus on findings from the interviews and data that remain relevant in this rapidly changing environment.

1. The United States and Australia should enact equivalency agreements that recognize that certain defense standards are close enough to be mutually acceptable, even if these standards are not made in common.

2. The United States and Australia should mutually recognize each other's accreditation standards regimes where the requirements are close enough to be functionally interchangeable.

3. A joint procurement vehicle, such as a pan-AUKUS panel, could be created to deepen AUKUS collaboration, especially in the key areas of AUKUS Pillar II.

4. The United States and Australia should fully embrace mutual recognition of security clearances, within necessary parameters.

5. The U.S. Congress should explore passing legislation which would connect the AUKUS and Canadian ITAR exemptions to allow cross-compatibility.

Cultural, Political, and Strategic Barriers Recommendations

Legislative and regulatory changes alone are often not sufficient to effectively overcome barriers to cooperation. Cultural change is often necessary to fully institutionalize new authorities or integrate regulations into everyday practice. In the absence of cultural change, new authorities and strategic objectives may be hindered by inertia. Cultural change can be further hampered by competing messages, such as when the United States' Buy American Office was launched just before AUKUS materialized (The White House, 2021).[3] Measures such as these create an acquisition culture of defaulting to what is known, streamlined, and easy. Overcoming this cultural default requires not just direction from senior leaders but also consistent reinforcement at all levels and periodic evaluation of success.

1. The DoD and ADOD should conduct an audit of the implementation and outcomes of industrial cooperation efforts in preparation for each annual Australia–United States Ministerial Consultation (AUSMIN).
2. The DoD and ADOD should furnish an annual report on the implementation of cross-national industrial integration initiatives to their respective legislatures.
3. The U.S. Department of State should conduct a rigorous and proactive outreach campaign to inform industry about the specific requirements of the 2024 new AUKUS waiver; Austrade should establish a pipeline to refer companies to the Defence Export Controls office to provide clear messaging and education to industry about ITAR rules and boundaries.

Economic Recommendations

Economic barriers to defense industrial cooperation are mainly the product of uncertain returns for vendors supporting international cooperation, which make it difficult to justify addressing the forms of friction introduced by borders. A common kind of friction is the difficulty of incorporating a foreign supplier into an established supply chain or the costs of establishing a subsidiary and building contacts in a distant foreign nation. Insufficient incentives undercut the

[3] These steps do not directly contradict each other because the U.S.-Australia reciprocal defense procurement arrangements mean that the two countries exempt one another from national preference laws.

role the U.S. and Australian defense industrial bases can play in support of defense cooperation. Barriers to international cooperation can resemble those of small or non-traditional U.S. vendors. The lack of a perceived "front door" access for AUKUS applicants complicates what would otherwise be a comparatively easy path for Australian vendors that have established a U.S. subsidiary.

1. Defense industry groups in the AUKUS countries should consider creating an AUKUS-focused consortium.
2. The governments of Australia and the United Kingdom should consider subsidizing the overhead costs of establishing these AUKUS-focused consortia for AUKUS Pillar II topics.
3. The U.S. acquisition workforce should use Other Transaction Authority arrangements for AUKUS acquisition coordination.
4. The Australian government should empower and provide additional funding to Austrade to enhance Australian industry understanding of the U.S. acquisition system and to aid Australian SMEs in establishing U.S. domestic subsidiaries.
5. The DoD should expand the training and education of its acquisition workforce to include the financial and bureaucratic complexities of working with international companies.
6. The DoD and ADOD should embrace Modular Open Systems Approaches to lower barriers to entry and encourage competition.

Conclusion

Enhancing defense industrial cooperation between Australia and the United States will take concerted efforts by government and industry from both nations. Ensuring that business practices and industry and government culture support cooperation will require both resources and a thoughtful requirement-setting process that enhances opportunities for collaboration. Simple solutions and single policy changes (e.g., "fix ITAR") alone will not yield the results desired. To enhance cooperation outcomes, both systems must commit to a longer-term plan for change management, with a focus on sharing the strategic vision, providing necessary resources and training, and continually identifying and addressing barriers. Measuring and tracking cooperative activities can provide a useful metric to assess whether policy changes are having the desired effect. Real and sustained change can only start once policymakers embrace a mindset that believes time is of the essence and that approaches the challenges of national and allied preparedness with a sense of urgency (Kotter International Inc., n.d.). Given new collaborations between Beijing, Moscow, Pyongyang, and Tehran, as well as the rapidly deteriorating security conditions globally and in the Indo-Pacific region specifically, allied deterrence must transform into a collective endeavor. Such an enterprise necessarily demands more integration and alignment of defense systems and industries. AUKUS provides a superb opportunity to expand defense industrial collaborations by revitalizing the U.S.-Australian alliance with a laser-like focus on industrial policy. The national security of both counties, and the stability of the region, might well depend on it.

References

Adjei, S., & Hendricks, C. (2022). Increasing defense contractor competition in a predominately sole source contracting environment [Thesis, Acquisition Research Program]. https://dair.nps.edu/handle/123456789/4529

AUKUS Defense Trade Integration Determination. (2024, August 15). United States Department of State. https://www.state.gov/aukus-defense-trade-integration-determination/

Austrade. (n.d.). Austrade, Australian Government. Retrieved September 5, 2024, from https://www.austrade.gov.au/

Australian Government, Department of Defence. (n.d.). Export controls framework: Legislation. business & industry. Retrieved September 12, 2024, from https://www.defence.gov.au/business-industry/exporting/export-controls-framework/legislation

Biden, J. (2022, October 12). National security strategy. The White House. https://www.whitehouse.gov/briefing-room/statements-releases/2022/10/12/fact-sheet-the-biden-harris-administrations-national-security-strategy/

Connolly, E., & Jackman, B. (2017). The availability of business finance.

Cook, C. R., Edel, C., Myler, P, Chynoweth, G, & Burton, B. (2024a, April 5). Strengthening Australia-U.S. defence Industrial cooperation: Keynote panel. CSIS Events. https://www.csis.org/events/strengthening-australia-us-defence-industrial-cooperation-keynote-panel

Cook, C. R., Edel, C., & Jeffrey, H. (2024b, March 5). Strengthening Australia-U.S. defence industrial cooperation: Keynote by Deputy Secretary Hugh Jeffrey. CSIS Events. https://www.csis.org/events/strengthening-australia-us-defence-industrial-cooperation

Cooper, R. C. (2024, September 15). On the third AUKUS anniversary, a toast to ITAR reform and a call to keep going. Atlantic Council. https://www.atlanticcouncil.org/blogs/new-atlanticist/on-the-third-aukus-anniversary-a-toast-to-itar-reform-and-a-call-to-keep-going/

Defence Australia. (2018). Defence export strategy. https://www.defence.gov.au/business-industry/industry-governance/industry-regulators/australian-defence-export-office/australian-defence-export-office/export/strategy

Defense Acquisition University. (n.d.). Defense cooperation. Retrieved September 17, 2024, from https://www.dau.edu/glossary/defense-cooperation

Department of Defence. (2024a, August 5). Global supply chain program. Australian Government: Defence. https://www.defence.gov.au/business-industry/industry-capability-programs/global-supply-chain-program

Department of Defence. (2024b, September 3). Changes to export controls. Commonwealth of Australia; Defence. https://www.defence.gov.au/business-industry/exporting/export-controls-framework/changes-export-controls

Department of Defence. (2024c, December 11). Australian defence export office policy and engagement. https://www.defence.gov.au/business-industry/industry-governance/industry-regulators/australian-defence-export-office/policy-engagement

DoD. AUKUS defense ministers' joint statement. (n.d.). Retrieved September 12, 2024, from https://www.defense.gov/News/Releases/Release/Article/3733790/aukus-defense-ministers-joint-statement/https%3A%2F%2Fwww.defense.gov%2FNews%2FReleases%2FRelease%2FArticle%2F3733790%2Faukus-defense-ministers-joint-statement%2F

Garamone, J. (2023, July 29). Australian, U.S. leaders say alliance is more relevant than ever. U.S. Department of Defense. https://www.defense.gov/News/News-Stories/Article/Article/3476054/australian-us-leaders-say-alliance-is-more-relevant-than-ever/https%3A%2F%2Fwww.defense.gov%2FNews%2FNews-Stories%2FArticle%2FArticle%2F3476054%2Faustralian-us-leaders-say-alliance-is-more-relevant-than-ever%2F

Greenwalt, D. W., & Corben, T. (2024, August 21). AUKUS enablers? Assessing defence trade control reforms in Australia and the United States. https://www.ussc.edu.au/aukus-assessing-defence-trade-control-reforms-in-australia-and-the-united-states

Henneke, G., & Stephens, R. (2024). AUKUS Pillar 2 critical pathways: A road map to enabling international collaboration [Special report]. Australian Strategic Policy Institute. https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2024-05/SR206%20AUKUS%20Pillar%202%20critical%20pathways_0.pdf

Henneke, G., & Stephens, R. (2024, May 9). AUKUS pillar 2 critical pathways: A road map to enabling international collaboration. http://www.aspi.org.au/report/aukus-pillar-2-critical-pathways-road-map-enabling-international-collaboration

Industry and Security Bureau. (2024, April 19). Export control revisions for Australia, United Kingdom, United States (AUKUS) enhanced trilateral security partnership. Federal Register. https://www.federalregister.gov/documents/2024/04/19/2024-08446/export-control-revisions-for-australia-united-kingdom-united-states-aukus-enhanced-trilateral

Kerr, P. K., & Casey, C. A. (2021). The U.S. export control system and the export control reform act of 2018.

Kotter International Inc. (n.d.). The 8-Step Process for Leading Change. Retrieved October 31, 2024, from https://www.kotterinc.com/methodology/8-steps/

McGinn, J. (2023, June 26). A “build allied” approach to increase industrial base capacity. George Mason University. https://business.gmu.edu/news/2023-06/build-allied-approach-increase-industrial-base-capacity

Moroney, J. D. P., Pezard, S., Thaler, D. E., Germanovich, G., Grill, B., McClintock, B., Schwindt, K., Adgie, M. K., Binnendijk, A., Connolly, K. J., Feistel, K., Hornung, J. W., Hottes, A. K., Kim, M., Nazha, I., Tarini, G., Toukan, M., & Zeman, J. (2023). Overcoming barriers to working with highly capable allies and partners in the air, space, and cyber domains: An exploratory analysis. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA968-1.html

Mulabdic, A., & Rotunno, L. (2022). Trade barriers in government procurement. European Economic Review, 148, 104204. https://doi.org/10.1016/j.euroecorev.2022.104204

State Department. (2024, August 20). International traffic in arms regulations: Exemption for defense trade and cooperation among Australia, the United Kingdom, and the United States. Federal Register. https://www.federalregister.gov/documents/2024/08/20/2024-18043/international-traffic-in-arms-regulations-exemption-for-defense-trade-and-cooperation-among

U.S. Customs and Border Protection. (2024, March 7). U.S. Government Agencies with Export Requirements. https://www.cbp.gov/trade/aes/pga/other-agencies

United States Department of State. (2023, July 10). Myths and facts about U.S. defense export controls: Fact sheet. https://www.state.gov/myths-and-facts-about-u-s-defense-export-controls/

Vaughn, B. (2023). Australia: Background and U.S. relations. Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R47378

The White House. (2021, January 25). President Biden to sign executive order strengthening buy American provisions, ensuring future of America is made in America by all of America’s workers. https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/25/president-biden-to-sign-executive-order-strengthening-buy-american-provisions-ensuring-future-of-america-is-made-in-america-by-all-of-americas-workers/

The White House. (2022, April 5). Fact sheet: Implementation of the Australia–United Kingdom–United States Partnership (AUKUS). https://www.whitehouse.gov/briefing-room/statements-releases/2022/04/05/fact-sheet-implementation-of-the-australia-united-kingdom-united-states-partnership-aukus/

Competing at the Upstream of Innovation: The US-China Balance in Critical Minerals

Nathan Picarsic—is a senior fellow at the Foundation for the Defense of Democracies, a senior visiting fellow at the Krach Institute for Tech Diplomacy, and a co-founder of Horizon Advisory. [nate.picarsic@post.harvard.edu]

Emily de La Bruyère—is a senior fellow at the Foundation for the Defense of Democracies, a senior visiting fellow at the Krach Institute for Tech Diplomacy, a non-resident fellow at the National Bureau of Asian Research, and a co-founder of Horizon Advisory. [ebde@alumni.princeton.edu]

Abstract

The United States and China are in the throes of a long-term, peacetime competition. That contest has, thus far, centered on science, technology, and industry. The means deployed have been non-kinetic: export controls, investment restrictions, market protections, and trade remedies. Critical and strategic materials – the upstream of legacy and emerging technological applications – have figured prominently in these peacetime salvos. China and the United States have very different capabilities in critical and strategic minerals. They also have very different approaches to the domain. This paper provides an overview of the asymmetries in strategic orientation defining the critical mineral postures of the US and China; the threats that those asymmetries pose to the United States; and the role that the defense acquisition system can play in facing down those threats.

China has an upper hand in critical and strategic minerals. Beijing has proven its willingness to use that upper hand offensively. And China is investing, disproportionately – vis-à-vis its broader science and technology program – in early-stage innovation in minerals and materials that could lock in the PRC's advantage and disrupt American downstream strengths. This reality poses a direct national security and economic security threat to the United States. Beijing's market control, pricing power, and distortive effects are such that extant market forces cannot resolve the threat within the current economic order. Despite the severity of this strategic challenge, the defense acquisition system can strengthen US defenses, support and direct early-stage research and development to enhance US strengths, and, ultimately, position to impose costs on China-tied supply chains.

Introduction

China treats supply chains – and, especially, the upstream of supply chains – as the core elements of geopolitical competition in today's globalized world. And Beijing considers its current positioning in critical supply chains to be a core asset in its arsenal for confronting the United States; a trump card vis-à-vis America's leadership in cutting-edge technologies.

Chinese government discourse is explicit about the country's supply chain strategy. The Chinese Ministry of Transport has stated that, "enterprise competition is no longer a competition among individual companies, but rather among supply chains." Xi Jinping himself explained in in 2016 that:

> If a company is heavily dependent on foreign countries for its core components, and if the 'major artery' of the supply chain is in the hands of others, it is like building a house on someone else's foundation. No matter how big and beautiful it is, it may not stand up to wind and rain, and it may be so vulnerable that it collapses at the first blow. (Xi Jinping, 2016)

The Chinese government – in both its discourse and its resource allocations – is also explicit about prioritizing critical and strategic minerals as it competes for supply chains. This prioritization encompasses access to strategic minerals, industrialization of them, and leapfrog innovation in both the materials science used to produce them and their applications. In order to enhance its power, Beijing wants independence in these materials and the geopolitical leverage that comes from the dependence of the rest of the world.

As early as 1986, the State Council published the Mineral Resources Law of the People's Republic of China, declaring that “the development, utilization and protection of mineral resources should adhere to the leadership of the Communist Party of China” to “implement the overall national security concept.” Over the decades since, Beijing has secured strongholds in global mineral supply. Beijing has also invested in cutting-edge research and development in strategic minerals intended to cement next-generation leadership in the field, and foster leapfrog capabilities in their applications. And China has shown willingness to use its strategic mineral capabilities for coercive ends. In 2010, amid a territorial dispute with Tokyo over the Senkaku Islands, Beijing temporarily ceased exports of select, critical rare-earth elements to Japan. (Keith Bradsher, 2010)

Both China’s prioritization of strategic minerals and its framing of them as offensive assets have only increased in recent years. In November 2024, the 12th Meeting of the Standing Committee of the 14th National People’s Congress adopted a new version of the Mineral Resources Law, to come into effect in July 2025. The new Mineral Resources Law makes clear that strategic minerals are inputs into national security: “Mineral resources are an important material basis for economic and social development, and the exploration and development of mineral resources are related to the national economy, people's livelihood and national security.”

And the actions that Beijing has taken in parallel with drafting and issuing the Mineral Resources Law make clear that as inputs into national security, strategic minerals are offensive as well as defensive assets. Since 2024, in response to US tariffs and technology restrictions, China has implemented its own export controls on critical minerals including gallium, germanium, and antimony as well, more recently, as restrictions on various rare earth elements. (Amy Lu, 2025)

Beijing’s ability and willingness to implement these export controls reflects an asymmetry vis-à-vis the United States. America has begun to recognize the importance of secure, independent supply chains, and of critical minerals in those, in its competition with China. But this recognition lags that of the PRC. And Beijing’s multi-decade advance -- as well as its centralized state system, industrial capacity, and natural resource advantages -- have allowed China to secure a clear upper hand in the strategic minerals contest. Moreover, the United States tends to focus on critical minerals as an area in which to play defense; to protect against Chinese dominance. Beijing by contrast uses strategic minerals for offensive ends.

This dynamic creates obvious threats for the United States. Beijing has leverage, over minerals critical for both security and commercial applications. The Chinese government has refined mechanisms for using that leverage for coercive effect. And China is investing in early-stage breakthroughs that could both lock in China’s upstream advantage and disrupt America’s downstream strength. Moreover, Beijing is adept at leveraging its pricing power, technological advantage, and full value-chain approach in critical minerals to undermine emergent US efforts to establish truly independent, domestic supply chains. China’s broad-based dominance grants Beijing varying degrees of veto power over American efforts to unleash market forces to solve for its current weaknesses.

And considering both the security relevance of the threat at hand and the impossibility of relying on China-distorted markets to face it down, defense acquisition processes and actors have important roles to play countering China’s critical minerals threat. Those roles should include doubling down on defense, as for example with stockpiling of critical minerals. They should also include investments in early-stage research and development. And the defense acquisition system can incorporate upstream vulnerabilities into program requirements and performance metrics to shift market incentives away from China.

The Critical Minerals Landscape: A US Disadvantage

Both Washington and Beijing have promulgated policies defining, respectively, “critical” and “strategic” minerals. In 2016, China’s Ministry of Land and Resources published the National Mineral Resources Plan. (Ministry of Land and Resources, 2016) That document presented a list of 24 strategic minerals, including energy, metallic, and non-metallic minerals. (Though the list also groups all “rare earth” together as one. Because there are 17 rare earth minerals, Beijing might more accurately be described as having identified 40 strategic minerals.)

A year later, pursuant to “Executive Order 13817A Federal Strategy to Ensure Secure and Reliable Supplies of Critical Minerals,” the US began to publish lists of “critical minerals.” The most recent such list, published in 2022, covered 50 minerals. (US Geological Survey, 2022)

If rare earth minerals are counted individually, 27 of the minerals on the US list of critical minerals also appear on China’s of strategic minerals, as reflected in the table below, while 23 are unique to the US list. And 12 of China’s “critical minerals” are captured in the US list. One obvious and notable discrepancy lies in China’s inclusion of “energy minerals” – or oil, natural gas, shale gas, coal, coal bed methane, and uranium – under strategic minerals, while the United States only captures metallic and non-metallic minerals.

Strategic and Critical Minerals, as Defined by China and the US[1]		
China: Strategic Minerals		**US: Critical Minerals**
Energy minerals	Oil, natural gas, shale gas, coal, coal bed methane, uranium	**Aluminum, antimony,** arsenic, barite, beryllium, bismuth, **cerium**,* cesium, **chromium, cobalt**, **dysprosium**,* **erbium**,* **europium**,* **fluorspar**, **gadolinium**,* gallium, germanium, **graphite**, hafnium, **holmium**,* indium, iridium, **lanthanum**,* **lithium**, **lutetium**,* magnesium, manganese, **neodymium**,* **nickel**, niobium, palladium, platinum, **praseodymium**,* rhodium, rubidium, ruthenium, **samarium**,* **scandium**,* tantalum, tellurium, **terbium**,* **thulium**,* **tin**, titanium, **tungsten**, vanadium, **ytterbium**,* **yttrium**,* zinc, and **zirconium**.
Metallic minerals	Iron, **chromium**, copper, **aluminium**, gold, **nickel**, **tungsten**, **tin**, molybdenum, **antimony**, **cobalt**, **lithium**, **rare earths**, **zirconium**	
Non-metallic minerals	Phosphorus, potash, **crystalline graphite**, **fluorspar**	

No matter the set of strategic or critical minerals adopted, China is clearly better positioned in terms of both access to the minerals themselves and production through midstream processing. The United States is more than 50 percent import dependent in 38 of the minerals that it has identified as critical. In five of the remaining 12 cases, there is insufficient data to assess US import dependence. The US is also 45 percent import dependent in copper and 93 percent in potash, both of which China defines as “strategic minerals,” though the US does not. And in 30 of its 50 critical minerals – including 25 of the 38 in which it is more than 50 percent import dependent – the US relies on China as one of its critical suppliers. By contrast, according to available figures, chromium is the only strategic mineral for which China is essentially completely reliant on foreign imports. (US Geological Survey, 2024).

[1] Bolded words are minerals that are listed as strategic or critical by both countries. Rare earths are marked with an asterisk.

US Import Dependence and Sources in Prioritized "Critical Minerals"			
Prioritized by	**Mineral**	**US Import Dependence (%)**	**Major US import sources (2020-2023)**
China/US	**Aluminum**	47	Canada, UAE, Bahrain, China
China/US	**Antimony**	85	China, Belgium, India, Bolivia
China/US	**Cerium**	80	China, Malaysia, Japan, Estonia
China/US	**Chromium**	77	South Africa, Kazakhstan, Canada, Finland
China/US	**Cobalt**	76	Norway, Finland, Japan, Canada
China/US	**Dysprosium**	80	China, Malaysia, Japan, Estonia
China/US	**Erbium**	80	China, Malaysia, Japan, Estonia
China/US	**Europium**	80	China, Malaysia, Japan, Estonia
China/US	**Fluorspar**	100	Mexico, Vietnam, South Africa, China
China/US	**Gadolinium**	80	China, Malaysia, Japan, Estonia
China/US	**Graphite**	100	China, Canada, Mexico, Mozambique
China/US	**Holmium**	80	China, Malaysia, Japan, Estonia
China/US	**Lanthanum**	80	China, Malaysia, Japan, Estonia
China/US	**Lithium**	>50	Chile, Argentina
China/US	**Lutetium**	80	China, Malaysia, Japan, Estonia
China/US	**Neodymium**	80	China, Malaysia, Japan, Estonia
China/US	**Nickel**	48	Canada, Norway, Australia, Brazil
China/US	**Praseodymium**	80	China, Malaysia, Japan, Estonia
China/US	**Samarium**	80	China, Malaysia, Japan, Estonia
China/US	**Scandium**	100	Japan, China, Philippines
China/US	**Terbium**	80	China, Malaysia, Japan, Estonia
China/US	**Thulium**	80	China, Malaysia, Japan, Estonia
China/US	**Tin**	73	Peru, Bolivia, Indonesia, Brazil
China/US	**Tungsten**	>50	China, Germany, Bolivia, Vietnam
China/US	**Ytterbium**	80	China, Malaysia, Japan, Estonia
China/US	**Yttrium**	100	China, Germany
China/US	**Zirconium.**	<25	South Africa, Australia, Senegal
US only	**Arsenic**	100	China, Morocco, Malaysia, Belgium
US only	**Barite**	>75	India, China, Morocco, Mexico
US only	**Beryllium**	0	
US only	**Bismuth**	89	China, Republic of Korea
US only	**Cesium**	100	Germany, China
US only	**Gallium**	100	Japan, China, Germany, Canada

US only	**Germanium**	>50	Belgium, Canada, China, Germany
US only	**Hafnium**	NK	Germany, China
US only	**Indium**	100	Korea, Japan, Canada, Belgium
US only	**Iridium**	NK	
US only	**Magnesium**	>75	Israel, Canada, Turkey, Czechia
US only	**Manganese**	100	Gabon, South Africa, Australia, Malaysia
US only	**Niobium**	100	Brazil, Canada
US only	**Palladium**	36	Russia, South Africa, Belgium, Italy
US only	**Platinum**	85	South Africa, Belgium, Germany, Italy
US only	**Rhodium**	NK	
US only	**Rubidium**	NK	
US only	**Ruthenium**	NK	
US only	**Tantalum**	100	China, Australia, Germany, Indonesia
US only	**Tellurium**	<25	Canada, Philippines, Japan, Germany
US only	**Titanium**	86	South Africa, Madagascar, Canada, Australia
US only	**Vanadium**	40	Canada, Brazil, Austria, South Africa
US only	**Zinc**	73	Canada, Mexico, Republic of Korea, Peru

A Difference in Strategic Orientation: Offense vs. Defense

Perhaps more important than the specific minerals identified by the US and China – and even relative dependencies in them – is the difference in the two sides' strategic orientations toward strategic and critical minerals. Washington approaches the field with a defensive posture. United States policy defines "critical minerals" as those with supply chains vulnerable to disruption; the United States invests to enhance its access to critical minerals but not to limit that of China, or other adversaries; and in its technological strategy, the United States de-prioritizes investment in research, development, and innovation related to minerals and materials.

Beijing's approach is the opposite. China's orientation toward strategic minerals is an offensive one. China's definition of "strategic minerals" includes both those that are weaknesses and those that are strengths. China has shown a capacity and political will to wields its mineral advantages for coercive effect. And China prioritizes investments in the cutting-edge of mineral and material research and development – in order to establish enduring leadership over the field.

The Definitional Divide

China's 2016 National Mineral Resources Plan defined strategic minerals based on their value to industry and to national security: They are the minerals necessary to "safeguard national economic security, national defense security and the development needs of strategic emerging industries."

The 2017 "Executive Order 13817A Federal Strategy to Ensure Secure and Reliable Supplies of Critical Minerals presented a contrasting US framework for identifying critical minerals. Like the Chinese definition, that framework hinges on importance to national security and industry. But it also hinges on vulnerability. Critical minerals, per the United States, are those that are "(1) are "essential to the economic and national security of the United States," (2) have supply chains that are "vulnerable to disruption," and (3) serve "an essential function in the manufacturing of a product, the absence of which would have significant consequences for our economy or our national security."

Those policy definitions, published at almost the same time by the US and China, align on the value that critical or strategic minerals provide; these are necessary inputs for both security and industry. But the definitions differ in their relative emphases on vulnerability. For a mineral to be included in the US framing of critical minerals, it must be vulnerable to supply chain disruption. It must be a defensive concern. That is not the case for China. Beijing includes in its definition not only minerals on which it depends on international players, but also those in which it dominates the global market or has offensive capabilities. The Chinese approach centers more fully for the objective importance of a given material – and leaves room to prioritize those minerals that Beijing can use to coerce.

That PRC orientation is not limited to the National Mineral Resources Plan. Researchers at the Institute of Mineral Resources under the Chinese Academy of Geological Sciences argued in a 2021 paper published in Acta Geoscientica Sinica, a journal associated with the China Geological Survey, that the criteria for judging the strategic value of a mineral should include not only its economic significance and import dependence, but also whether it has "international market advantages and certain bargaining power and have important uses in strategic emerging industries."

That difference between the US and Chinese orientations toward strategic and critical minerals is borne out in the lists of specific minerals identified by the two countries. Of the critical minerals on the US list, there is only one, beryllium, in which America is a net exporter. By contrast, Beijing includes rare earths in its set of strategic minerals, despite the country's' dominance in the field: China accounts for 60 and 87 percent, respectively, of global rare earth production and processing.

Contrasting Postures

The asymmetry in orientation is not simply rhetorical. It has also concretely informed the manner in which the two countries have wielded their relative mineral positioning. US activities in critical and strategic minerals have tended to be purely defensive – and largely reactive at that. Washington has focused on investing to shore up weaknesses and dependencies vis-à-vis China. And even in those areas where the United States *does* have an upper hand in critical mineral value chains, the country has at no point leveraged, or threatened to leverage, that advantage. For instance, even as America has begun to treat China as an adversary, and even as Washington has imposed export controls and trade barriers on Beijing, there has at no point been any threat to limit the export of beryllium, a critical input into aerospace, nuclear, and medical fields – and for which China depends on imports from the United States. (World Bank, 2025)

Beijing, on the other hand, has consistently, over the past quarter century, leveraged its supply chain advantages for offensive effect – and not just against the United States. As early as 2010, amid a territorial dispute with Tokyo over the Senkaku Islands, Beijing temporarily ceased exports of select, critical rare-earth elements to Japan. In hindsight, that move was an early clue as to how Beijing would wield influence globally. Over recent years, as tensions between China and the United States have escalated, Beijing has again and again retaliated against Washington by limiting the export of critical materials. In 2024, after the United States placed restrictions on the export of advanced semiconductor technology to China, China responded by imposing export restrictions on gallium, germanium, and antimony to the United

States. All three are critical inputs into semiconductors – and US-defined critical minerals – in which China holds a globally dominant position. And in 2025, after President Trump placed a new round of tariffs on imports from China, Beijing retaliated by implementing export controls on a host of minerals and mineral products, including samarium, gadolinium, terbium, dysprosium, lutetium, scandium, and yttrium, as well as their alloys and oxides.

Back in 2010, when restricting rare earth exports to Japan, China did so surreptitiously. Beijing denied at the time that it was leveraging international trade – and its positioning at the upstream of strategic value chains – for coercive, geopolitical effect. Today, Beijing is explicit about its activities. Beijing frames export controls of critical minerals very clearly as retaliation; Beijing describes its advantage in them as a trump card in evolving security and industrial competition.

