

SYM-AM-23-110



PROCEEDINGS  
OF THE  
TWENTIETH ANNUAL  
ACQUISITION RESEARCH SYMPOSIUM

---

THURSDAY, MAY 11, 2023 SESSIONS  
VOLUME II

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

**May 10–11, 2023**

**Published: May 1, 2023**

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.



ACQUISITION RESEARCH PROGRAM  
DEPARTMENT OF DEFENSE MANAGEMENT  
NAVAL POSTGRADUATE SCHOOL

The research presented in this report was supported by the Acquisition Research Program of the 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](mailto:arp@nps.edu)

[www.acquisitionresearch.net](http://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/>



ACQUISITION RESEARCH PROGRAM  
DEPARTMENT OF DEFENSE MANAGEMENT  
NAVAL POSTGRADUATE SCHOOL

# Table of Contents

---

<b>Welcome: David H. Lewis, VADM, U.S. Navy (Ret), Acquisition Chair, Acquisition Research Program.....</b>	<b>1</b>
<b>Keynote Speaker: Honorable Nickolas H. Guertin, Director, Operational Testing and Evaluation, Office of the Secretary of Defense .....</b>	<b>2</b>
<b>Panel 14. The Future of Navy and Army Acquisition .....</b>	<b>3</b>
<b>Panel 15. Perspectives on Portfolio Management.....</b>	<b>5</b>
Model-Based Approach in Defense Portfolio Management: Data Preparation, Analysis, and Visualization of Decision Spaces.....	7
Portfolio Management Structures: System, Capability, and Mission Portfolios.....	24
Joint All-Domain Command and Control (JADC2) Opportunities on the Horizon .....	42
<b>Panel 16. Technology-enabled Logistics &amp; Sustainment .....</b>	<b>58</b>
Acquiring Maintainable AI-Enabled Systems .....	59
Commercial and Defense Vendor Management: A Comparison of Competitive Procurement Below the Prime—Subcontract Competition—How Real is It?.....	68
Optimizing Operations and Logistics Support Using Opus Evo.....	81
<b>Panel 17. Software Acquisition Pathway .....</b>	<b>89</b>
Software Acquisition and the Color of Money .....	90
Crossing the Great Software Development Divide Within the DoN .....	<b>Error! Bookmark not defined.</b>
GBSD: A U.S. Department of Defense Enterprise Agility Pathfinder.....	<b>Error! Bookmark not defined.</b>
<b>Panel 18. Acquisition Workforce Perspective from DACMs/DATMs .....</b>	<b>110</b>
<b>Panel 19. The Acquisition Frontier.....</b>	<b>113</b>
Defense Acquisitions: DOD Should Take Additional Actions to Improve How It Approaches Intellectual Property.....	115
Social Engineering Impacts on Government Acquisition.....	132
Comparative Analysis of Pathways to Changeability .....	148
<b>Panel 20. Finding and Leveraging Sources of Asymmetric Advantage in Defense Acquisitions..</b>	<b>157</b>
Asymmetries and their Potential for Enduring Advantage .....	159
<b>Panel 21. Calculating Return on Investment.....</b>	<b>177</b>
Is it Ready? Quantifying the Maturity of Emerging Technologies.....	178



Management and Business Knowledge Representation for Decision Making: Applying Artificial Intelligence, Machine Learning, Data Science, and Advanced Quantitative Decision Analytics for Making Better-Informed Decisions.....	191
You Can't Wait for ROI to Justify Model-Based Design and Analysis for Cyber Physical Systems' Embedded Computing Resources.....	213
<b>Panel 22. Acquisition through Modeling &amp; Simulation .....</b>	<b>231</b>
The Design and Development of a Defense Acquisition Workforce Virtual Environments for Asynchronous Collaboration (VEAC).....	232
A Reference Architecture for a Policy Test Laboratory .....	236
Towards an Enterprise All-Domain M&S Environment for T&E: Overcoming M&S Challenges Within the DoD .....	251
<b>Panel 23. Next Generation Primes - Moving from Innovation to Fielding .....</b>	<b>259</b>
The Innovation Paradox—Merging Process with Disruptive Thinking to Accelerate Capability Transition to the War Fighter Through the Educational Innovation Capstone Process.....	260
Assessing the Effectiveness of Defense-Sponsored Innovation Programs as a Means of Accelerating the Adoption of Innovation Force Wide .....	270
Leverage AI to Learn, Optimize, and Wargame (LAILOW) for Strategic Laydown and Dispersal (SLD) of the Operating Forces of the U.S. Navy .....	287
<b>Panel 24. Digital Engineering in Test and Evaluation .....</b>	<b>296</b>
Proven Warfighting Capabilities Delivered at the Speed of Need .....	297
Shifting Left: Opportunities to Reduce Defense Acquisition Cycle Time by Fully Integrating Test and Evaluation in Model Based Systems Engineering .....	307
Using Digital Twins to Tame the Testing of AI/ML Systems .....	329
<b>Papers Only .....</b>	<b>347</b>
Guiding the Hands of Time: Toward Reliable Schedule Estimates .....	348
Training an Agile Acquisition Workforce to Combat Emerging Threats .....	362
Through the Looking Glass: Why EVM Is an Essential Risk Mitigation Measure for Decision Makers and Program Managers .....	<b>Error! Bookmark not defined.</b>



SYM-AM-23-110



PROCEEDINGS  
OF THE  
TWENTIETH ANNUAL  
ACQUISITION RESEARCH SYMPOSIUM

---

THURSDAY, MAY 11, 2023 SESSIONS  
VOLUME II

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

**May 10-11, 2023**

**Published: May 1, 2023**

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.



ACQUISITION RESEARCH PROGRAM  
DEPARTMENT OF DEFENSE MANAGEMENT  
NAVAL POSTGRADUATE SCHOOL

THIS PAGE INTENTIONALLY LEFT BLANK



ACQUISITION RESEARCH PROGRAM  
DEPARTMENT OF DEFENSE MANAGEMENT  
NAVAL POSTGRADUATE SCHOOL

## **WELCOME: DAVID H. LEWIS, VADM, U.S. NAVY (RET), ACQUISITION CHAIR, ACQUISITION RESEARCH PROGRAM**

---

**Vice Admiral David H. Lewis, USN (Ret.)** took the helm as the Naval Postgraduate School Chair of Acquisition in 2021 and led the Acquisition Research Program (ARP) in the Graduate School of Defense Management to connect NPS with leaders and policymakers in the acquisition community. Lewis graduated from NPS in 1988 with a Master of Science in Computer Science and returned to campus to replace the founding Chair of Acquisition, Rear Admiral, USN (Ret.) Jim Greene, who retired.

Most recently, Lewis served as Director of the Defense Contract Management Agency, managing over \$7 trillion in defense contracts. In this role, he oversaw the agency's efforts to ensure that supplies and services contracted for by the Department of Defense are delivered on time and in line with contract performance requirements.

During his career at sea, Lewis served as a communications officer, fire control and missile battery officer, and combat systems officer aboard destroyers and guided-missile cruisers.

Upon selection to flag rank in 2009, Lewis served as Vice Commander, Naval Sea Systems Command and then served four years as Program Executive Officer, Ships, where he directed the delivery of 18 ships and procurement of another 51 ships. From 2014-2017 he served as Commander, Space and Naval Warfare Systems Command where he led a global workforce of 10,300 civilian and military personnel who design, develop and deploy advanced communications and information capabilities.

Lewis's extensive experience in shipbuilding has given him a unique understanding of the full acquisition lifecycle. He has delivered ships as a program manager and program executive officer, then later sustained and modernized them as a fleet engineer and systems commander.



## **KEYNOTE SPEAKER: HONORABLE NICKOLAS H. GUERTIN, DIRECTOR, OPERATIONAL TESTING AND EVALUATION, OFFICE OF THE SECRETARY OF DEFENSE**

---

**The Honorable Nickolas H. Guertin** was sworn in as Director, Operational Test and Evaluation on December 20, 2021. A Presidential appointee confirmed by the United States Senate, he serves as the senior advisor to the Secretary of Defense on operational and live fire test and evaluation of 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. Most recently, he has 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 been in leadership of organizational transformation, improving competition, application of modular open system approaches, as well as prototyping and experimentation. He has also researched and published extensively on software-reliant system design, testing and acquisition. He received a BS 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 also a registered Professional Engineer (Mechanical).

Mr. Guertin is involved with his community as an Assistant Scoutmaster and Merit Badge Counselor for two local Scouts BSA troops as well as being an avid amateur musician. He is a native of Connecticut and now resides in Virginia with his wife and twin children.





## PANEL 14. THE FUTURE OF NAVY AND ARMY ACQUISITION

Thursday, May 11, 2023

9:05 a.m. –  
10:15 a.m.

**Chair: Michael Williamson, LTG USA (ret.)** Senior Vice President, Global Business Development & Strategy, Lockheed Martin

**Panelists:**

**Vice Admiral Francis Morley, USN**, Principal Military Deputy to the Assistant Secretary of the Navy (Research, Development and Acquisition)

**Lieutenant General David G. Bassett, USA**, Director, Defense Contract Management Agency

**Michael Williamson, LTG USA (ret.)**—is the senior vice president for Global Business Development & Strategy at Lockheed Martin Corporation. In this role, Williamson is focused on bringing integrated solutions to customers who rely on Lockheed Martin's capabilities and technologies to support their missions and address their most pressing needs. His responsibilities also include establishing comprehensive strategies across the enterprise that will enable future growth.

Previously, Williamson served as vice president and general manager for Lockheed Martin Missiles and Fire Control (MFC), where he was responsible for operational excellence, a diverse portfolio of products and business enabling initiatives.

He also previously served as vice president of Tactical and Strike Missiles for MFC. In this capacity, he managed significant programs in the areas of Hypersonic Weapon Systems, Close Combat Systems, Strike Systems, Precision Fires and Advanced Programs.

Williamson joined Lockheed Martin in 2017 following a distinguished career as a lieutenant general with the U.S. Army. He served as the principal military deputy to the assistant secretary of the Army for Acquisition, Logistics and Technology and director of Acquisition Career Management. He also served as a congressional fellow on Capitol Hill.

Williamson holds a bachelor's degree in business administration from Husson University, a master's in systems management from the Naval Postgraduate School, and a Ph.D. in business administration from Madison University. He is also a graduate of the Advanced Management Program at the Harvard Business School.

**Vice Admiral Francis Morley, USN**—is a native of Phoenix, Arizona. He earned a Bachelor of Science in Physics and a commission as an ensign from the Naval Reserve Officer Training Corps at San Diego State University. He is a graduate of the U.S. Naval Test Pilot School and holds a Master of Science in Aviation Systems from the University of Tennessee. He is a graduate of the Air Command and Staff College, Joint Forces Staff College, Defense Systems Management College, George Washington University National Security Studies Program and Harvard's Kennedy School of Government National and International Security Program.

In August 2021, he assumed responsibilities as Principal Military Deputy for the Assistant Secretary of the Navy Research, Development & Acquisition.

Morley has been recognized as Commander, Naval Air Force Atlantic Ship Handler of the Year and the Department of the Navy Program Manager of the Year. He has more than 3,500 flight hours and 750 carrier arrested landings. He has flown more than 35 different types of aircraft, including the F/A-18A-F, EA-18G, AV-8B, F-14, F-15, F-16 and MiG-29.



Lieutenant General David G. Bassett, USA—Army Lt. Gen. David G. Bassett is the director of the Defense Contract Management Agency, headquartered at Fort Lee, Virginia. As the director, he leads a Department of Defense agency consisting of more than 12,000 civilians and military personnel who manage more than 300,000 contracts, performed at 15,000 locations worldwide, with a total value in excess of \$7 trillion.

Bassett assumed leadership of DCMA on June 4, 2020. He came to the agency after serving as Program Executive Officer for Command, Control and Communications-Tactical (PEO C3T) since January 2018, where he was responsible for the development, acquisition, fielding and support of the Army's tactical network, a critical modernization priority.

Bassett is a graduate of the Army Command and General Staff College at Fort Leavenworth, Kansas, and a distinguished graduate of the Industrial College of the Armed Forces in Washington, D.C.



## PANEL 15. PERSPECTIVES ON PORTFOLIO MANAGEMENT

Thursday, May 11, 2023

10:30 a.m. –  
11:45 a.m.

**Chair: Brigadier General Frank J. Lozano, USA**, Program Executive Office, Missiles and Space

***Model-based Approach in Defense Portfolio Management: Data Preparation, Analysis, and Visualization of Decision Spaces***

Waterloo Tsutsui, Purdue University

Cesare Guariniello, Purdue University

Kshitij Mall, Purdue University

Frank Patterson, Georgia Tech Research Institute

Santiago Balestrini-Robinson, Georgia Tech Research Institute

Jitesh Panchal, Purdue University

Daniel DeLaurentis, Purdue University

***Portfolio Management Structures: System, Capability, and Mission Portfolios***

John Driessnack, University of Maryland

Caitlin Kenney, University of Maryland

***Joint All-Domain Command and Control (JADC2) Opportunities on the Horizon***

Roshanak Rose Nilchiani, Stevens Institute of Technology

Dinesh Verma, Stevens Institute of Technology

Philip S. Antón, Stevens Institute of Technology

**Brigadier General Frank J. Lozano, USA**—is the Program Executive Officer (PEO), Missiles and Space, Redstone Arsenal, AL. He is responsible for the development, production, fielding, sustainment, and international program aspects for assigned missile and space systems. BG Lozano assumed his current position August 2022.

BG Lozano assessed into the Army Acquisition Corps in 2001 and graduated with an MBA from the University of Texas at Arlington. He served with Lockheed Martin Missiles and Fire Control in Grand Prairie, TX as part of the Training With Industry (TWI) program.

After completion of Command and General Staff College, BG Lozano was assigned as the Assistant Product Manager for Project Manager Soldier Weapons, PEO Soldier, followed by an assignment as an Ammunition and Demolition System Acquisition Manager for the Special Operations Command (SOCOM) and the Army Research Development and Engineering Command (RDECOM).

In 2008, BG Lozano was assigned as a Department of the Army System Coordinator (DASC) for Tactical Missile Systems and Ballistic Missile Defense Systems. BG Lozano was selected to be a Special Assistant for the Army's Vice Chief of Staff, GEN. As the Special Assistant, he provided insight, advice, and counsel on Army acquisition programs crossing many different functional capability areas.

BG Lozano commanded the Product Management Office for Soldier Protective Equipment, PEO Soldier from 2011 until 2014. Afterwards, he was assigned to the Joint Staff, J-8 Capabilities and Acquisition Division. Upon graduation from the US Army War College, BG Lozano was assigned as the Project Manager for the Lower Tier Project Office, PEO Missiles and Space from 2017 until 2020, followed by an



assignment as the Integrated Fires and Rapid Capability Office PM. From April 2021 to May 2022 BG Lozano served as the ASA(ALT) Chief of Staff.

BG Lozano's operational and combat experience include deployments to Bosnia, Kuwait and Iraq. His awards and decorations include the Parachutist Badge, Ranger Tab, Legion of Merit, Bronze Star Medal, Joint Service Commendation Medal, the NATO Service Medal, the Army Staff Identification Badge, and the Joint Staff Identification Badge. He is certified in Program Management; Contracting; System Research; Planning and Engineering; and System Test career fields.

BG Lozano is married to the former Anne E. Yesconis of Dallas, TX and has three children: Olivia, Jackson, and Nicholas.



# Model-Based Approach in Defense Portfolio Management: Data Preparation, Analysis, and Visualization of Decision Spaces

**Waterloo Tsutsui**—is a Senior Research Associate in the School of Aeronautics and Astronautics at Purdue University, IN. 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 10 years, with the last position involving the research and development of lithium-ion battery cells for electric vehicles. Tsutsui’s research interests are systems engineering, mission engineering, energy storage systems, multifunctional structures and materials design, and the scholarship of teaching and learning. [wtsutsui@purdue.edu]

**Cesare Guariniello**—is a Research Scientist in the School of Aeronautics and Astronautics at Purdue University. He received his PhD in Aeronautics and Astronautics from Purdue University in 2016 and his master’s degrees in Astronautical Engineering and Computer and Automation Engineering from the University of Rome “La Sapienza.” Guariniello works as part of the Center for Integrated Systems in Aerospace led by DeLaurentis, and is currently engaged in projects funded by NASA, the DoD Systems Engineering Research Center (SERC), and the NSF. His main research interests include modeling and analysis of complex systems and SoS architectures—with particular focus on space mission architectures— aerospace technologies, and robotics. Guariniello is a senior member of IEEE and AIAA. [cguarini@purdue.edu]

**Kshitij Mall**—is a Post-doctoral Research Associate at the Center for Integrated Systems in Aerospace, Purdue University. He obtained his PhD and master’s degrees from the School of Aeronautics & Astronautics, Purdue University. He was a Post-doctoral Research Fellow in the department of Aerospace Engineering at Auburn University in 2019. Previously, he completed B. Tech. in Mechanical Engineering at JSSATE Noida, India and then worked for a year at Infosys Technologies Ltd. as a Computer Systems Engineer Trainee. His research interests lie in the areas of Systems Engineering, Atmospheric Flight Mechanics, Explainable Artificial Intelligence, and Human-Class Mars missions.

**Frank Patterson**—is a Senior Research Engineer at the Georgia Tech Research Institute (GTRI) in the Systems Engineering Research Division (SERD). His current research includes the application of state of the art of computational methods and tools to the design and analysis of complex systems. He is also experienced in the development, integration, and use of multi-disciplinary simulation, modeling, and analyses for the design of systems under uncertainty. He has more than 15 years of experience supporting the DoD and warfighter as an engineer across various domains. Patterson earned his bachelor’s degree, master’s degree, and PhD in Aerospace Engineering at Georgia Tech.

**Santiago Balestrini-Robinson**—is the head of the Methods and Analysis Developments Branch (MADB) in the Electronic Systems Laboratory (ELSYS) of the Georgia Tech Research Institute (GTRI). His primary area of research is the development of collaborative quantitative and qualitative decision support tools and frameworks. Balestrini-Robinson has led teams supporting multibillion-dollar military acquisition programs, requirements analysis studies for novel operational and materiel concepts, as well as the development of general frameworks to support collaborative and executable Model-based Systems Engineering. He earned a BS, an MS, and PhD in Aerospace Engineering from the Georgia Institute of Technology in 2003, 2006, and 2009, respectively.

**Jitesh Panchal**—is a Professor of Mechanical Engineering at Purdue University. He received his BTech from Indian Institute of Technology (IIT) Guwahati, and MS and PhD from Georgia Institute of Technology. He is a member of the Systems Engineering Research Center (SERC) Council. He is a recipient of 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. [panchal@purdue.edu]

**Daniel DeLaurentis**—is Vice President for Discovery Park District Institutes and Professor of Aeronautics and Astronautics at Purdue University. He leads the Center for Integrated Systems in Aerospace (CISA)



activities on research problem formulation, modeling, and systems engineering methods for aerospace systems and system-of-systems. DeLaurentis also serves as Chief Scientist of the U.S. DoD's SERC UARC to understand the systems engineering research needs of the defense community (primarily) and translate that to research programs that are then mapped to the nation's best researchers in the SERC's university network. He is a Fellow of the INCOSE and the AIAA. [ddelaure@purdue.edu]

## Abstract

The research team adapted a previously developed system-of-systems analytic workbench to address Integrated Acquisition Portfolio reviews via mission engineering analysis. The team illustrated the findings to date in developing decision-support tools tailored to the needs of these reviews and the insights they produce for improved acquisition outcomes. The essence of the prototype acquisition decision-support tools we are developing is a combination of portfolio optimization and mission engineering. We explore the interactions between candidate systems to acquire and existing systems to identify capability gaps and features of portfolios that optimally cover a family of mission threads. Moreover, we investigate the role of digital engineering in facilitating this process to shift the stakeholders' mindset from the traditional forms of acquisition decision-making to a predominantly model-based approach, from data preparation, analysis, and visualization of the decision spaces. Preliminary findings indicate that these approaches indeed do provide the stakeholders with a broader range of more accessible information, such as resource tradeoffs and cost sensitivity analysis. Longer-term goals include a more comprehensive model-based acquisition decision-support system, with consistent data definitions extracted from "authoritative sources of truth," thereby connecting all models with common data definitions.

## Introduction

The research team developed a pilot/prototype capability to enhance data-driven decision-making regarding acquisition and sustainment programs, motivated by the context of the Department of Defense's (DoD) Integrated Acquisition Portfolio Review (IAPR) process. As the DoD transforms its acquisition paradigm from centralized oversight of Acquisition Category (ACAT) 1D programs to decentralized oversight delegated across Components, the Office of the Under Secretary of Defense for Acquisition & Sustainment (OUSD(A&S)) must likewise shift its focus from traditional program oversight to enabling acquisition innovation and managing a portfolio of capabilities. OUSD(A&S) has made significant strides in acquisition innovation through the rollout of the Adaptive Acquisition Framework and Capability Portfolio Management. However, it has not fully realized the analytic capability necessary to underpin acquisition investment decisions with clear traceability to warfighter requirements.

The research team focuses on a portfolio-centric approach, which we implemented by enhancing and adapting an existing research product called the System-of-Systems Analytic Workbench (AWB). The AWB consists of several SoS tools, the primary of which are Robust Portfolio Optimization (RPO), Systems Operational Dependency Analysis (SODA), and Systems Developmental Dependency Analysis (SDDA). A significant part of the research effort described herein included enhancements to the AWB elements. More specifically, we upgraded the scripts and functions representing the various AWB elements to a set of qualified Python packages. These upgrades enabled ease of continued development of the packages and their capabilities. The upgrades also made the components of AWB more friendly for both developers and users while also easing any future burden of transitioning the tools to the sponsor and their designees.

The research team also explored the enablement of portfolio management from a mission engineering perspective. The two guiding principles for this research are 1) the demonstration of the viability of the Mission Engineering (ME) approach to support Joint acquisition decision-making and 2) the initiative for the development of a reusable Digital Engineering environment and methodology to support future Mission Engineering pilots, studies, and acquisition analyses. Furthermore, the research explored the transition from a paper-based



(PowerPoint modality) review of various portfolios (e.g., EW Portfolio, NC3 Portfolio, or ASuW Portfolio) to a more model-based review of the portfolios, addressing questions such as: What form should this take? What information is key from a leadership perspective? How do we ensure a holistic review without being overwhelmed with complexity and information? The research included engagement with selected mission portfolio managers to understand their priorities and challenges to enable evidence/data-based portfolio management.

This paper starts with an explanation of the AWB development, wherein we describe the overall description of the AWB tools, followed by a discussion on the development of the AWB tools. The paper also demonstrates the enhanced AWB using a notional anti-surface warfare (ASuW) application to illustrate its application to a non-trivial domain.

## **Analytic Workbench Development**

The AWB is a collection of methods and techniques developed by researchers at Purdue University within SERC projects starting in 2011. Due to the complex and multifaceted nature of SoS modeling and analysis, the most effective approach is to develop different methodologies, each addressing one specific aspect of SoS, for example, emergence due to interactions or portfolio-wide considerations. The AWB implements this approach by providing a set of tools developed on purpose for modeling and analysis of SoS.

The AWB addresses complexities associated with interconnections that exist across physical, functional, and developmental SoS hierarchies. The idea is to support the “top-down integration, bottom-up implementation” paradigm at the SoS level. The analytical tools in the workbench account for the complex and highly interconnected nature of the systems that constitute the overall SoS. The analytical tools allow the user to:

- Quantify performance and risk for individual systems, links, and of overall SoS;
- Assess the impact that changes to SoS architecture (add/remove links and/or nodes) will have; and
- Quantitatively identify optimal sets of architectural solutions given constraints on cost, performance, and risk.

When building tools to support decision-making in an SoS environment, the challenge is that such tools must address the technical and programmatic complexities of SoS, yet remain domain-agnostic. It is up to researchers to find the appropriate balance between the need for tools that can be used on a broad spectrum of applications in various fields and the need for tools that can be easily tailored to specific applications and user requirements.

This project focused on three tools from the AWB: Robust Portfolio Optimization (RPO), Systems Operational Dependency Analysis (SODA), and Systems Developmental Dependency Analysis (SDDA). Figure 1 shows the inputs and outputs of tools in the AWB and in the Decision Support Framework (DSF), the framework that used RPO and SODA sequentially.



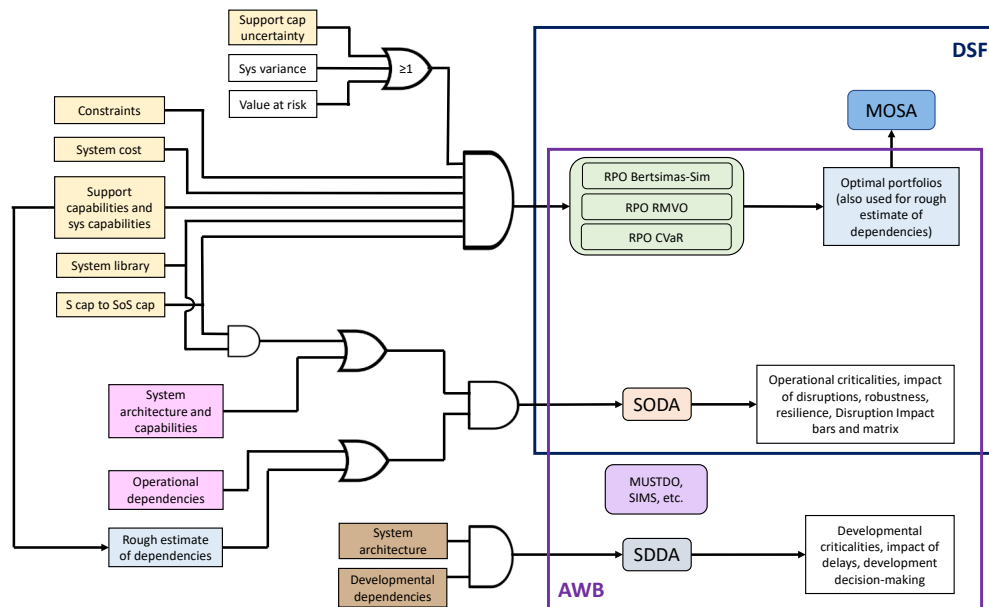


Figure 1. Inputs to AWB and DSF

## Robust Portfolio Optimization (RPO)

The SoS modeling and analysis problem can sometimes be described as a combinatorial problem to determine the most promising portfolio of individual systems in which to invest to achieve a certain capability. For instance, in space systems architecture, mission designers need to find the best combination of spacecraft, launch systems, launch windows, commodities to be transported, existing space systems such as the International Space Station (ISS), and other capabilities to achieve the goal of a long journey and possible settlement on other space bodies. The process of selecting the optimal portfolio considers the Life Cycle Cost (LCC) and the capability of individual systems. The process keeps both cost and certain types of risk (e.g., developmental cost) within an acceptable level while accounting for the impact of various forms of data uncertainty. Implementation and solution of RPO for a particular SoS design problem yields a set of Pareto optimal solutions (each solution is a portfolio of systems) corresponding to a user-defined risk aversion factor. RPO analysis allows an SoS manager to explore the design space of available options when designing a mission. Additionally, for a chosen portfolio, further desired analysis can be carried out using other AWB tools.

The risk can be characterized into three types: developmental, operational, and simulated. Based on these three types of risks, three flavors of RPO have been developed: (1) Robust Mean Variance Optimization (RMVO) includes developmental risks (Davendralingam & DeLaurentis, 2015; Rubinstein, 2002), (2) the Bertsimas-Sim method involves operational risks (Bertsimas & Sim, 2004; Davendralingam & DeLaurentis, 2013), and (3) Conditional Value at Risk (CVaR) addresses simulated risks (Davendralingam & DeLaurentis, 2014; Shah et al., 2015). Based on the problem at hand, the stakeholder needs to select a specific flavor of RPO.

## Systems Operational Dependency Analysis (SODA)

SODA methodology addresses the operational domain of an SoS by providing an analysis of the impact of dependencies between constituent systems on the propagation of the effect of disruptions (Guariniello et al., 2019). In SODA, a parametric model of system behavior is combined with a network representation for the system architecture. A small set of





parameters is used to simplify the dependencies between each system. These parameters were chosen to represent aspects of the dependency of the operability of a system on the operability of another system. The Strength of Dependency (SOD) represents a linearized operational dependency between systems in the case of minor disruptions. The Criticality of Dependency (COD) represents the loss of operability due to major disruptions. The Impact of Dependency (IOD) models the boundary between the small disruption regime and the major disruption regime.

Based on the parameters of the model, SODA can quantify the cascading effect of disruptions in the architecture and constitutes a quantitative method of risk analysis that can be used to expand the traditional risk matrix. The algorithm can also model partial failures, both deterministic and stochastic, and multiple paths of propagation within the model. SODA thus provides early-stage feedback for the architecture's design, reducing the amount of simulation and other verification methods required to ensure mission feasibility and to identify criticalities and areas of potential emergent behavior (Guariniello et al., 2019).

### **Systems Developmental Dependency Analysis (SDDA)**

SDDA is the counterpart of SODA in the developmental domain. It is a parametric model of developmental dependencies and constitutes an extension of PERT/CPM techniques which adds partial parallel development and partial dependencies.

The outcome of SDDA modeling and analysis is a quantitative assessment of the beginning and completion time of activities in a project (e.g., development of technologies, systems, or SoS capabilities), accounting for the combined effect of multiple developmental dependencies and of possible delays in the development of predecessors. The *lead time* (i.e., the amount of time by which a system can begin to be developed before a predecessor is fully developed) is calculated based on the dependencies and the performance of predecessors.

SDDA allows for deterministic or stochastic analysis. In deterministic analysis, an amount of delay is assigned to each system, and SDDA evaluates the resulting schedule. In stochastic analysis, the amount of delay in each system follows a probability density function with the resulting beginning and completion time of each system also as a distribution. SDDA identifies the most critical nodes and dependencies with respect to overall development time and delay propagation, important decision support for both system managers and the SoS architect. Results from the analysis are used to compare different architectures in terms of development time, risk, and capability of absorbing delays.

### **AWB Interoperability, Extensibility, and Usability Upgrades**

A significant part of this research effort included upgrades to the AWB. The original scripts and functions representing the various elements of AWB were upgraded to a set of qualified Python packages. This process included the implementation of industry-standard software control and revision processes and standards (GitHub Resources, 2022). These upgrades enable ease of continued development of the packages and their capabilities across academic, industry, and government teams. The upgrades also make the components of AWB more friendly for developers and users and ease any future burden of delivery.

A summary of the upgrades for the different components is outlined here:

- **Robust Portfolio Optimization (RPO):** RPO was upgraded to a fully Python-based application, removing the need for a MATLAB license. A set of input and output data control and validation methods was provided for interaction with RPO. RPO was also integrated into a controlled Python product with available pip and Anaconda packages. This process included the



addition of unit and integration testing, static code analysis, and implementation of CI/CD. The input for RPO was converted into a compact text-based file format called JavaScript Object Notation (JSON) from a Microsoft Excel datasheet. Furthermore, Jupyter Notebooks (Jupyter, n.d.) were used for adding the input data, running the Python-based code, and analyzing the results on a webpage in an interactive manner.

- **Systems Operational Dependency Analysis (SODA):** SODA was integrated into a controlled Python product with available pip and Anaconda packages. This process included the addition of unit and integration testing, static code analysis, and implementation of CI/CD.
- **Systems Developmental Dependency Analysis (SDDA):** SDDA was integrated into a controlled Python product with available pip and Anaconda packages. This process included the addition of unit and integration testing, static code analysis, and implementation of CI/CD.
- **AWB:** AWB was integrated into a controlled Python product with available pip and Anaconda packages. Applicable automation was developed for AWB to ease the control and installation of necessary dependencies (i.e., RPO, SODA, and SDDA) across platforms. This process included the addition of unit and integration testing, static code analysis, and implementation of CI/CD.

Several other specific upgrades were implemented to make it easier for users to develop appropriate data to define the problems AWB is meant to address. These upgrades also support the ongoing work to develop an appropriate user interface (UI) for building AWB problem data. In particular, the team also improved the RPO package by reimplementing the code surrounding the data and optimizer handling logic. This included changes to the interface used to connect to RPO. The team made RPO features functional again within a UI in Python so that users can pass parameters to RPO, and RPO output plots and results can be displayed. The whole AWB UI allows the selection of tools such as RPO, SODA, SDDA, etc. Custom images or logos are displayed based on the tool selected. Input to the tool can be a built-in example, custom-build scenario (TBD), or file import. The RPO tool has tabs for setup and output tabs which allows for input selection, input setup, and display analysis output. SODA and SDDA have an Interactive tab with functionality that will be implemented in the future.

For all AWB tools, a Links widget will allow dependencies between nodes to be defined. A directed graph is used to show the dependencies between systems. More options for SODA and SDDA can be selected.

The RPO Output tab plots Pareto frontiers for cost vs. SoS performance index. There is also a table of allocations that shows the numbers of individual assets at each cost point. The SODA Output tab can show either repair impact or failure impact plots. SDDA output shows the resulting schedule of development based on SDDA analysis.

A side effect of the new RPO updates was the breaking of a convenience feature within the tool suite which allowed for quickly using RPO results as input into SODA. Here, a user-selected RPO allocation is “automatically” fed into SODA without parameter adjustment by the user. Therefore, further work was done to reconnect RPO output to SODA analysis. This enables SODA output plots to appear in addition to RPO output within the new UI.

## RPO Data Validation Overview

The RPO software requires users to create instances of code classes (e.g., System, Capability) that have unique data requirements. Validating data is a common challenge in software development. The GTRI team has approached this problem by using JSON-based schema to capture information about data requirements so that only valid inputs are used to run scenarios within the RPO tooling. JSON Schema (a standard for developing these validations; JSON Schema, n.d.) is used. Generating the JSON Schema for all classes in the RPO library



would require significant manual effort to create an additional effort to update each time the class definitions (i.e., data models) are modified. To alleviate this, all RPO data models are defined using a Python standard library (data classes), which can be used to automate the creation of JSON Schema for each class. This allows continuous updating of the schema for validating instances of objects against the expected representation without requiring a manual definition of the schema.

## Schema Generation Script

While the schema for class instances is not expected to change frequently, RPO needs a method to automate the process. To aid developers, a “generate\_schema” script is included in the RPO scripts folder. Schema is checked into the repo to ensure that any changes undergo review by SERC developers. Once the script is called, users can commit changes, and all validation changes will propagate to the RPO tests. Invocation documentation can be found in the RPO README.

## Automation of SODA/SDDA Data from RPO Problem

A previous algorithm existed for converting the inputs used in RPO into those used in SODA and SDDA. Namely, the three dependency characteristic matrices of Strength of Dependency (SOD), Criticality of Dependency (COD), and Impact of Dependency (IOD). This pipeline utilizes the system support requirements and outputs to identify potential relationships among systems. For example, if System A produces Resource R, and System B requires resource R, then B is assumed to depend on A. Similarly, we also capture relationships among systems and capabilities. Namely, if System A has Capability C, then C is said to depend on A. In reality, it is also possible for a system to require a capability as well as a capability to require another capability, but this information is not possible to deduce from the inputs of RPO alone. Furthermore, while the algorithm is designed to provide sensible values for each relationship’s strength, criticality, and impact, it is generally accepted that expert opinion is necessary to achieve reliable results from SODA and SDDA analysis.

An expansion of this algorithm to approximate the inputs to SODA/SDDA has been designed as a proof-of-concept. The desire is to display approximated parameters to the user via an interactive data entry widget which will also provide the ability to correct them as needed. This project included the final stages of automating the algorithm and integrating it with the data structures used to run the Python implementation of RPO. An initial implementation of the data entry widget is being developed in parallel.

## Anti-Surface Warfare (ASuW) Problem

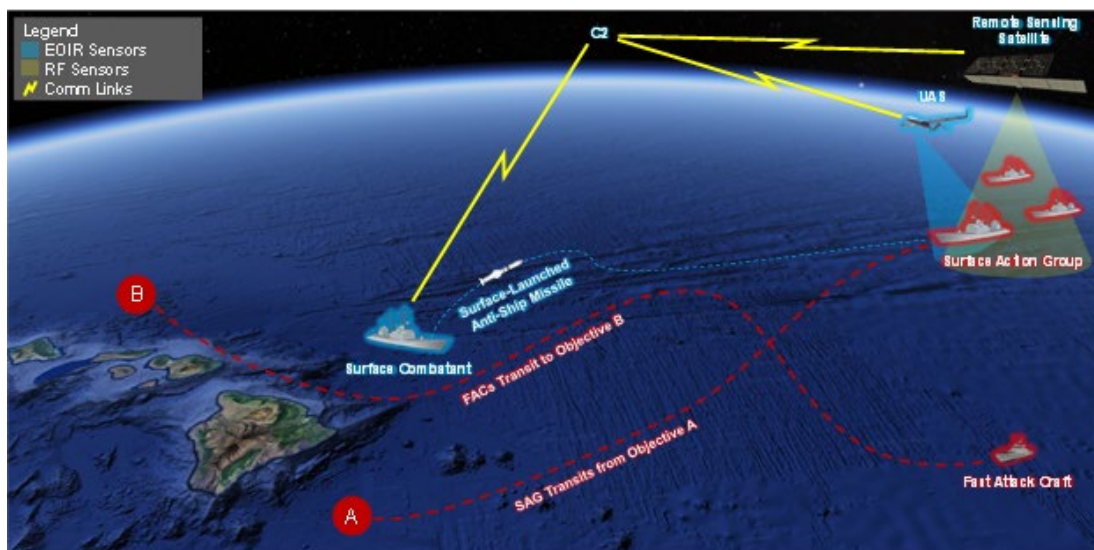
### Problem Formulation

The Anti-Surface Warfare (ASuW) problem selected is intended to be notional while remaining illustrative of an ASuW problem that may be relevant to the U.S. Navy ( Broadfoot et al., 2018; Kaymal, 2013; Neumann, 2021). The basic model has two surface threats traversing a body of water, (1) a Surface Action Group (SAG) and a Fast Attack Craft (FAC) group. The SAG is composed of surface combatants (e.g., frigates, destroyers, and cruisers). The FAC group is composed of small and fast (40+ kts) vessels. The blue force must complete the Find, Fix, Track, Target, Engage, Assess (F2T2EA) kill chain against these threats. Therefore, the blue force must include sensors, shooters, and a command and control element that coordinates and decides how to task the other elements. The sensors can be space, airborne, and surface.

For this problem, the blue architecture does not include subsurface elements (e.g., submarines and underwater arrays). The ASuW problem selected is intended to be notional



while remaining illustrative of an ASuW problem that may be relevant to the U.S. Navy. This is a realistic warfighting problem and one that is addressed by a complex portfolio of assets from across multiple platforms and weapons. The basic model has two surface threats traversing a body of water: a Surface Action Group (SAG) and a Fast Attack Craft (FAC) group. The SAG is composed of surface combatants (e.g., frigates, destroyers, and cruisers). The FAC group is composed of small and fast (40+ kts) vessels. The blue force must complete the Find, Fix, Track, Target, Engage, Assess (F2T2EA) kill chain against these threats. Therefore, the blue force must include sensors, shooters, and a command and control (C2) element that coordinates and decides how to task the other elements. The sensors can be space, airborne, and surface. For this example of the problem space, the blue force architecture does not include subsurface elements (e.g., submarines, underwater arrays). Subsurface elements could be added to this analysis at a later date without a change in the methodology. A simple basic architecture is depicted in Figure 2.



**Figure 2. OV-1 of the Simple Notional Anti-Surface Warfare Scenario**

A more comprehensive construct based on a richer kill web is depicted in Figure 3. This construct increases multidomain effects and interdependencies as it includes additional sensors (i.e., Maritime Patrol Aircraft [MPA], Helicopter, Radar from Surface Combatants) and shooters (i.e., MPA, Attack Aircraft, Helo). Sensors are categorized into two sets: Electro-Optical/Infra-Red (EO/IR) sensors and Radio Frequency (RF) sensors. In this scenario, EO/IR sensors are primarily used for target identification, while RF sensors are used for target detection and potentially for cueing other sensors. All data and concepts illustrated in this notional example are derived from open-source data, primarily wikidata.org.

All the data for the assets and weapons described below was obtained or derived from wikidata.org. The roles/responsibilities and the specific values are not intended to be overly accurate or complete but capture coarse-level capabilities that illustrate the potential tradeoffs that the Analytical Workbench can assess.

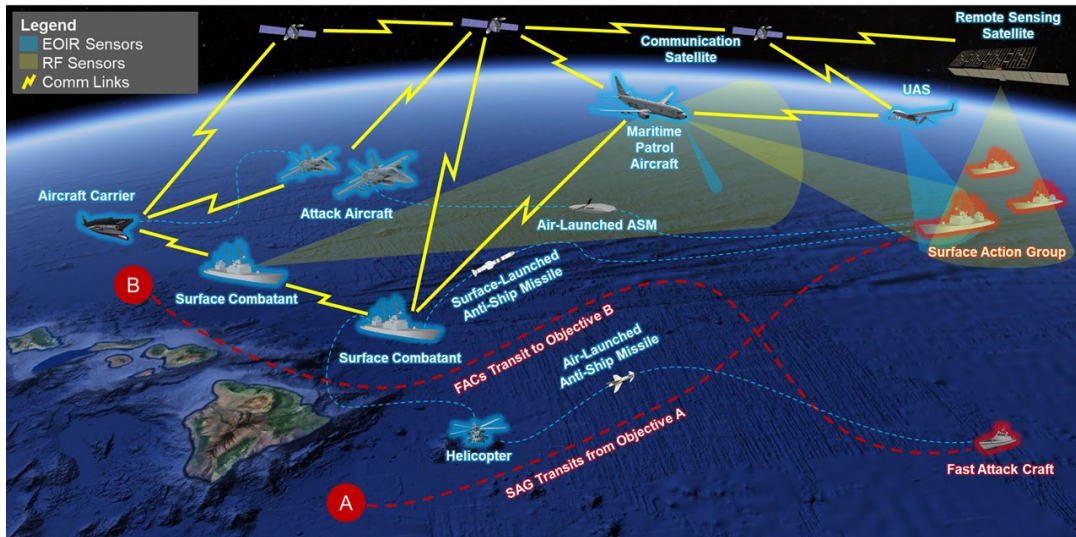


Figure 3. OV-1 of a More Comprehensive Anti-Surface Warfare Scenario

The blue force ASuW kill chain is based on the Find, Fix, Track, Target, Engage, Assess (F2T2EA) kill chain, with some simplifications to ensure the example remains unclassified.

- **Find:** This is the task of doing the initial detection of the red surface vessels. The result is a cue to other sensors to do the additional assessment of the potential target.
- **Fix:** In this formulation, the *fix* is primarily concerned with identifying the potential target. It requires distinguishing a target from its surroundings and is doctrinally described as “identifying an emerging target as worthy of engagement and determining its position and other data with sufficient fidelity to permit engagement.” This requires a cue from another sensor.
- **Track:** In this formulation, the process of formulating tracks for targets is highly abstracted. In reality, this task can require complex processes to fuse different sources of data and assess the error, maintaining custody of a target across one or multiple assets until a target solution is determined. Adding to the complexity, most systems today that can fix can also track. In reality, this task can require complex processes to fuse different sources of data and assess the error.
- **Target:** This step involves defining/selecting a capability to take action against an identified target, inclusive of the weapons, platforms with those weapons, other resources, and authorities. In the formulation defined for this effort, the targeting phase is highly simplified and presumed to be done with a high degree of certainty. In this formulation, the targeting phase is highly simplified and is presumed to be done with a high degree of certainty.
- **Engage:** For the purposes of this model, the engage phase primarily consists in launching a weapon against the target and evaluating a random chance of the weapon finding the target and killing it. This makes the problem tractable and able to produce outcomes measurable by the integrated set of methods. In reality, however, increased standoff ranges in contested environments inject time into the F2T2EA, something not accounted for in traditional kill chain analyses in general.
- **Assess:** the assessment phase of the F2T2EA kill chain is critical. However, for the purposes of this simulation, the process is highly simplified. Any surviving targets remain alive in the simulation and can be picked up by other sensors.

The assets required to complete the kill chain are listed in Table 1. The potential assets considered in the architecture are grouped by their domain. Weapons and personnel are the two



other categories of elements considered in the architecture mix analysis. For ease of understanding, the real names of assets are used, but all the properties and tactics, techniques, and procedures (TTPs) used for the assets are notional and unclassified.

**Table 1. Potential Assets for the ASuW Scenario**

Domain	Asset	Description
Space	Legacy Synthetic Aperture Radar (SAR) Satellite	Larger Space-based Remote Sensor that uses Synthetic Aperture Radar to detect surface vessels from their wake. Primary function: cueing.
	Small SAR Satellite	Smaller, more affordable, and less capable Space-based Remote Sensor that uses Synthetic Aperture Radar to detect surface vessels from their wake. Primary function: cueing.
	Electro-Optical/Infra-red (EO/IR) Imaging Satellite	EO/IR space-based remote sensing capability that may be able to identify surface vessels. Primary function: target identification. It may provide cueing but not the primary function.
	Communications Satellite	Space-based communications relay provides over-the-horizon communications.
Air	MQ-4C	Unmanned reconnaissance aircraft. Primary function: target identification, secondary function: target detection.
	P-8A	A Maritime Patrol Aircraft (MPA) can detect and identify surface targets.
	EA-18G	A Standoff Electronic Attack Aircraft that can passively detect surface targets.
	F/A-18E/F	An attack/fighter fixed-wing aircraft that can launch anti-ship weapons. Requires a CVN to launch from.
	MH-60S	A rotary wing aircraft that can detect, identify targets at close range, and launch short-range anti-ship weapons with limited lethality.
	F-35B	Short Take-off and Vertical Landing (STOVL) stealth aircraft can be operated from amphibious assault ships (e.g., LHA, LHD).
	F-35C	Carrier-capable fixed-wing stealth aircraft can only be launched from CVNs.
Surface	FREEDOM (LCS-1)	Mono-hull Littoral Combat Ship (LCS), a small, more affordable, but less capable surface combatant.
	INDEPENDENCE (LCS-2)	Multi-hull Littoral Combat Ship (LCS), a small, more affordable, but less capable surface combatant.
	ARLEIGH BURKE (DDG-51)	First generation (Flight I) of a modern missile-guided destroyer, no capability to support helo operations.
	MAHAN (DDG-72)	Second generation (Flight II) of a modern missile-guided destroyer, limited capability to support helo operations.
	OSCAR AUSTIN (DDG-79)	Third generation (Flight IIA) of a modern missile-guided destroyer, full capability to support helo operations.
	JACK LUCAS (DDG-125)	Future generation (Flight III) of a modern missile-guided destroyer, full capability to support helo operations and improved sensing/weapon systems.
	ZUMWALT (DDG-1000)	Best-in-class guided missile destroyer with improved sensors and weapon systems, signature management capabilities, but limited quantities of anti-ship weapons.
	TICONDEROGA (CG-47)	Legacy cruiser with moderate sensing capability but large missile capacity.
	BUNKER HILL (CG-52)	Modern cruiser with modern sensing capabilities and large missile capacity.
	WASP (LHA-1)	A small aircraft carrier that can support STOVL aircraft operations.
	AMERICA (LHA-6)	A small aircraft carrier that can support STOVL aircraft operations.
	FORD (CVN-78)	A large aircraft carrier that can support carrier-based aircraft operations.

Resource: ([wikidata.org](http://wikidata.org), n.d.)

As with the assets, the weapons are notional, with numbers obtained from unclassified sources. However, real weapon names are used to facilitate the understanding of the scenario and the results produced by the framework. The goal of the Anti-Ship Missile (ASM) weapon mix (Table 2) was to illustrate that a notional capability/cost tradeoff could be captured by the AWB and the discrete event simulation.

The conduct of the operations and employment of the different assets in the ASuW scenario model are dependent on a wide range of factors. This includes the physical



deployment of the different assets, the operation rules of engagement, and the actions of the red force actors (there may be more than one working at some level of coordination). For this analysis, we are assuming that the SAG and the FAC group are one red-force actor and are working in full coordination.

Operationally, for the blue force to be successful in the kill chain, their actions must result in the red force losing mission capability and/or deterring the red force from future engagement. Set-based methods containing tactical and intelligence input and assessment results will be required to evaluate the amount of reduction of red force capability needed for blue force success. Within the ASuW mission area, metrics of success are primarily the reduction of red force capability, weapons expended, and blue force casualties.

**Table 2. Anti-Ship Weapons**

Designation	Name	Launcher Domain	Range (nmi)	Speed (kts)	Cost (k\$)
AGM-114L	Hellfire	Air	6	864	150
AGM-119	Penguin	Air	100	633	800
AGM-158C	LRASM	Air	300	633	3,960
AGM-158D	JASSM-XR	Air	970	1026	1,500
AGM-84D	Harpoon	Air	50	461	500
AGM-84F	Harpoon	Air	170	461	600
AGM-84H/K	SLAM-ER	Air	150	461	3,300
BGM-109 Blk V	Maritime Strike Tomahawk	Surface	1350	493	1,409
RGM-184A	Naval Strike Missile	Surface	100	600	2,194
RGM-84F	Harpoon	Surface	150	461	600
RIM-174	Standard Extended Range Active Missile	Surface	130	2315	4,318

Resource: ([wikidata.org](http://wikidata.org), n.d.)

## Measuring SoS Capability

A relatively simple set of system capabilities was developed to measure how the various systems contribute to the overall ASuW scenario (Table 3). These are utilized by RPO and can be passed to supporting analysis (e.g., the Discrete Event Simulation developed using UPSTAGE; Arruda, 2018) for more detailed analysis. The better a system performs for each of these capabilities, the more likely it is to be allocated when it is SoS capability. These capabilities are intended to be notional and illustrative of the types of characteristics that may be used to assess how well an ASuW System-of-Systems performs. RPO performs optimization of system allocation against SoS performance. To facilitate this, five System of System Capabilities were defined for the overall scenario (Table 4). These SoS Capabilities are groupings of the individual system capabilities.



**Table 3. System Capabilities**

System Capability	Name	Measurement	Measurement Units
SC 1	Maritime Surveillance	Notional Area Surveillance Capability	1/3/9: Low/Med/High
SC 2	Identify Surface Contacts	Notional ID Capability	1/3/9: Low/Med/High
SC 3	Jam Ship Radars	Notional Jamming Capability	1/3/9: Low/Med/High
SC 4	Standoff Range	Weapon Range (nm)	nmi
SC 5	Disable Surface Combatant	P <sub>hit</sub> SC	%
SC 6	Damage Surface Combatant	P <sub>kill/hit</sub> SC	%
SC 7	Disable Fast Attack Craft	P <sub>hit</sub> FAC	%
SC 8	Damage Fast Attack Craft	P <sub>kill/hit</sub> FAC	%
SC 9	Quickness	Airspeed	kts
SC 10	Coverage	Flight Range	nmi
SC 11	Power Projection	Capacity	#

**Table 4. SoS Capabilities Defined for the Overall Scenario**

SoS-Capability	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11
SoS-Capability	Maritime Surveillance	Identify Surface Contacts	Jam Ship Radars	Standoff Range	Disable Surface Combatant	Damage Surface Combatant	Disable FAC	Damage FAC	Quickness	Coverage	Power Projection
ASuW Offensive											
ASuW Defensive											
ASuW Near Peer Actor											
ASuW Non-State Actor											
Maritime Awareness											

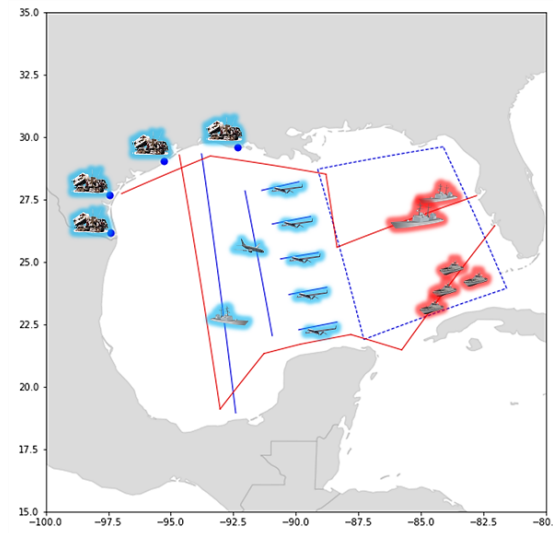
**ASuW-Specific Mission Analysis of Optimized Portfolios with UPSTAGE**

For this effort, the GTRI team utilized UPSTAGE to develop a more complex network for an ASuW mission, one incorporating both blue and red platforms and capabilities, to produce an analysis that would inform the AWB framework with improved mission fidelity. This will offer a much more complex representation than the tools without these dynamics being considered. Even so, the models for this effort are intended to demonstrate the general capability but will still fall short of real-world dynamic complexity. This unclassified implementation of UPSTAGE to the ASuW problem simulates a notional scenario where Red (Florida) is set to carry out a strike





mission against Blue (Texas) launch assets deployed along the coast (Figure 4). A Red Surface Action Group (SAG) and Fast Attack Craft (FAC) group move from their ports through northern and southern routes, respectively, to reach Blue's home shore.

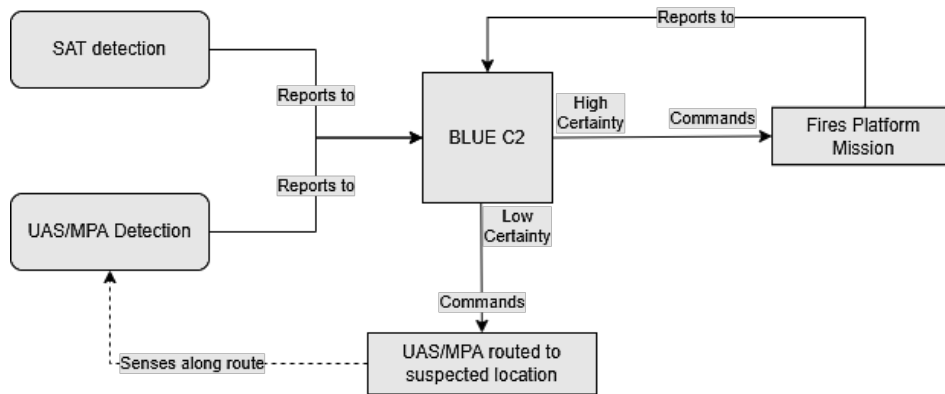


**Figure 4. Red (Florida) vs. Blue (Texas)**

Blue forces are arranged to provide a layered defense of their shore. Unmanned Aircraft Systems (UAS) fly patrol routes in the eastern portion of the Blue sea, while Maritime Patrol Aircraft (MPA) fly a patrol route to the west of the UAS. Further west, Blue naval assets such as DDGs and CVNs conduct patrols. The exact makeup of the Blue patrols and of the patrolling system's attributes are input into the UPSTAGE simulation. As Red forces move through Blue waters, Blue satellites may detect them in the dotted region in Figure 4. The satellites' detection capabilities are also inputs to the simulation, and it is possible that false targets can be found and reported to the Blue command.

Difficulty in analyzing parameterized force structures is parameterizing command and control (C2). UPSTAGE mitigates this difficulty through entity grouping and rehearsal features to support C2's selection of friendly assets based on user-defined capabilities.

The F2T2EA kill chain is abstracted in the ASuW simulation to follow this general flow (Figure 5). The Blue C2 will receive information from the systems given to it—based on a portfolio—and that information can have a variable certainty as a function of the system that performed the detection. Low certainty information will cause Blue C2 to follow up with UAS or MPA tasking to provide higher-quality track information. If the track quality is high enough, Blue C2 will initiate a fires mission from one of the available fires systems.



**Figure 5. F2T2EA Kill-Chain**

A fires mission will generally involve a flyout to a known or best-predicted position of a Red system using a platform with enough of a weapon class to ensure success. If only surface assets are available, they will be selected, but they are not preferred. Individual fires-capable systems are allowed to detect and fire on Red systems of their own volition. These include any land-based batteries or surface ships. The simulation will run until Red systems are destroyed, or they reach Blue shores. The time it takes to complete the scenario and the success/failure are the primary outputs. Secondary outputs can include resource usage, such as fuel and munitions, comms requirements, and other interactions.

### Initial ASuW Results

While the process of setting up the full ASuW problem formation is well in progress, initial results show some interesting trades. Running the problem through RPO and examining the results show a continuous improvement of the SoS performance score as the cost constraint is raised, as expected. On closer inspection of the allocations, more interesting results are seen. The parameters used for the initial results are listed in Table 5.

The results of the RPO run can be viewed in Figure 6, where SoS Performance Index is a non-dimensional measure indicating performance across all selected SoS Capabilities. For this initial example, all five of the SoS Capability measures were analyzed simultaneously. In the future, more nuanced results could be achieved by optimizing the SoS Capabilities individually.

The increase in overall SoS capability, as more money is spent, is a fairly obvious and expected result. The most noticeable trend in this chart is the divergence of performance at higher costs when more risk (lower conservatism) is allowed in the solution. More interesting results can be observed in the full allocation table, however. Table 6 shows how many of each system were purchased for the points plotted in Figure 6, a run of RPO on the ASuW scenario.

**Table 5. Initial Results Run Parameters**

	Minimum	Maximum	Steps
Cost (\$MUSD)	50.0	800.0	15
Risk (n.d)	0.2	1.2	3

From Table 6, it is apparent that the preferred low-cost solution (allocations 0–2) involves an investment in LCS ships carrying the Hellfire Longbow (AGM-114L), a currently experimental solution, with limited allocations of aircraft, DDGs, or dedicated anti-surface missiles like the AGM-84. However, as the cost constraint is relaxed and the optimizer can



afford more expensive systems, it quickly shifts to a solution based largely on amphibious assault ships, F-35Bs, and JASSM-XRs. In this middle range, RPO also demonstrates that Fix, Tracking, and Targeting can be largely based on unmanned assets.

As more money is allowed to be invested, the strategy again shifts to provide more SoS performance. This time once a full carrier is affordable, the investment strategy quickly switches to carriers, F/A-18E/Fs and AGM-84Ds. F-35Cs and JASSM-XRs are preferred if they can be afforded and mixed in with the F-18s as more money is allocated and if more risk is allowed. At that point, if more money is allocated, the same strategy is repeated, mixing in some Arleigh Burke destroyers until another carrier can be afforded.

This initial trend will be further analyzed as the team is able to integrate more tools (SODA, SDDA) with the ASuW scenario. As space domain-specific technology injection is integrated, the effects of satellite technologies on the allocations and analysis will be explored. Higher fidelity will also be executed via UPSTAGE to future predict how these different investment strategies might play out in a simulation. The team expects to include these results in the final report.

Even though the data used for this analysis is notional, some other interesting trends in the mix include the use of the Tomahawk Maritime Strike (MST) missile (i.e., BGM-109 BIK V). More conservative portfolios tend to use fewer MSTs as they have higher uncertainty in their ability to hit targets than the other missile options. AGM-84Ds tend to be preferred because of their cost-effectiveness and because the risk to the launching asset is not captured by the analysis.

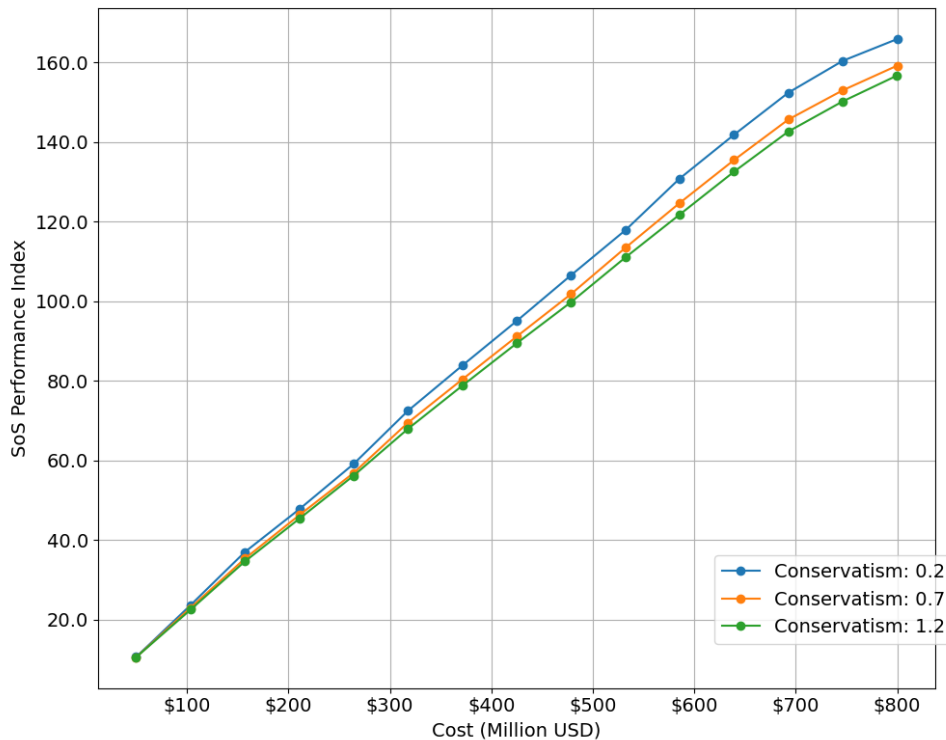


Figure 6. Initial Results: Cost vs. SoS Performance



**Table 6. Initial Results Allocations**

Alternative	0	1	2	3	4	5	39	40	41	42	43	44
Objective Value	10.7	10.7	10.7	22.6	22.2	21.8	159.6	157.6	156.5	169.5	166.7	164.6
Cost	\$ 50.00	\$ 49.94	\$ 49.96	\$ 103.53	\$ 103.53	\$ 103.57	\$ 746.40	\$ 746.36	\$ 746.42	\$ 800.00	\$ 800.00	\$ 800.00
Max Conservatism	0.2	0.7	1.2	0.2	0.7	1.2	0.2	0.7	1.2	0.2	0.7	1.2
Legacy SAR Satellite	0	0	0	0	0	0	1	1	0	0	0	0
Small SAR Satellite	1	1	1	2	2	2	5	5	8	8	8	8
EO/IR Satellite	0	0	0	0	0	0	0	0	0	0	0	0
Comm Satellite	2	2	2	4	4	4	20	20	20	20	20	20
MQ-4C	12	12	12	15	15	20	20	20	20	7	1	0
P-8A	4	4	4	3	4	4	4	4	4	4	4	4
EA-18G	0	0	0	0	0	0	0	0	3	1	1	1
F/A-18E/F	0	0	0	0	0	0	83	114	107	92	88	87
MH-60S	0	0	0	0	0	1	1	1	0	0	1	0
F-35B	0	0	0	8	8	0	0	0	0	0	0	0
F-35C	0	0	0	0	0	0	77	46	50	67	71	72
INDEPENDENCE (LCS-2)	0	1	1	0	0	1	0	0	0	0	0	0
FREEDOM (LCS-1)	1	0	0	0	0	0	0	0	0	0	0	0
ARLEIGH BURKE (DDG-51)	1	1	1	1	1	2	3	2	2	3	3	3
MAHAN (DDG-72)	0	0	0	0	0	0	0	0	0	0	0	0
OSCAR AUSTIN (DDG-79)	0	0	0	0	0	0	1	0	0	1	0	0
JACK LUCAS (DDG-125)	0	0	0	0	0	0	0	0	0	0	0	0
ZUMWALT (DDG-1000)	0	0	0	0	0	0	0	0	0	0	0	0
TICONDEROGA (CG-47)	0	0	0	0	0	0	0	0	0	0	0	0
BUNKER HILL (CG-52)	0	0	0	1	1	0	2	3	3	2	3	3
WASP (LHD-1)	0	0	0	1	1	0	0	0	0	0	0	0
AMERICA (LHA-6)	0	0	0	0	0	0	0	0	0	0	0	0
FORD (CVN-78)	0	0	0	0	0	0	2	2	2	2	2	2
AGM-84H/K	0	0	0	0	0	0	0	0	0	0	0	0
BGM-109 BIK V	5	5	5	0	0	28	145	174	179	211	228	230
RIM-174	0	0	0	0	0	0	0	0	0	0	0	0
AGM-158D JASSM-XR	0	0	0	27	25	0	99	61	60	77	66	61
AGM-158C LRASM	0	0	0	0	0	0	0	0	0	0	0	0
AGM-84D	6	6	6	6	7	7	174	236	221	191	183	113
AGM-84F	2	2	2	0	1	1	0	0	1	1	1	69
RGM-84F	8	8	8	16	16	16	48	40	40	48	48	48
AGM-119	0	0	0	0	0	0	0	0	0	0	0	0
RGM-184A (NSM)	1	1	1	0	0	1	0	0	0	0	0	0
AGM-114L	19	19	19	0	0	26	0	0	0	0	0	0
Navy Officer Personnel	64	66	67	133	137	109	864	847	863	849	851	854
Navy Enlisted Personnel	439	419	420	1277	1285	747	6668	6395	6409	6630	6655	6633
Navy Flight Personnel	9	11	12	15	19	15	171	173	180	183	175	174

Alternatives 6-38 were omitted from this figure

## Conclusions

The research team adapted a previously developed SoS-AWB to inform decisions in IAPRs. The AWB we developed and enhanced supports OUSD(A&S) for the rollout of the Adaptive Acquisition Framework and Capability Portfolio Management since our software suite can provide the analytic capability that is necessary to provide a solid foundation for acquisition investment decisions with clear traceability. These advanced prototypes provide broader insights (e.g., resource tradeoffs, cost-sensitivity analysis, and the most robust ASuW systems to be acquired in specific portfolios) for the stakeholder’s decision-making process. Future work could improve the tools to identify the following: how risk aversion affects portfolio optimization; technical dependencies among systems; developmental dependencies; and portfolio performance effects from stakeholder decisions. As a result, future work could assist in the activities for the new Acquisition Integration and Interoperability Office within OUSD(A&S).