Asymmetric Investment Patterns

The difference in the two sides' strategic orientation toward critical minerals is further borne out in how they invest in the field – and, especially, in relevant research and development (R&D). Broadly speaking, across tech areas generally, the United States tends to prioritize early stage R&D and corresponding innovation far more than does China. Beijing instead emphasizes refinement, application, and scaling of proven technologies. Government budget figures bear out this difference. In 2024, less than seven percent of China's R&D spending went to basic research. For the US, that figure is some three times as high: The US FY2023 budget allocated 23 percent of R&D dollars to basic research. (Ministry of Commerce, 2024)

But in critical and strategic minerals, and upstream materials more generally, the relative prioritization of basic research and development flips. The United States eschews basic R&D in the field. There is very little emphasis – in either US technology or US mineral policy – on upstream-relevant research and development. Instead, discussion of and policy in the field focuses on identification of, development of capacity in, and development of processing capabilities for resources known to be valuable for contemporary use cases. The same holds for investment. For instance, Department of Defense research, development, test, and evaluation (RDT&E) funding goes to component development and downstream complete systems, not to upstream inputs even in fields of investment focused on the earliest stage of basic science. (Congressional Research Service, 2022)

Beijing's investments follow a very different pattern. China, disproportionately vis-à-vis the rest of its technology program, invests in early-stage innovation in minerals and materials, with the stated ambition of capturing the leading heights in and shaping the direction of the field, as well as its downstream applications. This prioritization is spelled out in the 14th Five Year Plan for National Economic and Social Development – the guiding, central document detailing Beijing's ambitions and plans for the 2021 to 2035 period. That plan explicitly lists "basic materials" as a "key and core technology to prioritize," and "new materials" as a "pillar of the industrial system." It also asserts that China will focus on technological breakthrough in the field:

> We will promote breakthroughs in advanced metals and inorganic non-metallic materials such as high-end rare earth functional materials, high-quality special steels, high-performance alloys, high-temperature alloys, high-purity rare metal materials, high-performance ceramics, and electronic gases; we will strengthen the research and development and application of carbon fiber, aramid fiber, and other high-performance fibers and their composite materials and bio-based biomedical materials; and we will accelerate breakthroughs in key technologies of high-performance resins, such as metallocene polyethylene, and high-purity electronic materials such as photoresists for integrated circuits.

Resource allocations bear out this framing. More than 30 of China's National Key Laboratories focus specifically on metals and/or materials. And their work covers basic research and development. The National Key Laboratory of Rare Metal Special Materials, for instance, pursues "original theoretical and subversive application technology research," in order to "make technical breakthroughs in stuck neck targets" and "creat[e] a leading position in China and even in the world." The National Key Laboratory of New Technologies for Intensified Metallurgy of Nonferrous Metals funds both theoretical and applied work, including research on the "theory of efficient enrichment and enhanced separation of complex copper, lead, and zinc resources;" "basic research on solid waste resource utilization and urban mine development;" and "basic research on clean and low-carbon extraction and metallurgical technology of refractory copper resources using microorganisms." The National Key Laboratory of Research and Comprehensive Utilization of Bayan Obo Rare Earth Resources offers another case. That laboratory's mandate leans toward applied research and development. It is intended to focus on research, development, and utilization of existing rare earth resources in the Bayan Obo Region. But it pursues that work with an emphasis on developing new breakthroughs, for example in "new technologies for mining and smelting." (Innovation China, 2025)

In other words, while Beijing is generally content to be a fast follower in the global tech and industrial competition, China is investing to be not only a powerhouse but also a first mover in minerals and materials. Beijing is working to develop breakthroughs in the materials sciences that promise "leapfrog" or "overtaking" in downstream domains.

Conclusion: Implications of the Competitive Balance

The upstream of critical materials has become a new battleground in the race for technological supremacy. The United States and China both recognize the utility of critical materials for today's geopolitical and national security competitions. US INDO-PACOM Commander ADM Sam Paparo put a fine point on it in recent remarks about the risk of kinetic conflict with China: "You can't AI your way out of material deficiency." (John Grady, 2025)

Supply and sustainment of critical materials matters for the weapons systems and defense posture of today. It also matters for developing the determinative capabilities of tomorrow – those that carry security as well as commercial importance. And the US is at risk of losing out to China across the board.

China has developed an advantage in critical minerals and shown its willingness to use that advantage to coerce. At a first order and in the immediate, this means that China can create real battle damage in the exchange of peacetime salvos – like, for example, with export restrictions on gallium and germanium. And Beijing's threat is positioned only to grow in the future. Unlike most realms of technological competition, China eschews its typical "fast follower" approach to seizing advantage at the upstream. Critical materials constitute one outlier realm in which Beijing invests to generate breakthroughs. And that effort threatens to allow Beijing not only to wield advantage in critical materials for coercive effect today, but also to lock in such advantage for tomorrow and, ultimately, use it to unseat downstream US leadership. Meanwhile, the United States has largely remained in reactive and defensive modes as it attempts to catch up with Beijing's upstream lead.

China critical mineral positioning poses a next order threat as well. Beijing has positioned to stymy US efforts to defend, let alone to compete. Beijing is adept at leveraging its pricing power, technological advantage, and full value-chain approach in critical minerals to undermine US efforts to establish truly independent, domestic supply chains. China can, for instance, hold at risk access to equipment necessary for mineral processing. Or – bigger picture and more troubling – China can, right as the US develops nascent critical minerals capabilities, flood global markets to tank global prices, therefore pricing fledgling US players right out of the game.

This competitive balance, or imbalance, is such that the US needs to take immediate and drastic action. And considering both the security relevance of the threat at hand and the impossibility of relying on China-distorted markets to face it down, defense acquisition processes and actors have important roles to play in that action.

The US approach to the upstream of innovation needs to see a doubling down of defense. For example, the US should work to develop stockpiles of critical minerals – to include relevant processed oxides and alloys – to meet defense industrial base demand that China may otherwise place at risk. Those efforts will require coordination across policy, acquisition, and industrial base actors that each have roles to play in bolstering US and allied supply chains. Private stockpiles, refining and processing operations, and mines all, for example, could benefit from explicit offtake signals from defense acquisition programs coordinated through Defense Production Act or other relevant acquisition authorities.

At the same time, the US needs to move beyond the defensive. The US needs to take an offensive tack, and one that does not simply react to China's positioning and signaled or latent leverage. That offensive should include promotion of mineral- and material-relevant research and development, focused both on finding alternatives to China dependence and on enhancing US strengths. The US defense acquisition ecosystem has enormous potential for increasing early stage, basic research activities and directing those activities toward materials sciences domains that may propel critical mineral breakthroughs, all along the value chain from extraction to midstream processing to applications.

In addition, where possible, the US defense acquisition system can chip away at China's market and pricing power by imposing more costs on China-tied supply chains. US downstream manufacturers, including the defense industrial base, can serve as catalysts in the effort to develop independent supply lines by committing to upstream procurements from domestic sources. Acquisition processes can reinforce the incentives for such downstream alignment by incorporating upstream vulnerabilities into program requirements and performance metrics.

Reference List

Bradsher, Keith, "China is Said to Resume Shipping Rare Earth Minerals," *The New York Times*,October 28, 2010.

de La Bruyère, Emily and Nathan Picarsic, "Understanding How China Secured Its Chip Stack,"Hinrich Foundation, April 2025,https://www.hinrichfoundation.com/research/wp/tech/understanding-how-china-securedits-chip-stack/.

de La Bruyère, Emily and Nathan Picarsic, *Elemental Strategy: Countering the ChineseCommunist Party's Efforts to Dominate the Rare Earth Industry*, Foundation for Defenseof Democracies, February 2022, https://www.fdd.org/analysis/2022/02/10/elementalstrategy-countering-the-chinese-communist-partys-efforts-to-dominate-the-rare-earthindustry/.

de La Bruyère, Emily and Nathan Picarsic, "Two Markets, Two Resources: Documenting China'sEngagement in Africa," US-China Economic and Security Review Commission,November 2020, https://www.uscc.gov/sites/default/files/202011/Two_Markets_Two_Resources_Documenting_Chinas_Engagement_in_Africa.pdf.

Doshi, Rush, Emily de La Bruyère, John Ferguson, and Nathan Picarsic, “China as a ‘cybergreat power’: Beijing’s two voices in telecommunications,” Brookings Institution, April2021, https://www.brookings.edu/articles/china-as-a-cyber-great-power-beijings-twovoices-in-telecommunications/.

Grady, John, “INDOPACOM CO Paparo Outlines Risk of Western Pacific Conflict,” USNI, April10, 2025, https://news.usni.org/2025/04/10/indopacom-co-paparo-outlines-risk-ofwestern-pacific-conflict.

Jinping, Xi, “习近平在网信工作座谈会上的讲话全文发表[The Full Text of Xi Jinping’s Speech at the Forum on Cybersecurity and Informatization Work],” Xinhua, 2016.

Lu, Amy and Tony Munroe, "China Bans Export of Critical Minerals to US as Trade Tensions Escalate," Reuters, December 3, 2024.

Puko, Timothy, “U.S. Is Vulnerable to China’s Dominance in Rare Earths, Report Finds,” *Wall Street Journal*, June 29, 2020, https://www.wsj.com/articles/u-s-is-vulnerable-to-chinasdominance-in-rare-earths-report-finds-11593423003.

“2022 Final List of Critical Minerals,” US Geological Survey, February 2022.https://www.federalregister.gov/documents/2022/02/24/2022-04027/2022-final-list-of-critical-minerals.

“Department of Defense Research, Development, Test, and Evaluation (RDT&E): Appropriations Structure,” Congressional Research Service, September 2022,https://sgp.fas.org/crs/natsec/R44711.pdf.

“Executive Order 13817A Federal Strategy to Ensure Secure and Reliable Supplies of Critical Minerals,” December 20, 2017.

“Federal Research and Development (R&D) Funding: FY2024,” US Congress, April 10, 2025.

“The Role of Critical Minerals in Clean Energy Transitions,” International Energy Agency, May 2021.

"Outline of the PRC’s 14th Five Year Plan for National Economic and Social Development [中人民共和国国经济和社会发展第十四个五年规划和2035 年远景目标纲要]," March 13,2021.

US Geological Survey, “Mineral Commodities Summary,” 2024.

World Bank World Integrated Trade Solution, Accessed 2025.

"Zhengzhou University and CNMC (Ningxia) Dongfang Group Co., Ltd. Held a SigningCeremony for the Joint Construction of the 'National Key Laboratory of Rare MetalSpecial Materials' [郑州大学与中色（宁夏东方集团有限公司举行共建“稀有金属特种料全国重点实验室”签约仪式]", Zhongyuan Critical Metals Laboratory, March 8, 2023.

"白云鄂博稀土资源研究与综合利用全国重点实验室 [Bayan Obo Rare Earth ResourcesResearch andComprehensive Utilization National Key Laboratory]," Innovation China, March 12, 2025.

"有色金属强化冶金新技术全国重点实验室开放课题申请指南(2024年度) [Application Guide forOpen Topicsof the National Key Laboratory of New Technologies for IntensifiedMetallurgy of Nonferrous Metals(2024)]," Chinalco, May 21, 2024.

"我国研究与试验发展（R&D）经费投入稳步增长 [China's Research and Development (R&D)Funding Has Grown Steadily]," National Bureau of Statics, January 23, 2025.

"商务部 海关总署公告2025年第18号 公布对部分中重稀土相关物项实施出口管制的决[Ministry of Commerce and General Administration of Customs Announcement No. 18 of 2025 Decision on Implementing Export Control on Some Medium and Heavy RareEarth-Related Items]," China Ministry of Commerce, April 4, 2024.

"国内外战略性矿产厘定理论与方法 [Theory and Methods of Strategic Mineral Determination atHome and Abroad], Acta Geoscientia Sinica, 2021(2).

"全国矿产资源规划（2016—2020年）[National Mineral Resources Plan (2016-2020)]," Ministryof Land and Resources, 2016.

"全国矿产资源规划（2016—2020年）[National Mineral Resources Plan (2016-2020)]," Ministryof Land and Resources, 2016.

Improving Arms Sales, Technology Transfer, and Defense Industrial Cooperation with Allies and Partners

Audrey Aldisert—is a Research Associate in the Defense and Security Department at the Center for Strategic and International Studies (CSIS). Her analytical focus includes the defense industrial base, soft power initiatives, and the relationship between security and development. Prior to joining CSIS, she worked as an intern at the National Consortium for the Study of Terrorism and Responses to Terrorism and the State Department's Bureau of Counterterrorism. Aldisert holds a BA in international development studies and political science with a concentration in international relations from University of California, Los Angeles. [AAldisert@csis.org]

Cynthia R. Cook—is the Director of the Defense-Industrial Initiatives Group. She is a member of the editorial board for the Defense Acquisition Research Journal and is an Adjunct Professor at the Pardee RAND Graduate School. From 1997 to 2021, Cook worked as a Senior Management Scientist at RAND, where she served as the director of the Acquisition and Technology Policy Center. Previously, Cook was a Research Specialist at the Massachusetts Institute of Technology, working on the Lean Aerospace Initiative. She holds a PhD in sociology from Harvard University and a BS in management from the University of Pennsylvania. [CCook@csis.org]

Abstract

This paper examines barriers to effective international arms sales between the United States and its allies and partners. U.S. allies and partners frequently face barriers to receiving advanced technologies and military equipment because of cumbersome policies and regulations around Foreign Military Sales (FMS) and International Traffic in Arms Regulations. They also face delays in acquiring vital weapons systems and challenges related to integrating U.S. technologies into their armed forces. To illuminate these challenges, the Center for Strategic and International Studies will be presenting results from the first ever survey of member states of the Defense MOU Attaches Group (DMAG), the set of nations who have Reciprocal Defense Procurement Agreements with the United States. The survey identifies challenges and enablers allies and partners undergo when doing business with the United States. Topics span selling/receiving arms to/from the United States, complex and rigid U.S. export control policies, country-specific security and defense industrial goals, processes that enable technology transfer and weapons sales, and the utility of defense cooperation agreements and programs.

The implications and relevance of this project for the larger acquisition community lie in streamlining international defense procurement procedures, which is of paramount importance given today's global threat environment. The National Defense Industrial Strategy highlights the importance of working with allies and partners in one of its four strategic priorities, Economic Deterrence. While Foreign Military Sales offer the potential for allies to greatly increase their military power and for the United States to strengthen the overall bilateral relationship, the United States has a complex set of rules governing arms sales and dual-use technologies that are often too rigid and complex. These regulations, while important to ensure advanced U.S. military technology does not fall into the hands of hostile actors, also slow international acquisition processes to a point that threatens U.S. deterrence strategy and, by extension, the international, rules-based order.

Introduction

As the United States and its allies and partners face a more dangerous and uncertain world, the strategic imperative for cooperation has intensified. One of the greatest strengths of the United States has always been the nations' connections with allies and partners. This is underpinned by robust defense industrial cooperation, which strengthens partnerships, increases interoperability, and fills gaps in U.S. industrial capacity and capability. Working with allies offers an opportunity to surge production and contributes to deterrence. However, and in spite of the benefits, there are challenges limiting cooperation. Particularly for the nation's allies and partners, buying from the United States and working with U.S. industry on co-development and co-production can run into a wide range of regulatory and other types of barriers. While certain military capabilities, such as intelligence sharing, may be done with a relatively narrow set of nations, defense industrial cooperation offers a way of building connections to a broader range of allies.

This paper takes a fresh look at the question of the challenge of defense industrial cooperation through a direct survey of some of the United States' closest industrial partners, those with a Reciprocal Defense Procurement Agreement Memorandum of Understanding (RDP MOU). These agreements allow foreign industry to be considered domestic sources, granting the U.S. Department of Defense (DoD) greater and easier access to ally and partner technologies and supply chains. The survey covered a range of questions on the reasons for and the challenges with cooperation.

Survey respondents confirmed that building domestic capacity, deterring the threat, ensuring interoperability, and building regional capacity were all important goals, with building domestic capacity their main interest. The Technology Security and Foreign Disclosure (TSFD) and International Traffic in Arms Regulations (ITAR) were noted as particularly challenging processes when doing business with the United States, be it importing arms, co-developing or co-producing defense goods, or transferring technologies. However, respondents recognized that their home country processes added friction too. Document markings like Controlled Unclassified Information (CUI) and NOFORN were highlighted as challenges, with additional discussions revealing that the lack of clarity in how to dispute these markings was a source of frustration. RDP MOUs are widely acknowledged as being key facilitators for doing business with the United States; however, the exemptions they grant to participating countries are often not recognized by, or are even opposed by, U.S. program offices, U.S. congressional members, and the executive branch.

While none of these findings are particularly surprising, the survey approach extending a sample beyond the United States' largest and closest partners confirmed that the challenges are persistent. They also offer an initial baseline against which can be used to measure the impact of future changes in policy. As the United States faces a more dangerous world where near peer competitors are investing in and expanding the capacity of their own industrial bases, working with allies offers an effective way to strengthen partnerships and increase deterrence.

Background: The Goals and Challenges of Defense Industrial Cooperation

Defense industrial cooperation offers participating nations many benefits. As a subset of broader security cooperation efforts, defense industrial cooperation strengthens alliances and partnerships through the relationship building integral to working together on research and development and production and enhances military interoperability because of operating common platforms. A recent Defense Innovation Board report offered that cooperation is increasingly important because "The United States is no longer the leading source of progress across critical areas of defense related technology innovation, such as 5G, hypersonics, and electronic warfare, while our allies and partners increasingly lead in other areas, including semiconductors, directed energy, and quantum science. Cooperation is urgently required to ensure access to advanced technology" (Defense Innovation Board, 2024). Cooperation can improve supply chain resilience through the development of additional suppliers and can also build domestic defense industries as they participate in joint efforts that may have the potential for additional customers. The "build allied" advantages means that the United States and its allies and partners seek appropriate opportunities for defense cooperation. For other nations in particular, working with the United States has been desirable because of the opportunities to strengthen ties and benefit from advanced technology (McGinn, 2023).

However, any cooperation among nations with different priorities, political systems, funding cycles, and laws and regulations can be challenging to successfully accomplish. The desirability of working with the United States means that those challenges of working with that nation have been highlighted as examples of why it is so hard to execute. Some contractors have successfully de-coded the puzzle, but bureaucratic red tape creates high barriers to entry for domestic industry, let alone foreign industry. A few notable barriers include International Traffic in Arms Regulations (ITAR), Technology Security and Foreign Disclosure (TSFD), and Foreign Military Sales (FMS). We highlight these and other structural barriers in this research.

Given the number of frameworks and processes covered in the research, it is useful to start with brief definitions of frameworks for cooperation and regulations that create challenges. Table 1 summarizes the frameworks.

Table 1. Frameworks for Cooperation

Acronym	Full Name	Description
RDP MOU	Reciprocal Defense Procurement Memorandum of Understanding	Allows foreign industry to be considered a domestic source of defense equipment
SOSA	Security of Supply Arrangement	Bilateral agreements for prioritized contract performance
NTIB	National Technology and Industrial Base	Framework for enhancing defense integration and R&D collaboration
NDPP	NATO Defense Planning Process	Framework that aims to harmonize alliance force and capability planning activities
MIEA	Master Information Exchange Agreement	Framework for sharing technical and operational data
QA	Reciprocal Government Quality Assurance	Mutual recognition of quality assurance processes
DEF	Defense Exportability Features	Early incorporation of exportability features into defense systems
FCT	Foreign Comparative Testing	Program to test and evaluate foreign technologies

The frameworks and supporting processes facilitate cooperation, but there are also a number of processes, regulations and controls that challenge working together. Table 2 lists the U.S. export and arms sales regulations used to manage technology transfer and safeguard U.S. technology. It also includes the processes that handle delivering defense products to a partner nation, along with markings that can prevent partners from accessing information relevant to cooperation. Several of the more processes are also described in more detail below.

Table 2. U.S. Regulations and Processes

Acronym	Full Name	Description
ITAR	International Traffic in Arms Regulations	Governs defense items and services exports
TSFD	Technology Security and Foreign Disclosure	Governs technology transfer and information sharing
FMS	Foreign Military Sales	Government-to-government process for defense equipment sales
DCS	Direct Commercial Sales	Commercial-to-government process for defense equipment sales
ACEA	Arms Control Export Act	Provides authority for FMS and DCS, implemented through ITAR
EAR	Export Administration Regulation	Regulates export of dual-use items
CMMC	Cyber Maturity Model Certification	Cybersecurity standards for defense contractors
EDA	Excess Defense Articles	Program for transferring surplus U.S. military equipment
TPT	Third Party Transfer	Process to transfer U.S.-origin defense articles to third parties
IEA	Information Exchange Annexes	Specific agreements under MIEA for data exchange
NOFORN	Not Releasable to Foreign Nationals	Information classification restricting foreign access
CUI	Controlled Unclassified Information	Safeguards sensitive but unclassified information
USML	U.S. Munitions List	Catalog of defense articles regulated by ITAR
CCL	Commerce Control List	Catalog of dual-use items regulated by EAR

Value of Improving Arms Sales

Streamlining the arms sales process is critical to advancing the goals of defense industrial cooperation outlined above. Reducing bureaucratic inefficiencies and accelerating delivery timelines ensures that partners will receive the capabilities they need at the earliest opportunity. This strengthens alliances by reinforcing the United States as a reliable and responsible security partner. Efficient arms sales processes not only facilitate the fielding of U.S. platforms, weapons systems, and technologies to partner nations, but ensure that the United States is reaping the benefits of partner state-developed technologies. This improves interoperability and reduces friction during multinational operations, enhancing coordination and force effectiveness in complex operating environments.

Cooperation through arms sales also bolsters supply chain resilience by integrating additional suppliers and production lines across allied and partner nations. Supply chain diversification reduces dependency on single-source suppliers, decreases risks associated with domestic supply chain bottlenecks, and ensures continuity of material during crises. Overcoming barriers to cooperation and improving the structure and execution of arms sales is a critical component to defense industrial cooperation and collective security.

Notable Challenges

International Traffic and Arms Regulations: ITAR governs defense items and services to ensure sophisticated military technology—such as what is found on the U.S. Munitions List (USML)[1]—does not fall into the hands of U.S. adversaries or hostile actors. ITAR serves as the implementing framework for the Arms Export Control Act (22 U.S.C. 2778), which is overseen by the State Department's Bureau of Political-Military Affairs (U.S. Department of State, n.d.-b). ITAR is fundamental to safeguarding U.S. technologies and weapons systems but may also create risks to U.S. and coalition military readiness (Defense Innovation Board, 2024). ITAR regulations have contributed to a risk averse culture which has led to hesitancy in sharing technology even with its closest allies and partners. It prevents the United States from quickly proliferating advanced technologies to its friends and can impede U.S. warfighters from gaining access to advanced allied capabilities when foreign companies desire to avoid ITAR processes.

Technology Security and Foreign Disclosure: TSFD, similar in purpose to ITAR, manages the tradeoffs between building allied capabilities and safeguarding national security. TSFD policies aim to balance the risks associated with transferring sensitive and highly classified U.S. technology and information with the benefits of international cooperation (DAU, 2018). Navigating TSFD processes or "pipes" is challenging yet necessary for engagement with friendly nations. DoD program management offices (PMOs) must clear various TSFD pipes to include allied participation, which include International Cooperative Programs, FMS, DCS, and International Contracting (McGinn, 2023). U.S. industry is typically required to acquire TSFD approvals prior to requests for export approval, which strains PMOs and U.S. industry when pursing international cooperation efforts (McGinn, 2023).

Selling to Allies and Partners: Foreign Military Sales (FMS) is a process through which eligible foreign governments may purchase defense equipment and services from the U.S. government. FMS is a government-to-government process, whereas Direct Commercial Sales (DCS) is a commercial-to-government process. FMS is the largest U.S. security assistance program, aimed to help protect the economic health and security of allies and partners (DAU, n.d.-c). The Department of State determines what countries are eligible for FMS programs, while the Department of Defense executes the programs (Defense Security Cooperation University, n.d.-b).

In an FMS program, the foreign government is responsible for all of the costs associated with the sale. Purchased items can either come from DoD stockpiles or from new procurement, in which the DoD then enters a contract with a U.S. defense contractor to produce the items purchased. A single FMS case can contain hundreds of individual line items, span multiple commands and military departments, and take years or decades to fully deliver. FMS utilizes both Title 22 and Title 10 funds, each with its own set of rules. For some complex FMS programs, Congressional review and approval is required, which can significantly delay the FMS process. The threshold values for a sale to require Congressional oversight has not been adjusted in the last two decades, resulting in more cases subject to Congressional review now

[1] The USML includes a range of military items, such as firearms, explosives, military vehicles, aircraft, and classified technical data. The United States Munitions List, U.S. Code of Federal Regulations, Part 121, https://www.ecfr.gov/current/title-22/chapter-I/subchapter-M/part-121

than when the Arms Control Export Act (ACEA) was first implemented in 1976 (House Foreign Affairs Committee, 2024).

Tools to Overcome Challenges: The United States has a variety of specific agreements with allies and partners to enhance defense industrial cooperation. Defined in the U.S. Code, the National Technology and Industrial Base (NTIB) framework is aimed at enhancing collaboration in defense production, innovation, and supply chain with Canada, the United Kingdom, Australia, and New Zealand (Congressional Research Service [CRS], 2023). Security of Supply Arrangements (SOSA) are bilateral agreements allowing the U.S. and partner departments of defense to request "prioritized performance of contracts from companies in SOSA-signatory nations, and for SOSA signatories to request the same from U.S. firms" (CRS, 2021).

RDP MOUs are the broadest of these agreements, designed to promote rationalization, standardization, and interoperability with ally and partner nations. They grant the United States and allied countries reciprocal access to their respective defense markets. These agreements streamline procurement processes to enhance effective defense cooperation and establish transparency and openness to competition. RDP MOUs relax provisions from the 1933 Buy American Act that require the U.S. government to purchase supplies and finished goods domestically, otherwise requiring a waiver to buy internationally (Defense MOU Attaches Group, n.d.). Each agreement generally contains similar provisions, such as granting each party increased access to the other's defense procurement system, "removing barriers to trade, providing reciprocal treatment to industrial enterprises of the other country, or waiving 'buy national' laws" (GAO, 2024).

Assessing the Challenge of Cooperation

Research on defense cooperation very often focuses on a narrow set of allies with defense substantial trade with the United States. To get a broader perspective, the CSIS research team drew on the group RDP MOU nations using a survey that sought to identify the challenges and enablers countries experience when doing business with the United States. Topics spanned selling arms to and receiving arms from the United States, U.S. export control policies, country-specific security and defense industrial goals, processes that enable technology transfer and weapons sales, and the utility of defense cooperation agreements and programs.

The survey was completed by representatives from member states of the Defense Memoranda of Understanding Military Attachés Group (DMAG), with the support of DMAG leadership. The DMAG is a group comprised of defense officials and attachés from countries that have RDP MOUs or equivalent agreements with the United States. As of 2025, 28 countries have RDP MOUs with the United States: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Germany, Greece, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, Turkey, and the United Kingdom. Of these, 23 are members of NATO. South Korea, Brazil, and India are in ongoing RDP MOU negotiations with the United States. Given its DMAG "observer" status, South Korea was part of the sample population. Brazil and India are not.

Thirteen nations provided responses to the questions, and many offered additional optional comments. While survey responses were limited to one per country, that does not mean that each response was answered by only one country representative. In many cases, entire acquisition teams contributed to ensure the response reflected their country's broader approach rather than individual perspectives. However, variability inevitably exists due to differences in participants' experiences, expertise, and roles. As a result, while responses may represent a national viewpoint, they should not be interpreted as an official country position.

There have been calls from experts in recent years to increase available data in the realized benefits and challenges that procurement agreements bring to U.S. allies and partners. This survey seeks to help fill this quantitative information void by offering a dataset on U.S. bureaucratic processes that facilitate—or hinder—defense industrial cooperation. The results also offer a baseline against which future policy changes can be assessed.

Allied Country Defense Industrial Priorities

Survey recipients were asked to rate four key defense industrial cooperation goals on a scale from one to five—one being not important and five being extremely important.[2] The goals include building domestic industry capacity, deterring the threat, ensuring interoperability, and developing regional capabilities.

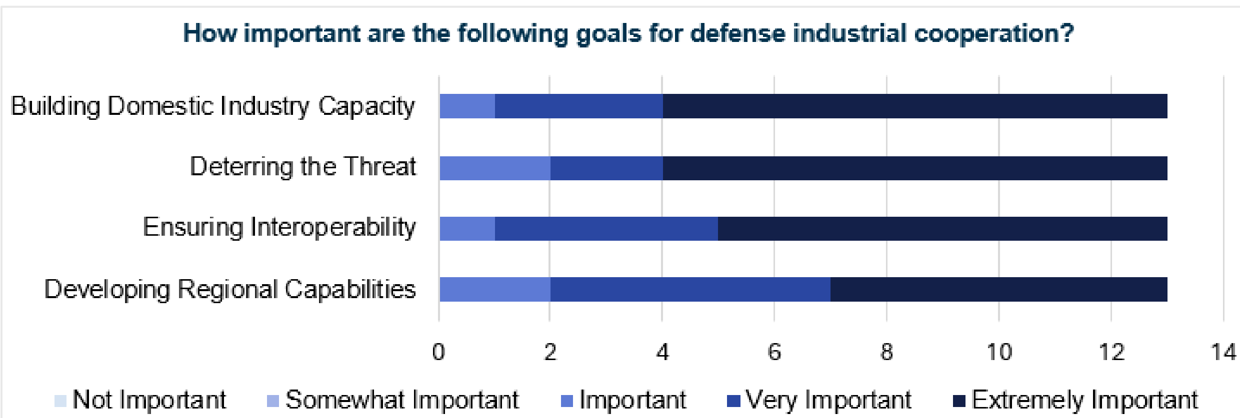


Figure 1. Goals for Defense Industrial Cooperation

All four goals were rated as at least important by all respondents; no participant rated any of these four goals as either not important or somewhat important. Building domestic industry capacity, deterring the threat, and ensuring interoperability were rated as extremely important by most, with nearly half rating developing regional capabilities as extremely important.

Defense industrial cooperation priorities are directly linked to operational requirements, meaning countries must ensure they can produce and sustain defense critical defense systems to meet their strategic needs. In some cases, this necessity leads to domestically indigenizing sovereign defense capabilities, even if it comes at the expense of international collaboration. As a result, respondents noted investments may pivot away from cooperative efforts and toward developing domestic industries and capacity that can independently support long-term defense readiness.

Respondents whose neighbors pose a direct threat to their national security note the importance of defense collaboration with allies as essential to extended deterrence, and a strong and integrated defense industrial base strengthens deterrence posture.

A country's size, natural resources, geographical location, and topography influence defense needs. Certain weapons systems or defense services are more conducive to specific terrain—a landlocked country has less need to heavily invest in naval capabilities, for instance. And smaller countries with long coastlines may invest in advanced undersea capabilities and

[2] Questions rated on a five-point Likert scale were categorized as the following: not important, somewhat important, important, very important, and extremely important.

specialty systems. This creates a strong, but niche, defense industrial base in specific domains, which can lead to strong dependencies on larger allies like the United States for key military platforms like fighter jets. Despite continued investment into developing domestic capabilities, allies with limited resources and specific geographies view regional and U.S. cooperation as a key supplement to defense industrial areas of national strategic importance. Thus, these allies may view developing regional capabilities as more favorable than others that have more independent defense industrial bases.

Legislative offsets were also noted as a priority among discussions with the DMAG and other defense acquisition stakeholders. Legislative offsets refer to the benefits—such as the economic, industrial, or technological advantages—that purchasing countries obtain if acquiring defense systems from the United States (Kenlon, 2020). These conditions of purchase pertain to both government-to-government or commercial sales of defense articles or services, and "compensation can include mandatory co-production, licensed production, subcontractor production, technology transfer, and foreign investment"(Bureau of Industry and Security, n.d.). Certain U.S. allies and partners legally mandate offsets to ensure economic and industrial benefits when purchasing defense systems from abroad.

U.S. Export Control Challenges

U.S. export control processes have been long cited by allies and partners as complex, slow-moving, and opaque (Corben & Greenwalt, 2023). These challenges can create uncertainty for foreign buyers, complicate defense cooperation procedures, and, in some extreme cases, incentivize partners to seek alternative suppliers.

Survey respondents were asked to rate seven key U.S. export control processes on a scale from one to five—one being not challenging and five being very challenging. The seven processes include ITAR, Export Administration Regulations (EAR), Foreign Military Sales (FMS), Direct Commercial Sales (DCS), Excess Defense Articles, TSFD, and Cybersecurity Maturity Model Certification (CMMC).

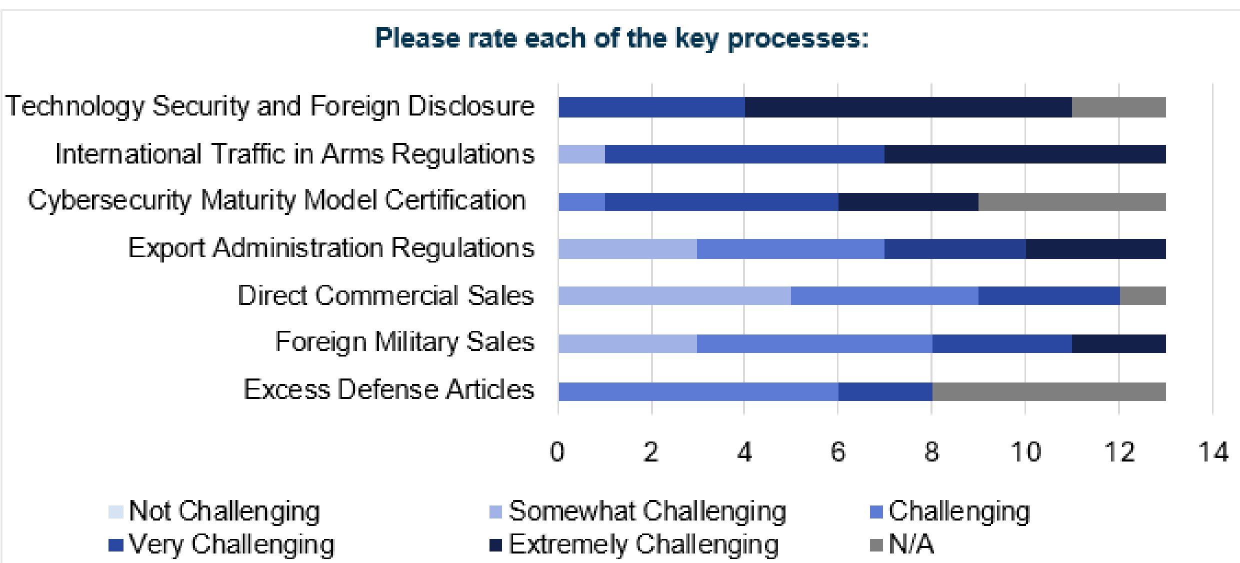


Figure 2. How Difficult Are Processes?

As shown in Figure 2, respondents deemed TSFD and ITAR to be among the most challenging hurdles to international procurement from the United States. No respondent rated any of these processes as not challenging.

ITAR is a complex export control system marked by bureaucratic red tape that U.S. allies and partners continuously express frustration over. ITAR processes are often perceived as too stringent and not conductive to the current era of geopolitical competition. DMAG respondents recognized the purpose and importance behind ITAR but simultaneously critiqued its rigidity and prolonged lead times. Because ITAR is such an expansive bureaucratic process, respondents noted that guidance from various U.S. authorities may be different or even conflicting. These barriers impede defense industrial cooperation and ultimately jeopardize U.S. and allied defense posture and readiness.

While ITAR does create significant challenge to international defense cooperation, it is part of a broader framework for partnerships. The Directorate of Defense Trade Controls (DDTC), responsible for administering ITAR, plays a role in facilitating collaboration by reviewing, and then subsequently approving, export licenses. Through this process, the DDTC ensures that defense technologies are transferred in a controlled and responsible manner, while simultaneously supporting co-development and co-production efforts with international partners. ITAR helps balance national security concerns with opportunities for technology innovation and collaboration with partner nations.

Some respondents offered additional nuance to the TSFD and ITAR processes. TSFD may be relatively opaque with a lack of clarity around which authorities do what or the pipeline of approval. However, it is not always a very challenging process. They felt ITAR suffers from the opposite problem—that it is a challenging process despite knowledge of the steps required for compliance.

Respondents noted that even though key U.S. export control and technology transfer processes are viewed as at least somewhat challenging, that does not mean that they are wrong or misguided. There is an understanding amongst ally and partner nations that these regulations exist for a reason despite their complexity. Allied nations have their own complex export control regimes that share the same objective of U.S. protection policies: to prevent sensitive technology and information from falling into the hands of unfriendly nations. There was no call by survey respondents to eliminate U.S. processes, but rather that a more transparent, streamlined system with predictable lead times could enhance cooperation and the benefits to both nations.

Given the difficulty of ITAR procedures and the various stakeholders involved, the research team asked DMAG how well they understand ITAR processes. Figure 3 shows the results.

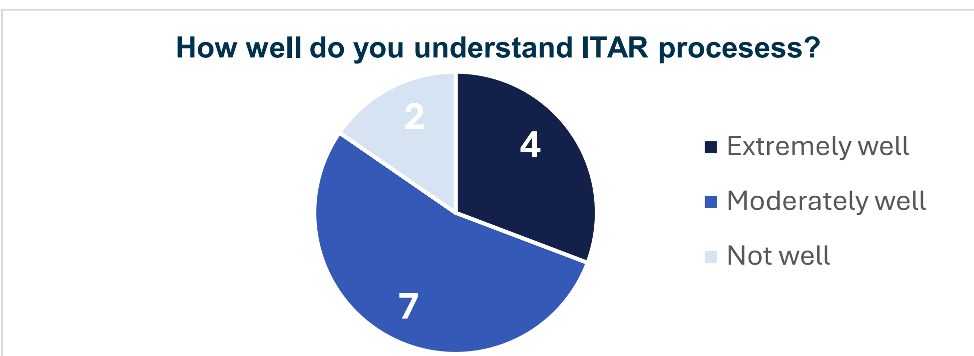


Figure 3. Knowledge of ITAR Processes

Most—but not all—respondents understand ITAR processes moderately well. The DMAG respondents—defense cooperation attaches and defense officials—are familiar and relatively well-versed in export control processes by nature of their profession, and it may be of strategic concern for the United States that only four respondents reported a confident understanding of its requirements.

One specific challenge of ITAR was raised in the comments. Specifically, when a foreign company manages to sell to the DoD, they often will set up production in the United States to manage the volumes. “But every product will be improved over time and new functions might be added. In this scenario the knowledge created in the US subsidiary will not flow back to the mother company due to ITAR. This is not a problem for the company. But it reduces interoperability and interchangeability.”

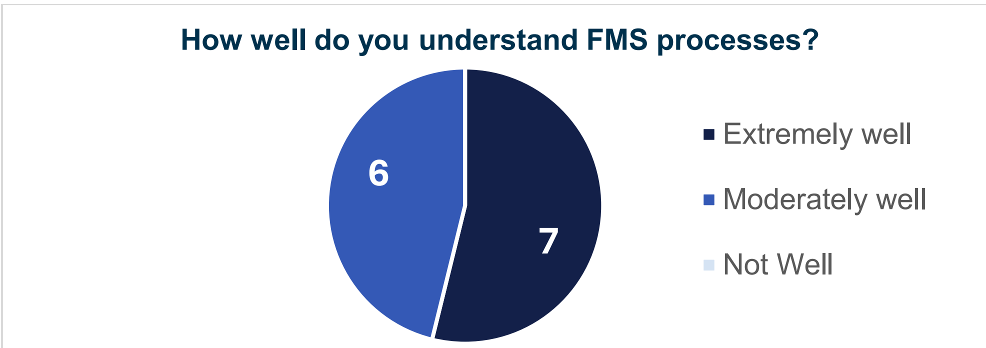


Figure 4. Knowledge of FMS Processes

Unlike ITAR, all respondents reported understanding FMS processes. Although FMS is not directly linked to international cooperation, as it primarily involves one nation’s government purchasing defense systems from the U.S. government, it serves an important role in helping nations achieve their domestic defense industrial goals. FMS enhances interoperability by allowing partner forces to operate using the same systems. FMS can improve regional capabilities by equipping partner nations with advanced technologies that bolster their defense readiness. This serves to enhance deterrence posture as partner nations are better equipped to defend against emerging threats. FMS may also stimulate the growth of a partner nation’s defense industrial capacity by facilitating local production of the acquired system upon license approvals.

The FMS process can be complex and cumbersome, which is why the Defense Security Cooperation University (DSCU) offers a foundational level FMS course that explores the essential components to military sales and transfers between the United States and partner nations. Students learn how to “plan, execute, and sustain the many complex and interrelated aspects of sales and transfers under the FMS program” (DSCU, n.d.-a). One respondent noted that this FMS course is no longer available for individuals in their office, meaning a growing number of foreign FMS officers lack a basic understanding of FMS processes in defense cooperation offices.

Foreign defense attachés lacking an adequate understanding of the FMS acquisition process not only impedes their home country’s ability to acquire U.S. defense systems efficiently, but also negatively impacts the United States directly. Delays in the acquisition process, or even reduced purchases of U.S. defense systems, result in the United States exporting fewer defense products and providing fewer services, which reduces industry sales and hampers the ability of the United States to interoperate with its allies and partners.

FMS is a complex process, and the literature indicates a variety of specific challenges. The survey included a question asking respondents to rate the challenges of a variety of FMS processes.

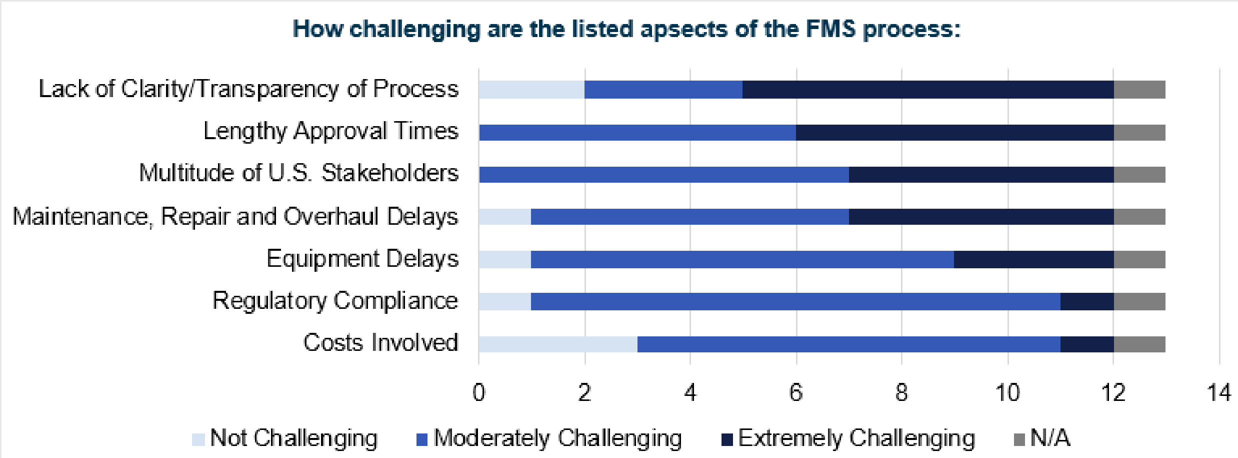


Figure 5. FMS Process Challenges

Most respondents rated the lack of clarity and transparency of the FMS process as the most challenging, closely followed by lengthy approval times and the multitude of U.S. stakeholders involved in the process. Most respondents rated equipment delays, regulatory compliance, and costs involved as moderately challenging. The costs associated with the FMS process were identified as the least challenging factor.

In the optional written section, respondents noted a few other challenges that were not listed. First, long periods of time are required for case closure despite service completion. FMS case closure occurs when “all material has been delivered, services have been performed, other requirements of the LOA have been satisfied, known financial transactions (including collections) have been completed, and the purchaser receives a final statement of account” (Saum-Manning et al., 2024). Prompt case closure minimizes the amount of administrative effort required for an unnecessary open case, which diverts resources from other priorities. Prolonged case closure, a common frustration among primary FMS customers, delays the release of excess purchaser funds (Defense Security Cooperation Agency, n.d.). This practice may erode the long-term willingness of partner nations to engage in our arms sales process, especially if certain material or systems can be purchased elsewhere. Secondly, a lack of workforce capacity within the defense industry can lead to increased costs for the production and delivery of defense systems. A workforce that does not meet demand may force partner nations to face higher prices and acquisition delays.

Challenges with coordination were also noted, in particular a fragmented approach when it comes to working with allies. The United States reviews every FMS case on a country-by-country basis, and NATO allies lack a centralized authority to streamline FMS coordination. This fragmented approach limits opportunities to optimize FMS outcomes for the broader strategic goals of the alliance. Exploring whether there are groups of countries for which FMS cases can be reviewed together could streamline the process for the United States and speed acquisition by allies.

One respondent noted that on occasion, borderline cases tended to linger as they are being reviewed, which they felt was because of U.S. government hesitation to rapidly decline case requests and rather opt for extensive deliberations to provide alternatives. They suggested that sometimes a faster decision, even if it was negative, would be preferred because it would reduce uncertainty.

Feedback from participants also raised the consideration that current thresholds for Congressional notifications also often hinder the efficiency of the export process. When the U.S. government plans to sell defense equipment, services, or technology to a foreign country, it must submit a notification to Congress that allows lawmakers a designated period to review the proposed sale (CRS, 2024). The requirement for Congressional notification on all sales, regardless of their scale or impact, can create unnecessary delays and administrative burdens, especially on standardized exports that the United States has historically been exporting to its allies and partners. Many of these notifications pertain to sales that are not sophisticated nor strategically sensitive. This slows down the acquisition process without significantly enhancing oversight or national security. It may be beneficial to recuse the notification thresholds to exclude routine transactions of small value. Moreover, thresholds should be updated regularly to account for inflation (Saum-Manning et al., 2024).

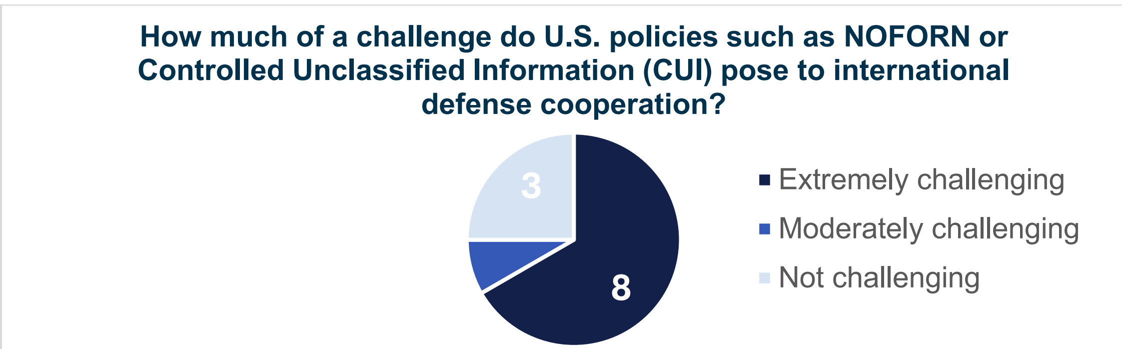


Figure 6. The Challenge of NOFORN and CUI

The challenge of U.S. categorization of information as Not Releasable to Foreign Nationals (NOFORN) or controlled unclassified information (CUI) markings was consistently mentioned as a barrier throughout the duration of this study in both discussions and survey responses, as displayed in Figure 6. These categorizations markings can create barriers to foreign partners' access to information and can hinder procurement or co-production processes. These restrictions can lead to delays in equipment delivery, licensing processes, and may negatively impact interoperability between allied forces.

Respondents note that NOFORN and CUI limit the ability of foreign contractors to compete for opportunities. In some cases, the information is made available, but without sufficient time for foreign contractors to develop a bid. One implication is that while it may increase U.S.-content, it may mean that the DoD is not accessing best-in-class technical solutions. Reforming protectionist policies demands not only regulatory changes, but cultural change to support systematic alteration in the way the DoD approaches classification markings. While the use of NOFORN to obstruct competition is illegal, respondents felt that it remained overused and hence impeded cooperative defense industrial efforts.

Respondents also noted Master Information Exchange Agreements (MIEAs), and subordinate Information Exchange Annexes, are extremely useful. MIEAs establish a reciprocal, balanced exchange of R&D between participating parties and authorize specified IEAs (*U.S./ROK Master Information Exchange Agreement*, n.d.). IEAs exchange R&D pertaining to specific technology or weapons development areas.

Ally and Partner Nation Export Control Challenges and Enablers

The challenge that regulations represent is not unique to the United States. All nations have regulations that have some impact on defense cooperation. Figure 7 reports the results of a comparison, with most respondents believing that working with the United States is no different or harder than working with other nations. Every nation has its own individual export control and technology transfer challenges, and export control challenges are not limited to just the United States.

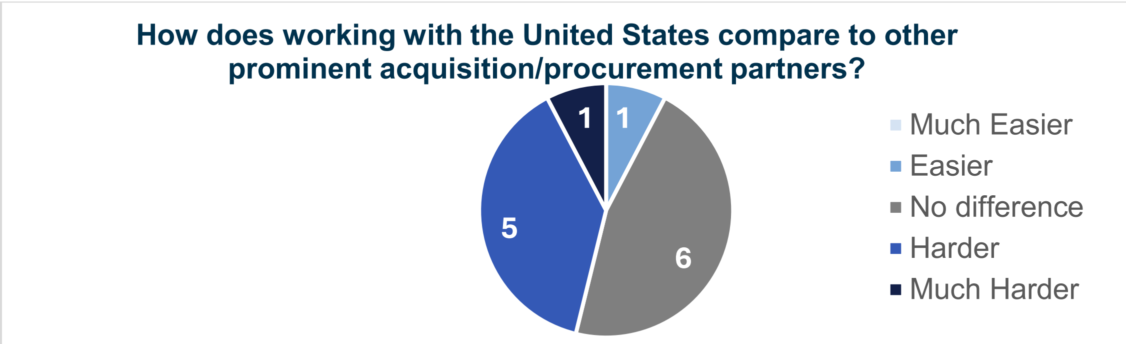


Figure 7. Comparing the United States to Other Partners

That said, the United States is the sole supplier of a number of advanced capabilities. This makes cooperation with the United States, and an understanding of its export control processes, mandatory for those who wish to acquire certain U.S. designed and produced weapons systems. And it means that U.S. regulations have an outsized impact on partners.

U.S. allies and partners also have their own set of export control and technology transfer processes that can hinder—but also enable—information sharing and arms sales. Survey respondents have varying perceptions of their home country's export control processes, but provided useful feedback on what mechanisms could help—or not help—facilitate their international procurement processes.

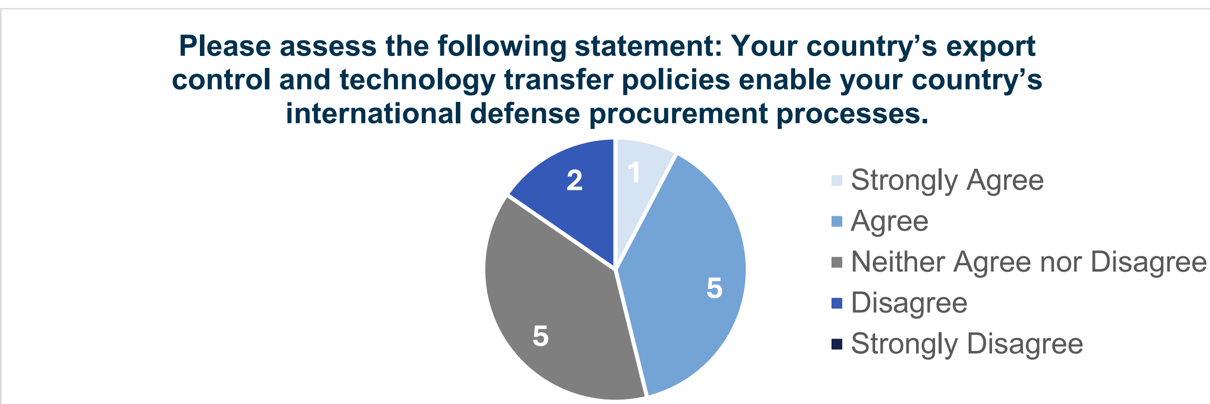


Figure 8. Assessment of Home Country Export Controls

Figure 8 offers a take on this. Most respondents are neutral or agree that their home country export control and technology transfer policies enable procurement processes.

Figure 9: The Impact of Home Country Export Controls on Doing Business With the United States – Significant Challenge

As shown in Figure 9, respondents are reasonably mixed on whether their own export controls pose a challenge to doing business specifically with the United States, but only three agreed that the challenge was on their side.

To gain further nuance on respondents' perspectives of their home country's export controls, the research team also asked whether their domestic export controls generate "friction" for their procurement and cooperation processes, so something less than a "significant challenge."

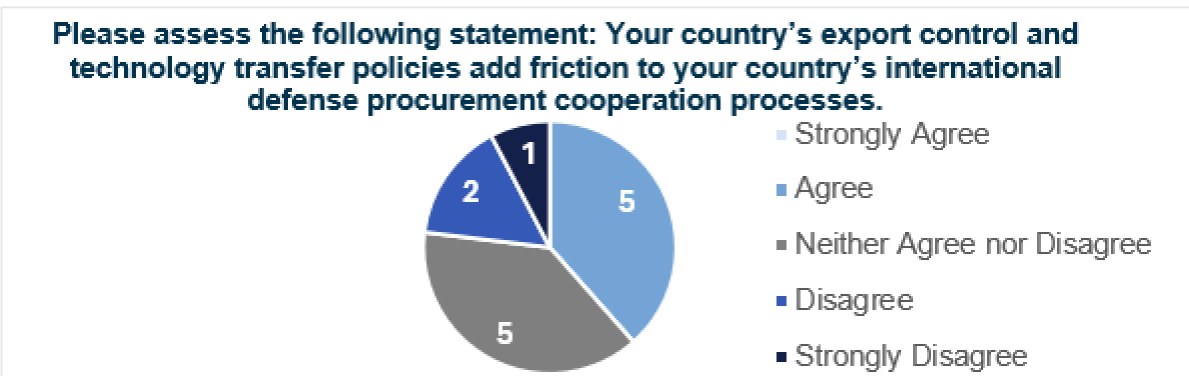


Figure 10. The Impact of Home Country Export Controls on Doing Business With the United States – Friction

Figure 10 shows that most respondents recognized that respondents are more likely to agree that their export control and technology transfer policies add friction to their procurement or cooperation processes.

The research team asked respondents to rate certain processes or agreements based on how beneficial they would be to facilitating trade defense trade with the United States, with the results reported in Figure 11. These are the Defense Production Act, AUKUS, reclassifying items from the USML to the Commerce Control List, and the NTIB. There was significant agreement that most of these would be useful.

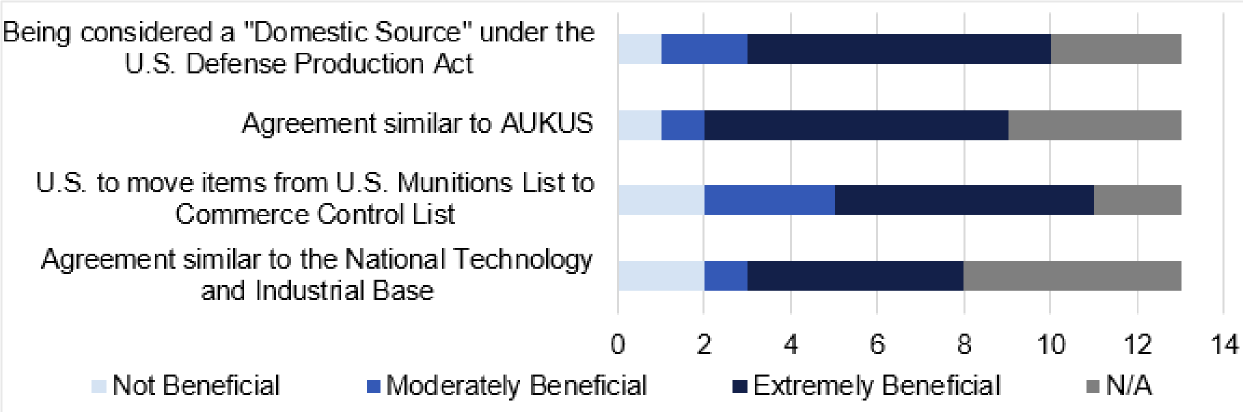


Figure 11. What Could Make Trade Easier?

Defense Production Act

Being considered a domestic source under the DPA was rated among the most beneficial programs by respondents. The DPA, passed in 1950, grants the president the authority to influence domestic industry and expand and expedient certain material required for national defense during emergency mobilizations (FEMA, 2024). Domestic industry may be called upon to expand the production and supply of material critical to national security or emergencies—President Donald Trump utilized the DPA to order General Motors to produce more ventilators and 3M to produce N95 masks during the Covid-19 pandemic, for instance (Siripurapu, 2021).

Title III of the DPA, the Expansion of Productive Capacity and Supply, authorizes "incentives to include loans, loan guarantees, direct purchases and purchase commitments, and the authority to procure and install equipment in private industrial facilities" (CRS, 2023b). Along with U.S. industry, Canada has been considered a domestic source since 1992 (DoD, 2024a). The FY2024 National Defense Authorization Act designated the United Kingdom and Australia to also be considered domestic sources and therefore eligible for DPA funds (National Defense Authorization Act, 2023). This means Canada, the United Kingdom, and Australia enjoy certain U.S. government benefits under certain conditions when able to provide essential defense materials and goods.