## Acknowledgments

The authors acknowledge financial support from the U.S. Department of Defense through SERC/AIRC on research task WRT 1049.5, contract no. HQ0034-19-D-0003 and report no. AIRC-2022-TR-007.

In addition, the authors thank the following individuals who have significantly contributed



to the development of the SoS tools in various SERC and AIRC projects. Contributing to the original development of the Analytic Workbench: Dr. Karen Marais, Dr. Payuna Uday, Dr. Navindran Davendralingam, Dr. Zhemei Fang, Mr. Rakshit Chandrasaha, Dr. Judith Dahmann, Mr. Scott Lucero. Supporting further development and improvement: Dr. Kris Ezra, Mr. Robert Campbell, Mr. Clint Hanthorn, Dr. Paul Grogan, Dr. Brian Chell, Mr. Benjamin Stanley, Mr. Corey Batchelder, Dr. Daniel Browne, Dr. Craig Arndt, Dr. Zachary Welz, Mr. James Arruda, Mr. Joel Stansbury, Mr. Nicholas Bollweg, Dr. Valerie Sitterle, Mr. Joshua Mathews, Mr. Lucas Karinshak, and Dr. Peter Korfiatis.

## References

- Arruda, J. (2018). UPSTAGE: Universal platform for simulating tasks and actors with graphs and events. In *86th MORS Symposium*.
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35–53.
- Broadfoot, M., Bush, C., Harpel, B. L., Lajoie, T., Laube, P. H., O'Grady, M. R., Overman, E. A., & Parcus, A. (2018). *Examining operational and design effects of MH-60S with enhanced weapon systems in anti surface warfare missions*.
- Davendralingam, N., & DeLaurentis, D. (2013). A robust optimization framework to architecting system of systems. *Procedia Computer Science*, 16, 255–264.
- Davendralingam, N., & DeLaurentis, D. (2014). An analytic portfolio approach to system of systems evolutions. *Procedia Computer Science*, 28, 711–719.
- Davendralingam, N., & DeLaurentis, D. A. (2015). A robust portfolio optimization approach to system of system architectures. *Systems Engineering*, 18(3), 269–283.
- GitHub Resources. (2022). *The fundamentals of continuous integration in DevOps*. <https://resources.github.com/devops/fundamentals/ci-cd/integration/>
- Guariniello, C., Grande, M., Brand, C., Durbin, L., Dai, M., Das-Stuart, A., Alexander, R., Howell, K., & DeLaurentis, D. (2019). Quantifying the impact of systems interdependencies in space systems architectures. *70th International Astronautical Congress*.
- JSON Schema. (n.d.). *JSON schema*. Retrieved March 31, 2023, from <https://json-schema.org/>
- Jupyter. (n.d.). *Jupyter notebook*. Retrieved March 31, 2023, from <https://jupyter.org/>
- Kaymal, T. (2013). *Assessing the operational effectiveness of a small surface combat ship in an anti-surface warfare environment*.
- Neumann, N. (2021). *The shifting threats driving anti-surface warfare (ASuW) capabilities*. Naval Technology. <https://www.naval-technology.com/features/anti-surface-warfare-asuw/>
- Rubinstein, M. (2002). Markowitz's "portfolio selection": A fifty-year retrospective. *The Journal of Finance*, 57(3), 1041–1045. <http://www.jstor.org/stable/2697771>
- Shah, P., Davendralingam, N., & DeLaurentis, D. A. (2015). A conditional value-at-risk approach to risk management in system-of-systems architectures. *2015 10th System of Systems Engineering Conference (SoSE)*, 457–462.
- wikidata.org. (n.d.). <https://www.wikidata.org/>



# Portfolio Management Structures: System, Capability, and Mission Portfolios

**John Driessnack**—Principal Investigator, University of Maryland, Project Management Center of Excellence. He is also a Professorial Lecturer at American University's Key Executive Leadership Program, a Professor at Defense Systems Management College, and President of the College of Performance Management. Driessnack owns Olde Stone Consulting LLC, a sole proprietorship. In recent years, a member of two ANSI standard development teams and currently on PMI's Standards Insight Committee. A retired USAF Lt Colonel, Driessnack's military experience includes six major programs, including senior SPM for Global Broadcast System. Research and consulting interests include institutional economics and modeling quantitative performance management. [jdriess@umd.edu]

**Caitlin Kenney, PMP, P.E**—is a Civil Engineering, Project Management PhD Candidate at the University of Maryland. She has an MS in Systems Engineering from UMD (2018) and a BS in Industrial Engineering from Northeastern University (2011). She has over 10 years of experience supporting DoD projects, including two DoD Agile Pilot programs for the Army Contracting Writing System (AWS) and the AEGIS BL 10 Software Development Program. She currently works for International Systems Management Corp as the Enterprise Agile Coach for the PEO IWS Forge Software Factory, supporting agile and DevSecOps project management initiatives for the U.S. Navy. [ckenney7@umd.edu]

## Abstract

Since 2008, ongoing attempts by the DoD to support decision-making across capability portfolio management have been unsuccessful. Proposed is a multidimensional portfolio structure schema, which utilizes ANSI Standard for Portfolio Management and is informed by ISO Standard for Building Information Modeling (BIM)s. The schema creates a non-hierarchical structure of three portfolio types across component PEOs representing product/platforms, component operational units representing capabilities, and combatant and supporting commander Operations Plans (OpsPlan) representing missions. The multidimensional nature of the structure allows for enhanced management insight and decision-making using structured performance management across the DoD Decision Support Systems (D2S2). Observations and challenges discussed range from the misalignment of Joint Capability Assessment (JCA) with the field use of Universal Joint Task List (UJTL) to not capturing cost estimate's quantitative risk data. The path forward outlines building a notional multidimensional programmatics model, which demonstrates how key data can be aligned with Mission Engineering and Systems Engineering models, allowing for full utilization of evolving Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques to enhance management insight and decision-making across the enterprise.

## Introduction

The University of Maryland, Project Management Center of Excellence, conducted research in support of Capability, Mission, and PEO (CMP) Portfolio Performance Analysis and Visualization task.<sup>1</sup> This research paper focuses on portfolio performance analyses and visualization across platforms, capabilities, and missions managed across DoD PEO portfolios. The research supports the National Defense Authorization Act (NDAA) Sec. 913 (FY18 NDAA) and Sec. 801 and 836 (FY22 NDAA), by identifying data-driven approaches to analytic insight at the program and portfolio levels.

This research, cognitive of the significant change in the defense acquisition environment over the past decade, looked at recent DoD attempts for portfolio performance analyses and visualization across PEO portfolios of systems and also capability and/or mission portfolios.

---

<sup>1</sup> 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 HQ003421F0480



Recent changes to support capability portfolio management, such as the Integrated Acquisition Portfolio Review (IAPR) efforts, are struggling to produce actionable advice for senior leaders. The “capability” structure utilized is loosely aligned with CJCS Joint Capability Assessment (JCA) structure. The newly created Chief Data and Artificial Intelligence Office (CDAO) has been gathering data with the Advance Analytics (ADVANA) systems started by the OSD Comptroller.<sup>2</sup> We observed a general frustration in attempts to use the classic program-centric acquisition structure and metrics along with decision tools, such as Spruill Charts, in a program or portfolio aggregate, whether at product/platform, capability, or mission views. Using the existing structure was not providing clarity to the ever-increasing complicated and complex<sup>3</sup> nature of the underlying SoS product/platforms and SoS decisions support management structure used to make decisions to achieve joint warfighter capabilities.

The Government Accountability Office (GAO) identified in 2015 that the Department of Defense (DOD) was not effectively using portfolio management to optimize its weapon system investments, as evidenced by affordability challenges in areas such as shipbuilding and potential duplication among some of its programs. Best practices recommend assessing investments collectively from an enterprise-wide perspective and integrating requirements, acquisition, and budget information, but several factors inhibit DOD’s ability to do so. (GAO, 2015b)

In 2019, the Section 809 Panel wrote extensively on how the DoD should move to a more industry-standard<sup>4</sup> approach for portfolio management for the PEO collection of product/platforms. Many of those recommendations have not been implemented; thus poor portfolio management structures and practices persist.

Many parts of the D2S2 have their own portfolio approach. The Chairman of the Joint Chiefs of Staff (CJCS) has Capability Portfolio Management Reviews based on JCAs as part of the requirements systems known as JCIDS.<sup>5</sup> The reviews are conducted by Functional Capability Boards (FCB), which also align with JCAs. USD(R&E) holds Technology Portfolio Management Review (TPMR) for 14 technology areas. USD(A&S) conducts Integrated Acquisition Portfolio Review (IAPR) based on an organizational structure poorly aligned to JCAs. The Cost Assessment and Program Evaluation office holds Strategic Portfolio Reviews (SPRs). There does not appear to be any alignment across these various portfolio reviews, making it difficult if not impossible to aggregate across the enterprise.<sup>6</sup>

Each part of the D2S2 is a decision system that takes a unique hierarchical approach. The use of a hierarchical portfolio structure method for integration and trade-offs can work, but not if the various structures cannot be aligned. The migration to network structure was recommended shortly after the department moved to the capability-based planning approach driven by the 2001 Quadrennial Defense Review and the subsequent Aldridge Study in 2003. At the time, it was recognized the complicated nature and the need for non-hierarchical approaches. The allocation of resources is

---

<sup>2</sup> The ADVANA effort was described to the research team by some OSD staff as a data mesh quickly turning into a data swamp.

<sup>3</sup> We separate the complicated from the complex as defined in Cynefin Framework by Dave Snowden.

<sup>4</sup> Project Management Institute (PMI) publishes an ANSI Standard for Portfolio Management. There is also an ISO standard in the 21500 series for project, program, and portfolio management.

<sup>5</sup> JCIDS is the Joint Capabilities Integrated and Development Systems, CJCSI 5123.011.

<sup>6</sup> Technology portfolios would flow directly into systems/platform portfolios.



not only at a given level and within a given concept of operations, but also across levels and configurations. Anyone who imagines that analysts can readily compute the relative worth of an additional fighter aircraft, missile launcher, or company of tanks probably has a simplistic and rigid notion of military operations and a correspondingly simple-minded way of comparing worth (e.g., by their relative lethality in a duck-shooting contest). It is better to adopt the spirit of portfolio analysis and recognize the role of multidimensional trade-offs and subjective judgments. This view may be heretical to operations researchers, but it is true nonetheless. (Davis, 2002)

The architecture/civil engineering/building industry has created an “Organization and digitization of information about building and civil engineering works, including building information modeling (BIM)” international standard, known as ISO 19650. A framework for pulling programmatic, engineering, and sustainment information together. The defense department has built the DoD Architecture Framework (DODAF), a structure “designed to meet the specific business and operational needs of the DOD.” It might be useful, but unlike the WBS structure in weapons acquisitions, it has not taken hold across products/platforms even as the network-centric approach has permeated throughout products/platforms. The DODAF structure is not used in any of the D2S2 systems.

The challenge at the enterprise level is no robust integrating structures across the various organizations within OSD and its components.<sup>7</sup> The DoD operates with an incomplete, diverse D2S2 across six or more decision systems, creating a complex framework for making decisions. A decade ago, it was noted, “It is arguably time for the strategic level of analysis to be revisited” (Davis et al., 2008). It is not going to be simple, but like any wicked set of problems, they need to break down into manageable challenges, which is what the three aligned portfolio structures can provide.

## Background

The Acquisition Innovation Research Center (AIRC) was created by the Secretary of Defense in September 2020 in response to 10 U.S.C. 2361(a) utilizing the DoD’s Systems Engineering Research Center (SERC) University-Affiliated Research Center (UARC).<sup>8</sup> Among the research tasks was an emphasis on portfolios/missions with a *Data-driven capability portfolio management pilot* to prototype capability to enhance data-driven decision-making regarding acquisition and support programs (UARC, 2021) Working from a previous effort to create a Model-Based Portfolio Analysis Capability for the Joint PEO for Chemical, Biological, Radiological and Nuclear Defense (JPEO-CMRND; WRT, 2020), Dr. Daniel DeLaurentis with his Purdue research team and other university partners

adapted a previously developed systems-of-systems analytic workbench (SoS-AWB) of analytic tools to create a decision-support prototype, effective for informing decisions in Integrated Acquisition Portfolio Reviews (IAPRs). These advanced prototypes provide a broader range of insights (e.g., resource trade-offs, cost-sensitivity analysis, etc.) for stakeholder decision making.

The report notes under a portfolio-centric approach:

---

<sup>7</sup> DoD components include OSD, CJCS, DoD Inspector General, Military Departments, DoD field activities, the Combatant Commanders, and some other minor organizations.

<sup>8</sup> University of Maryland is a member of UARC.





The Department of Defense (DoD) has an increasing focus on Mission Engineering (ME) analysis and architecture development for modernization decisions, including investments and prioritization related to requirement development and selection of capabilities to support various concepts of employment and technological improvements. Typically, however, systems engineering tools focus on the system itself. That is, the tools may not translate the complexities of mission engineering analysis into the configuration in a way that is both (a) meaningful to the requirements within the trade space of capabilities and (b) flexible, scalable, and configurable to integrate with other analyses. To this end, recommendations from an advisory panel suggested that the DoD approach should take a more holistic and portfolio-centric method for acquisitions rather than the current program-centric approach. In our prototype, systems and technologies are evaluated within an overall portfolio, exposing how each component plays a role in the realized capability while connecting the mission needs of warfighters with acquisition decisions. Continued development along these lines will eventually pave the way for the establishment of Acquisition Integration and Interoperability (AII), which should be based on mission and digital engineering, using data-driven methods (AIRC, 2022).

Though the concepts are solid, they have not evolved into usable tools for OSD decision-makers within the DoD Decision Support Systems (D2S2), which has evolved over the past 60 years, but is fundamentally the same structure of interfaced, but not aligned, decision systems. If it was a weapon system of systems, it would be considered poorly integrated and not interoperable. During the past 60 years, the underlying weapon and other products/platforms have grown more integrated and interoperable. As Vice Admiral Arthur Cebrowski noted at the end of the last millennium,

Network-centric warfare and all of its associated revolutions in military affairs grow out of and draw their power from the fundamental changes in American society. These changes have been dominated by the co-evolution of economics, information technology, and business processes and organizations, and they are linked by three themes:

- The shift in focus from the platform to the network.
- The shift from viewing actors as independent to viewing them as part of a continuously adapting ecosystem.
- The importance of making strategic choices to adapt or even survive in such changing ecosystem.

As the OSD and other DoD components explore portfolio management, which industry standards is also evolving, the DoD cannot ignore that it is fundamentally still in a major platform-centric management structure. As Admiral Cebrowski notes at the end of the article, as B. H. Liddell Hart said, “The only thing harder than getting a new idea into the military mind is getting an old one out” (Cebrowski & Garstka, 1998). The DoD should move toward a network of portfolios when managing product.

Much has been written on the need for the DoD to effectively use portfolio management, both for improving the DoD’s acquisition outcomes (GAO, 2007) as well as at an enterprise level to integrate DoD Decision Support Systems (D2S2; GAO, 2015a).



DOD attempted to standardize portfolios in the 2006 to 2008 time period. However, a former senior official who was involved in that effort said the mapping was “impossible” and that there was organizational resistance because the portfolios did not align with many decision makers’ areas of responsibility. Many of the enterprise- and service-level officials we interviewed said using a wide variety of constructs is necessary and sometimes beneficial given the different roles and perspectives of the organizations involved. However, when they want to analyze their portfolios from another perspective—for example, examining funding associated with joint capabilities areas—they have to go through extensive mapping exercises. (GAO, 2015a)

For programs, the Program Executive Officer has the requirement to balance risk, cost, schedule, performance interoperability, sustainability, and affordability of a portfolio of acquisition programs (GAO, 2007). The Section 809 Panel report provides almost a hundred pages on how PEOs could be more effective as empowered Portfolio Acquisition Executives (PAE) with a half dozen portfolio-specific recommendations (DTIC, 2019). The 809 Panel recommended a transition from the program-centric execution model to a portfolio model with an increased enterprise view, which meant various portfolio views led within the requirements structure. The panel also recommended implementing best practices for portfolio management. The effectiveness of D2S2 “must be assessed in terms of developing, delivering, and supporting defense systems that enable US dominance. ... For more than 50 years, the fundamental structure and focus of acquisition have been on MDAPs,<sup>9</sup> but the nature of capabilities has changed” (Ahern & Driessnack, 2019).

In 2008, the DoD published the Directive on Capability Portfolio Management, DoDD 7045.20. It establishes the policy to use capability portfolios following the “existing joint capability areas (JCA) structure.” The directive called for “Capability Portfolio Strategic Plans” and creating co-leads with “no independent decision-making authority.” The directive was not well implemented, nor has it been effective in using the Deputy’s Management Action Group (DMAG) or “ensuring alignment to strategic priorities and capability demands” (DOD, 2019).

The latest DoDD 5000.01<sup>10</sup> calls for “Capability portfolio management, mission engineering, and integration analysis using an effects/kill chain framework will be employed to assess the integration and interoperability of the SoS required to execute critical mission requirements.” Recently USD(A&S) has reorganized into so-called<sup>11</sup> Capability Portfolios whose “mission is to use Capability Portfolio Management to analyze, manage, and inform acquisition and resourcing decisions in platform and weapon portfolios” (DOD, 2023). Last year an Integrated Acquisition Portfolio Review (IAPR) was established but has also not been successful. As a result, in late January 2023, an Acquisition Integration and Interoperability (AI2) concept was outlined to be established within OSD(A&S) to:

- “Enable the delivery of integrated defense capabilities
- Drive adoption of threat-based mission thread analysis ...
- Acquisition portfolio reviews to drive resourcing and enterprise decision ...IAPR
- Establishes an OSD entity to align service-specific systems acquisition programs, prototypes, and S&T projects to deliver joint integrated capabilities.

---

<sup>9</sup> MDAP is a Major Defense Acquisition Program.

<sup>10</sup> DoD acquisition directive and instructions were significantly revamped from 2020 to 2022.

<sup>11</sup> We use the term “so-called” because the structure does not align with the CJCS JCA structure.



- In partnership with key stakeholders across OSD, the Joint Staff, and the Military Departments and Military Services, AI2 will deliver dedicated analysis, planning, resource recommendations, and portfolio management necessary to deliver joint capabilities across the Department.”

Overall, the approach does not address the enterprise level; nor does it establish a network structure of portfolio below the enterprise; nor does it adopt portfolio management best practices as recommended in the Section 809 panel and defined by ANSI standards for portfolio management.

## Research Tasks

The research effort is currently broken into three phases with a goal to “expand and enhance capability and performance management insights across DoD acquisitions program, including at Mission and Program Executive Officer (PEO) portfolio levels. Two main thrusts (1) focus on portfolio funding profile and (2) development of a portfolio executive dashboard to provide integrated data/views for missions, capability, and product/platform. Each phase expands both thrusts, which are perceived to be synergistic.

The team interfaced with a half dozen current OSD portfolio managers and participated in weekly OSD level meetings on improving data analytics to support and improve insights for the portfolio managers. Two comprehensive reviews were conducted in December 2022 and February 2023 with OSD staff as well as numerous other meetings to clarify observations and insights. The team also met with PEO IWS staff on several occasions as well as Navy and Air Force staff relative to those components data systems, such as the Assistant Secretary of the Navy (Research, Development, and Acquisition) (ASN(RD&A)), Information System (RDAIS), and Army/Air Force/Space Force Project Management Resource Tool (PMRT).

### Portfolio Funding Profile (Task 1a)

Portfolio Level Funding and Quantities Chart (see Table 1), commonly known as the “Spruill Chart,” named after Dr. Nancy Spruill, has been around for over 20 years (Woolsey, 2018). The chart is explained in detail in the DoD Cost Estimating Guide. The research team was tasked with creating a portfolio version by aggregating all the data for all the programs/systems within the portfolio, whether those portfolios are by PEO of Systems, Capabilities, or Missions. A notional minimum viable product (MVP), a wireframe mockup, using PEO IWS as an example for a portfolio-level dashboard was created and reviewed with OSD portfolio managers (Kenney & Kwapong, 2023b). The concept was to move beyond “charts” to more visually integrated data graphics. Further work was suspended due to data quality and access restrictions.



Table 1 The Program Funding and Quantities Chart

Program Funding & Quantities		Acquisition to O&S Cost Ratio						(BY 2019)	Curr Est	Δ Current	Δ Original	
		Total Required Acq (BY\$M): \$14,782 32%						PAUC:	581.9M	+4.6%	+10.2%	
(\$ in Millions / Then Year)		Prior	FY19	FY20	FY21	FY22	FY23	FY24	FY25	FY21-25	To Comp	Prog Total
<b>RDT&amp;E</b>		BLIs:										
Prior \$ (PB 20)			416.5	1,718.1	1,092.7	1,005.0	220.9	16.0	0.0	2,334.6	0.0	4,469.1
Current \$ (POM 21)			447.8	1,652.0	1,256.0	995.0	235.0	0.0	0.0	2,486.0	0.0	4,585.8
Delta \$ (Current - Prior)			31.3	(66.1)	163.3	(10.0)	14.1	(16.0)	-	151.4	-	116.7
Required <sup>1</sup> \$			488.1	1,635.5	1,205.8	985.1	251.5	0.0	0.0	2,442.3		4,565.8
Delta \$ (Current - Required)			(40.3)	16.5	50.2	10.0	(16.5)	-	-	43.7	-	20.0
<b>PROCUREMENT</b>		BLIs:										
Prior \$ (PB 20)			0.0	0.0	522.7	1,999.6	2,313.5	2,650.6	2,346.3	9,832.5	1,872.2	11,704.8
Current \$ (POM 21)			0.0	0.0	562.0	1,754.0	2,385.0	3,012.0	2,133.0	9,846.0	2,152.0	11,998.0
Delta \$ (Current - Prior)			-	-	39.3	(245.6)	71.6	361.4	(213.3)	13.5	279.8	293.2
Required <sup>1</sup> \$			0.0	0.0	562.0	1,859.2	2,385.0	2,861.4	2,026.4	9,694.0	1,974.1	11,668.1
Delta \$ (Current - Required)			-	-	-	(105.2)	-	150.6	106.7	152.0	177.9	329.9
<b>MILCON</b>		BLIs:										
Prior \$ (PB 20)			0.0	1.5	1.7	0.0	1.7	16.0	2.9	22.3	15.3	39.1
Current \$ (POM 21)			0.0	1.4	1.7	0.0	2.0	2.1	3.0	8.8	12.6	22.8
Delta \$ (Current - Prior)			-	(0.1)	(0.0)	-	0.3	(13.9)	0.1	(13.5)	(2.7)	(15.3)
Required <sup>1</sup> \$			0.0	1.5	1.8	0.0	2.0	2.2	3.3	9.3	12.6	23.4
Delta \$ (Current - Required)			-	(0.1)	(0.1)	-	-	(0.1)	(0.3)	(0.5)	-	(0.6)
<b>SYSTEM O&amp;M<sup>2</sup></b>		BLIs:										
Prior \$ (PB 20)			0.0	0.0	0.0	0.0	141.3	16.0	1,230.0	1,387.2	37,051.0	38,438.2
Current \$ (POM 21)			0.0	0.0	0.0	0.0	125.0	359.0	1,260.0	1,752.0	37,051.0	38,803.0
Delta \$ (Current - Prior)			-	-	-	-	(16.3)	343.0	38.0	364.8	-	364.8
Required <sup>1</sup> \$			0.0	0.0	0.0	0.0	118.8	362.6	1,318.7	1,800.1	35,198.5	36,998.5
Delta \$ (Current - Required)			-	-	-	-	6.3	(3.6)	(50.7)	(48.1)	1,852.6	1,804.5
<b>TOTAL</b>			416.5	1,719.6	1,617.1	3,004.5	2,677.3	2,698.6	3,579.1	13,576.6	38,938.5	54,651.2
Current \$ (POM 21)			447.8	1,653.4	1,819.7	2,749.0	2,747.0	3,373.1	3,404.0	14,092.8	39,215.6	55,409.6
Delta \$ (Current - Prior)			31.3	(66.2)	202.6	(255.5)	69.7	674.5	(175.1)	516.2	277.1	758.4
Required <sup>1</sup> \$			488.1	1,637.0	1,769.6	2,844.3	2,757.2	3,226.2	3,348.4	13,945.6	37,185.1	53,255.8
Delta \$ (Current - Required)			(40.3)	16.5	50.1	(95.3)	(10.2)	146.9	55.6	147.2	2,030.5	2,153.8
<b>QUANTITIES</b>			0	2	1	2	4	6	2	15	3	20
Current (POM 21)			0	2	1	2	4	6	2	15	3	20
Delta Qty (Current - Prior)			0	0	0	0	0	0	0	0	0	0
Required Qty <sup>3</sup>			0	2	1	2	4	6	2	15	3	20
Delta Qty (Current - Required)			0	0	0	0	0	0	0	0	0	0

We had three significant observations with this task. The first observation was with the required line and the point cost estimate nature of that line not representing the confidence level of the proposed required funding line. The second observation was simply understanding what products/platforms were within the portfolio, whether that was a PEO portfolio, a capability portfolio, or a mission portfolio. Finally, the concept of aggregating the individual program funding within appropriations would likely not be very useful as the movement of funds across programs is restricted by reprogramming rules.

### Capability/Mission Thread Portfolio Schedule/roadmap on PEO-IWS (Task 2a)

Under the concept of the OSD level Integrated Acquisition Portfolio Review (IAPR), a particular capability portfolio, made up of the product/platforms which provide the particular capability would be reviewed together, both individually and as an aggregate. Also, the capabilities used within selected mission threads would be reviewed to determine if end-user mission capabilities were being improved. As a result, the research team was tasked with developing a capability portfolio view that looked across systems/platforms within the capability.





in any standardized format that was traceable from year to year. Below the PE project code on the RDT&E document, the use of project code, there was no standardized structure on how descriptions, accomplishments, and plans were characterized. In many cases, project codes were not used. The use of AI/ML/NLP type techniques would not be productive to pursue at this time without some improvement in how the efforts were described in a more standardized approach to the project level.

The budget documents are structured for reporting, not for management of the program; thus, not characterized by a meaningful management structure, but rather for justification of dollars. In some cases, project # could present billions of dollars while other project # presented millions. There was no alignment with approved program WBS, nor capability, such as a JCA or UJTL item, nor a mission thread.

Overall, data management across a multidimensional portfolio will be a data challenge. To address portfolio data management, a review of NoSQL approaches, including key value pairs, is being explored to model not just programmatic data, but operational capabilities and missions (Kenney & Driessnack, 2023a). Traditional SQL approaches, in use today by most data systems within the DoD, require rigid, structured, relational databases. This approach is limited at scale for enterprise portfolios because it requires strict data formats which can be difficult to modify and prone to user error. As a result, data can be “lost” within these systems, making it difficult for portfolio managers to see the full picture. NoSQL approaches that interact with non-relational databases, such as column-oriented, document-oriented, key-value pairs, and graph databases, can be used for many-to-many relationships, such as multiple systems supporting multiple capabilities and missions.

## **Phases 2 and 3 Plan**

Phases 2 and 3 will expand to program performance management metrics, such as earned value, agile, and classic qualitative and quantitative risk metrics expanded to include constraints, assumptions, issues, risks, and opportunities (CAIRO). The collection of CAIRO data is known as challenge management. The team will also explore the use of artificial intelligence (AI) and machine learning (ML) along with use natural language processing (NLP) on written assessments.

Portfolio performance management metrics for capabilities and missions are not readily available. One OSD capability manager provided their own set of metrics utilized to assess the portfolio of programs with many being proxies, such as looking at program obligation and expenditure rates as an indication of portfolio health. The availability of program-level data below the MDAP was reported as almost impossible for the portfolio managers to obtain. The more the portfolio manager had subsystems, components, or modifications within their portfolio, the less visible the data. The team will review existing data in use, as well as propose a data and metrics framework to support the multidimensional portfolio reporting needed for capabilities and missions.

It is generally accepted that risk management, and explicitly quantitative risk management, is key to managing forward with data. Risk Management is looking into the future, understanding the CAIRO that provides an understanding of how the leadership/management should focus to make decisions today that affect not just the future plan, but the confidence of that plan. The DoD does this type of work within cost and schedule estimating and required contractors on higher cost-plus contracts to incorporate risks in estimates to complete. The challenge is that the data does not make it into any of the OSD or other DoD component management systems except at level 1 of the WBS.

Given the finding in Phase 1, the team has made recommendations for specific changes



to the approach in the future, which are discussed below in the Path Forward section.

## Challenges

Table 1 shows funding and quantities for prior vis current vis required for the Execution/Budget/POM years with the deltas (current—prior as well as current—required) across each type of appropriation. The chart provides a complete picture of the funding for a particular product/platform. It does not show changes from the prior couple of years or changes in requirements or changes in estimates.

### Cost Estimate Range and Risk Drivers Challenge:

All but one of the lines shown in Table 1 are available within the OSD Comptrollers data systems and ADVANA. The exception is the required lines. As outlined in the DoD Cost Guide, these lines represent the “Latest estimate of funds required to successfully execute a program, e.g., support the Warfighter and note simple math available budget TOAs. Typically, this would reflect the Will Cost estimate, CCP, or POE<sup>12</sup> that has not yet been validated by a component cost agency or the CAPE.” These estimates are not recorded in any database within OSD, nor the DoD components. The line typically represents to the management team, whether a program, MDA, or Service position, the funding requirements based on one of the cost estimates noted. A cost estimate by its nature would not be a point estimate by fiscal year. Plus, the cost estimate would go through phasing, in which the cost estimate is allocated across the fiscal years to ensure adequate budget authority. Depending on the appropriation, the phasing would be different, which RDT&E incrementally funded, and procurement fully funded. How sensitivity analysis and risks or opportunities, and uncertainty were addressed could also affect phasing. The research team held discussions with several current OSD portfolio managers across several capabilities. The common goal was to “assure the component was robustly funding the program.” This is hard to do when the level of confidence in cost and schedule are not documented in a manner that the data is readily available.

The data that characterized the estimate is critical to understanding the uncertainty in the program. “Without a risk and uncertainty analysis, the program estimate will not reflect the degree of uncertainty, and a level of confidence cannot be given about the estimate. Unless a range of costs is provided, decision-makers will lack information on cost, schedule, and technical risks, and will not have insight into the likelihood of executing the program within the cost estimate.” It goes on to note that “without an S curve, decision-makers will lack insight of what the likelihood of different funding alternative imply about program success” (GAO, 2020). The DoD Cost Guide provides a suggested S-Curve (Figure 2), as well as other formats for characterizing range. These practices are industry best practices, documented in both ANSI and ISO standards with clear characterizations of both qualitative and quantitative risks along with the designation of contingent and management reserves (PMI, 2021).

---

<sup>12</sup> CCP is component cost positions; POE is program office estimate.



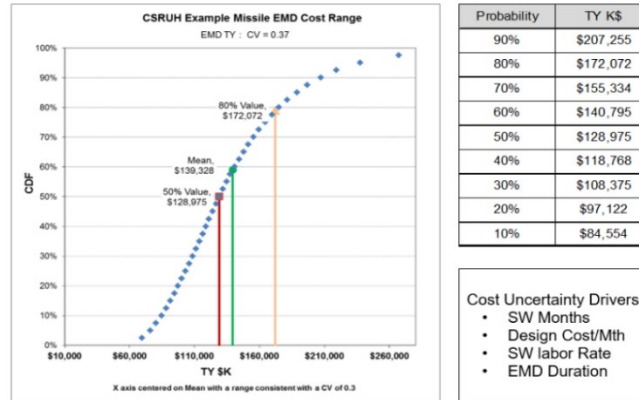


Figure 2. S-Curve example (notional)

The DoD Capability Portfolio Management goal is stated “to optimize capability investments across the defense enterprise (both materiel and non-materiel) and minimize risk in meeting the Department’s capability needs in support of strategy” (DOD, 2017). It is not possible to achieve with the Spruill Chart and the limited availability of data that characterizes the challenges at the program level and the aggregate at the portfolio levels.

### Network Schedule Challenge?

The second challenge is the lack of program or PEO level network schedules, or Integrated Master Schedules. Scheduling information, which includes dates for objectives and thresholds on key milestones, is usually submitted in PDF Gantt chart form with budget documentation or briefing slides attached to PMRT or other reporting tools. However, it typically does not include network schedules that model the program.

The DoD requires very detailed contractor schedules on programs implementing Earned Value Management, with data being submitted per data items descriptions (DID). However, our cursory review and discussion with OSD Portfolio Managers did not expose any network schedule models. One OSD portfolio manager understanding the value of such models was producing a network schedule gathering data from various sources on their own.

Without a schedule model, it is almost impossible to do a comprehensive quantitative risk analysis on schedule. OSD Cost Estimating guide calls for such analysis, but DOD 5000 series of directives and instructions do not specifically require it. This is counter to industry standards and the evolution within the industry. Within the construction and other industries, “Digital twin’ technology and planning and scheduling are integrated to form a planning and scheduling system based on digital twin” (Wang 2020). A similar approach is needed to address the complex scheduling network within DoD portfolios.

### What Is in the Portfolio Challenge?

The third challenge is how a particular portfolio is aggregated. Of the three proposed multidimensional portfolios, Products/Platform, Capability, or Mission, none of them have a clear characterization of the portfolio in any reference schema to allow easy aggregation. The programs/platforms portfolio has focused on Major Acquisition Defense Program (MDAP), using a program number (PNO). OSD is in the process of establishing a PNO, which is three-digit alpha-numeric, for all baselined programs at all levels. The PNO number is also used in budget documentation. However, this is not likely to solve the needs within the multidimensional portfolio structure.

An example would be tracking variants with alternative configurations as the PNO





designation is focused on a program, not the product or platform. The PNO designation is focused on how the DoD baselines and funds an effort. An example of this is GCSS Army has five entries, PNO N03, H41, 347, 402, and 501. PNO 501 is the current effort, GCSS-A Increment 2. In order to see the portfolio view of the capability, the capability manager has to look across multiple PNOs. Another example is the B-52. The B-52 is in DoD Acquisitions Visualization Environment (DAVE) over 25 times with numerous PNOs. This is because the platform has been modified via separate programs, meaning baselined with separate funding, numerous times over the decades. This makes data retrieval and analysis prone to error, as the relationships of the PNOs are not easily traceable within the existing data systems.

For capability portfolios, the CJCS Functional Capability Boards (FCB) under JCIDS, which manage the capability portfolios in DoDD 7045.20, poorly align with USD (A&E) capability breakouts. We could not find a Product/Platform to JCA alignment within the DoD data schemas. Missions Threads aggregated into any type of portfolio structure could also not be found. No comprehensive schema to align across the various D2S2 systems exists.

Industry uses projects as subsets of programs or portfolios. This three-tiered structure would be helpful in further breaking down programs within the DOD. The project term is used in budget documents, but that use is not aligned with any formal baselining of projects under the formalized baselined programs. The Portfolio to Program to Project is a governance breakdown structure (GBS), which is different from the Work Breakdown Structure. The DoD product is a system, like a jet engine, and the platform systems, the fighter aircraft. If F-16 is the platform, a system of systems, then the engine as a system was managed in a different program than the F-16, but relative to the F-16 is in the platform WBS. In the F-35, the engine is within the same programs, but the same governance structure.

Within the project/program management profession, a standard work breakdown structure is key for data collection and integration across not just engineering, but also cost, schedule, risks, and overall programmatic data. The DoD has had a standard WBS structure since 1968. MIL-STD-881 at one point was made a handbook in 1998 to reduce military standards. The handbook version moved back to a standard in 2011, as it became clear the flexibility of the handbook was not providing the appropriate level of standardization.<sup>13</sup>

### **Multidimensional System of Systems Challenge**

The only mention of system of systems (SoS) in DoDD 5000.01 is related to capability portfolio management, mission engineering, and integration analysis using an effect/kill chain framework that employs the **integration and interoperability** of the SoS required to execute crucial mission requirements. Integration and interoperability are bolded to remind ourselves of the new OSD initiative on Acquisition Interoperability and Integration (AI2). It is not about platform SoS but mission SoS. It is not mentioned in DoDI 5000.02 or DoDI 5000.88 on Engineering of Defense Systems. But it does show up in the OSD Mission Engineering Guide relative again to warfighter **integration and interoperability** of SoS.

We are using system and SoS in the broad sense. Industry defines SoS as a “Set of systems or system elements that interact to provide a unique capability that none of the constituent systems can accomplish on its own. Note: Systems elements can be necessary to facilitate the interaction of the constituent systems in the system of systems. Constituent systems can be part of one or more SoS. Note: Each constituent is a useful system by itself, having its own development, management goals and resources, but interacts within the SoS to provide the unique capability of the SoS” (Henshaw et al., 2023). The SEBoK noting the seminal work of Dr. Mark Maier (1998) postulated five key characteristics (not criteria) of SoS, noting

---

<sup>13</sup> Per discussion with Neil Albert, March 16, 2023 with John Driessnack



operational independence and managerial independence as the two principal distinguishing characteristics of SoS.

It should be useful to combine the management implications of portfolio management (Pfm) with the technical and capability implications of system of systems (SoS). Understanding in both cases, there can be sub-portfolios within portfolios and sub-systems or system of systems within a system of systems. For our goal of improving enterprise decisions relative to resources at the DMAG level, the complicated structure is broken into three sets of portfolios:

**Portfolio of products/platforms for the purpose of life-cycle management of those products/platforms.** This is the traditional System Program Manager (SPM) who works for a PEO. The SPM in many cases has a portfolio of systems that fit within a larger portfolio managed by the PEO. We need a governance breakdown structure (GBS) that manages products/platforms in which we use a work breakdown structure (WBS). The significant interchange between the GBS and WBS as products used in various platforms are managed under various governance schemas, which are not consistent. The schema creates its own system of systems.

**Portfolio of operational unit capabilities<sup>14</sup> for the purpose of managing the requirements relative to a family of similar products/platforms.** This is traditionally the component requirements officer who works with an overall military capability planning organization. Products do not have operational capability; this should refer to the military unit, the fighter squadron, not just the fighter platform. Most if not all of DOTmLPF-P structure needs to be considered. In this portfolio, we suggest the structure should follow how the DoD components are structured by operational units.

**Portfolio of combatant missions for the purpose of managing the missions within a combatant command's (CCMD) or combat support agency operations plans.** Here, the capable DoD component operational units are placed into a combatant or support unit structure to perform missions under an operational plan. The structure could follow OpsPlan structure.

### Operational Unit Capability Structure Challenge

The operational unit capability challenge was identified when it became clear that the capability portfolios within USD(A&S) and those within the CJCS organization did not align. Table 2 describes groupings of related capabilities that support strategic decision-making and capability portfolio management, including joint analyses of capability gaps, excesses, and major trade-off opportunities. The challenge is capability is defined by CJCS as “the ability to complete a task or execute a course of action under specified conditions and level of performance. This can be achieved through a combination of means and ways across doctrine, organization, training, leadership and education, materiel, personnel, facilities, and policy.” The keys in this definition are “TASK” and the reference to “DOTmLPF-P.”<sup>15</sup> There is not an emphasis on “materiel,” but the whole of the DOTmLPF-P. Materiel is defined as “all items necessary to equip, operate, maintain, and support military activities without distinction as to its application for administrative or combat purposes.”

---

<sup>14</sup> We will use operational capability to distinguish from technical capability of the product.

<sup>15</sup> DOTmLPF-P is defined in CJCSI 5123.01, the Charter of the JROC and Implementation of the JCIDS as Joint Doctrine, Organization, Training, materiel, Leadership and Education, Personnel, Facilities, and Policy (DOTmLPF-P). The instruction defines the Functional Process Owner (FPO) for each of the DOTmLPF-P process. For “materiel,” J-8 Force Structure, Resource & Assessment Directorate is the FPO and manages the overall JCIDS process.



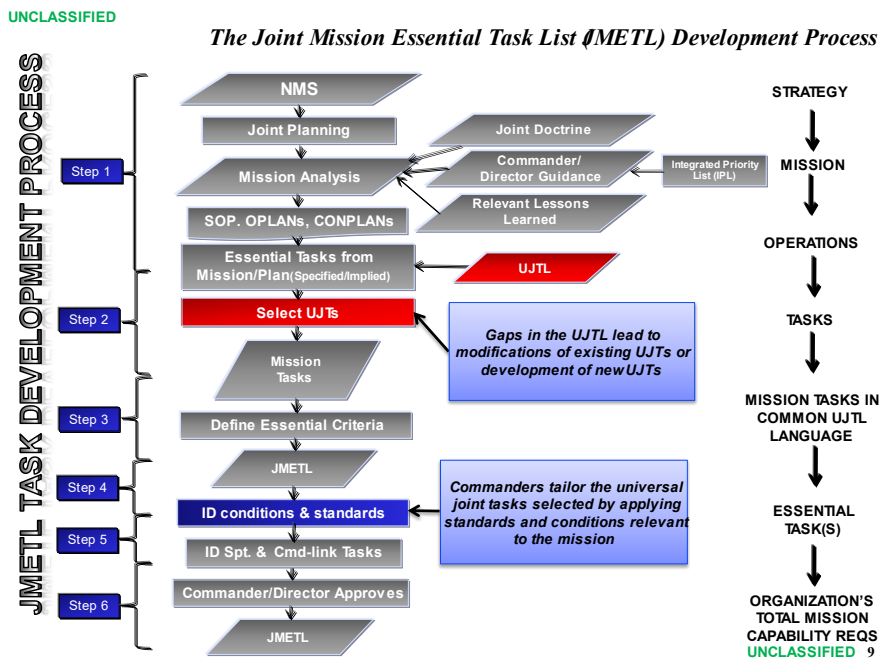
**Table 2 Tier 1 JCAs POC's**

<b>Tier 1 JCA(s)</b>	<b>Organization</b>
Force Integration	FI FCB
Battlespace Awareness (BA)	BA FCB
Force Application (FA)	FA FCB
Logistics (LOG)	LOG FCB
Command and Control	C4/Cyber FCB
Communications and Computers	C4/Cyber FCB
Protection	Protection FCB
Corporate Management and Support	Pending DEPSECDEF Assignment

The CJCS method to track tasks is through the Universal Joint Task List (UJTL), which is the authoritative common language for all approved joint tasks required for planning, readiness reporting, training and exercises, lessons learner processing, and requirements. “A universal joint task (UJT) is an action or activity assigned to a unit or organization to perform a specific function and/or provide a capability or resource. UJTs are based on extant joint capabilities, and they have a foundation in approved joint doctrine. Specifically, UJTs describe “what” joint organizations must do using common and joint terminology” (CJCS, 2022).

It appears as if the UJTL is to missions as a product-based WBS is to the product/platform. Both track capability, one helps with operational, the other with technical. JCAs are a management structure for the CJCS minimally aligned with the field, similar to the Capability Portfolio structure within OSD minimally aligned with the PEO/SPM structure in the acquisition community. This weak alignment inhibits any reasonable mapping of data from the governance structures to the actual efforts.

A Joint mission-essential task (JMET) is a mission task selected by a joint force commander deemed essential to mission accomplishment and defined using the common language of the Universal Joint Task List in terms of task, condition, and standard. See also condition; Universal Joint Task List. Source: JP 3-33. The UJTL is a key schema that could be used to map the capabilities of operational units with missions (see Figure 3).



**Figure 3 JMETL Development Process**

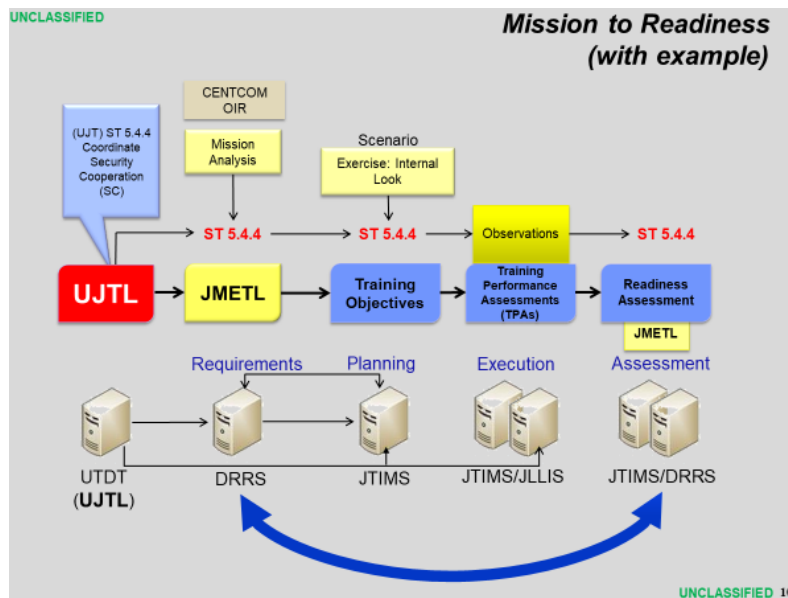


Going back to the F-16 example, in UJTL we could use OP 6.1.4 Conduct Defensive Counterair (DCA). The task is to “conduct defensive measures designed to neutralize or destroy enemy forces attempting to penetrate or attack through friendly airspace. JP 3-01.” The UJTL has six defined measures (see Table 3) for which to assess the capability. Within the “mission to readiness” example (see Figure 4), the UJTL sets up the structure not only for the JMETL, but also the training and readiness assessment. The UJTL then forms the structure for the Defense Readiness Reporting Systems (DRRS), which could provide meaningful assessment data.

**Table 3 UJTL Capability Assessment Measures**

Measures:		
M1	Minutes	To notify friendly counterair forces (to gain intercept position).
M2	Percent	Of joint security area (JSA) and joint operational area (JOA) in which friendly freedom of movement allowed.
M3	Percent	Of enemy air attacks detected early enough to allow engagement.
M4	Percent	Of enemy air defense targets successfully engaged.
M5	Percent	Of enemy aircraft penetrate air defenses.
M6	Percent	Of first-shot kills by friendly fighters in air-to-air combat.

The scope of the Phase 1 research topics helped identify five key challenges, which need to be addressed to complete Phase 1 and continue with Phase 2 and 3. To model across the product/platform we need a structure, like the UJTL, to create a comprehensive architecture for missions. The architecture requirements are to provide an aligned set of structures that will support the individual decision system data systems as well as the enterprise data systems within the D2S2 to enhance effective and efficient decision analytics.



**Figure 5 UJTL Common Thread, Mission to Readiness Example**



Next, we will create a notional example model of a multidimensional (not multilayers, as it is not a hierarchical challenge) structure of portfolios of products/platforms, and operational units, which are assigned to combatant and support commanders. The goal of the structure will be to represent the challenges across the D2S2 by creating a notional data set within each structure. The structure will allow the creation of models with enough standardization to allow useful information to flow from the lowest levels up to the Deputy's Management Action Group (DMAG) level. Additionally, the structures must have enough flexibility to be useful to the various level of managers within the governance structures. While much of the current structure already used within each domain will be considered, the research team anticipates changes will be needed to allow alignment across the portfolios.

Assessment within a Capability or Mission Portfolio will not be possible until the enterprise works off an aligned structure. In phase 2, the use of UJTL will be explored for capability along with a unit organizational structure. Hopefully, it can also be used as a common structure for mission thread assessment. We will need to consider the structure of the regional and functional combatant commanders along with supporting commanders and how Operational Plans are structured under CJCS policy.

The resulting multidimensional portfolio structure could be documented in a revised DoDD 7045.20, renamed the D2S2 Enterprise Portfolio Structure and Management. Today there is no DoD directive or instruction for the overall enterprise DoD Decision Support Systems. As IAPRs focus on Integration and Interoperability, it is not just an acquisition goal but should cut across the enterprise, which should be the value proposition of the Enterprise Portfolio. Under the current structure, the DEPSECDEF supported by the DMAG would be the Enterprise Portfolio Management team.

The Department of the Air Force (DAF) in preparation for the FY24 budget has taken a step in this direction under the Operational Imperatives initiative. The Operational Imperatives, breaking partly from the traditional PEO approach, grouped specific efforts, many of them programs of record, into operational capabilities focused grouping aligned on pacing challenges. The operational imperatives are aligned with the Joint Warfighter Concept, which appears, in some case, to have driven the groups away from traditional PEO buckets. The creation of a structure based on a strategy to achieve specific operational objectives (Figure 5) within and across the seven Operational Imperatives is an example of an enterprise approach (USAF 2023).

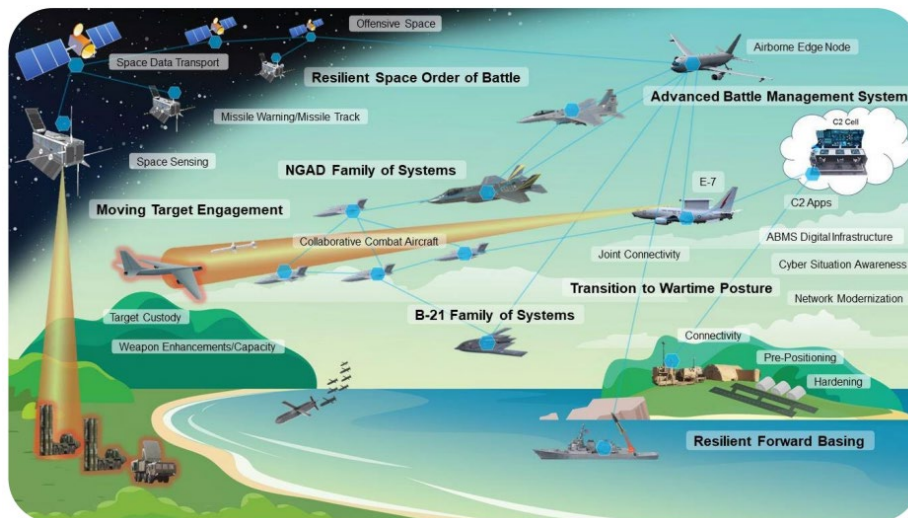


Figure 5 Operational Imperatives



The concept of portfolio management within the department needs to embrace the seven performance domains within the ANSI Standard for Portfolio Management, especially strategic and value management, so clear objectives for these sub-portfolios can be determined, and thus a performance management structure can be established to drive the appropriate measure that will allow data-driven management to those objectives. Moving to a multi (many) dimensions (measure in one direction) view of portfolio management under an enterprise portfolio structure for D2S2 decision-making will allow the DoD organizationally to form a structured network of teams with clear empowerment, which embraces John Kotter's Accelerate concept of a second system within a company that is organized in a network, which has shown a proven approach to accelerate strategic agility and strategic execution in a faster-moving world (Kotter, 2014). This would allow the enterprise to move to a network-centric management approach for decision-making on a network-centric JWC.

## **Notional Enterprise Decision Support Structure and Models**

It might seem unreasonable to do an enterprise-level decision support model. Too complicated with too many stakeholders, thus too complex. But with the creation of the Defense's Chief Digital and Artificial Intelligence Office (CDAO) and the creation of ADVANA, a successful approach is more likely. Key will be conceptually to take enterprise and portfolio level Analysis of Alternatives (AoAs). In defense acquisitions, AoA is an "assessment of potential material solutions to satisfy the capability need to be documented in the approved Initial Capabilities Document (ICD). The AoA focuses on the identification and assessment of potential materiel solutions, key trades between cost and capability, total life-cycle cost, including sustainment, schedule, concepts of operations, and overall risk." AoA typically leverages available data and documents "sufficient quality to support investment and acquisition decisions. ... Common or 'wash costs" (DOD, 2022). The key is paying attention to what are the differences in the alternatives.

Enterprise AoAs would take advantage of AoAs that are more focused on the acquisition level or organizational capabilities or combatant missions but would look beyond the individual system's decisions that have been set for a particular requirement and move to a more enterprise view. To accomplish this, a model structure will be needed to look across various PEOs (across DoD components) with various types of system program managers of various materiel/technology solutions (platform, product, sub-product, commercial, software, material/commodity, etc.), which are assigned as assets to component operational units. Those component operational units are assigned to various combatant components (regional, functional, supporting). In Phase 2, the team's goal is to develop an example model using notional data (unclassified) as a tool to demonstrate further possible decision analytics within and across the product/platform, operational capability, or mission portfolios as well as at the enterprise portfolio level.

## **Acknowledgements**

OSD Sponsor Dr. Brian B. Joseph, Deputy Director Data Analytics

OSD Portfolio Managers who participated in discussions and research reviews, including David Crim, Lisa Didden, Thomas Gehrki, Col Scott Helmore

ADVANA support from Patrick Lobner (BAH)

UMD Research Team at the Clark School of Engineering, Civil Engineering Department, Project Management Center of Excellence. The team includes Professor Gregory Baecher, Professor Qingbin Cui, John Johnson (Professional Program Manager PM COE), and Theodoxea Kwapong (Project Management Graduate Student).

We also appreciate the support from our host, Catalyst Campus for Technology & Innovation at their PEO IWS X Forge Software Factory facility.



## References

- Ahern, D., & Driessnack, J. (2019, June 30). *Implementing a multitier portfolio management structure for defense acquisition*. Olde Stone Consulting. <http://www.oldestoneconsulting.com/blog>
- AIRC. (2022). *Data-driven capability portfolio management pilot, technical report* (AIRC-2022-TR-007). Stevens Institute of Technology.
- Cebrowski, A. K., & Garstka, J. H. (1998). Network-centric warfare—Its origin and future. *Naval Institute Proceedings*, 124(1), 1139.
- CJCS. (2022). *Universal joint task list program*. Department of Defense. [https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/cjcsi\\_3500\\_02c.pdf?ver=xeYyY0JZpGbKjxlvaqY4kA%3d%3d](https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/cjcsi_3500_02c.pdf?ver=xeYyY0JZpGbKjxlvaqY4kA%3d%3d)
- Davis, P. K. (2002). *Analytic architecture for capabilities-based planning, mission-system analysis, and transformation*. RAND Corporation.
- Davis, P. K., Shaver, R. D., & Beck, J. (2008). *Portfolio-analysis methods for assessing capability options*. RAND Corporation.
- DoD. (2017). *Capability portfolio management* (DoDD 7045.20).
- DoD. (2019). *DoD directive: Capability portfolio management*. <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/704520p.pdf>
- DoD. (2022, January 22). *Analysis of alternatives cost estimating handbook*. <https://www.cape.osd.mil/files/Reports/AoACostHandbook2021.pdf>
- DoD. (2023). *Platform and weapon portfolio management*. Acquisition: Office of the Assistant Secretary of Defense. <https://www.acq.osd.mil/asda/pwpm/index.html>
- DTIC. (2019). *Section 809 panel*. <https://discover.dtic.mil/section-809-panel/>
- GAO. (2007). *Best practices: An integrated portfolio management approach to weapon system investments could improve DOD's acquisition outcomes* (GAO-07-388). <https://www.gao.gov/products/gao-07-388>
- GAO. (2015a). *Weapon system acquisitions: Opportunities exist to improve the Department of Defense's portfolio management* (GAO-15-466). <https://www.gao.gov/products/gao-15-466>
- GAO. (2015b). *Weapon systems acquisitions* (GAO-15-466).
- GAO. (2020). *Cost estimating and assessment guide: Best practices for developing and managing program costs* (GAO-20-195G).
- Gibeling, Keith E, Universal Joint Task List (UJTL) 101, May 12, 2022, Joint Staff J7, CJCS website, February 25, 2023.
- Henshaw, M., Jahmann, J., & Lawson, B. (2023). Systems of systems (SoS): Summary of ISO/IEC/IEEE 21839. In *SEBoK: Guide to the systems engineering body of knowledge*. Stevens Institute of Technology. [https://sebokwiki.org/wiki/Systems\\_of\\_Systems\\_\(SoS\)](https://sebokwiki.org/wiki/Systems_of_Systems_(SoS))
- Kenney, C., & Driessnack, J. (2023a). *Portfolio management with the Department of Defense: A data challenge*. Proceedings of the IISE Annual Conference & Expo 2023, May 2023.
- Kenney, C., & Kwapong, T. (2023b). *Capability, mission, and PEO (CMP) portfolio performance analysis and visualization, MVP A White Paper - February 2023* [unpublished working paper].
- Kotter, J. P. (2014). *Accelerate: Building strategic agility for a faster-moving world*. Harvard Business Review Press.
- Project Management Institute. (2021). *A guide to the project management body of knowledge, ANSI 899-001* (7th ed.).
- UARC. (2021). *Annual Report FY 2021* (Extramural Acquisition Innovation and Research Activities). Stevens Institute of Technology.
- USAF. (2023). *F24 PB rollout brief*. <https://www.saffm.hq.af.mil/FM-Resources/Budget/Air-Force-Presidents-Budget-FY24/>
- Wang, Y., & Wu, Z. (2020). Model construction of planning and scheduling system based on digital twin. *International Journal of Advanced Manufacturing Technology*, 109, 2189–2203. <https://doi.org/10.1007/s00170-020-05779-9>
- Woolsey, J. P. (2018). *Spruill charts a way forward*. DAU News. <https://www.dau.edu/library/defense-atl/blog/Spruill-Charts-a-Way-Forward>
- WRT. (2020). *Foundations for a model-based portfolio analysis capability* (Technical Report SERC-2020-TR-013). Joint Program Executive Office for Chemical, Biological, Radiological and Nuclear Defense (JPEO-CBRND)



## Joint All-Domain Command and Control (JADC2) Opportunities on the Horizon

**Roshanak Rose Nilchiani**—received her PhD in space systems engineering from the Massachusetts Institute of Technology (MIT) in 2005 and currently is a tenured Associate Professor at Stevens Institute of Technology. In the past two decades, her research has focused on modeling and quantifying complexity, and complex engineered and social systems' response to change. Her most recent research explores the relationship between system complexity, uncertainty, emergence, and risk with an emphasis on developing new quantitative models and frameworks for risk management. Her research group is working on formulating and modeling quantitative measures of complexity and increased entropy in complex networks and engineered systems such as social network resilience and complexity, network growth and collapse, technical risk assessment of acquisition programs, and complexity modeling in various lifecycle phases of space systems and programs. [rnilchia@stevens.edu]

**Dinesh Verma**—received a PhD (1994) and an M.S. (1991) in industrial and systems engineering from Virginia Tech, Blacksburg, VA. He served as the Founding Dean of the School of Systems and Enterprises at Stevens Institute of Technology from 2007 through 2016. He currently serves as the Executive Director of the Systems Engineering Research Center (SERC), a U.S. Department of Defense sponsored University Affiliated Research Center (UARC) focused on systems engineering research, along with the Acquisition Innovation Research Center (AIRC). Verma has authored over 150 technical papers, book reviews, technical monographs, and books. In addition to his publications, Verma has received three patents in the areas of life-cycle costing and fuzzy logic techniques for evaluating design concepts. He was recognized with an honorary doctorate degree (Honoris Causa) in technology and design from Linnaeus University (Sweden) in January 2007 and with an honorary Master of Engineering degree (Honoris Causa) from Stevens Institute of Technology in September 2008. [dverma@stevens.edu]

**Philip S. Antón**—is the Chief Scientist of the Acquisition Innovation Research Center (AIRC) at the Stevens Institute of Technology. As a Pentagon interface between the AIRC and the Department of Defense (DoD) acquisition ecosystem, he assesses the practical needs of the DoD, helps to envision and develop innovative acquisition research in the AIRC, and ensures the transition and application of AIRC results in DoD acquisition policies, guidance, practices, reports, and workforce development. Prior to this, Antón was a Senior Information Scientist at the RAND Corporation for 23 years, where he conducted research on acquisition and sustainment policy, cybersecurity, emerging technologies, technology foresight, process performance measurement and efficiency, aeronautics test infrastructure, and military modeling and simulation. Antón earned his PhD and MS in information and computer science from the University of California at Irvine, specializing in computational neuroscience and artificial intelligence. He holds a BS in engineering from UCLA, specializing in computer engineering. [panton@stevens.edu]

### Abstract

Joint All-Domain Command and Control (JADC2) is an enormous effort in information sharing—sense, make sense, and act—to empower joint force commanders in warfighting. This effort will take advantage of materiel and non-materiel solutions as well as modify existing policies, authorities, organizational constructs, and operational procedures. The goal of JADC2 is to empower the U.S. military to join forces to seize, maintain, protect, gain information and knowledge, and maintain decision advantage and superiority. There are several challenges and questions raised by experts in the DoD including but not limited to: the need for the portfolio management of JADC2-related efforts, the decision-making authority structure within JADC2, affordability and specific budget allocation, and technical maturity of the proposed technologies as well as optimal technical system design and lifecycle management. This paper looks at JADC2 through an academic/scientific lens to identify multiple opportunities in which academic institutions in various domains (engineering, sciences, and social sciences) can contribute to creating a state-of-the-art, Joint All-Domain Command and Control system.

**Keywords:** Joint All-Domain Command and Control, JADC2, Systems View, MBSE





## Introduction

The U.S. military operates in an ever-changing operational landscape, requiring quick adaptation to shifting circumstances. In such a dynamic environment, achieving and maintaining information superiority is of utmost importance. To this end, the Department of Defense (DoD) has established Joint All-Domain Command and Control (JADC2), an initiative, and concept aimed at improving Joint Force C2 capabilities (Hoehn, 2022). However, due to the significant diversity among the various sectors and departments within the DoD, the development and implementation of JADC2 require considerable effort to consider the distinct needs and perspectives of all stakeholders and agencies involved.

To guide and oversee the development and implementation of JADC2, a cross-functional team has been created which will work in collaboration with a Deputy Secretary of Defense–related staff that is comprised of Senior Executive Service (SES)–level members from various agencies, for example, the DoD, Office of the Secretary of Defense, Defense Advanced Research Projects Agency, Air Force, Army, and Navy (Hoehn, 2022). The main objective and focus of this team are to identify and implement command and control improvements in the form of an implementation plan.

This paper provides an overview of the current state of the JADC2 initiative, provides a set of suggestions, and identifies several opportunities to solve and improve some of the key challenges of JADC2 in multiple domains of technical, organizational, and data enterprise. This paper begins by providing an introduction and overview of the significant challenges pertaining to jointness and JADC2. It provides a brief overview of JADC2 history followed by a general conceptual overview of JADC2. The next section of the paper provides a set of technical and conceptual solutions and directions needed for research, development, and acquisitions of the technologies that would enable the DoD to achieve a resilient and elegant advanced solution to JADC2.

## History and Progress

Before the JADC2 initiative, distinctive command and control systems (C2) were owned and operated by each force independently (Hoehn, 2022; McInnis, 2021; Theohary, 2021; Woolf, 2021). Historically, each military service has developed and acquired its own unique tactical command and control network, often incompatible across weapons systems, platforms, and operating domains. As a result, decision time cycles and the transmission of critical time-sensitive data for decision making were slow, redundant, and organizationally stove-piped (*Advanced Battle Management System*, 2022) and domains of air, land, sea, space, and cyberspace were treated separately (DoD, 2022; Feickert, 2022; O'Rourke, 2021; Theohary, 2021) in addition to geographically separated command units (*Advanced Battle Management System*, 2022; DoD, 2022; *Doctrine for the Armed Forces of the United States*, 2013). While multiple command and control systems owned by different forces enabled highly specialized and effective solutions to be developed and implemented, it also required significant efforts on all fronts and limited the threat reaction capabilities and information sharing between all forces (Feickert, 2022).

The legacy C2 systems come with the disadvantage of potential susceptibility to adversaries' anti-access and denial attacks. The adversaries' anti-access/area denial (A2/AD) tactics, including electronic warfare, cyber weapons, long-range missiles, advanced air defenses, and GPS denial, can affect our operational ability and decision cycle that relies on sensors and technologies (*Advanced Battle Management System*, 2022; Friedman, 2019; *Joint Doctrine Publication 5 Command and Control*, 2012; Kreisher, 2001). In addition, current threats are not limited to individual domains anymore, which makes it difficult to counter with dedicated and partially isolated solutions. Consequently, DoD leaders have expressed the need to expand



access to information in an extensive approach to increase overall agility and preparedness for contingencies from different directions (*Doctrine for the Armed Forces of the United States*, 2013; *Jointness - A Selected Bibliography*, 1993; Kirtland, n.d.; *Transforming the Joint Force*, 2003; Woolf, 2021).