Though the DPA includes Canada, the United Kingdom and Australia as domestic sources, the degree to which the U.S. government can direct a foreign firm to produce under the DPA is more nuanced than with a purely domestic firm. The DPA primarily provides incentives (loans, guarantees, etc.) to encourage production, and the United States would be likely to work through diplomatic channels to encourage a foreign firm to increase production of critical goods during a crisis (Office of the Assistant Secretary of Defense, Industrial Base Policy, n.d.). The Defense Priorities and Allocation System (DPAS) implements Title I of the DPA under the Department of Commerce, and applies to all entities physically in the United States, regardless of foreign or domestic ownership (Department of Commerce, n.d.). However, foreign companies and foreign subsidiaries of U.S. companies are outside DPAS jurisdiction; therefore, the U.S. government cannot order an Australian, Canadian, or British firm to produce goods if it is not physically located in the United States (Department of Commerce, n.d.). The DPA provides a framework and financial tools for crisis production, but it requires a collaborative approach. There is a distinction between being considered a domestic source under the DPA and having an RDP MOU with the United States. The latter ensures allied and partner industry are considered domestic sources, waving obstacles associated with the Buy American Act and facilitating smoother access to U.S. defense contracts. But domestic sources under the DPA are utilized during times of national crisis, as firms are incentivized—and ordered—to produce a certain amount of goods or material necessary for national security or during times of crisis. This serves as a mechanism to rapidly mobilize the defense industrial base to ensure the United States has access to vital resources when traditional free-market principles are not sufficient.

An Agreement Similar to AUKUS

AUKUS is a trilateral security partnership between the United States, the United Kingdom, and Australia. It is designed to promote further information sharing and technology transfer and better integrate and diversify security-related supply chains and industrial bases (DoD, n.d.). AUKUS has two pillars, the first being to support the Royal Australian Navy in acquiring nuclear-powered submarines. The second pillar is focused on advanced technologies, including cyber, artificial intelligence, quantum, and undersea capabilities (DoD, n.d.).

To implement these two pillars, efficient procurement strategies between AUKUS member nations was required. Once their defense information protection systems, such as strengthening cybersecurity measures and harmonizing classification standards, were aligned with those of the United States, information sharing and technology transfer were simplified. This was reflected in the revisions made in the EAR and ITAR.

In April 2024, the BIS amended the EAR to facilitate license-free trade with Australia and the United Kingdom in furtherance of the AUKUS objectives. It removed certain “license requirements, expanded license exemptions, and reduced the scope of end-use and end-user-based license requirements for exports, reexports, and transfers (in-country) to or within Australia and the United Kingdom” (Federal Register, 2024b). The BIS estimates that $7.5 billion in trade with Australia and the United Kingdom were subject to these previous license regulations (Bureau of Industry and Security, 2024).

The DDTC made similar changes to ITAR that enable the license-free transfer of commercial defense trade for Australia and the United Kingdom. With certain limitations, authorized users between AUKUS members require no license or other approval for the “export, reexport, retransfer, or temporary import of defense articles, the performance of defense services, or engaging in brokering activities”(Exemption for Defense Trade and Cooperation among Australia, the United Kingdom, and the United States, 2024). This rule also allows for an expedited export licensing process for defense articles or services to Australia, the United Kingdom, and Canada.

One DMAG participant likened AUKUS membership to having a “fast pass” or “carpool lane” through ITAR, streamlining defense cooperation with the United States. However, for other nations seeking privileges comparable to those enjoyed by the United Kingdom and Canada, reform is necessary not only within the U.S. system but also within their own domestic frameworks. Certain nations expressed a desire to be a part of AUKUS pillar two, even if with specific technologies only, such as hypersonic, missile, and undersea capabilities.

Moving Items From the USML to the CCL

ITAR governs the U.S. Munitions List (USML), which is a list of defense-related articles, services and technologies designated as critical to U.S. national security. The Directorate of Defense Trade Controls (DDTC), within the U.S. Department of State, is responsible for administering ITAR. The DDTC must approve export licenses for items on the USML in order to prevent U.S. adversaries from obtaining advanced technologies critical to U.S. military advantage.

The Commerce Control List (CCL) is a list of dual-use items that have military but also commercial applications. The Export Administration Regulations, enforced by the Bureau of Industry and Security within the U.S. Department of Commerce, governs the CCL. Items on the

CCL are typically less restricted than their counterparts on the USML and only sometimes require a license (unlike items on the USM list which always do). CCL items requiring licenses include sensitive technologies such as semiconductors and aerospace components.

In further conversations with the DMAG, countries expressed this solution of moving items from the USML to the CCL was an underrated solution to complex U.S procurement procedures.

National Technology Industrial Base

The NTIB is an agreement between the United States, Australia, the United Kingdom, New Zealand, and Canada that establishes joint national security and dual-use research and development initiatives as well as production and maintenance related activities (CRS, 2023a) While some respondents offered that it would be extremely beneficial, it was noted that NTIB has done little more than enabling limited information exchange. NTIB lacks funding and does not change standing policy, which limits its contributions to improved defense industrial cooperation processes. It does not address the inefficiencies baked into various export control regimes—such as ITAR, EAR, FMS, the Canada ITAR waiver for unclassified goods control, and the Australia-UK Defense Trade Treaties—that foreign companies are subjected to depending on where they are based, some of which are located in all NTIB countries. Ensuring compliance to these various regulations requires “an army of lawyers and clerks, burning up a significant amount of resources” (Greenwalt, 2019).

Benefits of Cooperation

One reason nations seek to engage in defense industrial cooperation with the United States is to “uplift” their domestic industry. Selling to the United States was viewed as extremely important to home country industry by all but one of the respondents, as shown in Figure 12.

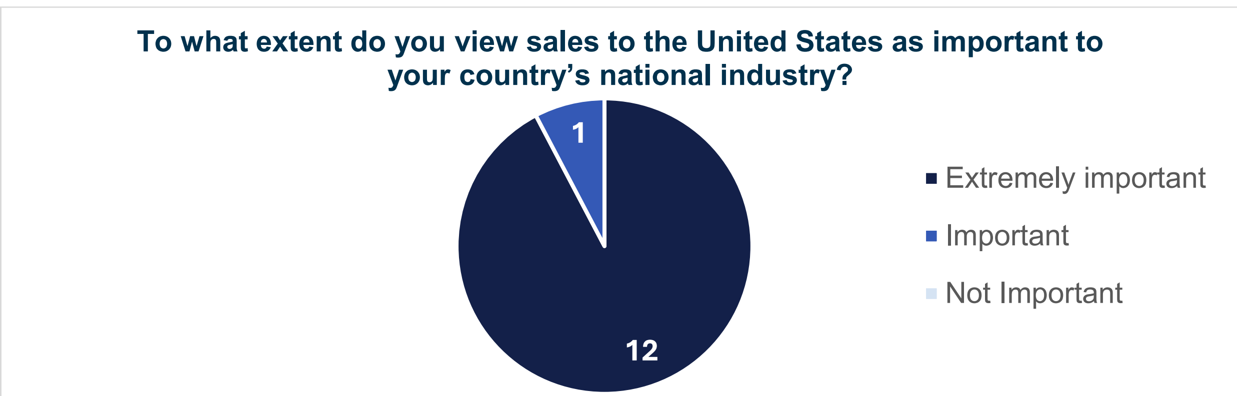


Figure 12. Uplifting Domestic Industry

These findings are unsurprising given the United States has the largest arms market in the world, making the United States a critical enabler to partner nation industrial development strategy. Beyond the sheer size of the market, there are other advantages. Selling to the United States serves as a forcing function for nations to align their modular standards to those of the United States, making it more likely to be a steady customer for domestically produced defense products and services. If partner nations build for exportability with operating systems that are compatible with those of the United States, the option to at least export to the United States will always be there. This allows partner nations to deploy systems that are interoperable with U.S. systems, strengthening coalition and joint operation efforts. Second, the United States has a high trust value; that is, it acts in good faith to honor agreements and will reciprocally provide high-quality, dependable, and compatible defense products and services to its partner nations.

Other identified factors make the United States a less valued customer. Survey respondents noted that they may feel compelled to purchase from elsewhere if newly developed capabilities were made available by European Union member states or other allied nations, especially if that country's export processes and technology transfer policies were easier to navigate. Countries also face pressure to spend domestically; investing in internal capabilities and capacity means a more independent and indigenized industrial base. These incentives may include the desire to foster local innovation, reduce reliance on foreign suppliers, decrease unemployment rates by boosting job opportunities, and develop and maintain technologies critical for safeguarding national security. Other respondents noted the importance of speed—and an oft-cited shortcoming of allied procurement of U.S. systems (Chindea et al., 2024).

During the discussing roundtable, participants mentioned that a drastic shift in U.S. trade policy with punitive tariff measures could lead to considering other sources of acquiring defense capabilities.

Defense Cooperation Agreements and Programs

There are a variety of defense cooperation agreements and programs that serve to enhance defense industrial cooperation and more easily facilitate technology transfer, including RDP MOUs, SOSAs, and NATO membership. While RDP MOUs are critical enablers to defense industrial cooperation and grant qualifying countries broadened access to the U.S. defense market, there remains institutional and regulatory hurdles that RDP MOU member nations are subjected to despite their contractual exemptions.

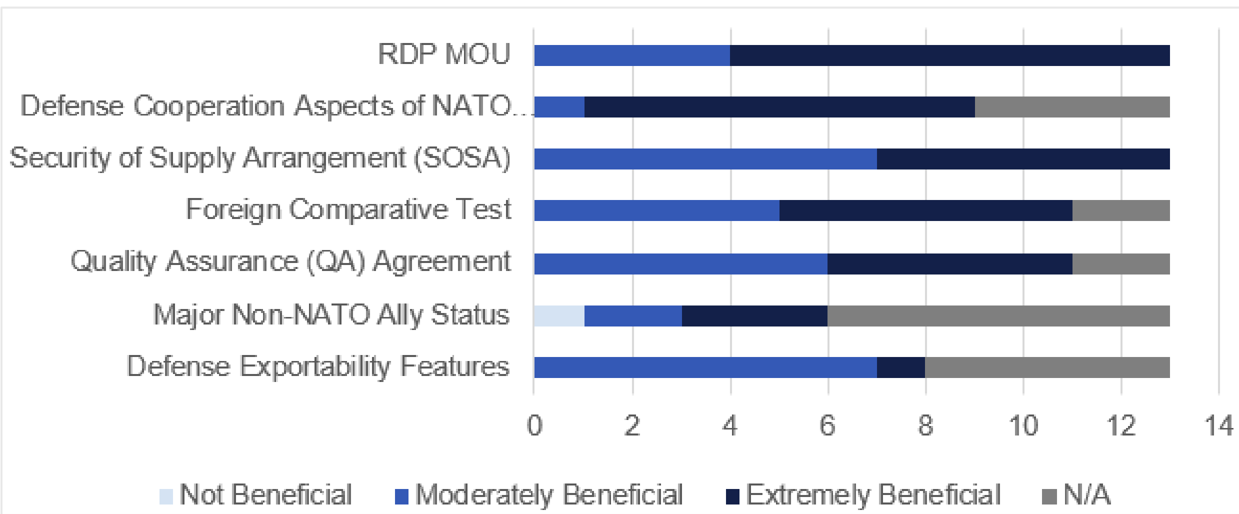


Figure 13. Value of Different Kinds of Agreements

Respondents were asked to rate defense cooperation agreements and programs on the basis of how much each benefits their home country's ability to do business with the United States. RDP MOUs were rated as the most beneficial, with nine respondents rating RDP MOUs as extremely beneficial. It should be noted that this may be a case of selection bias, since potential respondents were identified as being part of an organization comprised of RDP MOU-holding countries and that the sample that chose to respond to the survey may be more invested in the agreement.

Second to RDP MOUs, the defense cooperation aspects of NATO were positively viewed by all survey respondents. NATO has a set of programs to enable nations to work together on acquisition. There have been several joint acquisition programs, including NATO Alliance Ground Surveillance, NATO Sea Sparrow Consortium, and NATO Multinational Multi

Role Tanker and Transport Fleet (McGinn, 2023). In 2021, NATO established the Defence Innovation Accelerator for the North Atlantic (DIANA) to integrate and deliver new technologies to NATO forces. DIANA primarily focuses on "big data, artificial intelligence, autonomy, quantum, biotechnologies and human enhancement, energy and propulsion, novel materials and advanced manufacturing and aerospace" (NATO, n.d.).

NATO also has the NATO Defence Planning Process (NDPP) which allows Allies to harmonize their force and capability planning activities. It facilitates the interoperability of forces and ensures they are properly equipped and supported to undertake missions without compromising the readiness of Allies' national militaries (NATO, 2022a). NDPP is responsible for identifying requirements for NATO forces and supports capability development and acquisition (NATO, 2022a). NATO also has the NATO Support and Procurement Agency (NSPA), which delivers capabilities, logistics support, and procurement frameworks to member nations (NATO, 2022b). It also supports the weapons system lifecycle management (NATO, 2022b).

Security of Supply Arrangements (SOSAs) were viewed as generally beneficial. SOSAs allow the United States and participating nations to request priority supply of defense goods and services (Office of the Under Secretary of Defense for Acquisition and Sustainment – Industrial Base Policy, n.d.). For instance, the United States can request foreign industry to prioritize delivery under DoD contracts, subcontracts, or orders, and vice versa (Office of the Under Secretary of Defense for Acquisition and Sustainment – Industrial Base Policy, n.d.). SOSAs allow for streamlined procurement processes and may be viewed more favorably by U.S. program offices having already established a security of supply framework. They ensure partner nations are prioritized when supply shortages or geopolitical tensions arise (Office of the Under Secretary of Defense for Acquisition and Sustainment – Industrial Base Policy, n.d.). However, SOSAs are voluntary or "best effort" frameworks and therefore more about confidence building (DoD, 2024b). This diminishes their utility as binding international agreements obligate signatories to invoke the terms of the agreement.

Foreign Comparative Testing (FCT) was also viewed favorably. The FCT program allows the United States to satisfy its defense needs more quickly and cost efficiently by testing the technologies developed by allies and partners with high Technology Readiness Levels to better equip U.S. operational forces and satisfy U.S. defense needs (*Foreign Comparative Testing*, n.d.). This accelerates U.S. government acquisition from foreign industry, circumventing traditional acquisition pathways that typically include domestic capability development and lengthy and costly R&D investments (Foreign Comparative Testing, n.d.). FCT allows the United States to test partner national technologies, capabilities, and weapon systems prior to definitively procuring these systems, following a "try before you buy" model (Foreign Comparative Testing, n.d.). This approach allows roughly a third of foreign vendors to either directly partner with U.S. industry or at the very least establish a U.S. presence (Foreign Comparative Testing, n.d.). As of January 2024, 1,297 technologies from partner nations were assessed, and 307 technologies were procured/acquired into U.S. forces (Foreign Comparative Testing, n.d.).

The United States has Reciprocal Government Quality Assurance (QA) agreements with six countries: Czech Republic, Finland, South Korea, Poland, Romania, and Slovak Republic. QA agreements ensure defense products and services meet U.S. military specifications through a set of standardized procedures for testing, inspection, and certification. This reduces the risk of defective parts in critical defense systems and streamlines defense procurement processes—products certified under nations who have a QA with the United States are more readily accepted by the United States and its partners who share interoperability standards with the United States.

Defense Exportability Features (DEF) is the practice of encouraging DoD program management to design and develop exportability features early in a program's lifecycle. Designing for exportability earlier in the program's lifecycle can facilitate exports, for example by incorporating technology protection earlier in the design process to avoid expensive retrofits and costly and time-consuming redesigns to meet export control and partner-specific requirements (DAU, n.d.-b). DEF facilitates business with the United States by making U.S. defense systems more export-friendly, reducing costs for foreign buyers, and improving interoperability with allies and partners. DEF also simplifies FMS processes by pre-engineering exportable versions of systems, reducing delays caused by technology transfer restrictions.

Challenges of RDP MOUs

While the intent of RDP MOUs may be to facilitate defense trade, many respondents offered that U.S. government stakeholders were less supportive of their function. Most respondents did not think the "Buy America" exemptions are well recognized within program offices, as shown in Figure 14. Furthermore, Figure 15 shows about half of the respondents think acquisition program offices are leery of the "Buy America" exemptions they offer.

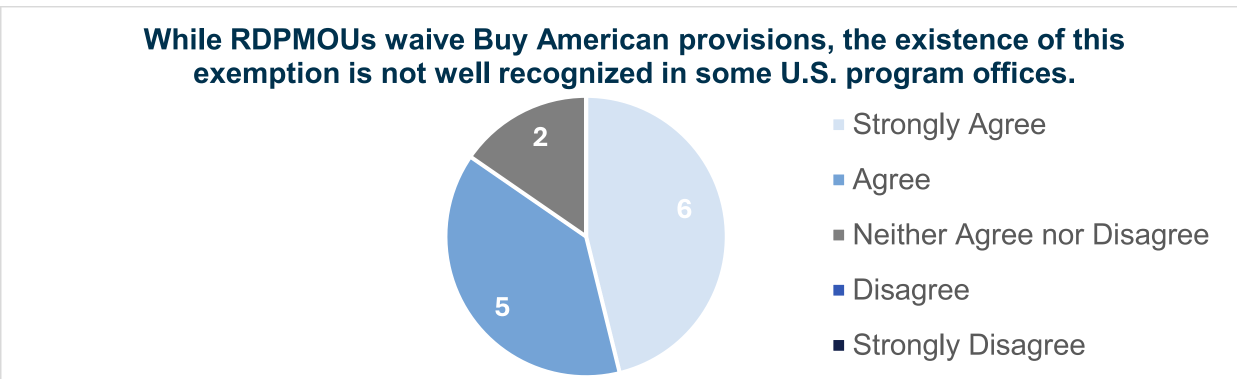


Figure 14. Perspective on Program Offices – Recognition

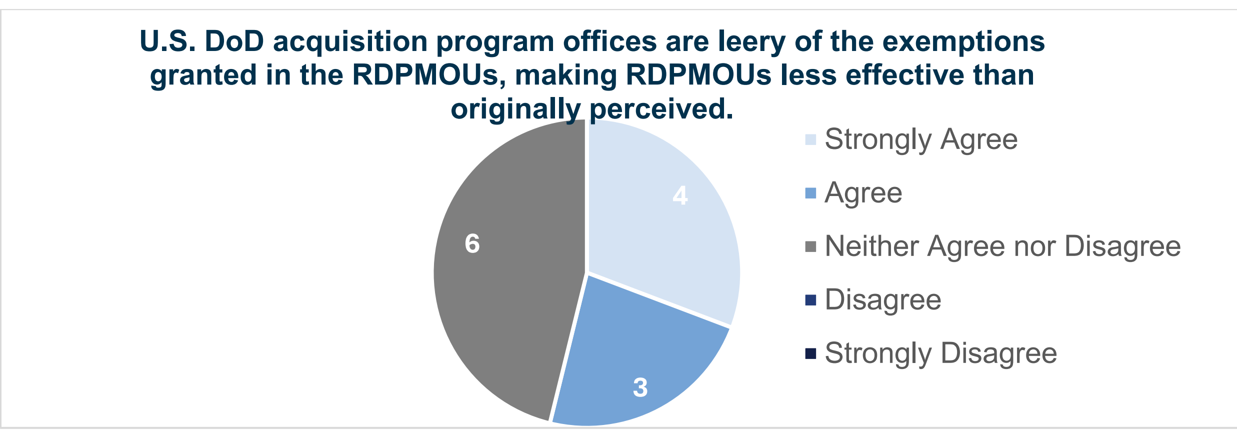


Figure 15. Perspective on Program Offices – Leery of the Exemptions

There are a variety of potential reasons for this. There may be misconceptions among acquisition contracting officers as to what exactly RDP MOUs are. And respondents suspect that program offices find it easier to default to purchasing U.S. goods and services. As one interviewee noted, "no one gets fired for buying American."

The 2024 National Defense Industrial Strategy fails to mention RDP MOUs, demonstrating a lack of awareness and understanding of the benefits RDP MOUs provide the

United States (Joint Chiefs of Staff, 2024). However, there have been welcome changes made in the Defense Federal Acquisition Regulation Supplement (DFARS), which are responsible for implementing RDP MOU exemptions (GAO, 2024). In recent years, the DFARS have become more inclusionary of RDP MOU provisions and have more systematically integrated these agreements into the broader defense acquisition framework (Federal Register, 2024a).

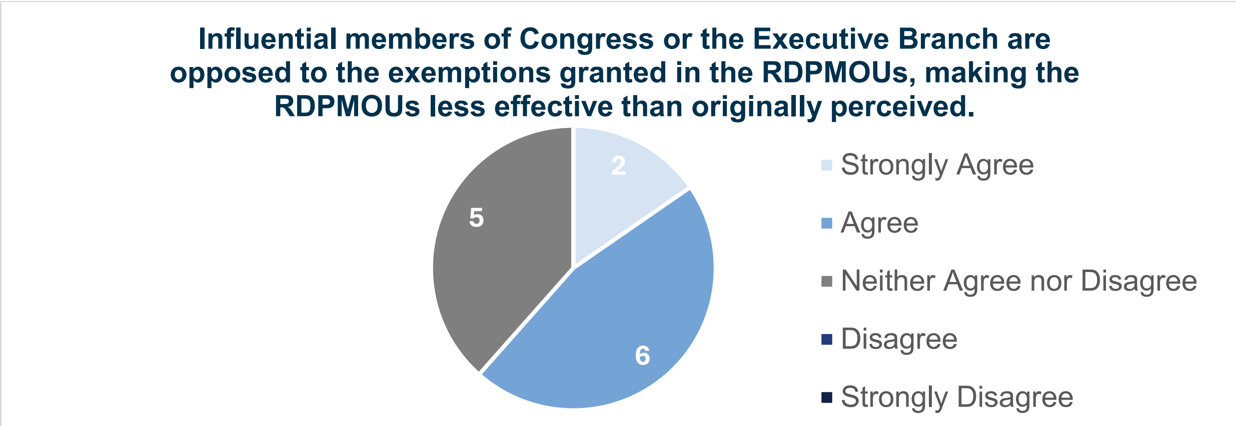


Figure 16. How Do Partners Think Congress Views RDP MOUs

According to Figure 16, participants generally are neutral or agree that Congress or the Executive Branch are opposed to the exemptions granted in the RDP MOUs. Congressional debates are frequently centered on protecting American industry without recognizing that these international agreements consider signatory country defense industrial bases as complementary to, rather than competitive with, U.S. defense manufacturing. This may lead legislators to undermine the cooperative defense industrial relationships that enhance mutual security capabilities.

Section Three: Insights and Conclusions

A lack of defense industrial integration between the United States and its allies means that potential improvements in capability are not identified and executed, which contributes to vulnerabilities that potential adversaries can exploit to wield their influence across the global strategic landscape. The United States is not prepared to solely ramp up production to meet current demand in the near term, and some capabilities may take a decade or longer to build. Without an integrated defense industrial base, allied nations will be less effective in the development, production, and sustainment of critical military capabilities—and ultimately, struggle to fight together. Partner countries may turn to non-allied nations, or even adversary suppliers, to support their basic defense needs.

Arms sales and technology transfer play a large role in ensuring the United States and its allies are properly equipped to build competitive advantage. Though export controls are required—and necessary—for any state that has a defense industry, they are designed to protect a nation’s technological advancements and intellectual property. They form the hallmark of ensuring sophisticated weapons systems do not fall into the hands of hostile actors. However, complex and lengthy export control processes may limit important partnerships that underpin deterrence.

There is a compelling business case to also be made for deeper U.S. defense industrial base integration with allies and partners. Integration extends beyond traditional arms sales and technology transfers to include co-production and co-development. While these collaborative ventures can be complex and often require higher upfront investment, they offer long-term financial benefits after achieving economies of scale and ultimately reducing per-unit costs. Additionally, deeper defense industrial collaboration positions the United States to profit from arms sales to partner nations.

The survey found that countries with an RDP MOU with the United States are hopeful that the agreement will help uplift their own industrial bases through increased cooperation and sales. A lack of consistency across administrations was identified as a limit here, with a survey response of, “As the government changes every four years, new policies such as the National Defense Industrial Strategy promoted by the current government often lose their momentum. Therefore, a defense industry cooperation policy that can be sustainably kept is needed.” There is a persistent tension in the United States between “Build American” regulations and industrial cooperation. This is even more salient given the Trump administration’s focus on tariffs as an instrument of economic policy. Every administration should remember that the benefits of cooperation, which can include increased sales as well as closer ties and enhanced interoperability, should not be forgotten in the face of the pressure to onshore.

Standing in the way of cooperation are a variety of regulations, which are designed for important functions like limiting technology proliferation to adversaries, but do create delays and uncertainly. ITAR and TSFD are the most challenging export control processes. Document markings of CUI and NOFORN should be carefully managed to ensure that they do not needlessly limit competition. Periodic reviews of the policies themselves to ensure that they are appropriately limiting technology proliferation without causing undue delay would be useful. There is also an incomplete and uneven understanding of U.S. government regulations on the part of allies and partners, including those who are DMAG participants. As these individuals play an enhancing bilateral defense cooperation, this knowledge gap may lead to unnecessary delays. Formal training offered by the United States could help facilitate both arms sales and cooperation. Strengthening the requirement to design for exportability in appropriate systems would also facilitate defense trade.

Another option to reduce the regulatory burden relates to the fact that every bilateral arrangement requires a separate review, even if two close allies are buying the same equipment. Allies working as a group to procure U.S. systems or the United States combining reviews could both be a structural solution to speed the processes. As one survey respondent suggested “The US should encourage allies to work together when they procure the same systems. If the same system is sold to several nations in a region, all NATO-members, the US should not wait for Third Party Transfer (TPT) requests for them to be able to cooperate but rather encourage this and push out that license. This will increase the total allied capability.” Creating a joint structure for FMS reviews is another option.

RDP allies were consistent in their feedback that MOUs are not particularly well understood at program offices, which limits the ability of the DoD to draw on the expertise of allies. A recent Defense Innovation Board report addressed this directly, suggesting “all DoD program managers should be trained on the RDP MoU and additional Buy American waivers and exemptions. In addition, the office that negotiates these waivers must be empowered to inform and educate the DoD contracting and acquisition workforce on the proper use of these existing authorities” (Defense Innovation Board, 2024). Consistent education as part of required acquisition certifications would address this challenge.

Bibliography

Assistant Secretary of Defense for Mission Capabilities. (n.d.). *Foreign comparative testing.* https://ac.cto.mil/pe/fct/

Bureau of Industry and Security. (n.d.). *Offsets in defense trade – Frequently asked questions*. https://www.bis.doc.gov/index.php/documents/pdfs/1677-offsets-in-defense-trade-faqs-final/file#:~:text=Offsets%20are%20industrial%20compensation%20practices,%C2%A7%202751%2C%20et%20seq

Bureau of Industry and Security. (2024, April 18). *Commerce significantly streamlines export controls for Australia and the United Kingdom advances goals of the Aukus enhanced trilateral security partnership*. https://www.bis.gov/press-release/commerce-significantly-streamlines-export-controls-australia-and-united-kingdom

Chindea, I. A., Moroney, J. D. P., Webber, S., Brin, I., Thaler, D. E., Rhoades, A., Grill, B., Cormarie, P., & Lashendock, J. (2024, July 24). *Aligning strategic priorities and foreign military sales to fill critical capability gaps*. RAND. https://www.rand.org/pubs/research_reports/RRA2438-2.html

Congressional Research Service. (2023a, March 30). *Defense primer: The national technology and industrial base*. https://sgp.fas.org/crs/natsec/IF11311.pdf

Congressional Research Service. (2023b, October 6). *The Defense Production Act of 1950: History, authorities, and considerations for Congress*. https://crsreports.congress.gov/product/pdf/R/R43767

Congressional Research Service. (2024, August 13). *Arms sales: Congressional review process*. https://crsreports.congress.gov/product/pdf/RL/RL31675#:~:text=Under%20the%20Arms%20Export%20Control,or%20services%20valued%20at%20%2450

Corben, T., & Greenwalt, W. (2023, May 17). *Breaking the barriers: Reforming US export controls to realise the potential of AUKUS*. United States Studies Center. https://www.ussc.edu.au/breaking-the-barriers-reforming-us-export-controls-to-realise-the-potential-of-auku s

Defense Acquisition University. (n.d.-a). *Foreign comparative testing*. https://www.dau.edu/cop/navaltest/resources/foreign-comparative-testing

eDefense Acquisition University. (n.d.-b). *International acquisition—Defense exportability features program*. https://www.dau.edu/acquipedia-article/international-acquisition-defense-exportability-features-program

Defense Acquisition University. (n.d.-c). *International acquisition—Foreign military sales (FMS)*. https://www.dau.edu/acquipedia-article/international-acquisition-foreign-military-sales-fms

Defense Acquisition University. (2018, February). *Defense Acquisition University (DAU) teaching note technology security and foreign disclosure (TSFD)*. https://www.dau.edu/sites/default/files/Migrated/CopDocuments/DAU%20Tech%20Security%20Foreign%20Disclosure%20%28TSFD%29%20Teaching%20Note%20-%20DSMC-I.pdf

Defense MOU Attaches Group. (n.d.). *Defence cooperation*. https://www.dmagdc.info/

Defense Security Cooperation Agency. (n.d.). *Case reconciliation and closure*. https://samm.dsca.mil/chapter/chapter-1 6

Defense Security Cooperation University. (n.d.-a). *FMS-151 foundational foreign military sales*. https://dscu.edu/course-catalog/course/12286

Defense Security Cooperation University. (n.d.-b). *Foreign military sales (FMS)*.