The JADC2 initiative and the proposed shared infrastructure would reinforce and enhance the effectiveness of all armed forces and services. Such a shared foundation allows for simultaneous and consecutive operations, as well as continuous integration of capabilities across all domains. In recent years, major efforts have been undertaken to join specific areas of operation and exploit the advantages of combined information and technology, such as the AirLand Battle concept (Kirtland, n.d.), DARPA's Mosaic Warfare program, the Air Force Advanced Battle Management System (ABMS), the Navy's project Overmatch, and the Army's Project Convergence (Congressional Research Service, 2021b). Jointness efforts have also been reported in various forms in other countries, such as the Netherlands and India (Birch et al., 2020; Congressional Research Service, 2022; Nardulli et al., 2003). Additionally, JADC2 tests were conducted in 2019 and 2020 (McInnis, 2021). Due to the disproportionate increase in complexity, growing connectedness of networks of sensors, and novel and sophisticated joint technologies exceeding human cognitive capabilities, no particular solution has been widely implemented as of today.

Several challenges have been identified as follows:

- More approval steps are required to integrate multiple domains (Builder et al., 1999)
- Planners have insufficient expertise in or access to information on relevant multi-domain operations (Builder et al., 1999)
- Increased dependence on multi-dimensional operation communication systems (Builder et al., 1999)
- C2 legacy systems incompatibilities
- Presence of a single-domain or service-centric mindset as well as cultural and organizational biases (Builder et al., 1999)
- Integrating multiple domains increases risks to unifying efforts
- Managerial aspects and budget allocation (Alberts & Hayes, 2006)
- Interservice conflicts and competition (Alberts & Hayes, 2006)
- Overlapping organizational structures (Hoehn, 2022)

Such challenges not only affect the technical or cultural feasibility of JADC2 but also pose congressional challenges to budgeting and funding this major effort (Congressional Research Service, 2021b). The JADC2 program will address and respond to these challenges.

### **Joint All Domain Command and Control Concept and Framework**

The JADC2 envisions Joint Force command and control capabilities for the future. It aims to establish a warfighting capability that can effectively sense, interpret, and respond at all levels and phases of the war, across all domains, and in collaboration with partners. The ultimate goal is to provide information advantage with unprecedented speed and relevance (Alberts & Hayes, 2003; Kirtland, n.d.). The JADC2 strategy employs a System-of-Systems approach, which integrates various capabilities, platforms, and systems, and is aimed at accelerating the implementation of necessary technological advancement and doctrinal change in the Joint Force C2. JADC2 will enable the Joint Force to use vast volumes of data and convert them to information and knowledge, employ automation and AI, utilize a secure and resilient, and adaptable infrastructure, and act inside an adversary's decision cycle (Builder et al., 1999). To address these efforts, an implementation plan has been developed and a team appointed to oversee the process. This team consists of cross-functional SES-level members



from the areas of Combatant Command together with Services, Defense Agencies, as well as Joint and OSD staff.

### Sense, Make Sense, and Act

**Sense:** To ensure the usability and usefulness of joint data for all forces/services and Joint Force Commander, a common and shared sensing methodology and information management technologies are required. The approach requires that information collection and provision in an operational environment can be conducted and delivered to the receiving. JADC2 implements a novel data-sharing approach in combination with advanced information management technologies. These networks are created based on federated data “fabrics” and enable the Joint Forces to achieve information that can be used for decision-making. Through sensing and integration, it is possible to “discover, collect, correlate, aggregate, process, and exploit data from all domains and sources (friendly, adversary, and neutral)” and “share the information as the basis for understanding and decision-making” (Kirtland, n.d.).

**Make Sense:** The process of making sense involves analyzing, understanding, and predicting the operational environment as well as the adversary and friendly force actions. In this phase, data is transformed into information, and information churns into knowledge. Making sense requires the ability to fuse, analyze, and render validated information from all domains and the electromagnetic spectrum. One major requirement in this phase is to provide secure as well as accessible information execution. The capabilities developed by JADC2 will leverage Artificial Intelligence and Machine Learning (ML) to accelerate the joint force commander’s decision cycle (Builder et al., 1999; Kirtland, n.d.). The technical and procedural advancements will also significantly enhance the Joint Force’s ability to operate in a C2 degraded environment.

**Act:** To “Act” is to make and disseminate decisions to the Joint Force and its mission partners. This phase combines the human elements of decision-making with the technical means to perceive, understand, and predict the actions and intentions of adversaries, and take action. This step includes decision analysis, conveying the decision, and the execution phase. Novel decision support applications will be implemented between Joint Forces through advanced, resilient, and redundant communication systems, an accessible and comprehensive transport infrastructure, and flexible data formats to enable the rapid, accurate, and secure dissemination of decisions. “Act” also means providing the Joint Forces with proper training. Using a Mission Command approach, subordinate commanders are empowered to act with confidence and authority through understanding a senior commander’s operational intent while retaining the ability to act when communications linkages are broken or when the urgency of operations precludes the time necessary to seek guidance. Mission Command provides the Joint Force the agility and trust needed to seize the initiative and maintain information and decision advantage (Kirtland, n.d.).

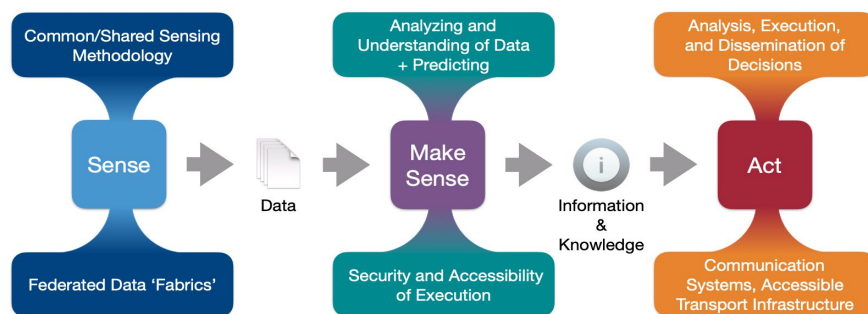


Figure 1. JADC2 Action Chain and Process



## Lines of Effort (LOEs)

The JADC2 strategy is organized around five LOEs to guide Department actions in delivering capabilities, including data enterprise, human enterprise, technology enterprise, integrating with nuclear C2 and C3, and modernizing mission partners' information sharing. Each LOE is guided by an Office of Primary Responsibility represented by senior Flag/SES persons that can raise issues, interact with, and support the Joint Requirements Oversight Committee through its Joint Capability Board (Congressional Research Service, 2021a).

**LOE 1: Establish Data Enterprise** - The first LOE addresses the data structures and infrastructure. As a strategic asset, data must be effectively managed by JADC2 to enable it to seize, maintain, and protect information and decision advantage. To accelerate the decision-making process, joint forces must be able to discover and access any data and information from all warfighting domains at all levels of warfare. The following key data standardization objectives have been identified as critical to JADC2:

- Establishment of minimum metadata tagging criteria
- Adoption and use of standardized data interfaces
- Implementation of common data availability and access practices
- Incorporation of data security best practices
- Establishment of JADC2 conformant Information Technology (IT) standards
- Continued application of data strategic objectives (Visible, Accessible, Understandable, Linked, Trustworthy, Interoperable, Secure). (Kirtland, n.d.)

**LOE 2: Establish the JADC2 Human Enterprise** - The second LOE addresses the human and organizational performance in command-and-control capabilities using innovative tools such as Artificial intelligence and Machine Learning. This LOE is also tasked with reforming, realigning, or creating organizations with the structure, agility, and resources to more effectively combine the physical and informational strength of the Joint Force and its mission partners such that they are capable of exercising effective control of the Joint Information Advantage (JIA) operations (Kirtland, n.d.). The human enterprise will also address the professional development and training of the leaders as well as guide and support the development of JADC2 aspects of policies, concepts of operation (CONOPS), doctrine, and tactics, techniques, and procedures (TTPs) to optimize the advantages gained through new JADC2 capabilities.

**LOE 3: Establish the JADC2 Technical Enterprise** - The third LOE addresses enhanced shared situational awareness, synchronous and asynchronous global collaboration, strategic and operational joint planning, real-time global force visualization and management, predictive force readiness and logistics, real-time synchronization and integration of kinetic and non-kinetic joint and long-range precision fires, and enhanced abilities to assess Joint Force and mission partner performance (Kirtland, n.d.). The technical enterprise is required to provide secure, worldwide communications networks with sufficient speed and bandwidth to meet warfighting needs. LOE 3 also addresses the transport infrastructure of the JADC2, as well as essential minimum features necessary to ensure continuous C2 capability (communications system resiliency and diversity, multi-level security, elimination of single points of failure).

**LOE 4: Integrate NC2/NC3 with JADC2** - JADC2 will have the capability to collaborate with nuclear C2 communication, and therefore the requirements for NC2 should be considered at the technical and human enterprise level (Kirtland, n.d.).

**LOE 5: Modernize Mission Partner Information Sharing** - The last LOE describes the institutional interoperability needs and organizational architecture for JADC2. The Joint Force Commander will establish and maintain a common understanding of the operational



environment through shared situational awareness with mission partners. Such integration is realized when data from each partner's C2 systems can be accessed, viewed, and acted upon by every other approved partner (Kirtland, n.d.). However, some challenging tasks in this LOE include emerging missions, large coalitions, and evolving technologies that present ongoing obstacles to achieving this goal.

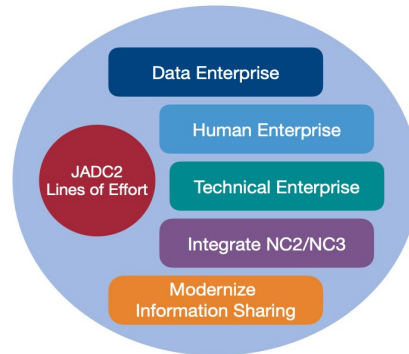


Figure 2. JADC2 Line of Efforts

**Capabilities**

- Connecting all sensors of C2, scaled information sharing
- Network based, cloud-like environment
- Data and interoperability standard driven
- Shared intelligence
- Secure and agile
- Resilient in degraded environment
- Unity of effort in capability development

**Issues raised**

- Technical maturity of the proposed technologies
- Affordability and budget allocation
- Decision making authority across various domains
- Portfolio management needs

**Guiding principles**

1. Information Sharing capability improvements are designed and scaled at the enterprise level
2. Joint Force C2 improvements employ layered security features
3. JADC2 data fabric consists of efficient, evolvable, and broadly applicable common data standards and architectures
4. Joint Force C2 must be resilient in degraded and contested electromagnetic environments
5. Department development and implementation processes must be unified to deliver more effective cross-domain capability options
6. Department development and implementation processes must execute at faster speeds.

Figure 3. Summary of JADC2 Guiding Principles, Capabilities, and Issues Raised

## Systemic and Lifecycle View of JADC2: Opportunities on the Horizon and Required Research

The Joint All-Domain Command and Control (JADC2) concept suggests joining sensors from all military services—Air Force, Army, Marine Corps, Navy, and Space Force—into a single network. Toward achieving this goal, the Department of Defense is pursuing the integration of a few emerging technologies including automation and artificial intelligence, cloud environments, and new communications methods. However, to integrate and infuse multiple new technologies into large legacy System-of-Systems (SoS), a systems and lifecycle approach



is essential to assure a sophisticated, cost-effective, low-risk, and highly capable, unique system that would provide an unparalleled unique set of capabilities to our military services.

There are multiple organizational, technological, sociocultural, and enterprise layers in JADC2 that are in perpetual interactions. The requirements for the JADC2 System of Systems are to integrate legacy systems into novel, disruptive, and cutting-edge technologies that need to be working smoothly together in a highly reliable, efficient, and cost-effective manner. Therefore, the authors propose a systemic approach to identify the opportunities and risks of such a complex system to assure the success of this great endeavor. In this section, the authors propose multiple systemic and lifecycle clusters of opportunities and risks that JADC2 is facing and provide direction of research and solutions for each identified opportunity.

This paper discusses five clusters of opportunities. The first opportunity is the need for novel culturally centered interoperable collaborative mechanisms between services/forces to ensure the formation of best practices in collaboration between the Air Force, Army, Navy, Marine Corps, Space Force, and other departments of defense services. The second opportunity discusses the imminent need for innovation and research in decision science, scenario analysis, and socio-culturally informed game theory modifications. The current game theory application is limited to rational and consistent actors, and the United States often is facing adversaries that are partially rational/or irrational and may have limited consistency in their behavior. The third opportunity discusses the need for complexity management of the growing network of interconnected sensors, decision-makers, and shooters. As the legacy system of sensor networks from all forces are united, the risk of excess network complexity rises and therefore there is an essential need for a resilient architecture for connecting legacy networks. The authors suggest a Universal Translator flexible network of hardware and software to connect all existing and future heterogeneous networks of sensors and assets. The fourth opportunity discusses the need for a novel and strong portfolio management framework for JADC2 Acquisition Programs (to manage, optimize, integrate, and fund JADC2-related projects and acquisition programs). JADC2 consists of multiple acquisition programs at software, hardware, and organizational level that are infused with current legacy and existing systems asynchronously and therefore would require high-level portfolio management to orchestrate multiple projects and tasks over the JADC2 lifecycle. And finally, opportunity 5 discusses the need for requirements and MBSE for JADC2 as an SoS in the following domains: materiel, non-materiel, policies, authorities, organizational constructs, and operational procedures (Nilchiani, 2022).

### **Opportunity 1: Create Novel Culturally Centered Interoperable Collaborative Mechanisms Between Services**

All services and forces in the Department of Defense possess unique cultural and organizational heritage, history, and communication styles, and their assets are composed of legacy systems as well as the latest state-of-the-art in various technologies. One of the JADC2 lines of efforts (LOEs) is composed of human enterprise which involves the human and organizational aspects of the JADC2 implementation. However, the question remains what is the best organizational structure for the most optimal cooperation and collaboration between forces in JADC2? What potential force structure changes will be necessary to meet JADC2 requirements (Congressional Research Service, 2021b)?

The Department of Defense needs a unique one-of-a-kind approach to joining forces that recognizes the individuality and organizational identities of each joining organization, unique traditions, and values across various forces and departments. A successful collaborative solution calls for organic and optimal cooperation of different departments and forces while minimizing interdepartmental conflicts. Such a novel solution would require studies and research based on state-of-the-art organizational research on identity, historical and



anthropological studies of values and traditions of each of the forces, and proposing organic solutions that have emerged from voluntary and mutually agreed-upon collaborations. The JADC2's jointness factors and human enterprise needs to provide a unique organizational solution/blueprint that cannot be solved by technology alone. Figure 4 summarizes the first opportunity and relevant recommendations.

Suggestion: Invest in a unique, long-term, culturally informed solution/organizational blueprint of jointness that has dynamic longevity, versus limited, short-term “solutions” that do not solve core equities, roles, and functions.

Needed Academic Research: Organizational theory, Incentives to motivate jointness, organizational anthropology, and psychology to find the best and unique jointness and collaboration architectures.

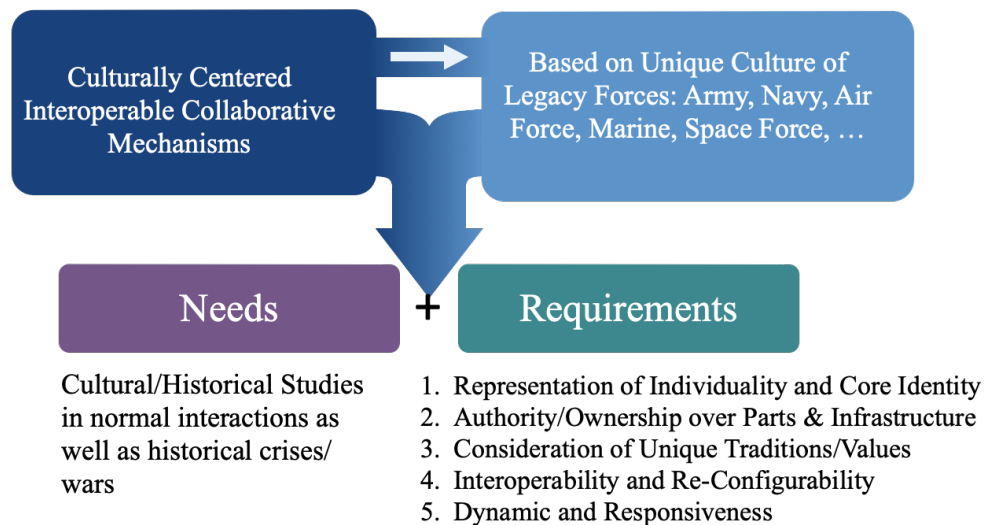


Figure 4. Conceptualization of Needs and Requirements for Organizational Solutions/Blueprints

## Opportunity 2: Need for Innovation in Decision Analysis and Socio-Cultural Game Theory Modifications

At the core of JADC2, there are three actions of sense, making sense, and act on the collected and processed data from the sensors that are interwoven. Sensing and integrating provide the ability to discover, collect, aggregate, and process data from all domains and sources. Then using Machine Learning (ML) and Artificial Intelligence (Builder et al., 1999), the data gets fused, analyzed, and validated. The last step involves a critical decision-making process which is the focus of the suggestion on opportunity 2.

In the Act phase, joint forces engage in making and disseminating decisions to join forces and mission partners. A large portion of the tasks in this phase is to combine the following:

- Human elements of decision making
- Technological means to perceive, understand, and predict the actions and intentions of adversaries and take actions.

Game Theory explains the dynamics of situations where decision makers interact (Priebe et al., 2020) and has been used for decades in decision-making processes. The scientific focus of game theory addresses political, economic, and biological topics and phenomena predominantly (Priebe et al., 2020). The first major advances in game theory were



made by Borel (1927; Alkire et al., 2018) in the 1920s together with von Neumann, who also later published one of the milestone publications in game theory, *Theory of Games and Economic Behavior* (Spirtas, 2018). More recently, game theory in economics has advanced dramatically by two Nobel laureates, John F. Nash (Michael et al., 2017) and John C. Harsanyi (1967).

Within the game theory, models are set up to represent the overall circumstances and dynamics which, four main aspects are defined: first, the decision makers, who are often considered players; second, the strategies and actions that each player/decision maker can choose; third the possible results and outcomes, that are linked to the action and strategic choices of the players; fourth, the payoffs respectively for each player in conjunction with the outcomes/results (Rapoport, 2012). In addition to these aspects, the players and decision makers within the scenarios are considered individually rational, meaning that the judgment of the payoffs in each player's perspective is rational and ordered, in addition to the assumption that each player assumes the other players to be rational (Rapoport, 2012). As a result, the players in the game can factor their knowledge and assumptions about other players into their strategy and can choose accordingly. Game theory allows for logical analysis of interest conflict situations as well as cooperation and therefore defines the theories of rational decision making in conflict situations (Lawlor, 2007).

Yet, the current approaches in decision analysis and game theory fall short of integration and use in JADC2. Game theory assumes rational and consistent actors/adversaries as the basis for strategies and decision analysis suggestions in conflict situations. However, not all actors/adversaries in game theory are "Rational." There is a critical need for novel research in socio-cultural game theory modification. This new science of decision analysis should take into account irrational and inconsistent players among adversaries from different socio-cultural backgrounds and create a modified game theory that strategizes based on new information.

In line with the need for modification of game theory, there is also a need for blueprints/systemic knowledge of adversaries' cultural norms, traditions, and mindsets, such as the underlying cultural norms and strategies presented in Sun Tzu (Bass et al., 2014; *JNT-501S Introduction to Joint Operations: Curriculum*, 2019) and to find the best decision analysis methodologies that take into account cultural differences, values, and approaches. *The Art of War* has been the authoritative military and political guide in the Far East for many centuries and translated and used in the West for the past century. There is a need for academic research to translate the principles of *The Art of War* into abstract rules and heuristics and create a framework that can enable a deep understanding of adversaries' actions and suggest the best strategies in action for JADC2. As an example, the five essentials for victory from Sun Tzu can be interpreted as follows:

- 1) Timing of the fight is essential (suggestions for minimizing engagement and optimizing the timing of decision points)
- 2) the ability to handle superior as well as inferior forces (scalability and ability to engage with adversaries of various scales and capability of forces)
- 3) applying the same operational principles across ranks in forces
- 4) preparation and taking adversaries when unprepared (which will point at surveillance and intelligence and accumulation of patterns and blueprints of operation)
- 5) military capacity and scalability of operations.





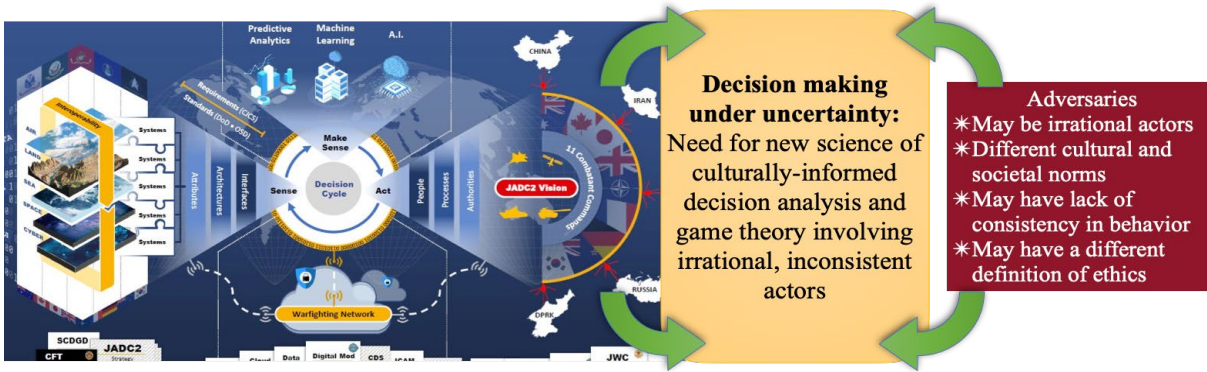


Figure 5. Conceptualization of the Need for Advancement in Modified Game Theory

Academic research that enables these advances are the following but not limited to the organizational theory, incentives to motivate jointness, game theory and modification for irrational and inconsistent actors/adversaries and extracting and understanding operational principles of various actors/adversaries that are culturally informed such as Sun Tzu.

### Opportunity 3: Universal Translator System for Federated Heterogeneous Networks of Sensors: Complexity Management of the Growing Network of Interconnected Sensors, Decision Makers, and Shooters

One of the core technical challenges of JADC2 is the integration of the highly technical legacy sensor networks that are managed and operated by all forces/services. Each service owns a state-of-the-art in intelligence, surveillance, and reconnaissance sensor network that are not necessarily interoperable with other forces' tactical networks. Sense and integration are the ability to discover, collect, correlate, aggregate, process, and exploit data from all domains and sources (friendly, adversary, and neutral) and share the information for decision-making. The requirements for effective data integration must be considered from the earliest stages of data sharing and security and applied across the warfighting domains to deliver rapid collection, fusion, and customization of data (Kirtland, n.d.).

The heterogeneity of the tactical networks and assets of the DoD poses some challenges as well as great advantages to JADC2. The technological solutions for integrating multiple heterogeneous tactical networks are numerous, and many are suboptimal. Each solution uses a specific systems architecture and a combination of technical hardware and software solutions. However, in choosing the best technical solution for integrating a network of sensors, two systems characteristics are of critical importance: flexibility and complexity. Flexibility is the ability of the system to respond to various internal and external changes in a timely and cost-effective manner and is therefore critical for the JADC2 network of sensors, as various scenarios may rise that would need a prompt rearrangement of the interconnected networks. Increased complexity in architecture and technical solutions can also contribute to a fragile network that is prone to errors and attacks on the network, and therefore the complexity of the technical solution should be controlled (Chullen & Nilchiani, 2021; Nilchiani & Pugliese, 2017; Priebe et al., 2020; Pugliese et al., 2018).

As JADC2 looks for the best technical solutions for merging the network of sensors, there are multiple factors that should be considered:

- Need for compartmentalization and federation of complex networks, especially to accommodate the culturally centered interoperable collaborative mechanisms.



- Need for firewalling (protection by isolating from the rest of the networks) and multi-layered security of critical portions of the network, if the need arises (e.g., to separate service-specific functions from joint functions, or if the network goes under attack by an adversary)
- Ownership and management of the integrated networks of sensors: The choice between equal ownership on all interconnected networks versus keeping the primary ownership of each network by forces and sharing when needed (military Services, allied, and coalition)
- How to avoid vulnerabilities from monolithic jointness? Should the heterogeneity of each network remain intact?
- How to isolate adversaries sabotaging efforts, firewall their attacks on our networks, and respond?
- How to avoid and halt intentional/malicious propagation in the network? Noise propagation can delay sensor reading and interpretation of results and affects the effective decision-making process.

The excess network complexity and connecting leads to risks of errors (error propagation and from cross-Service misunderstandings) and vulnerability to attacks from adversaries. The technical solution should address managing complexity on a regular basis and incorporate flexibility and the ability to reconfigure the heterogeneous networks of sensors if the necessity arises. Multiple DoD initiatives related to JADC2 efforts have been working on technical solutions, including Mosaic Warfare (DARPA), Advanced Battle Management System (ABMS; Air Force), Project Convergence (Army), Project Overmatch (NAVY), Fully Networked Command, Control, and Communications (FNC3; Office of the Secretary of Defense), and Fifth Generation (5G) Information Communications Technologies (DoD Chief Information Office). DARPA's Mosaic Warfare program has specifically focused on the need for flexibility and responding to ever-changing environments and scenarios and therefore studying solutions that are responsive to rearrangement and change in situations and environments rapidly.

### **Technical Solution: Universal Translator System for Federated Heterogeneous Networks of Sensors (Rosetta Stone)**

The technical solution for joining networks from all forces (Army, Navy, Air Force, Marine, Space Force) requires achieving a system-of-systems that is more resilient, flexible, and responsive to demands and produces greater information and insights in different scenarios that the DoD is facing. Often, over connecting all sensors and assets of all forces/services could pose some substantial problems including but not limited to 1) slowdown in sensor and information transfer, 2) increased risk of errors and issues in the collection and transfer of data, 3) network vulnerability in the face of cyber attacks and loss of ability to swiftly isolate and contain attacks.

The authors suggest the exploration of a novel concept of a universal translator infrastructure. This Universal Translator would consist of a combination of embedded hardware and software distributed nodes that will act as the interface translator between federated network sensors and assets across all five forces/services as well as all DoD agencies. Figure 6 shows the Universal Translator network concept.

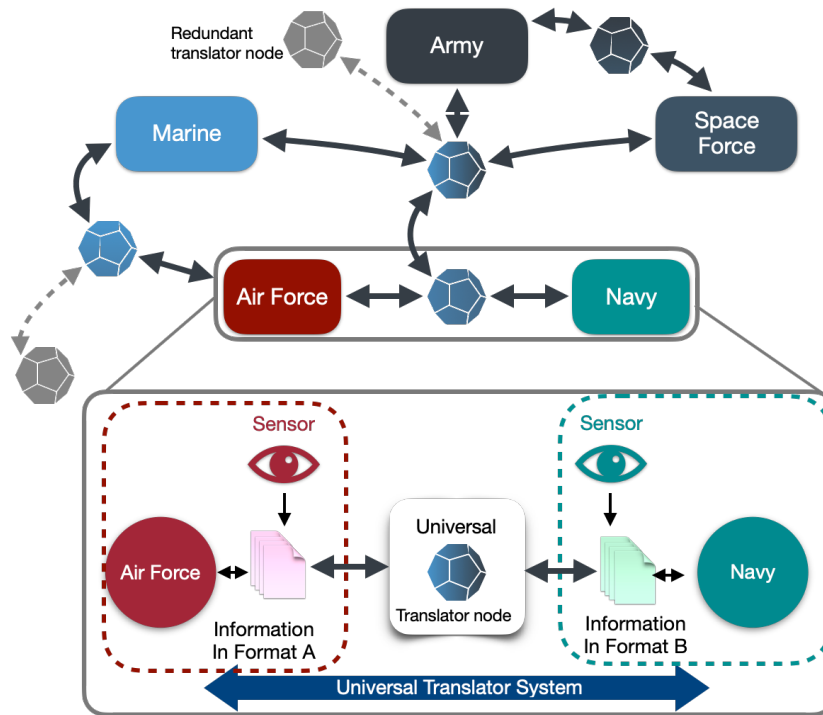
The characteristics of the Universal Translator are as follows:

- Translation between different legacy systems takes place at certain physical hardware and software nodes which are positioned between two or more heterogeneous networks of sensors, belonging to different forces.
- There will be no need to invest in making all sensor assets from different forces into a unanimous frequency and standard. The Universal Translator will provide the translation



between networks, and legacy systems can continue their normal functions with no interruption.

- Universal Translator can consist of multiple nodes as well as redundancies that can operate as a fractionated network of translators and can be easily expanded, modernized, and upgraded with the latest state-of-the-art in technological advances in the future, and rearranged to create new network functions and topography on demand. This concept will provide a high degree of flexibility, adaptation, and upgradability, as well as an added layer of security and protection for all assets and sensors across all forces.
- Universal Translator nodes will act as a bottleneck between two separate networks and can act as a firewall mechanism. If necessary to turn off or isolate a sensor network under attack, certain translator node(s) can be turned off which will revert the isolated network to its original function.
- The Universal Translator network can be embedded with various security layers, giving each force's network extra protection and the ability of Mosaic Warfare (DARPA) novel network rearrangement and protocols.
- Each force can yet command their original assets (network of sensors) as the primary owner of the assets as well as share their data through permission and activation of the Universal Translator to the other forces. Data from various forces can be shared without the need to share the detailed blueprint and architecture behind each network.



**Figure 6. Concept of Universal Translator/Rosetta Stone Infrastructure With a Detailed View of the Universal Translator Infrastructure and Software Translating Data Between Two Agencies/Forces (Nilchiani, 2022)**

Suggestion: Invest in a Universal Translator system for federated heterogeneous networks of sensors that can preserve service-specific functions yet interface seamlessly with joint functions and also operate independently from the rest of the network if under attack.

## Opportunity 4: Portfolio Management of JADC2-Related Acquisition Programs

In *Joint All-Domain Command and Control: Background and Issues for Congress* (Hoehn, 2022), there are several clusters of questions raised regarding managing JADC2-related efforts, budget, cost estimates, and requirements. Among those questions were JADC2 spending priorities, initiatives as well as management of JADC2-related efforts. The solution to managing multiple JADC2-related efforts is to adopt the best practices in portfolio management from the industry and create a comprehensive DoD portfolio management framework to manage multiple efforts. By studying the best of industry innovations on portfolio management, innovative System-of-Systems, and enterprise-level frameworks can be created that empower joint staff of JADC2 to manage, optimize, integrate, and fund JADC2.

JADC2 consists of multiple efforts in data, human, and technical enterprise that fit within hardware, software, business, and major acquisitions. Dealing with multiple concurrent capability acquisitions needs a System-of-Systems-based framework that integrates multiple programs, and a portfolio management approach that funds, manages, and integrates multiple potentially asynchronous acquisition programs for JADC2. The portfolio management framework will need to incorporate the shared governance structure (architecture of governance) for JADC2-related projects.

Academic Research: Portfolio management framework for multiple acquisition programs, Shared governance architecture

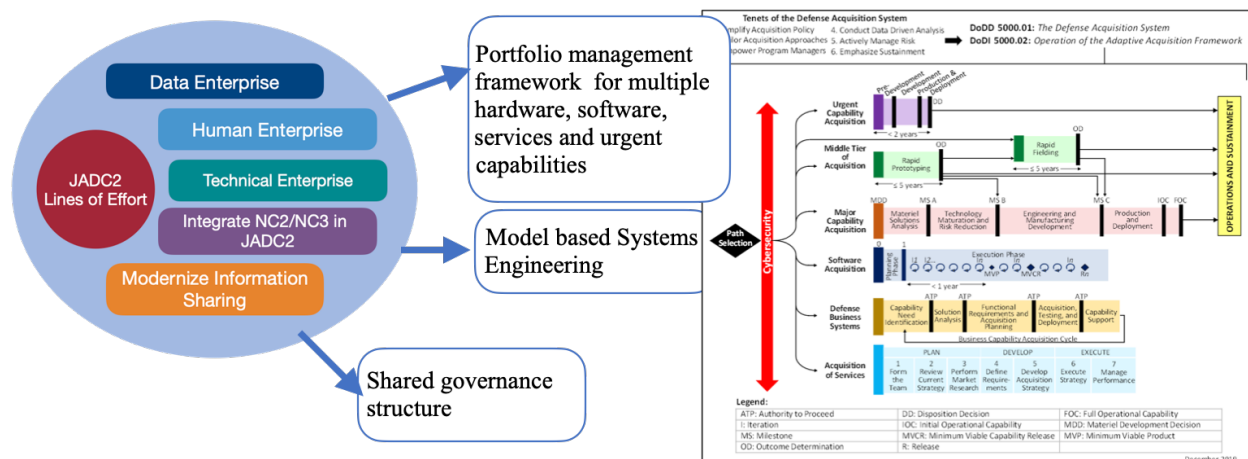


Figure 7. Conceptualization of the Need for Portfolio Management for All JADC2-Related Acquisition Programs

## Opportunity 5: Needs, Requirements, and MBSE for JADC2 as a System of Systems

Joint All-Domain Command and Control (JADC2) is a System of Systems that operates in domains of materiel, non-materiel, policies, authorities, organizational constructs, and operational procedures and therefore in need of systems engineering tools and methodologies to assure the most optimal system of systems. Model based systems engineering (MBSE) can help in responding optimally to categories of questions regarding the acquisition of individual technologies/programs, lifecycle management, and periodic upgrade and infusion of novel technologies to JADC2, as well as ownership and management of various sensors and assets in JADC2.

MBSE can help identify joint-specific systems, needs, and requirements, and guide the acquisition process for a portfolio of programs and technologies. Technical questions about how



sensor networks of various forces and services will be connected, the architecture, and the concept of operation are enabled by the systems approach. MBSE can also find optimal solutions to lifecycle-related questions of JADC2 including identifying new disruptive technologies and integration with current legacy systems, as well as complexity management of the growing interconnected sensor and asset networks of JADC2.

MBSE can also provide suggestions and solutions for network ownership and architecture for various assets. For example, can jointness be achieved and implemented successfully while respecting primary ownership of each force over their assets/sensors? Using the concept of universal translator, each service can retain its primary ownership and command over its assets and sensors and share a secondary ownership of all assets on a need basis. In extreme scenarios, the primary owner can sever their assets from the rest of the network to protect their assets or other services assets and operate independently if need be.

## Summary

This paper provides an overview of the current state of the Joint-All Domain Command and Control and suggests a set of recommendations and opportunities through the lens of academic research and development (R&D). This set of opportunities emphasizes the need for research and development and gaps in knowledge, technologies, procedures, and capabilities that can empower JADC2 as a resilient, agile, adaptive, and strong shared command and control platform.

The following opportunities were proposed in the paper: *opportunity 1*: novel culturally centered interoperable collaborative mechanisms between forces (organizational and cultural studies); *opportunity 2*: necessity for innovation in decision analysis and game theory (modified based on adversaries' socio-cultural nuances ); *opportunity 3*: need for complexity management and best system architecture design for the growing network of interconnected sensors, decision makers, and shooters (the authors suggests a Universal Translator network concept of hardware and software to connect all existing and future heterogeneous network of sensors and assets of the DoD, which will empower rearranging, reorganizing, expanding, and infusing the latest advances in technologies as they become available); *opportunity 4*: need for a novel, strong portfolio management framework of JADC2 Acquisition Programs (to manage, optimize, integrate, and fund JADC2 related projects and acquisition programs); and *opportunity 5*: need for Model Based Systems Engineering (MBSE) for JADC2 as an SoS in domains of materiel, non-materiel, policies, authorities, organizational constructs, and operational procedures.

## Acknowledgement

*The authors would like to thank the support of the Department of Defense on this research effort. 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#0309.*

## References

- Advanced battle management system* (978-0-309-68621-1). (2022).  
<https://nap.nationalacademies.org/catalog/26525/advanced-battle-management-system-needs-progress-challenges-and-opportunities-facing>
- Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge*.  
[http://edocs.nps.edu/dodpubs/org/CCRP/Alberts\\_Power.pdf](http://edocs.nps.edu/dodpubs/org/CCRP/Alberts_Power.pdf)
- Alberts, D. S., & Hayes, R. E. (2006). *Understanding command and control*.  
<https://apps.dtic.mil/sti/citations/ADA457162>



- Alkire, B., Lingel, S., Baxter, C., Carson, C. M., Chen, C., Gordon, D., Hanser, L. M., Menthe, L., & Romano, D. M. (2018). *Command and control of joint air operations in the Pacific: Methods for comparing and contrasting alternative concepts* (9780833098085). [https://www.rand.org/pubs/research\\_reports/RR1865.html](https://www.rand.org/pubs/research_reports/RR1865.html)
- Bass, B. K., Bartels, D. K., Escalante, S. A., Fenton, D. R., & Rathgeb, K. J. (2014). *Overcoming joint interoperability challenges*. National Defense University Press. <https://ndupress.ndu.edu/JFQ/Joint-Force-Quarterly-74/Article/577545/overcoming-joint-interoperability-challenges/>
- Birch, P., Reeves, R., & Dewees, B. (2020). *Building the command and control of the future from the bottom up*. Metamorphic Media. <https://warontherocks.com/2020/01/building-the-command-and-control-of-the-future-from-the-bottom-up/>
- Borel, É. (1927). Sur les systèmes de formes linéaires à déterminant symétrique gauche et la théorie générale du jeu. *Comptes rendus de l'Académie des Sciences*, 184, 52–53.
- Builder, C. H., Bankes, S. C., & Nordin, R. (1999). *Command concepts: A theory derived from the practice of command and control*. RAND. [https://www.rand.org/pubs/monograph\\_reports/MR775.html](https://www.rand.org/pubs/monograph_reports/MR775.html)
- Chullen, C., & Nilchiani, R. (2021). Infusion complexity: Understanding the need to measure infusion success of advanced technologies into complex systems. *2021 IEEE International Systems Conference (SysCon)*.
- Congressional Research Service. (2021a). *Defense primer: What is command and control?* <https://crsreports.congress.gov>
- Congressional Research Service. (2021b). *Joint all-domain command and control: Background and issues for Congress*. <https://crsreports.congress.gov>
- Congressional Research Service. (2022). *Joint all-domain command and control (JADC2)*. <https://crsreports.congress.gov>
- Doctrine for the Armed Forces of the United States*. (2013). A. F. o. t. U. States.
- DoD. (2022). *Summary of the joint all-domain command & control (JADC2) strategy*.
- Feickert, A. (2022). *U.S. special operations forces (SOF): Background and issues for Congress* (RS21048). <https://crsreports.congress.gov/product/details?prodcode=RS21048>
- Friedman, B. H. (2019). *Bad idea: Management jointness in DoD*. Center for Strategic and International Studies. <https://defense360.csis.org/bad-idea-management-jointness-in-dod/>
- Hansanyi, J. (1967). Games with incomplete information played by Bayesian players, I: Basic model. *Management Sci*, 14(3), 159–182.
- Hoehn, J. R. (2022). *Joint all-domain command and control: Background and issues for Congress*. <https://crsreports.congress.gov/product/details?prodcode=R46725>
- JNT-501S introduction to joint operations: Curriculum*. (2019). [https://fairchild-mil.libguides.com/Introduction\\_To\\_Joint\\_Operations](https://fairchild-mil.libguides.com/Introduction_To_Joint_Operations)
- Joint doctrine publication 5 command and control*. (2012).
- Jointness - A selected bibliography*. (1993). <https://apps.dtic.mil/sti/pdfs/ADA272189.pdf>
- Kirtland, M. A. (n.d.). <https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Chronicles/joint.pdf>
- Kreisher, O. (2001). The quest for jointness. *Air Force Magazine*. <https://www.airforcemag.com/article/0901joint/>
- Lawlor, M. (2007). *Military jointness grows up*. AFCEA International. <https://www.afcea.org/content/military-jointness-grows>
- McInnis, K. J. (2021). *Defense primer: Commanding U.S. military operations*. <https://crsreports.congress.gov/product/details?prodcode=IF10542>
- Michael, K., Siman-Tov, D., & Yoeli, O. (2017). Jointness in intelligence organizations: Theory put into practice. *Cyber, Intelligence, and Security*, 1(1). <https://www.inss.org.il/wp-content/uploads/2017/03/Jointness-in-Intelligence-OrganizationsTheory-Put-into-Practice.pdf>
- Nardulli, B. R., Cliff, R., Bensahel, N., Rostker, B. D., Pirnie, B. R., Gordon, J., IV, Orletsky, D. T., Hosek, S. D., Peltz, E., & Halliday, J., e. a. (2003). *The U.S. Army and the new national security strategy*. RAND. [https://www.rand.org/pubs/monograph\\_reports/MR1657.html](https://www.rand.org/pubs/monograph_reports/MR1657.html)
- Nilchiani, R. R. (2022). *Joint all-domain command and control (JADC2) technical challenges and research opportunities* [Report]. [https://document.acqirc.org/publication\\_documents/reports/1670425668.JADC2\\_REPORT.pdf](https://document.acqirc.org/publication_documents/reports/1670425668.JADC2_REPORT.pdf)
- Nilchiani, R. R., & Pugliese, A. (2017). *A systems complexity-based assessment of risk in acquisition and development programs*.



- O'Rourke, R. (2021). *U.S. role in the world: Background and issues for Congress* (R44891). <https://crsreports.congress.gov/product/details?prodcode=R44891>
- Priebe, M., Ligor, D. C., McClintock, B., Spirtas, M., Schwindt, K., Lee, C., Rhoades, A. L., Eaton, D., Hodgson, Q. E., & Rooney, B. (2020). *Multiple dilemmas for the joint force: Joint all-domain command and control*. [https://www.rand.org/pubs/research\\_briefs/RBA381-1.html](https://www.rand.org/pubs/research_briefs/RBA381-1.html)
- Pugliese, A., Enos, J., & Nilchiani, R. (2018). *Acquisition and development programs through the lens of system complexity*.
- Rapoport, A. (2012). *Game theory as a theory of conflict resolution* (Vol. 2). Springer Science & Business Media.
- Spirtas, M. (2018). *Toward one understanding of multiple domains*. RAND. <https://www.rand.org/blog/2018/05/toward-one-understanding-of-multiple-domains.html>
- Theohary, C. A. (2021). *Defense primer: Cyberspace operations* (IF10537). <https://crsreports.congress.gov/product/details?prodcode=IF10537>
- Transforming the joint force: A warfighting concept for great power competition*. (2003). Defense Media Activity - WEB.mil. <https://www.pacom.mil/Media/Speeches-Testimony/Article/2101115/transforming-the-joint-force-a-warfighting-concept-for-great-power-competition/>
- Wolf, A. F. (2021). *Defense primer: Command and control of nuclear forces* (IF10521). <https://crsreports.congress.gov/product/details?prodcode=IF10521>



## PANEL 16. TECHNOLOGY-ENABLED LOGISTICS & SUSTAINMENT

Thursday, May 11, 2023	
10:30 a.m. – 11:45 a.m.	<p><b>Chair: Raymond Jones, COL USA (Ret.),</b> Chair, Department of Defense Management, Naval Postgraduate School</p> <p><b><i>Acquiring Maintainable AI-Enabled Systems</i></b> Shane Kohtz, USMA/ACI Iain Cruickshank, USMA/ACI</p> <p><b><i>Commercial and Defense Vendor Management: A Comparison of Competitive Procurement Below the Prime – Subcontract Competition – How real is it?</i></b> Lt Col Daniel Finkenstadt, USAF, Naval Postgraduate School Kyle Braunlich, Air Force Lifecycle Management Center Peter Guinto, Resilinc</p> <p><b><i>Optimizing Operations and Logistics Support Using Opus Evo</i></b> Gustaf Solveling, Systecon North America John Verbanick, Systecon North America</p>

**Raymond Jones, COL USA (Ret.)**— retired as a Colonel from the U.S. Army in 2012 and is a Professor of Practice within the Graduate School of Defense Management at the U.S. Naval Postgraduate School in Monterey, CA. He also serves as a Guest Lecturer for the IDARM Program within the Institute for Security Governance (ISG), Defense Security Cooperation Agency (DSCA). 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.





# Acquiring Maintainable AI-Enabled Systems

**MAJ Iain Cruickshank, USA**—is a Functional Area 49 (Operations Research/Systems Analysis) officer in the U.S. Army. He is currently a senior research scientist at the Army Cyber Institute. He has previous assignments with the Army's Artificial Intelligence Integration Center, the 780th Military Intelligence Brigade, and the 101st Airborne Division. He holds a PhD in Societal Computing from Carnegie Mellon University, which was obtained as a National Science Foundation Graduate Research Fellow, and an MS in Operations Research from the University of Edinburgh, which was obtained as a Rotary Ambassadorial Scholar. [iain.cruickshank@westpoint.edu]

**MAJ Shane Kohtz, USA**—is a Functional Area 51 (Acquisition) officer in the U.S. Army. He is currently a cyber research manager at the Army Cyber Institute. He has previous assignments with the Missile Defense Agency, the Army's Program Executive Office Intelligence Electronic Warfare & Sensors, the 1st Infantry Division, and the 101st Airborne Division. He holds an MBA from the Naval Postgraduate School with a focus in Systems Acquisition Management. He is a member of the Army Acquisition Corps and holds a DAWIA Advanced certification in program management. [shane.kohtz@westpoint.edu]

## Abstract

The Army and other services are quickly entering into an age where many, if not all, acquisitions programs will need to contend with acquiring Artificial Intelligence (AI)-enabled systems. While there has been research on how to acquire the data or model for an AI-enabled systems, sustainment considerations have been overlooked. Given the importance of sustainment for any acquisition program of record—both in terms of cost and in terms of program effectiveness—it is imperative that the Army, and the rest of the DoD, plan for AI-enabled system maintenance. To address this gap, this paper proposes a framework and practices that draw on best practices from industry, program maintenance, and Machine Learning Operations (MLOps) to integrate AI maintenance into a product support strategy and Life Cycle Sustainment Plan. The framework outlines necessary components for sustainable AI and considers varying levels of maintenance to reduce operation and sustainment costs.

## Introduction

Technology on the battlefield will increasingly need to become data centric and automated to have a tactical advantage over adversaries' technologies; AI will be an integral part of future warfare (NSCAI, 2021). The United States Department of Defense's (DoD) primary solution to this capability gap is a significant investment into Artificial Intelligence (AI) and, AI's primary driver, Machine Learning (ML). For example, in preparation for fiscal year 2023, the Department of Defense requested \$1.1 billion to further research and development of the immature AI and ML technology (DoD, 2022a). AI will be part of many future systems that we will acquire and upgrade; by 2045 it will probably be a standard component of every major piece of military equipment (NSCAI, 2021). As these technologies mature, and are incorporated into systems and programs, they then need to be maintained. While the defense acquisition community has started considering data (Nagy, 2022), use cases (Guariniello, 2021), and hardware for AI-enabled systems, there is little to no thought on how the sustainment of these AI-enabled systems will work for major programs. Thus, while the DoD has invested heavily into maturing AI and ML for future AI-enabled systems, its less clear how the defense acquisition community could maintain and sustain these AI-enabled systems.

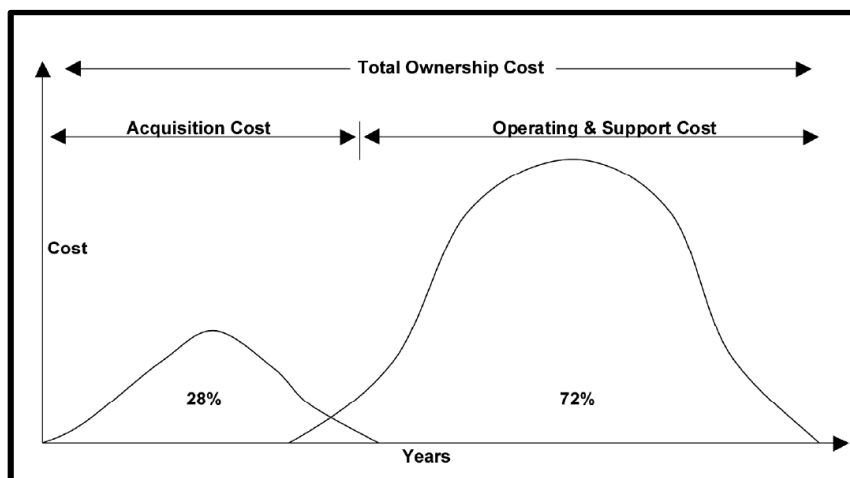
This paper proposes a paradigm, along with recommendations for program offices, to utilize when planning the acquisition strategy of an AI-enabled program of record. We first outline the importance of maintenance planning in a program and why AI-enabled systems need maintenance. We then discuss the main considerations in planning for the maintenance of an AI-enabled system. These maintenance considerations are necessary to inform the strategy to meet sustainment requirements known as the Product Support Strategy (PSS) and Life Cycle



Support Plan (LCSP) for a program of record (OUSD(A&S), 2021). The early planning for the acquisition logistics strategy prevents the possibility of a program breach or uncaptured costs later in the program. AI-enabled systems will become more prevalent on the future battlefield while the sustainment planning occurs now.

## Background

*Maintenance planning in a program of record.* Maintenance is one of the most critical aspects of a major acquisitions program. Maintenance considerations occur early in the life cycle of a program of record, and early sustainment decisions have a long-term effect during the operations and sustainment phase of a program (DoD, 2016). Why is sustainment planning important early in the acquisition life cycle? The acquisition community has known for years that operation and sustainment costs account for the majority of a program’s total ownership costs; in fact, 72% of the total ownership costs occur during the program’s operation and sustainment phase (Schinasi, 2003). Figure 1 illustrates how a program costs are distributed across an acquisition program’s life cycle. Operation and sustainment planning slightly improved in recent years. The O&S Cost Management Guidebook stated, “in the December 2014 Selected Acquisition Reports (SARs), on average, 67% of the reported costs are attributable to O&S” (DoD, 2016). Despite the slight improvement, most costs for a program remain during operations and sustainment.



**Figure 1. Nominal Life-Cycle Cost of Typical DOD Acquisition Program with a 30-Year Service Life (Schinasi, 2003).**

In addition, when requirements are approved, nearly 85% of operation and sustainment costs are known with less than 10 % of the life cycle costs spent (Schinasi, 2003). Figure 2 illustrates the importance of early planning with systems for AI/ML requirements. AI/ML capable systems are early in the technology maturation process with substantial investments, but the majority of sustainment costs are already determined. Program offices must proactively plan and determine the Product Support Strategy (PSS) at program inception and then the Life Cycle Sustainment Plan (LCSP) at the first acquisition milestone, Milestone A, even though the sustainment of AI enabled systems may be unknown currently (OUSD[A&S], 2021).

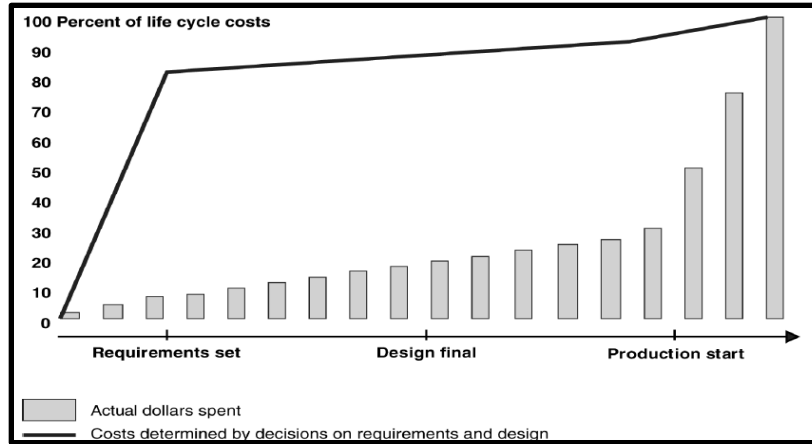


Figure 2. Percent of Operating and Support Costs Determined at Various Points in the Acquisition Process (Schinasi, 2003).

*AI-enabled systems and their maintenance.* AI-enabled systems, like any other piece of technology, require maintenance. An AI-enabled system consists of traditional software and, possibly, hardware, depending on the purpose of the system in addition to AI components of the system. AI components often require several hardware and software dependencies, often called a stack (Moore, 2018). Figure 3 illustrates the AI stack. One of the critical elements of the AI components, and, really, what makes the entire system an AI-enabled system are the ML models. The ML models enable the system to engage in automated behaviors and activities that typically require human levels of perception or reasoning; they are the “brain” of the AI-enabled system. These ML models, much like every other component of the AI-enabled system, also require maintenance.

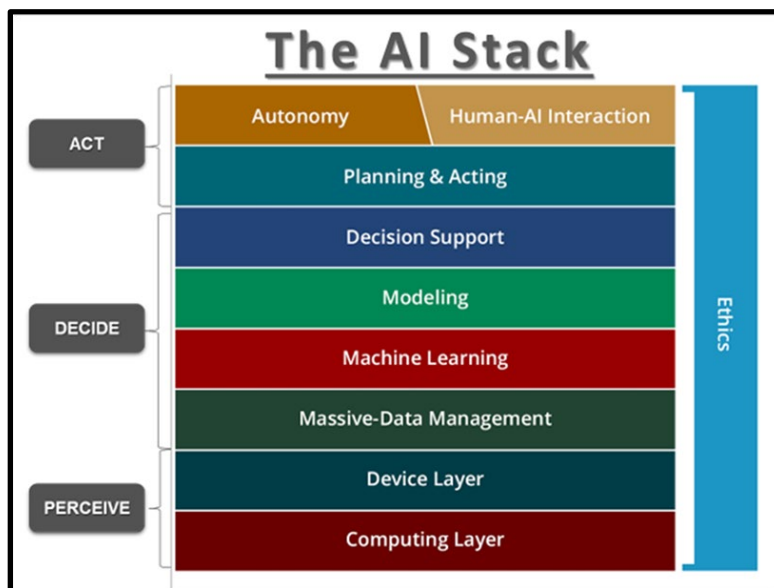


Figure 3. Carnegie Mellon University's AI Stack, Depicting the Necessary Components of an AI-Enabled System (Moore, 2018).

ML models, despite their potential, still suffer from several issues that necessitate frequent maintenance. ML models, by their nature, learn correlations useful to a certain task from the data that is presented to them. Thus, these models could have performance issues if

the data presented to the model when in use is different than the data it was trained on (i.e., Out-of-Domain Data problem; Patrino, 2019). As an example of this, a computer vision ML model, which is meant to detect certain vehicles from a ground perspective, can fail when something as simple as the background, or biome, is different between the model's training data and where the model is used (e.g., urban versus rural setting). ML models can also suffer from issues like model drift (Talby, 2018), data drift (Evidently AI, 2021), concept drift (Patrino, 2019), or even changing of hardware, like sensors, which all greatly affect ML model performance. In addition to those issues which naturally arise, ML models can also be directly attacked via Adversarial ML, which will also seriously degrade ML model performance (Talby, 2018). Finally, it should be noted that many of these issues are unique to ML and ML-enabled systems; changing of something like the background of images does not affect the hardware or software of a traditional, digital system. Thus, ML models have their own inherent issues which necessitate maintenance for those ML models, which over and above the maintenance for traditional hardware and software systems.

While ML models suffer from several issues, which can greatly affect their performance, dealing with these issues frequently requires far less resources and know-how than the initial development of the ML model. Maintaining ML models in use in the real-world (i.e., model deployment) can often be handled with a collection of updating and monitoring processes, which are collectively part of the industrial ML paradigm of MLOps (Treveil et al., 2020). MLOps, at its core, is a set of practices which aims to productionize ML systems (Treveil et al., 2020). Figure 4 depicts the core components and relationships of MLOps. While the principles and practice of MLOps are still an active area of research, three practices that are a mainstay of MLOps are the monitoring of data and models in production, the continual updating of models in response to changes, and having model maintenance take place with model operation (Treveil et al., 2020). These are an integral part of MLOps because they are how organizations and businesses can use ML models despite their inherent issues. Thus, key to the use of ML models in the real world and in production systems in the MLOps paradigm is having in place the right tools and practices to monitor an ML model and its data as well as the correct steps to update ML models, as close to operation as is feasible.

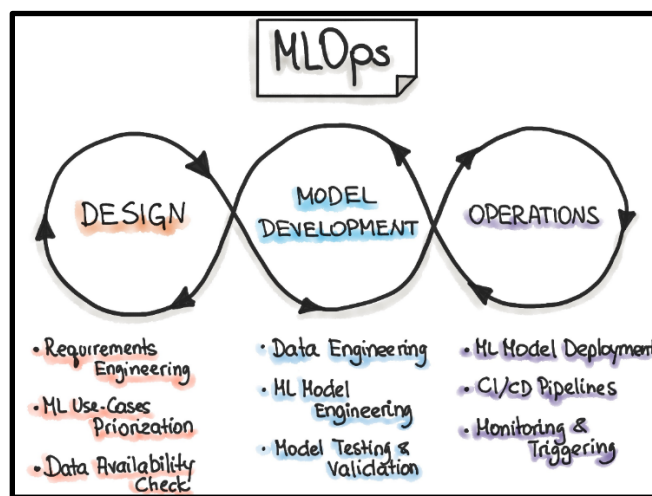


Figure 4. Core Components of MLOps and Their relationships (Visengeriyeva, 2023).

Of note in the MLOps paradigm is *model retraining*. Ideally, model retraining consists of running all of the steps required to train an ML model, but with a new dataset; model retraining should not require any changes to the code—just changes to the weights of the model (Patrino, 2019). This type of maintenance generally needs to occur anytime the data changes, and an

updated training data set is available (Evidently AI, 2021). Thus, this type of maintenance generally comes in two forms, periodic and dynamic (Evidently AI, 2021). Periodic retraining is when there are known changes in the data that will occur, such as quarterly or yearly changes in business practices. Whereas dynamic retraining occurs any time there are changes in the data generation process, such as collecting in an adversarial environment (e.g., detecting credit fraud) or a naturally dynamic process (e.g., labeling objects in imagery). When it comes to dynamic retraining, it can occur on widely variant time scales depending on the ML application; some ML models need to be updated daily, while others need only be updated monthly or yearly (Evidently AI, 2021). Regardless of the frequency of ML model retraining, all experts on the subject of using ML models in the real world agree that this process is a must for any ML-enabled system. Thus, model retraining is a necessary part of any ML model and may need to occur daily.

## Considerations for Maintaining an AI-Enabled System

When it comes to taking AI-enabled maintenance into program planning, there are a couple of key considerations. These considerations should inform program offices when they perform a Product Support Business Case Analysis (PS BCA) that informs the PSS and LCSP (DoD, 2022b). The PS BCA evaluates potential alternatives for sustainment to include organic, contractor, or a ratio mix of support that informs a decision on the program's sustainment strategy (DoD, 2014). The PSS and LCSP are updated at each acquisition milestone; however, as highlighted earlier, nearly 85% of the sustainment costs are determined when requirements are set (Schinasi, 2003). An understanding of the requirements and maintenance "touch time" of AI/ML systems is imperative during the strategy development phase to properly plan and budget sustainment. This maintenance of ML models is in addition to all the hardware and software underlying the AI stack, which are necessary to run the ML models. Such a requirement can enable possible project scenarios wherein the ML model is a sub-product, or product-within-a-product, of a larger AI-enabled system. Overall, in addition to the maintenance requirements of software and any hardware, there are also requirements for the maintenance of the AI components that should address any intellectual property, data, and ML models.

*Intellectual Property and Data.* A critical component to the PS BCA, PSS and LCSP is a program's Intellectual Property (IP) Strategy for sustainment planning. DODI 5000.91 (Product Support Management for the Adaptive Acquisition Framework) states "the IP strategy identifies, and acquisition contracts should secure, sufficient technical data, manuals, and publications to enable informed Government decisions to acquire maintenance and repair through Government organic capability and/or contractor-provided solutions" (OUSD[A&S], 2021). The role of data rights is even more critical for AI enabled systems given the amount of maintenance required on a routine basis. Program offices may be unaware of the type of data required to conduct organic maintenance because AI is an emerging technology.

The Defense Federal Acquisition Regulation Supplement outlines government rights for data, which are unlimited rights, government purpose rights, or limited rights (GSA, 2023). Program offices must understand these rights in acquisition planning and contract negotiation for AI/ML enabled systems. A recent RAND study noted that government program offices did not understand data rights, which had long term impacts on sustainment planning. Vendors would leverage the "proprietary" label and utilize court systems to maintain data rights in a weapon system for follow on sustainment. As a result, the government typically would not want to go through the elaborate court proceedings and thus acquiesce to the vendor's claims concerning data rights (RAND, 2021). The RAND case study highlights the importance of data rights when planning weapon system sustainment, and the lessons learned are imperative since AI-enabled systems require a substantial amount of touch time for maintenance.



*ML Model Maintenance Considerations.* There are a few different paradigms to approaching maintenance for ML models. Much like sustainment for other components of a system, the maintenance of an ML model can use both contract and organic service support alternatives. At the one end of the spectrum is the ML model maintenance being performed solely by contract. This means contractors would be responsible for all of the tasks of model maintenance including data and model monitoring, development of test and evaluation metrics, development of model retraining procedures, model updating (i.e., performing the model retraining procedures), model retirement and replacement, and model governance (i.e., making sure any ML model is meeting necessary guidelines and regulations). A particular version of the contractor only approach in use is the ML-as-a-Service (MaaS) model. The MaaS model usually works through application programming interfaces (APIs), whereby the contractor has full responsibility for the model, to include initial development and maintenance, and a user just sends data to an API to use the ML model. This type of model is currently used by companies like OpenAI and by organizations like the XVIIIth Airborne Corps and often works on a pay-per-usage type of pricing scheme.

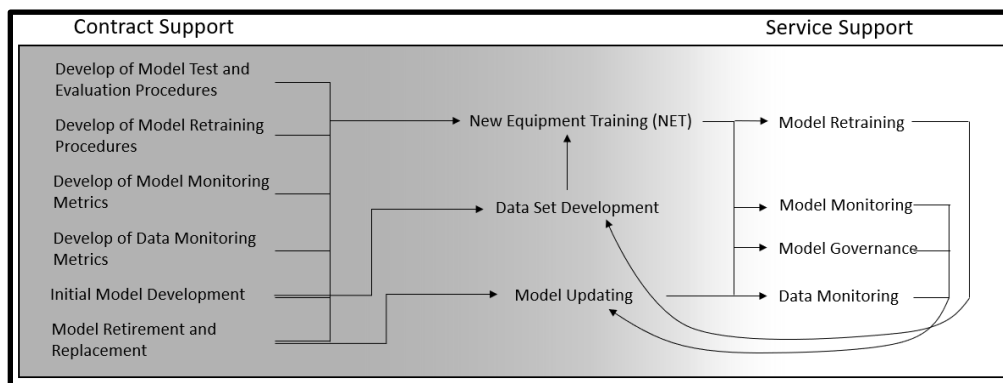
While the contractor-only approaches present the simplest approach to maintenance planning, they have serious pitfalls that must be considered. For the MaaS model, despite the simplicity of this model, much like any other pay-per-use pricing scheme (e.g., cloud services, SaaS), it can quickly become exorbitantly expensive if there is a lot of use of the service. Additionally, it requires connectivity back to the API to work. So, if the AI-enabled system is meant to work in austere environment or have a lot of usage on the ML-models, going through a MaaS model may be overly costly. Additionally, having contractors perform all the functions of ML maintenance ignores the hard-learned lessons behind the MLOps paradigm; namely the operation of the ML model has been separated from its maintenance and development. A primary reason why MLOps places the development and maintenance of ML models so close to the running of ML models is that these models require constant monitoring and frequent updating (Treveil et al., 2020). In fact, one form of updating, model retraining, can occur as frequently as daily for an ML model in production in an adversarial and dynamic environment. As with our previous computer vision example of detecting objects from a ground perspective, the ML model would need to be, at a minimum, retrained every time the biome changes (e.g., moving from rural to urban) and every time an organization wants to detect a new or different set of objects. Conceivably, such a change in an ML model's operating environment could occur several times over the course of a single operation for a military unit. Thus, given the frequent nature of ML model maintenance, having contractors provide all this maintenance could be cost prohibitive.

At the other end of the spectrum is a service only solution, where servicemembers and DoD civilians are responsible for all of the aforementioned ML model maintenance tasks. While this certainly presents some potential for cost savings in terms of maintenance, the Army and DoD may lack the skill sets in house, in sufficient numbers, to perform some maintenance functions. This is especially true for maintenance functions like designing a test and evaluation scheme for both the ML model and its data as well as determining the right model retraining procedures (e.g., active learning, fine-tuning, etc.). These types of maintenance tasks often take a seasoned data scientist with domain area expertise and, often, advanced education. That said, some of the maintenance tasks actually require very little education and can be learned with suitable training. For example, actually performing model updates, given a guide to the model's retraining procedures, is trainable task that does not require an advanced educational background. Thus, planning to do the full spectrum of model maintenance in house may be infeasible, given constraints on in house ML expertise.



## Conclusion and Recommendation

The acquisition of AI-enabled technologies that will be successful for military operations must have sustainment of their ML models taken into primary consideration. ML models have critical fragilities that require monitoring and updating. What is more the typical frequency of ML model retraining for dynamic and adversarial environments makes it prohibitive for this type of maintenance to be done by contractors. Fortunately, if an AI-enabled system is properly implemented, monitoring and retraining ML models can be a trainable task that can be performed in house. So, it is vital that we acquire AI-enabled systems that allow for this in house maintenance if that AI-enabled system is going to be useful for military operations. As such, we recommend a hybrid approach to ML model sustainment planning, that leverages expertise from contractors, but relies on servicemembers for execution of the maintenance. Figure 5, details the sustainment tasks and which component should be responsible for them.



**Figure 5. ML Model Sustainment Tasks in a Hybrid Maintenance Plan with Associated Dependencies Between Contractor and Service Maintenance Tasks.**

When it comes to the actual amount of effort expended on these maintenance tasks, those in the service support region are the equivalent of field-level maintenance (DoD, 2022). Those tasks are the ones most frequently done and the tasks that can address most issues with ML models in use. Whereas those within the contract support, namely model retirement and replacement, as well as some that are a shared task, like model updating, would be depot-level maintenance (DoD, 2022). These tasks should only be needed periodically and to address major issues with the ML model.

Along with our proposal of a hybrid maintenance model for AI-enabled systems, we also propose the following points be part of any program planning:

- **Data Rights:** Program offices, looking to have ML models in their programs, may negotiate limited rights for implementation of the ML models since government operators would be doing the model retraining and monitoring. However, since the deliverables will most likely come from mixed funding, the program offices should, at a minimum, negotiate for government purpose rights of the technical data and deliverables. This approach will give the program office flexibility in the future if they decide to change the sustainment strategy.
- **ML Model Touch-Time Analysis:** As has been mentioned within this paper, ML models, the brain of any AI-enabled system, require model retraining for various reasons. The amount of model retraining for any given ML model is highly context dependent; it can vary from daily retraining up to monthly or even yearly (Evidently AI, 2021). Thus, as part of the PS BCA, there needs to be a retraining requirements analysis. This analysis should, at a minimum, consider how often the data environment for the AI-enabled

system predictably changes, whether it will be used in an adversarial environment (i.e., data environment where people generating the data attempt to change data generation patterns to fool the system), and how often the data generation process changes physical locations (i.e., a sensor moves from one geographic region to another). With the information from this analysis, a program office can have a much better estimation of the maintenance cost requirements. We also note that this type of analysis is fruitful grounds for future, impactful research.

In conclusion, as the Department of Defense invests heavily in emerging AI technology, the acquisition community must prioritize maintenance and sustainment considerations. Early and knowledgeable sustainment planning for a new technology such as AI and ML is imperative considering 85% of operation and sustainment costs are determined in the requirement development stage (Schinasi, 2003). This research proposes a new paradigm and provides a usable framework for the acquisition and sustainment strategy development of a maintainable AI-enabled system. ML models have critical fragilities that drive the need for substantial maintenance on AI-enabled systems. The proposed framework's maintenance considerations serve as a starting point for program offices to evaluate alternatives in the Product Support Business Case Analysis for informed decision-making on Product Support Strategy and Life Cycle Sustainment Plan. The necessary technical data, data rights, training, and a mix of organic and contractor maintenance support are important inputs when developing the Product Support Strategy. This research recommends a mixed sustainment strategy for contractor deliverables and depot-level maintenance while service members execute field-level maintenance for data monitoring and model retraining, monitoring, and governance. Future research can focus on maintenance touch time frequency in a complex operational environment to inform AI maintenance requirements further. Nonetheless, AI-enabled system sustainment planning is crucial and should start now.

*The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.*

## References

- Camm, F., Whitmore, T. C., Weichenberg, G., Sheng, T. L., Carter, P., Dougherty, B., Nalette, K., Bohman, A., & Shostak, M. (2021). *Data rights relevant to weapon systems in Air Force special operations command* (Report No. RR-4298-AF). RAND.  
[https://www.rand.org/pubs/research\\_reports/RR4298.html](https://www.rand.org/pubs/research_reports/RR4298.html)
- DoD. (2014). *DoD product support business case analysis guidebook*.  
[https://www.dau.edu/tools/Lists/DAUTools/Attachments/127/Product-Support-Business-Case-Analysis-\(BCA\)-Guidebook.pdf](https://www.dau.edu/tools/Lists/DAUTools/Attachments/127/Product-Support-Business-Case-Analysis-(BCA)-Guidebook.pdf)
- DoD. (2016). *Operating and support cost management guidebook*.  
<https://www.dau.edu/tools/Lists/DAUTools/Attachments/126/Operating-and-Support-Cost-Management-Guidebook.pdf>
- DoD. (2022a). *Defense budget overview: United States Department of Defense fiscal year 2023 budget request*.  
[https://comptroller.defense.gov/Portals/45/Documents/defbudget/FY2023/FY2023\\_Budget\\_Request\\_Overview\\_Book.pdf](https://comptroller.defense.gov/Portals/45/Documents/defbudget/FY2023/FY2023_Budget_Request_Overview_Book.pdf)
- DoD. (2022b). *Product support manager guidebook*.  
[https://www.dau.edu/tools/Lists/DAUTools/Attachments/129/Product-Support-Manager-\(PSM\)-Guidebook.pdf](https://www.dau.edu/tools/Lists/DAUTools/Attachments/129/Product-Support-Manager-(PSM)-Guidebook.pdf)
- Evidently AI. (2021, July 01). *When to retrain a machine learning model? Run these 5 checks to decide on the schedule*. <https://www.kdnuggets.com/2021/07/retrain-machine-learning-model-5-checks-decide-schedule.html>
- General Services Administration. (2023, March 1). *Defense federal acquisition regulation supplement (DFARS) 227.7103-5 government rights*. <https://www.acquisition.gov/dfars/227.7103-5-government-rights>





- Guariniello, C., Balasubramani, P., & DeLaurentis, D. (2021). A system-of-systems approach to enterprise analytics design: Acquisition support in the age of machine learning and artificial intelligence. *Proceedings of the Nineteenth Annual Acquisition Research Symposium*, 205–217. <https://dair.nps.edu/handle/123456789/4519>
- Moore, A., Hebert, M., & Shaneman, S. (2018). The AI stack: A blueprint for developing and deploying artificial intelligence. *Proceedings of SPIE Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*. <https://www.ri.cmu.edu/publications/the-ai-stack-a-blueprint-for-developing-and-deploying-artificial-intelligence/>.
- Nagy, B. (2022). Tips for CDRLs/requirements when acquiring/developing AI-enabled systems. *Proceedings of the Nineteenth Annual Acquisition Research Symposium*, 218–241. <https://dair.nps.edu/handle/123456789/4587>
- National Security Commission on Artificial Intelligence. (2021). *Final report*. <https://www.nsc.ai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>
- Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD[A&S]). (2021, November 4). *Product support management for the adaptive acquisition framework* (Department of Defense Directive 5000.91). Department of Defense. <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500091p.PDF?ver=qk1sICU3Y0c1aclDocWyJA%3d%3d>
- Patruno, L. (2019, June 10). *The ultimate guide to model retraining*. <https://mlinproduction.com/model-retraining/#:~:text=Quickly%20changing%20training%20sets%20might,require%20monthly%20or%20annual%20retraining>
- Schinasi, K. V. (2003). *Best practices: Setting requirements differently could reduce weapon systems' total ownership costs*. Government Accountability Office.
- Talby, D. (2018, June 5). *Lessons learned turning machine learning models into real products and services*. <https://www.oreilly.com/radar/lessons-learned-turning-machine-learning-models-into-real-products-and-services/>
- Treveil, M., Omont, N., Stenac, C., Lefevre, K., Phan, D., Zentici, J., . . . Heidmann, L. (2020). *Introducing MLOps*. O'Reilly.
- Visengeriyeva, L., Kammer, A., Bär, I., Kniesz, A., Plöd, M., & Eberstaller, S. (2023, March 30). *MLOps Principles*. <https://ml-ops.org/content/ml-ops-principles>



# **Commercial and Defense Vendor Management: A Comparison of Competitive Procurement Below the Prime—Subcontract Competition—How Real is It?**

**Lieutenant Colonel Daniel Finkenstadt, USAF**—is an Assistant Professor at NPS. Finkenstadt completed his PhD in Marketing from Kenan-Flagler Business School, has more than 20 years of defense contracting experience, and is a graduate of NPS. His research interests are perceived service quality, value, business-to-government markets, professional services (knowledge-based services), and non-traditional government contractor motivations. He has published articles in the NCMA CM Magazine, Defense Acquisition Review Journal, Journal of Purchasing and Supply Management, International Journal of Operations and Production Management, the Milbank Quarterly, and the Harvard Business Review. He is the Principal Investigator for the new Simulation and Ideation Lab for Applied Science (SILAS) at NPS.

**Kyle Braunlich**—is a Contract Cost/Price Analyst and Contracting Officer for the USAF at Wright Patterson Air Force Base. He leads acquisition teams through complex, large dollar procurements for major weapons systems and programs. Braunlich has held multiple contracting positions spanning command, control, communications, computers, intelligence, cyber, and career field management. Braunlich has an MBA from the State University of New York Polytechnic Institute. He also served as an adjunct professor lecturing business courses to include Financial Management (Corporate Finance), Financial Accounting, Macroeconomics, and Human Resources Management at the Mohawk Valley Community College.

**Pete Guinto**—is the President of Government, Defense, and Aerospace at Resilinc and has past work in Federal acquisition, law, emergency medical response, firefighting, and respiratory therapy. For the USAF, Guinto held positions as a Chief of Contracts, Chief of Career Field Management, a contracting officer, cost/price analyst, program manager and a procurement analyst across assignments at WP-AFB, the Pentagon, and Randolph AFB. Guinto worked on the AF/DoD/Federal COVID Supply Chain Task Forces and for many AF weapon systems offices. Guinto has undergraduate and a law degrees from the University of Akron and has executive education certificates from Wharton, Kennedy School, Darden, Kellogg, Smeal, and McCombs.

## **Abstract**

This research looks at how the rates of competition at the subcontractor level compare to commercial norms across a wide data set. A quantitative analysis of a large number of commercial parts (~5 million) compare to a statistically similar number of parts from the DoD will be conducted to compare how frequently items are single or sole sourced in each space. The findings will help assess whether the rate of subcontract competition is similar or dissimilar and the degree to which acquisition strategies may need to be adjusted to account for those differences. Then, a qualitative study will be performed assessing the differences and similarities in the data. Generally, acquisition in the DoD leans heavily on competition to drive improvements to cost, schedule, and performance. GAO reports (<https://www.gao.gov/products/gao-15-484r>), reports by the DoD (<https://media.defense.gov/2022/Feb/15/2002939087/-1/-1/1/STATE-OF-COMPETITION-WITHIN-THE-DEFENSE-INDUSTRIAL-BASE.PDF>) and from news outlets (<https://www.defensenews.com/pentagon/2022/04/12/kathleen-hicks-warns-of-substantial-decline-in-defense-industrial-base-competition/>) have all pointed to reduction in competition in the defense industrial base. These sources look primarily to competition at the prime contract level and with very large subcontracts that trigger reporting requirements. Currently, the Competition in Contracting Act (CICA), signed into law in 1984, is the driving force behind using competition as a driver for fair prices in Government acquisition. In that same year, “The Japanese Way” (<https://hbr.org/1984/07/simple-truths-of-japanese-manufacturing>) was brought into mainstream manufacturing in the United States and management practices that encouraged lean manufacturing and closer relationships with single and sole source vendors. The qualitative analysis of the results will be used to assess the business and vendor management strategies



deployed by both commercial and defense acquisition personnel with a focus on enriching a more sophisticated understanding of both competition and collaboration within the vendor base.

## **Subcontract Competition—How Real is It?**

### **Introduction and Research Question(s):**

This research explores variation in rates of competition at the subcontractor level between commercial norms and Department of Defense (DoD) industrial base norms across large data sets from both markets. A quantitative analysis of a large number of commercial parts (~2.3 million) are compared to a large number of parts from the DoD (over 29,000 line items/1.3 million discrete parts) was conducted to compare how frequently items are single or sole sourced or competitively sourced in each space. This exploratory information gives way to a list of recommendations and future research agenda(s) presented by the authors.

Acquisition regulation and policy in the DoD leans heavily on competition to drive improvements to cost, schedule, and performance. The Competition in Contracting Act (CICA), signed into law in 1984, drives the use of competition as a means to obtain fair and reasonable prices in Government acquisition. In that same year CICA was enacted, “The Japanese Way” was brought into mainstream manufacturing and general business management in the United States (Weiss, 1984). This new way of managing suppliers landed in stark contrast to CICA, encouraging lean manufacturing and closer relationships with smaller contractor pools. Generally, this management style leverages longer duration contracts and partnering with suppliers to achieve cost, performance, and schedule improvements through collaborative improvements versus the constant threat of competitive replacement.

This presents a stark contrast in DoD acquisition policy and commercial management trends. Given this contrast, we should expect the rate of competition with suppliers to be higher in the defense base than in the commercial space. However, reductions in the supply base and less scrutiny placed on CICA compliance for prime contractors, in comparison to Government acquirers, could lower the rate of competition in the defense industrial base.

GAO reports (OUSD/A&S, 2022), the DoD, and news outlets (Gould, 2022) have all pointed to a reduction in competition and number of contractors in the defense industrial base. They primarily assess competition at the prime contract level and with very large subcontracts subject to reporting requirements. The availability of data assessing competition rates for subcontracts and materials purchased by prime contractors is generally more difficult to acquire and analyze than for prime contracts and high dollar subcontractors.

This study provides some indication of the differences in the rate of competitive sourcing in commercial and defense markets with a focus on enriching a more sophisticated understanding of both utilization of competition and collaboration within the contractor base for both sectors. This study finds similarities and differences in both market subcontract competition rates. The DoD sample subcontract competition rates are higher overall, yet discrete programs show as low or lower subcontract competition rates as the commercial market based on our data samples. What follows is a discussion of the history of competition in federal public procurement, how the commercial market is faring in terms of competition, an analysis and discussion of the data and initial propositions and areas for further research to explain this phenomenon.

## **Competition in Contracting Act and Subcontract Competition**

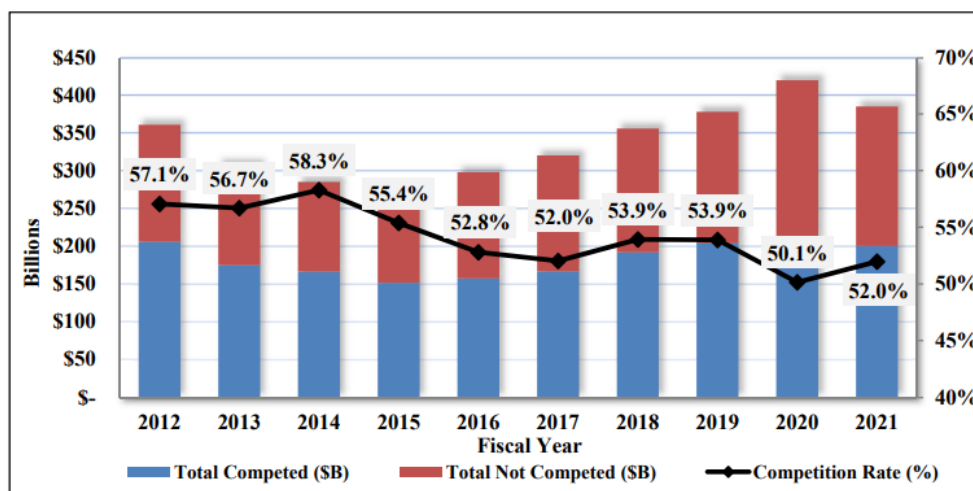
The three primary objectives of public procurement are generally seen to be meeting requirements on time, obtaining value for money, and maintaining public trust (Finkenstadt & Hawkins, 2016; Gilbert et al., 2009). Competition provides more opportunities for public



agencies to meet these goals. Competition has been found to save money, curb cost growth, promote innovation, provide more small business opportunity, provide insights into industrial capability and capacity, and maintain integrity in the expenditure of public funds (Cohen, 1983; OUSD/A&S, 2022). Competition in federal contracting has been around since the early 1800s (Cohen, 1983). Despite this requirement, researchers found that, as of 1982, a majority of federal procurements were completed via non-competitive acquisition (Cohen, 1983).

The Competition in Contracting Act (CICA) was enacted in 1984 in response to these competition concerns based on legislation brought forth by Senators Roth, Levin, and Cohen in 1982. It mandated the use of full and open competitive acquisition procedures unless an exemption was authorized by law. Currently the Federal Acquisition Regulation Part 6 governs competitive mandates for federal acquisition and provides for only seven exceptions to this rule: 1) only one responsible source, 2) unusual and compelling urgency, 3) industrial mobilization, 4) international agreement, 5) authorized by statute, 6) national security, or 7) public interest. Soliciting for full an open competition does not always mean that the government will receive multiple offers, in many cases they only receive one offer to evaluate. This may be the result of market conditions, consolidation activities, or could be the result of the government’s own solicitation methods or requirements definition (GAO, 2010).

Figure 1 from OUSD/A&S’s 2022 report on the competition in the defense industrial base shows the ten-year trend in competition rates in the DoD’s acquisition portfolio. While OUSD/A&S reports this as a “relatively stable” pattern (OUSD/A&S, 2022, pg. 3) it shows two interesting patterns regarding competition: 1) the overall competition rate has been in decline since 2015 and 2) in 2020, the year of COVID’s initial outbreak, we saw the highest dollars spent under the lowest competition rate in a decade. This may mean that we are losing ground on defense competition in general while also being unable to respond to massive supply chain disruptions using competitive procedures.



Note: Dollars shown in billions

Figure 1. Ten Year Trend for DoD Competitive Actions Against the Fiscal Year Budget. (OUSD/A&S, 2022).

While the DoD specifically has focused on increasing competition at the prime contract level with questionable success, it, like many federal agencies, have never really focused on competition below the prime contract level. Subcontract competition is not a primary area of interest for many federal acquisitions. FAR Part 44 does mandate the flow down of a subcontract competition clause for subcontracts that exceed the simplified acquisition threshold, but consent to subcontract and randomized audits are the only practical means the government



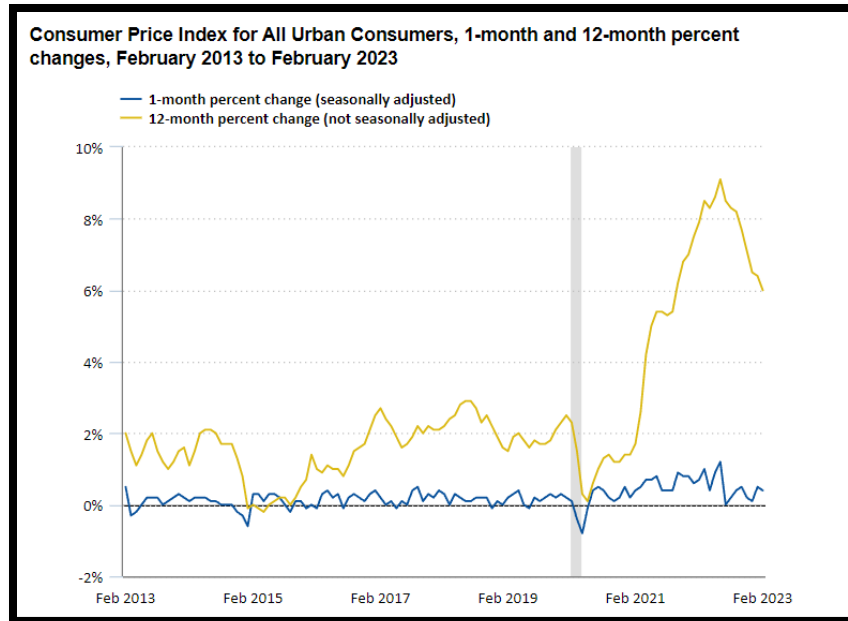
has to enforce that flow down. Consent to subcontract is not expressly required if the contractor has an approved purchasing system and even then, only for cost-reimbursement, time-and-materials, labor-hour, or letter contracts, and also for unpriced actions (including unpriced modifications and unpriced delivery orders) under fixed-price contracts that exceed the simplified acquisition threshold (FAR Subpart 44.2).

Insight into subcontracts in general is poor within the federal government procurement system. Since 2006 federal contractors have been required to submit information about first-tier subcontracts to the Federal Funding Accountability and Transparency Act (FFTA) Subaward Reporting System (FSRS). This reporting was mandated by the FFATA of 2006 and consisted of a phased approach that was to have all subcontracts in excess of \$30,000 reported to the government by October 2015. However, this requirement does not require deeper tier reporting for subcontracts and does not include any requirement to show or report competition at subcontract levels. Therefore, the requirement to flow down competition requirements in certain instances in one part of its regulations consequently misses the opportunity to obtain valuable data on the state of those first-tier competitions via other parts and contractors. Publicly available data on government subcontracts in USASpending.gov is supposed to leverage this data as well. However, our exploration of the data finds it to be extraordinarily poor and inaccurate. As an example, a query of first-tier subcontracts from FY17 to present was conducted in April 2023 and found that the data reported subcontract values in certain years at levels that exceeded the federal budget entirely. As an example, FY20 reports one quintillion dollars in reportable first-tier subcontracts yet only \$670.6 billion in prime obligations.

### **The State of Commercial Competition and the “Japanese Way”**

Competition in commercial exchanges seem to be dwindling as well. The White House Counsel of Economic Advisors released a post in 2021 pointing to market consolidations in the American economy. They list a series of issues such as food packaging market concentrations, domination of commercial air by four major firms and limited to zero localized competition for broadband services for many Americans (Boushey & Knudsen, 2021). Macro-economic evidence is mounting that market power is growing on the part of a small group of consolidated firms such as record high corporate profits during a time of hampered innovation and suppressed wages (Autor et al., 2020; Boushey & Knudsen, 2021). Studies of mergers in commercial markets show that price increases related to mergers can be substantial (Ashenfelter et al., 2014; Kwoka, 2017). They find that the average price effects of mergers are around 7.2%. Figure 2 shows that this price effect is essentially in line with inflation rates since 2021 (BLS, 2023). Unsurprisingly, we have price inflation during a time of mass market consolidation and supply chain disruption. But is a lack of competition always bad for prices? Should we treat all competition at every level the same?





**Figure 2. CPI for All Urban Consumers from February 2013 to February 2023. (BLS, 2023).**

Ironically, in the same year that the federal Government enacted CICA, effectively establishing competition as the primary means of controlling cost, price, performance, and schedule, “The Japanese Way” was brought into mainstream manufacturing and general business management in the United States (Weiss, 1984). Often, business management professionals in the United States focus on lean manufacturing as the core principle of The Japanese Way and, of late some have blamed this principle for supply chain disturbances, like those experienced during COVID, where greater inventories would have prevented disruptions. This however is far from reality, as deployment of The Japanese Way as demonstrated by the Toyota Production System proved during the pandemic that strategic contractor management deployed in combination with lean led to greater resiliency (Shih, 2022).

A critical element of The Japanese Way is the different notion of supplier management. This presents an alternative strategy to one of constant competition among contractors utilized to control supplier pricing. Instead of utilizing short-term contracts leveraging competition to drive supplier behaviors, long-term contracts with smaller pools of contractors and close collaboration with contractors are utilized as the primary means of improving supply base outcomes. By 1995, MIT Sloan Management Review published these findings:

Supplier-customer relationships in the United States are changing rapidly. Where once contracts were short-term, arm’s-length relationships, now contracts have increasingly become long term. More and more, suppliers must provide customers with detailed information about their processes, and customers talk of “partnerships” with their suppliers. (Helper & Sako, 1995)

A traditional Japanese supply system, keiretsu, was used and modified to manage contractors strategically for long-term success. This system is based on trust, cooperation, education and long-term commitment with suppliers. In these cases, contracts are longer in duration with generally ambiguous terms. These contracts mostly enforce specific requirements to reduce cost over time with some share of cost efficiencies retained by suppliers, long-term relationships and cost targets set by leveraging the global marketplace, and procurement of



integrated systems of components (instead of individual parts) to encourage contractor development of high-quality products (Aoki & Lennerfors, 2013).

This system does not fully discard competitive processes. Instead of short-term competitions for individual lots of parts, companies like Toyota use long-term relationships and future products as the primary application of competition. Contractors invest their money, people, and technology development into cost, schedule, and performance improvement for the purpose of securing future long-term contracts.

Many firms view dual sourcing primarily as a way to drive down costs by making suppliers compete in bidding wars. But Toyota takes a very different view: Having two suppliers means it can enjoy resilient capabilities. (Shih, 2022)

Toyota also looks at dual sourcing differently and its approach is very distinct from multi-sourcing that is used by some firms. Multi-sourcing involves buying the same component from multiple sources—say the steering wheel for Model A from contractors 1 and 2. This creates direct competition for a larger share of the requirements for Model A. Generally, Toyota opts to dual source for steering wheels with a different framework. Contractor 1 may produce a steering wheel for Model A and Contractor 2 may produce a steering wheel for Model B. Though there is not direct competition for the current models, both contractors will be indirectly competing to be the preferred contractors for next year's models and for new vehicles under development. This effectively increases resiliency and incentivizes product and manufacturing innovations for future models. At the same time, both contractors can maximize the efficiencies of scale on current production to keep current model costs as low as possible. The massive supply chain problems that have proliferated the news since 2020 has many firms now considering multi-sourcing as a means to increase supply chain resiliency and reduce dependence on overseas markets that may be at risk based on other factors than cost, such as geopolitical or weather-related disruptions. But what does the current data show?

## Data and Methods

We conducted an exploratory analysis of competition at sub-tier levels within commercial and defense markets. Our analysis utilizes data from the Air Force's Life Cycle Management Center Cost and Pricing Division and compared it with commercial data from the Resilinc Corporation, a multi-tiered supply chain mapping firm. Resilinc works with its clients and the clients' contractors to map out multiple tiers of supply networks, to proactively identify and manage risks, to provide continuous monitoring at all tiers of the network, and to use a common communication platform to rapidly respond to and fix disturbances when they occur. They have found that, in collecting the necessary mapping data, it is critical that parts and sites are viewed independently, even when they come from the same contractor. 3M is a good example of why. They manufacture in at least 37 countries with 70 manufacturing sites in the United States alone (3M, n.d.). Supply chain risks vary dramatically by country, whether due to geopolitical, weather, climate or natural disaster risks. Supply chain risks also vary dramatically by sourcing strategies for individual components. Components that are sole sourced, single sourced, or competitively sourced all carry very different risk profiles. Sole sourced components, due to specific nature of use or highly tailored design, can only be produced by an individual contractor and generally carry the greatest risk. Single sourced items can or are produced by more than one contractor but are deliberately sourced from one contractor to leverage greater economies of scale and to simplify contractor management (i.e., the Japanese way). These components generally carry less risk than sole source items. Competitively sourced parts can be acquired from two or more contractors for two general reasons. The first is for parts that are less complex and easily producible by multiple contractors in the marketplace. The second, is to reduce risk for critical



items of supply or to maintain competition where selecting only one contractor, due to highly tailored design, would effectively eliminate the competitive base for future procurement. Competitively sourced components generally carry the lowest risk. By collecting data for more than 800,000 contractors and clients, Resilinc has a very large volume of parts and the sourcing methodology utilized for those parts from the commercial space.

The Air Force Life Cycle Management Center (AFLCMC) executes and manages contracts for a myriad of major weapons systems platforms and programs: manned and unmanned aircraft platforms, command, control, communication, intelligence and networks, battle management systems, digital platforms, major service acquisitions, air munitions, propulsion, cyber security, business systems and multiple Federally-Funded Research and Development Centers (Air Force Materiel Command, n.d.). This includes more than \$45 billion in annual contract obligations for acquisition and sustainment activities of weapons systems platforms and programs. These contracts require cost and pricing data analysis that give some insight into the competitive nature of sourcing for parts below the prime contract level. We utilized data from 11 major sources covering three major weapons programs that fall under two major prime contractors in the DoD industrial base. The two major prime contractors are large businesses, have held and currently hold contracts across multiple services within the DoD including other major weapon systems, and have commercial business segments. The utilized data from bill of material (BOM) data files consisting of more than 29,000 line items and 1.3 million discrete parts accounting for more than \$3.6 billion in material. The BOM data files had between seven and 407 unique suppliers per file. The BOM data files included multiple lot buys for one major weapon system that is also procured and used by other services within the DoD. The utilized data from the AFLCMC did not identify if the subcontract data was for sole source or single source parts. However, it did distinguish parts sourced competitively versus those sourced non-competitively. The basis for the BOM data files were largely determined based on estimating methodologies to include historical data.

We compare rates of competition within BOM data files to Resilinc's supply chain mapping alternative sourcing data. This allows us to compare relative competition and sourcing strategy for DoD major weapons systems with the private marketplace's day-to-day competition strategies for commercial supply chains.

We treat Resilinc's "sole" and "single" supplier data counts as non-competitive data as these suppliers are selected and then utilized for future sourcing for the parts they are assigned. The single sourced parts can be sourced from other contractors, so even without direct competition on individual lots, the threat of contractor replacement creates some competitive cost control. However, we cannot confirm that such pressures exist in our data, so we elected to only treat their "multi-sourced" data as competitive data as a comparison to the DoD data. These parts demonstrate the potential for competitive pressures within each customer supply chain profile. The Air Force data contains labels indicating competition vs. the mix of available sources. The data is either coded with price basis codes that denote competition or directly list parts price basis as "competition." We compare these rates for competitive Air Force parts data within the BOMs to Resilinc's multi-sourced data within their customer repositories.

## **Analysis and Results**

The Air Force data consisted of 11 separate program efforts across three weapons platforms, two major defense primes, and one foreign military sales data set. Table 1 below shows a summary of the number of line items, parts and suppliers by program data file analyzed, how many line items were competitive in nature, the total value of those actions and the amount and percentage of that subcontract bill of materials (BOM) that was based on competitive pricing. The BOM data utilized represents various programs during the course of

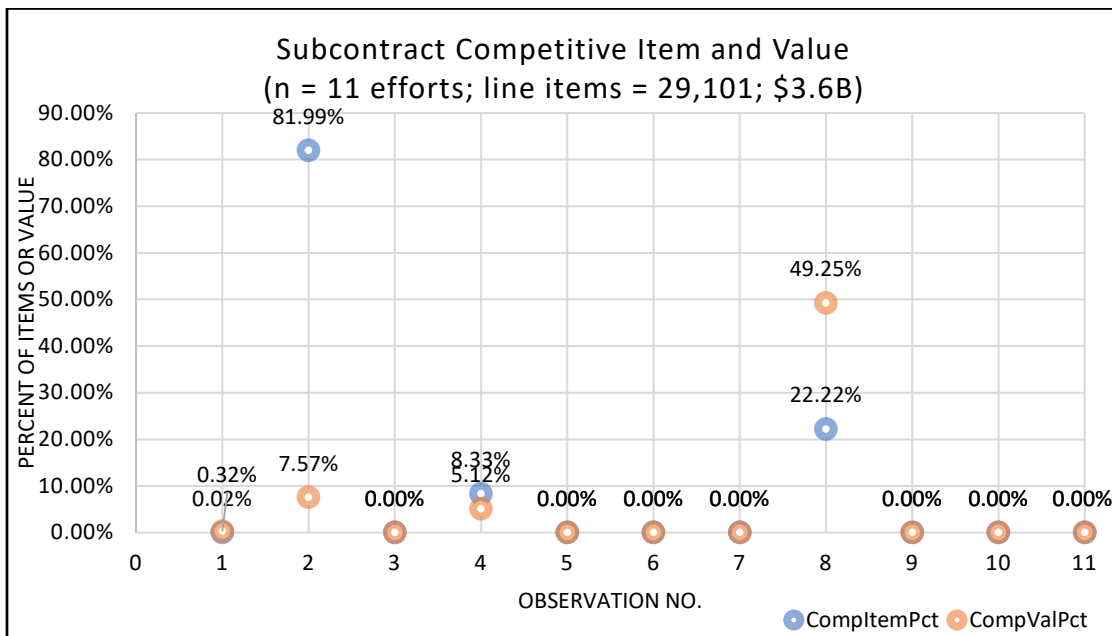




negotiation of contracts, each at different lifecycle stages. As a result, the sourcing strategy listed by part is not perfectly representative of what may have occurred during contract performance or that of prior or future lifecycle stages. The BOM data utilized does represent what the contractor deemed to be a sufficient basis for price reasonableness, which was overwhelmingly non-competitive. Some of the data is from Undefined Contract Actions (UCAs) where performance and procurement for the effort was underway during negotiations, in these instances the BOM data does reflect actual performance. The 11 programs have been numbered with names removed to protect the anonymity of the firms. As can be seen in Table 1 and Figure 3, most of the items were acquired using non-competitive means and a majority of the program values are not based on competitive procurements. Figure 3 provides a scatter plot of the relative percentage of competitive line items and value by each of the 11 program initiatives analyzed.

**Table 1. Air Force Program Analysis Summary Data**

Prgm Effort	Suppliers	Parts	Line Items	Competitive	CompPct	TotalVal	CompVal	CompValPct
1	407	652781	12780	2	0.02%	\$ 486,708,547.00	\$ 1,560,695.00	0.32%
2	212	19360	3703	3036	81.99%	\$ 47,434,562.21	\$ 3,591,030.45	7.57%
3	135	48064	1215	0	0.00%	\$ 357,927,004.70	\$ -	0.00%
4	12		12	1	8.33%	\$ 15,971,026.00	\$ 817,655.00	5.12%
5	124	19697	4267	0	0.00%	\$ 156,155,972.00	\$ -	0.00%
6	7		8	0	0.00%	\$ 36,351,733.00	\$ -	0.00%
7	113	8587	1195	0	0.00%	\$ 186,930,689.00	\$ -	0.00%
8	26		54	12	22.22%	\$ 782,950,000.00	\$ 385,610,000.00	49.25%
9	160	472722	3435	0	0.00%	\$ 829,173,269.00	\$ -	0.00%
10	123	90506	2384	0	0.00%	\$ 392,424,764.04	\$ -	0.00%
11	25		48	0	0.00%	\$ 353,623,000.00	\$ -	0.00%
<b>Total (Avg for Pct)</b>	<b>1344</b>	<b>1311717</b>	<b>29101</b>	<b>3051</b>	<b>10.23%</b>	<b>\$ 3,645,650,566.95</b>	<b>\$ 391,579,380.45</b>	<b>5.66%</b>



**Figure 3. Subcontract Competitive Percentages from AFLCMC Program Bill of Materials**

Across a sampling of the 11 program initiatives, we see a low competitive value percentage (equal to total dollars in competitive awards divided by the total value of all costs in



the BOM). There is one outlier that had 49% of total BOM value competed. It was of smaller item counts (only 54 items) but was for the second largest BOM cost listed at \$782.9 million (roughly 10.6% of observed costs in this study). Further analysis would be required to determine why this one spare buy for a large, new weapons platform in an earlier life cycle stage utilized more competitive sourcing than others, especially given that it was by a prime and program that utilized very little competitive buying at the subcontract level for other initiatives analyzed. We are aware that it was for a program on UCA that was experiencing major cost control issues.

The Resilinc data includes part sourcing labeling for 2,308,781 parts. This data spans industry verticals including High Tech, Consumer Electronics, Life Sciences, Medical Devices, Semiconductors, Pharma, Auto, Industrial, Healthcare, and Aerospace. In total, very little direct competition is utilized in the commercial space, with only .84% of components multi-sourced. Additionally, very few sole source parts are present (.63%) indicating that even when direct competition is not utilized, the ability to replace contractors is likely. We cannot determine this level of alternative sourcing from the Air Force data. The very high degree of single-sourced parts (98.53%) in Resilinc data is indicative of The Japanese Way utilized in the commercial space, particularly as it relates to moving away from direct competition as the primary means of controlling cost.

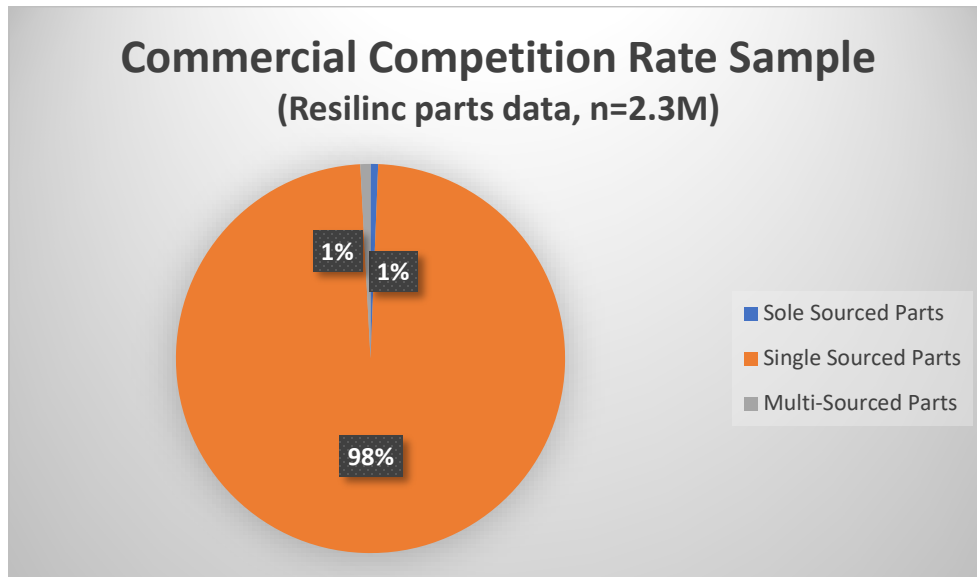


Figure 4. Commercial Competition Sample from Resilinc System

Table 2. Resilinc Parts Data

Sole Sourced Parts	14,513	0.63%
Single Sourced Parts	2,274,801	98.53%
Multi-Sourced Parts	19,467	0.84%
<b>Total Parts</b>	<b>2,308,781</b>	
Total Non-Competitive	2,289,314	99.16%
Multi-Sourced Parts	19,467	0.84%
<b>Total Parts</b>	<b>2,308,781</b>	



## Discussion

The challenges of today and emergent threats of tomorrow continue to “transcend all boundaries and limits” (Liang & Xiangsui, 2020). These challenges will further shape the modernization of existing major weapons systems and the acquisition of future ones. The current acquisition landscape in the DoD comprised of low rates in competitive sourcing and resultant higher weapon system life cycle costs translates to lost capability that could be used to further increase our lethality and readiness. Acquisition strategies and teams within the DoD will need to consider competitive sourcing, especially for long-term acquisitions, at both the prime and sub-level in order to make capability (value) driven decisions. Likewise, contractors in the defense industrial base must also adapt to commercial approaches with strategic investments in their supply chains to maintain a competitive edge and create value. Sustainment costs alone within the DoD continue to represent an unsustainable “70% of the average weapon system’s total life cycle cost” (Serbu, 2022) and largely consist of non-competitive contracts including some contractor-locked major weapons systems. Moreover, matching the pace of our adversaries in modern warfare demands the mix of contractors within the defense industrial base also transcend traditional boundaries to deliver both agile and superior capabilities. This also requires services across the DoD to enhance traditional knowledge, skills, and abilities to foster enterprise-wide innovative and value-driven strategies capable of acquiring and sustaining major weapons systems at the “speed of relevance” (Dowling & Johnson, 2019). Essentially, knowing and acting when necessary, versus a day late and a dollar short.

In 2021 the GAO reported that estimated F-35 aircraft operation and sustainment costs across the services had increased to nearly \$1.27 trillion over 66 years—roughly \$6 billion beyond affordability thresholds—and further found “cost reductions become increasingly difficult as the program grows and matures” (GAO, 2021a). Consideration of future sustainment costs are generally not transparent or accurate when acquiring a major weapons systems leading to additional total life cycle costs and lost capability. One way to potentially reduce risk in future long-term major weapons systems acquisitions could be to adopt the traditional Japanese supply system, keiretsu, by incentivizing greater competition and sharing in cost efficiencies with the prime-level contractor. This could also further increase supply chain resiliency and diminishing manufacturing source(s) during sustainment. The first step in managing the defense industrial base is to discover what it looks like. Without better sub-tier visibility and more standardized reporting procedures we will be unable to systematically capture, track and take action on the competition rates below the prime contract level.

The DoD is examining “sustainment health strategies” (Serbu, 2022) and other initiatives to improve sustainment costs through competition in future acquisitions. Initiatives currently discussed include securing intellectual property at the forefront. Securing intellectual property would provide continuous opportunities for greater competition and strategic sourcing, in whole or part, throughout a weapon system’s life cycle and thus have a downward pressure on total life cycle costs. However, while it may sound appealing this approach could also inadvertently result in increased costs during the acquisition of major weapons systems as contractors try to recoup revenues typically realized during sustainment associated with IRAD or other proprietary investments.

The USAF Collaborative Combat Aircraft (CCA) program (Tirpak, 2022), an autonomous collaborative platform variant, initiated by Air Combat Command (Hadley, 2022) and captured in the Honorable USAF Secretary Frank Kendall’s “Operational Imperatives” (Department of the Air Force, n.d.) is another initiative that could prove successful in placing a downward pressure on a weapons systems total life cycle cost and increase competition. Shorter life cycle periods would naturally result in an increased frequency of requirements being competed.



Our data analysis shows that some defense program efforts exceed the competition rates seen in the commercial market, while others demonstrate similarly low rates of competition below the prime level. As stated earlier, there does not seem to be a high rate of direct competition to create cost control in the commercial space, with single-sourced parts used in 98.53% of our sample. We did find that the data fidelity and analysis for DoD data was far lacking compared to commercial data. The DoD needs to start tracking single vs. sole to understand the amount of indirect competition present. The DoD should consider whether CICA has become outdated and if better contractor engagement, longer term contracts, and partnership in the contractor base could provide better long-term outcomes. Additionally, absent this type of engagement, it is important to consider how adoption of lean inventory management provides an incomplete adoption of The Japanese Way and introduces threats to resiliency in the defense industrial base. This study does offer interesting insights into how we monitor and analyze this data. This study found that analyzing discrete BOMs for competition is extremely difficult as none of the programs used similar data labeling and structural methodologies. Further, open-source data on federal and defense subcontracts is lacking and inaccurate and oversight has paid sufficient attention to subcontract data reporting issues in open-source data (GAO, 2014, 2021b, 2021c).

### **Limitations and Areas for Further Research**

We admit that our analysis is a simple exploratory view of DoD and commercial markets. However, the samples do cover a large number of parts and spend which can offer direction for future research in this area.

#### **Limitation 1: The DoD data only included current Air Force programs**

Future Research: Future studies should explore similar trends in subcontract competition rates for other services and departments to determine if low competition rates and data issues are systemic to the federal government, DoD, or agencies themselves. Future research should also explore whether subcontract competition data is richer and more prevalent when new programs are being competed.

#### **Limitation 2: DoD data does not contain sufficient details to analyze the differences in rates of sole vs. single sourced procurements**

Future Research: Further research is needed within the DoD to assess whether single and sole source decisions in the supply base are tracked. Competition methods and sourcing during weapons systems life cycles should also be assessed given that diminishing manufacturing sources for parts/items generally increase toward the end of a weapon systems life cycle. This would improve the analysis of the state of sourcing in the defense base. If not being tracked, the DoD should start tracking single vs. sole sourcing to assess the supply chain management practices in the defense base and determine if more modern contractor management techniques should be instilled into large defense contractors through more hands-on management by Government personnel, through contractual terms, or potentially changes to current regulations. For example, DoD contractors are already required to maintain a database for each part number to include the name of the supplier, amongst other items. The increased administrative cost related to the monthly or quarterly delivery of this data would be insignificant compared to how it could better shape future acquisition strategies. With the current DoD drive for innovation, it is important to understand if the Toyota Production System's proven method of utilizing single sources for different components to exploit production scale, enhance contractor performance through indirect competition, and incentivize product and manufacturing innovation is part of the current DoD contractor management framework.



### **Limitation 3: This analysis does not contain sufficient qualitative data to determine why we see certain patterns of competition in commercial and defense data.**

Future Research: The DoD should consider whether CICA has become outdated and if better contractor engagement, longer term contracts, and partnership in the contractor base could provide better long-term outcomes. Additionally, absent this type of engagement, it is important to consider how adoption of lean inventory management provides an incomplete adoption of The Japanese Way and introduces threats to resiliency in the defense industrial base. As noted earlier, the program data with the highest level of competition in our sample was for a major weapons system that was facing incredible costs pressures. Further, analysis is needed to determine if the high levels of competition were a result of these cost control issues or a cause.

The DoD states that competition matters and improves outcomes. However, roughly half of the defense budget is executed using non-competitive prime awards. If the DoD truly believes competition matters perhaps they should gain a higher fidelity view of it at sub-tiers to truly understand the defense industrial base, its competitive pressures and how they lead to variations in mission outcomes. But, again, step one is visibility. You can't measure what you can't find, and our initial analysis shows that the DoD is walking in the dark when it comes to subcontract data.

### **References**

- 3M. (n.d.) 3M central Europe region—Companies and Sites. <https://multimedia.3m.com/mws/media/1576821O/zahlen-und-fakten-englisch.pdf>
- Air Force Materiel Command. (n.d.). *Air Force life cycle management center*. <https://www.afmc.af.mil/About-Us/Fact-Sheets/Display/Article/2218380/air-force-life-cycle-management-center/>
- Aoki, K., & Lennerfors, T. T. (2013, September). The new, improved keiretsu. *Harvard Business Review*. <https://hbr.org/2013/09/the-new-improved-keiretsu>
- Ashenfelter, O., Hosken, D., & Weinberg, M. (2014). Did Robert Bork understate the competitive impact of mergers? Evidence from consummated mergers. *The Journal of Law & Economics*, 57(S3), S67–S100. <https://doi.org/10.1086/675862>
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020, May). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2), 645–709. <https://doi.org/10.1093/qje/qjaa004>
- Boushey, H., & Knudsen, H. (2021, July 9). The importance of competition for the American economy. *Counsel of Economic Advisors Blog*. <https://www.whitehouse.gov/cea/written-materials/2021/07/09/the-importance-of-competition-for-the-american-economy/>
- Bureau of Labor Statistics. (2023, March 20). *Consumer Price Index up 0.4 percent over the month, 6.0 percent over the year, in February 2023*. <https://www.bls.gov/opub/ted/2023/consumer-price-index-up-0-4-percent-over-the-month-6-0-percent-over-the-year-in-february-2023.htm>
- Cohen, W. S. (1983). The competition in contracting act. *Public Contract Law Journal*, 14(1), 1–39.
- Department of the Air Force. (n.d.). *Operational imperatives*. [https://www.af.mil/Portals/1/documents/2023SAF/OPERATIONAL\\_IMPARITIVES\\_INFOGRAPHI C.pdf](https://www.af.mil/Portals/1/documents/2023SAF/OPERATIONAL_IMPARITIVES_INFOGRAPHI C.pdf)
- Dowling, J., & Johnson, R. (2019, December 11). *Learning at the speed of relevance*. Defense Acquisition University. <https://www.dau.edu/library/defense-atl/blog/Learning-at-the--Speed-of-Relevance>
- Finkenstadt, D. J., & Hawkins, T. G. (2016). #eVALUate: Monetizing service acquisition tradeoffs using the quality-infused price© methodology. *Defense Acquisition Research Journal*, 23(2), 202–230.
- GAO. (2010). *Opportunities exist to increase competition and assess reasons when only one offer is received* (Report No. GAO-10-833).
- GAO. (2014, June 30). *Data transparency: Oversight needed to address underreporting and inconsistencies on federal award website* (GAO-14-476).



- GAO. (2021a, July 7). *F-35 sustainment: DoD needs to cut billions in estimated costs to achieve affordability* (GAO-21-439). <https://www.gao.gov/products/gao-21-439#:~:text=Publicly%20Released%3A%20Jul%2007%2C%202021,has%20steadily%20increased%20since%202012>
- GAO. (2021b, November 8). *Opportunities exist to further improve the information available on USAspending.gov* (GAO-22-104702).
- GAO. (2021c, December 16). Opportunities exist for treasury to further improve USAspending.gov's use and usefulness (GAO-22-104127).
- Gilbert, D., Schapper, P. R., & Veiga-Malta, J. N. (2009). Framework for the management and reform of public procurement. In K. V. Thai, *International handbook of public procurement* (Chap. 4). CRC Press.
- Gould, J. (2022, April 12). Kathleen Hicks warns of "substantial decline" in defense industrial base competition. *Defense News*. <https://www.defensenews.com/pentagon/2022/04/12/kathleen-hicks-warns-of-substantial-decline-in-defense-industrial-base-competition/>
- Hadley, G. (2022, December 14). Air Force leaders: CCA is about capability not just cost. *Air & Spaceforces Magazine*. <https://www.airandspaceforces.com/air-force-leaders-cca-is-about-capability-not-just-cost/>
- Helper, S. C., & Sako, M. (1995, April 15). Supplier relations in Japan and the United States: Are they converging?. *MIT Sloan Management Review*. <https://sloanreview.mit.edu/article/supplier-relations-in-japan-and-the-united-states-are-they-converging/>
- Kwoka, J. (2017). The structural presumption and the safe harbor in merger review: False positives or unwarranted concerns? *Antitrust Law Journal*, 81(3), 837–872. <http://www.jstor.org/stable/26425580>
- Liang, Q., & Xiangsui, W. (2020). *Unrestricted warfare* (pp. 4–11). Albatross Publishers.
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2022). *State of competition within the defense industrial base*. Department of Defense. <https://media.defense.gov/2022/Feb/15/2002939087/-1/-1/1/STATE-OF-COMPETITION-WITHIN-THE-DEFENSE-INDUSTRIAL-BASE.PDF>
- Serbu, J. (2022, November 1). *Pentagon plans new initiatives to tackle long term sustainment costs*. Federal News Network. <https://federalnewsnetwork.com/defense-main/2022/11/pentagon-plans-new-initiatives-to-tackle-long-term-sustainment-costs/>
- Shih, W. C., (2022, November 15). What really makes Toyota's production system resilient. *Harvard Business Review*. <https://hbr.org/2022/11/what-really-makes-toyotas-production-system-resilient>
- Tirpak, J. A. (2022, October 13). ACC and USAF HQ are crafting collaborative combat aircraft requirements. *Air & Spaceforces Magazine*. <https://www.airandspaceforces.com/acc-and-usaf-hq-are-crafting-collaborative-combat-aircraft-requirements/>
- Weiss, A. (1984, July). Simple truths of Japanese manufacturing. *Harvard Business Review*. <https://hbr.org/1984/07/simple-truths-of-japanese-manufacturing>



# Optimizing Operations and Logistics Support Using Opus Evo

**Gustaf Solveling**—is the principal analyst and senior software developer at Systecon North America with more than 10 years of analytics and software development experience. Solveling has 10 years of logistics support system optimization experience as consultant and decision support system developer and has served as a project manager and technical leader for software development projects, integration projects, and data science projects. He has a PhD in Industrial Engineering and a Master of Science in Operations Research from Georgia Institute of Technology and a Master of Science in Communication and Transportation Engineering from Linkoping University. [Gustaf.Solvering@systecon.us]

**John Verbanick**—is a senior analyst with three years' experience at Systecon North America. He served 15 years as a Weapon Systems Officer in the United States Air Force. After leaving the military Verbanick went to work as a Manufacturing Engineer supervisor with Honeywell Aerospace, working on manufacturing process control, supply chain improvements, and process improvements based off Six Sigma analysis. He has a Bachelor of Science in History from the United States Air Force Academy and a Master of Military Operational Arts and Science from the Air Command and Staff College. [John.Verbanick@systecon.us]

## Abstract

Understanding the interplay between the reliability and maintainability of a fleet of complex systems, the logistics support organization, and the operational scenario, is vital from a short-term tactical perspective, as well as a strategic long-term Life Cycle Cost (LCC) perspective, as each of these areas has a direct impact on one another. A common method to analyze and evaluate the performance of the overall scenario, as well as getting insights into problem areas, bottlenecks, and to perform analysis-of-alternatives is to use discrete event simulation.

In this paper we present a methodology to extend a discrete event simulation tool with inherent optimization capabilities. Using established heuristic optimization techniques, we perform simulation driven optimization that optimizes parameters in the modeled scenario. Optimized parameters typically include:

- sparing strategies such as inventory levels and locations
- resource quantities and location
- deployed system quantities to fulfill mission requirements
- scheduled maintenance times.
- transportation and resource schedules.

A case study is presented that utilizes Opus Evo, an application that extends the commercial off the shelf Opus Suite with capability to perform heuristic optimization using simulation.

**Keywords:** Heuristic Optimization, Tactical Logistic Planning

## Introduction

Predicting and optimizing mission capability and readiness for a system requires knowledge and data in a range of areas that each have a direct impact on the outcome. The reliability of the components making up the system, the maintainability of the system, the responsiveness of the support organization, and the operational tempo are all factors that contribute to, or inhibit, readiness. Furthermore, mission capability and readiness are always associated with a cost and understating the relationship between cost and readiness is important, especially when optimizing readiness given budget constraints.

To represent the modeled scenario it is important to have a suitable domain model. An appropriate domain model simplifies data entry, promotes the understanding of the model, is compatible with established standards, etcetera, but in the end, the data in the domain model is



used for analytics which provides insights and recommendations. In the Domain Model section we present characteristics of a domain model that can be used for evaluation and optimization of system availability and mission readiness.

Opus Suite is a suite of software applications that is used for predictive analytics of complex technical systems together with its operational characteristics and its logistics support network. One of the core applications in the suite is SIMLOX, a discrete event simulation tool for predicting mission performance and readiness over time. The objective of a simulation tool is typically to evaluate the behavior of a system given stochastic parameters and stochastic dynamics, where the analysis is limited to simulation of one scenario at a time.

In this paper we present a method that enables *optimization* of any entity in the domain model. Enabling optimization makes it possible to not only analyze and evaluate the performance of the overall scenario, but also identify improvements in a systematic way without a manual “trial and error” process. The method is based on evolutionary algorithms, and we show how any numeric data element in the domain model can be optimized. In the next section, the domain model is introduced with examples on what can be optimized using the proposed methods. Following that, the application in which the method has been implemented is presented together with an algorithm description. In the last section we present a case study where the application, which goes by the name of *Opus Evo*, has been utilized to optimize a pack-up kit for deployed operations of an aircraft system. The methodology presented in this paper is general in nature and does not depend on a specific set of tools or applications. For the proof of concept and for the case study, the Opus Suite has been utilized to provide a domain model and optimization evaluator.

## Domain Model

A domain model to support mission readiness and system availability optimization requires representation of data in a number of categories. Examples of the categories are:

- Product breakdown
- Reliability
- Task
- Corrective and preventive maintenance event
- Maintenance capabilities
- Operation profiles
- Functional breakdown/reliability block
- Mission characteristics
- Inventory
- Resources

Within each category there are typically several entities with associated attributes. In general attributes can be of any data type, but for use with the evolutionary algorithm presented in the next section, attributes to be optimized need to be numerical values. An example of an entity that can be optimized is seen in Table 1, where a typical objective is to maximize readiness by determining inventory levels and locations subject to a budget constraint.





**Table 7. Entity in Domain Model**

Entity:	Inventory Level
Key:	Item Identifier
Key:	Location
Attribute:	Nominal Stock Level
Attribute:	Item Cost
Attribute:	Storage Cost

As will be seen in subsequent sections, the proposed methodology enables optimization of any entity in the domain model. Examples of scenarios that can be optimized are listed below. In all examples below, the objective is to maximize mission capability and readiness.

- Maximize mission capability by optimizing inventory levels subject to budget constraints.
- Maximize mission capability by optimizing resource quantities and locations subject to budget constraints.
- Maximize mission capability by optimizing the mix of inventory vs. resources given budget constraints.
- Minimize deployed system quantities while achieving a specified mission capability level.
- Maximize mission capability by optimizing the time of maintenance given specified maintenance windows.

## Technical Solution

### Evolutionary Algorithms

Evolutionary algorithms are heuristic optimization algorithms inspired by processes in nature. The algorithms are applicable to many optimization problems, as the algorithms typically only require evaluation of the objective function, often referred to as the fitness function, to determine the quality of a single solution for the problem at hand. Although the principles of all evolutionary algorithms are the same, the methods performed in the general steps may differ, which creates several types of evolutionary algorithms. Due to the nature of the optimization problems represented in the domain model, the evolutionary algorithm type that have been tested and implemented is *differential evolution*. Differential evolution is especially suited for problems where the variables to be optimized are numerical values (in contrast to binary values typically used for genetic algorithms).

The basic algorithm steps of all evolutionary algorithms are (Simon, 2013):

1. Randomly generate the starting sample set
2. Evaluate the fitness function of all samples
3. Select the best samples to keep for reproduction (parents)
4. Combine and create new samples from the parents (offspring)
5. Replace least fit samples with new offspring (survival of the fittest)
6. If termination criterion is not reached, go back to step 2, otherwise terminate, and return the sample with the best fitness as the solution.

The calculations specific to differential evolution take place in steps 4 and 5 in the algorithm above, where new samples are generated according to the following procedure (Price et al., 2005):

1. For each sample,  $\bar{x} \in \mathbb{R}^n$ , in the population, create a new sample,  $\bar{y} \in \mathbb{R}^n$ , according to:
  - a. Select three samples from the population, distinct from  $\bar{x}$  and from each other.



Call these  $\bar{a}$ ,  $\bar{b}$ , and  $\bar{c}$ .

- b. Determine a subset  $J \subseteq \{1, \dots, n\}$ , such that  $|J| \geq 1$  (at least one in the vector dimension will change).
  - c. For each  $j \in J$  let  $y_j = a_j + F \times (b_j - c_j)$
2. If the fitness function applied to  $\bar{y}$  gives a better value than for  $\bar{x}$ , replace  $\bar{x}$  with  $\bar{y}$  in the set of samples.

The subset  $J$  in step 1b is determined using a constant called *crossover probability* and the mutation in step 1c uses a constant  $F$  called *crossover factor*. For further details on the differential evolution algorithm, see Price et al. (2005).

The standard differential evolution algorithm is extended with a taboo list so that samples whose fitness function value is already known do not need to be evaluated a second time.

## Opus Evo

The application in which the proposed methodology has been implemented is referred to as *Opus Evo*. Opus Evo is an application within Opus Suite that enables optimization of any attribute, or combination of attributes, within the domain model. The optimization is performed using the differential evolution algorithm presented above. The application is made up of several components, which are listed below. For each component, a general description is provided together with additional details that are particular to Opus Evo.

*The core differential evolution algorithm:* Given a representation of the problem in form of a sample and a fitness function the algorithm will find a solution whose fitness function value is not worse than that of any other solution discovered. A global optimum cannot be guaranteed for evolutionary algorithms. Note that the algorithm itself does not need to know anything about the problem being optimized, other than the representation of a sample and what fitness function to use. The steps of the algorithm are described in The Evolutionary Algorithms section.

*A domain representation, problem data, and data storage:* This is where the problem that is being solved is modeled. A general domain model is described in the Domain Model section. In Opus Evo, the existing Opus Suite domain model and data storage is leveraged. Thus, a model instance that already exists in Opus Suite can be optimized in Opus Evo.

*A domain to vector mapping:* The evolutionary algorithm requires the optimized attributes in a vector representation, but the algorithm does not need to know what the values in the vector represent. To achieve this, it is necessary to have a mapping from entity attributes in the domain model to the internal vector representation. In Opus Evo, the mapping of entity attributes to the vector representation is provided through a text file specified by the user. In the variable declaration, bounds on the variables can be provided. The mapping is necessary to translate the two representations, e.g., when evaluating the fitness function or when to present the best solution to the end user.

*A fitness function:* The purpose of the fitness function is to determine the fitness, or quality, of a given solution. In the context of the evolutionary algorithm, a fitness function is a black box, which takes a vector of values as input, and return a single- or multi-dimension objective value as output. Opus Evo uses the simulation tool SIMLOX for fitness evaluation. The optimization algorithm can use any fitness function, but using SIMLOX has several advantages:

- SIMLOX is an evaluator/function that already can handle all aspects of the domain model. Thus, the same fitness function can be used regardless of what entities being optimized.



- The result of SIMLOX is mission capability or system availability, which is typically the desired objective for an optimization in this domain.
- Discrete event simulation works well with distributed computing, where evaluation of individual samples can be distributed. Furthermore, replications within a simulation run can be distributed and processed in parallel.

The fitness function can support multi-dimensional objectives, so it is possible to simultaneously optimize different dimensions (e.g., cost and readiness). However, it has been observed that the algorithm progress to an optimal solution quicker if a single objective is considered.

*A computing resource orchestrator.* Evaluation of samples during an iteration of the algorithm can be distributed and performed in parallel, as each sample is independent from all other samples. Thus, the performance of the algorithm scales linearly with the number of processors available. Opus Evo includes support for distributed processing using the Message Passing Interface (MPI; Microsoft, 2022). For the case study presented in the next section, a cluster with a total of 70 logical processors distributed over two servers was utilized.

The domain representation and the fitness function are presented in the context of Opus Suite, but the application is general and any domain representation that meets the criteria in the Domain Model section can be used, as well as any fitness function that is able to evaluate a solution can be used.

## Case Study Air Force Deployment

The case study being presented deals with looking at optimizing spare parts with requirements outside the typical cost vs. availability tradeoff. This problem required transporting equipment to a new location along with all the spare parts for that equipment. However, this location has limited access and constraints in space available to set up the proposed operation. Given these parameters the problem further required that the equipment to be operational over a 20-day period, given the requirement that the availability rate exceeds 98% of the desired operating window.

For this problem, an initial package of spares had already been created with an operational design to last for a 20-day period while providing a high availability rate, but these off-the-shelf packages are based on an equipment usage rate twice that of what is expected over this proposed excursion. The packages also assume that no restrictions exist based on transportation and storage space for the required parts. A typical modeling approach would provide an optimized solution based on a cost vs. availability curve. With this solution curve, further analysis would be required to find and isolate the numerous solutions that do not meet an optimal curve point. With this problem, the introduction of a new factor to analyze becomes necessary, in this problem that constraint is the dimensional data of all the spare parts that are modeled. Typically, the modeling solution focuses solely on maintenance significant items (MSIs), or those items which have been determined to be necessary to keep the equipment in a maintenance up and operational status based off historical failure rate analysis of all the components of the equipment.

Addressing this problem within the newly defined constraints, one must start with a basic cost vs. availability model of the system. SIMLOX is designed to perform this analysis and an existing model of the equipment existed that could serve as a baseline for the new solution. From this baseline one can take the provided solution, utilize the Opus Evo application, and run the model with the dimensional restrictions to develop a solution that becomes acceptable to the requirements and limitations that were laid out.



The case study involved the movement of 24 pieces of equipment with each typically requiring a full transportation unit to move and have available. The transportation cost of 24 transport units was determined to be too much and unacceptable to the problem's end state needs. The problem had another complication in that the new location would not be able to accommodate the footprint of equipment that these transportation units would bring with them. As such, a determination and innovative approach is required to achieve operational success which would provide capability while reducing the dimensional limitations that are typical given the movement of equipment.

The requirement that the equipment have the capability to operate every day for the 20-day period with four systems in operation at any given time while those not in operation requiring a two-hour pre-operating period and two-hour post maintenance period provided an added challenge to the problem set. A total of 1,200 hours (about 20 days) of actual operational coverage provided by the equipment is necessitated to achieve operational success. These restrictions are half of what the off-the shelf spares package provides, so the overall question became what is required to achieve success while reducing the maintenance and storage footprint to the max extent possible?

Typically, if one were to create and run this model, it could be accomplished through traditional modeling means. However, this approach adds a cost associated which is the time to compute. As discussed, to find the best fit solution you would first optimize the model and then take that solution as a baseline. Once this baseline is established you can proceed in one of two ways. The first solution requires a degree of time through trial-and-error, running iterations with slightly different data points to narrow down the result into one that fits the operational requirements. This requires looking at the data from each iteration and performing calculations on the results outside modeling software. The impact of this is after every modeling run the modeler is required to choose outcomes that bring the solution in line with requirements. The added time of this becomes excessive to the modeling process. As one can imagine, this considerable time cost to run the iterations and analyze the results to find the best solution for the problem makes meeting a shortened timeline unacceptable to requirements.