Department of Commerce. (n.d.). *Defense Priorities and Allocations Systems (DPAS) frequently asked questions*. https://www.bis.doc.gov/index.php/documents/sies/3553-dpas-faqs-11062024-final-locked/file

DoD. (n.d.). *AUKUS: The trilateral security partnership between Australia, U.K. and U.S.* https://www.defense.gov/Spotlights/AUKUS/

DoD. (2024a, May 14). *DOD releases open announcement through other transaction authority for U.S. and selected international partners*. https://www.defense.gov/News/Releases/Release/Article/3774005/dod-releases-open-announcement-through-other-transaction-authority-for-us-and-s/

DoD. (2024b, July 10). *Defense Innovation Board optimizing innovation cooperation with allies and partners*. https://innovation.defense.gov/Portals/63/20240710%20DIB%20Allies%20and%20Partners%20Study%20FINAL.pdf

Exemption for Defense Trade and Cooperation among Australia, the United Kingdom, and the United States, C.F.R. § 126.7 (2024, August 20). https://www.ecfr.gov/current/title-22/section-126.7

Federal Register. (2024a, February 15). *Defense Federal Acquisition Regulation Supplement: DFARS Buy American Act requirements (DFARS Case 2022-D019)*. https://www.federalregister.gov/documents/2024/02/15/2024-01220/defense-federal-acquisition-regulation-supplement-dfars-buy-american-act-requirements-dfars-case

Federal Register. (2024b, April 19). *Export control revisions for Australia, United Kingdom, United States (AUKUS) enhanced trilateral security partnership; Correction*. https://www.federalregister.gov/documents/2024/05/08/2024-10079/export-control-revisions-for-australia-united-kingdom-united-states-aukus-enhanced-trilateral

FEMA. (2024, September 26). *Defense Production Act*. https://www.fema.gov/disaster/defense-production-act#:~:text=The%20Defense%20Production%20Act%20(DPA,from%20the%20domestic%20industrial%20base

GAO. (2024, December 20). *International trade: Agencies should improve oversight of reciprocal defense procurement agreements*. https://www.gao.gov/assets/gao-25-106936.pdf

Greenwalt, W. (2019). Leveraging the national technology industrial base to address great-power competition: The imperative to integration industrial capabilities of close allies. *Atlantic Council*. https://www.atlanticcouncil.org/wp-content/uploads/2019/04/Leveraging_the_National_Technology_Industrial_Base_to_Address_Great-Power_Competition.pdf

House Foreign Affairs Committee. (2024, February 7). *Foreign military sales tiger task force: Report*. https://foreignaffairs.house.gov/wp-content/uploads/2024/02/2.7.24-FMS-TIGER-Task-Force-Report-1.pdf

Joint Chiefs of Staff. (2024). *National Defense Industrial Strategy*. DoD.

Kenlon, F. (2020, February 18). *Offset fundamentals*. Defense Acquisition University. https://www.dau.edu/sites/default/files/Migrated/CopDocuments/DAU%20Offset%20Fundamentals%20Presentation%2002112020.pdf

McGinn, J. (2023, June 26). *A "build allied" approach to increase industrial base capacity*. George Mason University. https://business.gmu.edu/news/2023-06/build-allied-approach-increase-industrial-base-capacity

National Defense Authorization Act for Fiscal Year 2024, Pub. L. No. 118–31, H.R. 2670 (2023). https://www.congress.gov/bill/118th-congress/house-bill/2670/text

NATO. (n.d.). *DIANA*. https://www.diana.nato.int/about-diana.html

NATO. (2022a, March 31). *NATO defence planning process*. https://www.nato.int/cps/ua/natohq/topics_49202.htm

NATO. (2022b, April 22). *NATO Support and Procurement Agency (NSPA)*. https://www.nato.int/cps/em/natohq/topics_88734.htm

Office of the Assistant Secretary of Defense, Industrial Base Policy. (n.d.). *Defense Production Act title III*. DoD. https://www.businessdefense.gov/ibr/mceip/dpai/dpat3/docs/DPA-TitleIII-Overview.pd f

Office of the Under Secretary of Defense for Acquisition and Sustainment – Industrial Base Policy. (n.d.). *Security of supply*. Retrieved April 25, 2023, from https://www.businessdefense.gov/security-of-supply.html

Saum-Manning, L., Marquis, J. P., Chindea, I. A., Elinoff, D., Pigott, T., & Brennan, E. (2024, August 8). *Optimizing foreign military sales roles, responsibilities, and authorities*. RAND. https://www.rand.org/pubs/research_reports/RRA2631-1.html

Siripurapu, A. (2021, December 22). *What is the Defense Production Act?* Council on Foreign Relations. https://www.cfr.org/in-brief/what-defense-production-act

U.S. Department of State. (n.d.-a). *Agreement between the Department of Defense of the United States of America and the Ministry of National Defense of the Republic of Korea concerning exchange of research and development information.* https://2009-2017.state.gov/documents/organization/135112.pdf

U.S. Department of State. (n.d.-b). *Directorate of Defense Trade Controls*. https://www.state.gov/bureaus-offices/under-secretary-for-arms-control-and-international-security-affairs/bureau-of-political-military-affairs/directorate-of-defense-trade-controls-pm-ddtc/

PANEL 13. STRATEGIC APPROACHES TO DEFENSE ACQUISITION: RICKOVER'S LEGACY, MOSA, AND DATA INTEGRATION

Wednesday, May 7, 2025	
1240 – 1355 PT 1440 – 1555 CT 1540 – 1655 ET	**Chair: Hon. Nickolas H. Guertin,** Former Assistant Secretary of the Navy for Research, Development, and Acquisition (ASN RDA) ***Would Admiral Rickover's method still work in today's complex acquisition and S&T landscape?*** Dr. Jerry Kim, EW Specialist, Envisioneering Inc. ***A Modular Open Systems Approach (MOSA) to Enable Technology Transition*** Dr Kelly Alexander, Chief Systems Engineer, System Innovation ***Using Data Analytics and Dashboards in a Research Organization Environment for Project Management*** David A. Lechner, Head, Signal Identification Section, Naval Research Laboratory

Hon. Nickolas H. Guertin—was sworn in as Assistant Secretary of the Navy for Research, Development, and Acquisition (ASN RD&A) on December 20, 2023. A Presidential appointee confirmed by the United States Senate, he lead the Department of the Navy's (DON) Research, Development, Acquisition, and Sustainment programs and the DON's contracting community. Prior to this role, Mr. Guertin served as the senior advisor to the Secretary of Defense on operational and live fire test and evaluation of Department of Defense weapon systems as the Director, Operational Test and Evaluation.

Mr. Guertin has an extensive four-decade combined military and civilian career in submarine operations; ship construction and maintenance; development and testing of weapons, sensors, combat management products including the improvement of systems engineering; and defense acquisition. He has also performed applied research for government and academia in software-reliant and cyber-physical systems at Carnegie Mellon University's Software Engineering Institute.

Over his career, he has led organizational transformation, improved competition, and increased application of modular open-system approaches, prototyping, and experimentation. He has also researched and published extensively on software-reliant system design, testing, and acquisition. He received a Bachelor of Science in Mechanical Engineering from the University of Washington and an MBA from Bryant University. He is a retired Navy Reserve Engineering Duty Officer, was Defense Acquisition Workforce Improvement Act (DAWIA) certified in Program Management and Engineering, and is a licensed Professional Engineer (Mechanical).

Mr. Guertin vacated the position of Assistant Secretary of the Navy for Research, Development and Acquisition in January 2025.

Would Admiral Rickover's Method Still Work in Today's Complex Acquisition and S&T Landscape?

Jerry T. Kim, PhD—Jerry Kim received his BS in Mathematics from the United States Naval Academy (Annapolis, MD) in 2000, MS in Physics from the Naval Postgraduate School (Monterey, CA) in 2007, and MS in Mathematics and PhD in Mathematics from Rensselaer Polytechnic Institute (Troy, NY) in 2012 and 2015, respectively. He received his commission in the U.S. Navy as a Surface Warfare Officer with an Engineering Duty Officer (EDO) Option in 2000. He then went to serve as the First Lieutenant on the USS SIMPSON FFG-56 from 2000–2003. From 2005–2007, he served as the Combat Systems Officer on the USS ROBIN MHC-54. He was re-designated as an Engineering Duty Officer in 2004. He then served as the Ship Superintendent for Japan Regional Maintenance Center (2007–2008) responsible for the dry docking and maintenance. He served as the Deputy Shore Branch Head/Project Controller (2008–2010) for SPAWAR SYSTEMS FACILITIES PACIFIC YOKOSUKA, JAPAN, interfacing with the U.S. Seventh Fleet, program office and the warfare center for the planning and execution of naval programs. He then worked at the U.S. Naval Research Laboratory for the Radar Division from 2013–2014 and the Tactical Electronic Warfare Division from 2014–2016 as a government researcher. He joined the MITRE Corporation from 2016–2021. He was a researcher and engineer for various IR&D programs in signal processing, radar, EW, communications, and cyber. He was the principal investigator for novel EW-Cyber research, and he provided program support for Navy and Marine Corps acquisition programs. He served as the Department Chief Engineer overseeing products and advising government partners. He joined Envisioneering Inc. serving as a senior EW specialist to the Office of Naval Research in 2022. [jkim@envisioneeringinc.com]

Abstract

The easy access of advanced and capable microelectronics has lowered the barrier for technologists to participate in fields that were traditionally secluded for state actors, the consequences of which have inspired the public to focus its resources to compete without boundaries and at great speeds. Over the last several decades, technology innovation has moved from being defense-led research to commercial-led research, resulting in the ubiquitous presence of advanced sensing, signal processing, and amplifying technologies which have placed large stresses on defense systems. The demand for transitioning advanced technologies for the defense environment has surpassed the capacity of what the traditional acquisition and science and technology (S&T) communities can provide. This paper addresses some of the S&T challenges that ADM Rickover faced when transitioning the nuclear reactor technology, discusses the impacts of the Goldwater-Nichols Act and the current landscape, provides suggestions for contracting and reforms needed, and provides some lessons from history and applies it to present times.

Current Speed of Technology

In the past, there were huge technical and capital barriers for acquiring parts and integrating them into large systems to perform tasks that require resources from a government or a modern-day wealthy aristocrat. Hence the Department of Defense (DoD) only had to be concerned with large nation states in technological competition. However, the speed at which technology moves today, from an idea to the marketplace, has increased drastically due to advancements in additive manufacturing, advanced modeling software, microelectronics, and access to private capital. Over the last several decades, technology innovation has moved from being defense-led research to commercial-led research, resulting in the ubiquitous presence of advanced sensing, signal processing, and amplifying technologies which have placed large stresses on defense systems.

For example, today, the arrival of software defined radios (SDRs) have wrought havoc on the Electronic Warfare community. Advanced waveform generation can be done through a USB-based, small-form-factor waveform generator that can be purchased for $2,000 at Signal

Hound. An SDR that can be purchased for under $1,000 on Amazon can be configured to interact with an arbitrary waveform generator to broadcast any type of waveform for any application (communication, radar, etc.). A small USB-based SDR (Figure 1) can be purchased for a few hundred dollars. On plugging an antenna into it, one can receive the radio frequency (RF) to generate a very decent waterfall diagram or even to perform time-difference-of-arrival (TDOA) calculations, if one combines multiple SDRs. Moreover, any developer can create an application to display the calculations.

Often these SDRs are regularly updated in firmware and software as fast the market can deliver them. There is a growing community of SDR users filled with hobbyists and people from academia, industry, and government. These SDRs are getting more and more capable every year with no evidence of slowing down in the advancements of microelectronics and signal processing. Furthermore, the sharing of code on platforms like GitHub has significantly increased the speed of development in signal analysis for classification, signal modulation for communication and sensing, just to name a few. Many software developments are then tested in realistic environments through experiments by users around the world. This activity is a type of crowd-source development that might move the technology at speeds that have never been seen before. Combining an SDR with current advancements in artificial intelligence (AI) and machine learning (ML) may lead to an explosion of advanced concepts and capabilities in this domain. The current development speeds of technologies such as this are alarming.

Will our current legal structure and defense acquisition framework be ready to handle what lies ahead in the 21st century?

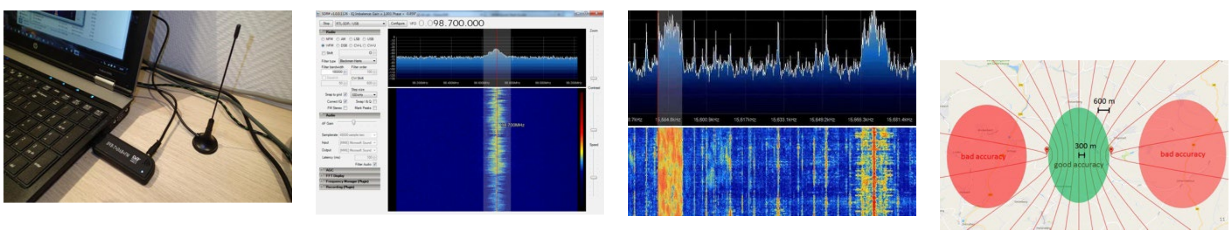


Figure 1. Current Market SDR Capability

Large Innovative Technology has Large Consequences

One day a disgruntled commanding officer of a nuclear submarine wrote a letter to ADM Rickover, stating that if nuclear propulsion was so harsh in its demands, the navy would do well to find another propulsion system. "Rickover, tossing the letter aside for a moment remarked that technology brings its own discipline, a truth he was not sure society understood. What he meant was quite simple: the stronger the forces of nature harnessed by a technology, the more discipline was needed by those who design, build, operate, and maintain the products of technology" (Duncan, 1990).

ADM Rickover realized that the more revolutionary and more powerful a technology becomes, it needs to be controlled by the highest discipline exercised by a strong technical group, which itself was the product due to this discipline. He realized as early as 1946 that nuclear technology came with an immense responsibility and that to incorporate nuclear technology in the fleet could not be done through the normal navy or industrial organization (Duncan, 1990, p. 279). If a technology was so critical that a failure would lead to catastrophe, then having the "discipline of technology" becomes paramount, and the criticality is proportional to consequences of the failed technology.

It should be of no surprise that a technology which has a massive impact would also come with the greatest critics and inertia. The level of difficulty and effort to advance a technology to a program of record increases nonlinearly. A technology not only has to overcome technical challenges, but it also must overcome political challenges. The more impactful a technology becomes, the more important the politics and storytelling are.

ADM Rickover not only acquired Congressional support despite the pushback from the Naval bureaucracy, but he also took great care in recruiting, training, and creating the conditions under which people were able to perform at their best potential. He protected his people from the red tape and viciously guarded his time. He would assume full responsibility for what he would consider political and was zealous in protecting his people from any distractions. He maintained that discipline and kept the highest standards for himself and the organization that he ran (Duncan, 1990). The higher the impact the technology had for the fleet, the more disciplined and the more persistent he had to be in order to transition the technology.

Department of Defense S&T Lexicon

Figure 2. Technology Readiness Level—Budget Activity Map

Figure 2 comes from the DoD 500 Acquisition Guidebook and shows the mapping between a Budget Activity (BA), which is a type of the Research, Development, Test & Evaluation (RDT&E) funds, and the Technology Readiness Level (TRL). A TRL assesses the readiness of the technology in question. Both terms are used when assessing a technology and are shared lexicon across the services.

A Short Survey of Atomic Physics

Figure 3 is a short survey of scientific accomplishments in atomic physics. Since it is impractical to cover the immense amount of major discovery and work done in this area, only a few discoveries are selected to provide an overall appreciation of how much basic research (6.1) and applied research (6.2) had to be done for ADM Rickover to make that first naval nuclear reactor. He rested on the shoulders of giants who made gargantuan scientific contributions. He leveraged their findings and applied his engineering prowess, grit, and resilience to design and develop the prototype of the Navy's first nuclear-powered submarine.

Henry Cavendish's discovery of the hydrogen atom to Enrico Fermi's first controlled nuclear chain reaction are shown in Figure 3. The list of discoveries in Figure 3 is not meant to be an exhaustive but rather is intended to give the reader an appreciation of what level of discoveries was needed before the nuclear reactor became possible. By attempting to categorize scientific discoveries, the discussion of advancements in Atomic Physics can be made in the DoD framework. The demarcation from 6.1 to 6.2 was made on the basis that the activities in physics were starting to shift from trying to gain the fundamental knowledge of the atom to a more applied exploration of the atom. After Henry Cavendish's discovery of the hydrogen atom in 1766, more than 130 years elapsed before J. J. Thompson's discovery of the electron in 1897. Generations of physicists and mathematicians had to develop the mathematical tools to explain the physical world that they were observing. From 1897 onwards into the twentieth century, scientists were starting to use the knowledge that they had gained to applications by performing experiments, developing theories to explain the experiments, and then developing the mathematical tools to predict and explain the nature of the atom reliably. These efforts crossed into the 6.2 world and were filled with numerous scientific ventures as the world became fascinated with the unseen world.

Figure 3. Survey of Atomic Physic

Heinrich Hertz was discovering the photoelectric effect, and Albert Einstein explained its phenomenon using the concept of quanta of light, which later was influential in the development of quantum theory. Ernest Rutherford discovered the alpha and beta particles emitted by uranium. Niels Bohr presented the quantum model of the atom, and Arnold Sommerfeld built on that by replacing circular orbits with elliptical orbits. Robert Millikan defined the fundamental unit of an electric charge. Louie de Broglie suggested that electrons would have wave-like properties in addition to particle-like behaviors. Werner Heisenberg, Max Born, and Pascual Jordan developed the quantum matrix mechanics. Erwin Schrödinger improved the work showing that the wave and matrix formulations of quantum theory were mathematically equivalent. Max Born then showed the probabilistic nature of the wavefunctions. A collaboration between Max Born and Robert Oppenheimer introduced the Born-Oppenheimer approximation. Subsequently, a series of major discoveries of coordinated scientific work led to Eugene Wigner to develop the theory of neutron absorption by the atomic nuclei, which then led to Enrico Fermi making the first controlled nuclear chain reaction in 1942. It was this broad coordination across international lines and research interests among scientists that allowed for Robert Oppenheimer to know how to put the team together who would understand the known physics at the time to develop the fission bomb for the Manhattan Project.

The takeaway is that it took 45 years to make the first controlled nuclear reaction after J. J. Thompson's discovery of the atom, which itself was based on the previous 130 years of 6.1 research on theoretical fundamentals. During those 45 years of very productive scientific coordination and endeavors to 6.2 research, no one could have predicted having the entire world at war. Fortuitously, the products of investments in 6.1 and 6.2 were in place, so that the development of the first atomic bomb was possible. From the first controlled nuclear chain reaction in 1942, which would be considered 6.3 by the current DoD definition, to the test of the first fully functional nuclear bomb in 1945 was three years. By placing immense national resources, the working prototype of an atomic bomb crossed over to 6.4 in three years. For the reactor, it would take 15 years from 1942 to 1957 until the first working civilian nuclear reactor became fully operational at the Shipping Port Atomic Power Station. ADM Rickover understood the what impact this technology would have for the Navy.

ADM Rickover made the case to Chief of Naval Operations ADM Chester Nimitz, who understood what this technology would bring to the Navy and made a strong case to the Secretary of the Navy (SECNAV) John L Sullivan. ADM Rickover received his charter and worked with the scientists at the Oak Ridge Laboratory to develop the nuclear reactor. The USS *Nautilus* (SSN-571) completed its epic journey submerged through the North Pole in 1958 with the newly developed reactor. From Enrico's demonstration in 1942 at the 6.3 level to the fully operational submarine commissioned in 1954 at the 6.5+ level took 12 years. It was a valiant, political effort to cross the "valley of death"[1] to a fully operational submarine. And the nuclear submarine catapulted the United States to a significant technology advantage over her adversary, so far ahead that the Soviet Union would end up playing catch-up for the rest of the Cold War. Furthermore, because of this technological breakthrough in nuclear propulsion, the United States still holds this key advantage in nuclear propulsion against any near peer adversaries.

Challenges Faced by Rickover

"Nothing worthwhile can be accomplished without determination. In the early days of nuclear power, for example, getting approval to build the first nuclear submarine—the Nautilus—was almost as difficult as designing and building it. Many in the Navy opposed building a nuclear submarine." –ADM Rickover

The Cold War was in full swing, and the United States had a great advantage with the large technology gap provided by forward submarine forces that were persistent and quiet, with powerful propulsion. For strategic arms to work well, the ships had to have the ability to have long on-station times. The Manhattan Project proved that a controlled chain reaction could be achieved, and the proof of concept was demonstrated with a working prototype. However, the engineering journey needed to be done for ship propulsion, which still needed the development of reactor fuel with long life and high integrity. To accomplish that, materials that could withstand

[1] The "valley of death" in research refers to the challenging phase where promising technologies or ideas struggle to transition from initial research to commercialization due to funding gaps, regulatory hurdles, and other factors.

intense and prolonged radiation, the development of coolant to remove heat quickly from the system, and a long list of other capabilities were needed.

Although ADM Rickover knew nothing about the Manhattan Project, he was mesmerized by its achievement. At the Oak Ridge Laboratory, he followed many key scientists, some of them many decades younger than him. He would listen to their explanations of complex atomic physics and would study the equations on the blackboard. It was a humbling experience for him as he did not have this background. He was thoroughly impressed by their command of the technical knowledge, but he had issues. For example, at the outset of the nuclear propulsion program, he discovered, to his dismay, that the scientists on whom he was so keen had no awareness of the principles and standards of safety and reliability. For ADM Rickover, it was a disappointment, but the lesson to be taken here is that the scientists who are trying to solve 6.1/6.2 problems are not trained to be engineers. They are trying to understand the fundamental physics, not engineering a critical system to be used in a real-world environment. Consequently, the Daniel reactor project was never built (Duncan, 1990, p. 192).

ADM Rickover had to start over and implement an arduous program of study, interviewing, and learning. He realized that there had to be a change in mindset. He wasn't against science or scientists. But he learned that "scientific truth was not engineering truth." The worldview and approaches were different. One was about discovery, and the other was about engineering a practical working system. Since the 6.2/6.3 work had stopped, it was time to learn from the scientists and transfer that knowledge to seasoned engineers who would apply the necessary rigors to design and engineer a safe and effective reactor plant. It was time to bring in industry—Bettis Atomic Lab under Westinghouse Electric Company was selected to start designing and building the reactor. The takeaway is: in order to transition, the science should stop and the engineering must begin.

ADM Rickover also had internal Navy inertia. The Navy was concerned about competing against the other services in delivering strategic arms against the Soviet Union. They did not understand the novel technology of the reactor and were pushing back against ADM Rickover. But about the same time, the Soviet Union's ADM Gorshkov was pushing hard towards a nuclear hegemony with the submarines to create an uncomfortable technology gap with the United States. This gap never did materialize because ADM Rickover drove the nuclear reactor technology into being through his grit and vision. The Navy leadership was focused on nuclear weaponry, not in developing nuclear engines. According to Captain Edward L. Beach in 1947, the Navy was focused on countering the U.S. Air Force's claim that only the long-range bombers could deliver nuclear weapons. The Navy felt that it would lose its mission to the Air Force. This sort of fierce inter-service rivalry would be addressed decades later by the Goldwater-Nichols Act, as the inter-service rivalry had led to serious failures in military operations.

Therefore, the Navy focused all its resources into the development of nuclear weapons for aircraft carriers. Developing a nuclear submarine was very much a secondary objective. ADM Rickover was faced with the largest resistance and criticism coming from the Navy as he understood what this technology would bring to the Navy, and how the nuclear-powered Navy could alter the tide of the Cold War. The demonstration of the nuclear reactor and its ability to provide propulsion, giving the Navy a submarine fleet that could be persistent on station around the world was clear, but that was not interesting to the Navy. Though the reactor technology had crossed over to 6.4, the novel reactor technology required a champion. ADM Rickover found an advocate in ADM Chester Nimitz, which incurred the ire of the Naval leadership.

Even after his victory, ADM Rickover had to "defend" the nuclear reactor technology from the Navy, as the expenses of building ships and submarines using this propulsion technology were significant and the operating and maintenance instructions of the reactor were extremely strict. Although the completed journey of the USS *Nautilus* clearly showed the importance of nuclear propulsion, with the reactor technology at TRL 9, the future was not certain until it received the support of Congress. The compelling story was to make all surface and submarines powered by a nuclear reactor. This approach was costly, but the value was clearly there. However, there was a new competing radar technology for surface ships called AEGIS which would provide air defense against Soviet threats. This new important technology impacted the agenda of making the Navy be powered by nuclear reactors. It was too expensive to build both an AEGIS system and a reactor on surface ships. The choices were two AEGIS ships with no reactors or one AEGIS ship with a nuclear reactor. From the budgeting point of view, it was better to have two AEGIS ships (Duncan, 1990).

Having a technology achieve maturity is insufficient for a transition to a fielded system. Transition requires an unrelenting champion. Once reactor technology crossed over to 6.4, transitioning it became, as Rickover would put it, “political;” that is, the difficult problems of scheduling, budgets, stakeholder adoption, etc., needed to be overcome. Those problems were not technical in nature, but they were needed for the continued movement of the technology through the bureaucratic system. ADM Rickover had many challenges, but he did not face the legislative and regulatory burdens that many current innovators face when transitioning novel technology.

A New World Under the Goldwater-Nichols DoD Reorganization Act of 1986

The Goldwater-Nichols Department of Defense Reorganization Act of 1986 was signed into law on October 1, 1986. The chairman of the Joint-Chiefs-of-Staff General (David C. Jones) started the process to push for reforms in the DoD, but the House Armed Services Committee did not have much interest. However, through Senators Barry Goldwater and Samuel Nunn, the Senate Armed Services Committee pushed for the legislation to make major reforms within the DoD (Locher, 2002). National leaders understood that a reform was needed within the DoD as there were fierce rivalries among the services that led to technology duplications. President Ronald Reagan requested the Packard Commission in 1985 to perform a study to provide recommendations to reform the DoD, which fed into the creation of the Goldwater-Nichols Act. The legislation was to reduce inter-service rivalries and address many of the inter-service problems. The Packard Commission addressed serious acquisition problems where systems were acquired within the services that were not able to interoperate. The Goldwater-Nichols Act was a response to series of military failures and discovery of much fraud and waste. The need for the legislation became apparent after series of joint operation failures, such as: (1) the SS *Mayaguz* incident during the Fall of Saigon, where a joint rescue mission resulted in casualties from lack of coordination: (2) Operation Urgent Fury in October 1983 in Grenada, where there were significant joint cooperation issues between the Army and Navy; and (3) Desert One, a 1980 Iranian Hostage Rescue mission that ended with various aborted missions leading to fatal accidents from lack of joint cooperation.

The legislation created the following significant changes (Bond et al., 2016).

- Clear military chain of command from operational commanders (i.e., combatant commanders) through the Secretary of Defense (SECDEF) to the President.
- Service Chiefs are responsible for training and equipping forces, while explicitly clear that they were not in the operational military chain of command.
- Chairman of the Joint Chiefs of Staff (JCS) was elevated above the other service chiefs, being the military advisor to the President.
    - Creation of the Vice Chair position.

- Required military personnel entering strategic leadership roles to have experience working with their counterparts from other services.
- Creation of an organization for the services to collaborate when developing capability requirements and acquisition programs, thereby reducing redundant procurement programs. This established the position, Undersecretary of Defense for Acquisition.

The legislation created the Undersecretary of Defense for Acquisition [USD (A)] and consequently created the Program Executive Offices (PEOs) for the services. It also created the Vice Chairman of the Joint Chiefs of Staff, a position that presided over the Joint Requirements Oversight Council (JROC). The Vice Chairman also held the Vice Chair position for the Defense Acquisition Board (Locher, 2002). This law made significant changes in the military. The service chiefs were no longer involved in military operations but rather in the training and equipping of the services, and consequently, they controlled the requirements process which is defined by Joint Capabilities Integration and Development System (JCIDS). The creation of the program offices for the acquisition of the capabilities for each service fell under the service secretaries to the Undersecretary of Defense. This consequential legislation impacted three DoD processes—Planning, Programming, Budgeting and Execution (PPBE); the Defense Acquisition System (DAS), which is defined by the DoD Instruction 5000 series; and JCIDS. The law directed services to share technology and development efforts through the USD (A). The intention was to streamline what the services were doing so that duplication would be reduced while increasing procurement efficiency (Locher, 2002). This law was further amplified through the Defense Acquisition Workforce Improvement Act (DAWIA), the Weapons Systems Acquisition Reform Act (WSARA), and the National Defense Authorization Act of 1987.

The impacts were consequential. The entity that did the asking of the technology was now separate from the entity acquiring the technology. Not only did the entity asking for the technology have the ability to ask for a capability, but it could also get to direct the "how" through the JCIDS process. In a 1974 talk, ADM Rickover made some comments that it was not wise for the military staff to dictate the "how". He cited a few examples in history where this did not end with good outcomes. His position was that the CNO and his staff were trained military experts not technical experts. They would not know how to dictate the "how" and would then have to expand the staff in order to perform this task. The CNO and his staff would be distracted to be executing on the acquisition mission where he believed that that should be the function of the SECNAV while leaving the warfighting doctrine and warfighting to the CNO (Rickover, 1974). The legislation removed that capability. The actual warfighting was to be done through the combatant commands. The legislation created the senior acquisition executive who was answerable to the service secretary, and the requirements process rested in the hands of the service chiefs. The service chiefs get to play a technical role to drive the acquisition function from the service secretaries.

Consequently, the JCIDS process supported the JROC and the Chairman of the JCS by identifying, assessing, validating, and prioritizing joint military capability requirements. It was meant to be a transparent process that allowed for the JROC to balance the demands of the military.

Thirty years later, the DoD is still struggling with trying to transition important technologies to the warfighters. The late Senator John McCain, Chairman of the Senate Arms Services Committee, said in 2015, "It was about 30 years ago that Goldwater-Nichols was enacted, and the one thing we are committed to is a thorough and complete review of Goldwater-Nichols Act" (McCain & Thornberry, 2015). The law is designed for the Cold War, which was a contest between the two superpowers under stated agreed-upon rules, but is it sufficient for the 21st century?

The JCIDS process involved the service chiefs in the technical direction of their requested technology; while the service secretaries became more involved in acquisition at much higher TRLs, restricting themselves largely to budget, schedule, and performance. The result over time has been that the technologists who once lived under the SECNAV were less needed and demands for technical people within the CNO and OPNAV increased, burdening the CNO with more tasks. The different chains of command and authorities, along with the distributed nature of requirements and acquisitions, have diluted the responsible party of making the technology to transition.

Technology does not understand organizational structure, nor does it care about the laws that command it to comply. Technology only understands physical laws and obeys only the demands of nature. By requiring those who ask and those who acquire to be separate personalities, it became necessary that the two entities must find a delicate balance, further constrained by the budgeting process of PPBE. Consequently, the action officer and the requirement officer in the POM process of the PPBE must agree with the current leadership visions and policies. And since those positions are transient in nature, the technology that has been in demand under one leadership could shift to different priorities. Therefore, the desire to take on higher risks to write requirements for novel technologies has been curtailed by the demand to comply with the current vision of the leadership.

In the 1980s, the findings of the Packard Commission showed that reforms were needed in the DoD, and the Goldwater-Nichols Act was hailed as the most significant legislation that changed the way the DoD operated. The creation of the Undersecretary of Defense for Acquisition and the corresponding Senior Acquisition Executives changed the way the DoD controls and manages procurement. And this in turn made a significant impact on the S&T community who attempts to transition innovative technology to the warfighter. ADM Rickover did not have to navigate through this new landscape.

PPBE Process

The PPBE process is the process by which the DoD acquires its funds to execute within its charter. To acquire the funds from Congress, various policy and procedural documents are associated with preparing, submitting, and defending the annual Program Objective Memorandum (POM) submission. The POM process is calendar driven and is often myopic in nature; that is, the POM addresses only the budgeting cycle. Innovations require longer cycles and a long steady plan. Much of the woes in S&T can be traced back to the long-term nature of doing research and development and the short-term nature of the budget cycles.

The Action Officers (AO) that serve the CNO in putting the POM together must understand that the POM belongs to the CNO and must align to the CNO's visions and priorities and must help deliver the POM to the Secretary of the Navy. An AO is generally not seasoned in the PPBE process and often requires going through two POM cycles to become effective or to acquire a "journeyman" understanding of being a requirements officer (RO; Blickstein et al., 2016). Because of the transience of the CNO, the priorities continue to change, and many S&T programs were cancelled on the whim of the CNO. In order to advance bleeding edge technology, there needs to be stability and continuity. ADM Rickover would often quip that he had to protect the nuclear reactor power program from the Navy (Duncan, 1990). The technical challenges of the technology do not change. However, with every new POM cycle, the technology that is being developed must bend to the demands of the changes set by each new CNO.

Figure 4 is an illustrative way to understand the value propositions of innovation and technology. The left side indicates the basic and applied research areas. On the right of the 6.4 line are the engineering and development areas. The innovation vectors on the side represent the level of novelty. The technology velocity vector represents how quickly the technology moves. On the left in the basic and applied research world, the technology moves slowly as the methods of science are applied towards exploration and understanding. The technology vector on the right is high because it is an engineering problem, resulting in a prototype using sound engineering principles. The technology moves quickly, and systems are built. The performers are different in the 6.1/6.2 space and the 6.4+ space. Typically, academia and service laboratories are involved in basic research, and industry is on the right side of the diagram.

There is a tendency to move towards areas of high-technology velocity areas. The stakeholders prefer the technologies binned in the 6.4+ areas. The products are polished, and the delivery times usually can fit into a POM cycle. The impact of the technology would be acceptable but not outstanding. The opportunities for technology to surprise or leap far ahead of the competitors as ADM Rickover did with the nuclear reactor would not be realized under the current paradigm. To realize groundbreaking technologies, a more holistic approach is required, and PPBE, JCIDS, and DAS must be strongly aligned.

JCIDS

When researchers are developing new innovative ideas, there is a heavy emphasis to transition the technology. The technology should map to a capability gap, or it could address an urgent needs statement. But what is not usually clear is how the innovation makes an impact at the warfare level while it is still on the left side of 6.3 line. The JCIDS process generates requirements. More in-depth mission engineering tools are needed that can connect the technology to the mission, which can be shared with the researchers developing the technology. According to Freeman, mission engineering involves forecasting the performance of future capabilities to inform future requirements and acquisition priorities (Freeman et al, 2024).

For future successful S&T transitions to the warfighter, the author feels that there needs to be a focus on the warfighting doctrine of tomorrow that would create the warfighter needs and that those needs need to be translated to technical problems that scientists and engineers can solve. Presently, OUSD R&E describes mission engineering in a five-part process:

1. Frame the mission problem
2. Characterize the mission (e.g., mission sets)
3. Model the mission architectures
4. Perform analysis and evaluate tradeoff
5. Document results and recommendation

Figure 4. Transition Under the DoD S&T Framework

The mission engineering tools should be advanced and developed in all warfare areas to identify the key capability gaps. In their paper, Freeman et al. also discussed an AI integrated strategy. The advantages are well articulated in the paper, and there is no doubt that a more effective way to generate warfighting requirements and technical requirements is needed.

Many technologies that have been developed sit on shelves today because of the lack of adoption by the warfighters. Much of these woes can be traced back to untraceable requirements (i.e., the warfighters did not "ask" for them). Many advanced concepts and innovations do make it to 6.3 through stakeholder interests, but the difficulty in crossing the "Valley of Death" is two-fold: the lack of funding and a lack of compelling narrative to make the jump. Often, the warfighters may not know how to ask for innovative technologies or merely do not want to ask for them (change the status quo). The success of ADM Rickover was his ability to communicate a compelling narrative on how the nuclear reactor mattered for the Navy to key stakeholders. Naval nuclear reactors came into being not because of the Navy, but despite it. When ADM Rickover got naval nuclear reactors to 6.4+, the value of the nuclear reactor was instantly apparent when the USS *Nautilus* completed its journey through the North Pole. The champion must be able to visualize this reality and to articulate this message when the innovation is still at 6.3. The transition process comes with high risks, and the champion needs to be able to assess and accept those risks, to build a phenomenal team, and to maintain the incredible "discipline" required to mature the technology.

Novel Technology Adoption Requires a Champion

The nuclear reactor was not the only technology that came with bureaucratic inertia. In its infancy, radar technology proved to be extremely critical during World War II. It was a primary contributor for the Germans losing the air campaign in the Battle of Britain. The Royal Air Force was able to detect and engage the Luftwaffe and contributed to the victory of the battle. After World War II, the world shifted to a Cold War which was a quiet war between the United States and the Soviet Union. The fierce competition between the two nations led to the development of the phased array radar concept around 1949, and there were many proponents of the concepts. For example, MIT Lincoln Laboratory started developing the phased array around 1958 (Fenn et al., 2000).

Figure 5 shows an artist's concept of the phased arrays that were being developed in the nation's laboratories, post-World War II. From the early concept to the development of the phased array was 10 years, spanning from late 6.2 to 6.3. The phased array concept was not easily adopted, and it came with fierce resistance from key radar figures at the time such as Merrill Skolnik, with as many critics as advocates. Although it may not be a debatable item today, the technology was at a crossroad of being stored away in a warehouse or being transitioned as part of an advanced radar system.

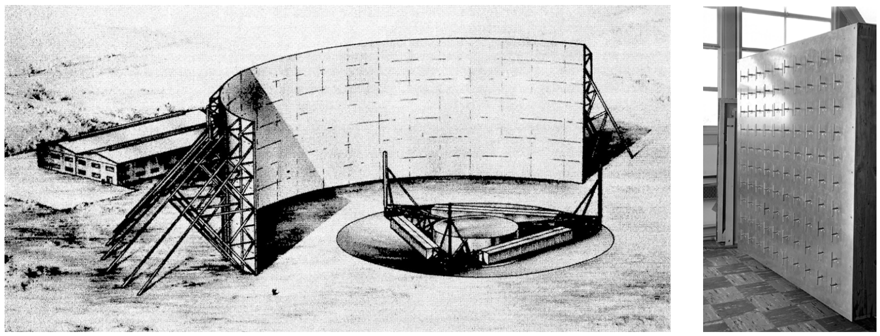


Figure 5. (Left) 1950s Era Hybrid Phased Array Radar Combining Mechanical and Electrical Steering. (Right) An Early L-band Dipole Phased-Array Test Bed Developed by the Sperry Rand Corporation, Used in the Lincoln Laboratory Array Investigation During the 1960s (Fenn et al., 2000)

The transition of the phased array was not clear and had similar transition challenges as did ADM Rickover with the nuclear reactor. Getting the technology across the "Valley of Death" required a champion. Getting it over the 6.3 line to a full system required both the technical and political maneuvering as was the case for ADM Rickover with the Navy reactors. The Navy, at the time, required an advanced air defense shield, and RADM Frederic S. Withington delivered a report to the SECNAV on May 15, 1965, recommending five major items—(1) a phased array S-Band radar to search and track air targets, (2) six slaved X-band radars for illumination and fire control, (3) a digital control system compatible with the Naval Tactical Data System, (4) a standard missile that could be directed in flight, and (5) a dual rail-launcher. With this report, the case for the phased array radars was set in stone, and a prime was selected to develop the radar, despite the fierce resistance of the technology and the lower TRL at the time this report was made to the secretariat.

Understanding the impact that the phased array would have, RADM Wayne E. Meyer had a slogan "AEGIS at sea" in 1971 (Meyer, 2008). His mindset was "build a little, test a little, learn a lot." He was committed to pushing this high-risk 6.3 technology over the "Valley of Death" to 6.4 and beyond. Since the phased array was novel, many engineering challenges needed to be retired. For example, developing the phase shifters and the amplifiers that combined the power out of each array were highly technical, high-risk challenges. In 1975, RADM Meyer became the founding Program Manager for the AEGIS Shipbuilding Office (PMS-400), and he implemented rigorous system engineering discipline throughout the organization and was into the program details in much the same way that ADM Rickover was for the nuclear reactor program. In his later years, he attributes much of the transition of the AEGIS radar program to the people that he was able to muster to execute the program.

In his 2008 interview, a year before his passing, RADM Meyer stated that there were many engineering challenges and many critics of the AEGIS program. The costs of these systems were too high, and it was not clear that the funds would be there to execute the program. What he realized then was that if he had the right people, the development could be done in a cost-effective way. The engineering challenges and the cost challenges were so risky that he had to either abandon the program or he and his team needed to be all in, even to the point of complete and utter failure. When he created the team, he had some criteria. First, he did not want anyone other than those who would volunteer into the program. He did not want critics. He wanted believers of the phased array system. Second, he wanted people who were willing to risk their careers if they failed. People had to rise to the occasion or sink with the ship. There was no middle ground. He needed people with the right attitude.

At Moorestown, he had his sailors and officers come in through the front door of the facility. They had to come in uniform as living examples for why the engineers were building the radar components. Lockheed Martin's slogan even today says, "We never forget who we are working for." He wanted to make sure that everyone is working towards a common goal. If the AEGIS program failed, he knew the Navy was going to come after them. Failure was not an option for him.

Technology that would revolutionize the warfighter and significantly catapult the capabilities of the warfighter ahead of their adversaries requires the kind of commitment to the programs that ADMs Rickover and Meyer had for the program. However, technology does not understand bureaucracies or man-made laws. It only obeys its underpinning physical laws. The realization of the technology needs to have the same commitment and energy from the engineers that make them. Crossing the 6.4 threshold requires overcoming the technical challenges that come with such technology. Furthermore, advancement beyond 6.4 requires a champion, who needs to understand the nuances of the PPBE, JCIDS, and DAS processes and to navigate the technology through politics and bureaucracy. The more significant the technology, the more committed and risk taking the champion needs to be.

Alignment of PPBE, JCIDS, and DAS

A proposed strategy for transitioning game-changing technology is to align the business processes of PPBE, JCIDS, and DAS. Figure 6 is a Venn Diagram of these three business processes. In the intersection is the war doctrine, technical requirements, capacity requirements, S&T, acquisition, retirement prioritization, correctly resourcing the requirements, and steady and reliable execution of the budget.

When researchers try to determine what novel technologies to develop, they operate from their own worldview and try to align that with the potential sponsor. The difficulty is that if the technology is too novel, the potential sponsor has no requirement or use case for it. Even if a sponsor is deeply technical and understands a technology, the employment of the technology becomes a problem. Lost opportunities occur because researchers do not understand what capabilities the warfighter truly needs. Researchers have technically deep skills, but they typically do not have the background to understand how a technology could be used in an operational environment.

Figure 6. Aligning Business Processes

The war doctrine on how we are going to fight the future war needs to be articulated properly, which in turn needs to be translated to technical requirements that the scientists and engineers can understand. The technical requirements would then need to be checked with capacity requirement. How many do we need? This was ADM Rickover's dilemma. Should the entire surface and submarine Navy have gone with nuclear propulsion? If not, how many ships should have nuclear propulsion? The capacity requirements should be mapped back to the war doctrine. Then there should be a requirement prioritization. n the PPBE and JCIDS processes, the requirements owner should prioritize the requirements and appropriately resource them.

The S&T organizations can ingest the requirements and distribute them to their performers for providing high-risk solutions, and the acquisition entities, whose mission is to field the technology, would tamper and manage those risks when transitioning the technology, resulting in more stable and reliable systems. Doing so requires strong discipline across the entities with clear communication and well understood expectations.

Defense Contracting Strategy

Choosing the right contracting strategy has significant impact in the movement of the technology along the maturation levels. Figure 7 shows the various FAR and non-FAR based contracting strategies from the Defense Acquisition University (DAU). Normally, the Broad Agency Announcement (BAA) has been used for basic and applied research. A BAA is an announcement for potential performers, but once published, dialogue is not encouraged. Among the contracting strategies, the Commercial Solutions Opening (CSO) may be a good strategy to use for 6.3–6.4+ work. CSO is a merit-based market-driven source-selection strategy for soliciting commercial solutions that align with the government's requirements. Like other traditional commercial solicitations, it involves competitive methods. Multiple potential performers submit their proposals to address solicited requirements; however, its focus is to attract businesses and institutions that are not the traditional partners with the U.S. government. CSO was authorized by Section 879 of the FY17 National Defense Authorization Act (NDAA).

Section 803 of the FY22 NDAA codifies CSO authority in 10 U.S.C. Section 3458, where "Commercial Solutions Opening (CSO) is a non-Federal Acquisition Regulation (FAR) based solicitation authority for acquiring innovative and commercial solutions."

CSO differs from other contracting approaches such as the Other Transaction Authority (OTA), which is often regarded as the flexible approach in engaging industry partners for research, technology development, and prototype projects. OTA is a legally binding agreement, whereas CSO functions as a solicitation method aimed at acquiring available commercial products. The emphasis for an CSO is to have a solicitation that leads to a contract award as a fixed-price contract or an Other Transaction (OT) agreement. CSO can lead to an OT contract or FAR-based fixed-price contract.

CSO can be a FAR-based contract or opt out to be non-FAR-based, which allows degrees of freedom in the procurement strategies (Defense Acquisition University, 2025). Its purpose is to reduce the barrier of entry for many participants not accustomed to the defense market through simpler contract terms, a streamlined application process, fast-track evaluation timelines for solutions briefs, normally within 30 calendar days of topic closure, and generous negotiable intellectual property rights. The strength of a CSO is that the evaluation is based on merit and on how a solution best solves the problem, rather than on a competitive forum for choosing a "winner."

The CSO approach provides a great way of soliciting higher TRL technology to be matched with technology that has been developed by the government with the cost reliability of fixed-price contracts that would lead to a prototype and production. CSO could be a way to move at commercial speeds, since the process is straightforward. A problem statement is provided, and the potential performer can present a white paper, which is a minimal effort for the vendor compared to the traditional government solicitation methods. The government evaluates the proposal on merit instead of comparing with competitors' proposals. Subsequent to the evaluation, the 2nd stage is an interactive phase, where the vendor provides information on its higher TRL technology with an appropriate cost estimate and the government elaborates on the use case. The 3rd and final stage involve the government generating a statement of work and negotiating prices and terms before reaching a prototype OT agreement or a fixed-price FAR agreement. CSO requires dialogues unlike other contractual processes.

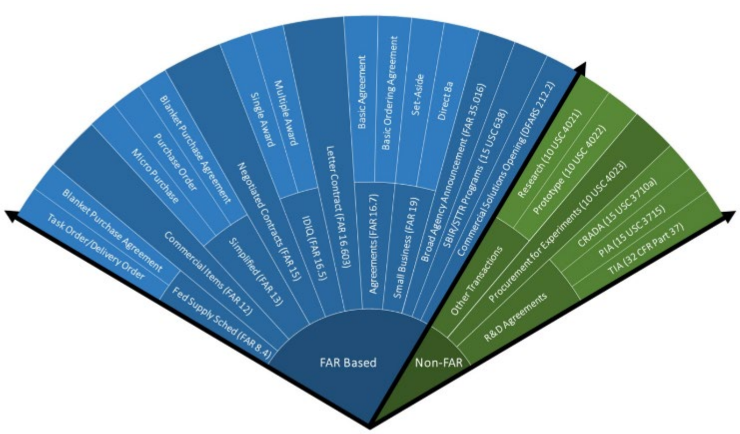


Figure 7. Defense Contracting Cone
(Defense Acquisition University, 2025)

Figure 8. A Potential Contracting Strategy for transition

In Figure 8, the traditional methods of grants and BAAs can be used for a 6.1–6.3 level of effort. They could be accomplished through grants or the service laboratories. These vehicles can be coordinated efforts between academia and government laboratories. If an innovation has reached a certain level of maturity, a service laboratory could shepherd the innovation under an appropriate venue to 6.3. At this point, there would be an industry collaboration to move the innovation forward. If an innovation cannot follow the traditional transition path, the idea is to provide a more relaxed Intellectual Property (IP) strategy. This strategy may involve using an OT or a CSO in the contracting approach, allowing the performers to own the IP and further the technology through commercial or government investments. Doing this may help build the industrialization of the U.S. commercial sector to take on tough DoD challenges.

RCA was able to take on the original AEGIS radar work in the 1970s because the core structure was present to develop the phased array. RCA's long-term production work in the commercial sector was ready to take on the highly risky technical venture of creating AEGIS. The gap between 6.3 and 6.4 is the "Valley of Death," and the available funds for moving the multitude of requisite technologies across it are limited. Developing a licensing strategy or IP strategy to allow private equities to be involved may help many of the technology investments to bridge this gap with efficacious results at commercial speeds.

Goldwater-Nichols Act Revisited?

The current speeds at which technology advances have placed severe stress on defense acquisition systems. By the time a solution is acquired and placed in the hands of the warfighters, the solution has often become obsolete. The machinery that runs the DoD's acquisition is highly stressed to deliver capability on time. Some temporary workarounds and waivers exist for getting technology through the rigorous JCIDS, DAS (DoD 5000), and PPBE processes, but they are insufficient.

These processes are a result of the Goldwater-Nichols Act of 1986 and its derivative laws and regulations. The legislation fixed many joint problems that existed prior to 1986. The Act's solution to joint problems was to create the combatant commands with operational control of the joint forces. By doing so, the service chiefs started to pick up or share roles that the service secretaries used to hold. The service chiefs hold the JCIDS process, and the service secretaries hold the DAS process; they both interact with the PPBE process.

Consequently, the size of a service chief's staff like the CNO's OPNAV has grown to execute the roles and responsibilities that it inherited (Rickover, 1974). Furthermore, with the CNO's staff being transient, there is little continuity over time. New action officers must re-learn already established lessons. What was an exciting new technology program loses its air in the sails when a new CNO shifts the sail. The PEOs under the service secretaries are limited to the type of technologies that they can acquire. When novel technologies must navigate through such uncertainties and instability, only a precious few can transition unless they are championed by the grit and commitment of the likes of ADMs Rickover and Meyer. But even for them, reproducing their results within the current DoD framework might not be possible. It is not clear whether the current framework with all the program management tools would have permitted the provisions of the nuclear reactor budget to pass; the cost analysis might have killed the program.

In light of current challenges and national leaders speaking of acquisition reforms, it may be time to stand up a commission analogous to that of the Packard Commission to look into the impact of the Goldwater-Nichols Act and to take a serious look at the current construct to see if it is optimized to deliver technology. Joint operational failures in the second half of the 20th century resulted in a commission that led to the Goldwater-Nichols Act. Perhaps the many acquisition failures could result in another commission.

A good example of a joint acquisition problem is the F-35 program. Requirements creep and management can be traced back to service-driven priorities (Air Force, Navy, Marines). The service chiefs were pushing for their own capabilities but not necessarily looking at the trade-offs. The GAO's 2021 report on the F-35 flagged poor coordination among service leaders and acquisition officials, with costs ballooning to more than $1.1 trillion. Had this been pre-Goldwater-Nichols, the service secretaries would have had more leverage to temper the service chief's ambitions.

Conclusion

Developing highly complicated technology is a difficult and complex process. There must be a relentless and rigorous pursuit of it. There is an adage that ADM Rickover would say at his public speaking engagements. People, not organizations, make things happen. The responsibility and roles must be placed on the right people.

Francis Duncan writes of the Discipline of Technology in his biography of ADM Rickover:

> Many times he tried to express this thought: “Technology knows no rank”; “Technology will not yield to leadership”; “Technology will not obey an order”; and “You can't argue with technology.”
>
> The aphorisms might have little direct meaning for a manufacturer of many everyday products, or for most people doing paperwork in offices, but to men developing products at the forefront of an advanced technology they cannot be so easily set aside. The success of the naval nuclear reactor means that the organization must adapt to the technology, and not the technology to the organization. . . . The discipline of technology raises moral and ethical questions. Technological development undertaken as a profit-making venture can bring about circumstances involving ethical considerations when

goals slip far beyond their schedules and when cost estimates soar far over budget. The operation of highly complex machinery without proper maintenance and timing can also raise similar questions. The discipline of technology can make sad reading in the balance sheet and in the annual report to the stockholders. But so can newspaper headlines about accidents caused by the poor design of a component or the faulty training of the operator.

> Rickover was convinced that the discipline of technology was essential to the survival of society. He thought it unfortunate that those who benefited most from technology usually accepted its benefits without question, indeed almost as a right. No force penetrated more deeply into a society than technology nor was more active in transforming it. Yet the dangers of technology and its flawed products raised serious questions. A society based on technology but alienated from it was dangerously divided. . . . But more important, the discipline of technology conferred upon an individual the greatest challenge of all—acceptance of responsibility. . . . Unless you can point your finger at the man who is responsible when something goes wrong, then you have never had anyone really responsible. (Duncan, 1990, pp. 293–294)

Laws can change. New regulations can be passed. Technology will be indifferent to laws and managerial systems. To develop complex novel technologies, responsibility must be placed on the right leaders who have the vision and the commitment to carrying through to the end. It is not enough to be a great researcher or technologist. To have transitions of major types, it requires the “discipline of technology,” and some of the takeaway from ADM Rickover are:

1. The 6.1 and 6.2 investments must be made for groundbreaking technology. The reactor was realized due to more than a century of basic and applied research. Though the timelines are long, the opportunities for great technologies can be realized.
2. In order to transition, engineering with rigor must be done once the science is understood.
3. Technology requires an unrelenting champion with a compelling story.
4. Warfighters and technologists need to be engaged with minimal bureaucracy in between.
5. Clearly defined responsibilities and roles are needed.
6. Commitment.
7. Create a work force that is agile, committed, and risk-taking with minimal “distractions.”

When interviewed by Paul Stillwell in 2008, RADM Meyer was in the twilight of his life and passed the following year at age 83. Of all the things that he could have discussed (cross-field amplifiers, waveguides, array construction, etc.), he chose to talk most about the people who built the AEGIS radar, because he felt that people not organizations get things done. The people were there because they wanted to be there, and they were willing to take the program to its finish even at the risk of their own careers (Meyer, 2008). The riskiest technology requires the greatest sacrifices.

References

Blickstein, I., Yurchak, J. M., Martin, B., Sollinger, J. M., & Tremblay, D. (2016). Navy planning, programming, budgeting, and execution. Rand Corporation.

Bond, D. P., Davis, S. M., & Pearsall, A. D. (2016). The Goldwater-Nichols Act of 1986. Monterey: Naval Postgraduate School.

Bond, D. P., Davis, S. M., & Pearsall, A. D. (2016). The Goldwater-Nichols Act of 1986: 30 years of acquisition reform. Naval Postgraduate School: Graduate School of Business & Public Policy.

Defense Acquisition University. (2025, March 29). The contracting cone. https://aaf.dau.edu/aaf/contracting-cone/

Demirjian, E., Eskridge, T. C., Walton, A. L., Sandell, D. L., Shah, N. N., Bailey, R., & Henderson, J. (2024, March 1). Superteams: The development and effect of artificial Intelligence reammates on organizational performance. Naval Engineers Journal, 136(1–1), 313–316 (14).

Duncan, F. (1990). Rickover and the nuclear navy: The discipline of technology. US Naval Institute.

Fenn, A. J., Temme, D. H., Delaney, W. P., & Courtney, W. E. (2000). The development of phased-array radar technology. Lincoln Laboratory Journal, 12(2), 321–340.

Freeman, L., Hobson, S., Moreland, J., Lucero, S., Werner, J., & O'Toole, R. (2024, March 1). Accelerating implementation of critical joint warfighting concepts and capabilities. Naval Engineers Journal, 136(1–2), 41–49 (9).

Locher, J. R. (2002). Victory on the Potomac: The Goldwater-Nichols Act unifies the Pentagon (1st ed.). Texas A&M University Press.

McCain, J., & Thornberry, M. (2015, October 20). Sen. John McCain and Rep. Mac Thornberry discuss the National Defense Authorization Act. Discussion hosted by the Brookings Institution, Center for 21st Century Security and Intelligence. https://www.brookings.edu/wp-content/uploads/2015/10/20151020_mccain_thornberry_ndaa_transcript.pdf

Meyer, R. W. (2008, July 16). 2008 interview with RADM Wayne E Meyer about the AEGIS Project. (P. Stillwell, Interviewer)

Milham, L., Smith, B., Riddle, D., Alexander, K., & Hobson, S. (2024, March 1). The empowerment of the T&E workforce through enterprise learner data. Naval Engineers Journal, 136(1–2), 293–304 (12).

Rickover, H. G. (1974, August 30). The role of engineering in the Navy. National Society of Former Special Agents of the FBI.

Rickover, H. G. (1981, November 5). Doing a Job.

A Modular Open Systems Approach (MOSA) to Enable Technology Transition

Dr. Kelly Alexander—Chief Systems Engineer | System Innovation, Contractor Support to OUSD(R&E) SE&A. [kelly.d.alexander12.ctr@mail.mil]

Ms. Monique Ofori—Systems Engineering Manager | SAIC, Contractor Support to OUSD(R&E) SE&A. [monique.f.ofori.ctr@mail.mil]

Ms. Nadine Geier—Systems Engineering Director, OUSD(R&E) SE&A. [nadine.m.geier.civ@mail.mil]

Abstract

Technology Transition is referred to as the “valley of death” due to commonly experienced lack of successful transition to the next phase of system development. The high risk of technology development can cause a delay or cancellation that can be mitigated by Modular Open Systems Approaches (MOSA) principles. MOSA enables technology transition by providing a framework for integrating, upgrading, and replacing components with minimal disruption. By addressing these MOSA principles early in the development cycle, technology transition is more predictable and manageable:

(1) Modular Architecture: Focuses on modularity offer plug-and-play capability, where system components adhere to defined standards and interfaces. Modularity also supports incremental upgrades, enabling individual modules to be updated or replaced as technology evolves and parallel development for specified (potentially high-risk) components.

(2) Interface Management with Consensus Based Open Standards: (a) Well-Defined Interfaces that rely on widely recognized, consensus-based open standards, ensuring that new technologies from different vendors integrate effectively, reducing development and integration challenges. (b) Open Standards facilitate component reuse, which reduces integration time and can also reduce lifecycle costs by increased competition.

(3) Enabling Environment that promotes Model Based Systems Engineering (MBSE) tools and processes with access to data to enhance interoperability and options in configuration.