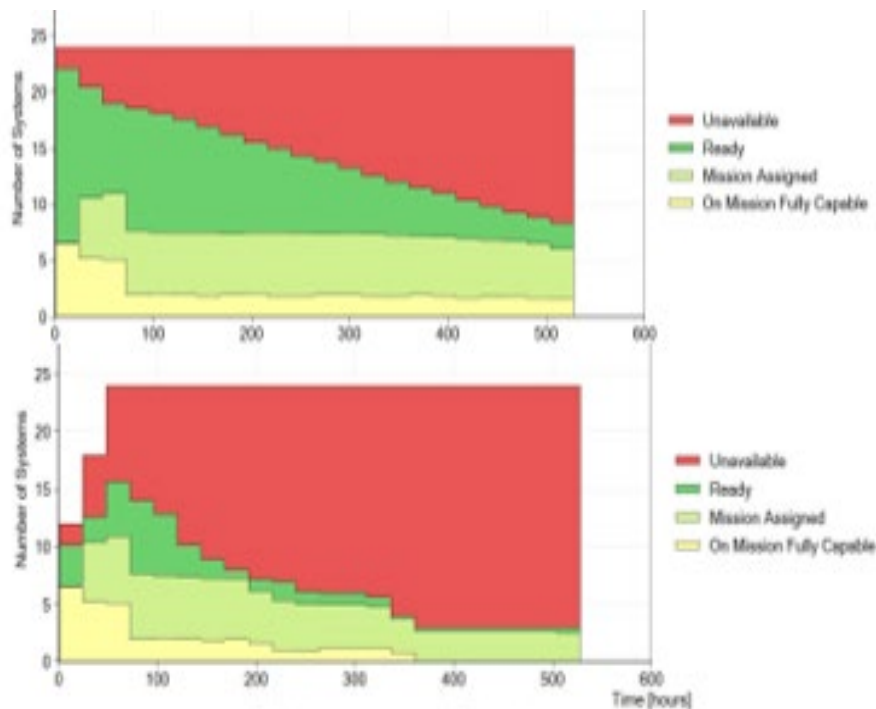
The second option builds off the first, but it is to run two additional iterations to the baseline model and then narrow down the solution set utilizing interpolation of the results to find a working solution, this would be useful to save time in a smaller model, however the size and scope of this problem makes this prohibitive in application, requiring the modeler to make assumptions that may or may not be realistic to the problem at hand. This approach, like the previous requires taking the results from three runs and interpolating the results to find an acceptable solution. This approach brings artificiality into the modeling process and does not guarantee an optimized result.

For this problem, the baseline model requires approximately eight hours of computing time, with each follow-up iteration requiring the same time cost. This requires the modeler to spend at a minimum 24 hours, and possibly days of computational time looking at different iterations and factors. Computational time alone is an issue; however, the time required to analyze the results and determine a workable solution could take additional weeks to sort. The time sink that these approaches bring make these two solutions unworkable. If time is not a factor, one can brute force and with some luck achieve an acceptable solution in a matter of weeks/months. However, for this problem the requirement is a working solution within days. The timeline for this problem is thus prohibited in the traditional modeling capability and at best using these means one can attain a rough solution and perform a post event analysis between our solution and what they utilized.



With Opus Evo one can run numerous iterations over a brief period with the new constraints utilizing heuristic simulation on the baseline model. The new tool allows one to use the baseline model and then utilize machine learning and the power of heuristic simulation provided through Opus Evo. This tool allows different iterations of the baseline model to be run concurrently, utilizing a third factor outside the traditional cost vs. availability. Given an appropriate expanded data set, one can utilize this approach and optimize a model based off any third factor. With the right data this software permits the modeler to develop an optimized solution in a fraction of the time that it would take to perform the task manually. One can now utilize this capability for the problem of creating a solution meeting their requirements in both system availability, reduced maintenance, and storage footprint in a shortened timeline. To validate the solution in both effectiveness and time indicators, the results were compared between the modeled solution and the planned solution the problem utilized to determine if the new Opus Evo software indeed could optimize the problem while meeting the operational requirements.

When we compared the modeled operational results with those provided as the user solution, the modeled results produced a stark contrast in system availability, number of parts required, and cost associated with the scenario. The top chart in Figure 1 displays that utilizing Opus Evo the modeled solution ensured a greater number of systems available and in operations throughout the entire 20-day (480-hour) window whereas the solution provided Figure 1 bottom chart shows the equipment becoming unsustainable at the 15-day(360-hour) point of their operations.



**Figure 1. Simulation Results Opus Evo on the Top; Customer Solution on the Bottom**

The contrast in solutions shows a distinct advantage to utilizing Opus Evo. The modeled solution on the top shows that system availability was maintained at eight systems at any given time throughout the entire 20-day period. Whereas the solution we were given drops to three systems available at the 15-day point. This savings in capability alone shows that in modeling through Opus Evo, we can optimize the availability of the systems and meet the problem requirements of four systems in operation at any given time, while also enabling an added surge



capacity in usage. The savings in this aspect alone from the modeled solution over the provided solution is enough to prove out the software capabilities.

To further prove the ability of Opus Evo, the ability to meet the problem requirements, in post action analysis it was determined that the problem solution planned to under deploy 55 MSI parts as well as over deployed by 296 parts over what the modeled solution provided. In total by utilizing Opus Evo, it was determined that one could meet the requirements with a 38% reduction in parts, a 56% reduction in cost, and a 50% increase in equipment availability.

## References

- Microsoft. (2022). *Microsoft MPI*. <https://docs.microsoft.com/en-us/message-passing-interface/microsoft-mpi>
- Price, K., Storn, R. M., & Lampinen, J. A. (2005). *Differential evolution: A Practical approach to global optimization*. Springer.
- Simon, D. (2013). *Evolutionary optimization algorithms*, Wiley.



## PANEL 17. SOFTWARE ACQUISITION PATHWAY

Thursday, May 11, 2023

10:30 a.m. –  
11:45 a.m.

**Chair: James L. Day Jr.**, Deputy Director, Surface Warfare Division, Chief of Naval Operations (OPNAV N96)

***Software Acquisition and the Color of Money***

Jeff Dunlap, Naval Postgraduate School

***Crossing the Great Software Development Divide within DoN***

Chris Johnson, NIWC Pacific

Amanda George, NIWC Pacific

David Jenkins, NIWC Pacific

**James L. Day Jr.**—has served the United States Navy in multiple capacities for over 29 years. He was selected to the Senior Executive Service in August, 2020.

Mr. Day currently serves as the Deputy Director of the Surface Warfare Division (OPNAV N96B) on the Chief of Naval Operations' staff. In this capacity, he supports the development, integration, and resourcing of a roughly \$20 billion annual portfolio spanning the manpower, training, sustainment, modernization, and procurement requirements of the Navy's Surface Warfare systems (ships, weapons, combat systems, and sensors).

Prior to joining OPNAV, he served as the Director, Production, Deployment and Fleet Readiness within the Program Executive Office Integrated Warfare Systems (PEO IWS). Mr. Day also served as the Deputy Program Manager for the Guided Missile Frigate Construction Program (PMS 515). While serving as DPM, Mr. Day led the accomplishment of several major program milestones including the award of one of the largest full and openly competed ship construction contracts in Navy history.

Previously, Mr. Day served on the staff of the Assistant Secretary of the Navy, Research, Development and Acquisition (ASN RD&A) providing oversight and guidance for all Navy, Above Water Sensors and Laser Weapon Systems. Mr. Day also served as the Principal Acquisition Program Manager for LCS Combat Systems, the lead for DDG 51 Combat System, Ship Integration and the lead for AN/SQQ-89(V) Advanced Development and Production in PEO IWS. He has also held multiple leadership positions at the Naval Undersea Warfare Center, Newport, RI. Mr. Day is a veteran of the United States Navy having served from 1992-1998 and achieving the rank of Sonar Technician, Second Class.

Mr. Day received his BS in Computer Information Systems from Roger Williams University and his Master of Business Administration from the University of Maryland Global Campus. Mr. Day is also a graduate of the Executive Program Managers Course at Defense Acquisition University and the Executive Leadership Certification Program from Cornell University.



## Software Acquisition and the Color of Money

**Jeff Dunlap, CAPT USN (Ret.)**—is a Faculty Member at the Department of Defense Management at the Naval Postgraduate School (NPS). Dunlap has over 25 years of experience in the Department of Defense (DoD) as an acquisition professional and has led several software-intensive programs.

As part of his service to NPS, Dunlap provides mentorship and thesis advice to military and civilian students researching software changes needed within the DoD to increase timeliness and value to the warfighter.

Dunlap has a BS from Virginia Tech at Blacksburg. He received an MS in engineering from NPS and Defense Acquisition University ACAT I PM certifications. [jeffrey.dunlap@nps.edu]

### Abstract

The current Department of Defense (DoD) acquisition budgeting process provides funding visibility to Congress for hardware-intensive systems from requirement generation to ultimate disposal. Unfortunately, a square peg in a round hole quandary has occurred with a funding mismatch as modern software-intensive systems are required to comply with traditional funding appropriation breakout categories (aka colors of money). The 2019 Defense Innovation Board (DIB) Software and Acquisition Practices (SWAP) report identified the funding challenges of continuous software development and stated, “Colors of money doom software projects.”

In the fiscal year (FY) 2020 National Defense Authorization Act, Congress created a pilot program with a new appropriation category for software-intensive DoD programs (BA-8). The challenge to the DoD is to prove via quantifiable metrics that a single appropriation of funds enables speed-to-capability deliveries in the software pilots. Other contributing factors made it difficult to discern the effects of BA-8, as revealed by the pilot program metrics, which highlighted potential future study areas. Regulations and policies regarding funding that do not consider the continuous delivery of software capability to the user after the fielding event milestone can lead to confusion about the appropriate appropriations to use and their timing.

### Executive Summary

Software acquisition, development, and support practices within the Department of Defense (DoD) had not fundamentally changed since the implementation of the “waterfall” model in the 1975 DoD Directive 5000.29. In 1987, The Defense Science Board Task Force on Military Software recommended a shift away from waterfall software practices (more common in hardware-intensive programs) to an iterative prototype (agile) lifecycle model. The seminal report of 2019 from the Defense Innovation Board (Software and Acquisition and Practices [SWAP]) was fundamental in describing a tailored, software-specific pathway that guided acquisition change. The end goal of the software change recommendations from the SWAP report was to empower acquisition professionals to deliver relevant and secure capabilities at the “speed to need” using modern software practices found in the commercial sector.

The acquisition and sustainment process for fielding and supporting software-intensive systems changed with the software pathway of the 2020 DoD Instruction (DoDI) 5000.02 (Operation of the Adaptive Acquisition Framework). Although simplifying acquisition policy has eliminated numerous obstacles, the budgeting system (PPBES) categories do not effectively apply to software development that follows iterative and continuous approaches. In contrast, acquisition strategies for incremental waterfall software development programs aligned closely to the budget appropriations spending categories of development, production, and operations & sustainment phases.





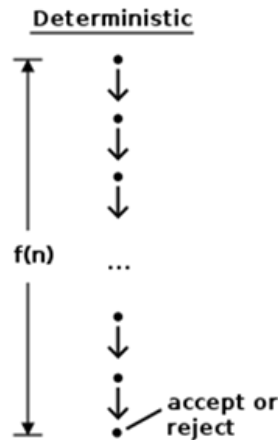
Included in the 2019 SWAP report was the recommendation to create a new appropriation category that would allow software-intensive programs to be funded as a single budget appropriation item since “software is never done.” The model for a modern continuous software acquisition is that there is no separation between development, production, and operations & sustainment. In 2021, the “bleached” or “colorless” appropriation pilot (Budget Activity 8 [BA-8] for software and digital technology pilot programs) began with nine DoD software-intensive programs. The BA-8 pilot provided a single appropriation that could be utilized for any legitimate expenditure.

Intuitively, it makes sense that programs developing continuous capability in an agile fashion would need money that has no restrictions (such as development, testing, or maintenance). Metrics to validate and assess the effectiveness of these pilots would allow the DoD to understand the impact of BA-8 on delivering capabilities “at the speed of relevance.” Since BA-8 is not limited to programs using the Software Pathway, understanding the secondary and tertiary impacts on capability delivery became important in determining what effect PPBES had on software-intensive programs. Data collection and metrics for the BA-8 pilots were required quarterly by Congress. An early assessment by the Under Secretary of Defense for Acquisition and Sustainment (USD A&S) for the effectiveness of BA-8 occurred in the 18th month of implementation. The results indicated that the pilot programs could not demonstrate the singular value of “colorless money.” Other factors were discovered to have as much influence on software acquisition as BA-8. The absence of measurable criteria to exhibit to Congress the worth of a solitary appropriation could hinder the establishment of a sustainable budget classification for software-intensive programs.

### **Delivery Speed is not a Characteristic of the Deterministic Waterfall Approach**

The waterfall development method’s primary utilization was in software engineering for decades. This approach follows a linear sequential path, where each phase of the software development process must be completed before moving to the next phase. The method begins with requirement gathering, followed by design, implementation, testing, deployment, and maintenance. Its rigid structure ensures that each phase must be finished before the next phase can begin, making it difficult to make changes later in the process. Despite its limitations, the waterfall method is still used in some projects where requirements are well-defined, and flexibility is unnecessary. Software acquisition, development, and support practices within the DoD have not fundamentally changed since implementing the “waterfall” model in the 1975 DoD Directive 5000.29. The waterfall process is not inherently good or bad, as its effectiveness depends on a clear understanding of the requirements in advance, which must be known and remain unchanged. Waterfall methodology is very deterministic, where all the software’s functions or features are understood in advance, and the entire software is either accepted or rejected at verification (Figure 1).





**Figure 1. Deterministic Software Development**

Deterministic programs often lack the flexibility and agility to deliver capabilities quickly. Software programs that follow the deterministic waterfall approach cannot deliver capabilities quickly for several reasons. Some of the most common causes include

1. **Outdated technology:** Software developers often build software using outdated technology. As a result, integrating new features and functionalities becomes difficult and time-consuming.
2. **Complexity:** Software programs themselves often exhibit a high level of complexity and are hard to modify. The reason for this is that numerous developers with their preferred programming languages, tools, and methodologies may have contributed to the development of the program over time.
3. **Lack of documentation or code:** DoD software programs may have little or no documentation or software code due to intellectual property or data rights missing in the contract deliverable. Even if the code is available, it is difficult for other developers to understand how the software works and how to modify it without the original developer's documentation.
4. **Limited resources:** software programs may not have the resources or budget to invest in modern software development practices. The delivery of new capabilities can be slowed down due to this.
5. **Technical debt:** Over time, software programs may accumulate technical debt, which refers to the cost of maintaining and updating software that was not properly designed or developed in the first place. Technical debt can slow development and make adding new functionality difficult without introducing new bugs or issues.

Before 2020, the DoD's acquisition framework did not encourage a modern software development methodology, whereby contracts with the Defense Industrial Base are awarded without complete visibility of the requirements. Currently, the DoD's budgeting process (PPBES) still mandates deterministic knowledge of the total acquisition requirements, as well as its timing and cost, regardless of its development approach.



## Modern Software Development Enables Speed

In the early 2000s, the private sector shifted away from traditional heavyweight methods of developing software-intensive systems, such as the waterfall approach. The need to meet market demands for speed and capture customers by delivering working software drove this pivot. The rise of lightweight software development methods such as Agile and the automation software tools needed to build, integrate, test, and deploy continuously began the DevOps/software factory concept. The DoD began to question whether it could take a page from the private sector and refactor how software is developed and deployed to the customer (warfighter) to meet the “speed to capability” demand signals to maintain the warfighter advantage. The Defense Science Board has stated over the years that shifting to a modern software process is not a technology issue but a process and culture question.

A self-evaluation of the software acquisition process in 2019 by the Under Secretary of Defense for Acquisition and Sustainment (USD A&S) reinforced that the DoD is a performance-based bureaucracy that focuses on time, schedule, and budget to evaluate the performance of its programs. The DoD’s acquisition strategy was guided by the capability-based assessment process, also known as JCIDS, to counter future threats to the national security mission. This requirements-based process provided justification inputs into the budgeting process (PPBES), which produces a current and future year budget forecast (5 years into the future). When comparing the commercial marketplace and decisions that the private sector often makes on software development capability investments to dominate competitors, it becomes evident that there is a great divide between the two processes.

DoD Acquisition Guidance underwent a significant process change in 2020: the Adaptive Acquisition Framework (AAF; DoDI 5000.02). The objective of this modification was to provide the end user with prompt and cost-effective solutions that are efficient, appropriate, durable, and environmentally sustainable. Following this release, the Software Acquisition Pathway (SWP; DoDI 5000.87) further defined the purpose to facilitate rapid and iterative delivery of software capability (e.g., software-intensive systems or software-intensive components or sub-systems) to the user (Figure 2). The SWP Characteristics were similar to the commercial marketplace where the user became the focus.

This pathway integrates modern software development practices such as Agile Software Development, Development, Security, Operations, and Lean Practices. Small cross-functional teams that include operational users, developmental and operational testers, software developers, and cybersecurity experts leverage enterprise services to deliver software rapidly and iteratively to meet the highest priority user needs. These mission-focused, government-industry teams leverage automated tools for iterative development, builds, integration, testing, production, certification, and deployment of capabilities to the operational environment. (Office of the Under Secretary of Defense for Acquisition and Sustainment [OUSD(A&S)], 2020a)

The DoD defined this shift in software acquisition procedures as a rapid, iterative approach to software development that reduces costs, technological obsolescence, and acquisition risk. However, because many software acquisition programs involve either applications or embedded software, there are differences in planning and execution timing. In addition, to expedite speed to capability execution and get to quick wins, several steps required by traditional capability acquisition programs were relaxed or eliminated. Unfortunately, the PPBE funding process was unaffected by this acquisition initiative and still follows the cold war era processes with little ability to flex based on emerging threats. The DoD’s move towards a non-deterministic software architecture is rapidly growing in practice.



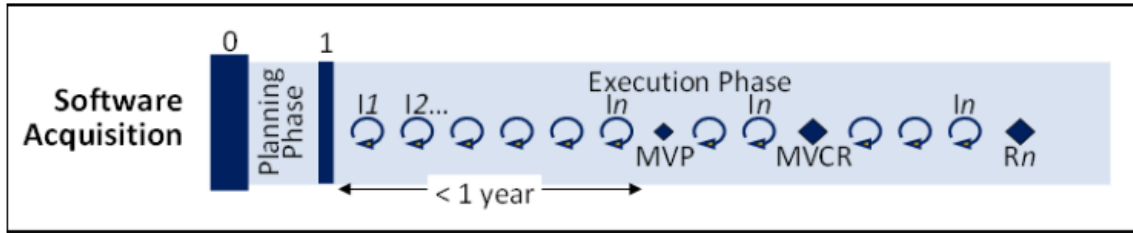


Figure 2. The Software Acquisition Pathway: Iterative Development of Application Software

## Monolithic Architectures Have Inertia to Change

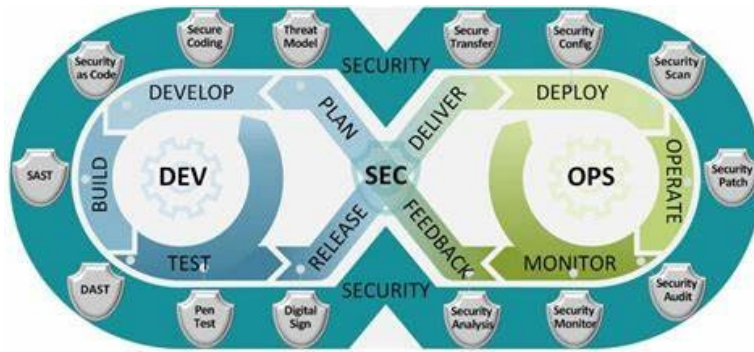
Most of our weapon systems' offensive and defensive capabilities are software-controlled, and the ability to respond to opportunities and threats requires software updates at the tactical edge. The DoD built most of its warfighting software using a Monolithic Architecture. Monolithic Architecture is a traditional software design approach that involves developing an application as a single, self-contained unit. This architecture is characterized by tightly coupled program components or functions, meaning they are highly dependent on each other and tightly integrated into the overall application. This design approach makes developing, deploying, and maintaining the software more manageable, as everything is contained within a single codebase. It's like building a large, complex building with all the rooms and floors interconnected and dependent on each other.

The downside of this approach is that it can limit scalability and flexibility, as changes to one component may affect the entire application, and it can be challenging to add new features or scale the system as it grows in complexity. Monolithic Architecture requires all associated components to be present for code execution or compilation and for the software to run. Moreover, modifying a single program component may necessitate modifying other software elements, leading to the entire application requiring recompilation and testing. Such a process can consume a significant amount of time and hinder the agility and swiftness of software development teams. An example of this problem is the delay in enhancements/updates to the shipboard combat systems software (AEGIS/SSDS), which typically exceeds 6 years to get to the warfighter and, by administrative procedure, never updated while deployed operationally (PEO IWS X).

## Modern Software Architecture Brings Speed to Delivery

The DevSecOps process (software factory) is the big buzzword in DoD software acquisition. Looking at Figure 3, there is a continuous process of engagement with the development team (software coders), security experts (helping the coders learn and verify best practices), and the users (operators of the system). The development and operational environments are closely related and connected through telemetry, enabling health and status reporting with user feedback. Agile software coding principles and culture within the software factory remain fundamental to the team's success.

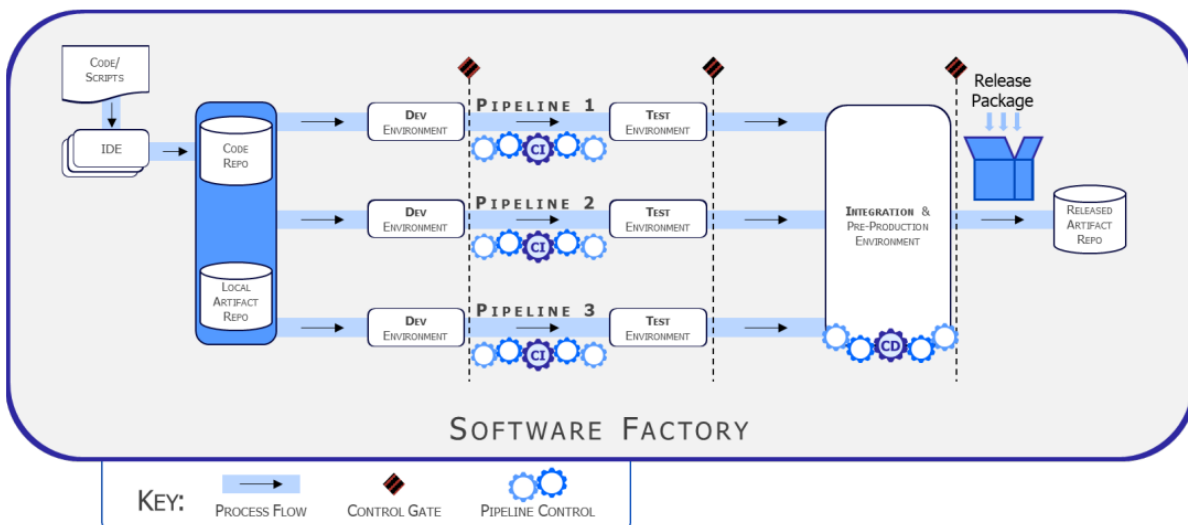




**Figure 3. DevSecOps Distinct Lifecycle Phases and Philosophies**

Most people can intuitively grasp the conceptual value of connecting software coders to the system’s users, as this enables feedback and integrates security experts into the DevSecOps process. However, the concepts of modern software architecture go far beyond the figure of connecting rings. Understanding how a software factory can continuously deliver cyber-secure software to the tactical edge is also connected to the “colors of money” discussion. The DoD Chief Information Officer (CIO; 2021) has provided a DevSecOps Strategy Guide as a starting reference to what the advantages of a software factory may bring.

The software factory is the “Dev” component of DevSecOps (Figure 4). It pertains to the processes where software developers continuously integrate and test their code in a secure cloud environment. The “pipelines” produce software applications of self-contained functionality, also known as “containers.” Container applications are lightweight (<10 megabytes) and bundled in a release package. Under the traditional acquisition approach, the software is compiled into machine language and provided as a single monolithic package containing multiple features, typically exceeding 10 gigabytes in size. Every time a change is made or added to the software, the entire monolithic package has to be recompiled and re-installed. The software factory differs substantially from the monolithic waterfall method because it does not compile functionality into a single software package. Instead, each container application executes specific functions upon receiving a request from an orchestrator and terminates afterward.



**Figure 4. Software Factory Construct**

The software factory achieves modern architecture by breaking the traditional monolithic into discrete domains and orchestrating containers to perform these domain services. Microservice architecture is an architectural style that structures an application as a collection of container services. In a rapidly changing environment where maintaining warfighting dominance is crucial, the microservice architecture enables organizations to deliver large, complex applications quickly, frequently, reliably, and sustainably. Figure 5 shows the conceptual difference between a monolithic and microservice architecture. The key advantage of the microservice architecture is the speed at which users can add/modify capability. Development teams can rapidly deploy individual software components without redeploying the entire application. The DevSecOps Software Factory follows a pipeline process to develop new containerized software, which involves testing, integration, and release into the repository. The size of the software matters, as bandwidth is often a limiting factor at the tactical edge or in a contested spectral environment.

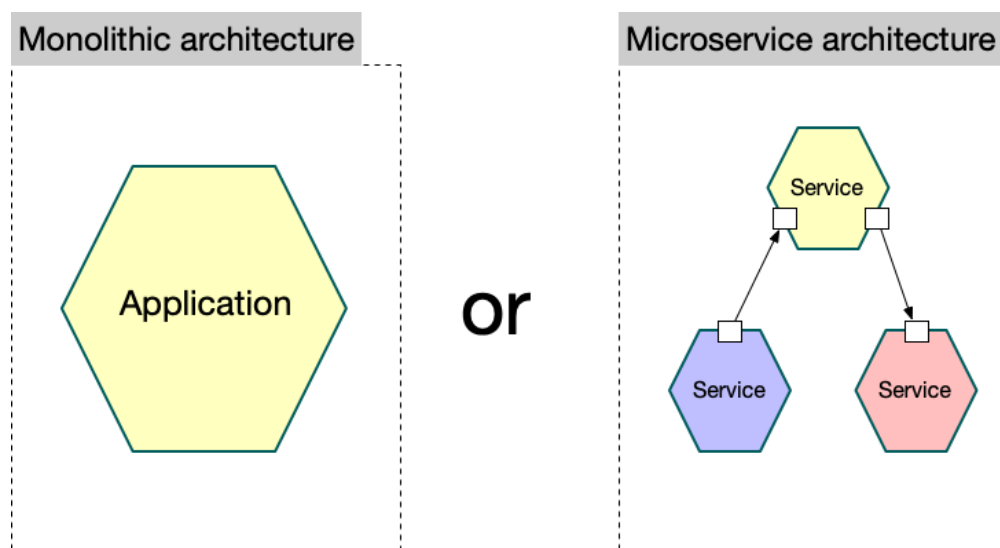


Figure 5. The Conceptual Difference Between Software Architectures

### Acquisition Categories of Money Doom Software Factories?

Traditionally, acquisition programs have followed the same development pattern over the last 40 years. Although there has been encouragement to tailor the acquisition pattern to reduce wasted effort, pathways did not exhibit variances based on the product being developed until the implementation of DoDI 5000.02 in 2020. The Major Capability Acquisition pathway is typical of how appropriation categories of money are programmed into the budget. As the product progresses through the milestones (MS A/B/C), funding is primarily for Research, Development, Test, and Evaluation (RDT&E). Procurement funds are the dominating category spent after MS C, and once fielded, the category shifts to Operations and Sustainment (Figure 6). Unfortunately, the budgeting process (PPBES) sees all acquisition pathways progressing through these funding categories.

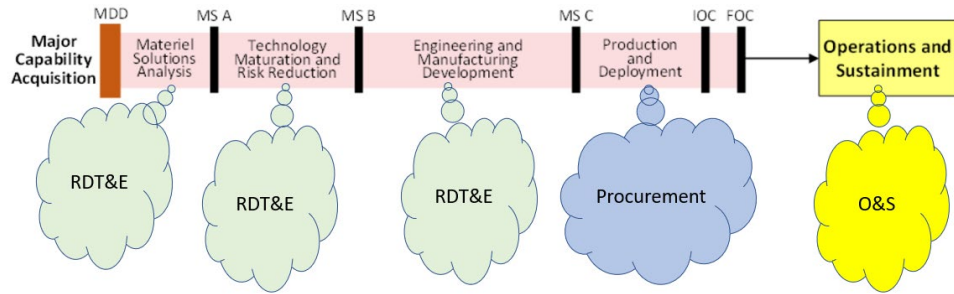
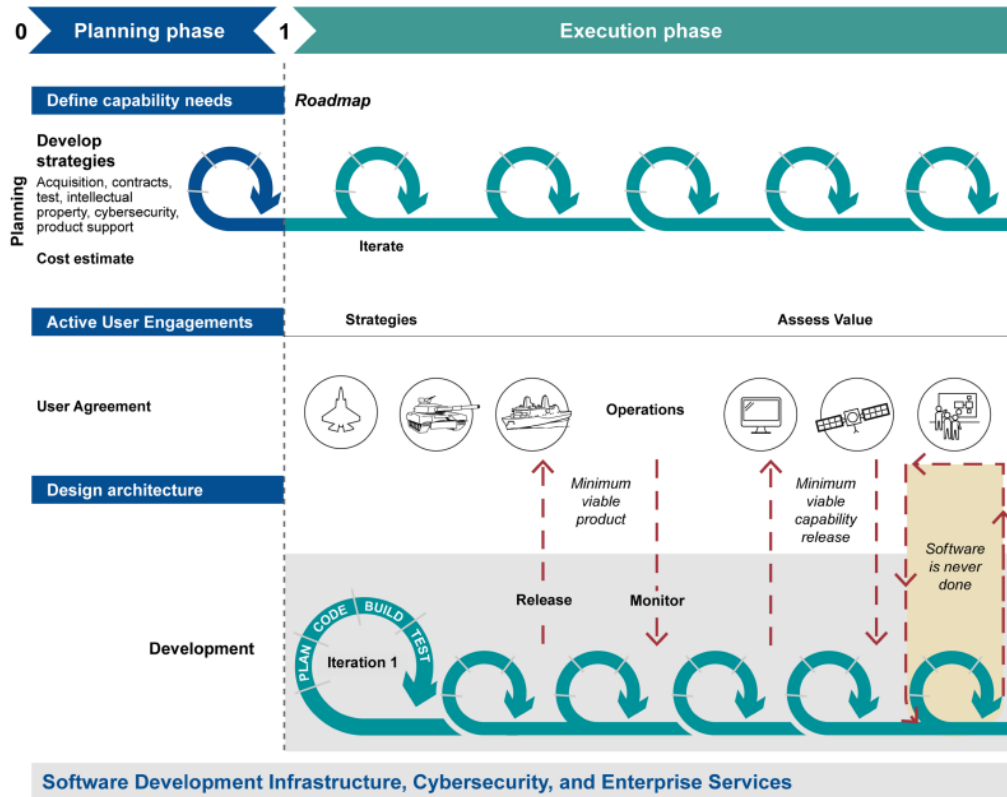


Figure 6. Principle Acquisition Categories of Funding for MCA

The Software Acquisition Pathway is detailed in DoDI 5000.87 and illustrates what is occurring during the two phases: planning and execution (Figure 7).



Source: Department of Defense Instruction 5000.87 (October 2020). | GAO-21-105298

Figure 7. The Software Acquisition Pathway

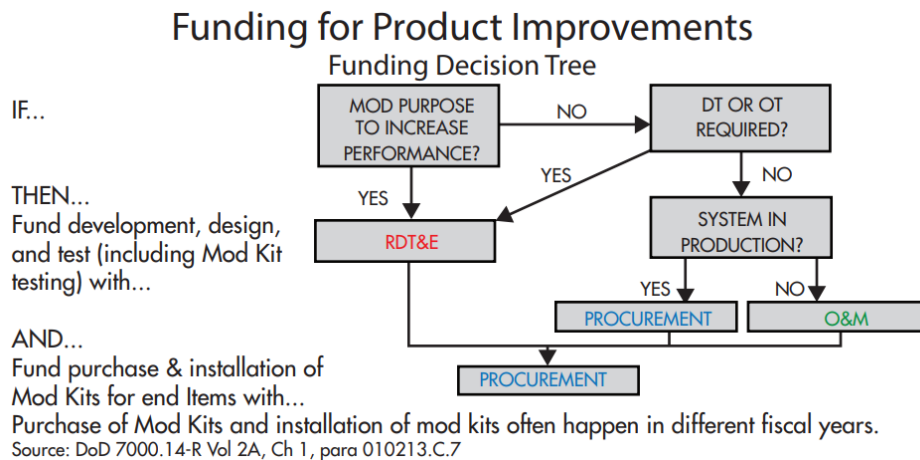
Gone are the funding triggers for product procurement, operations and sustainment from the software acquisition pathway, and in their place, an iterative capability development that is never done. Executing a program following this pathway presents a PPBES dilemma on what category of money is needed and when. The operational funds in the MCA pathway are triggered upon capability deployment to users with specific task guidelines:

Types of expenses funded by O&M appropriations generally include DoD civilian salaries, supplies and materials, maintenance of equipment, certain



equipment items, real property maintenance, rental of equipment and facilities, food, clothing, and fuel. (Defense Acquisition University, 2023)

Funding category and timing decisions are usually based on performance improvement or testing requirements, as illustrated in Figure 9. This flowchart is valuable for the MCA pathway but has little relevance when executing continuous delivery in the Software Pathway. If a new software container is developed that increases the fielded software system’s performance, RDT&E is required using this flowchart. Procurement dollars are needed if a software correction is made to a container and the system is in production. The challenge with “colors of money” in software development lies in determining whether the release of the minimum viable product marks the start of production or the point at which the system is in service. Trying to fit the concept of “colors of money” into software projects is akin to fitting a square peg in a round hole, which can lead to project delays. The reasons for specific congressional guidance on how money is spent make sense only from an accountability perspective. But because software is in continuous development (it is never “done”), colors of money (besides RDT&E) tend to reduce the agility to obligate funds when reprogramming is required. We need to create pathways for “bleaching” funds to smooth this process for long-term programs (DIB, 2019).



**Figure 9. Funds Management Platinum Card Decision Tree**

### Metrics on BA-8 Show Other Influencers

Congress established a pilot program in FY2021 to provide a single appropriation BA-8 (colorless money) to several software-intensive programs. This new appropriation category for software capability delivery has no separation between RDT&E, production, and sustainment. This initiative was a multi-year pilot to collect and analyze metrics to inform a final recommendation to make this an enduring appropriation.

Congress required quarterly BA-8 metrics from the USD A&S. Their FY2021 report to Congress stated that they did not consider BA-8 a silver bullet. Although BA-8 is expected to address some critical challenges faced by programs adopting commercial software development practices, it is not a comprehensive solution to all their problems. Metrics assist leadership in comprehending the effectiveness of pilot programs implementing BA-8. The metrics that OSD picked for the pilot programs in FY2021 are shown in Table 1. These metrics are adapted from the Google DevOps Research and Assessment (DORA) team and are used by DevOps teams to measure their performance and find out whether they are “low performers” to “elite performers.”





**Table 1. BA-8 Pilot Metrics Collected**

Factor Influence on Measure	Product Delivery Lead Time	Release Frequency to Operational Environment	Deployment Frequency to Production	Mean Time to Restore	Change Fail Percentage
BA-8 Single Appropriation	High	Medium	Medium	Medium	Low

The USD A&S acknowledged in the fourth quarterly report to Congress for FY2022 that while there is compelling evidence of improvements provided by BA-8 for the pilot program, it is primarily qualitative. Quantitative measures were utilized to measure the influence of BA-8 based on traditional commercial software factory metrics.

Product Delivery Lead time in a software factory (DevSecOps) measures how much time has elapsed between committing code and deploying it to production, tracking the time spent on implementing, testing, and delivering changes to the codebase. BA-8 positively influenced product delivery lead time, indicating the ability to move quickly through the process. Product delivery lead time, for example, has other high-influence items: total funding, developer staffing, developer skill, development environment, test facilities, developmental and operational test support, and system complexity, as seen in Table 2.

The USD A&S report for FY2022 states that numerous factors, not just BA-8, have an equivalent or more significant impact on metric outcomes. Therefore, it was difficult to quantify the effect of BA-8 in isolation precisely. Table 2 provides a comprehensive analysis of the factors influencing software program activities by considering these additional variables.

**Table 2. Software Factory “Other” Quantitative Factors Having Influence**

Factor Influence on Measure	Product Delivery Lead Time	Release Frequency to Operational Environment	Deployment Frequency to Production	Mean Time to Restore	Change Fail Percentage
Total Funding	High	Medium	Medium	Medium	Medium
Developer Staffing	High	Medium	High	High	High
Developer Skill	High	Low	Low	Medium	High
Development Environment	High	Medium	Medium	High	Low
Test Facilities	High	High	High	Medium	High
Developmental & Operational Test Support	High	Low	High	Medium	Low\
Time to get Authority to Operate	Low	Medium	High	Low	N/A
Capability Complexity	High	High	High	High	High
User Ability to Accept Releases	N/A	N/A	High	N/A	N/A
Contracting Methods	Medium	Low	Low	N/A	N/A

Metrics, such as product development lead time and deployment frequency into production, help teams understand their overall engineering performance. In addition, they



provide the software program with an objective way to measure and improve software delivery. Metrics help DevSecOps teams quickly identify bottlenecks and inefficient processes in their development pipeline and create a plan to improve their daily work (Software.com, 2023). Several quantitative factors identified in the BA-8 Pilot play a crucial role in delivering high-quality software to the user within the deadline and are of significant value beyond the BA-8 efforts.

## **Software Metrics That Add Value**

Commercial Software teams use modern iterative software methods to emphasize development using fixed cost and time, with flexible requirement estimates. Defining all of the software requirements at the program's start is impractical, as this is counter to agile software non-deterministic development methodology. Current software cost estimation and reporting processes and procedures in the DoD have proven to be time-consuming, highly inaccurate, and time late. Metrics of Earned Value Management for software development cannot match the continuous capability delivery and maintenance velocity of DevSecOps. Metrics that align with the DevSecOps approach and offer continuous visibility into program progress are necessary.

The SWAP report recommends that projects develop metrics that measure value to the user (or customer satisfaction), which involves close, ongoing communication with users. How this metric of "user value" is calculated is undefined in the BA-8 Pilot. In the commercial sector, many agile software teams use broader business indicators to gauge overall performance and product quality. The software factory doesn't directly own or collect data for these metrics since they represent customer satisfaction, value delivery, and flexibility.

The measurement of cost and performance for software factories are automated within the infrastructure tools and report continuous speed and cycle time, cybersecurity vulnerabilities, code quality, and functionality to assess, manage, and justify terminating a software program (if needed). In addition, software code performance metrics address issues such as deployment rate and speed of delivery, response, and recovery from outages, and can be automatically generated continuously.

## **Future Funding of Software Programs Uncertain**

Congress did not authorize additional BA-8 pilots in FY2023 due to the perceived lack of quantitative metrics from the USD A&S. All of the Senior Department Software Acquisition Executives provided qualitative inputs for the BA-8 Pilot, Fourth FY2022 Quarterly Report to Congress, and the Army's comment on the value of BA-8 funding (OUSD A&S, 2022) was particularly relevant to this paper.

Given the modern and ever-changing software environment, the legacy funding model of RDT&E, procurement, and O&M makes it difficult to effectively and efficiently acquire and develop software. With the Army's need to remain competitive and defeat near-peer adversaries, the Army must be able to rapidly secure, enhance, and maintain software. Legacy software development practices cannot keep up with the pace of change required to address the ever-changing threat landscape. They also establish clear lines between software development (new capabilities) and software maintenance (cyber and software fixes) activities. This division of activities aligned well with current funding models; development = RDT&E and maintenance = O&M.

With the advancements of cloud computing, Agile software development, and Development, Security, and Operations (DevSecOps), everything is



integrated and must operate at a rapid pace. These modern software practices do not distinguish between software development and software maintenance. The software is viewed as a product that is continuously evolving. These practices involve adding capability, fixing software problems, and cyber-securing software with a single team as part of a single software delivery. This cultural and technological change removes the line between software development and software sustainment, making it challenging to fund those activities separately with different appropriations. Without the use of BA-8, it will require a very cumbersome and difficult process to identify exactly the number of hours each team member spends on adding capability (RDT&E) and fixing problems (sustainment – O&M).

From a qualitative perspective, the services agreed with the SWAP report view of the value of colorless money. The understanding that software is no longer a monolithic delivery and that capability can be delivered to the warfighter at the tactical edge in lightweight application containers is a quantum jump forward. However, the genuine concern ultimately resides with the funding needed for the software factory itself.

### **Maybe Treat Software Factories as an Enduring Service?**

The DoD's issues with various appropriation categories could be addressed by adopting existing best practices in the private sector by establishing software factories as an enduring service. Software Factories established per the DoD software modernization strategy of 2021 should combine Cloud-based computing and use an assembled set of software tools enabling developers, users, and management to work together on a daily tempo to achieve delivery of a minimum viable product. The software development continues until a minimum viable capability is released into the user community. Funding for the software factory becomes either a time of material or a level of effort contract expense for labor that ebbs and flows as the software factory continues to add user-desired capabilities during the execution phase. In addition, the software factory itself has expenses such as software tool licenses and government salaries. As the number of features to be coded and delivered decreases over time, the software factory can either start new tasking from another pillar program or reduce the workforce to keep a core capability while the software is in user operation. Whether it is a new capability, fixing a deficiency, or cyber vulnerability, it is colorless money.

Software factories provide significant value as an enduring service for software development. By providing a consistent, standardized approach to software development, software factories can help to increase productivity, improve collaboration and quality, reduce risk, and provide ongoing support and maintenance for software applications. Additionally, software factories can benefit individual developers, enabling them to work with new technologies and tools and improve their skills over time. As software development becomes increasingly complex and demanding, software factories may play an essential role in enabling teams to work more efficiently and effectively and deliver high-quality software applications.

### **Concluding Thoughts**

Pilot results are essential in confirming study assertions and making necessary adjustments to achieve desired results. Congress and the DoD have been aware since the Defense Science Board study of 1987 that monolithic waterfall acquisition of software takes too long, is too expensive, and exposes warfighters to an unacceptable risk by delaying their access to the software needed to ensure mission success. Software is responsible for most of the capabilities in our weapon systems and applications that provide command, control, and



communications. The DoD realized a different “adaptive” acquisition process was needed to speed technology and capability delivered to the warfighter. The USD A&S in 2020 provided the framework leadership needs to drive change in software acquisition. By leveraging commercial practices, experience, and tools, the DoD implemented the DevSecOps (Software Factory) initiative to support software as an enduring and evolving capability that is continuously improved throughout its lifecycle. Modifying statutes, regulations, and processes will not accomplish the enduring result needed to prioritize speed and continuous capability delivery to the warfighter. The color of money pilot (BA-8) and the measures of effectiveness have revealed that other software factory influences, if addressed proactively, may have a positive synergistic effect on delivery velocity.

The BA-8 Pilot Software-intensive programs picked all had one thing in common; none were in receiving funding other than RDT&E appropriations. The ability of a software program to estimate years in advance in PPBES exactly how much RDT&E, Procurement, and O&M dollars are needed to support the software is an inexact art. Reprogramming funding between appropriations is non-trivial and highly time-consuming. A new threat that emerges to a fielded system or an opportunity to exploit an enemy’s weakness may be lost in the bureaucracy of money exchange. Bleached or colorless money in the hands of the software factory would allow enhanced containerized applications to be developed, tested, and sent to the tactical edge in days, if not hours.

Software is an enduring and evolving capability that must be supported and continuously improved throughout its lifecycle. DoD’s acquisition process and culture need to be streamlined for effective delivery and oversight of multiple types of software-enabled systems, at scale, and at the speed of relevance. Optimizing for software is the path forward. (DIB, 2019)

## References

- Defense Acquisition University. (2023, March). *Operations and maintenance (O&M) funds*. <https://www.dau.edu/acquipedia/pages/ArticleContent.aspx?itemid=339>
- Defense Innovation Board. (2019, May). *Software is never done: Refactoring the acquisition code for competitive advantage*. Department of Defense. [https://media.defense.gov/2019/Mar/14/2002101480/-1/-1/0/DIB-SWAP\\_STUDY\\_REPORT.PDF](https://media.defense.gov/2019/Mar/14/2002101480/-1/-1/0/DIB-SWAP_STUDY_REPORT.PDF)
- Department of Defense Chief Information Officer. (2021, March). *Department of Defense enterprise DevSecOps strategy guide*. <https://dl.dod.cyber.mil/wp-content/uploads/devsecops/pdf/DoDEnterpriseDevSecOpsStrategyGuide.pdf>
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020a, January 23). *Operation of the adaptive acquisition framework* (DoD Instruction 5000.02). Department of Defense. <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500002p.pdf?ver=2020-01-23-144114-093>
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020b, October 2). *Operation of the software adaptive acquisition pathway* (DoD Instruction 5000.87). Department of Defense. [https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500087p.PDF?ver=virAfQj4vLgN1JxpB\\_dpA%3D%3D](https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500087p.PDF?ver=virAfQj4vLgN1JxpB_dpA%3D%3D)
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2022, November). *BA-8 pilot, fourth FY22 quarterly report to Congress*.
- Program Executive Office Integrated Warfare Systems X. (2022, July 28). *ICS overview brief*.
- Software.com. (2023, March). *Key engineering metrics in software delivery*. <https://www.software.com/devops-guides/engineering-metrics>



# Crossing the Great Software Development Divide within DoN

**Chris Johnson**—is the Naval Information Warfare Center Pacific Division Head for the Command and Control Systems Division. Chris was the co-creator of the NRDE Commercial Cloud and led the development of the Collaborative Software Armory, now the Overmatch Software Armory (OSA). He is a retired Operations Specialist Chief Petty Officer in the US Navy. Mr. Johnson holds a Bachelors of Science in Information Technology and a Master's degree in Software Engineering from California State University Fullerton. [christopher.e.johnson40.civ@us.navy.mil]

**Amanda George**—is the Naval Information Warfare Center Pacific Business Portfolio Manager for Business and Force Support (BFS). The BFS Portfolio encompasses a wide range of projects providing innovative and agile business and security solutions to the Warfighter. As the Business Portfolio Manager, Amanda supports the BFS projects by engaging with industry and stakeholders, building internal collaboration between projects, and providing support for strategic planning. Amanda holds both a Master's Degree and Bachelors (Phi Beta Kappa) from the University of California, San Diego (UCSD). [Amanda.j.george7.civ@us.navy.mil]

**David Jenkins**—is the Deputy for Business Development for Naval Information Warfare Center Pacific's Command and Control Systems Division. In this capacity, David applies his extensive knowledge of contracting, marketing, and project management to support the development of hardware and software-based command and control systems. David holds a Bachelor's degree in International Business from San Diego State University and a Master's Degree in Business Administration from National University. [david.w.jenkins5.civ@us.navy.mil]

## Abstract

The US Navy must undergo a software development cultural transformation in order to address outdated considerations associated with pushing its development efforts successfully into the future. The current processes to develop software and acquire key capabilities needed to develop software, and the culture that these processes produce is slowing down the Navy's ability to provide crucial technologies to the warfighter quickly. Naval Information Warfare Center Pacific (NIWC Pacific) has been on the forefront of utilizing cutting-edge software development techniques and enabling technologies, and thus has some key lessons learned to share with the acquisition community. This paper will look at three (3) key areas: 1) Acquisition processes via contracting 2) Software development security approaches, and 3) The Navy's financial ownership of software development. This paper will explain why they are key, and provide recommendations to transform the way ahead for the US Navy in its software development efforts.

## Introduction

“In this era of competition and race for digital dominance, we cannot settle for incremental change. The Department must join together to deliver software better and operate as a 21st century force.”

- Department of Defense Software Modernization Strategy (2022, p. ii)

The DoD's Software Modernization Strategy succinctly summarized the importance of DoD's ability to deliver software, saying “fighting and winning on the next battlefield will depend on DoD's proficiency to rapidly and securely deliver resilient software capabilities” (2022, p. 1). The key to this is the focus on rapid and secure delivery of software. If the DoD and the Navy can't deliver software rapidly, it will be too late to support the fight. If they can't deliver resilient capability, then we will never succeed in a contested cyberspace. The DoD's Software Modernization Strategy then goes on to identify a practical approach to “unify efforts across DoD and partner with industry-leading software institutions to produce a portfolio of best-in-class software capabilities enabled by DoD processes” (2022, p. 2). While this is in fact necessary, it's not a sufficient approach. Current acquisition and security processes developed to acquire and



secure large-scale physical vessels, vehicles, and machinery, along with a lack of financial commitment from the Department of Navy are major blockers that severely degrade the ability of the Navy to support the rapid pace of product delivery required to defeat our adversaries. The largest blocker, however is a cultural divide within the Navy and DoD surrounding whether and how to adopt a new, agile, resilient software development mindset.

The US Navy must undergo a software development cultural transformation in order to address outdated considerations associated with pushing its development efforts successfully into the future. The current processes to develop software and acquire key capabilities needed to develop software, and the culture that these processes produce is slowing down the Navy's ability to provide crucial technologies to the warfighter quickly. Naval Information Warfare Center Pacific (NIWC Pacific) has been on the forefront of utilizing cutting-edge software development techniques and enabling technologies, and thus has some key lessons learned to share with the acquisition community.

## **Acquisition Processes via Contracting**

Providing capability to warfighters and meeting program requirements is often less a technical challenge than one of acquisition. Taking a quick glance around displays at a trade conference, such as WEST 2023, you will see showcased many new capabilities primed for transition into Navy programs. Industry experts in building custom Government applications are looking for teaming opportunities with Navy customers to develop products. Indeed, industry is poised and ready to help solve some of the Navy's biggest technical challenges, but in order to partner, a contract has to be issued. Long contract lead times, however, are working against accelerating delivery of new capability to the warfighter.

The current acquisition system was created to facilitate the acquisition of large-scale physical procurements of everything a military might need from bullets to an aircraft carrier. This process was designed to reduce the risk inherent in procurement. This risk reduction focus has created a culture in the acquisition community that highly prioritizes set requirements. Sacrificing agility in favor of risk reduction is fundamentally opposite to the culture we need. We need a contracting approach that can move agilely, so we can implement industry solutions at the speed industry is creating them. We need a contracting workforce that is empowered to apply the best, tailored contracting approach for the procurement need. We need a contract acquisition environment that matches agile software development principles.

At the present, the acquisition environment is not conducive to working as swiftly as need requires. These days, software development most often occurs on a "two pizza" sized team, building small applications that can be created over a short (less than a year) time frame, in an environment where requirements are not stable and require day-to-day interaction with the warfighting customer to solidify the design and function of Minimum Viable Product (MVP). The Navy's contracting approach should reflect this dynamic. Instead of the current contracting standard requiring highly specific and well-defined requirements, "good enough" requirements and evaluation criteria should be the goal in such an agile development environment, where risk-taking by both the Government and their industry partner should be encouraged and even rewarded.

When agility is required, the time to award to an industry partner must be short in nature. At many Military commands, spending months on requirements development is the norm. Identifying requirement specifics to an acceptable level of detail and documenting justifications for changes to standard processes significantly increase the time needed to execute a contract. Spending months nailing down requirements should be an outlier, not the norm. A contract needs to capture the essence (i.e. most important aspects) of the requirements at the



development onset, and it should allow for evolving and emerging requirements all the way through to MVP; it need not eliminate all risk nor capture all requirements needed in perpetuity. Task order award timelines, in particular, must reflect product teams' realities, and should be measured in weeks not months.

### **What do we Need to do to Change the Mindset?**

So how do we, both requirements owners and contract owners, change our mindset? The solution is likely two-pronged: educate the local contract and technical personnel to effectively communicate agile software development requirements, while also providing nontraditional contracting approaches to increase speed and access to nontraditional contract partners. At a local level, NIWC Pacific has focused on working with our technical and contracts personnel to ensure that they each understand the requirements of modern software development. The DoD has made significant efforts, laid out in the 2022 DoD Software Modernization Strategy to "make the acquisition lifecycle and the funding of software programs more agile." (pg. 13) In addition, NIWC Pacific has been looking at options like the Information Warfare Research Project (IWRP) program using Other Transaction Authority (OTA) to bring in prototypes and small efforts more quickly. While this is happening to a degree locally, the larger issue is shifting the culture of contracting away from the idea that a procurement team (to include both requirements owner and contracting team) must spend months on getting requirements exactly right before making a purchase. Agile software development needs smaller increments of improved capability, rolled out at scale with a rapid refresh rate. Such a cultural shift is challenging, but it could happen if it gained momentum at multiple commands and within the larger DoD community.

### **Software Development Security Approaches**

"Hey, my code is secure because I filled out the RMF spreadsheet!" - said no one, ever.

While you likely laughed, this statement reflects how we currently manage software security requirements. A large part of what we do as software developers is focused on filling out certain Risk Management Framework (RMF) artifacts, as if by doing so we have secured our code and made our systems safer. In years past, after programs went through months of coding development that software would be integrated into the larger-scale system to ensure module-to-module interoperability. Concurrent to this activity happening, the security evaluation would begin by a separate engineering team tasked to ensure that DoD RMF guidance locked down security vulnerabilities, in accordance with policy, to ensure no High Risk vulnerabilities were present which would prevent the application from acquiring the all-important Authority to Operate (ATO). All-too-often, the activity of locking down these identified security vulnerabilities created new issues that prevented the applications from performing as intended. The application would then be banished back into development for re-work, re-test, then more re-work, etc. This was not only frustrating and slow for all parties involved, but it has also been hugely expensive, as can be attested by virtually everyone that has ventured to release a DoD application.

Another primary concern with software development associated with security approaches, was the common approach of packaging up enterprise or infrastructure elements (i.e. databases, service buses) and incorporate them into their new application package. This gave a false promise to programs that doing so would assure success when going live in the production environment. Unfortunately the opposite often held true. More often than not, the applications instantly inherited the vulnerabilities of the associated service or infrastructure dependency, and once that happened, they were doomed at being able to successfully achieve an up-check security accreditation. Removing such dependencies proved equally complex and

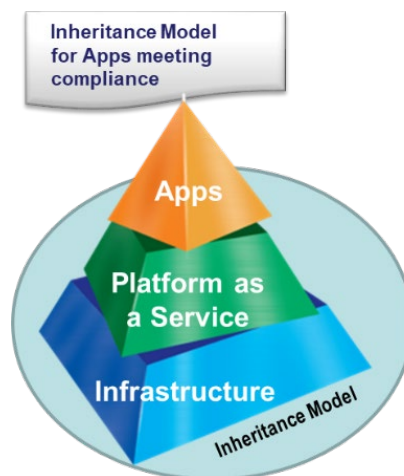


difficult, and many programs were cancelled or abandoned over such issues, sometimes resulting in negative publicity for the Navy.

Our desire as a forward-leaning software community is to provide a way to gain early, accurate insight into the security posture of software during the build phase of development. So early, in fact, that immediate feedback can be internalized by our development teams and immediately address high and moderate risk issues. Today's Development-Security-Operations (DEVSECOPS) environments are designed with tooling that provides a wide array of testing that covers many aspects of security issues and vulnerabilities. Results, then, are consolidated and analyzed while still in the software factory to evaluate any potential false positives, and at essentially the same time ready replacements can be offered. These powerful and always changing tools help alleviate the potential for wasted work years chasing issues that were minor, at best.

These tools provide the developers a direct pointer to the very lines of problematic software code, allowing them to focus their effort on writing better code rather than searching through endless lines of code for an identified vulnerability. In some cases, these tools take it even a step further, and provide not only the exact line of code that is broken, but also can implement a prescribed correction to fix the issue. Such prescribed fixes can save developer time from having to research security fix precedence on something that may be unique or something they are not familiar with. Such an approach is a more secure way to address the cyber posture of our software; and, once again, this can all be done in near real-time. Using these tools, evaluators and decision makers who control fielding decisions allowing production software to be fielded can quickly see how secure the software is within minutes of going through the build phase. In fact, actual fielding decisions could be made months ahead of time, providing for hour-to-hour fielding decisions if required. During the transition, and for those "old-schoolers" that still want a spreadsheet to review, these automated tools can provide those at the touch of a "print report" button.

In addition to helping software developers go faster, the other part of the software security approach challenge is how we as a community adopt a reciprocity approach that would help lessen the burden of fielding systems today. In a current production model using cloud, there are three levels: 1) Infrastructure as a Service (IaaS), 2) Platform as a Service (PaaS), and 3) Mission Applications. Each of these, by definition, also carry a certain level of requirement for cyber accreditation. The majority of the requirements for any ATO reside within the IaaS. All core components of the infrastructure are contained within the IaaS, such as databases, enterprise services, etc.



**Figure 7 Tri-Tier Security Inheritance Model**





In Figure 1, infrastructure is the base of everything that is required for the system and applications to operate. In a typical production example, such as a shipboard environment, the CANES infrastructure serves as the shipboard IaaS, and is the holder of the ATO for the environment. It also contains the bulk of associated RMF security requirements. The other dependent tiers - the PaaS, and Mission Applications - rely on the IaaS to have these items available and then leverage those items for their ATO package.

Next, the PaaS contains the RMF security items necessary to operate the services and elements associated with that PaaS. Again, building on the RMF inheritance model, this enables the mission applications to focus on delivering just the mission component without having to worry about adding in missing services necessary to communicate with the IaaS or other PaaS components. This keeps the Mission Application requirements for RMF lightweight, and easily achievable when tools within DEVSECOPS pipelines can readily provide security insight and accelerate the development-to-delivery timeline. The Navy needs program offices that *trust* the results provided by the software pipeline tooling. This is the very essence of what RMF (emphasis on the “R”) means, and such trust is imperative to delivering code to the warfighter quickly. Trusting these tools to do the job they are built for would represent a healthy shift in culture that will result in an accelerated delivery of software capability and fixes to the Fleet. All too often, the Naval community is not willing to allow for *any* risk, and everything down to lower-level vulnerabilities have to be accounted for with a Plan of Action and Milestone. While these reports are important, very few of them have anything to do with the actual operating posture of our applications and how secure they are.

## Navy’s Financial Ownership of Software Development

Software development is a highly skilled complex task; building a viable software application is equally as hard as building a ship, flying an airplane, or designing a submarine. Similar to building our physical assets—ships, jets, missiles—it is an endeavor that takes tremendous resources of people and specialized tools. In addition, like a Navy shipyard, software development relies on a set of infrastructure tools that need to be provided to the software developers.

For years the Navy has relied on industry building Navy software, and for years we have struggled with the fact that much of that software is self-contained. It does not interoperate well with other applications delivered by other vendors. The code platform might be different, the services might be unique, or the design pattern might not fit the architecture. We typically discover these problems the minute we try to integrate some application into a larger system.

Industry controlled software development was a paradigm long overdue to shift, and recently has, toward Government-owned and operated environments, using the best available industry-developed tools. Requiring each developer to develop within the same Government environment is an acceleration method to ensure that software can interoperate with other applications sooner in the lifecycle. Having a common set of tools helps to confine or bin the development environment, ensuring that there is a smaller risk of language and service diversity inside the applications. To do this, however, requires the Government to be willing to undertake the cost of owning and maintaining those development environments.

A typical environment consists of a cloud service provider, code repositories, cyber security tools, agile project management tools, and engineering to ensure that all the tools are set to the correct security level, are accessible to users, and are integrated. Unfortunately, the tools required to produce these environments are not “plug and play,” particularly given the Government’s security requirements. Therefore, each environment comes with a financial cost that is split between Government labor, contractor labor, and licensing fees.



Given this cost, the Navy must choose to either fund the entire software development pipeline, or share that cost among the users of the software development pipelines. The DoD and the Navy have chosen the latter, enabling and encouraging a multitude of software development pipelines to form. As the GAO states in a recent report (2022), “[a]s of August 2022, the DOD has established 29 software factories across the department” (p. 19) Each of the software pipelines is supported, in large part, by the programs and projects utilizing their services.

The cost of each software factory is initially borne by the organization that needs it. As the RAND Corporation addressed extensively in its 2020 report “Personnel Needs for the Department of Air Force Digital Talent A Case Study of Software Factories”:

Because the software factories are start-up organizations, funding for most of the factories at the time of this study (FY 2020) appeared to be ad hoc, with the parent or owning organization providing initial funds and billets but the majority of funding coming from customers that use their software development and platform capabilities. Although software factory missions primarily focus on serving their parent or owning organization, they also have customers that expand beyond the owner, including the broader DoD.(p. 2)

The software factories spread the cost of the engineering support and tool licenses across all the users. This model, where the command provides an initial investment, then the programs pay as they need the services, creates a significant challenge. If there aren’t enough programs that utilize the unique offerings of a software pipeline, economies of scale are not achieved, resulting in the pipeline becoming unaffordable for projects.

The cost of operating in a software development pipeline for small software development efforts is often much greater than their resourcing. In the face of large price tags, small software development efforts often “go without” such pipelines and an ad hoc, build-it-as-you-go approach, often resulting in unforeseen costs and fewer capabilities. The culture needs to change so that these small projects can identify the software development pipelines as a clear requirement to their programs, and thus budgeted for. Smaller programs have to accept as a given that top-tier DevSecOps is a “must have” rather than a “nice to have.”

Larger-sized programs often have a different response to the large price tag associated with software development pipelines; they simply decide that it would be easier and cheaper to build their own software development infrastructure. These projects view the large cost associated with a software pipeline, then underestimate the costs of procuring, integrating, and securing their own environment. These programs then end up spending a great deal more than anticipated due to unrecognized engineering requirements, unanticipated ATO accreditation requirements, and license fees. These larger-sized programs need to recognize up front the inherent benefits of using an existent software development pipeline.

Many pipelines have been created to address unique challenges that DoD software developers may face, such as where the software will deploy, be it a fighter jet, an unmanned system, or a Naval ship. One NIWC Pacific-developed software pipeline, for example, is designed to support deployment of software on Naval vessels. This specific pipeline is tailored toward delivering software to afloat units where software delivery is extra complicated due to low bandwidth availability of shipboard networks, and the potential for at-sea degraded communications. This can and does present the software developer extra challenge, since they must ensure that not only must they pack all the features necessary into only the containers destined for a shipboard delivery, but also not eat up precious satellite bandwidth for prolonged periods. Building and maintaining such an environment is a tremendous undertaking. It requires unique understanding of the problem space, resources that understand not only how



DEVSECOPS works, but also how to rapidly accredit and release software that is secure and of low risk to the warfighter. Unfortunately, these unique requirements also come with a strong demand for specialized tools and strong engineering support for the software development environment, and ultimately, a large price tag.

At the moment, the Navy is currently dealing with a number of disparate software pipelines, each with a small base of users, procuring licenses in silos, and developing their own technical expertise. Given the increasing importance of software in all parts of the Navy, it makes business sense for the Navy to start viewing support of software development pipelines as a larger platform, with a dedicated funding stream to ensure the software pipelines are secure, agile, and don't cripple smaller programs with cost. Getting to this ideal requires the Navy to stop thinking of software development as an add-on or enabler, and start seeing it as a key asset in the force. Software pipelines may not be as tangible and photogenic as our fighter jets and aircraft carriers, but they are just as important in ensuring we will prevail in any future conflict.

The Navy needs to commit to "owning" their various software factories to truly realize the power they could bring to rapid delivery of capability, interoperability, and common service adoption. Such benefits easily justify the investment. The Navy engineering community universally understand the need to have such development environments, as well as how to manage and operate them. The aforementioned NIWC Pacific-software pipeline was created utilizing scant innovation dollars, as were most other pipelines. The Navy now needs to commit programmed dollars toward adopting and maintaining these software factories at scale.

## Summary

We are writing our software code in an agile manner, but we aren't acquiring it, securing it, or financing it in a similar manner. By moving our software engineering community to an agile mindset, it exponentially improved the speed and quality of the code we produce. Now, we need to move our acquisition, security, and investment communities towards that same agile mindset! We can't buy, secure or produce software to combat 21st century problems with a 20<sup>th</sup> century purchase, security and investment mindset.

## References

- Keller, K.M, Lytell, M.C., & Bharadwaj, S. (2022). *Personnel Needs for Department of the Air Force Digital Talent: A Case Study of Software Factories*. RAND Corporation. [https://www.rand.org/content/dam/rand/pubs/research\\_reports/RRA500/RRA550-1/RAND\\_RRA550-1.pdf](https://www.rand.org/content/dam/rand/pubs/research_reports/RRA500/RRA550-1/RAND_RRA550-1.pdf)
- Office of the Deputy Secretary of Defense. (2022). *Department of Defense Software Modernization Strategy*. <https://media.defense.gov/2022/Feb/03/2002932833/-1/-1/1/DEPARTMENT-OF-DEFENSE-SOFTWARE-MODERNIZATION-STRATEGY.PDF>
- United States Government Accountability Office. (2023). *Software Acquisition: Additional Actions Needed to Help DoD Implement Future Modernization Efforts*. <https://www.gao.gov/assets/gao-23-105611.pdf>



## PANEL 18. ACQUISITION WORKFORCE PERSPECTIVE FROM DACMS/DATMS

Thursday, May 11, 2023	
12:45 p.m. – 2:00 p.m.	<p><b>Chair: James P. Woolsey</b>, President, Defense Acquisition University</p> <p><b>Panelists:</b></p> <p><b>Ronald R. Richardson, Jr.</b>, Director, U.S. Army Acquisition Support Center and Director, Acquisition Career Management (DACM)</p> <p><b>Marianne Lyons</b>, U.S. Navy Director, Acquisition Talent Management (DATM)</p> <p><b>David Slade</b>, U.S. Air Force Director, Acquisition Career Management (DACM)</p> <p><b>Otis Lincoln</b>, 4th Estate Director Acquisition Career Management (DACM)</p>

**James P. Woolsey**—is President of the Defense Acquisition University (DAU), a position he has held since January 2014. In that role, he is responsible for delivery of learning products through the DAU regions, the Defense Systems Management College, and the College of Contract Management; curriculum development; online learning programs; learning technology; and library services for a major Department of Defense corporate university. DAU is strategically located within five geographical regions across the country and provides a global learning environment to develop qualified acquisition, requirements, and contingency professionals who deliver and sustain effective and affordable warfighting capabilities.