Background

The successful transition of emerging technologies into operational systems is a critical challenge in modern defense system development. Systems development often begins with a science and technology (S&T) development effort to mature a technology solution. Manufacturing objectives and sustainment are generally included in the effort, but the primary rationale for completing S&T development phase is an assessment of the technical maturity. Technical maturity is determined in accordance with the DoD technology readiness assessment guidelines (Office of Systems Engineering and Architecture & Office of the Under Secretary of Defense for Research and Engineering [OUSD(R&E)], 2025) which defines the parameters for the transition to the advanced systems development phase that is the entrance to an acquisition program. The transition from an S&T program to an acquisition program or technology transition is often referred to as the “valley of death” because of the commonly experienced lack of successful transition between these phases.

Based on prior studies, there are several ongoing efforts to improve the success rate of technology transition across the DoD. The DAU lists 19 separate programs or activities that are intended to improve or influence the success rate of this transition (Defense Acquisition University [DAU], n.d.) However, none of these programs explicitly addresses the role that a Modular Open Systems Approach (MOSA) can bring to the Defense Innovation Ecosystem. The principles of MOSA can provide a framework to improve the transition success within the context of DoD S&T and Acquisition.

A MOSA provides a structured methodology to enhance adaptability (scalability and upgradability), interoperability, and lifecycle affordability. MOSA enables technology transition by providing a framework for integrating, upgrading, and replacing components. Aligning technology development and technology transition strategies with MOSA principles as a technical framework can improve integration, reduce obsolescence risks, provide opportunities for competition and accelerate innovation adoption. This paper will discuss the current technology transition concerns, followed by a description and discussion of MOSA principles that can enable technology transition, and finally, propose a MOSA Aligned Technology Transition Framework (MA-TTF) that can be implemented within the current technology development strategy documentation.

Current State and Challenges of Technology Transition

The DoD takes technology transition seriously and has worked to study and address the issues and challenges; however, integrating these solutions into larger, more complex defense systems is still a significant contributor to program delays. The current DAU website lists 19 ongoing technology transition management programs that target the technical and funding challenges identified in the various studies. While technology transition has technical, cultural and business challenges, the discussion below is focused on technical and related challenges.

Older studies conducted by the Government Accountability Office (GAO, 2007, 2013) indicate challenges that align with business, technical and culture (workforce) elements. In one of the more recent studies, conducted in 2023, the Commission on Defense Innovation Adoption recommends 10 steps to increase transition of technologies and innovation, including using modular approaches to development efforts to leverage common components and align technology and acquisition portfolios. In this study, the “modular approaches” are in reference to architecting partitioned components or subsystems that allow individualized technology development with the intent to reduce risk by building options for substitutions where technical maturity lags or to allow refresh for new technology (McNamara et al., 2024). This approach is an enabler of MOSA that addresses challenges related to technology transition. Based on several similar reports, a summary of areas of technology transition that cause delays in DoD acquisition programs includes the following:

- **Integration**: Systems can require extensive modification and testing to incorporate new technologies once the S&T effort is complete, resulting in increased costs and transition delays.
- **Proprietary components** that may become unavailable or unsupported during an upgrade or development effort. Proprietary systems also restrict access to alternative suppliers, increasing costs and reducing innovation opportunities.
- **Custom-built, closed systems** require significant investment in development, integration, and sustainment

This list is not exhaustive but focuses on areas that can be addressed by MOSA.

How MOSA Addresses These Challenges

MOSA is defined by a business and technical approach that is based on five key principles: establish an enabling environment, employ modular design, designate open interfaces, use widely available consensus-based standards and certify conformance (shown in Figure 1). Implementing these MOSA principles into the S&T effort counters many of the technical challenges of technology transition by focusing on modular architecture, especially those associated with integration. Moreover, because many decisions that impact the architecture occur during the S&T phase, using a MOSA aligned architecture has a positive impact on innovation and technology upgrades that may occur later in the system lifecycle.

The Office of the Under Secretary of Defense for Research and Engineering (OUSD[R&E]) recently released an *Implementing a Modular Open Systems Approach in Department of Defense Programs* guidebook that provides a robust discussion of MOSA and how it brings value to programs (Office of Systems Engineering and Architecture & OUSD[R&E], 2025). Each of the military departments have also released implementing guidance. A summary of the advantages that MOSA brings is below and depicted in Figure 1.

Facilitates Interoperability through: (a) Well-Defined Interfaces: MOSA relies on widely recognized, consensus-based open standards, ensuring that new technologies from different vendors integrate effectively, reducing development and integration challenges. (b) System Integration: Technologies developed for one system can be transitioned to another with minimal modifications, enhancing flexibility and operational efficiency along with scalability and upgradability. MOSA's modular architecture approach offers plug-and-play capability, when system components adhere to defined standards and interface modularity also supports incremental upgrades, enabling individual modules to be updated or replaced as technology evolves, ensuring a continuous and smooth transition without manageable impact on other parts of the system.

Incorporates Innovation and Tech Refresh Opportunities: By leveraging open standards, MOSA mitigates vendor lock-in, fostering a competitive environment that stimulates innovation among technology providers. MOSA also enables rapid adoption of innovative solutions from diverse sources, promoting technology refresh.

Enhances Competition and Affordability by Re-Usability of Common Components: MOSA can accelerate the development cycle by facilitating component reuse, which reduces the time needed to develop and deploy new capabilities. Modular architectures support parallel development, allowing specified (potentially high-risk) components to be developed, tested, and validated concurrently. This concurrent approach speeds up technology transitions and deployment.

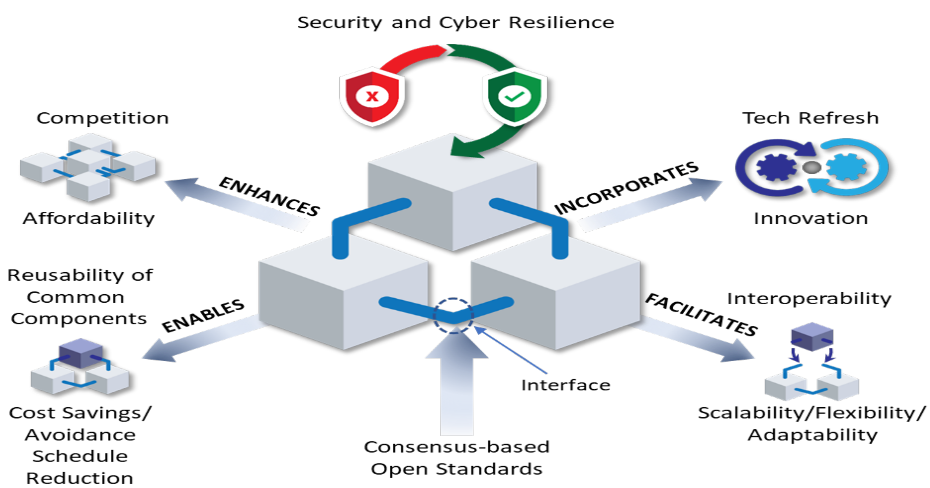


Figure 1. MOSA Pillars and Benefits (Office of Systems Engineering and Architecture & OUSD[R&E], 2025)

While this paper will not include a discussion on software specifically, MOSA also enables DevSecOps Integration and Agile practices that support Continuous Integration/Continuous Deployment (CI/CD). Modular designs in software enable frequent, minimally disruptive updates and seamless integration with legacy systems, contributing to more efficient, secure, and resilient development cycles. Additional information on the ongoing OSD effort for SW Modernization can be found at the OSD R&E website (cto.mil).

Aligning Technology Development Challenges With MOSA Benefits

Connecting the technology transition challenges with MOSA pillars and benefits shows the synergy between the two. MOSA provides the opportunity to balance and address technology transition challenges, as shown in Table 1.

Table 1. MOSA Benefits Aligned to Technology Transition Challenges

Technology Transition Challenge	MOSA Benefit	MOSA Pillar
Integration challenges that slow capability deployment – Rigid architectures and lengthy integration and interoperability efforts delay the fielding of critical capabilities	**Interoperability through modularity** – Standardized interfaces and modular design enable faster integration of new components.	Standardized interfaces Modularity Open standards
Reliance on Proprietary Components– can result in components that may become unavailable or unsupported	**Technology Refresh & Lifecycle Agility** – Open architectures allow for incremental upgrades and replacement of obsolete parts without overhauling entire systems.	Open architecture Modularity
Custom built closed systems Proprietary systems restrict access to alternative suppliers, increasing costs and reducing innovation	**Open Market & Competition –** open interfaces foster a competitive supplier base, driving innovation and cost savings **Rapid Fielding & Iterative Upgrades** – Open, modular systems support incremental enhancements and faster certification through predefined compliance criteria.	Open interfaces Modular Open Standards

The alignment of technology transition challenges with MOSA benefits and pillars lead to the components of a MOSA aligned technology transition framework (MA-TTF). The MA-TTF can be used to bridge the gap between research, prototyping, and an acquisition program. The framework emphasizes architecture-driven development along with interface management and reliance on an enabling environment that supports MBSE. While MOSA offers a counter to many of the technology transition challenges, this paper proposes a focus on three key MOSA principles: modular architecture, interface management and an enabling environment that incorporates MBSE.

MOSA Aligned Technology Transition and Framework (MA-TTF)

Accompanying MOSA principles with MBSE as the means to manage the baseline and consider options for transition that enable interoperability and integration results in a framework that can bridge the technology transition “valley of death.” The establishment of MOSA in the technology development strategy can be implemented as discussed below and shown in Figure 2.

The MA-TTF is based on an architectural approach that relies on MOSA technical principles, modular architecture, and identifying and managing key interfaces that relies on open standards. The MOSA Implementation Guide (MIG) emphasizes these key steps to incorporate a MOSA into the program strategy. The following into the technology development strategy:

1. **Architecture Development & Management planning** that uses MBSE tools to manage model-based systems engineering (MBSE) to consider optional solutions.
2. **Interface Management** that supports integration & interoperability goals to include reuse strategies such as product line architecture.
3. **Plan for Consensus-based Open Standards** that facilitate future upgrades and technology refresh cycles.
4. **Architecture Development & Management planning** that uses MBSE tools to manage Use model-based systems engineering (MBSE) to consider optional solutions.
5. **Interface Management** that supports integration & interoperability goals to include reuse strategies such as product line architecture.
6. **Plan for Consensus-based Open Standards t**hat facilitate future upgrades and technology refresh cycles.

If the technology is software focused, establish a DevSecOps Pipeline for Technology Insertion by implementing continuous integration, verification, and cybersecurity measures that utilizes the MOSA driven architecture. Current SW Modernization ongoing in the DoD is implementing these principles on some programs using the SWA pathway.

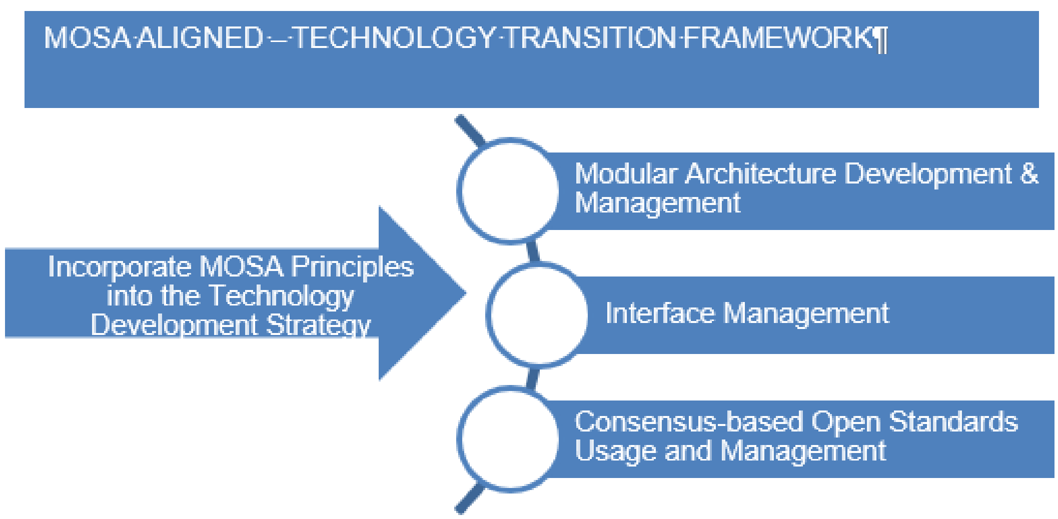


Figure 2. MA-TTF

Conclusion

As these MOSA principles are applied to technology development, the technology transition into a system or an integrated capability is more predictable and provides for strategies that allow for future competition and technology refresh. Incorporating an architecture-driven strategy that is rooted in MOSA principles early in the lifecycle ensures that validation pipelines (reuse) and sustainment planning are better accommodated. Most importantly, integration and interoperability are enabled, thus reducing risk early in the development cycle. Overall incorporation of a MA-TTF can result in more predictable and manageable technology transition.

Additional information on MOSA is available at the Systems Engineering and

Architecture website (https://www.cto.mil/sea/pg/), including the MOSA implementing guidance that provides the DoD community of stakeholders including Military Services, Civilians, and DoD contractors with information to support a MOSA as part of the defense program acquisition lifecycle. Each of the Military Departments have also released MOSA guidance that provides service specific approaches to MOSA implementation. Additionally, the DoD Standardization office has established a database of open standards that facilitates their use and implementation into programs (https://www.dsp.dla.mil/Publications/DSP-Journal/News-Display/Article/4117175/new-mosa-enabling-standard-in-assist/).

This paper is a current look at MOSA and how it can enable system development during the S&T phase and create long term benefits. However, future work should include a deeper look at programs and assess the impact based on examples using this approach or a similar architecture driven development cycle.

References

Defense Acquisition University. (n.d.). *Technology transition and transfer programs*. www.dau.edu

GAO. (2007). *GAO review of technology transition practices* (GAO-07-061).

GAO. (2013). *Defense technology development: Technology transition programs support military users, but opportunities exist to improve measurement of outcomes* (GAO-13-286).

Gunderson, C. (2014). *A platform across the valley of death: Tech transition via open enterprise information system development*. Naval Postgraduate School.

Kotila, B., Drezner, J. A., Bartels, E. M., Hill, D., Hodgson, Q. E., Huilgol, S. S., Manuel, S., Simpson, M., & Wong, J. P. (2023). *Strengthening the defense innovation ecosystem*. RAND. https://www.rand.org/pubs/research_reports/RRA1352-1.html

McNamara, W. M., Modigliani, P., MacGregor, M., & Lofgren, E. (2024). In M. T. Esper & D. L. James (Chairs), *Commission on Defense Innovation Adoption: Final report*. Atlantic Council. https://www.atlanticcouncil.org/in-depth-research-reports/report/atlantic-council-commission-on-defense-innovation-adoption/

Office of Systems Engineering and Architecture & Office of the Under Secretary of Defense for Research and Engineering. (2025, February). *Implementing a modular open systems approach in Department of Defense programs*.

Office of the Under Secretary of Defense for Research and Engineering & Office of Systems Engineering and Architecture. (2025, February). *Technology readiness assessment guidebook*.

Using Data Analytics and Dashboards in a Research Organization Environment for Project Management

David A. Lechner, PhD—is a senior scientist at the Naval Research Laboratory with over 40 years of experience in program management and defense program research. He spent 15 years managing EW and Sonar programs at NAVSEA, then over two decades in private industry developing technology and managing programs. He received his PhD in 2019 from the Catholic University of America in Mechanical Engineering and Acoustics, is a former Washington chapter President of the Acoustical Society of America, and has taught Physics as an Adjunct Professor at Anne Arundel Community college. [David.A.Lechner.Civ@US.Navy.mil]

Austin Tamez—is a Senior Project Manager at Envisioneering Inc with over 10 years of management experience. He is a former U.S. Navy Lieutenant and served on the ballistic missile submarine SSBN 735 USS Pennsylvania and taught as a submarine tactics instructor for several years. He received a BS in Operations Research from the Unites States Naval Academy and was certified as a Nuclear Engineering Officer by Naval Reactors. He currently works with Naval Research Laboratory Code 5720 helping to manage advanced R&D programs, and helped implement and improve the Power BI PM Tool designed at NRL. [Austin.R.Tamez.ctr@US.Navy.mil]

Abstract

The use of data analytics tools offers significant new opportunities for acquisition programs in the Department of Defense (DoD). In particular, the use of modern tools like Power BI and Tableau provide data platforms to organize, visualize, and track program data. This paper will review how data dashboards built on the Navy's Flank-speed platform were developed for experimental use by one particular section of a research organization, Code 5720 within the U.S. Naval Research Laboratory (NRL). The results are largely transferrable across the DoD, as the tools use Microsoft's Power BI and the dashboard modules are instantly replicable. We review the motivation for building the platform, review its capabilities in tracking technical goal alignment with "north star" objectives, how it tracks expenses and estimated costs at completion, and how the visual schedule data is easily updated and understood. We will review how long it took our team to build the dashboard, what it takes to instantly copy it onto other platforms, and what we do to maintain the data using simple MS Excel files. We will show how a potential organizational management dashboard might look, and review how the use of dashboards is improving our branch's operations and a potential data analytics framework that crosses the full acquisition life cycle.

Note: This work reflects an experimental use of data tools in one specific branch of NRL and does not reflect the views or usage by the entire organization.

Research Issue/Problem Statement

Defense program management requires management of financial, schedule, technical, and quality status and progress. Historically programs were planned by writing lengthy documents. Those documents were then approved by senior management as stand-alone volumes, using the knowledge and experience (and limitations and bias) of the approver. Typical attempts to improve program oversight have included adding more reporting, typically via added documents and reporting of schedule and financial data in greater detail. Without tracking (good and bad) project data, there is no way to compare the planning data and results metrics as the plans and technical specifications are contained in static documents. It should be possible to prevent project failure if we are able to look backwards and determine common elements in failed projects. NRL 5720 has been developing a set of tools that convert plans, presentations, and tables of financial data into an integrated planning and program management tool. Although designed for R&D projects, it is generally extendable to the development, engineering, and production phases of a program. This paper presents the motivation, design, and development history of the planning tool.

Motivation for Building a Data Analytics Tool for Project Management

NRL 5720 is a branch at NRL with about 30 projects, and we wanted to improve management of our technical planning, schedules, financial status, and staffing. We built the early version of this prototype in 2023 using PowerPoint slides to design and storyboard the tool, and used the format from decade-old quarterly program reviews as our starting point. We started using the prototype in 2024 for one project, then recently expanded it to all of our projects in 2025. Starting with effectively zero institutional knowledge about how to construct a Power BI dashboard, a total of about 6 staff months of effort have been invested in the tool so far, reflecting how easy these tools are to use. Over time the expectation is that the data will accumulate, allowing new uses in comparative and predictive analytics.

There were several questions we wanted the tool to answer, primarily:

- What are the strategic technical objectives of our research?
- Are the projects on track to meet technical objectives?
- Are the projects on track to complete on promised schedules?
- Is our cashflow OK? Will we run out? Are we fully spent for expiring funds?
- Do we have enough work for our staff?
- Do we have enough staff? Do we need to hire? When? How many?
- How do the sections of our branch compare in metrics? Which needs help?
- Who are our current sponsors? Who are the future sponsors?

Project Management Tool Capabilities

The project management (PM) tool we developed shows project managers and organizational leadership data as a series of tabs on a web page, but is designed to show the health of our projects and our organization in a dashboard style, just like the dashboard of a car shows its health while driving. Just like a car, the data is color coded when possible, with green lights for good data and yellow and red for bad. This does focus the viewer on problem areas, and allows faster review of good projects and data.

The dashboards are designed to replace the typical PowerPoint slides that dominated staff meetings in the past, and allow instant viewing of data without having to enter data, copy graphics, and paste them on a slide. The data is always available, and the visual data can be refreshed as often as 5 times a day. This does not eliminate the need to have data updated, but by pulling the data from Excel files we provided a simple and commonly-known tool for users to edit the data. We estimate that a project leader can update the data for their project in about 10 minutes would use a monthly refresh cycle. A second paper included in this panel discusses the ability to pull data from the Navy's Enterprise Resource Program (ERP) into Power BI, which allows a team to develop their own financial data tools with even less need for manual data entry.

The tool also has several other dashboards that were developed to mimic tables that we used to show for new projects. One showed which other projects preceded or were related to the new project. The second summarizes the credentials of the people that will be working on the projects.

The following figures are provided, along with a brief explanation, to illustrate the capabilities of this tool. Note that the data shown was created for illustration only, and does not represent any actual research projects at NRL (at least not yet!).

Branch Overview Dashboards

The organizational data dashboards are intended to show the current and primary technical objectives of the organization and how well it is doing in the metrics that we set for branch. This includes metrics as how well are we meeting the technical objectives, are we delivering our projects in a timely manner, how are our finances, and describe the key attributes of the branch such as staffing levels and financial health.

The "Branch Overview" dashboard (Figure 1) provides a report on multiple aspects of a research branch's "health" for senior leadership. It shows status on metrics important in an R&D organization and financial data. The data is "sliceable" by sponsor type, section code, and even by individual project.

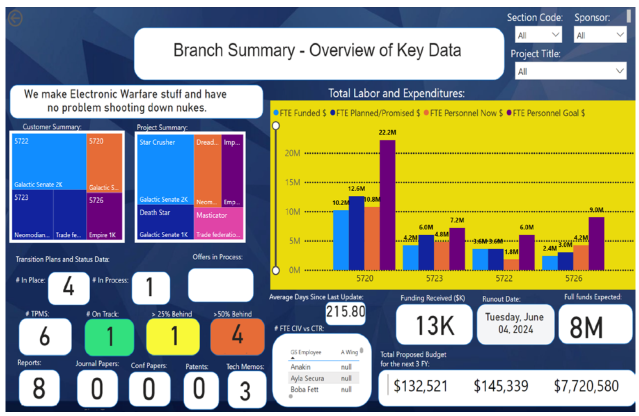


Figure 1: Branch Overview Dashboard, provides an overview of key metrics on one page

Note: It summarizes our customers with a box chart, our three sections with a bar chart, and key organizational metrics including financial health

The data entry for the dashboard is done with simple Excel files, as shown in Table 1. We will not show these tables for the other dashboards, as they duplicate the visual information in the dashboard. Users simply open the Excel file and updating the data, and within a scheduled time the dashboards will show the new data (we set the updates as 5x per day, at time periods 8,10,12,2, and 4).

Table 1 and 2: Examples of Excel Data Used for Branch Overview Data

Work Unit	TPM	Percentage of TPM Metric Met	Met or Exceeded	Greater than 25% behind	Greater than 50% behind	Has a TPM(1 for yes 0 for n
57X002	Decrease mold buildup in the	110	1	0	0	1
57X003	Pcc Improves by 0.25	50	0	0	1	1
57X004	Decrease time needed for ta	0	0	0	1	1
57X001	Laser color will be greener by	75	0	1	0	1
57X003	Floatiness of the struts will i	10	0	0	1	1
57X005	Shield strength increased by	33	0	0	1	1

Branch Code	FTE Goal	FTE Now	FTE Planned	FTE Funded	FTE Goal $	FTE Now $	FTE Planned $	FTE Funded $	Full FTE Funds Expe
5723	12	8	10	7	$ 7,200,000	$ 4,800,000	$ 6,000,000	$ 4,200,000	$ 10,200,000
5726	15	7	5	4	$ 9,000,000	$ 4,200,000	$ 3,000,000	$ 2,400,000	$ 5,400,000
5722	10	3	6	6	$ 6,000,000	$ 1,800,000	$ 3,600,000	$ 3,600,000	$ 7,200,000
5720	37	18	21	17	$ 22,200,000	$ 10,800,000	$ 12,600,000	$ 10,200,000	$ 22,800,000

Note: The data for the dashboards are maintained by users using simple Excel spreadsheets that are easily updated by users. Once the files are updated and closed, the changes are automatic.

Technical Performance Metrics Dashboard

The Technical Performance Metrics (TPM) dashboard (Figure 2) is designed to reflect the key technical goals of the organization, which can then be used to flow-down to and be referenced by projects at a lower level. This may be a dashboard developed by a strategic planning process. It would be useful during management reviews to understand how a group is doing in achieving and underachieving its goals. Some goals may be secretive, thus may use code-names. Some metrics may be secretive, and thus may use relative percentage values. The goals are sometimes referred to as the “North Stars” of the team, and allow everyone to see whether the progress is good or not and work towards those goals.

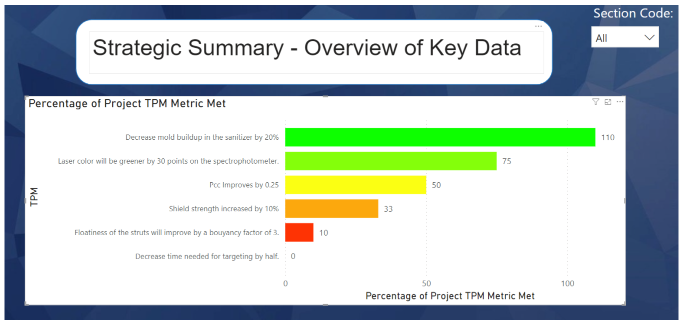


Figure 2: Technical Performance Dashboard Example

Note: This dashboard allowed tracking of our “North Star” technical objectives.

Staff Planning

The staff planning dashboard (Figure 3) was developed to allow us to manage our people, as their time is truly our most valuable asset. They allow a planner, project lead, or individual to see the staff loading over a planning time (1 year for our branch). The data is shown in four bar-graphs, which show whether a project is fully staffed or not and whether people are fully tasked or not. This is a key tool for task and staff planning.

In our branch we try to be sure that each project has more than one person assigned to it, and each person has a primary and secondary project that they support. This dashboard allows proactive workload planning. By graphically displaying an individual’s workload next to a project’s necessary workload, a planner can better coordinate how to plan researcher’s time when they work on more than one project.

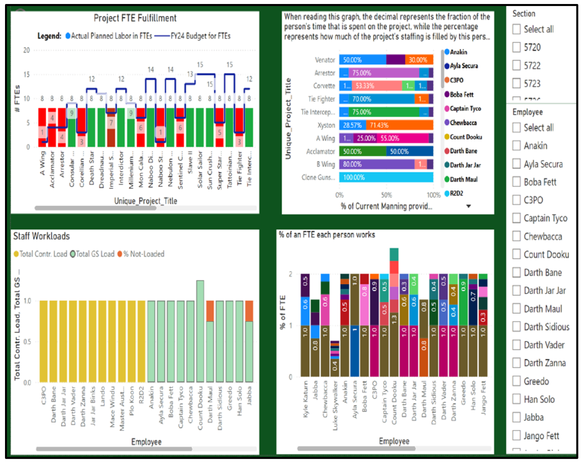


Figure 3: Staffing Plan Dashboards

Note: These dashboards helped show that we had enough work for all of our people, and enough people for all of our projects.

Each of the four staff planning views can also be viewed individually, simply by clicking a button in the corner, as shown in Figure 4. By showing this data, we see which staff members need work and which are over-tasked. We also can see which projects are under-staffed. In the future a similar graphic will be built showing how actual staff labor charging rates in the last 30 or 60 days compare to planned allocations, thus identifying problems where costs will over-run or work will under-run.

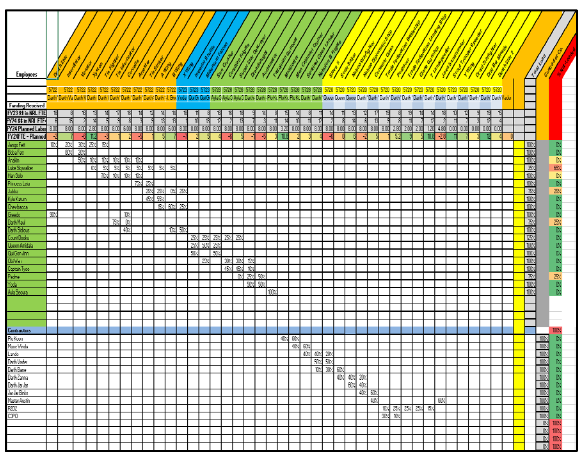


Figure 4: Staffing Summary Dashboard

Note: This one provides an overview of how all the staff is allocated between every project.

Project-Specific Dashboards

The project-level dashboards show data specific to a single project, but in a consistent format. They are organized to provide a consistent view of data across projects, explain them to management and visitors, and track the health in project management terms (technical, cost, schedule, and quality metrics).

In the first and summary level view of a project we recreated the common “Quad Chart,” showing a graphic illustration of the project, technical summary, schedules, and financial data (Figure 5). It can be used to review and compare new project proposals or projects that were approved and funded. On a monthly basis the project lead is asked to update the technical issues and schedule elements. The financial data should be retrievable from an enterprise financial tool with little or no effort. In the future, a semi-automated link to enterprise tools should enable daily data updates, but getting good data is often harder than getting the latest data. The Power BI software tool will automatically pull data elements from individual fields in the Excel file, and uses a single project number as a "key" that identifies the data for that project.

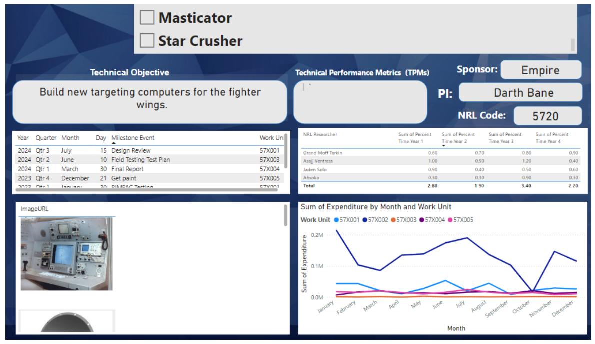


Figure 5: Project Summary Dashboard based on the Common Quad Chart

Note: This one provides a project specific view for each project.