He previously served as the first Deputy Director for Performance Assessments (PA) in the office of Performance Assessments and Root Cause Analyses (PARCA). In standing up the PA organization, he created the processes and practices that allowed it to perform its statutory responsibility of assessing the progress of all Major Defense Acquisition Programs. The new office also made a substantial contribution to re-invigorating the Defense Acquisition Executive Summary process and provided the Under Secretary of Defense for Acquisition, Technology and Logistics with unique analyses to give him improved visibility into the status of the MDAP portfolio.

Mr. Woolsey was previously an Assistant Director in the Cost Analysis and Research Division of the Institute for Defense Analyses. His responsibilities included management of the division's cost analysis and research, and leadership of a wide range of cost and acquisition studies. His work included a congressionally-directed cost benefit analysis of the F-35 alternate engine, an evaluation of KC 767A lease prices, C-5 re-engineering costs and benefits, F-22 production readiness, Joint Air-to-Surface Standoff Missile costs, and space launch alternatives. Mr. Woolsey also served on a Defense Science Board Task Force on long-range strike.

Mr. Woolsey's other previous positions include service as a structures engineer for F/A-18 aircraft at Naval Air Systems Command, and work as an engineer for Lockheed Martin airlift programs in Marietta, GA.

**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 (3), the Meritorious Civilian Service Medal (2), the Civilian Service Achievement Medal, the Army Staff Identification Badge, and the Order of Military Medical Merit (O2M3). 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.

**Marianne Lyons**—Since April 2019 Ms. Lyons has served as the Department of the Navy Director, Acquisition Talent Management (DATM). She is the Navy and Marine Corps' lead for the professional development and management of the DoN's over 70,000 civilian and military acquisition workforce. Ms. Lyons is the chief advisor to the Assistant Secretary of the Navy for Research, Development, and Acquisition, and guides all matters relating to initiatives and other strategic efforts that improve the acquisition workforce through education, training, and career management. She began her career with the Navy in 1989 as a naval architect and progressed to ship design management. In 2003, she transitioned to Program Management and later became an Action Officer at the Office of DASN Ships for the Auxiliary and Amphibious Ships portfolio. Prior to the DATM she was the Deputy Program Manager for the LPD 17 Amphibious Transport Dock Ship Program in PEO Ships. Ms. Lyons has a Civil Engineering degree from Virginia Tech and a Masters in Business from the Florida Institute of Technology. She is PM Advanced and ETM Practitioner DAWIA certified.

**David Slade**—is the Director of Acquisition Career Management, Assistant Secretary of the Air Force for Acquisition (SAF/AQH). Mr. Slade is responsible for the integrated management of the acquisition workforce across all functional areas. He provides acquisition human resources policy and strategic planning while managing the training and development of civilian and military acquisition personnel Air Force-wide. Additionally, Mr. Slade ensures Air Force compliance and implementation of the Defense Acquisition Workforce Improvement Act (DAWIA) through management of the Acquisition Professional Development Program (APDP) and the Defense Acquisition Workforce Development Fund (DAWDF). Mr. Slade is also designated as the Career Field Manager for both military and civilian Scientists, Engineers, and Acquisition Program Managers. His team also provides personnel management services for the SAF/AQ Headquarters Staff.

Mr. Slade received an aerospace engineering degree from the University of Colorado and was commissioned through the Reserve Officer Training Corps in 1983. Following pilot training, he served as a forward air controller, flying the O2-A and an F-15C and AT-38 instructor pilot. He served as a Commander at the Squadron and Group levels. As a command pilot with over 3,600 flying hours, he flew 32 missions over Iraq during Operation DESERT STORM and has participated in Operations NOBLE EAGLE, NORTHERN WATCH and SOUTHERN WATCH.

Prior to his current assignment, Mr. Slade served as Director of Assignments, Headquarters Air Force Personnel Center, Randolph Air Force Base, TX. He was responsible for the assignment of more than



65,000 officers below the grade of Colonel and 285,000 enlisted personnel below the grade of Chief Master Sergeant.

Mr. Slade retired from active duty, in the rank of Colonel, after 29 years in November 2012 and entered Civil Service in January 2013.

**Otis Lincoln**—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 Contract Specialist and a warranted Contracting Officer on several procurements supporting multiple Directorates across DIA, he continued to expand his aperture within the acquisition community moving into the project and program management realm. In multiple capacities, he was responsible for the successful planning and execution of various multi-million dollar programs that included increasing acquisition exposure to industry, training and career development of the agency's acquisition workforce as well playing an integral part of the hiring and placement of new acquisition members and set career paths in the finance and acquisition field. Mr. Lincoln has also utilized his Defense Acquisition Workforce Improvement Act (DAWIA) expertise in support of the Navy Systems Management Activity (NSMA) having served as their DAWIA Program Director overseeing and managing their workforce by expanding their training, certification, and career development. 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. Currently, he serves as the Director, Acquisition Career Management for the 4th Estate (32 defense agencies/field activities) with oversight of statutory training, professional credentialing, continuous learning, and career development for more than 31,000+ acquisition workforce members.



## PANEL 19. THE ACQUISITION FRONTIER

Thursday, May 11, 2023

2:15 p.m. –  
3:30 p.m.

**Chair: Captain Andrew S. Gibbons, USN**, Deputy Chief Engineer for Space Capabilities, NAVWARSYSCOM 5.0 Office of the Chief Engineer

***Defense Acquisitions: DOD Should Take Additional Actions to Improve How It Approaches Intellectual Property***

Timothy J. DiNapoli, U.S. Government Accountability Office

Nathan Tranquilli, U.S. Government Accountability Office

Holly Williams, U.S. Government Accountability Office

***Social Engineering Impacts on Government Acquisition***

Katie Hyatt, MITRE Corporation

Zack Levenson, MITRE Corporation

***Comparative Analysis of Pathways to Changeability***

Aditya Singh, George Washington University

Zoe Szajnfarber, George Washington University

**Captain Andrew S. Gibbons, USN**— graduated from the U. S. Naval Academy in 1993 with a bachelor of science degree in mechanical engineering.

In February 1994, he reported to USS Mount Vernon (LSD-39) as well deck officer and Aft Engineering plant division officer. While stationed onboard USS Mount Vernon, he qualified as a surface warfare officer. He next reported to Assault Craft Unit Five in Camp Pendleton, Calif., where he qualified as a Small Craft officer-in-charge.

In January 1998, Gibbons reported to Naval Postgraduate School in Monterey, Calif., where he received his masters of science in mechanical engineering. While there, he was selected to be an engineering duty officer. During this time, he received a master's of science in software engineering from National University.

In 2004, Gibbons reported to Naval Information Warfare Systems Command (NAVWAR), Installations and Logistics Directorate as the assistant program manager for Installations for Pacific Fleet ships. A year later, he stood up the installation management responsibilities for Carrier and Air Integration program office (PMW-750).

In January 2007, he deployed to Baghdad as an individual augmentee in support of Operation Iraqi Freedom. While there, he was assigned as the director of fielding for vehicle-borne Counter Radio-Controlled Electronic Warfare systems employed by Multi-National Corps-Iraq.

From 2008 to 2011, Gibbons was officer-in-charge of NAVWAR Systems Facility Pacific, Japan. He also served as Naval Sea Systems Command Naval Systems Engineering Directorate's strike force interoperability officer for Undersea Warfare. While officer-in-charge, he was responsible for all NAVWAR Enterprise activities within the Far East area of responsibility.

Gibbons reported to Program Executive Officer, Command, Control, Communications, Computers and Intelligence in February 2011 to assume duties as the assistant program manager for the Consolidated Afloat Networks and Enterprise Services program. In September 2013, he became the deputy program manager for Ship Integration program office (PMW-760). In December 2013, Gibbons became the acting deputy program manager for the Tactical Networks program office (PMW-160). From September 2014 to November 2017, he served as the program manager for PMW-750.



Gibbons assumed command as the program manager for Communications and Global Positioning System (GPS) Navigation program office (PMW/A-170) in November 2017. He is responsible for the development and fielding of the Navy's satellite communications, tactical communications, and GPS navigation systems across all naval platforms.

Gibbons is a member of the Defense Acquisition Corps and is Level III certified in Program Management, Production, Quality & Manufacturing and Engineering.

His awards include the Legion of Merit, Bronze Star, three Meritorious Service Medals, two Navy Commendation Medals, four Navy Achievement Medals and various campaign and service awards.





# Defense Acquisitions: DOD Should Take Additional Actions to Improve How It Approaches Intellectual Property

**Timothy J. DiNapoli**—is a managing director at the U.S. Government Accountability Office. [dinapolit@gao.gov]

**Nathan Tranquilli**—is an assistant director at the U.S. Government Accountability Office. [tranquillin@gao.gov]

**Holly Williams**—is a senior analyst at the U.S. Government Accountability Office. [williamshn@gao.gov]

## Abstract

The Department of Defense (DOD) acquires and licenses intellectual property (IP) for its cutting-edge weapon systems. Yet, the DOD often does not acquire the IP it needs to operate and maintain those systems, which can lead to surging costs later. In 2019, the DOD assigned specific IP responsibilities to organizations within the department. However, we found the DOD had not fully addressed how the IP Cadre—the DOD's new group of specialized experts—will fulfill all of its responsibilities. The IP Cadre faced uncertainty in three areas: (1) The DOD planned to provide the director of the IP Cadre and his team in the Office of the Secretary of Defense (OSD) with funding for five positions through Fiscal Year 2023. IP Cadre members told us the temporary positions could present future staffing obstacles. (2) The members of the IP Cadre at the OSD expect to tap into a larger pool of IP experts across the DOD. However, the DOD had not detailed how the OSD team will work with these experts. (3) DOD officials said the department lacked sufficient expertise in two key areas—IP valuation and financial analysis. We made four recommendations to the DOD. The DOD concurred with all four recommendations. Our original report is accessible at [www.gao.gov/products/gao-22-104752](http://www.gao.gov/products/gao-22-104752).

**Keywords:** Department of Defense, intellectual property, IP Cadre





## Methodology

In this report, we (1) examine issues addressed in the Department of Defense's (DOD's) intellectual property (IP) instruction, (2) examine the extent to which the DOD has implemented the IP instruction, (3) assess the Defense Acquisition University's (DAU's) efforts to improve IP training, and (4) describe the DOD's efforts to develop a capability to track the IP the department has acquired and licensed. We reviewed guidance, reports, and documentation on IP issues; interviewed DOD personnel, military officials, and industry groups; and reviewed the existing regulatory and agency frameworks related to IP.

## Background

Companies protect their IP in several ways, including through the use of patents, trademarks, copyrights, and trade secrets. See Figure 1 for more details on these types of IP categories.



			
<b>Patents</b>	<b>Trademarks</b>	<b>Copyrights</b>	<b>Trade secrets</b>
Provide exclusive rights to make, use, import, sell, and offer for sale an invention for up to 20 years	Protect words, names, symbols, sounds, or colors that distinguish goods and services	Protect works of authorship, such as software, writings, music, and works of art that have been tangibly expressed	Protect information that companies keep secret to give them an advantage over their competitors

Note. The source of the data is Government Accountability Office (GAO) analysis of U.S. Patent and Trademark Office guidance. See GAO (2021c) for the original figure.

**Figure 1. Types of Intellectual Property**

Congress has enacted several laws related to IP over the past several decades.<sup>1</sup> For example, in 1980, Congress passed the Patent and Trademark Law Amendments Act, which addressed patent rights in inventions made with federal assistance. The act addressed the rights of small businesses, universities, and other nonprofit organizations and generally gave them the right to retain title to subject inventions, provided they adhered to certain requirements. A subject invention was defined as any invention of the contractor conceived or first actually reduced to practice in the performance of work under a funding agreement. In 1983, an executive order stated that it granted to all contractors, regardless of size, the title to patents made in whole or in part with federal funds (Reagan, 1983). The following year, Congress passed the Defense Procurement Reform Act, which required that regulations address rights in technical data, including procedures to validate any proprietary data restrictions asserted by contractors.

The Federal Acquisition Regulation (FAR) and Defense Federal Acquisition Regulation Supplement (DFARS) implement these laws and provide the basic regulatory framework governing how the DOD may license and acquire contractor IP.<sup>2</sup> For example, these regulations describe how the government may obtain technical data rights and licenses to computer software.<sup>3</sup> In general, using another entity’s IP requires permission, and the government typically uses licenses to obtain permission and define the scope of its rights to use a particular contractor’s IP. The federal government also obtains data rights when the development of IP was funded by the government—in whole or in part—and the types of data rights obtained by the government generally depend on how the IP was developed and funded.<sup>4</sup> Federal acquisition regulations established data rights, organized in three categories in Figure 2.<sup>5</sup>

<sup>1</sup> In this report, we use the definition of intellectual property from DOD (2019): information, products, or services that are protected by law as intangible property, including data (e.g., technical data and computer software), technical know-how, inventions, creative works of expression, and trade names.

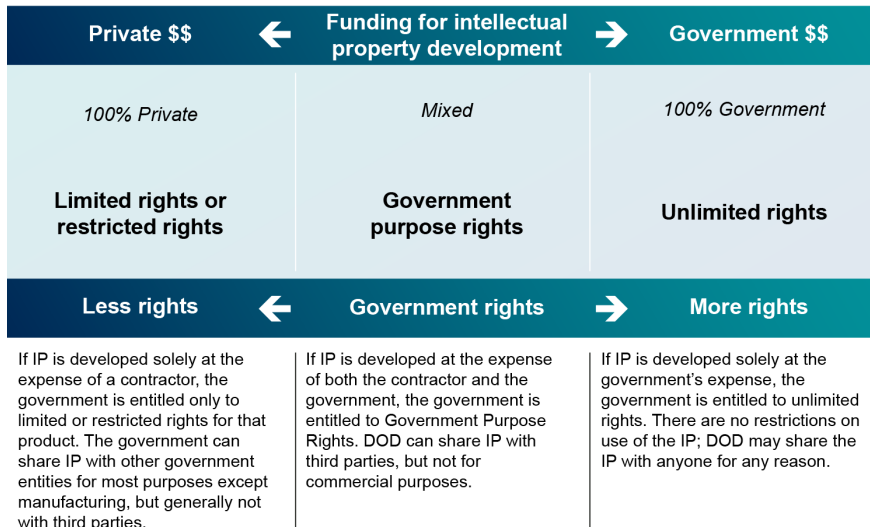
<sup>2</sup> For example, see 10 U.S.C. §§ 2320 & 2321; DFARS § 252.227.71 (Rights in Technical Data); DFARS § 252.227.72 (Rights in Computer Software and Computer Software Documentation); and DFARS 252.227-7013, -7014, -7015, -7017, -7018, -7019, -7026, -7027, -7030, and -7037.

<sup>3</sup> Technical data includes any recorded information of a scientific or technical nature such as product design or maintenance data and computer software documentation. Computer software includes executable code, source code, code listings, design details, processes, flow charts, and related materials. See DFARS 252.227-7013, -7014.

<sup>4</sup> Data rights are also determined by whether the item, process, or software is commercial or noncommercial, and the purpose of the data in question.

<sup>5</sup> The government obtains technical data and license rights to use IP assets in accordance with the FAR, agency supplements to the FAR, and any specifically negotiated licenses in the contract. These rights control how the government can use, disclose, or reproduce contractor-owned information.





*Note.* The source of the data is GAO analysis of DOD documentation. The table does not represent every license right available to the DOD within federal acquisition regulations. "Limited rights" refer to those rights in technical data, and "restricted rights" refer to those rights in noncommercial software. See GAO (2021c) for the original figure.

**Figure 2. Types of License Rights for Intellectual Property**

Regardless of the source of funding used for IP development, the government obtains unlimited rights to form, fit, and function data and data necessary for operation, maintenance, installation, and training purposes. Not included within those exceptions are detailed manufacturing or process data (DMPD), including the steps, sequences, and assembly used by manufacturers to produce an item.

### Recent Congressional Action to Improve How the DOD Acquires and Manages IP

In recent years, Congress included numerous requirements in national defense authorization acts (NDAA) for the DOD to assess and improve how it acquires and manages IP, including technical data needed to manufacture equipment or systems. For example, the National Defense Authorization Act for Fiscal Year 2016 directed the DOD to establish an advisory panel of industry and government experts—known as the 813 Panel—to provide recommendations to help ensure that statutory and related regulatory requirements pertaining to technical data were structured to best serve the interests of taxpayers and the national defense. Among other things, the 813 Panel found that two-thirds of system life-cycle costs typically occur in a system's sustainment phase; thus, it is critical for federal agencies to identify the necessary IP and licenses during source selection to thoroughly assess proposals during competition. We similarly reported that a weapon system's operating and support costs account for approximately 70% of a weapon system's total life-cycle cost (GAO, 2018).

The Fiscal Year 2016 NDAA also directed the DOD to commission an independent review of its regulations and practices addressing the use of IP rights of private sector firms, among other things. In a May 2017 report to Congress, the Institute for Defense Analyses (2017) found that there are often only two or three capable suppliers for key DOD systems, and that providers have a great deal of leverage in IP negotiations once a selection is made. The report stated that, given the long-term value of these contracts, contractors sometimes bid low under the assumption that they will secure profitable sustainment opportunities in the future. Figure 3 includes details of IP-related provisions from recent NDAs and actions taken to address them.



	Selected requirements	Outcomes
NDA FY 2016 Public Law 114-92	<b>Section 813:</b> Required DOD to establish a government-industry panel to review 10 U.S.C. §§ 2320 and 2321 regarding rights in technical data.	<b>Section 813:</b> Section 813 panel submitted its report to Congress in November 2018.
	<b>Section 821:</b> Amended Title 10 by adding § 2431a, requiring acquisition strategies to include IP for major defense acquisition programs, major automated information systems, and major systems.	<b>Section 821:</b> Implemented in the DOD 5000 series guidance.
	<b>Section 875:</b> Directed DOD to conduct an independent review and provide a report on its regulations, practices, and sustainment requirements related to government access and use of private sector IP, among other things.	<b>Section 875:</b> Institute for Defense Analyses conducted an independent review, and issued its report in May 2017.
NDA FY 2017 Public Law 114-328	<b>Section 809:</b> Amended 10 U.S.C. § 2320 regarding technical data rights relating to interfaces, including major systems interfaces, when funded with private or mixed funding.	<b>Section 809:</b> Statutory amendments currently being implemented by several open Defense Federal Acquisition Regulation Supplement cases.
	<b>Section 844:</b> Directed DOD to review decisions regarding IP requirements for major defense acquisition programs.	<b>Section 844:</b> MITRE Corporation issued a report in November 2017.
NDA FY 2018 Public Law 115-91	<b>Section 802:</b> Amended Title 10 by adding § 2322, which directed DOD to develop policy on the acquisition or licensing of IP, and to establish the IP Cadre.	<b>Section 802:</b> DOD Instruction 5010.44 was issued in October 2019, and DOD subsequently established the IP Cadre.
NDA FY 2020 Public Law 116-92	<b>Section 801:</b> Authorized DOD to conduct a 3-year pilot program to assess mechanisms to evaluate IP.	<b>Section 801:</b> Pilot began and DOD issued the first of three reports to Congress in March 2021.
	<b>Section 838:</b> Amended FY18 NDAA § 802, and directed DOD to submit a report that describes the leadership structure of the IP Cadre and the activities and efforts undertaken by the IP Cadre.	<b>Section 838:</b> DOD issued its report in April 2020.
NDA FY 2021 Public Law 116-283	<b>Section 801:</b> Directed each service acquisition executive to report to Office of the Secretary of Defense leadership on how it is addressing operation and sustainment risks associated with access to IP.	<b>Section 801:</b> DOD issued its report in August 2021.
	<b>Section 802:</b> Amended 10 U.S.C. § 2337 to require each applicable system has an approved life-cycle sustainment plan that includes IP.	<b>Section 802:</b> DOD is updating its guidance.
	<b>Section 804:</b> Amended 10 U.S.C. § 2320 regarding the type of technical data rights the government will acquire pertaining to modular system interfaces developed either exclusively at private expense or with mixed funding.	<b>Section 804:</b> DOD is updating its guidance and regulations.

DOD = Department of Defense      IP = Intellectual property      OSD = Office of the Secretary of Defense  
FY = Fiscal year      NDAA = National Defense Authorization Act      USC = United States Code

Note. The source of the data is GAO analysis of the NDAs for Fiscal Years 2016–21. See GAO (2021c) for the original figure.

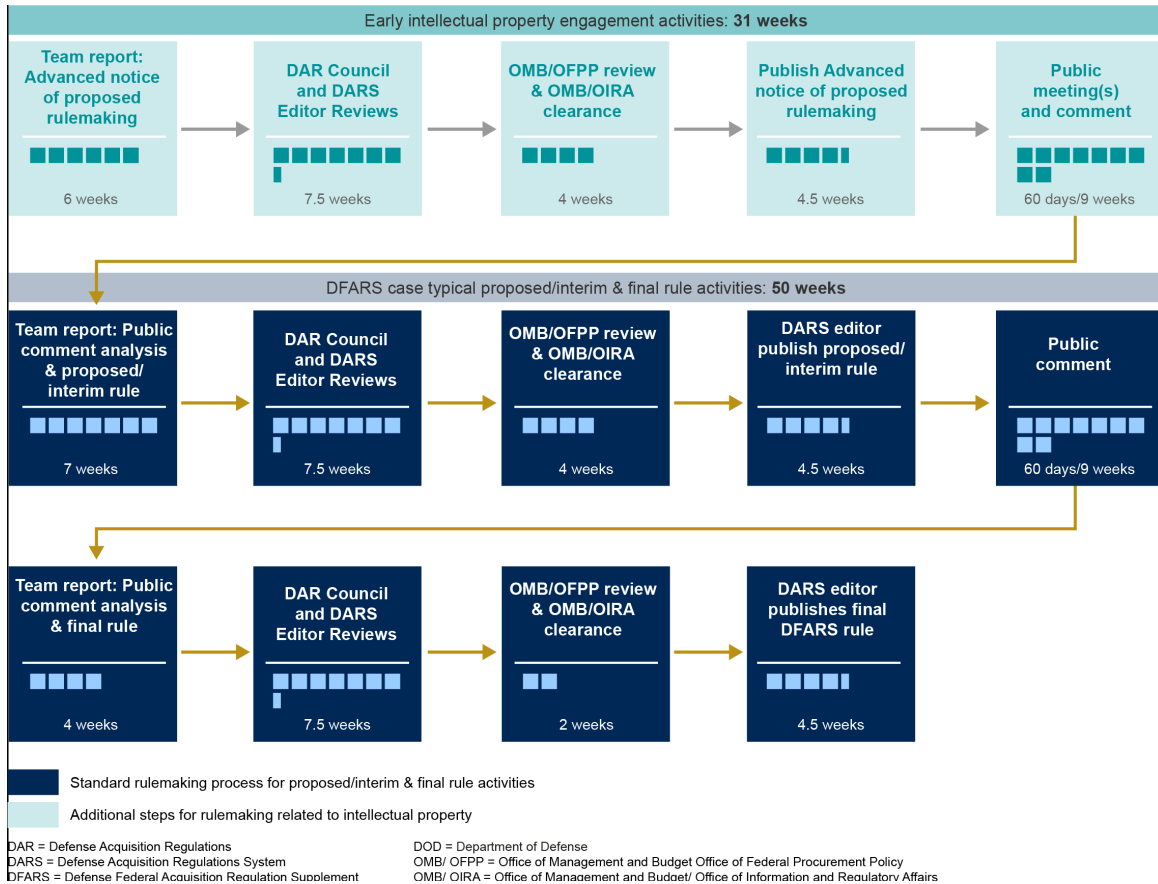
**Figure 3. Summaries of Key IP-Related NDAA Provisions from Fiscal Years 2016–2021**

NDAA provisions, including those related to IP, can result in changes to federal or agency acquisition regulations. Regulatory changes to the FAR and DFARS occur through the federal rulemaking process, which includes opportunities for private sector representatives to provide input on how regulations should be updated. The DOD has a dedicated team—the Patents, Data, and Copyrights Team, chaired by the Director of the IP Cadre—that oversees regulatory changes involving IP in the DFARS. That team is currently working on eight proposed regulatory changes related to IP—based mostly on NDAA direction—including changes involving specially negotiated licenses and small business data.<sup>6</sup>

<sup>6</sup> A specially negotiated license is required when the standard data rights arrangements defined in the FAR, DFARS, or by a commercial entity are modified by mutual agreement between a contractor and the government.



We previously reported that regulatory changes involving complex topics like IP often take longer than the DOD's standard 12-month process (GAO, 2019). The DOD extended the time frames of the process to make the DFARS changes recommended by the Section 813 Panel to provide industry and the public additional opportunities to provide input early in the process. See Figure 4, which illustrates the extended rulemaking timeline.



Note. The source of the data is GAO analysis of DOD documentation. The extended process applies to DFARS changes recommended by an advisory panel of industry and government experts that the DOD established in response to the NDAA for Fiscal Year 2016. This panel is commonly known as the 813 Panel. See GAO (2021c) for the original figure.

**Figure 4. DOD's Extended Rulemaking Timeline for Selected Regulatory Changes Involving Intellectual Property (IP)**

### Prior GAO Reporting

Over the past 30 years, we have reported on the complexities of acquiring IP and associated rights—particularly technical data—for weapon systems (GAO, 1991, 2002, 2006, 2010, 2011). When IP rights are not acquired—because, for example, needs were not assessed—consequences may include sustainment cost growth, maintenance challenges, and the inability to competitively purchase follow-on systems and spare parts. We found that the military departments have experienced each of these consequences due to a lack of technical data or data rights. For example,

- In July 2006, we reported that a lack of technical data rights for several Army weapons systems disrupted sustainment plans intended to achieve cost savings and meet legislative requirements for depot maintenance capabilities (GAO, 2006). For example, when acquiring the Stryker family of vehicles, the Army did not obtain technical data



rights needed to develop competitive offers for the acquisition of spare parts and components. Following the initial acquisition, the program analyzed alternatives to the contractor's support strategy and attempted to acquire rights to the manufacturer's technical data package, which describes the parts and equipment in sufficient technical detail to allow the Army to use competition to lower the cost of parts. The contractor declined to sell the Stryker's technical data package to the Army. According to an Army Audit Agency report, the project office stated that the cost of the technical data, even if available, would most likely be prohibitively expensive at that point in the Stryker's fielding, offsetting any cost savings resulting from competition.

- In September 2014, we reported that the F-35 program did not acquire technical data needed to compete a subsequent award of the F-35 or its subsystems under its previously awarded system development contract (GAO, 2014). We also reported that program officials did not have an understanding of the technical data rights the DOD owned, what technical data rights it might still need, or how much it would cost to acquire those data rights to support the future sustainment of F-35 aircraft. We recommended that the F-35 program should, among other things, develop a long-term IP strategy that identifies (1) current levels of technical data rights ownership by the federal government, and (2) all critical technical data rights and their associated costs. The DOD concurred with the recommendation and stated that the program planned to address these technical data rights issues as part of the program's future sustainment strategy. However, in July 2021, we found that the F-35 program still does not have a comprehensive understanding of the technical data rights it currently owns, what technical data rights it may still need, or how much it will cost to acquire data needed to support F-35 sustainment (GAO, 2021a, 2021b).
- In March 2020, we found that a lack of technical data contributed to sustainment problems for several Navy ship programs, and that focusing on sustainment earlier in the acquisition process could save billions of dollars (GAO, 2020).
- Navy officials stated they did not have a clear understanding of all the IP needed until ship systems broke and Navy maintainers could not repair the systems with the IP available to them. Navy ship maintainers told us that once a ship is delivered it is often too late to implement strategies or agreements with manufacturers to get the IP needed to fully sustain the ship systems at an affordable price. We made several recommendations to the Navy, including that the Assistant Secretary of the Navy for Research, Development, and Acquisition should ensure that all shipbuilding programs develop and update life-cycle sustainment plans, in accordance with DOD policy, to demonstrate how they will affordably operate and maintain ship classes during sustainment. According to the DOD's acquisition policy in place at the time of our review, shipbuilding programs should document IP strategies early in acquisition planning to assess technical data needs and to determine what IP deliverables and license rights the program must acquire from contractors (DOD, 2013, 2018). The Navy agreed with this recommendation but has not addressed it yet.

### **DOD's IP Instruction Highlights Six Core Principles but Does Not Address DOD's Ability to Obtain Detailed Manufacturing or Process Data**

The DOD integrated existing IP guidance and requirements, highlighted six core principles, and set a department-wide expectation for DOD personnel to prioritize IP planning early in the acquisition life cycle in its 2019 IP instruction (DOD, 2019). According to military officials, the IP instruction is helpful for setting expectations, but it does not address the DOD's ability to pursue DMPD, which the department often needs to repair and competitively re-procure its weapons systems.



## **DOD's IP Instruction Integrated Existing IP Guidance and Requirements and Highlighted Six Core Principles**

In developing the IP instruction, the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD [A&S]) integrated existing requirements from prior DOD guidance into a single document. The IP instruction applies specifically to IP that is acquired, created by or for, or used by or on behalf of the DOD for purposes relating to the acquisition, operation, maintenance, modernization, and sustainment of defense products and services.<sup>7</sup> Prior requirements included the DOD's 5000 series acquisition guidance and the DOD Open Systems Architecture–Data Rights Team *IP Strategy Guidance* (DOD, 2013; Open Systems Architecture-Data Rights Team, 2014). These earlier documents, for example, require program managers to establish and maintain an IP strategy as part of their acquisition planning, and to identify and manage IP-related issues throughout the program's life cycle.

The IP instruction also presented six core principles that are rooted in laws, regulations, and earlier DOD guidance:

1. Integrate IP planning fully into acquisition strategies to account for long-term effects on competition and affordability.
2. Ensure acquisition professionals have relevant IP knowledge for their official duties to support critical, cross-functional coordination during IP acquisition planning.
3. Negotiate specialized IP deliverables and associated license rights when doing so more effectively balances DOD and industry interests than standard license rights.
4. Communicate clearly and effectively with industry regarding IP expectations and sustainment objectives.
5. Respect and protect IP funded by both the private sector and the government.
6. The government must ensure delivery of IP deliverables and corresponding licenses.

The IP instruction further identified roles and responsibilities for key DOD organizations and important elements of IP strategies, such as identifying system interfaces and considering use of specially negotiated licenses and modular open systems approaches. It also emphasized a department-wide expectation that DOD personnel should prioritize IP planning early—specifically during the initial phases of the acquisition life cycle—when DOD has the most leverage to obtain the IP rights it needs at a fair and reasonable price through competition.

To develop the IP instruction, OUSD (A&S) indicated that it solicited input from relevant DOD offices, including acquisition and sustainment offices from each of the military departments. OUSD (A&S) also established an IP working group that reviewed and implemented stakeholder comments and considered industry input obtained during the proceedings of the 813 Panel. The working group consisted of a cross-functional team with experts on requirements, acquisition, sustainment, research and development, engineering, and training from OSD, the military departments, and other DOD components.

### **DOD's IP Instruction and Department-Wide Guidance Do Not Directly Address DOD's Ability to Acquire Detailed Manufacturing or Process Data**

While the IP instruction emphasizes the importance of acquiring and licensing IP early in the acquisition process, officials from the IP Cadre and military departments stated that the instruction and department-wide guidance do not address the DOD's ability to acquire DMPD. According to these officials, some DOD personnel believe that the current regulations prevent them from requesting DMPD the department often needs for sustainment activities. However, IP Cadre officials told us that DOD personnel are, in fact, allowed to request these data. IP Cadre

---

<sup>7</sup> DOD Instruction 5010.44 does not apply to patent licensing or other technology transfer of U.S. government-owned IP or technology covered by DOD Directive 5535.03 and DOD Instruction 5535.8, or branding and trademark licensing by DOD components covered by DOD Directive 5535.09 and DOD Instruction 5535.12.



officials told us that the misunderstanding hinders cost-effective re-procurement and sustainment of DOD systems.

The 813 Panel report and IP Cadre officials attributed this misunderstanding, in part, to tensions in the regulatory framework governing IP. In June 1995, the DOD issued DFARS sections that implement two parts of the *U.S. Code* related to the acquisition of DMPD.<sup>8</sup> IP Cadre officials told us that the first DFARS section establishes that the DOD *cannot condition a contract award* on a vendor granting rights to DMPD, which they said may discourage DOD personnel from requesting it. According to the same officials, the second section, however, emphasizes what actions the DOD may take to acquire DMPD. Members of the IP Cadre told us that the DOD *can consider the effects* of acquiring rights to DMPD during source selections, and that these considerations are a more effective negotiation tool in a competitive environment. This position is consistent with findings from the 813 Panel. The panel reported that vendors' data deliverables and associated licenses should be considered during source selection, and that the DOD would not be forcing vendors to give up any license rights in violation of statute by asking that IP costs be included in the proposal (National Defense Industry Association, 2018).

The 813 Panel further found that the DOD's past source selections often did not include an evaluation factor for IP, particularly technical data and associated license rights. As a result, the DOD did not evaluate the value of IP during proposal evaluation. IP Cadre officials told us they want DOD personnel to be equally familiar with both DFARS sections and to use a balanced approach when considering the acquisition of DMPD. IP Cadre officials also want DOD personnel to evaluate the cost of requested IP deliverables and license rights during source selection in the ways that the regulations permit. However, the 2019 IP instruction does not reference either DFARS section or clarify the DOD's ability to acquire DMPD.<sup>9</sup>

IP Cadre officials told us the instruction does not address DMPD because DOD instructions generally do not address specific, individual challenges. They said that other types of guidance often address these types of challenges. However, we found that the DOD's current department-wide guidebook for acquiring IP rights from commercial companies also does not address how DOD officials can consider the effects of acquiring rights to DMPD during source selections (Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, 2001). In an April 2020 report to Congress, the DOD identified that it plans to publish a new department-wide IP guidebook intended to explain IP-related regulations and policies (Office of the Under Secretary of Defense for Acquisition and Sustainment, 2020). However, the report did not identify whether the guidebook will address how government personnel may pursue DMPD during source selections. Members of the IP Cadre told us that they expect the DOD will publish the guidebook in the first quarter of Fiscal Year 2022 and that they believe it should address common misunderstandings related to DMPD.

Standards for Internal Control in the Federal Government state that management should internally communicate information necessary to achieve objectives. In developing the next iteration of its guidebook, DOD leadership, specifically the Under Secretary of Defense for Acquisition and Sustainment, has an opportunity to clarify how DOD personnel should account for the two DFARS sections addressing DMPD and, ultimately, improve the re-procurement and sustainment of DOD systems.

---

<sup>8</sup> See DFARS § 227.7103-1(c) and § 227.7103-10(a)(5) implementing 10 U.S.C. §§2320, 2321. Congress provided limited exceptions for technical data, allowing for unlimited government rights in "form, fit, and function" data and technical data necessary for "installation, operation, maintenance, or training" purposes. See 10 U.S.C. § 2320(a)(2)(A)(i). However, Congress excluded contractors' protected manufacturing data, known as "detailed manufacturing or process data." See 10 U.S.C. § 2320 (a)(2)(C)(ii).

<sup>9</sup> We found that a 2015 Army guide cites both DFARS sections and clarifies that, while government personnel cannot require additional data rights from vendors, they can evaluate the effect of offered rights for technical data and computer software. However, this guidance has limited visibility across DOD. See U.S. Army Product Data & Engineering Working Group, 2015.





## DOD Is Taking Steps to Implement the IP Instruction but Has Not Fully Identified How the IP Cadre Will Meet Its Assigned Responsibilities

The DOD's IP instruction assigns specific responsibilities to several organizations within the department, including the DOD's Office of General Counsel, DAU, the military departments, and the DOD's new IP Cadre. We found that, while these organizations are working to meet their responsibilities, the DOD has not yet determined how the IP Cadre will fulfill all of its assigned responsibilities. In particular, the DOD has not ascertained whether the IP Cadre, whether by itself or in coordination with other entities within the DOD, has the capacity to conduct IP valuation or provide program support. Additionally, the DOD has not determined how the IP Cadre will be funded and staffed in the future.

### Organizations Identified in DOD's IP Instruction Are Taking Steps to Meet Their Responsibilities

The DOD's IP instruction identifies specific responsibilities for the Assistant Secretary of Defense for Acquisition, the DOD's Office of General Counsel, and the president of DAU. Our review of documentation provided by the DOD and interviews with cognizant DOD officials found that these organizations are taking various actions to meet their responsibilities. See Table 1.

**Table 1. Actions Taken to Address Key Responsibilities Established in DOD's IP Instruction**

DOD official/office	Responsibilities	Examples of actions taken
Assistant Secretary of Defense for Acquisition (ASD[A])	Serve as senior DOD official overseeing development and implementation of DOD IP policy and guidance	ASD(A) appointed a Director of the IP Cadre, with responsibility for department-wide implementation of DOD IP policy and guidance.
	Manage a cadre of experts (IP Cadre) in IP acquisition and licensing Coordinate the IP Cadre's development and activities	ASD(A) also established a support team under the Director of the IP Cadre, consisting of four temporary government positions and eight support contractors.
Office of General Counsel	Provide legal advice and services in support of DOD's IP instruction and in support of the IP Cadre	DOD General Counsel assigned a staff member to the team supporting the Director of the IP Cadre, as Associate General Counsel for IP, to advise and support IP acquisition, licensing, and management.
President of Defense Acquisition University (DAU)	Develop and update curricula and reference materials (in coordination with the IP Cadre)	DAU collaborated with the IP Cadre to develop new IP training and update existing IP training. In addition, DAU
	Provide IP training	<ul style="list-style-type: none"> <li>finalized a 5-year strategic plan for IP training;</li> </ul>
	Continuously improve and tailor IP training	<ul style="list-style-type: none"> <li>established an IP Community of Practice web portal; and</li> <li>established a foundational IP credential using DAU's online IP courses.</li> </ul>

Note. The sources of these data are GAO analysis of DOD Instruction 5010.44, DOD responses to a structured checklist, and related documentation.



Additionally, the DOD’s IP instruction identifies several specific responsibilities for the military departments, such as incorporating IP planning into acquisition strategies and source selections. DOD officials told us that the military departments are leveraging DOD and component-specific guidance to consider IP factors during source selections and to incorporate IP planning into their acquisition strategies, among other things. Table 2 provides examples of actions the military departments have taken to meet requirements from the IP instruction, according to DOD officials and our review of documentation provided by the DOD and the military departments.

**Table 2. Examples of How Military Departments Are Addressing Responsibilities Established in DOD’s IP Instruction**

<b>Responsibilities from IP instruction</b>	<b>Air Force approach</b>	<b>Army approach</b>	<b>Navy approach</b>
Ensure program personnel engaged in all stages of the acquisition life cycle have relevant knowledge of IP matters, as appropriate.	Air Force established component-specific IP guidance that sets an expectation for personnel at all stages of the acquisition life cycle to be familiar with relevant IP policy and guidance.	Army established component-specific IP guidance that directs staff at all stages of the acquisition life cycle to follow best practices for negotiating customized IP agreements with industry.	The Navy follows DOD guidance and component-specific acquisition guidance for program reviews and acquisition strategy approval processes to ensure that relevant personnel consider and use appropriate IP techniques and practices.
Incorporate consideration of types of IP deliverables and associated license rights into source selection evaluation factors and as negotiation objectives in sole-source awards, as appropriate.	Air Force IP guidance identifies IP as a source selection evaluation factor and directs contracting personnel and program officials to review and validate contractors’ restrictive assertions, when appropriate.	Army IP guidance directs staff to identify the types of IP and license rights needed and to consider including availability and delivery of identified data and rights as a source selection evaluation factor.	Navy open architecture guidance directs personnel to consider IP deliverables as part of proposal evaluation and for source selection.
Incorporate IP planning elements into acquisition strategies, emphasizing long-term analysis and planning during the earliest phases of the program, and preserving flexibility in the program sustainment strategy.	Air Force IP guidance addresses early IP planning, involving cost and benefits analysis, and the Air Force uses tools such as checklists and approval processes to ensure that proper IP planning has occurred.	Army guidance establishes that acquisition strategies should include IP strategies and notes that they should be developed as early as possible and continuously updated to reflect evolving conditions and needs over a system’s life cycle.	Navy uses the DOD’s Adaptive Acquisition Framework policy—and is in the process of updating its own acquisition guidance—to direct acquisition personnel to include a technical data plan in a program’s IP strategy.
Communicate clearly and effectively with industry on IP matters early in the program life cycle.	Air Force IP guidance directs personnel to communicate IP needs and strategies to vendors and to use tools such as checklists to ensure IP matters are considered when communicating with vendors.	Army guidance states that Army personnel should communicate with industry early in the acquisition process and share appropriate information from IP strategies.	Navy follows the DOD’s acquisition planning procedures, which require program offices to document their IP goals; Navy commands also have practices for sharing IP goals with vendors via industry days and draft solicitations.

Note. The sources of these data are GAO analysis of DOD Instruction 5010.44, DOD responses to a structured checklist, and related documentation including *Air Force Data Rights Guidebook* and Army Directive 2018-26.



## DOD Has Not Identified Strategies or Resources for the IP Cadre to Fully Meet Its Assigned Responsibilities

The DOD's IP instruction identifies several responsibilities for the IP Cadre that involve strategic activities and providing program support. See Table 3.

**Table 3. IP Cadre Responsibilities in DOD's IP instruction**

Type of responsibilities	Responsibilities
Strategic activities	Interpret and provide counsel on laws, regulations, and policies relating to IP Coordinate with DAU, academia, and industry to improve IP training Facilitate coordination and consistency across the DOD for determining the IP deliverables and rights necessary for operation, maintenance, modernization, and sustainment
Program support	Advise and assist acquisition programs with the development of acquisition, product support, and IP strategies Conduct or assist acquisition programs with financial analysis and valuation of IP Assist acquisition programs in drafting solicitations, contracts, or other transactions Address management of IP deliverables and IP rights to create a competitive environment Assist program interactions with contractors, including negotiations on solicitations and awards Conduct or assist acquisition programs with mediation if technical data are not delivered or do not meet contract terms

*Note.* The source of these data is GAO analysis of DOD Instruction 5010.44.

In addition to the responsibilities identified in Table 3, the DOD's IP instruction directs the ASD(A) to ensure that the IP Cadre is adequately staffed to provide seven areas of expertise:

1. Acquisition,
2. Contracting,
3. Engineering,
4. Law,
5. Logistics,
6. Financial analysis, and
7. Valuation.

The DOD has provided some information on its strategy for the IP Cadre to meet its responsibilities in two reports to Congress (Office of the Under Secretary of Defense for Acquisition and Sustainment, 2020, 2021). For example, these reports identify certain planned activities and provide information about the IP Cadre's existing areas of expertise. However, the DOD has not yet detailed

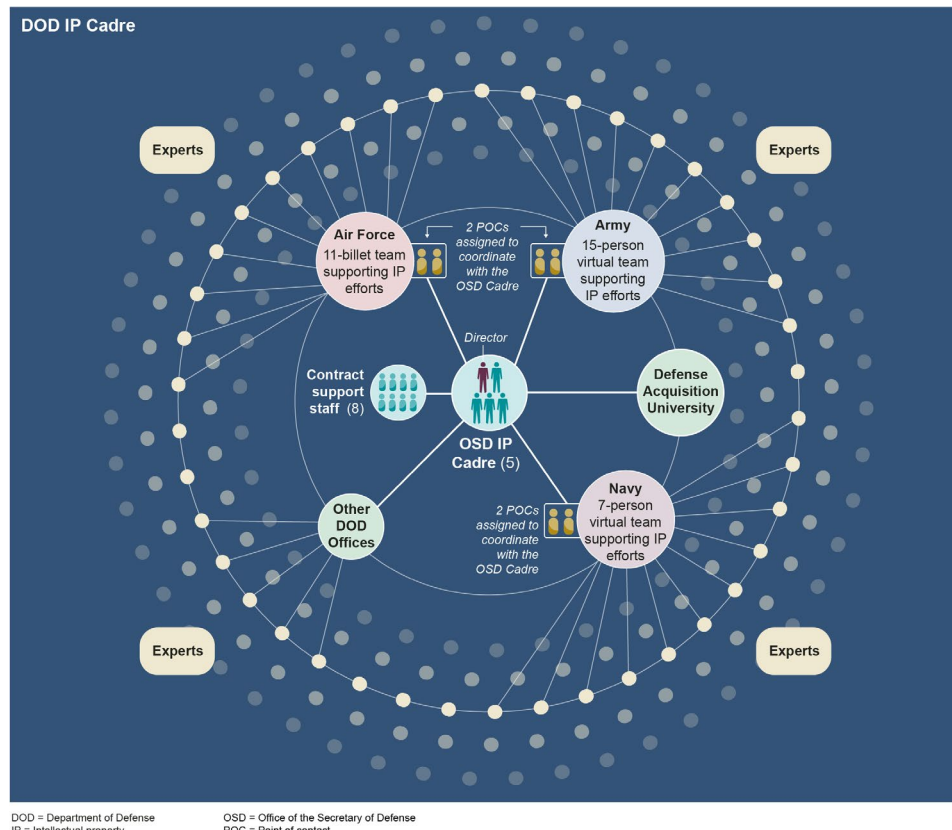
- how the IP Cadre will provide program support,
- how the IP Cadre will provide two key areas of expertise, and
- future funding and staffing needs for the IP Cadre.

### Program Support

The IP instruction assigns the IP Cadre responsibility for providing support to programs, such as assisting with the development of acquisition planning and product support planning.



The IP Cadre director told us that the IP Cadre will work to meet this responsibility through the federated structure described in the two reports to Congress. Specifically, in April 2020 and March 2021, the DOD described the IP Cadre’s organizational structure as a federated model that involves two cadres: the five-billet OSD IP Cadre situated in OUSD (A&S), which is part of a larger, less clearly defined network of DOD IP experts that span the entire department (Office of the Under Secretary of Defense for Acquisition and Sustainment, 2020, 2021). According to DOD officials, from October 2019 to September 2021, the DOD primarily focused on establishing the OSD IP Cadre. Figure 5 presents the IP Cadre’s proposed federated structure, including the OSD IP Cadre’s central role, contracted support staff, DAU, and dedicated points of contact at the military departments.



Note. The source of this figure is GAO analysis of DOD documentation. In addition to the IP Cadre, DAU coordinates with military departments, industry, academia, and the public on its intellectual property training and learning materials. See GAO (2021c) for the original figure.

**Figure 5. Proposed Federated Structure for DOD’s IP Cadre**

Under this approach, the five OSD IP Cadre members expect to tap into a much larger pool of IP experts from among the thousands of personnel that make up the DOD’s acquisition workforce. Members of the OSD IP Cadre expect that the members of the larger DOD IP Cadre will provide many of the program-support functions identified in the IP instruction and that these personnel will contribute in that capacity in addition to their current responsibilities. The IP Cadre director said that this approach maximizes DOD resources, allowing the five-person team to leverage its expertise across the department—primarily by conducting strategic activities such as interpreting laws, developing DOD-wide guidance and tools, and coordinating with DAU—while relying on military department staffs to support their own acquisition programs, as they have in the past. The members of the OSD IP Cadre plan to support programs when requested to do so. As of July 2021, the director of the IP Cadre told us the OSD IP Cadre had



provided support to four acquisition programs and eight other DOD offices, but indicated that members of the larger DOD IP Cadre will be principally responsible for supporting programs.

OSD IP Cadre officials told us more work is needed to refine how members of the OSD IP Cadre and the larger DOD IP Cadre will work together. For example, these officials told us that detailed staffing and resourcing requirements for the OSD IP Cadre and the military departments have not yet been identified.

### **Areas of Expertise**

DOD officials have efforts underway to increase expertise in two of the seven areas required by the IP instruction: IP valuation and financial analysis. Members of the OSD IP Cadre told us the military departments, including the offices proposed to be part of the larger DOD IP Cadre, currently lack sufficient expertise in those areas. In its April 2020 report to Congress, the DOD described its plan to leverage an ongoing 3-year pilot program that is assessing, in part, mechanisms for determining the value of IP.<sup>10</sup> The pilot program will study valuation strategies used by one major Army weapon system and three smaller Navy programs to identify practices that can be shared across the DOD and incorporated into department-wide guidance. The pilot program will also involve the collection and analysis of data across the DOD and outreach to industry, academia, and other nongovernmental entities. Further, OSD IP Cadre officials told us that they plan to work with the Defense Pricing and Contracting directorate on financial analysis matters, although they recognize that those experts generally do not provide the program-specific financial analysis or IP support assigned to the IP Cadre in the DOD instruction. OSD IP Cadre officials told us more work is needed to determine the level of workforce resources needed to meet those responsibilities.

### **Future Funding and Staffing for the IP Cadre**

In the Fiscal Year 2018 NDAA, Congress authorized the DOD to use the Defense Acquisition Workforce Development Account (DAWDA) to staff the IP Cadre for up to 3 years. In Fiscal Years 2020 and 2021, DOD officials told us that the department used \$4.7 million in DAWDA funding on IP Cadre staffing and activities. According to IP Cadre officials, the DOD planned to use available DAWDA funding to pay the salaries for four of the five OSD IP Cadre billets through July 2023. However, OSD IP Cadre officials told us these four billets were created as temporary billets, and that DOD leadership has not yet converted them to permanent billets. The director of the IP Cadre told us that securing permanent billets beyond July 2023 is the top risk to the IP Cadre's current framework. OSD IP Cadre members told us the temporary nature of their positions was a disincentive when they were assessing the employment opportunity, and they suggested that it could present an obstacle in future attempts to staff the OSD IP Cadre.

While the DOD has developed a conceptual framework intended to guide its operations, we found that the department has not yet detailed how the IP Cadre will meet its broad responsibilities or determined whether it has the capacity to do so. IP Cadre officials told us they plan to assess further the framework and the associated implementation plans and resource requirements. Office of Management and Budget (OMB) Circular A-11 states that performance planning, human capital planning, and budget processes should jointly support an agency's implementation of goals and objectives by establishing refined strategies and resource allocations, among other things (Office of Management and Budget, 2021). Until the DOD determines how the IP Cadre will meet its responsibilities and the resources needed to do so, the DOD will be at increased risk of not implementing a key element of its IP strategy.

---

<sup>10</sup> The NDAA for Fiscal Year 2020 authorized the DOD to conduct a 3-year pilot program assessing mechanisms for evaluating IP, including its monetary value.



## **DAU Is Working to Improve IP Training, but Its Strategic Plan Lacks Priorities, and the IP Cadre Has Not Specifically Identified Which DOD Personnel Should Take the Training**

To guide its efforts to improve its IP training, DAU developed a 5-year strategic plan that identified more than 60 activities that DAU could pursue. However, resource constraints limit DAU's ability to pursue all of them, and the plan does not prioritize these activities past 2023. Additionally, the DOD's IP instruction states that DOD personnel with a role in supporting IP acquisitions should receive IP training, but officials from the military departments told us additional clarification from the IP Cadre on which personnel specifically should receive IP training would be beneficial.

## **DAU Is Updating and Expanding IP Training, but Its Strategic Plan Does Not Prioritize Activities**

DAU developed a 5-year strategic plan for improving IP training after a comprehensive review of its IP and data rights courses and training materials and based upon recommendations from IP Cadre staff and other DOD stakeholders. To implement parts of that plan, DAU has undertaken several efforts. For example, DAU introduced a foundational IP credential in September 2020, based on seven existing IP training courses. The credential is intended to provide learners with a general understanding of a range of IP topics. DAU is currently in the process of updating those IP courses to reflect legislative and policy changes from the past 5 years. The DAU IP learning director told us that DAU tentatively plans to complete those updates by June 2022. DAU also plans to develop topical IP credentials and other IP training materials. Additionally, DAU created an IP community of practice web portal that visitors can use to identify DAU's IP-related training courses. This web portal serves as one of the OSD IP Cadre's primary conduits for disseminating IP resources (Defense Acquisition University, n.d.). For example, we found that as of August 2021, the portal contained over 40 documents, including recent IP-related policies, a collection of IP and data rights best practices, templates, and videos.

The strategic plan also includes more than 60 other activities related to IP training. Proposed activities include creating or updating specific IP training courses and collaborating with industry groups to develop IP-related learning resources. This aligns with our discussions with the IP Cadre, officials within the military departments, and representatives from industry groups, who identified a number of areas where additional training could be helpful. For example, officials from the OSD IP Cadre and military departments told us that DOD personnel responsible for activities across the acquisition life cycle would benefit from training tailored to their roles. In practice, for example, this training could enable engineers who develop technical requirements to work with logisticians who plan sustainment activities to determine what IP deliverables are necessary to maintain a system. In turn, program managers and contracting staff could use that information to assess risks and costs related to IP before awarding a contract. Industry groups also told us that DOD personnel often do not understand their roles in acquiring IP, and that more tailored training could help them better engage with industry to identify appropriate IP and strategies for obtaining it. Additionally, industry groups told us that DOD personnel could benefit from training to help them negotiate IP transactions with smaller and less experienced firms, particularly when using Other Transaction Authorities (OTAs) to enter into agreements with specially negotiated licenses for IP.<sup>11</sup> OSD IP Cadre and DAU officials told us that this additional training content could be delivered through courses on OTAs,

---

<sup>11</sup> Other Transaction Authorities allow the DOD to enter into agreements "other than" standard government contracts or other traditional mechanisms. Agreements under these authorities are generally not subject to federal laws and regulations applicable to federal contracts or financial assistance, allowing agencies to customize their other transaction agreements to help meet project requirements and mission needs (10 U.S.C. § 2371b).



specially negotiated licenses, Small Business Innovation Research and Small Business Technology Transfer programs,<sup>12</sup> and Modular Open Systems Approaches.<sup>13</sup>

However, DAU officials told us that DAU's ability to execute all the potential activities, including creating or updating courses that it identified in its strategic plan, is limited by resource constraints. DAU's strategic plan identifies seven priority issue areas, which DAU plans to address through December 2022. However, DAU has not identified which activities it will fund after that time frame (i.e., from January 2023 through December 2025, the end date for the strategic plan). The DAU learning director for IP told us DAU has not prioritized activities for Fiscal Year 2023 and beyond because the OSD IP Cadre has not yet identified which activities DAU should prioritize during that period.

The DOD's IP instruction directs DAU and the IP Cadre to collaborate on developing and improving IP training. Further, OMB Circular A-11 states that agencies should identify priorities supporting strategic objectives and that strategic plans should provide the context for budget planning (OMB, 2021). Until the OSD IP Cadre provides DAU with updated priorities, there is increased risk that DAU will not use its limited resources to develop and deliver the highest priority IP training.

### **OSD IP Cadre Has Not Yet Identified Who Specifically Should Receive IP Training Within the Military Departments**

The DOD's IP instruction states that the heads of components with acquisition authority—such as the military departments—shall ensure that personnel engaged in all stages of the acquisition life cycle have relevant knowledge of IP matters, laws, and regulations. The IP instruction also tasks the director of the IP Cadre with supporting the development of training requirements for the acquisition workforce. Officials representing the Directors of Acquisition Career Management (DACM) at the Army and Air Force told us that they need additional guidance from the IP Cadre to identify the specific individuals within key career fields who should receive IP training or pursue the IP credential. They also noted that training that targets its audience is more meaningful for the workforce. For example, according to Army and Air Force DACM officials, it would be more useful to have logisticians who contribute to life-cycle sustainment plans take the IP training, rather than requiring that all logisticians do so.

This position on targeted training is consistent with November 2020 guidance from the OUSD (A&S) and the president of DAU. That guidance sets an expectation that DAU should design training and credentials for people who need specific knowledge and skills at the time they need them (Woolsey & Shaffer, 2020). The DACM officials told us that they would be positioned to track whether the targeted personnel completed the courses, using the personnel's individualized training plans, if the OSD IP Cadre more specifically identified which DOD personnel should receive IP training or credentials. Until the director of the IP Cadre provides this guidance, however, the DOD is at increased risk that personnel that should be receiving IP training will not receive it when they would benefit from it most.

## **Conclusions**

The DOD's IP instruction highlights core principles and integrates guidance and requirements for acquiring and licensing IP. However, the instruction and other DOD-wide

---

<sup>12</sup> The Small Business Innovation Research and Small Business Technology Transfer programs encourage domestic small businesses to engage in federally sponsored research efforts with the potential for commercialization.

<sup>13</sup> DOD's modular open systems approach (MOSA) is to design systems with highly cohesive, loosely coupled, and severable modules that can be competed separately and acquired from independent vendors. This approach allows the department to acquire warfighting capabilities, including systems, subsystems, software components, and services, with more flexibility and competition. MOSA implies the use of modular open systems architecture, a structure in which system interfaces share common, widely accepted standards, with which conformance can be verified.



guidance do not address misconceptions about the DOD's ability to pursue detailed manufacturing or process data. This affects the department's ability to manage costs by competing requirements for weapons systems over time, including operation and maintenance requirements. The also has not yet established the refined strategies, staffing plans, and resource requirements needed for the IP Cadre to fully meet its broad responsibilities set forth in the department's IP instruction. The DOD also has opportunities to further improve IP training by ensuring that DAU prioritizes the development and delivery of high-priority IP training, and by identifying personnel that would benefit most from receiving IP training and credentials for their roles.

## Recommendations for Executive Action

We make four recommendations to the DOD:

1. The Under Secretary of Defense for Acquisition and Sustainment should ensure that the DOD's planned guidebook on IP clarifies how DOD personnel can pursue detailed manufacturing or process data.
2. The Secretary of Defense should determine the collaboration, staffing, and resources needed, both within the OSD and across the components, to execute the DOD's proposed federated approach for the IP Cadre.
3. The Assistant Secretary of Defense for Acquisition should ensure that the director of the IP Cadre collaborates with the president of DAU to prioritize IP-related tasks that DAU should undertake between 2023 through 2025.
4. The Assistant Secretary of Defense for Acquisition should ensure that the director of the IP Cadre develops additional guidance to help component heads and DACMs identify the DOD personnel in key career fields that would benefit most from receiving IP training and credentials.

## References

- Defense Acquisition University. (n.d.). *Acquisition community connection: Intellectual property (IP) & data rights*. Retrieved October 25, 2021, from <https://www.dau.edu/cop/IPDR/Pages/Default.aspx>
- Defense Procurement Reform Act of 1984, Pub. L. No. 98-525, 98 Stat. 2492 (1984). <https://www.govinfo.gov/content/pkg/STATUTE-98/pdf/STATUTE-98-Pg2492.pdf>
- DOD. (n.d.-a). *Rights in computer software and computer software documentation* (DFARS § 252.227.72). Retrieved April 2, 2023, from <https://www.acquisition.gov/dfars/subpart-227.72-rights-computer-software-and-computer-software-documentation>
- DOD. (n.d.-b). *Rights in technical data* (DFARS § 252.227.71). Retrieved April 2, 2023, from <https://www.acquisition.gov/dfars/subpart-227.71-technical-data-and-associated-rights>
- DOD. (2013). *Operation of the defense acquisition system* (DOD Instruction 5000.02). [https://www.acq.osd.mil/fo/docs/DSD%205000.02\\_Memo+Doc.pdf](https://www.acq.osd.mil/fo/docs/DSD%205000.02_Memo+Doc.pdf)
- DOD. (2018). *The defense acquisition system* (DOD Directive 5000.01). <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/500001p.pdf>
- DOD. (2019). *Intellectual property acquisition and licensing* (DOD Instruction 5010.44). <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/501044p.PDF>
- Exec. Order No. 12591, 3 C.F.R. (1987). <https://www.archives.gov/federal-register/codification/executive-order/12591.html>
- GAO. (1991). *Defense procurement: Acquiring technical data for spare parts procurement*. <https://www.gao.gov/products/nsiad-91-313>
- GAO. (2002). *Intellectual property: Industry and agency concerns over intellectual property rights* (testimony). <https://www.gao.gov/products/gao-02-723t>
- GAO. (2006). *Weapons acquisition: DOD should strengthen policies for assessing technical data needs to support weapon systems*. <https://www.gao.gov/products/gao-06-839>
- GAO. (2010). *Intellectual property: Agencies progress in implementing recent legislation, but enhancements could improve future plans*. <https://www.gao.gov/products/gao-11-39>





- GAO. (2011, May 11). *Defense acquisition: DOD should clarify requirements for assessing and documenting technical-data needs*. <https://www.gao.gov/products/gao-11-469>
- GAO. (2014). *F-35 sustainment: Need for affordable strategy, greater attention to risks, and improved cost estimates*. <https://www.gao.gov/products/gao-14-778>
- GAO. (2018). *Weapon system sustainment: Selected Air Force and Navy aircraft generally have not met availability goals, and DOD and Navy guidance need to be clarified*. <https://www.gao.gov/products/gao-18-678>
- GAO. (2019). *Defense acquisitions: DOD needs to improve how it communicates the status of regulation changes*. <https://www.gao.gov/products/gao-19-489>
- GAO. (2021a). *F-35 sustainment: DOD needs to cut billions in estimated costs to achieve affordability*. <https://www.gao.gov/products/gao-21-439>
- GAO. (2021b). *F-35 sustainment: Enhanced attention to and oversight of F-35 affordability are needed* (testimony). <https://www.gao.gov/products/gao-21-505t>
- GAO. (2021c). *Defense acquisitions: DOD should take additional actions to improve how it approaches intellectual property*. <https://www.gao.gov/products/gao-22-104752>
- Institute for Defense Analyses. (2017). *Department of Defense access to intellectual property for weapon systems sustainment*. <https://www.ida.org/-/media/feature/publications/d/de/departement-of-defense-access-to-intellectual-property-for-weapon-systems-sustainment/p-8266.ashx>
- National Defense Authorization Act for Fiscal Year 2016, Pub. L. No. 114-92, 129 Stat. 726 (2015). <https://www.congress.gov/114/plaws/publ92/PLAW-114publ92.pdf>
- National Defense Industry Association. (2018). *2018 report government-industry advisory panel on technical data rights*. <https://www.ndia.org/-/media/Sites/NDIA/Policy/Documents/Final%20Section%20813%20Report>
- Office of Management and Budget. (2021). *Preparation, submission, and execution of the budget* (Circular No. A-11). <https://www.whitehouse.gov/wp-content/uploads/2018/06/a11.pdf>
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020). *Report to Congress on intellectual property policy and the cadre of intellectual property experts, section 838 of the National Defense Authorization Act for Fiscal Year 2020*. [https://www.dau.edu/cop/IPDR/\\_layouts/15/WopiFrame.aspx?sourcedoc=/cop/IPDR/DAU%20Sponsored%20Documents/Lord\\_IP-Policy-IP-Cadre-Sec-838-FY20-NDAA\\_Report-to%20Congress%20May%202020.pdf&action=default](https://www.dau.edu/cop/IPDR/_layouts/15/WopiFrame.aspx?sourcedoc=/cop/IPDR/DAU%20Sponsored%20Documents/Lord_IP-Policy-IP-Cadre-Sec-838-FY20-NDAA_Report-to%20Congress%20May%202020.pdf&action=default)
- Office of the Under Secretary of Defense for Acquisition and Sustainment. (2021). *Report to Congress on pilot program on intellectual property evaluation for acquisition programs, section 801 of the National Defense Authorization Act for Fiscal Year 2020*. <https://www.acq.osd.mil/asda/ae/docs/IP%20Evaluation%20Pilot.pdf>
- Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics. (2001). *Intellectual property: Navigating through commercial waters. Issues and solutions when negotiating intellectual property with commercial companies*. <https://apps.dtic.mil/sti/pdfs/ADA400207.pdf>
- Open Systems Architecture–Data Rights Team. (2014). *Intellectual property strategy guidance*. DoD.
- Patent and Trademark Law Amendments Act of 1980 (Bayh–Dole Act), Pub. L. No. 96-517, 94 Stat. 301535 (1980). <https://www.govinfo.gov/content/pkg/STATUTE-94/pdf/STATUTE-94-Pg3015.pdf>
- Reagan, R. (1983). *Government patent policy* [President’s memorandum to the heads of the executive departments and agencies].
- U.S. Army Product Data & Engineering Working Group. (2015). *Army data & data rights (D&DR) guide: A reference for planning and performing data acquisition and data management activities throughout the DOD life cycle*. U.S. Army. [https://www.acq.osd.mil/asda/dpc/cp/policy/docs/pa/Army\\_Data\\_and\\_Data\\_Rights\\_Guide\\_1st\\_Edition\\_4\\_Aug\\_2015.pdf](https://www.acq.osd.mil/asda/dpc/cp/policy/docs/pa/Army_Data_and_Data_Rights_Guide_1st_Edition_4_Aug_2015.pdf)
- Woolsey, J. & Shaffer, A.R. (2020). *Defense Acquisition University reform: The intersection with Back-to-Basics*. Defense Acquisition University and Deputy Under Secretary of Defense for Acquisition & Sustainment. <https://asc.army.mil/web/wp-content/uploads/2020/11/DAU-Reform-The-Intersection-with-BtB-16NOV2020.pdf>



# Social Engineering Impacts on Government Acquisition

**Kathleen Hyatt**—a Lead Systems Engineer at The MITRE Corporation, supports Intelligence Community, Defense, and Federal sponsors throughout all phases of the acquisition life cycle including systems development. She has more than 12 years of experience in acquisitions and systems engineering. She holds an MS in Systems Engineering, George Washington University; an MS in Accounting and Finance Management, University of Maryland; a BA in English, University of Maryland; and a certificate in Procurement and Contracts Management, University of Virginia. [kbell@mitre.org]

**Zack Levenson**—Contract Analyst at The MITRE Corporation, supports Intelligence Community, Defense, and Federal sponsors throughout their acquisition life cycles. He has four years of experience in acquisition and contracting and is a former Subcontract Administrator. He holds an BA in Political Science from West Virginia University and a Certificate in Contract Management from Villanova University. [zlevenson@mitre.org]

## Abstract

Social engineering is the activity of attempting to manipulate users or employees to reveal sensitive data, obtain unauthorized access, or unknowingly perform fraudulent activity, and it is increasingly becoming a problem for the U.S. government Contracting and Acquisition community. Even though there are improvements in technology that make both online and offline environments safer, the human factor is still a significant vulnerability. This is especially prevalent within the Government Acquisition community, where much of the labor is not automated, and therefore relies on human actors.

Sensitive information that is collected can be used as intelligence by nation state adversaries; it can enable fraudulent financial activity; and it can be deployed to interfere, influence, and disrupt sovereign national activities. Privileged access can also be leveraged—even without theft of information—as an avenue through which actors can travel to attack computer systems in kinetic ways to disrupt operations, damage equipment, or even harm personnel. The U.S. government is not immune to this issue, losing hundreds of millions of dollars over the last decade due to social engineering attacks.

This paper addresses the impacts that social engineering can specifically have on U.S. government Contracting and Acquisition organizations, such as threats to the supply chain and deepfakes. Recommendations will also be made for how agencies can both recognize and prevent social engineering attacks from occurring, thus preventing damage, disruption, compromise, and the loss of resources.

## Executive Summary

Information is valuable. Knowledge is power. Because of the utility of information, those with ill intent work with steadfast discipline to extract data from those with privileged access through a variety of means. Sensitive information that is collected can be used as intelligence by nation state adversaries, it can enable fraudulent financial activity, and it can be deployed to interfere, influence, and disrupt sovereign national activities. Privileged access can also be leveraged—even without theft of information—as an avenue through which actors can travel to attack computer systems in kinetic ways (e.g., overspinning a centrifuge in a nuclear facility causing them to self-destruct) to disrupt U.S. government (USG) operations, damage equipment, or even harm personnel.

Therefore, information security is vital to prevent an adversarial advantage on multiple fronts and to ensure the security of U.S. and allied personnel and assets. Government employees can be unknowingly manipulated to provide valuable information and access to harmful actors which can cause varying degrees of damage in multiple areas. This paper highlights ways in which social engineering attacks can be used to manipulate the government acquisition ecosystem to detrimental effect.

Social engineering activities are prevalent within the government acquisition community because so



much of the labor is not automated, and therefore relies on human actors. For instance, as a part of most government acquisition operations, there is an individual Contracting Officer (CO), an industry official, and additional unsuspecting support staff who can potentially be manipulated to facilitate unauthorized access and/or fraudulent activity. This can happen to anyone and can vary in severity. The purpose of this paper is to educate practitioners and provide threat mitigation recommendations to the government acquisition community.

Note that many elements of social engineering as a discipline of adversary activity overlap with traditional Human Intelligence (HUMINT) and Cyber-HUMINT tactics, techniques, and procedures (TTPs); however, for the purposes of this paper and audience, prospective distinctions and similarities between these various categories of operations will not be called out. Additionally, for the sake of clarity and consistency of lexical terms used, this paper will focus on the concept of social engineering in the context of information security.

There are central themes to many social engineering attacks, and many attacks are conducted using a hybrid approach combining one or more of the types of attacks outlined in the Operational Social Engineering Attacks section.

Knowing that anyone can become a victim, this paper recommends both proactive offensive approaches and defensive approaches to counter-act the attempt at manipulation in the hopes of minimizing vulnerabilities in government acquisition and preventing the loss of information and millions of dollars.

## Definitions

For the purposes of understanding this document, the following terms are defined to clarify intent and scope.

**Social Engineering:** The act of deceiving an individual into revealing sensitive information, obtaining unauthorized access, or committing fraud by associating with the individual to gain confidence and trust (NIST, 2020).

**Government Acquisition:** The act of acquiring by contract with appropriated funds of supplies or services (including construction) by and for the use of the federal government through purchase or lease, whether the supplies or services are already in existence or must be created, developed, demonstrated, and evaluated (FAR, 2023).

## Introduction

Social engineering is increasingly becoming a problem for the USG. Even as there are advances in technology that create more secure online and offline operating environments, a significant vulnerability continues to be the human factor. Social engineering is the activity of attempting to manipulate users or employees to either reveal sensitive data, obtain unauthorized access, or unknowingly perform fraudulent activity. The USG is not immune to this issue and has lost hundreds of millions of dollars over the last decade due to social engineering attacks as detailed below.

This paper addresses the impacts that social engineering can specifically have on USG contracting and acquisition such as threats to the supply chain and deepfakes. Recommendations will also be made for how agencies can both recognize and prevent social engineering attacks from occurring, thus preventing damage, disruption, compromise, and the loss of resources.

## Background

Adversary-directed threats to U.S. systems, information, and personnel—including HUMINT operations, cyber-attacks, signals intelligence collection, and cyber-enabled espionage—have long plagued the Western national security enterprise. However, as the overarching rise of technology in society widens the attack surface on which adversaries can conduct operations, social engineering as a threat has also evolved in conjunction with these larger changes. In today's operational context, social engineering can manipulate a



plethora of individuals and technical access points to facilitate the fraudulent provision of information, the success of a network intrusion, and/or the execution of an influence, interference, or kinetic operation. Where cyber-attacks center on infrastructures and networks, social engineering attacks focus on the actors who control and access those networks.

Humans remain an unpredictable variable in maintaining cybersecurity, and therefore, are a common target for attackers. Technical attacks are typically easier for information security and counterintelligence (CI) entities to plan for given that these processes are often repeatable and predictable. However, it is much more difficult for human activities to be seen as reliably consistent in terms of TTPs because where computers and infrastructures might be the same, no two humans behave, react, or think in precisely similar ways. Where one person might be able to anticipate and recognize a social engineering attack, a different person might perceive an attacker's intrusion attempt to be an innocuous or friendly act and thereby unknowingly allow the attacker to access the information they seek.

Social Engineering attacks are typically more psychological than they are technological. Instead of using sophisticated hacking techniques or in-depth knowledge of computers, they rely on tricking people into giving away information. Cybercriminals that engage in social engineering are digital con artists, gaining vulnerable people's trust to steal money or data easily. (Partida, 2020)

Another reason that social engineering TTPs are growing (O'Reilly, 2021) in popularity with attackers is that they are generally perceived by users to be low-cost, high reward tools within the larger kit of computer exploitation options. For example, it might unnecessarily burden a given Advanced Persistent Threat (APT) [group](#) to design a complex, highly surreptitious, and deeply intrusive malware delivery package when a simplified socially engineered mass malware spam campaign can achieve the same objective of initial network access. Additionally, using social engineering techniques to gather information about a user could make it much easier and faster for that attacker to ascertain a user's password to access the system. In these cases, it often doesn't matter how sophisticated the security guarding the network is if the attacker is able to target the user and manipulate them into giving away credentials without realizing what they're doing.

While social engineering operations can result in gathered reconnaissance information that can then feed and shape the design of a network intrusion set, there is a prospective cyclical nature to many of these operations where the data gathered from a network intrusion can then feed additional tailored social engineering manipulations should the adversary wish to gain access to other hardened networks. That said, the sheer depth and breadth of publicly available online information sometimes eliminates the need for any intrusion set to precede a social engineering operation; this is because attackers can take commonly accessed information and twist it in a way that is advantageous for them. Simply put, social engineering attacks can take many forms depending on the context and needs of the attackers. This threat is especially present in the government acquisition arena. For example,

the fact that GovCon Co. is a prime contractor on a certain Government contract is generally available to the public; a press release, website news item, social media profile, or other public information may show that Subcontractor Co. is a subcontractor to GovCon Co. on that prime contract; and a simple LinkedIn or Facebook search may reveal that John Smith is a contracts manager or billing representative for Subcontractor Co. A fraudster need only create a domain and email address such as "jsmith@subocntractorco.com" to facilitate his or her scheme. Many individuals, when processing invoices, may not notice the misspelling in the domain name. They simply changed the bank account information and issued payment. The result? Hundreds of thousands of dollars in losses, and limited recourse to recover what was lost. (Mazza & Feinberg, 2020)

In fact, there are many examples of acquisition social engineering attacks that do not involve cyber intrusions at all. An example occurred in North Carolina in 2019 when the state government lost over \$1.7



million to a social engineering scheme. The government office was approached by what appeared to be a legitimate contracting business hired for the construction of a school.

They possessed allegedly valid licenses and all the required paperwork needed to establish an account and have funds transferred. The fraudulent actors were able to create such accurately forged papers because of publicly available information gathered on similar legitimate businesses. Possessing this convincing cover, the threat actors were then able to gather privileged information that enabled the theft of funds. The county in which this social engineering attack occurred was very clear to state that this was not a cyberattack, and the loss of funding was the direct result of an unintentional information leak. “The county was not hacked. It was not a cybersecurity [incident]. This is a case of a spoofed identity in which somebody posed as a vendor, provided seemingly valid documentation, and signed approvals” (Ropek, 2019).

## Life Cycle of a Social Engineering Attack

Social engineering has continued to grow as a persistent threat to U.S. businesses and government entities over the past decade. As the attacks have grown in frequency (O’Reilly, 2021), so has the understanding of how these attacks typically arise and evolve over time. As seen in Figure 1, researchers now depict and organize social engineering attack techniques in to four phases of adversary execution:

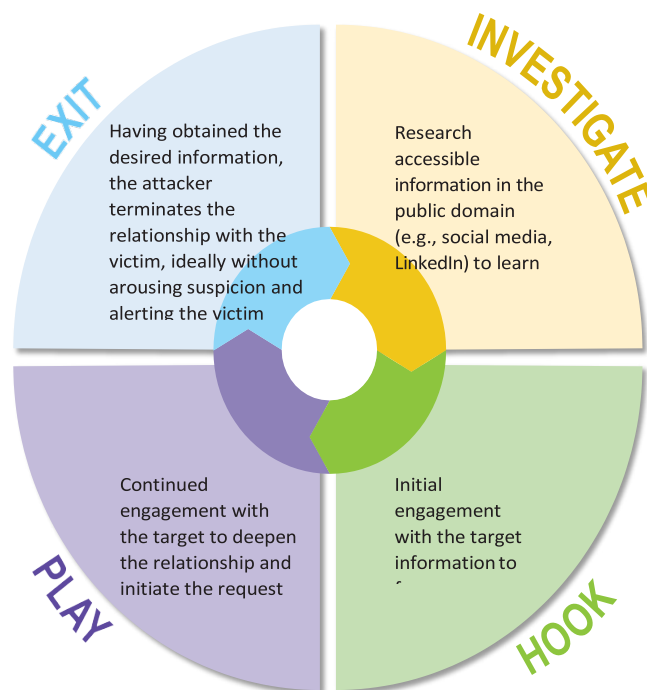


Figure 1. Life Cycle of a Social Engineering Attack<sup>1</sup>

1. **Investigation**—The initial stage in which the attacker already has an intended goal in mind and selects their victim(s). Once they know their target, they begin gathering background information (oftentimes information that the target has already released willingly through open channels) and decides on their preferred attack method (discussed further in the Operational Social Engineering Attacks section).
2. **Hook**—The stage where the initial interaction with the target occurs in the effort to gather the needed information. This includes preparing a cover story if needed and knowing how to

<sup>1</sup>Figure 1 represents a synthesized description of adversary behavior based on a collection of several previously published graphics detailing the attack cycle.

maintain control of the interaction to ensure the needed information is successfully obtained.

3. **Play**—The stage in which execution and continuation of the socially engineered manipulation occurs; this is where humans are influenced, coaxed, pressured, or unwittingly fooled into provide sensitive information or access. The duration of this stage can be long or short, depending on the type of social engineering attack used, but implies that the attacker will have the patience to play the long game and will engage with the target multiple times if needed. In some cases, the attacker might even use multiple techniques to gather as much valuable information from the target as possible.
4. **Exit**—The final stage in which the attacker generally ends the interaction with the victim in a natural way so as not to arouse any suspicion. This social engineering framework allows for the threat actor to cycle back into stage one for further investigation and manipulation should the adversary require additional information not gathered during the previous engagement(s).

Using the steps in the social engineering attack lifecycle, the attacker is able to retrieve all of the information they need without the target being aware that they have divulged valuable information. The target's lack of awareness about their own inadvertent support is what makes these targeting techniques so dangerous.

## Social Engineering Attacks

### Cognitive Exploitation

Procurement and acquisition play an essential role in a majority of government projects, and it should not be overlooked that social engineering activities can negatively affect this foundational element of the defense enterprise. Social engineering attacks are uniquely targeted at the human decision-making process. As Sherman and Arampatzis (2018) discuss in their article “Social Engineering as a Threat to Society,” the biggest challenge that makes humans (and therefore government employees) susceptible to social engineering attacks are cognitive biases. Cognitive biases refer to the ways that humans process information and how decisions are affected. Not everyone interprets information in the same way, and therefore it can be difficult to predict how humans will react in a given situation. Social engineering attackers capture this reality and use it to their advantage when collecting information from targets.

An example of this this cognitive bias is the tendency for the human brain to group similar memories or repetitive actions together, to the point where the brain almost goes into autopilot. If you read the previous sentence again, you may notice that an additional “this” has intentionally been included as a display of this bias in action. For many, the brain has self-corrected the error without registering that an additional word was present. Biases like this could impact contract and acquisition activities because it is a field where similar processes are repeated over and over, and it becomes possible for smaller and inaccurate details to go unnoticed.

As previously mentioned above, attackers can emulate domain names, email addresses, and other information easily based on information that is gathered electronically. An example could be processing invoices in a system, which Contracting Officer Representatives (CORs) must do quite often. The repeatable process begins to put the COR on autopilot and the COR could easily overlook pertinent information and submit payment to an attacker through human error. The social engineers who are looking to conduct attacks are aware of this and are prepared to take advantage as best they can. As discussed later in the Recommendations section, this is one area where advanced technological aids (particularly those relating to artificial intelligence-enabled “suspicious activity” detection) can be of particular use in terms of threat mitigation.

### Principles of Influence

Social engineering attacks tend to focus on the exploitable elements of human cognition and behavior in an attempt to manipulate workers. Robert Cialdini identified several of these characteristics in his work, *Influence: The Psychology of Persuasion*, which he refers to as the six principles of influence. These include:



1. **Reciprocity**—This refers to the tendency of people to return a favor when something is done for them. An example of this can be seen in marketing when businesses offer free samples or trial runs before requesting commitment to buy. An acquisition-salient example of this could manifest as a CO awarding a contract to an industry partner in return for monetary, professional, and/or personal benefits.
2. **Commitment and Consistency**—Commitment can be a powerful motivator and refers to the fact that once people say they are going to do something, they feel personally obligated to ensure it is completed. Sometimes they will continue with an activity even if the original intent has changed, or if its completion will no longer have an impact. Given that acquisition professionals are, as described later in this paper, often hyper-cognizant of their professional reputation, an example of this principle in action might include a CO prioritizing essential contract actions over good security practices.
3. **Social Proof**—This principle states that humans are more likely to conduct activities that they see others doing. This includes people who avoid being the first person to do something in case it results in failure or issues. An example of this might include the disincentive that a CO has to be the first (and possibly only) individual to identify and call out contract fraud.
4. **Authority**—Most people have a natural respect for authority and those in positions of power, and often reflexively comply instead of questioning the orders given to them by those types of figures. An example of this might include a CO receiving a call from a higher echelon of authority—a Department of Justice official or that CO's supervisor—whereby orders are given to provide sensitive source selection information.
5. **Liking**—This refers to the tendency for people to be more likely to listen to commands and follow directions that come from people that they like. It is easier for people to want to please those they have a higher opinion of; a desire to do one's best to ensure that the other person likes them in return is a related effect. An example of this might include a bad actor impersonating an individual known to be close friends with an influential contract manager in order to sway the requirements and outcomes of given contract awards.
6. **Scarcity**—Lastly, if people perceive that something is scarce, they believe it to be more valuable, and naturally will make more of an effort to obtain it even if that is not true. Scarcity might lead to people buying more items than they actually need or spending more than is necessary to obtain the items. An example of this might include a commercial organization being manipulated to believe that they are likely to win a valuable and highly competitive contract if they provide extensive PII.

All six of the principles of influence create opportunities for staff to be exploited by social engineering attackers. According to the Association of Government Accountants (AGA), there are many ways in which those principles can be exploited, and that staff can be targeted (AGA Tools and Resources, n.d.).

### **Operational Social Engineering Attacks**

Table 1 below shows many types of social engineering attacks and examples of how they can manifest in the operational environment.



**Table 1. Social Engineering Techniques and Examples**

Social Engineering Technique	Definition	Example
Phishing	As one of the most popular social engineering attack types, phishing scams are email and text message campaigns aimed at creating a sense of urgency, curiosity, or fear in victims. It then prods them into revealing sensitive information, clicking on links to malicious websites, or opening attachments that contain malware.	Ubiquiti Networks, a manufacturer of technology for networking, lost almost \$40 million dollars in 2015 after a phishing attack. It is believed that an employee email account was compromised in Hong Kong. Then, hackers used the technique of employee impersonation to request fraudulent payments, which were made by the accounting department (Gatefy, 2021). From an acquisition perspective, an example may look like email response to an Request for Information (RFI) that contains a corrupted word document therefore installing malware on to the CO's computer.
Elicitation	A subtle approach used to gather information from users through basic social interactions and research into a user's online and social media presence.	Hackers stole millions of Social Security numbers and thousands of credit and debit card numbers from the South Carolina Department of Revenue in 2012. Employees fell into scams by sharing their usernames and passwords with criminals. After that, with credentials in hands, the hackers gained access to the state agency's network (Gatefy, 2021). From an acquisition perspective, an example may look like a CO who is talking to co-workers in a public place, and inadvertently discloses sensitive contract information to a person listening in on their conversation.
Pharming	Redirecting web traffic from legitimate sites to malicious clones/fraudulent IP addresses. This ploy can be leveraged to create fake sites, upload content, monitor traffic, or hack official corporate systems (Barnett, 2022). For example, an attacker can use malicious code to monitor user web activity to trigger a redirect to a spoofed banking site. When a user enters their bank domain into the browser address bar, the pharming code hijacks the user's activity and redirects the browser to an attacker-controlled website with the same look and feel as the official bank account. Users rarely look at the domain in the browser's address bar, so it's an effective attack to steal user financial data, including their credentials (Proofpoint, n.d.).	"A number of news stories have emerged in recent years of corporations being attacked in this way, including instances of official corporate subdomains being hijacked to redirect to content including malware, pornography, and gambling-related material. Subdomains of the Xerox website, for example, were used in 2020 to drive traffic to sites selling fake goods, taking advantage of the trusted reputation of the official corporate domain to boost the search-engine ranking of the malicious content. In another case in 2019, GoDaddy® shut down 15,000 abused subdomains that drove a massive spam campaign geared towards the sale of counterfeits" (Barnett, 2022) From an acquisition perspective, an example may look like a website masking the Wide Area Workflow, the DoD's invoicing, payment, reporting, and contract information portal, would allow an unsuspecting contractor or government official to give proprietary, sensitive, and financial information to a bad actor.
Framing	The tactic used to frame a situation by asking leading questions or phrasing statements in such a way that they focus on the target's unique biological and cultural influences to create a level of comfort and familiarity. That familiarity is then leveraged to manipulate targets into sharing sensitive information or otherwise enabling access to systems.	If an attacker wants to obtain information on a certain type of security device they might ask, "Where can I get some info on security devices?" or, "What resources are there available to help me find information on security devices that can handle XYZ protocols?" If trying to obtain personal information from a secretary who has a family photo out an attacker can ask, "What is your child's name?" That direct question may close the door quickly. The secretary may answer it, but it may not allow for additional inquiry. Whereas "Is this your oldest child?" may elicit not only a positive response, but a plethora of information about other children she may have (Influencing Others, n.d.). From an acquisition perspective, an example may look like a bad actor or curious industry contractor could solicit information from a CO such as source selection information, future acquisitions, or vendor performance to influence stock trading or investment opportunities to enrich themselves.
Pretexting	A premeditated attack in which a person constructs an elaborate story to place a user in a tense and urgent situation in which they might disclose information they normally would not disclose. Pretexters can impersonate co-workers, police officers, bankers, tax authorities, clergy, insurance investigators, etc. Impersonating a person of authority or someone with a right-to-know lays the groundwork for applying pressure onto targets which thereby provide needed information. The pretexter must typically prepare answers to questions that might be asked by the victim. Sometimes, an authoritative voice, an earnest tone, and an ability to think on one's feet are all that is needed to create a pretextual scenario.	The most common example of a pretexting attack is when someone calls an employee and pretends to be an individual in a position of power, such as the Chief Executive Officer (CEO) or a staff member on the information technology (IT) team. The attacker convinces the victim that the scenario is true and collects the information that is sought (Nadeem, 2022). From an acquisition perspective, an example may look like a CO receiving a call from a person posing as an FBI agent requesting small bits of information on a specific program's vendors to aid in an investigation, which the CO complies with. If the program is sensitive, this information on which vendors are working the program can be used by adversaries to target and attempt to exploit these unsuspecting businesses.





Cold Calling/ Vishing	This is the simple act of gathering information by making unsolicited phone calls, sending voice messages, and leaving voicemails as a means to make contact; these acts are conducted in ways that initially seem to amount to insignificant interactions, but small pieces of information about a person gathered separately over time are often combined to form a valuable profile to be used by attackers.	Social engineers can mimic recognizable phone numbers and caller ID names to gain trust. Voicemail recordings, automatic “out of office” replies, and other volunteered information can also be leveraged to collect PII. As a hypothetical example, a social engineer could leverage an “out of office” reply to form the following elicitation email: “Hi Dan, I hope Erica is enjoying her vacation in the Bahamas. Since she won’t be back until July 31st, she directed me to you to answer my questions.” A confident opening is all a social engineer needs to appear as a credible source (Access Systems, 2019). From an acquisition perspective, an example may look like a CO who is targeted of specific “new” business pitches and their products/solutions where a CO reveals slowly what is interesting to them one product at a time, framing a picture of what the agency may be procuring in the future.
Gaslighting	This technique involves psychologically manipulating a target to the extent that they begin to question their own logic, opinions, and/ or sanity. This is an aggressive technique where attackers will do their best to lie, misdirect, and confuse people into providing information unwittingly in support of a social engineer’s operation.	One example involves asking questions with unimportant answers to create the opportunity for the attacker to get aggressive and fluster the employee to the point that they will offer any information they can to attempt to calm down the attacker and end the confrontation. Criminals and foreign actors can use gaslighting to change perceptions, behaviors, and actions. Gaslighting also stifles discussion and dissent because it attacks conviction and surety of a person’s knowledge and beliefs. Gaslighting must tear down an individual in order to manipulate and control them (McGuinness, 2020). From an acquisition perspective, an example may look like a CO receiving a call from a person posing as a vendor who is requesting confirmation of financial data. The CO may reply that the information has already been sent, but the fake vendor insists that they never received the information and threatens to call their supervisor. This immediately makes the CO question their past actions and resend the requested financial data, giving it directly to the fake vendor.
Client/Vendor Impersonation Fraud	This technique involves a social engineer posing as a client or vendor in order to gain sensitive information through a conduit of trust; phishing and other techniques can be used to collect information to build a more sophisticated cover-for-action and cover-for-status.	“An employee receives a phone call from an individual who he believes to be a genuine supplier. The fake supplier advises that his bank details have changed, and payment is to be made to a new account. Going through procedure, the employee advises that the request must be received in writing via email or on company letterhead. The employee later receives an email from what appears to be the legitimate supplier complete with the supplier’s signature at the foot of the email. The employee proceeds to change the bank details and a payment is issued. Sometime later, the genuine supplier requests payment, indicating that the original payment was never received. Further investigation will identify that the earlier request was fraudulent.” Due to a social engineering and Business Email Compromise (BEC) scam, Cabarrus County, in the United States, suffered a loss of \$1.7 million in 2018. Using malicious emails, hackers impersonated county suppliers and requested payments to a new bank account. According to the investigation, after the money was transferred, it was diverted to several accounts. In the emails, the scammers presented apparently legitimate documentation (Gatefy, 2020). From an acquisition perspective, the above example demonstrates how a bad actor can pose as a legitimate company and target a less seasoned acquisition professional.
Client/Vendor Impersonation Fraud	This technique involves a social engineer posing as a client or vendor in order to gain sensitive information through a conduit of trust; phishing and other techniques can be used to collect information to build a more sophisticated cover-for-action and cover-for-status.	“An employee receives a phone call from an individual who he believes to be a genuine supplier. The fake supplier advises that his bank details have changed, and payment is to be made to a new account. Going through procedure, the employee advises that the request must be received in writing via email or on company letterhead. The employee later receives an email from what appears to be the legitimate supplier complete with the supplier’s signature at the foot of the email. The employee proceeds to change the bank details and a payment is issued. Sometime later, the genuine supplier requests payment, indicating that the original payment was never received. Further investigation will identify that the earlier request was fraudulent.” Due to a social engineering and Business Email Compromise (BEC) scam, Cabarrus County, in the United States, suffered a loss of \$1.7 million in 2018. Using malicious emails, hackers impersonated county suppliers and requested payments to a new bank account. According to the investigation, after the money was transferred, it was diverted to several accounts. In the emails, the scammers presented apparently legitimate documentation (Gatefy, 2020). From an acquisition perspective, the above example demonstrates how a bad actor can pose as a legitimate company and target a less seasoned acquisition professional.



Fake Office Fraud	An attack in which the perpetrator will pose as a staff member from an office—usually one of authority—to threaten repercussions; this activity is often combined with a sense of urgency so as to not give the victim time to consider their actions.	<p>“A midlevel finance employee is the only person remaining in the office on a Friday evening when she receives a phone call from an individual who identifies himself as the company’s CEO. He explains that a major acquisition is about to take place, but it must close tonight, and he can’t get in touch with anyone else on the finance team to process the payments. The employee explains that she only has authority to transfer funds of up to \$50,000 and that no one else is in the office to countersign the transfer. The CEO grows increasingly irate with the employee for refusing to transfer the funds because she does not have the authority. He repeatedly tells her that he’s granting her the authority. Eventually the CEO persuades her to circumvent the established procedure by issuing multiple \$50,000 transfers totaling \$500,000.” (Arthur J. Gallagher &amp; Co., 2016)</p> <p>From an acquisition perspective, an example may look like a CO receiving a call from someone posing as the Office of the Inspector General, suddenly forcing them to reveal source selection material or proprietary information.</p>
Funds Transfer Fraud (FTF)	A type of social engineering attack in which government agencies think they are doing business with a legitimate company, when in actuality they are sending funds directly to attackers.	<p>FTF (aka BEC) has become a very popular form of social engineering attack given that if the targeted business does not have the proper protocols in place to verify the legitimacy of the vendor, they can potentially send large payments, once or even several times, resulting in significant losses. “According to the Federal Bureau of Investigation (FBI)’s 2019 Internet Crime Report, complaints revealed an uptick in BEC scams by a considerable margin. The FBI found BEC to be the most damaging type of cybercrime in 2019. BEC losses averaged \$75,000 per complaint, phishing, smishing, and vishing accounted for \$500 per complaint, and ransomware averaged \$4,400 per complaint.” (Cyber Armada Team, 2020)</p> <p>From an acquisition perspective, an example may look like a CO receives an unsolicited bid from someone posing as a vendor advertising a scarce resource. Due to the need for services during an urgent and compelling situation, the CO fails to verify the legitimacy of the vendor.</p>
Lawyer Impersonation	This technique involves a social engineer posing as an attorney or legal figure in order to gain sensitive information through a conduit of trust and often urgency; phishing and other techniques can be used to collect information to build a more sophisticated cover-for-action and cover-for-status.	<p>“An employee receives a phone call from someone posing as an attorney and claiming to be handling confidential or time-sensitive information. These scammers typically initiate contact at the end of the business day or work week to coincide with the close of business of international financial institutions.” (Arthur J. Gallagher &amp; Co., 2021)</p> <p>From an acquisition perspective, an example may look like receiving a fake data request from the agency legal office or the Government Accountability Office, and fulfilling the data call, which gives away trade secrets, proprietary information, or source selection information.</p>
Deepfake Deceptions	The use of “synthetic media” enabled by artificial intelligence to simulate a specific person’s appearance and/or voice via video or audio recording; this can be used to deceive victims into divulging information or performing an action.	<p>In 2019, a fake recording of a CEO’s voice was used to instruct an employee to transfer money to an international account. “The recording was left as a voicemail to the subordinate, who obeyed the fraudulent instructions and sent \$243,000 to the attackers” (Slater, 2021).</p> <p>From an acquisition perspective, an example may look like a CO receives a phone call from a bad actor using a synthetic voice manipulator to pose as the director of their department, requesting the immediate purchase of a specific item that can only be found on one website. Due to the low value of the product, which is below the micro-purchase threshold, no approvals and little documentation is needed, handing the money directly to the criminal.</p>
Browser Notification Hijack	A technique whereby social engineers insert notification script, malware, and/or influential messaging into web browser or website notifications; this requires that the target be convinced or manipulated into “allowing” notifications (e.g., engineers can disguise subscription consent as another action, they can switch the “accept” and “decline” buttons on subscription alerts, etc.).	<p>According to a Review Geek publication in March 2022, an affiliate of the website outlined what was perceived to be a pop-up computer virus pretending to be anti-virus software; however, these messages were actually malicious browser notifications from a website and as such, could not be removed with legitimate anti-virus software (Heinzman, 2022).</p> <p>From an acquisition perspective, an example may look like a CO’s weekly check of the file transfer where status reports are uploaded by contractors suddenly offers to push notification when a new file is submitted. When the CO clicks yes to save time, malicious code is downloaded onto their computer.</p>
<p>Additional social engineering techniques not mentioned in detail include: Spear Phishing, Vishing, Whaling, Smishing, Baiting, Piggybacking/Tailgating, Quid Pro Quo (i.e., tech support scams), Honeytraps (deceptive and/or false romance scams), Scareware, and Watering Hole attacks.</p> <p>The FBI’s 2021 Internet Crimes Report showed that “phishing (scams via email to induce recipients to share sensitive information), vishing (voicemail phishing), smishing (SMS text phishing) and pharming (using malicious code on the victim’s device to redirect to an attacker-controlled website) were the top forms of cybercrime in 2021” (Watson, 2022).</p>		



## Emerging Technology Integration and Autonomous Execution

With the advancement of artificial intelligence (AI), machine learning (ML), and Internet of Things technology, many emerging and aforementioned social engineering techniques have the power to be partially or fully automated from end-to-end giving rise to a compounding threat of “social engineering at scale” with significantly fewer human resources burdened to execute operations. Examples include attackers training AI and its algorithms to target specific types of files so that they can home in on the metadata of these files. Reporting on how ML can be leveraged to bolster the toolkit of cyber-criminals’ notes:

As in the case of phishing or infection preparation, hackers may use the [machine learning] classifying algorithms to characterize a potential victim as belonging to a relevant group. This means that after having collected thousands of emails, a hacker sends malware only to those who would click on the link. Thus, the attacker reduces the chances of early detection of the planned attack. Numerous factors may assist here. For example, the hacker can separate the users of social networking sites who write about IT from those focused on “food-and-cats” topics. The latter group might be unaware of threats. Various clustering and classification methods from K-means and random forests to neural networks can be used in this case on top of the [natural language processing] (NLP) analysis, which should be applied to victim’s posts on social networks. (Polyakov, 2019)

AI-enabled chatbots—often leveraged by IT help desks—can also be turned around by social engineers to seek out and extract sensitive PII from customers in need of technical assistance; in this way, an illegitimate chatbot posing as one tied to a legitimate business could be deployed at a target to extract data, but it is also possible that social engineers could pose as the very target they seek to extract data about when speaking to legitimate chatbots and use collected PII to access account information through the authentic automated help desk. In so many ways, bad actors’ opportunity for operational growth in this area is dangerously promising.

## Implications of Human Error in the Context of Emerging Technology

Social engineering techniques center around the unpredictable (e.g., difficult for bureaucracies to systematically mitigate) and malleable (e.g., exploitable) actions of humans, and one unavoidable fact is that humans tend to make mistakes. It does not matter if the mistakes are large or small, it only matters that social engineering attackers know that if they can create the right circumstances, they can increase likelihoods that humans will make the kinds of mistakes that will benefit their agenda. This likelihood expands sufficiently when social engineers attack in numbers (all it takes is one human’s error to open the network’s flood gates) and those numbers expand dramatically when enabled by advanced technology that pushes the social engineering operational tempo to an exponential scale. Spoken more bluntly, if 50,000 targets are attacked within one government agency every day (a scale potentially to be enabled by AI/ML tools) with social engineering techniques that are programmed to change and enhance themselves as neural networks learn more about the targets’ interests, habits, and behaviors (purposed to exploit the varying possible weak) it is not just high, it is oftentimes all that is needed for operational success and security disaster. For example, an employee can be trained to recognize a fraudulent email and phishing attempt that instructs them to “click this link for more information.” However, under the right situation where the employee is pressed for time due to multiple deadlines and when emails stress the urgency with wording such as “this must be done immediately,” the employee is more likely to make the mistake of clicking the link and opening a connection for the attacker. This is especially prevalent in acquisition as contracting professionals always have more work than time and are often working under extremely tight deadlines and heavy amounts of stress. Prognostic horizon analyses—and even diagnostic assessments of the more current threat—would not be sensational if they articulated that the threat was compounded by the prospect that new technologies are significantly increasing the quantity of human targets that can be hit (and the rate at which they can be attacked) therefore raising the threat level in unprecedented ways.



## Impacts on Procurement from Social Engineering Attacks

Another unique impact to businesses and government agencies that affects procurement activities is the loss of reputation. In procurement, reputation and past performance play a critical role in how many other businesses will want to engage in partnerships and relationships with a given entity. If a business entity is consistently unable to defend against social engineering attacks, it could cause them to lose future contract awards. "Perhaps the most damaging side effect of any data breach is a tarnished reputation. A Ponemon Institute study found that 65% of surveyed consumers lose trust in a business after a data breach. Furthermore, 27% ended their relationship with a company, and stock prices fall an average of 5% after a breach" (Partida, 2020). Social engineering attacks are dangerous because even if the monetary damage done to the business is small, the impact to a damaged reputation and future business lost can be severe. Since government agencies frequently rely on contractors to achieve their missions, contractors who have access to secure government assets are consistently vulnerable. If a contractor is impacted by a social engineering attack, it may have an adverse effect on the future government acquisitions and procurement process as well as put the mission in jeopardy. Additionally, disruptions to existing business relationships with contractors add to the overhead acquisition cost and make for a less efficient and more costly acquisition ecosystem. For example, losing a contract relationship due to social engineering attacks necessitates remedial market research to identify and select a new contractor, as well as follow-on contractor vetting, contractor surveillance, and training of new contractor staff.

## Indirect Losses to the Government

In addition to the threat of losses from direct social engineering attacks, indirect effects can be seen in supply chain disruptions which can have sizable downstream impacts on government operations. Mainly, due to the sheer number of contracts and operations that large businesses and government agencies interact with to purchase services and supplies, there is an increased likelihood of feeling the effects of social engineering attacks either by direct intrusion or by second- and third-order proxy. Because there are so many variables, there is a greater chance that somewhere down the supply chain, there is a vulnerability that can be exploited. Once one company in the supply chain is impacted, those effects can be seen by all other companies who do business with the exposed entity.

## Recommendations

Awareness is a primary challenge in social engineering attacks. However, so is the need to defend personnel, networks, and assets with techniques that match or outgun the sophistication of emerging social engineering attacks; to do so would be to act on the advice of counterintelligence/cybersecurity professionals and leaders that have historically been tasked with defending against tier-one threats to the U.S. defense enterprise. In order for acquisition staff to make efforts to prevent these social engineering attacks, they need to first be made aware of the threat and the ways in which they might be vulnerable. As mentioned above, the biggest challenge with addressing social engineering prevention is the vast differences in staff, i.e., a static set of techniques for making staff understand and prevent these attacks will not work for everyone. Some factors that must be included when developing different training processes include the employees' skill level, time in the work force, and internet usage (Aldawood et al., 2020); managers can go further to include factors such as trending attack techniques, promotion incentives or rewards for thwarted attacks, creative engaging "war game" exercises, and/or more flexibility provided to staff for detecting threats despite project deadlines. These are all factors that can impact someone's understanding, concern, and applicability of social engineering prevention measures. These factors have been broken into two categories, defensive/vulnerabilities and offensive/proactive.

## Defensive Factors and Vulnerabilities

Defensive factors should be implemented to ensure that staff and systems at any business or government agency are well postured to recognize social engineering attacks, know how to prevent them, and know what to do if an information leak should occur. The acquisition community is inherently outward facing



because they are the bridge between industry and the government. This makes them a unique target because of their need to interact outside the cyber-security perimeter of the government, their publicly available contact information, and their access to sensitive information. The following factors should be addressed and researched to ensure the best chance of repelling and identifying a social engineering attack:

- **Security Skill Level**—The degree to which the acquisition professional is familiar with common security practices and procedures. When assessing an employee’s security skill level, the government or specific agency may tailor training processes after asking:
  - Does the employee have the ability to determine whether something doesn’t seem right? If so, do they know how to appropriately respond?
  - Does the employee have a USG security clearance? Those with a clearance are more likely to think twice about engaging in risky behavior due to the additional training related to counterintelligence, manipulation, and risks associated with doing cleared work. Those without a clearance may need more in-depth training.
- **Time in the Work Force**—An employee’s level within the company (e.g., entry level, journeyman, or senior) could also be a factor as they will have different levels of responsibility and familiarity with established policies and procedures. For example, some employees who have been through years and years of training may be less likely to pay attention to new security measures because of the belief that they don’t need to learn anything new. Alternatively, some experienced employees may, because of that practical wisdom, be postured to recognize common schemes deployed at acquisition professionals. When developing social engineering training, the government should consider:
  - Does the employee’s knowledge of the work/ office environment unintentionally cause them to be a target? All employees, no matter age or time in workforce should be required to attend annual training for cyber security and social engineering threats which includes an assessment.
- **Internet Usage**—Internet usage is a part of every acquisition professional’s day-to-day activity, but some employees will be more familiar with it than others. That familiarity might be beneficial, but it also might become detrimental depending on how knowledge is applied; for example, experienced internet users who visit many sites and have higher activity levels may be more likely to accidentally click links they should not, or to enter a password to a site that gives an attacker back door access to a system. While the government has some ability to block some undesired websites, attackers are getting smarter and are creating duplicate sites that can be hard to detect. Acquisition employees must be trained on how to navigate the internet, particularly when conducting market research, opening documents from RFIs, or browsing social media. Regarding internet usage, ask:
  - Does the employee confidently use the internet? Users who have become accustomed to routine or repetitive web activity (e.g., visiting the same sites over and over) might become too comfortable and pay less attention to crucial security measures, or they may fall victim to the aforementioned “autopilot” cognitive bias.
  - Does the employee know how to recognize a legitimate website vs. a duplicated or imitation one? Does the employee know how to properly read a URL and detect a spoofed address?
- **Cybersecurity**—2021 figures from research firm IDC indicate that the COVID-19 pandemic has coincided with a spike in many forms of network intrusion (many of which can be and have been enabled by social engineering techniques); the same research notes that in response to such phishing, DNS hijacking attacks, and other forms of compromise, many institutions have turned to zero-trust cybersecurity initiatives to mitigate threats.
- A zero-trust model is a security framework that fortifies the enterprise by removing implicit trust and enforces strict user and device authentication throughout the physical and logical network ecosystem (for example, increased requirements for two-factor authentication); following this model of security and/ or asking whether elements of this model could be employed within government



and contractor networks is a discussion worth initiating within the many acquisition subcommunities. Other deployable elements of healthy cybersecurity and cyber awareness might include using a trusted, legitimate Internet Service Provider, paying for higher grade antivirus software, making device updates mandatory and monitored, and training employees to verify the legitimacy of website certificates, doublecheck URLs and website spellings, and to look for a locked padlock icon within their browsers when working both in the office and at home.

Increased awareness of acquisition social engineering and training to recognize these attacks is a significant step that government agencies and companies can take toward better whole-of-system security. There are four signs that employees need to be on the lookout for when recognizing a social engineering attack:

1. The attacker will request something of value such as money, account passwords, or financial information. If anyone is asking for information that is known to be sensitive, that should immediately set off red flags that something about the situation is not right.
2. The attacker may imply or state that they wish the interaction to be secret or private. Even when operating in environments where information can be “need-to-know,” if the requester asks for the interaction to be private, the employee should ask why. If it’s not something that can be told to managers, it is not something the employee should be doing.
3. The attacker will try to rush the interaction so that the employee does not have sufficient time to think through the request or involve others that may detect the malign activity.
4. The attacker may pose as someone from a position of authority or influence. As mentioned above in the Principles of Influence section, a deference to authority is one of the six principles of influence and suggests that humans are more likely to agree with something without question if it comes from someone in a higher position than themselves.

### **Active Defense Measures and a Proactive Approach**

Training of employees is essential to successfully limiting the effects of social engineering attacks, however there are additional avenues that can be explored to assist employees.

U.S. businesses and government agencies should explore emerging technologies (including AI and ML tools) to assist them. For example, AI software can be implemented to flag emails that come to employees from an external address or with misspelled address information. This is sometimes seen by denoting “EXT” (external) at the heading of external emails or by adding a red banner or bold lettering to signal to the employee to take a closer look at the email and the source. Because COs constantly receive emails originating from external email addresses, they may become saturated with “EXT” which may cause no heightened awareness. In addition, internal emails testing employee knowledge and comprehension with rewards for success should be implemented. All of these measures can be put in place to help prevent social engineering attacks before they can occur.

The rise in social engineering as enabled by emerging technology also begs for a commensurate rise in sophisticated active defense research and development and execution. As mentioned in further detail below, many experts within the cybersecurity and counterintelligence industry view this goal as one that requires not just a new layer of tools and services, but one that will, over time, be best served by a paradigm shift in culture and organization management. Advancements and emerging methods in this field are more likely to positively impact the threat landscape—and secure assets—when they seek to focus on the multiple stages of manipulation (e.g., investigation, hook, play, exit) and the specific tactics currently employed by adversaries (note that this alludes to a need for threat managers to shift defensive measures in accordance with attack vectors over time and to develop automated and continuously retailored defensive tools that can be used against emerging and anticipated threats). Some techniques and elements of a forward-leaning defense posture could include the following:

- Advanced risk measurement and reporting tools—Risk can be measured in various ways (citing a litany of commercial platforms that provide this capability as a service), but



sophisticated tools often leverage best practices from deep learning neural networks and combine data points from security awareness, user and group security performance (e.g., following a phishing security test), past breaches, high value and high threat job functions, network security scores, adversary intent scores (based on target asset worth and accessibility), adversary capability scores, recent threat intelligence on known bad actors, and the like.

- Forward leaning network security—This applies to software, firmware, and hardware as standard pillars of defensive systems, but it should include corporate efforts to go beyond standard defensive cybersecurity practices (as those mentioned in the Defensive Factors and Vulnerabilities section) by prioritizing the hiring of a capable and engaged IT security department intimately familiar with the latest threats, emerging best practices, and an intent to collaborate with and train the workforce with engaging and exciting training regiments instead of dull and mandatory annual online courses. This IT team should be tasked to ensure that the latest AI- and ML-enabled tools are integrated as force multipliers into the IT infrastructure.
- Advanced and innovative approaches to deflect, defeat, and deter adversary operations Perhaps in partnership with the government and private industry technology partnerships, acquisition leadership should consider:
  1. Embracing experimentation as a test bed of prospective methods defending against social engineering; to date, no one method has proven to provide a fool proof defense against social engineering, thus allowing the acquisition community the time and resources to test new methods that will generate ground-up tools and procedures tailored to that community's needs. An example of this may include running experiments to analyze which security training module leads to a more informed workforce and more secure asset holdings; instead of allowing the leadership to focus on training compliance numbers, run three segregated training methods within three areas of operation, take note of defense successes in the aforementioned "risk measurement" metrics, and employ the leading practice.
  2. Reward innovative defense-focused ideas; experienced acquisition practitioners are postured to know their systems and target surface more than outsiders peering in. While the latest tools to be leveraged may rightly source from tech-focused outside organizations (thereby justifying deep collaboration), the specifics of where and how adversaries are targeting acquisition systems is likely to source from two areas: threat intelligence professionals and acquisition professionals on the inside working on the operational floor. Incentivizing (financially, organizationally, and culturally) the internal workforce to begin identifying, reporting, and offering solutions in response to these real-time threats heeds current digital transformation wisdom ("transforming a system requires transforming the system within it" [Leshchinskiy & Bowne, 2022]) and would give personnel a sense of empowerment over their own procedures (in the context of many project-burdened staffers being further taxed by mandatory training modules); this would also segue well into the following recommendation.
  3. Integrally collaborate with emerging technology-focused organizations working in the area of social engineering and network security solutions; innovators leading the movement toward greater system security are beginning to employ AI- and ML-enabled tools and creative low-cost solutions against many of the threats articulated in this paper, often viewing upfront costs as valuable investment. Examples of such solutions include:
    - Integrating honey trap/Potemkin Village targets within a defending system to lure attackers into areas without sensitive assets (such initiatives work to deflect, defeat, and deter threat actors, while data from collected threat intelligence can be leveraged to identify threats and signatures that may arise again in future operations).
    - Leveraging automated, AI- and ML-enabled threat detection, reporting, and mitigation; this can take the form of funneling attackers to a hollow Potemkin network, a "vulnerable and publicly accessible"



chatbot posing as an acquisition officer, or ML-enabled detection software that repurposes data artifacts from threat signatures to search for and block new or recurring threat actors. Providing discovered signatures or bad actors that continue to operate to threat intelligence professionals would also provide the intelligence workforce the opportunity to penetrate these social engineer networks to collect information on their intended future targets and techniques (information that can be cycled back to acquisition practitioners to enable a more intelligent and more tailored defense). This list of active defense tools enabled by emerging technology grows by the day; empowering IT managers to leverage the latest in Commercial Off the Shelf and automated products and services (many of which are provided by leading cybersecurity firms that enjoy preexisting, vetted relationships with the government) will bring the acquisition community into a league of modern defense. Reduce the attack surface, restrict task burdens, and shift organizational focus onto security where possible. Commensurate with the degree that acquisition and defense leadership seeks to increase security against social engineering threats, opportunities exist to limit the number of acquisition compliance activities required to complete an acquisition task; less online activity (where many acts provide many opportunities for threat actors to interact with and compromise acquisition systems) and reduced task burdens (where personnel are less distracted from security duties by the number of perfunctory duties) tend to reduce multiple forms of online threats posed to organizations.

## Conclusion

Social engineering attacks are an increasing challenge to businesses and government agencies across the U.S. Acquisition professionals have constant interaction with both internal and external stakeholders such as government acquisition and technical teams and industry contractors. This creates a unique situation of prospective exploitation that not only threatens sensitive governmental and commercial data but also funds, personnel, proprietary ideas, and democratic institutions. As social engineering continues to grow as a threat, so must the prevention and mitigation techniques put in place against them. While social engineering attackers continue to layer in more sophisticated tools and tradecraft, the USG and its acquisition community must level up into a forward leaning position ahead of them to outsmart and outgun the threat. With a final spirit of optimism, we remind our readers that the suite of technology and skills that underpins adversary capability advancements is the same toolkit that can enable a well-postured defense of tomorrow.

## References

- Aldawood, H., Alashoor, T., & Skinner, G. (2020). Does awareness of social engineering make employees more secure? *International Journal of Computer Applications*, 177(38), 45–49.  
<https://doi.org/10.5120/ijca2020919891>
- Access Systems. (2019). *The top ten social engineering tactics you need to know*.  
<https://www.accesssystems.com/blog/the-top-10-social-engineering-tactics-you-need-to-know>
- AGA Tools and Resources. (n.d.) *Social engineering*. AGA. Retrieved April 4, 2023, from  
<https://www.agacgfm.org/Resources/intergov/FraudPrevention/FraudMitigation/SocialEngineering.aspx>
- Arthur J. Gallagher & Co. (2016). *Social engineering fraud*. Wisconsin Association of School Boards.  
[https://www.wasb.org/wp-content/uploads/2017/04/20161219\\_aigallagher\\_social\\_engineering\\_fraud.pdf](https://www.wasb.org/wp-content/uploads/2017/04/20161219_aigallagher_social_engineering_fraud.pdf)
- Barnett, D. (2022, April 14). *The world of the subdomain*. CSC Digital Brand Services Blog.  
<https://www.cscdbs.com/blog/the-world-of-the-subdomain/>
- Cyber Armada Team. (2020). *Social engineering threats to the supply chain during COVID-19*. Cyber Armada Blog.  
<https://blog.cyber-armada.com/articles-and-resources/social-engineering-threats-to-the-supply-chain-during-covid-19>
- Federal Acquisition Regulation (FAR), 48 C.F.R. 2.101 (2023).
- Gatefy. (2021, June). *Ten real and famous cases of social engineering attacks*. <https://gatefy.com/blog/real-and-famous-cases-social-engineering-attacks/>
- Heinzman, A. (2022). That computer virus you can't remove might be a browser notification. *Review Geek—Make Gadgets Fun Again*. <https://www.reviewgeek.com/111106/that-computer-virus-you-cant-remove-might-be-a-browser-notification/>
- Influencing Others. (n.d.). [www.social-engineer.org](http://www.social-engineer.org). <https://www.social-engineer.org/framework/influencing-others/framing/>





Leshchinskiy, B., & Bowne, A. (2022). *Digital transformation is a cultural problem, not a technological one*. War on the Rocks. <https://warontherocks.com/2022/05/digital-transformation-is-a-cultural-problem-not-a-technological-one/>

Mazza, P., & Feinberg, M. (2020). *Social engineering fraud: 4 Steps every company needs to take right now*. JDSupra. <https://www.jdsupra.com/legalnews/social-engineering-fraud-4-steps-every-29657/>

McGuinness, T. (2020). What is the purpose of gaslighting? *LinkedIn*. <https://www.linkedin.com/pulse/what-purpose-gaslighting-tim-mcguinness-ph-d->

Nadeem, M. S. (2022). *Social engineering: What is pretexting?*. Malifence.com. <https://blog.mailfence.com/pretexting/>

National Institute of Standards and Technology (NIST). (2020). *Digital Identity Guidelines* (Special Publication 800-63-3).

O'Reilly, L. (2021). Social engineering threats rose 270% in 2021—Indicating a shift to multi-channel phishing attacks as apps and browsers move to the cloud. *Slashnext*. <https://www.slashnext.com/blog/social-engineering-threats-rose-270-in-2021-indicating-a-shift-to-multi-channel-phishing-attacks-as-apps-and-browsers-move-to-the-cloud/>

Partida, D. (2020). Social engineering cyberattacks and how they're affecting businesses. *Security Infowatch*. <https://www.securityinfowatch.com/cybersecurity/article/21203580/social-engineering-cyberattacks-and-how-theyre-impacting-businesses>

Polyakov, A. (2019). *Machine learning for cybercriminals*. Medium. <https://towardsdatascience.com/machine-learning-for-cybercriminals-a46798a8c268>

Proofpoint. (n.d.) What is pharming? Retrieved April 4, 2023, from <https://www.proofpoint.com/us/threat-reference/pharming/>

Ropek, L. (2019). *Social engineering attack nets \$1.7M in government funds*. GovTech. <https://www.govtech.com/security/social-engineering-attack-nets-17m-in-government-funds.html>

Sherman, J., & Arampatzis, A. (2018, July 18). *Social engineering as a threat to societies: The Cambridge Analytica case*. <https://www.realcleardefense.com/articles/2018/07/18/social-engineering-as-a-threat-to-societies-the-cambridge-analytica-case-113620.html>

Slater, D. (2021). *7 new social engineering tactics threat actors are using now*. CSO Online. <https://www.csoonline.com/article/3613937/7-new-social-engineering-tactics-threat-actors-are-using-now.html>

Watson, R. (2022). Cyberattacks are gaining momentum. *Grand Rapids Business Journal*. <https://grbj.com/news/technology/cyberattacks-are-gaining-momentum/>

## Appendix A: Acronyms

AGA	Association of Government Accountants
AI	Artificial Intelligence
APT	Advanced Persistent Threat
BEC	Business Email Compromise
CEO	Chief Executive Officer
CI	Counterintelligence
CO	Contracting Officer
COR	Contracting Officer's Representative
EXT	External
FAR	Federal Acquisition Regulation
FBI	Federal Bureau of Investigation
FTF	Funds Transfer Fraud
HUMINT	Human Intelligence
IT	Information Technology
ML	Machine Learning
NLP	Natural Language Processing
PII	Personally Identifiable Information
RFI	Request for Information
TTPs	Tactics, Techniques, and Procedures
U.S.	United States
USG	U.S. Government



# Comparative Analysis of Pathways to Changeability

**Aditya Singh**—is a PhD student at The George Washington University, studying Engineering Management in the Department of Engineering Management and Systems Engineering, where he previously received his BS in systems engineering and economics. Currently, he is a Doctoral Fellow on National Science Foundation Research Traineeship Program: “Co-Design of Trustworthy AI Systems. [asingh25@gwu.edu]

**Zoe Szajnfarber**—is a Professor and Chair of Engineering Management and Systems Engineering at The George Washington University. Her research seeks to understand the fundamental dynamics of innovation in technology-intensive governmental organization, as a basis for decision-making. She received her bachelor’s degree in Engineering Science from the University of Toronto. Szajnfarber conducted her graduate work at the Massachusetts Institute of Technology, earning dual master’s degrees in Aeronautics & Astronautics and Technology Policy and a doctorate in Engineering Systems. [zszajnfa@gwu.edu]

## Abstract

Through an examination of three cases of change in the U-2 platform, this paper compares three pathways to changeability: form changes, operational changes, and cyber changes. Each pathway can lead to change in similar properties of a system but have varying levels of performance and time to implement. For each pathway, we describe the design mechanisms necessary to implement change in that pathway. We analyze the trade-off between performance or extent of change and agility or speed of change and find that form changes offer the highest degree of changeability but take the longest time to implement. Operational changes offer the least degree of changeability but are far quicker to implement. Cyber changes lie in between these two pathways. Understanding the design choices needed and the underlying trade-off of each pathway can enable decision-makers to better select a pathway to change when the need arises. This comparative analysis is especially useful since literature has thus far examined each of these pathways in isolation, not as different paths to the same goal.

## Introduction

Complex engineered systems (CES), such as aircraft and ships, often entail protracted design phases and lengthy lifecycles. This gap between system conceptualization and system retirement introduces a great deal of uncertainty over the system lifecycle as new needs arise as the gap grows. To guard against this inherent uncertainty, CES are often required to be changeable, meaning that they can change in response a change in the operating environment. Design for changeability literature has typically focused on mechanisms that make changing the physical form of the system easier. Previous work identified that system users can change *how* the system is used to maintain value in a changing operating environment without risky and expensive form changes. Software design literature has also examined how software can be designed to more easily incorporate changes after the initial design phase. These three pathways to changeability, form, operational, and cyber, have not been connected in the design for changeability literature and have not been compared to each other in terms of agility and performance. This paper shows that form, operational, and cyber changes can be leveraged to achieve similar types of change and compares the speed of implementation and performance each type using three cases of change in the U-2 platform.

## Literature Review

Design for changeability literature is concerned with how systems maintain value in the face of changing operating environments. Changeability is an umbrella term that



captures many strategies for how systems can change in response to a change in operating conditions (Fricke & Schulz, 2005). Four key strategies are adaptability, flexibility, scalability, and modifiability. Adaptable systems initiate change through internal change agents, while flexible systems initiate change throughout external change agents. Automatic software updates are an example of an internal change agent, while a technician modifying a system is considered an external change agent. Scalability refers to change the level of some system parameter, like bandwidth. Modifiability refers to the ability to move system parameters from agent to agent, such as using a dongle to connect a new subsystem to an existing computer (Ross et al., 2008). There are several more strategies, collectively referred to as the -ilities (de Weck et al., 2012) (Beesemyer, 2012) (Ross & Rhodes, 2019), but they are not covered for brevity and relevancy.

These strategies need specific mechanisms to be implemented. Changeability mechanisms are specific design choices that enable these strategies to be carried out. One of the most popular mechanisms is modularity, which involves a one-to-one mapping of function and module. Modules are loosely coupled with each other and the rest of the system, but modules themselves are often comprised of tightly coupled components (Baldwin & Clark, 2000). Modules rely on common interfaces to be easily swapped in and swapped out. While modularity continues to be a popular changeability mechanism in industry, modularity often comes at the cost of design optimization and performance of the system (Höltkä et al., 2005).

Real options are another popular mechanism for changeability. Stemming from finance, real options in engineering are contract tools that give system buyers the right but not the obligation to implement a change in the future (de Neufville, 2003). A classic example of real options is a parking garage where system buyers might include an option to add additional floors to the structure at some point in the future. This requires an upfront investment in the option, to make the foundations stronger to accommodate the potential change, and can be executed in the future if the buyers decide there is enough demand to justify the execution cost (de Neufville et al., 2006). Real options are rarely executed perfectly as technical, logistic, and organizational delays can create a gap between when the option is executed and when the option is fully implemented. The value of real options degrades as implementation delays arise (Sapol & Szajnfarber, 2020).

Margin, the excess of a system property beyond its required level, is another significant change mechanism. Margin has been tied mostly to evolvability, the transfer of common system traits from generation to generation (Allen et al., 2016; Tackett et al., 2014). Building in margin for design is related to adding in safety margin, which is a common practice in many fields like civil engineering (Eckert & Isaksson, 2017). Previous work identified margin as a key enabler of modularity and flexibility as well (Singh & Szajnfarber, 2022), but modern systems face many design requirements that require physically optimized design. Physically optimized design means an elimination of margin, which can limit the amount of form changes a system can accommodate.

Literature identified that changing *how* the system is used can enable changeability (Mekdeci et al., 2015). These operational changes can even provide systems with new capabilities, thus avoiding risky, expensive, and/or time consuming changes to the form of the system (Singh & Szajnfarber, 2022). Operational changes are often generated by system users, who are considered to be agents of changeability within the system (Cox, 2017). While changing how a system is operated has been shown to be a mechanism of changeability, it is still limited by the form of the system. Users can only do so much with the system that they have. This creates a need to change the system without extensive form changes, which can be accomplished through cyber changes.



While changeability literature has largely focused on form changes, there have been some considerations of changeability through software. In their seminal paper, Fricke and Schulz describe how automatic software updates could be a mechanism for achieving adaptability (Fricke & Schulz, 2005). Since then, others have created and discussed changeability as it relates to software, primarily relating to software quality (Brown et al., 2022). For example, researchers have discussed the maintainability of a software system, which is further subdivided into the repairability and modifiability of said system (Chen et al., 2018). Modifiability, the ability of a system to accommodate a change, is most closely related to how systems can add capabilities (Bachmann et al., 2007). Reducing coupling, a strategy to create modular systems, is also a key technique in software design. Delaying binding time, when a flexible software feature becomes fixed (Krisper & Kreiner, 2016), and increasing cohesion within modules to reduce overall module complexity are also key strategies within modifiability. Specific design mechanisms for each of these sub-strategies have been discussed in literature (Bachmann et al., 2007).

Many studies in software changeability are focused on the repair of these systems. Even those that are focused on adding or enhancing capabilities often cite software evolution and the pace of change in software as a key motivation for why change is needed. This is due, in part, to most of these studies focusing on software systems and not cyber-physical systems specifically.

Helen Gill coined the term cyber-physical systems, defining them as “systems with integrated computational and physical capabilities that can interact with humans ... and expand the capabilities [of] the physical world through computation, communication, and control” (Baheti & Gill, 2011, p. 161). Cyber-physical systems are deployed in very different environments than software only systems and face different change motivators. Cyber-physical systems have been identified as key platforms for changeability since the incorporation of several types of systems increases the trade space of changes that can be implemented and increases the number of experts due to the variety of systems found in cyber-physical systems today (Colombo, 2016).

Nevertheless, changeability literature falls short on analyzing how software design can enable new capabilities in cyber-physical systems. While software design literature has detailed mechanisms to achieve modifiability and other changeability mechanisms, changeability literature has failed to appropriately appreciate cyber pathways to change, especially in terms of adding or enhancing new capabilities in the field. Complex engineered cyber-physical systems, like many of today’s air and spacecraft, face less pressure to change from market forces and technological evolution, and face more pressure from changing operating environments over long lifecycles. Responding to these changes by adding and enhancing capabilities using software will be an important capability for complex engineered cyber-physical system operators and needs further investigation into how it can be enabled and how it is implemented.

## Methods

To investigate cyber pathways for changeability and compare them to other pathways of change, we examine three instances of change where U-2 targeting, imaging, and sensing capabilities were updated. Aircraft are a prime example of cyber-physical systems as modern jets are becoming more cyber reliant, while still relying on their physical form to accomplish their tasks. Cyber components of aircraft are often used to interface with physical components and can enable certain capabilities. Fricke and Schulz (2005) characterized systems that have a well-defined core function but highly variable secondary functions, have long lifecycles but rapid technology integration requirements, operate in a



system of systems environment, and have high deployment and maintenance costs as those that are best suited for changeable architecture (p. 7). Older military aircraft fit these criteria and have substantial publicly available information that is not available for commercial or modern military aircraft.

One instance is of a form change implemented through the Agile Pod system, another instance is of an operational change implemented during Desert Storm, and the final instance is of a software change implemented recently. Table 1 presents a summary of the three cases of change in the U-2 platform analyzed in this paper. We analyzed what necessitated the change, the extent of the change implemented, and the time required to implement the change. Through this analysis, we find that there is a trade-off between the extent of the change that is implemented and speed at which it can be implemented. Form changes are the most extensive, providing the highest degree of change but requiring the most amount of time to implement, while operational changes offer the lowest degree of change but require the least amount of time to change. Cyber changes lie in between form and operational changes on the extent and speed trade-off axis. There is a delay in developing software, but implementation can be instantaneous if over-the-air updates are enabled. Each case is discussed further in this section. For each pathway of change, we also discuss the upfront design requirements to implement, if any.

**Table 8: U-2 Results Table**

	<i>Need for Change</i>	<i>Extent of Change</i>	<i>Speed of Change</i>
<i>AgilePod (Form)</i>	Need to integrate multiple sensors & cameras onto U-2 and quickly swap equipment for different missions	Modular pod created that can swap different sensors in and out; leverages common mechanical and electrical interfaces	Useable prototype delivered in 18 months
<i>H-Cam (Operational)</i>	Request for higher resolution on intelligence images from H-cam on U-2; H-cam operates at an angle to capture maximum amount of ground	Camera angle changed to straight down for higher resolution; new flight routes developed	Changes implemented in a matter of days after camera angle was mechanically changed and new flight routes were planned
<i>Kubernetes (Cyber)</i>	Need to account for new types of targets not planned for originally	Improved automatic targeting algorithm developed and installed	Software created in weeks, implemented instantly over-the-air

## Differences in Implementing Different Pathways

### U-2 Agile Pod (Form)

Intelligence, surveillance, and reconnaissance (ISR) is a core requirement of the United States Air Force (USAF) which is comprised of several different missions, each with their own equipment needs. This variety of mission and associated equipment creates a difficult logistical environment since not all aircraft are able to accommodate each piece of equipment. The Air Force realized the need to enable aircraft to swap in and swap out ISR equipment easily and quickly (Trevithick, 2018a). To meet the challenge, USAF developed a pod made up of several compartments ranging in size that can be reconfigured to accommodate a variety of ISR equipment. Several iterations of the pod, known as the AgilePod, have been created to match different requirements, primarily focused on size to accommodate what the aircraft can hold and what the aircraft needs for each mission. AgilePod uses common interfaces and creates a single physical and electrical interface that can be mounted on aircraft pylons (Nine et al., 2019; Shirey et al., 2017).