Project Schedule Dashboard

This dashboard provides a classic Gantt chart to visualize the project schedule (Figure 6). This is based on David Bacci's Deneb script, which is available online. It allows a project lead to plan start and stop dates for each task, task dependencies, and the percentage completion for each reporting interval. This view shows the team's progress on a project and the inter-dependencies between tasks (Bacci, 2023; Payton, 2024). The data is entered using an Excel spreadsheet, shown in Table 2.

Figure 6: Project Schedule Dashboard

Note: This provides the classic Gantt view of a schedule and allows relatively easy visualization of progress on each task.

Table 2: Data Entry File for a Project Schedule

id	task	Progress	phase	milestone	Assigned T	start	end	completion	dependenci	Work Unit
1	1. Find a bounty hunter	1	Software Code		Jar Jar Binks	10/24/2023	11/15/2023	100		57X001
2	2. Tell the bounty hunter I want someone	1	Software Code		Jar Jar Binks	10/24/2023	11/15/2023	100		57X001
3	3. Show a picture of who I want.	1	Software Code		Jar Jar Binks	10/24/2023	11/15/2023	100		57X001
4	4. Show the name of who I want	1	Software Code	TRUE	Plo Koon	10/24/2023	11/23/2023	100		57X001
5	5. tell the bounty hunter not to fail	1	Test		Yoda	11/23/2023	11/25/2023	100		57X001
6	6. track the bounty hunter	1	Software Code		Plo Koon	10/24/2023	12/20/2023	80		57X001
7	7. follow the bounty hunter	0.1	Software Code		Yoda	10/24/2023	12/20/2023	60		57X001
8	8. ensure men are trained on how to fight bounty hunters	0.1	Software Code	TRUE	Yoda	11/25/2023	1/6/2024	40	1,2,3,4	57X001
9	9. give the crew lots of guns and turbo lasers	0.3	Test		Anakin	1/6/2024	1/8/2024	40	5,6,7	57X001
10	10. lose track of the bounty hunter	0	Software Code		Yoda	11/25/2023	1/13/2024	0	1,2,3,4	57X001
11	11. reveal you only pretended to lose track of the bounty hunte	0	Test		Anakin	1/13/2024	1/15/2024	0	8	57X001
12	12. show up unannounced at critical moment in fight	0	Software Code		Anakin	1/6/2024	2/10/2024	0	5,6,7	57X001
13	13. attempt to take prize from bounty hunter	0	Documentation		Anakin	2/10/2024	2/12/2024	0	9	57X001
14	14. lose prize to bounty hunter	0	Documentation	TRUE	Anakin	2/12/2024	3/11/2024	0	9	57X001
15	15. pay the bounty hunter	0	Documentation		Anakin	2/12/2024	3/11/2024	0	9	57X001
16	16. accept the hostage from the bounty hunter.	0	Software Code		Anakin	3/11/2024	3/13/2024	0	11	57X001
17	17. kill the bounty hunter and take your money back.	0	Software Code		Anakin	2/10/2024	3/30/2024	0	9	57X001
18	18. find out the hostage has escaped the garbage sector	0	Software Code	TRUE	Anakin	2/10/2024	3/30/2024	0	9	57X001
19	19. look for the hostage again	0	Software Code		Anakin	3/30/2024	4/1/2024		14	57X001
20	20. hear alarm bells in the deathstar	0	Test	TRUE	Anakin	10/24/2023	4/3/2024		Basic Date	57X001
21	21. evacuate the death star	0	Software Code		Anakin	4/3/2024	4/5/2024		15	57X001
22	22. discover the hostage is your mom	0	Software Code		Anakin	3/11/2024	5/6/2024		10,11	57X001
23	23. cry	0	Test	TRUE	Anakin	5/6/2024	5/8/2024		16	57X001
24	24. ask your grandpa if he wants to rule the galaxy together	0	Software Code		Anakin	3/11/2024	5/13/2024		10,11	57X001
25	25. find out she would rather spend time with people who beat	0	Software Code		Anakin	5/13/2024	5/15/2024		17	57X001
26	26. blow up grandpa's home planet	0	Test	TRUE	Anakin	12/24/2023	12/29/2023		Basic Date	57X001
27	27. sell tickets to a movie titled "how I blew up my grandpa's h	0	Software Code		Anakin	12/24/2023	12/29/2023		Basic Date	57X001
28	28. become rich fom movie and build more death stars	0	Test	TRUE	Anakin	12/29/2023	4/27/2024		18	57X001
29	29. blow up all the suns in the galaxy	0	Software Code		Anakin	5/15/2024	6/5/2024		19	57X001
30	30. start a flashlight company	0	Test	TRUE	Anakin	6/5/2024	6/19/2024		20	57X001
31	31. get even richer on flashlight company	0	Software Code		Anakin	6/19/2024	7/10/2024		MST10	57X001
32	32. miss your grandpa	0	Test	TRUE	Anakin	7/10/2024	8/9/2024		21	57X001
33	33. hire bounty hunter to see if your grandpa was on his home	0	Management		Anakin	8/9/2024	9/8/2024		MST11	57X001
34	34. find out he wasn't	0	Management		Anakin	9/8/2024	9/15/2024		MST12	57X001
35	35. bounty hunter returns and reveals they are your grandpa	0	Software Code		Anakin	9/15/2024	10/6/2024		MST13	57X001
36	36. make movie about your life	0	Integration		Anakin	10/6/2024	10/12/2024		19	57X001
37	37. Profit	0	Test	TRUE		10/12/2024	10/14/2024		20	57X001

Note: The schedule data is maintained in an Excel file.

An additional dashboard was created that tracks task schedules in a table and project summary data in summary quad-chart format (Figure 7). The dates in this table are the same as those used in the Gantt graphical view, but provides more detailed schedule data along with some basic project data.

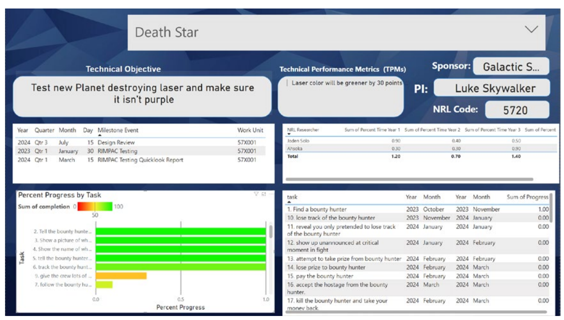


Figure 7: Quad-Chart Summary with Detailed Schedule Table

Note: This provides project summary data along with detailed schedule data, useful for detailed schedule planning meetings.

Contract Management Dashboard

Another dashboard (Figure 8) was created to help manage work by a contracted partner, and provides project leads an idea of what projects are using the contract, how long current funding will last, and when added funding should be put on a contract to ensure work continuity.

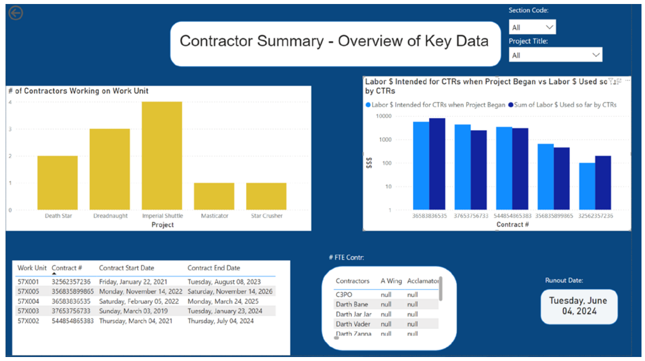


Figure 8: Contract Management Dashboard

Note: Provides an overview of contract work and funding status.

Proposal Review Dashboards

The PM Tool also includes several other data dashboards built to recreate several standard slides used in the annual NRL proposal process, including the project summary, team credentials, and related projects dashboards. Examples of these are shown in Figure 9, Figure 10, and Figure 11.

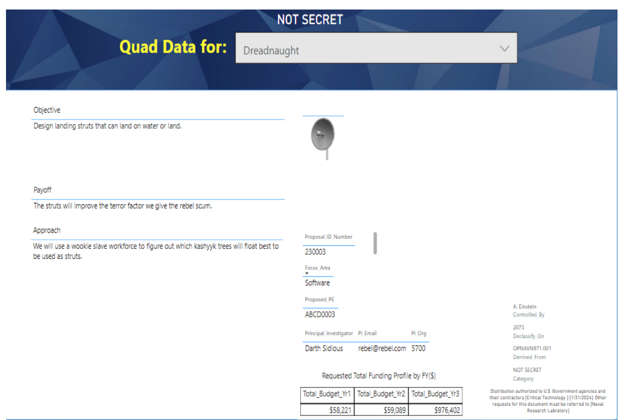


Figure 9: Proposal Summary Dashboard

Note: This provides project summary data on a proposed project.

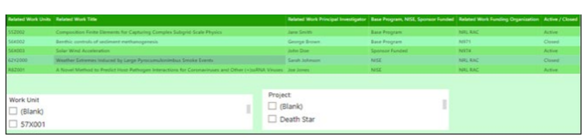


Figure 10: Related Work Units Summary

Note: Shows what other projects are related to the proposed effort.

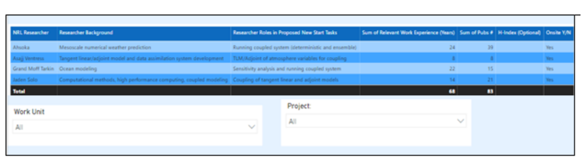


Figure 11: Team Credentials Dashboard

Note: Provides information on the proposing team to consider their credentials and experience.

Project Financial Dashboard

The last group of data dashboards we are experimenting with are financial data dashboards (Figure 12). These are meant to show how a project is spending money versus allocated funds, whether it will meet a budget or not, and compare several different metrics on financial health. It also provides overall branch financial data, and calculates cashflow and runout dates.

Figure 12: Financial Summary Dashboard

Note: Provides additional details on project and branch financial data to assess the work backlog and funding status.

A Potential Organizational Management Dashboard

Our project focused on what our branch needed, which was a tool to manage multiple research projects. As a part of the initial project we also prototyped a data dashboard that might be used to manage an organization of many branches at the leadership level. This would not necessarily involve just aggregating the lower level data, although some data fields may use that approach to compile summary statistics. At the organization level other factors are also important, such as performance metrics for functional support teams. The result is shown in Figure 13, which is also another example of the prototyping method using a power-point slide.

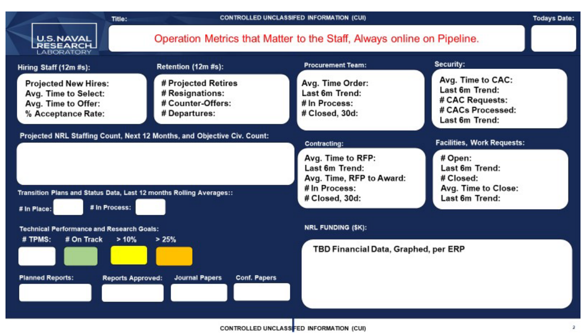


Figure 13: Hypothetical Organizational Dashboard

Note: shows potential organizational and functional performance metrics in one glance.

The command level dashboard could be used to provide the full organization insight into the metrics in each part of the organization, including all of the functional support departments. In some areas, pulling data from programs such as the docflow system would enable automated data input. It could be used to motivate the organization to improve their metrics via public awareness, and provide project teams critical data to enable planning efforts (e.g., procurement lead times or hiring timelines). Ultimately the choice of metrics and data content would be customized for each organization, and negotiated with the leadership team to show both current performance metrics and aspirational performance behaviors.

Data Security

One of the concerns we had in creating these dashboards was the overall data security, as the platform was hosted on Flankspeed and our division plans and project data would be stored there. Power BI and Flankspeed have security controls that limit who can see the dashboards and who can edit the data files. We also had to ensure that the site was properly marked, as it had Controlled Unclassified Information on it, and did not contain Personal Identifiable Information on individuals.

Building the Dashboard

To build the dashboards we started first by examining the data that is currently briefed to track project status, data obtained for new proposals, prior management data requests, then conducted online research for best practices in tracking project status and progress. We used senior staff members with both industry and academia experience to review recommended project management strategies and data visualization methods. Then the dashboards were prototyped using simple PowerPoint tools. These were then reviewed in a briefing mode with other managers. This design activity took about 2 months. After this the dashboards were prototyped in Power BI as that license was already included in our Microsoft Office toolset. The prototyping process took several months, and the only tricky part was the display of images and integration of certain graphical widgets. Overall it only took us about 6 months to build several working dashboards and iterate the design once. Power BI is a 5th-generation style environment and easily learned by any programmer or technical staff member. Conventional wisdom is that these projects are not technically difficult, and once completed the problem is getting users to accept and use a new business practice ("20% Technical, 80% Social" is how one senior staff member described these efforts). The harder part of designing a tool like this is leveraging useful senior managers to direct the content, and not generate too much detail so that the tool remains simple to use and focuses on important data.

Replication on Other Platforms

One of the strengths of building a tool like this is that it is easily replicated and modified for other organizations. The files generated at NRL can be copied into a small zip file and shared on any other licensed Power BI server (which covers most of the Navy and DoD). NRL is able to share the basic PM-Tool files on request and is seeking internal approvals to post them on a shared server. In some cases, an organization may benefit by hiring a contractor that has experience in building Power BI dashboards, in particular if a whole new set of dashboards is being developed. Besides building the tool however, the social problems start once it is introduced. To overcome these NRL hosted an overview briefing on the project and why the tool was being used, then had short individual training sessions with the primary users, one on one. We then built a how-to manual and posted it as a web page on our SharePoint site, right next to the PM Tool page. An online video is also planned. Even with all of this, the introduction of such a tool across a large organization is a significant effort. For a single branch of 32 engineers the planning was modest. Generally, such projects start with a single team, like ours, as a beta site. Then the introduction starts to scale up to a division of hundreds, a full site, and then the full organization. At least 6 months for each phase is recommended, possibly more.

Maintaining the Data Using Excel

The problem that some teams have reported in using data dashboards is that the creation effort is too easy, and multiple complex dashboards are created that then must be maintained and fed with data. This can be a problem. In the long-term maintaining data sets is an expensive proposition. In an organization with 2,000 staff members, if each update takes even 10 minutes per month that is 333 staff hours, and at a nominal $200 per hour becomes almost $800,000 for the staff time. This author has seen tools that take an hour per month to update, which would clearly be a huge expense in a large organization.

This tool was designed to take about 10 minutes per month, but save more than that in avoiding time spent developing typical program review slides. Some modest engineering project will spend several days of effort every quarter for program reviews, typically involving senior technical staff to prepare and dry-run the material. Maintenance effort must also be simple, so that staff does not become frustrated.

Any requirement to retype data from one table into another must be avoided at all costs, as that is not only prone to error but also extremely costly in staff time. Power BI includes tools to massage data from one format into another, making the ingest of data from an enterprise tool relatively simple if the right reports are available. When used, that allows relatively simple data update cycles and staff can focus on interpreting the data which is the higher value use of their time.

Improvements in Organizational Efficiency

This is a work in process, but preliminary results to date are promising. PM Team Tool users are reporting that the tool is providing a simple method to organize tasks and keep track of progress. Section managers reported typical difficulty in starting to use the tool, even in the case where the tool designer had to use it, but increased ease once the how-to pages were posted. One noted that for a small project it was important to create project schedules with about a dozen tasks at a time, and not get too detailed.

Initially usage of this tool spread slowly, with about one project per month being added in the first few months, but this also allowed the development engineer time to debug the tool and refine the data base. Use for project reviews has just started. A follow-up article is planned to discuss additional metrics in usage, acceptance, changes, and quality rates, but the primary metric for success is that projects are more clearly attaining technical objectives and delivery results on time and within budgets.

The use of dashboards and metrics are also ideal complements to six-sigma type efforts to improve organizational performance. The steps in a six-sigma or process improvement cycle are to a) establish performance metrics for the organizational team, b) monitor the metrics, c) set improvement goals jointly with the team leads, d) implement strategies to improve the performance levels, e) measure the new performance, and then f) repeat the process in cycles and over time.

Examples of methods for improving performance are:

a) Identifying bottlenecks and either enable parallelism, add staff, or automate activity with better information support to approvers and improve the throughput rate
b) Tracking organizational error rates, and enable low-error teams to bypass approvals
c) Monitoring approval actions with no value (i.e., very low correction or rejection rates) and eliminate those approval points
d) Invest capital for new tools or equipment
e) Restructuring a team
f) Changing the business processes, or
g) Delegating authority for activities

Use supplier quality certification strategies, where error rates are monitored after the fact, and authority increased to lower-level managers with low error rates, and taken away if errors go up.

Sharing Status Data with Customers

One aspect of this tool is that it could be used to replace monthly project status reporting. Instead of receiving a monthly report or briefing, the tool allows any (and only) registered site members to log-in and view the data. It is also relatively easy to do briefings using the data dashboards. Customers reported positive experiences after this approach was used in briefing status. It is also relatively easy to create special dashboards that use the same data, but provide outward-facing views of the project status for customers to see at any time. Another approach might be to package the data files on a regular basis and email them to the customer. In general, this approach should reduce project management and reporting costs for federal contracts. To enable this, a new DD-1423 Data Item Description would be needed, along with an easy to download set of Power BI and Excel files. This is a logical extension of this type of tool.

Other Lessons Learned

Be careful what data is being collected—data is not free, and collection has a cost. Know and understand the cost of collection. Maintaining data over the long term is expensive. Focus on key strategic metrics and maintain those well. Do not go metric-crazy and create a complex system. Simpler is better in this domain.

Behavior adapts based on data collected, and people are creative. Sometimes this resembles whack-a-mole. As an example, software testing metrics are collected to show error rates during final tests, so teams create an additional step for testing before the final test to detect errors earlier. The final test rate falls, but this is now taking longer as there are two steps.

Monitor metrics for maturity and utility, and discard those that are no longer needed or show constant levels of performance. Some metrics reflect fundamental business rules or the result of capital systems productivity, and do not change.

Understand quality performance methods—start out with frequent measurement of new projects or teams, and if performance remains consistently good back-off the measurement rates to a random audit or eliminate it.

Use color wisely—green is good and above a goal or within acceptable tolerance. Red is bad, outside of a defined tolerance limit or below acceptable limits. Define metrics so that “up” is good and “down” is bad. Be consistent in how data is displayed. As an example, use test success rates (99% is “good”) instead of error rates (going “up” from 1% to 8% is actually “bad”).

Understand statistical sample methods, confidence levels, and long-term averaging of data. In some cases, a small sample focus may be a problem, such as measuring costs on a daily basis and a large amount of procurement orders are placed in one day. Similarly, if the measurement for cost is monthly on a large project, a small error rate could seem hidden or portend a large change in final estimated cost.

R&D Data Analytics Within the Overall Acquisition Process

The PM Tool in use at NRL 5720 is just a portion of the overall data set needed to define, research, develop, produce, test, field, and support a program or capability within the DoD. The location of the research project PM Tool within the overall defense acquisition system is shown in Figure 14, marked with the red outline, but elements of it could be used in the requirements and budgeting phase, material acquisition phase, and even by sustainment organizations. Each aspect of this system is important, and could benefit in its own way from similar application of data dashboards, metrics, six-sigma improvement, and comparative and predictive data analytics.

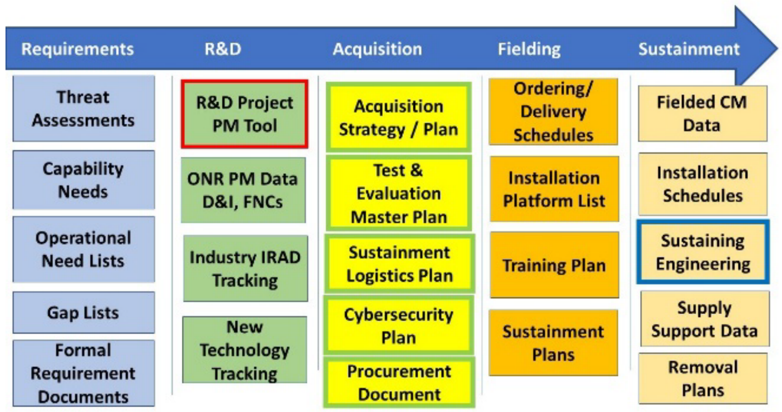


Figure 14: An Integrated Framework for Data Base Acquisition Analytics that Spans the DoD Acquisition Life Cycle

Note: The box outlined in Red indicates the PM tool for research projects discussed in this paper. A data schema for several of the acquisition functions (in yellow) was developed for Sharepoint Lists in prior work by this author (Lechner). Sustaining Engineering financial data (blue border) is the subject of another paper on this panel (Dunn).

This figure is also material-centric, and possibly the application of data dashboards for other parts of the doctrine, organization, training, materiel, leadership and education, personnel, and facilities (DOTMLPF) structure have already started in other DoD organizations. This will be an ongoing process for the DoD, but their use could be very pervasive within the next several years with some modest leadership encouragement, low-level advocacy, and organizational sharing.

Although the PM Tool in use at NRL 5720 is of general use in managing many types of technical projects there or in other warfare centers and development commands, it is easily modifiable to suit individual needs of other types of organizations. On a grander scale, it would be extremely useful to the overall DoD if the data analytics used across multiple programs and organizations used a common data schema and hierarchical structure.

A common data schema could allow the DoD to compare performance metrics between organizations, then leverage improvement ideas from the best organization. It would allow after-action analysis of projects that succeed and those that had problems, and eventual use of artificial intelligence tools to recognize leading indicators using predictive analytics and prevent problems as they occur or before it becomes too late.

An example of a common data structure used for the material acquisition phase of a program and was developed by this author for a project using SharePoint and Excel in 2018. That schema was designed to replace the Acquisition Plan document with an online data set, and similar schema were created for the Test and Evaluation Master Plan, Acquisition Strategy, Logistics Sustainment Plan, Cybersecurity Plan, and Installation planning documents (Lechner, 2018).

The 2018 demonstration aimed to create data entry forms like TurboTax for creating programs, and views like Amazon shopping for comparative analytics and program approvals. Initial testing indicated that it took a program team less than a day to fill in the data and allowed the AP data for a sample data set of a program to be reviewed and approved within 2 weeks. Those types of metrics, if extended and successfully scaled, could allow the programs to be created in months not years, and the elimination of hundreds of support jobs at each acquisition command. The data schema was built as a flat-file however, and the prototype used SharePoint forms, and thus would need to be reorganized into a star format and rehosted to allow use with a modern business intelligence tool.

Table 3: A Portion of the 2018 Acquisition Plan Schema

Title	Data Type	Title	Data Type
AP Appendix #	Text	8b Relevant Copyrights	Text
Project Title	Text	9.1a Funding Years (List 5 Years)	Currency
Document Date	Date	9.2a RDT&E Funding, Year 1	Currency
Acquisiton Category Level (ACAT)	Number	9.2b RDT&E Funding, Year 2	Currency
Technical Project Manager (PM)	Name	9.2c RDT&E Funding, Year 3	Currency
Approval Date, Technical Manager	Date	9.2d RDT&E Funding, Year 4	Currency
Local Warfare Center Project Office Manager	Name	9.2e RDT&E Funding, Year 5	Currency
Approval Date, Local Contracting Officer	Date	9.3a Procurement Funding, Year 1	Currency
Approval by HCA, PEO, or DRPM	Name	9.3b Procurement Funding, Year 2	Currency
Approval Date, HCA, PEO, or DRPM	Date	9.3c Procurement Funding, Year 3	Currency
Approved, Chief of Contracting Office	Name	9.3d Procurement Funding, Year 4	Currency
Approval Date, Chief of Contracting Office	Date	9.3e Procurement Funding, Year 5	Currency
Approval, Contracting Officer	Name	9.4a Construction Funding, Year 3	Currency
Approval Date, Contracting Officer	Date	9.4b Construction Funding, Year 4	Currency
Approval, Program Manager	Name	9.4c Construction Funding, Year 5	Currency
Approval Date, Program Manager	Date	9.5a Maintenance Funding, Year 3	Currency
1a. Prototype Status?	Yes/No	9.5b Maintenance Funding, Year 4	Currency
1c. MDAP Status	Yes/No	9.5c Maintenance Funding, Year 5	Currency
2. Technology Focus Area	Pull Down	9.6 Total Funding, All Types & Years	Currency
3. DOD Product of Services Descriptive Code	Number	9.7a RTI Phase 1 Est. Cost	Currency
4. Program Office (Code)	Text	9.7b RTI Phase 2 Assment Cost	Currency
5. Lead Contracting Office (Code)	Text	9.7c T&E Services Cost	Currency
6.1 Statement of Need	Text	9.7d Production Costs, All Years	Currency
6.2a Historical Summary, Background	Text	9.7d Procurement Profile, by Year	Currency
6.2b Historical Summary, Requirement	Text	9.7e RTI Eng. Services Costs, All years	Currency
6.3a Previous Contract	Text	10a CPARS Contact	Name
6.3b Previous Contractor	Text	10b Email of DPARS Contact	Email
6.3c Previous Contract Type	Text	10c CPARS Contact Phone #	Phone
6.3d Previous Contract Code	Text	11a MDA TDS Approval Date	Date
6.3e Prev. Contract, Quantity	Number	11b MDA Acquisition Strategy	Date
6.3f Prev Contract Award Date	Date	11d MDA Peer Review Date	Date
6.f Prev Contract, PoP End Date	Date	11e Purchase Request Receipt Date	Date
6.3g Prev Contract, Total Value	Currency	11h White Paper & Draft RFP Release Date	Date
6.3h Prev Contract, Competitive?	Yes/No	11j Full RFP Date	Date
6.4a RTI or FAST-Lane Rationale	Text	11l Proposal Due Date	Date
6.4b Govt. Resources Needed	Text	11m Contract Award Date	Date
6.4c Collaborating Integrator	Text	11o RTI Phase 1 Completion Date	Date
6.4d Integrator Funding, RTI Ph.1	Currency	11p MS-B Requirements Complete	Date
6.4e Integrator Funds, RTI Ph.2	Currency	11g Pre-EMD Peer Review Complete	Date
6.4f Integration Funding PMO Contact	Name	11c Actual AS approval	Date
7.1 RTI Project Description	Text	11f Actual PR Receipt Date	Date
7.2 Likely Quantity	Currency	11i Actual WP Release Date	Date
7.2a Estimated Unit Cost	Currency	11k Actual Proposal Receipt Date	Date
7.3a Primary EMP Name	Text	11n Actual Contract Award Date	Date
7.3b P-EMP Objective	Number	12a Prospective Sources	Text
7.4a Secondary EMP Name	Text	12b Other Considered contracts	Text
7.4b S-EMP Objective Value	Number	12c Small business Set-Aside Only?	Yes/No
7.5 EMP Tradeoffs	Text	13 Risks	Text
8a Relevant Patents	Text	14 GFP	Text
		15 GFI	Text

Note: This is a part of the data set the author tested in 2018 to replace an AP document with a data set for a program in EMD phase. The test showed that the data set could be created in less than a day.

Summary and Conclusions

NRL Code 5723 has developed a data dashboard on Power BI that can be easily adapted by other DoD users for project management. The data is maintained in simple MS Excel files and easily updated. Creating project instances takes a few hours, one time. Maintaining the data is less, projected at less than 10 minutes per month, or perhaps a quarter hour. The research project data views are also useful for procurement offices and easily adapted to show contract planning, project execution, schedules, and technical success (or not!). The approach is scalable to organizational metrics and management (e.g., procurement office, hiring, and financial execution).

We found that, similar to most data analytics projects, the technical development was the simple part of the problem, whereas the user adoption, training, and follow-up was the challenging part. Another finding was that the Power BI tools are very scalable and easily replicated to other organizations and programs. Finally, we found that the Power BI tools provide a simple method to track ongoing program status and organize project data across an organization, allowing comparison of program metrics and easier and real-time visualization of program data. This positive experience reinforces projections made by many advocates that using data analytics will provide significantly better tools for project management in the DoD, and that their use allows an evolving Acquisition Data Analytics approach that could effectively replace the bulk of the paper and briefings that make up the current DoD Instruction 5000.02 Defense Acquisition System.

References

Bacci, D. (2023). *A collection of advanced dataviz examples using Vega, Vega-Lite, Deneb and Power BI.* Github Site. https://github.com/PBI-David?tab=repositories

Dunn, K. (n.d.). *Streamlining financial reporting at NUWC Keyport: A centralized Sharepoint and Power Suite System for improved data access and analysis, Naval Undersea Warfare Center*.

Lechner D. (2018). *Using database and workflow methods for statutory and regulatory acquisition documents*. The MITRE Corporation.

Payton, C. (2024). *How to use David Bacci's Deneb template: Power BI Gantt w/ milestones, dependencies, and phases.* Youtube https://www.youtube.com/watch?v=tOhAxulPAoc

Acquisition Research Program
Department of Defense Management
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net