Recently, the Air Force awarded KEYW a contract to develop an AgilePod to accommodate a variety of ISR equipment. The pod was delivered in prototype form to the Air Force within 18 months (Alia-Novobilski, 2016; Cogliano, 2015; Trevithick, 2018a). A recent iteration of an agile pod was installed on an aircraft in a hangar at Wright-Patterson Air Force Base in Ohio for testing in a matter of weeks, showing how rapidly these AgilePods can enable new capabilities (Alia-Novobilski, 2018). Once installed, swapping ISR equipment becomes tantamount to swapping out ordnance on a fighter jet. The AgilePod was installed on U-2s, and contract vehicles have been created to develop new sensors for the AgilePod family of sensors (Trevithick, 2018b).

While AgilePod is one of the most agile and flexible systems in the Air Force acquisition pipeline, design and development took over a year, and a fit test took weeks. The test was conducted in the United States, but if AgilePod needed to deploy to an international field, additional logistic constraints and delays would arise. AgilePod provides a useful baseline for implementing rapidly needed capabilities even though it is not a fully fielded system on the U-2. Modular systems that provide new capabilities have been shipped to the field without full testing in the past, as noted in previous with the GPU-5/A sent to Desert Storm (Singh & Szajnarfarber, 2022; Smith, 2021).

### **U-2 Camera Positioning (Operational)**

Desert Storm was the largest U-2 operation in U.S. military history, providing key intelligence and targeting information to allied forces. U-2s operating in Desert Storm and Desert Shield carried a variety of sensors and cameras, including the High Resolution 329 camera (H-cam). The H-cam's normal concept of operations is to place the camera at an angle in the gyrostabilized compartment to provide the maximum amount of coverage. Those in the field relying on the data needed greater resolution for the H-cam data to be useful. To accomplish this, "Lieutenant Colonels Lafferty and Spencer ... decided to revise the H-camera's procedures" by shooting the camera straight down instead of at a coverage maximizing angle (Cross II, 2014, p. 41). This required technicians to reposition the camera in the compartment and required planners to redevelop the flight paths to accommodate for the loss of aerial photography coverage area. Through these operational changes, U-2 operators and intelligence officers were able to greatly improve image quality, over what the camera was advertised as offering, without having to acquire a new camera system (Cross II, 2014).

### **U-2 Targeting Software (Cyber)**

A U-2 recently received an over-the-air update that improved the aircraft's automatic targeting system (Trevithick, 2020). The update is the first time that military software was updated on an aircraft while the aircraft was in flight (Insinna, 2020). In-flight updates were made possible by Kubernetes, an open-source software containerization system developed by Google and donated to the Cloud Native Computing Foundation. Kubernetes enables developers to automate a large degree of testing and development through software modularization and reuse (Trevithick, 2020). To use Kubernetes, system functions need to be decoupled so that developers can quickly swap software modules without affecting the entire system. This type software module to system function mapping is the same modularity strategy employed by designers of physical systems. While software modules are mapped to system functions and loosely coupled with each other, the modules are tightly coupled within themselves as each Kubernetes module has all dependencies and libraries within the module. Being able to quickly swap modules in software and hardware are very similar in their design requirements, but they require extremely different logistical considerations to implement (Insinna, 2020). Kubernetes was installed on the U-2's existing computers without the need for new electronics or avionics. Following the U-2 over-the-air update, the



Kubernetes system was installed on F-16s in 45 days showing how rapidly software open architecture can be installed on a system (Chaillan, 2019). While complete function to module mapping was not completed in this 45-day span, F-16s were subsequently able to receive an over-the-air update that provided new electronic warfare data files. The update was initiated from an Air Force base hundreds of miles away from where the F-16 was flying when it received the update (F-16 System Program Office, 2021).

## Analysis

U-2s have shown that changeability can be achieved through form, operational, and cyber pathways of change. Each case covered related to some aspect of ISR for the same system, showing that each pathway could be used in the same context. The extent of changeability for each pathway of change was quite different. Form changes require the most extensive logistical requirements, with physical systems needing to be procured, produced, and shipped for installation. The AgilePod that was recently developed required 18 months to get to the prototype phase, showing how time-consuming physical system development can be, even when the product is based on an existing product framework. Even if the physical equipment needed is already produced, shipping and installation can introduce heavy tolls on logistical capacity, with large potential for severe delays (Sapol & Szajnfarder, 2020). Equipment like the AgilePod represent a best-case situation for form changes, as it leverages existing common interfaces and is designed to be extremely modular. Being able to add or swap equipment creates the largest trade space of possible changes, creating a trade-off between agility and the extent of changeability. This trade-off is reversed with operational changes.

Operational changes can be implemented very easily with system users changing how their system is used without extensive changes to the form of the system. Conceptualizing the change and training enough to ensure that new concepts of operation are effective require a highly variable amount of time but are generally much faster than implementing a new form change, as seen in the U-2 H-cam change and as noted in previous case study work (Singh & Szajnfarder, 2022). Adding to the agility of operational changes is that they do not require upfront design considerations. Systems need to be designed to easily accommodate future form changes but do not require such design considerations. While extremely agile, operational changes are restricted in degree of change they can create in a system. Operational changes that aim to improve capabilities or gain new capabilities in the field are generally initiated when system users face an urgent need and do not have time to wait on a form change to be initiated and implemented. This means that system users have to work with the system they have, not the system they want. While the H-cam change showed how changing how a system is used can increase its capabilities even beyond what system designers were willing to advertise, operational changes are still constrained by the physical limitations of their physical systems.

Cyber changes are a newer pathway of change that seem to be in the middle of form and operational changes on the agility and extent of changeability continuum. Similar to form, software requires system design choices that enable future changes to be easily implemented. The case discussed in this paper leveraged software modularity as a key strategy for changeability, requiring many of the same design considerations as physical modularity including loose coupling between modules and tight coupling within modules. A key difference, however, is when systems can be made modular. Decoupling physical components is far more difficult than decoupling software systems and this can be done after the fact, as the U-2 and F-16 software components were not explicitly designed with software modularity in mind. Physical systems are more defined by their initial design than



software systems, representing a timeline shift in when these design choices need to be made.

In terms of agility of implementation, software has been created and installed on platforms like the F-22 through over-the-air updates in a matter of just 60 days (Hadley, 2022). When software is already created and need to be transmitted, over-the-air updates enabled almost instantaneous implementation. This is not to say that software implementation does not require extensive logistical capabilities to be in place. The F-16 update used a satellite to implement, and other platforms hoping to take advantage of the agility of over-the-air updates need to have reliable access to transition and enough computing power available to implement. If these capabilities are in place, cyber changes can be implemented rapidly, but if they are not, cyber changes would require systems to return to a central depot, making them more akin to slower form changes.

In terms of extent of changeability, the limits of cyber change for CPS are being pushed constantly. Recently, Tesla and Mercedes released optional software updates that could be implemented over-the-air that would make their cars faster, meaning that software changes can impact the maximum physical performance of a system (Gerken, 2022). Making cars faster and improving targeting software are both examples of improving a system's existing capabilities, but the F-16 change represented "the first time a fighter aircraft has received a software update and gained new capability all while in flight" (F-16 System Program Office, 2021). As software is increasingly used to control and manipulate physical system properties, the trade space of changes that can be implemented through software-only changes will increase. Additionally, software updates may have unique interactions with other forms of change. For example, battery optimization software might be able to create margin in power supply where there was none before, enabling physical changes that take advantage of newly created margin.

## Conclusion

By examining three cases of change, we showed that form, operational, and software changes enhanced capabilities in the same mission area for the same platform. We additionally examined the design choices required to implement each change, the speed at which the change was implemented, and the extent of the change. Through this examination, we reveal the trade-off between agility and extent of change. Form changes are least agile but have the highest extent of changeability and require upfront design considerations. Operational changes are the most agile but have the least extent of changeability as system users must work within the constraints of the system. These changes do not require upfront design choices. Cyber changes lie in between form and operational changes on the agility and performance trade-off axis. Implementing cyber changes in the field requires modular design, but modularity can be superimposed on existing cyber physical systems after production. Additionally, over-the-air updates require infrastructure investments to relay updates from some location to the system in the field. If proper design and infrastructure is in place, cyber change implementation is only delayed by the time required to develop software. For practitioners, understanding these pathways and their associated trade-offs can enable better decision making about the type of change that should be undertaken based on the extent and urgency of the change needed.

## References

- Alia-Novobilski, M. (2016, December 28). *AgilePod 'reconfiguring' ISR mission*. Wright-Patterson AFB. <https://www.wpafb.af.mil/News/Article-Display/Article/1038723/agilepod-reconfiguring-isr-mission/https%3A%2F%2Fwww.wpafb.af.mil%2FNews%2FArticle-Display%2FArticle%2F1038723%2Fagilepod-reconfiguring-isr-mission%2F>





- Alia-Novobilski, M. (2018, January 2). *AFRL's AgilePod shows ISR versatility during Scorpion fit test*. Wright-Patterson AFB. <http://www.wpafb.af.mil/News/Article-Display/Article/1406999/afrls-agilepod-shows-isr-versatility-during-scorpion-fit-test>
- Allen, J. D., Mattson, C. A., & Ferguson, S. M. (2016). Evaluation of system evolvability based on usable excess. *Journal of Mechanical Design*, 138(9). <https://doi.org/10.1115/1.4033989>
- Bachmann, F., Bass, L., & Nord, R. (2007). *Modifiability tactics* (Technical Report CMU/SEI-2007-TR-002). Software Architecture Technology Initiative. <https://doi.org/10.21236/ADA472581>
- Baheti, R., & Gill, H. (2011). Cyber-physical systems. *The Impact of Control Technology*, 161–166.
- Baldwin, C. Y., & Clark, K. B. (2000). *Design rules: The power of modularity*. MIT Press.
- Beesemyer, J. C. (2012). *Empirically characterizing evolvability and changeability in engineering systems* [Thesis, Massachusetts Institute of Technology]. <https://dspace.mit.edu/handle/1721.1/76092>
- Brown, M., Dey, S., Tuxworth, G., Co, J., Bernus, P., & de Souza, P. (2022). An ility calculation for satellite software validation. *2022 IEEE Aerospace Conference (AERO)*, 1–20. <https://doi.org/10.1109/AERO53065.2022.9843603>
- Chaillan, N. (2019, November 22). *How the Department of Defense moved to Kubernetes and Istio*. KubeCon + CloudNativeCon North America 2019, San Diego, California. <https://www.youtube.com/watch?v=YjZ4AZ7hRM0>
- Chen, C., Lin, S., Shoga, M., Wang, Q., & Boehm, B. (2018). How do defects hurt qualities? An empirical study on characterizing a software maintainability ontology in open source software. *2018 IEEE International Conference on Software Quality, Reliability and Security (QRS)*, 226–237. <https://doi.org/10.1109/QRS.2018.00036>
- Cogliano, J. (2015, June 12). KEYW Corp. Tapped to develop more affordable, flexible aircraft spy pods. *Dayton Business Journal*. <https://www.bizjournals.com/dayton/news/2015/06/12/exclusive-firm-tapped-to-develop-more-affordable.html>
- Colombo, E. F. (2016). *Open innovation meets changeability: Strategic design analyses for cyber-physical industry platforms* [PhD, Politecnico di Milano]. <https://www.politesi.polimi.it/handle/10589/117765>
- Cox, A. (2017). *Functional gain and change mechanisms in post-production complex systems* [The George Washington University]. <https://www.proquest.com/openview/f0c8db7bd87e0fdc086e13fbe0ba523a/1.pdf?cbl=18750&pq-origsite=gscholar>
- Cross II, C. F. (2014). *The dragon lady meets the challenge: The U-2 in Desert Storm* (p. 125). 9th Reconnaissance Wing Historian.
- de Neufville, R. (2003). Real options: Dealing with uncertainty in systems planning and design. *Integrated Assessment*, 4(1), 26–34. <https://doi.org/10.1076/iaij.4.1.26.16461>
- de Neufville, R., Scholtes, S., & Wang, T. (2006). Real options by spreadsheet: Parking garage case example. *Journal of Infrastructure Systems*, 12(2), 107–111. [https://doi.org/10.1061/\(ASCE\)1076-0342\(2006\)12:2\(107\)](https://doi.org/10.1061/(ASCE)1076-0342(2006)12:2(107))
- de Weck, O. L., Ross, A. M., & Rhodes, D. H. (2012). *Investigating relationships and semantic sets amongst system lifecycle properties (Iilities)* [Working Paper]. Massachusetts Institute of Technology, Engineering Systems Division. <https://dspace.mit.edu/handle/1721.1/102927>
- Eckert, C., & Isaksson, O. (2017). Safety margins and design margins: A differentiation between interconnected concepts. *Procedia CIRP*, 60, 267–272. <https://doi.org/10.1016/j.procir.2017.03.140>
- F-16 System Program Office. (2021, July 31). *F-16 receives in-flight software update during recent flight test*. Air Force. <https://www.af.mil/News/Article-Display/Article/2715206/f-16-receives-in-flight-software-update-during-recent-flight-test/https%3A%2F%2Fwww.af.mil%2FNews%2FArticle-Display%2FArticle%2F2715206%2Ff-16-receives-in-flight-software-update-during-recent-flight-test%2F>
- Fricke, E., & Schulz, A. P. (2005). Design for changeability (DfC): Principles to enable changes in systems throughout their entire lifecycle. *Systems Engineering*, 8(4), 342–359. <https://doi.org/10.1002/sys.20039>



- Gerken, T. (2022, November 24). *Mercedes-Benz to introduce acceleration subscription fee*. BBC News. <https://www.bbc.com/news/technology-63743597>
- Hadley, G. (2022, September 1). F-22 flies with third-party apps, new open software architecture. *Air & Space Forces Magazine*. <https://www.airandspaceforces.com/f-22-flies-with-third-party-apps-new-open-software-architecture/>
- Höltkä, K., Suh, E. S., & de Weck, O. (2005). Tradeoff between modularity and performance for engineered systems. *DS 35: Proceedings ICED 05, the 15th International Conference on Engineering Design, Melbourne, Australia, 15.-18.08.2005*, 449–450.
- Insinna, V. (2020, October 19). *US Air Force sends software updates to one of its oldest aircraft midair*. Defense News. <https://www.defensenews.com/air/2020/10/09/the-air-force-updated-the-software-on-one-of-its-oldest-aircraft-while-it-was-in-the-air/>
- Krisper, M., & Kreiner, C. (2016). Describing binding time in software design patterns. *Proceedings of the 21st European Conference on Pattern Languages of Programs*, 1–15. <https://doi.org/10.1145/3011784.3011811>
- Mekdeci, B., Ross, A. M., Rhodes, D. H., & Hastings, D. E. (2015). Pliability and viable systems: Maintaining value under changing conditions. *IEEE Systems Journal*, 9(4), 1173–1184.
- Nine, J., Shirey, R., Thompson, B., George, A., Cunningham, J., & Mason, A. (2019). Open adaptable architecture (OA2) for agile intelligence surveillance reconnaissance (ISR). *Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation 2019*, 11015, 102–112. <https://doi.org/10.1117/12.2518819>
- Ross, A. M., & Rhodes, D. H. (2019). Ilities semantic basis: Research progress and future directions. *Procedia Computer Science*, 153, 126–134. <https://doi.org/10.1016/j.procs.2019.05.063>
- Ross, A. M., Rhodes, D. H., & Hastings, D. E. (2008). Defining changeability: Reconciling flexibility, adaptability, scalability, modifiability, and robustness for maintaining system lifecycle value. *Systems Engineering*, 11(3), 246–262. <https://doi.org/10.1002/sys.20098>
- Sapol, S. J., & Szajnarfarber, Z. (2020). Revisiting flexibility in design: An analysis of the impact of implementation uncertainty on the value of real options. *Journal of Mechanical Design*, 142(12). <https://doi.org/10.1115/1.4047682>
- Shirey, R. G., Borntreger, L. A., Soine, A. T., & Green, D. M. (2017). Blue Guardian: Open architecture intelligence, surveillance, and reconnaissance (ISR) demonstrations. *Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation 2017*, 10205, 123–129. <https://doi.org/10.1117/12.2263864>
- Singh, A., & Szajnarfarber, Z. (2022). Understanding post-production change and its implication for system design: A case study in close air support during Desert Storm. *Naval Engineers Journal*, 134(3), 87–95.
- Smith, P. (2021). *Fairchild Republic A-10 Thunderbolt II: The “Warthog” ground attack aircraft*. Pen and Sword.
- Tackett, M. W. P., Mattson, C. A., & Ferguson, S. M. (2014). A model for quantifying system evolvability based on excess and capacity. *Journal of Mechanical Design*, 136(5), 051002. <https://doi.org/10.1115/1.4026648>
- Trevithick, J. (2018a, January 3). *USAF uses Textron’s Scorpion Jet as the latest testbed for its modular sensor pod*. The Drive. <https://www.thedrive.com/the-war-zone/17352/usaf-uses-textrons-scorpion-jet-as-the-latest-testbed-for-its-modular-sensor-pod>
- Trevithick, J. (2018b, September 10). *USAF plans to drastically boost flexibility of U-2s and RQ-4s by adding modular “Agile Pods.”* The Drive. <https://www.thedrive.com/the-war-zone/23489/usaf-plans-to-dramatically-boost-flexibility-of-u-2s-and-rq-4s-by-adding-modular-agile-pods>
- Trevithick, J. (2020, October 19). *U-2 spy plane got new target recognition capabilities in first ever in flight software updates*. The Drive. <https://www.thedrive.com/the-war-zone/37131/u-2-spy-plane-got-new-target-recognition-capabilities-in-first-ever-in-flight-software-update>



## PANEL 20. FINDING AND LEVERAGING SOURCES OF ASYMMETRIC ADVANTAGE IN DEFENSE ACQUISITIONS

Thursday, May 11, 2023	
2:15 p.m. – 3:30 p.m.	<p><b>Chair: Todd Harrison</b>, Managing Director of Metrea Strategic Insights and non-resident Senior Advisor at CSIS</p> <p>Asymmetries and their Potential for Enduring Advantage Todd Harrison, Metrea Strategic Insights</p> <p><b>Panelists:</b></p> <p><b>Jacquelyn Schneider</b>, Hoover Fellow at the Hoover Institution and an affiliate with Stanford's Center for International Security and Cooperation</p> <p><b>Mandy Vaughn</b>, CEO of GXO Inc. and former President of VOX Space</p> <p><b>Jennifer “JJ” Snow</b>, Operating Partner at Metrea Discovery Partners</p> <p><b>Dr. Derek Tournear</b>, Director of the Space Development Agency (SDA)</p>

**Todd Harrison**—is a non-resident senior associate with the Aerospace Security Project and Defense Budget Analysis program at the Center for Strategic and International Studies (CSIS). Mr. Harrison is currently the managing director of Metrea Strategic Insights, where he leads the company’s efforts to conduct innovative, insightful, and pathfinding research. Prior to joining Metrea in May 2022, Mr. Harrison was a senior fellow and the director of the CSIS Defense Budget Analysis program and Aerospace Security Project, where he led the Center’s efforts to provide in-depth, nonpartisan research and analysis of defense funding, space security, and air power issues. Mr. Harrison joined CSIS from the Center for Strategic and Budgetary Assessments (CSBA), where he was a senior fellow for defense budget studies. At both CSIS and CSBA, Mr. Harrison authored numerous publications on trends in the defense budget, military space systems, threats to space systems, civil space exploration, defense acquisitions, military compensation and readiness, and military force structure, among other topics. Mr. Harrison previously worked at Booz Allen Hamilton, where he consulted for the U.S. Air Force on satellite communications systems and supported a variety of other clients evaluating the performance of acquisition programs. Prior to Booz Allen, he worked for AeroAstro Inc., a small start-up developing advanced space technologies, and as a management consultant at Diamond Cluster International. Mr. Harrison previously served as a captain in the U.S. Air Force Reserves and is a graduate of the Massachusetts Institute of Technology with both a BS and an MS in aeronautics and astronautics. He is currently a non-resident senior associate at CSIS, a member of the National Security Space Association Board of Advisors, and an adjunct faculty member at the Johns Hopkins School of Advanced International Studies.

**Jacquelyn Schneider**—is a Hoover Fellow at the Hoover Institution and an affiliate with Stanford's Center for International Security and Cooperation. Her research focuses on the intersection of technology, national security, and political psychology with a special interest in cybersecurity, autonomous technologies, wargames, and Northeast Asia. She is a non-resident fellow at the Naval War College's Cyber and Innovation Policy Institute and was previously a senior policy advisor to the Cyberspace Solarium Commission. Dr. Schneider was a 2020 winner of the Perry World House-Foreign Affairs Emerging Scholars Policy Prize. She is also the recipient of a Minerva grant on autonomy (with co-PIs Michael Horowitz, Julia Macdonald, and Allen Dafoe), a University of Denver grant to study public responses to the use of drones (with Macdonald), and a grant from the Stanton Foundation to study networks, cyber, and nuclear stability through wargames. Dr. Schneider is an active member of the defense policy community with previous positions at the Center for a New American Security and the RAND Corporation. Before beginning her academic career, she spent six years as an Air Force officer in South Korea and Japan and is currently a reservist assigned to US Space Systems Command.



She has a BA from Columbia University, MA from Arizona State University, and PhD from George Washington University.

**Mandy Vaughn**—is one of the most innovative and influential leaders in the US space industry. She founded GXO in 2021 to drive innovation, unlock new opportunities, and address emerging threats. Prior to founding GXO, at Virgin Orbit, Mandy worked on the LauncherOne program for government and commercial customers—including NASA and OneWeb. She served as president and CEO of VOX Space (a Virgin Orbit subsidiary) with a specific focus on the national security launch market. Additionally, Mandy has worked within General Dynamics Mission Systems, the Air Force as an active duty officer, and Kinsey Technical Services to advance next-generation innovations in GPS, satellite communications, defense, and security. Mandy holds a BS in Mechanical Engineering and an MS in Aeronautics and Astronautics, both from MIT.

**Jennifer “JJ” Snow**—brings 24 years of experience as a technologist, strategist and innovation officer operating across a diverse set of technologies and projects in support of the Department of Defense, Interagency, and Allied partners. She has served as an advisor on technology and innovation projects with John Hopkins University, Stanford University, Georgetown University and the University of Wisconsin-Madison. JJ’s work has been highlighted by seniors at the White House, the National Security Council, and Special Operations Command as well as international leadership across academia and the private sector. Her insights have been crucial to partners seeking to understand emergent risks and opportunities in technology influenced environments. As a military officer, JJ has completed tours in Iraq and Afghanistan including a seven month deployment where she was responsible for 67 locations across the country and \$365M in assets. She has served with Air Force Special Operations Command, Joint Special Operations Command, U.S. Special Operations Command, the National Security Agency and the Pentagon. JJ is a Distinguished Graduate of the Naval Postgraduate School. JJ is a Strategic Advisor and Partner at Verus Advisory. She also serves as the CTO for The Mentor Project and is an Innovation Advisor for MoonMark Space.

**Dr. Derek Tournear**—is currently the Director of the Space Development Agency, within the Office of the Under Secretary of Defense for Research & Engineering (OUSD(R&E)). SDA will unify and integrate space capability development and deployment across the department to achieve the DoD space vision while reducing overlap and inefficiency. He previously served as the Assistant Director for Space, responsible for developing the research and engineering roadmap to address future gaps in the DoD space architecture. Dr. Tournear previously held leadership roles in industry, most recently the director for the Harris Space & Intelligence (SIS) research & development. SIS was a \$2B business focused on providing advanced technical solutions addressing the top National Security threats from underwater to outer space. Prior to industry, Dr. Derek Tournear was a Senior Program Manager (SNIS-HQE) at the Intelligence Advanced Research Projects Activity (IARPA) in the Office of the Director of National Intelligence (ODNI). At IARPA, Dr. Tournear served as a senior scientist for space activities and space technologies in the Office of Smart Collection. Prior to IARPA, Dr. Tournear was a Program Manager for the Defense Advanced Research Projects Agency (DARPA), Tactical Technology Office and Strategic Technology Office. At DARPA, he initiated and directed a large portfolio of program, with an emphasis on sensors and space. He has professional experience at Los Alamos National Laboratory (LANL) managing intelligence and defense programs. During his time at LANL, he initiated a new field of study in gamma ray optics, and developed sensors for nuclear material detection. Dr. Tournear has a Ph.D. in physics from Stanford University and a B.S. from Purdue University. In 2010 he received an “Outstanding Alumnus” award from Purdue University and a 2008 DARPA award for “Outstanding Accomplishments in a Systems Technology Area.” Dr. Tournear is a 2011 recipient of the Secretary of Defense Medal for Exceptional Public Service, and a 2012 recipient of the Office of Director of National Intelligence Award for Exceptional Public Service.



# Asymmetries and their Potential for Enduring Advantage

**Todd Harrison**—is the Managing Director of Metrea Strategic Insights. Prior to joining Metrea in May 2022, Harrison was a senior fellow and the director of Defense Budget Analysis and the Aerospace Security Project at the Center for Strategic and International Studies (CSIS). He joined CSIS from the Center for Strategic and Budgetary Assessments (CSBA), where he was the senior fellow for Defense Budget Studies. At both CSIS and CSBA, Harrison authored numerous publications on trends in the defense budget, military space systems, threats to space systems, civil space exploration, defense acquisitions, military compensation and readiness, and military force structure, among other topics. Before joining the think tank community, Harrison worked as a consultant to Air Force Space Command while at Booz Allen Hamilton, as a program and product manager at space startup AeroAstro Inc., and as a management consultant at Diamond Cluster International. Harrison served in the U.S. Air Force Reserves and is a graduate of the Massachusetts Institute of Technology with both a BS and an MS in aeronautics and astronautics. He is currently a non-resident senior associate at CSIS, a member of the National Security Space Association Board of Advisors, and an adjunct faculty member at the Johns Hopkins School of Advanced International Studies where he teaches classes on the defense budget and military space systems.

## Abstract

The 2022 National Defense Strategy calls for a renewed focus on identifying and leveraging asymmetries to better direct investments in ways that will yield enduring military advantage. The pursuit of asymmetric advantage, however, is not new and has been part of military strategy for centuries. This paper—a preview of a more comprehensive forthcoming paper from Metrea Strategic Insights—uses examples from nature and military history to develop a framework for assessing the potential of an asymmetry to provide enduring military advantage. The framework consists of five key factors: how immutable the source of the asymmetry is, how difficult it is to copy or counter, at what level of effect the asymmetry is anchored and how applicable it is across the spectrum of operations, the degree to which it builds on other underlying asymmetries, and how well it scales. The paper applies the framework to assess three example areas of competition that are often touted as potential asymmetries: ubiquitous ISR, hypersonic weapons, and commercial innovation. The paper finds that asymmetries vary significantly in their ability to endure, the degree to which they maximize leverage, and their potential to scale effects exponentially. The framework presented can help inform which asymmetries are best aligned with defense strategy and how defense resources can be most effectively and efficiently applied.

## Introduction

The pursuit of asymmetric advantage has long been recognized as a critical factor in shaping the outcomes of strategic competition. While asymmetries are abundant, finding asymmetries that can produce significant and enduring advantages can be challenging. More than two thousand years ago, the Germanic chieftain Arminius leveraged asymmetric advantages in his choice of terrain and operational decision making to defeat a better armed and numerically superior Roman force (Goulding, 2000). More recently, Ukrainian forces have used a variety of asymmetric means to withstand a much larger Russian force, leveraging access to Western weapons, intelligence, and financial resources and having a more determined and defiant populus, to name a few. While the technology and character of war has profoundly changed over the centuries, the fundamentals of finding and leveraging asymmetric advantages remain relevant to the strategic discourse today.

The Third Offset Strategy, which came to prominence in the second term of the Obama Administration, was based largely around the idea of finding and exploiting asymmetric advantages. The “offset” in the strategy’s name refers to previous efforts in the 1950s and 1970s to offset the Soviet military’s quantitative advantage using asymmetric means. In the 1950s, the



United States exploited its asymmetric advantage in nuclear weapons, fielding a nuclear force capable of delivering a massive retaliatory strike sufficient to deter Soviet aggression. The 1970s offset strategy relied on an advantage in precision strike to offset a numerically larger Soviet conventional force. This asymmetry had the added advantage of luring the Soviets into a costly arms race, where they needed to either invest large sums of money in modernization to keep pace with the qualitative advantage of U.S. forces or build an even larger conventional force to overcome these advantages with mass. While the Third Offset arguably never fully congealed into a specific strategy, its basic premise was a continuation of the offsets pursued in the 1950s and 1970s. It sought to find and leverage asymmetric advantages in emerging technologies, such as artificial intelligence (AI), machine learning (ML), and autonomous vehicles, in combination with new organizational and operational constructs, to offset Chinese and Russian advances in conventional military capabilities (Gentile et al., 2021).

Current defense strategy seeks to exploit asymmetries in several ways. The 2022 National Defense Strategy says the military will use asymmetric approaches for deterrence, selectively share asymmetric capabilities with allies and partners, and leverage fundamental asymmetries in the American economy, culture, and system of government to “build enduring advantages” (DoD, 2022). Its predecessor document, the 2018 National Defense Strategy, cited the value of allies and partners as an “asymmetric advantage that no competitor or rival can match” (DoD, 2018). Moreover, the DoD’s recently published *Technology Vision for an Era of Competition* says the Defense Department will “maximize our asymmetric advantages by partnering with the larger innovation ecosystem, from industry to universities and to laboratories, allies and partners” (Office of the Undersecretary of Defense for Research and Engineering, 2022). These strategy documents and the historical examples cited highlight how asymmetries can be a powerful tool to counter or offset the technological, numerical, or operational superiority of an adversary.

This paper is a part of a more comprehensive forthcoming capstone study by Metrea Strategic Insights that develops an overarching theory of victory for the United States and its allies and partners that is rooted in asymmetries. As an excerpt from that study, this paper presents a framework for identifying and evaluating asymmetries to better assess the military advantage they can provide and their potential to endure over time. It begins with an exploration of asymmetries in nature to develop an understanding of asymmetries from first principles and why they matter. It uses historical examples of asymmetries in military competition to highlight the factors that affect how enduring they can be, the degree of leverage they can provide, and their potential to scale exponentially. Based on these examples, the paper presents a framework for evaluating and comparing asymmetries. It concludes by applying the framework to assess example areas of competition that are often touted as potential asymmetric advantages.

## Asymmetries in Nature

Symmetry can exist in many forms and to different degrees. An object can be symmetric top to bottom, left to right, or front to back. It can have translational symmetry, which means it appears the same if the observer moves from side to side. It can have rotational symmetry, which means the object appears the same if the observer rotates it about a central axis. A sphere is the most perfectly symmetric object in three dimensions, and we see many examples of sphere-like shapes throughout nature. The Earth, sun, moon, and many other celestial bodies in the macroscale universe appear spherical at a distance. As Frank Close writes in his book, *Lucifer’s Legacy: The Meaning of Asymmetry*, “The fact that the entire cosmos has a common feature implies that there is something deeply encoded in the laws of nature that makes it like this” (Close, 2000, p. 13). The general shape of celestial bodies is a result of the gravitational force that attracted clumps of matter together over millions of years. While the gravitational force diminishes in proportion to the inverse square of the distance between two objects, it acts uniformly in all directions between all objects regardless of the orientation of the objects that are interacting.



Asymmetry is the lack of symmetry, and like symmetry itself, it is a matter of degree. Something can be symmetric in one way but asymmetric in other ways. Examining celestial bodies in the macroscale universe more closely reveals that they are not perfect spheres and are in fact asymmetric in many respects. The moon is covered in irregularly distributed craters, with many more on the side of the moon facing away from Earth (Jones et al., 2022). The Earth itself has irregularly shaped land masses, mountains, valleys, and polar ice caps, and it bulges slightly around the equator. The sun is a highly dynamic system with asymmetric eruptions of energy in the form of coronal mass ejections. These irregularities are the result of more complex interactions in nature which involve other forces that, unlike gravity, do not act uniformly in all directions. The electromagnetic force, for example, attracts objects of opposite charge and repels objects of like charge, and this is the dominant force at the molecular level governing how atoms come together to form the gases, solids, liquids, and more complex structures around us. Time is itself asymmetric because it only progresses in one direction. As observed by Sir Arthur Stanley Eddington and others since, the asymmetry of time leads to many other forms of asymmetry and irreversible processes, such as the fact that heat spontaneously flows from hot to cold and not the reverse (Eddington, 1948). Without these fundamental asymmetries in nature, the universe as we know it would have never sprung into existence. As one scientist notes, “In physics, to be symmetrical is to be immune to possible changes” (Livio, 2012). Asymmetry creates the potential for change—the power to shape, affect, and evolve.

At increasing levels of complexity in the natural environment, asymmetries become more interesting and consequential, and in some instances, they lead to distinct advantages. Louis Pasteur, in a paper on asymmetry at the molecular level, remarked that, “Life as manifested to us is a function of the asymmetry of the Universe and of the consequences of this fact” (Salam, 1990). At the microbiological level, relatively simple single-celled organisms like bacteria evolved to have asymmetric shapes and growth patterns at different phases in their life cycles. When some single-celled organisms divide, for example, the two parts are not identical. This allows for a “selective advantage” to accrue over time. Research has shown that organisms can use asymmetric division to purge damage found in individual cells, such as misfolded proteins, from an overall population. As one journal article notes, “by biasing damage segregation into one cell upon division, a relatively damage-free daughter enjoys higher fitness at the expense of the aging cell, thereby increasing the overall damage tolerance of the population” (Kysela et al., 2013). The asymmetry of cell division proves to be an enduring advantage in nature in part because it only progresses in one direction—once a cell divides and the advantage accrues, it cannot be reversed. Moreover, it is a self-perpetuating advantage because cells that divide asymmetrically are more likely to survive and propagate.

More complex forms of life can exhibit more complex forms of asymmetric advantage. The family of fish known as Flatfish (or *Pleuronectidae*), which includes Flounder, Halibut, and Sole, is perhaps one of the most peculiarly asymmetric animals to have ever evolved, as shown in Figure 1. Charles Darwin commented on this type of fish in *The Origin of Species*, noting that the advantage of their “flattened and asymmetrical structure” is evident by the fact that several species of this family are “extremely common” in the wild (Darwin, 1872, p. 240). The flatfish begins life as a typical fish with bilateral (left-right) symmetry, but as it grows and matures one of its eyes migrates from one side of the head to the other side. As a result, these fish spend most of their adult lives swimming sideways with both eyes on the same side of their head (Skeptic’s Play, 2009). As Darwin notes, “the chief advantages thus gained seem to be protection from their enemies, and facility for feeding on the ground” (Darwin, 1972, p. 240).





**Figure 8. A Peacock Flounder with Asymmetric Eyes.**  
(© cherylvb / Adobe Stock).

The brain also evolved to be asymmetric in both its shape and function. In his book, *The Master and His Emissary: The Divided Brain and the Making of the Western World*, Iain McGilchrist explores how the shape and function of the human brain evolved to be highly asymmetric. Like other animals, the human brain is divided into hemispheres, but it “appears to have been twisted about its central axis,” with the left hemisphere wider towards the back and the right hemisphere wider towards the front (McGilchrist, 2012, p. 23). This asymmetry in shape also corresponds to an asymmetry in function that allows the brain to attend “to the world in two ways at once” (McGilchrist, 2012, p. 30). The left brain specializes in activities that require narrow and focused attention, whereas the right brain specializes in keeping track of the broader context of the environment and how one relates to that environment. McGilchrist goes on to explore how the asymmetric division of functions in the human brain manifests itself in society and culture (McGilchrist, 2012, p. 431).

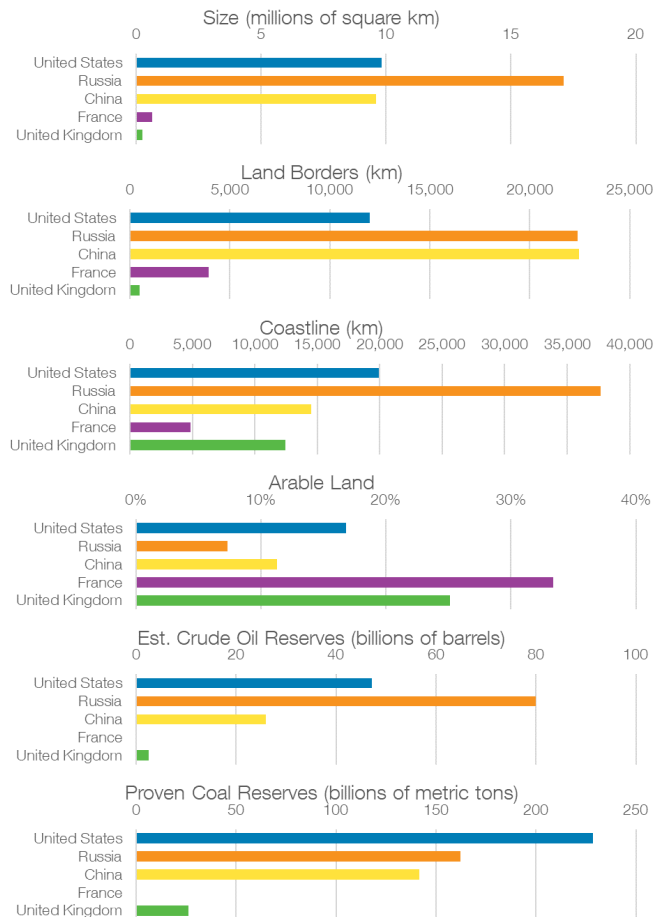
Another asymmetry in nature that can have far reaching effects in human interactions is geography. Macroscale forces acting over millions of years, such as the movement of tectonic plates, volcanic eruptions, and erosion, created the oceans and the land masses humans inhabit. The geography of the Earth, and the access to resources it conveys, is a fundamental asymmetry among nations. Differences in geography mean that no two nations are alike in the resources available to them or in the interdependencies they share with other nations for trade, diplomacy, and culture. Nations vary greatly in size, climate, water and mineral resources, arable land, access to natural trade routes (e.g., rivers and oceans), and proximity to rivals. History has shown that these geographic factors directly influence the character of a nation, its economy, and the security challenges it faces (Diamond, 1999, p, 25).

Figure 2 illustrates some of the geographic and natural resource differences among the permanent members of the United Nations Security Council, as an example (CIA, n.d.). While these five nations enjoy comparable diplomatic status and have historically been major powers in the world, fundamental differences in their geography influenced how they evolved over time and the strategic position they find themselves in today. For example, both Russia and China have extensive land borders, which requires significant resources for ground forces and border defenses. While Russia has extensive energy resources in the form of coal and crude oil reserves, it has relatively little arable land, leading it to become a major energy exporter and food importer.





China is relatively poor in energy and arable land (relative to its population and level of industrialization), and it is dependent on imports of both. The resource limitations imposed on China and Russia by nature incentivized each to adopt zero-sum, neomercantilist policies designed to make “asymmetric economic gains at the expense of competitors” (Ziegler & Menon, 2014). In contrast, the United States has friendly nations on its northern and southern land borders, large oceans protecting its coastlines to the east and west, and access to sufficient energy (mainly coal) and arable land. As a result, the United States has often had the luxury of choosing to engage with other nations when it is mutually beneficial and in accordance with its own values (Biden, 2021).



**Figure 9. Selected Geographic Factors for the UN Permanent Five**

As this discussion has shown, asymmetries play an important role in nature. They arise from the fundamental laws of the universe, and at the most basic level they are the result of imbalances and directionality in nature. Processes that are irreversible, from cellular division to volcanic eruptions, effectively “lock in” asymmetries and prevent nature from returning to an entropic path of decay toward absolute symmetry. Asymmetries that are advantageous in nature tend to be self-perpetuating, such as the selective advantage provided by the asymmetric division of functions within the brain. Moreover, it becomes evident from nature that asymmetries can work synergistically together over time to create an additive and, in some cases, exponentiating advantage. Asymmetries give us the ability to affect the environment around us, and to maximize those effects we must find ways to maximize the asymmetries that exist in relation to what we are attempting to affect.



## Asymmetries in Military Competition

Military competition is often asymmetric in each party's perception of their relative standing. If adversaries shared the same assessment of their relative standing, the rational course of action for the weaker side would be to regroup and find a new approach rather than commit itself to a competition in which it is unlikely to prevail (Farley, 2012). What is often found in practice is a David versus Goliath situation where the weaker side (or the side that perceives itself to be at a disadvantage) pursues an asymmetric strategy. Rather than trying to match its opponent plane for plane or tank for tank, it instead finds ways to compete that its opponent finds difficult to match or counter. Over time, however, the stronger opponent will attempt to rectify this by negating the asymmetry and regaining the advantage. For this reason, asymmetries in military competition are often "transient phenomena" (Krajewski, 2012). This section presents historical examples of asymmetric strategies at three different levels of effects: the New Look (or the First Offset) at the strategic level; precision strike (or the Second Offset) at the operational level; and the improvised explosive device (IED) at the tactical level. These examples provide insights into the factors that determine how much advantage an asymmetry is likely to provide and how enduring that advantage is likely to be.

### The New Look

World War II upended the world order in many ways. It left the British and French empires in demise, and it brought two of the most powerful militaries in the world, Japan and Germany, to their knees in unconditional surrender. This allowed the United States and the Soviet Union to quickly ascend as the world's two superpowers. Moreover, in the immediate aftermath of the war, the United States held a temporary monopoly on the most powerful weapon ever devised—the atomic bomb. The post-World War II period quickly became a struggle between the free nations of the West and the communist nations of the East, and the Korean peninsula became a focal point in this struggle. President Eisenhower took office in the middle of the Korean War and sought to quickly bring the conflict to a conclusion. During the presidential campaign of 1952, Eisenhower made clear that he viewed the Korean War as a grave error in U.S. foreign policy, saying "There is a Korean War—and we are fighting it—for the simplest of reasons: because free leadership failed to check and to turn back Communist ambition before it savagely attacked us." Later in the same speech he pledged to bring the war to "an early and honorable end" (Eisenhower, 1952).

As he worked to disentangle the U.S. military from the Korean War, Eisenhower recognized that the United States could not afford to match the Soviet Union in conventional forces or in prolonged proxy wars like Korea. The Soviet military maintained a much larger ground force, a necessity to defend its extensive and vulnerable land borders. Estimates at the time (which later proved to be misleading) suggested that the Soviets had 175 Army divisions that were "fully manned, fully armed, and combat-ready" (Bitzinger, 1989). This was roughly three times the conventional ground forces the United States and its allies possessed. Eisenhower believed the cost of matching the Soviets division for division would ultimately handicap the U.S. economy. Instead, he sought an alternative strategy that would "offset both the Soviet's advantage in conventional troops and their nascent nuclear arsenal" (Gentile et al., 2021, p. 9). This strategy—what became known as the New Look—was captured in National Security Council Paper 162 and later revised in NSC 162/2. In more recent years, this strategy has been referred to as the First Offset.

The New Look was based on the concept of massive retaliation—the idea that the United States would deter Soviet aggression at the strategic level by building a large and resilient nuclear arsenal that could survive a first strike and still deliver a devastating counterattack. It was rooted in the belief that "the only way to win the next world war is to prevent it" (Eisenhower, 1956). The strategy was particularly appealing at the time because it leveraged an asymmetric advantage the United States held in nuclear technology, and the cost of fielding and maintaining a nuclear



capability for massive retaliation was significantly less than matching the Soviet military division for division. Moreover, the Soviets could not easily defend against a nuclear attack because it would require a level of air and missile defense technology that was not yet within reach.

By the end of the 1950s, however, Eisenhower's New Look came under increasing criticism. The Army struggled throughout the 1950s to adapt to the new strategy, with a failed attempt to reorganize into "Pentomic Divisions" designed to operate on a nuclear battlefield. The strategy also became more symmetric over time as the Soviet Union reached rough parity with United States in its own nuclear forces. Massive retaliation was no longer a credible threat to deter lower-level Soviet aggression once it had a secured second-strike capability of its own. What was needed, opponents of the strategy argued, was a more flexible set of options for how the United States could respond to aggression—what became known as the Kennedy administration's Flexible Response strategy (Gentile et al., 2021, pp. 10–11).

### **Precision Strike**

By the 1970s, the buildup of Soviet nuclear forces had effectively eliminated the asymmetry the United States sought to exploit in the New Look. Nuclear parity made it possible to negotiate arms control treaties to limit the size of each nation's nuclear arsenal, and it created a stable deterrence posture (what became known as mutually assured destruction or MAD) that prevented a nuclear exchange. Under the first Strategic Arms Limitation Talks (SALT) treaty, the United States was limited to 1,054 Intercontinental Ballistic Missile (ICBM) silos and 710 Submarine Launched Ballistic Missiles (SLBM) launch tubes while the Soviet Union was limited to 1,618 ICBM silos and 950 SLBM launch tubes. The treaty did not limit bombers or the total number of warheads, and each side's nuclear forces remained more than sufficient to deliver a devastating second strike (Kimball, 2022). This rough symmetry at the nuclear level brought the focus of the U.S.-Soviet military competition back to conventional forces.

In the 1970s, the Soviet Union continued to have numerical superiority in its conventional forces. Moreover, the United States' elimination of the draft in 1973 made it more costly than ever before to field a force comparable in size to the Soviets (Comptroller General of the United States, 1978). Senior defense officials, particularly Defense Secretary Harold Brown, Undersecretary of Defense for Research and Engineering William Perry, and Director of Net Assessment Andrew Marshall, concluded that a new strategy was needed to offset the Soviet military's quantitative advantage and deter an armored assault across Europe—what later became known as the Second Offset (Gentile et al., 2021, pp. 12–13).

The idea behind the Second Offset was to shift the competition into an area where the United States would enjoy an asymmetric advantage: its ability to rapidly develop and operationalize innovative new technologies and operational concepts. As William Perry noted in congressional testimony, "Precision guided weapons . . . have the potential of revolutionizing warfare," and "greatly enhance our ability to deter war without having to compete tank for tank, missile for missile with the Soviet Union" (Gentile et al., 2021, p. 15). The strategy called for using precision guided weapons and advanced delivery systems (such as stealthy aircraft) in combination with innovative concepts of operation, such as Active Defense and AirLand Battle, to give the United States a qualitative advantage—a force multiplier that would allow a relatively smaller number of U.S. forces to defeat a much larger adversary.

The technologies and doctrine developed as part of the Second Offset were on full display in the 1991 Gulf War and later in the conflicts in Bosnia and Kosovo. These conflicts demonstrated the powerful effects that could be generated through the combined use of space systems, precision guided weapons, and stealthy aircraft, among the many other advanced weapon systems employed. The dramatic success of air power in these conflicts unwittingly exposed one of the weaknesses of precision strike—it is only an advantage if it is supported by precision



intelligence. For example, several months into the Kosovo air war, Air Force commanders worried that they were running out of good targets and that it was becoming increasingly difficult “to find and demolish the dispersed Yugoslav troops and equipment that remain in Kosovo without unintentionally striking civilians who are often mixed in with them” (Harris & Graham, 1999).

China, Russia, and other potential adversaries took note of the asymmetric advantage the United States held and adjusted their own strategy, doctrine, and investments accordingly. As former Director of National Intelligence Mike McConnell surmised, adversaries like China “concluded from the Desert Storm experience that their counterapproach had to be to challenge America’s control of the battle space by building capabilities to knock out our satellites and invading our cybernetworks” (Gardels, 2010). To counter the United States, China developed anti-access/area denial (A2/AD) capabilities, such as robust and integrated air defense networks and long-range ballistic missiles and cruise missiles, to keep U.S. forces at range. Both Russia and China developed a suite of anti-satellite (ASAT) capabilities, from direct-ascent ASAT weapons to satellite jammers and laser dazzlers, to disrupt American intelligence collection, navigation, and communications capabilities, making it harder to sense and coordinate actions (Johnson et al., 2022). And in parallel, both nations developed and fielded precision strike capabilities of their own, making the asymmetric advantage more symmetric.

### **Improvised Explosive Devices**

When the United States went into Afghanistan in 2001 and launched its invasion of Iraq in 2003, it enjoyed numerous advantages over the Taliban and Iraqi military. While China and Russia were working to undermine the Second Offset strategy, the asymmetric advantage of precision strike continued to work well in Iraq and Afghanistan. Neither adversary was able to mount significant resistance to the initial U.S. invasion force, and military commanders predicted a speedy victory in both conflicts.

However, an opponent does not need to be a major power or even a nation-state to find and exploit asymmetries in military conflict. Rather than quick victory, what ensued in the years and decades that followed was a roiling insurgency that found the United States engaged in irregular warfare in both conflicts. While the power disparity between the United States and the former government, tribal, and sectarian groups that resisted occupation was immense, the U.S. military struggled to adapt to this new form of warfare and suppress the insurgencies. A key weapon used by insurgents was the improvised explosive device (IED). As some scholars have noted, “the IED is a near perfect weapon system for balancing this power disparity” (Amoroso & Solis, 2019). The Government Accountability Office (GAO) found that by July 2008, “about 75 percent of casualties in combat operations in Iraq and Afghanistan were attributed to improvised explosive devices” (Sullivan, 2009).

The IED and the ever-present threat of IEDs had many impacts on the conflicts. It restricted freedom of activity for U.S. forces, making it more difficult and riskier to move supplies and personnel around the battlefield. It meant that there were no clear front lines in the conflict and no sanctuary. It undermined the ability of U.S. forces to provide security and basic support services to the population, making the United States look like a weak and unreliable partner. Perhaps most notably, IEDs had a disproportionate cost impact on the United States (Amoroso & Solis, 2019). These relatively inexpensive weapons forced the U.S. military to initiate a rapid acquisition program to field armored vehicles capable of protecting servicemembers from IEDS and to acquire a variety of other counter-IED technologies. From fiscal year 2006 through 2011, these efforts cost some \$58 billion in total and arguably shifted DoD’s acquisition attention away from longer term threats and modernization needs (Russell, 2012).



## A Framework for Evaluating Asymmetries

The examples presented of asymmetries in nature and in military competition highlight the key factors that should be considered when comparing the relative potential of asymmetries. This section organizes these factors into a framework of five questions that can be used to evaluate asymmetries. The factors these questions explore are: the immutability of an asymmetry's source, how difficult it is to restore symmetry by copying or countering, the level of effects at which an asymmetry is anchored and the spectrum of operations it supports, an asymmetry's ability to leverage other underlying asymmetries, and its ability to scale. The first two questions assess the ability of an asymmetric advantage to endure, the third and fourth questions assess the degree to which it maximizes leverage, and the fifth assesses its potential for exponentiality. Throughout this discussion, Blue refers to the party attempting to use an asymmetry (whether friend or foe) and Red refers to the party an asymmetry is being used against.

### 1. How immutable is the source of the asymmetry?

Perhaps the most important factor to determining whether an asymmetry is likely to endure is the immutability of its source. The most immutable sources of asymmetry arise directly from the laws that govern the physical universe. At higher levels of complexity in the environment and in the interactions within that environment, the factors that give rise to asymmetries can themselves be more variable and changing, making the asymmetry less likely to endure. The asymmetry of nuclear weapons relative to conventional weapons, for example, arises directly from immutable differences in the fundamental forces of nature that govern nuclear reactions (i.e., the strong nuclear force) and chemical reactions (i.e., the electromagnetic force). Geography is immutable in many ways but not entirely. While major shifts in geography tend to occur over many millennia, rising ocean levels threaten to reshape coastlines and reclaim low-lying islands within decades. Countries blessed with an abundance of fossil fuels may have an advantage today, but those resources can become depleted, and the relevance and utility of these resources can change over time. Technology is an often-touted source of asymmetry, and it played a key role in the Second Offset strategy (Metz & Johnson, 2001, p. 10). But technology is not immutable because it is always advancing and changing, and the rate at which technology advances is increasing as more people and organizations have access to the tools and resources needed to produce new technologies (Roser, 2023).

### 2. How difficult is it to copy or counter?

The endurance of an asymmetric advantage also depends in large part on how difficult it is for Red to revert the competition to symmetry by coping or countering the asymmetry. In the example of nuclear weapons in the New Look strategy, the Soviet Union recognized the asymmetric advantage the United States held in these weapons and worked to quickly build up its own nuclear forces. Even though the source of the asymmetry was the immutable laws of physics, it ceased to be asymmetric when the Soviets were able to create a comparable capability for themselves. While this undermined the asymmetry, it did not completely undermine the strategy because it led to a stable and roughly symmetric equilibrium between the two superpowers in the nuclear dimension of the competition. Financial, scientific, and political barriers may make an asymmetry difficult for Red to steal or independently develop, as has proven to be the case for most other nations seeking nuclear weapons. Even if Red can copy an asymmetry, it may not be able to generate symmetric effects because of other underlying asymmetries, such as geography. Asymmetries that rely on specific technical knowledge or the idea itself being kept secret to prevent it from being copied will only endure for as long as that secrecy can be maintained—or until Red makes the same discovery independently.

Red can also seek to restore symmetry by developing effective counters to an asymmetry. Counters can include protective measures an adversary takes that seek to limit the effects an asymmetry has, or they can involve active measures that attempt to disrupt or degrade its use.



Ideally for Blue, the counter will be exponentially more difficult than the asymmetry itself and require substantial resources that detract an adversary from building other military capabilities, making it more likely that the asymmetric advantage will endure. For example, the direct counter to the nuclear-armed ICBMs in the New Look strategy was a missile defense system that could intercept these missiles in flight—a much more challenging technical problem than the ICBMs themselves. An asymmetric advantage that is critically dependent on other supporting capabilities can also create vulnerabilities an adversary can use to counter the asymmetry. The U.S. military's failure to adequately protect its ISR and command and control systems, particularly in space, created a vulnerability for its precision strike asymmetric advantage that adversaries have sought to exploit (Harrison et al., 2021, pp. 1–2).

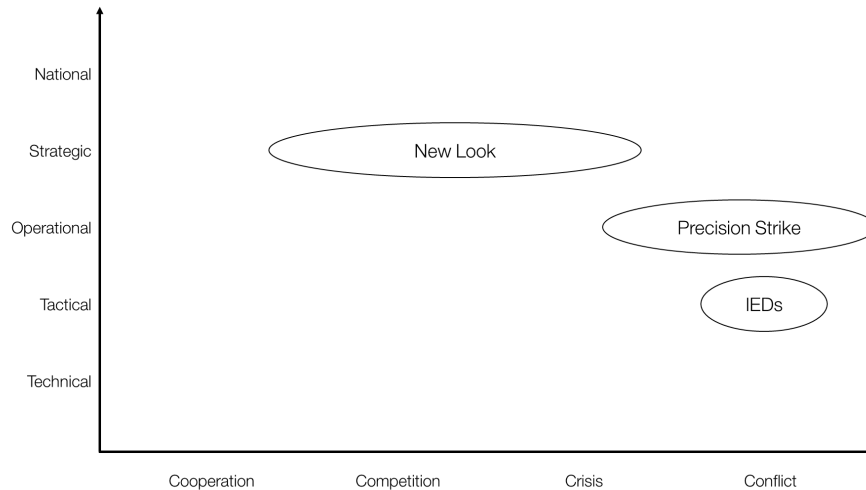
As Red works to counter an asymmetry, Blue can also take actions to counter Red's counters. For example, the IEDs used by insurgents in Iraq and Afghanistan led the U.S. military to field more heavily armored vehicles and to develop better technology to detect and disable IEDs before they could detonate. Insurgents countered these counters by developing a variety of fuses and detonators that were more difficult to defeat and using shaped charges to penetrate thicker plates of armor. As one writer has noted, the most effective counter to IEDs proved to be at a higher level of effect than the IED itself—a “change in relationship between U.S. troops and the local population made the greatest difference in overall security conditions, including with the IED” (Shell, 2017).

### **3. At what level of effect is the asymmetry anchored, and how applicable is it across the spectrum of operations?**

The degree of leverage an asymmetry provides depends in part on the level of effects at which it is anchored. The level of effects can be thought of as a continuum that begins at the scientific and technical levels and rises through the tactical, operational, strategic, and national levels. While some have noted that “the most common form of asymmetry resides at the operational level,” asymmetries can exist at many other levels as well (Metz & Johnson, 2001, p. 9). The New Look strategy of massive retaliation was anchored at the strategic level, while the asymmetry of IEDs used by insurgents in Iraq and Afghanistan was anchored at the tactical level. The higher the level at which an asymmetry is anchored, the more leverage it can provide.

A related dimension is how applicable an asymmetry is across the spectrum of conflict. The spectrum of conflict can be thought of as an orthogonal dimension to the level of effects that spans from cooperation through competition, crisis, and conflict, as shown in Figure 3. The more broadly applicable an asymmetry is across the spectrum, the more useful it becomes in practice. Asymmetries are particularly advantageous if they can deter a situation from escalating to crisis or conflict. Massive retaliation, for example, proved to be an effective deterrent at the strategic level across competition and crisis because it created a strong incentive for an adversary not to escalate. Precision strike, however, was not an effective deterrent at the operational level in competition and crisis. Even after its effectiveness in conflict was demonstrated in the 1991 Gulf War, the credible threat of U.S. military intervention with its precision strike advantage was not sufficient to compel belligerent forces in Bosnia or Kosovo to back down. If the overarching strategy is aimed at deterrence, Blue's focus should be on finding asymmetries that can be applied left of crisis and at higher levels of effect. Asymmetries that rely on secrecy or surprise for their effective employment, and therefore cannot be revealed in advance of their use, may only be applicable in crisis or conflict and have little (if any) effect in cooperation and competition.





**Figure 10. Example Diagram of Level of Effects Versus Spectrum of Operations**

#### **4. Does it leverage other underlying asymmetries?**

Asymmetries that build synergistically on other asymmetries can greatly enhance the leverage they provide and the complications they create for an adversary. For example, the precision strike asymmetry the United States sought to exploit in the Second Offset leveraged an immutable geographic asymmetry between the United States and the Soviet Union, namely that the Soviets had a much larger and precarious border to defend in the land, sea, and air domains. By some estimates, Soviet spending on air defenses exceeded that of the United States by a factor of ten (Lepingwell, 1989). U.S. advances in stealthy aircraft, such as the F-117A fighter and B-2 bomber, complicated an already difficult air defense challenge for the Soviet military and negated many of the investments they had already made. Layering asymmetries with other asymmetries can create an integrated advantage that is greater than the advantages each provides independent of one another. However, layered asymmetries that are interdependent on one another (meaning they cannot function separately) can introduce vulnerabilities Red may seek to exploit.

#### **5. How well does it scale?**

The ability of an asymmetry to scale is a critical factor in how much advantage it can provide. As an asymmetric advantage is exploited in larger numbers or to a higher degree, it may have diminishing or increasing returns on its effectiveness (i.e., non-linear scaling). There may also be thresholds beyond which its effectiveness abruptly changes (i.e., step functions). Nuclear weapons, for example, quickly reach a point of diminishing returns once a nation has enough weapons to deliver a secure and devastating second strike against all adversaries. Having more weapons beyond this point provides less and less incremental advantage. It is also important to understand how costs scale with an asymmetry—both the costs it imposes on Red and the costs of the asymmetry itself for Blue. Ideally for Blue, the costs imposed on Red will increase faster than the costs Blue incurs. It could be a linear relationship between the two (e.g., the costs imposed are X times more than the costs incurred), or it could be a more complex non-linear relationship. There may be a point at which the costs imposed and the costs incurred cross over one another, as in a square-cube law relationship, making an advantage become a disadvantage (or vice versa). These scaling dynamics are also connected to the ability of Red to counter the asymmetry, Blue's counter to Red's counter, Red's counter to Blue's counter of Red's initial counter, and so forth through some number of  $n$ -counter cycles. An asymmetry that grows stronger (i.e., becomes harder for Red) as  $n$  increases scales favorably for Blue, whereas one that gets weaker (i.e., becomes easier for Red) as  $n$  increases does not.



## Applying the Framework

In any discussion of historical examples, one enjoys the advantage of hindsight because the outcomes of these examples are already known—another example of the time asymmetry at work. In the case of this analysis, historical examples of asymmetries in military competition serve to highlight the strengths and weaknesses of each asymmetric approach and to construct a framework for comparing asymmetries. However, when it comes to applying this framework forward to evaluate future asymmetries, the time asymmetry is a distinct disadvantage because we know much less about the future than the past. But our understanding of the future, however crude and imprecise it may be, is immensely more valuable to the decisions that must be made today to prepare for whatever future awaits. This section applies the framework developed for evaluating asymmetries to areas that are often touted as potential asymmetric advantages. The aim is to assess the relative potential of these asymmetries and aid our understanding of how military resources can be most effectively applied today. As with any complex assessment, the five factors assessed in the framework do not simply combine into a single metric to determine the best asymmetry overall. Rather, they work together to provide a more comprehensive understanding of the conditions under which one asymmetry may be better than another.

## Ubiquitous ISR

The large-scale deployment of increasingly capable remote sensing satellites and highly proliferated terrestrial and airborne surveillance technology is creating an unprecedented level of transparency on Earth with a reach that transcends national borders. Satellites are particularly well-suited for broad area surveillance because of their altitude, freedom of overflight, and the regularity of orbits. A satellite in a sun synchronous orbit, for example, traverses from pole to pole as it orbits the Earth and, depending on its sensor suite, can sense across broad parts of the spectrum, from visible and infrared light to radio frequency signals. As more satellites are added to constellations, the revisit rate (i.e., the time between satellites passing over a given point on Earth) continues to go down and the amount of data collected continues to climb. Of course, the increase in remote sensing capabilities that are making ISR more ubiquitous is not limited to satellites—a variety of airborne and networked terrestrial-based sensors, ranging from drones to traffic cameras, are also increasing the ISR capabilities available. Remote sensing systems can use active sensing, such as synthetic aperture radar that can see through clouds and at night, or passive sensing, such as electro-optical imagery that relies on reflected sunlight or RF sensing that detects, geolocates, and characterizes stray radio emissions. A key enabler of ubiquitous ISR is the software that automates the processing of raw data into intelligence products and combines data from multiple space-based, airborne, and terrestrial sensors to create a near-real time view of the Earth and what is happening on it. This trend in ISR is also extending in the opposite direction, with more sensors pointed away from Earth at objects in space. The unprecedented level of insight into what is happening on and around the Earth has the potential to create an information asymmetry. The asymmetry is not that one side will have more information than the other; rather, the asymmetry is that a nation seeking to conceal activities within or beyond its own borders may no longer be able to do so.

- Immutability of the Source:

The underlying sources of ubiquitous ISR are both the physics involved in remote sensing and the software technology that enables automated processing of sensor data into actionable intelligence. The physics of remote sensing is immutable because it is based on the fact that electromagnetic signals (light, radio waves, etc.) naturally radiate outward in free space, making it possible to observe them from a distance. In contrast, the use of AI and ML software to increasingly automate the processing and exploitation of data is not immutable; it is rapidly changing and advancing.





- **Ability to Copy or Counter:**  
The capabilities that enable ubiquitous ISR are widely available (including commercial space remote sensing systems), and they are not difficult for a state or non-state actor to access. The benefit derived from copying this capability, however, can be asymmetric depending on who it is being used against (see Synergies below). There are many ways an adversary can counter ISR systems, such as camouflage, concealment, deception, blinding sensors, spoofing sensors, and disrupting the communications systems that support information dissemination. This naturally leads to what is likely a perpetuating hider-finder competition.
- **Level of Effects and Applicability Across the Spectrum:**  
Ubiquitous ISR is anchored at the tactical level because its primary effect is to enable a better near-real time understanding of what is happening in the battlespace and broader environment. It is applicable across the full spectrum of operations, from cooperation through conflict—although its use in conflict may be curtailed depending on how resilient the enabling capabilities are to attack.
- **Synergies with Other Asymmetries:**  
Ubiquitous ISR can act synergistically with other asymmetries that exist between different societies and forms of government. The advantage for Blue is much greater when used against a Red government that attempts to maintain tight controls on information within and beyond its borders. When used against an open and free society, however, ubiquitous ISR is less likely to reveal information that was not already known. Moreover, the very structure of free and open societies allows them to benefit from greater knowledge of themselves (e.g., more effective and transparent enforcement of laws).
- **Ability to Scale:**  
The effects of ubiquitous ISR increase in a linear fashion at first (e.g., doubling the number of sensors doubles the information collected), but at scale it produces diminishing returns because sensors begin to overlap with one another in time, space, or spectrum. The added value from each additional observation diminishes as more of the observations contain redundant information (e.g., multiple pictures of the same car in the same parking lot around the same time).

## Hypersonic Weapons

The development of hypersonic weapons has been a priority for the U.S. military for several years. The 2018 NDS specifically cited hypersonic weapons as one of the “technologies that ensure we will be able to fight and win the wars of the future.” By definition, hypersonic weapons fly more than five times the speed of sound—much faster than conventional cruise missiles—and they can be more maneuverable than ballistic missiles, making them more difficult to track and intercept. Part of the push to accelerate the development of these weapons is a perceived gap with Russia and China, which are reportedly more advanced in hypersonic weapons. As Michael White, principal director for hypersonics in the Office of the Undersecretary of Defense for Research and Engineering, noted in public comments, this gap in capability “presents a battlefield asymmetry and timescale that we simply cannot allow to stand” (Cronk, 2021). According to the Congressional Research Service, the DoD is investing in multiple hypersonic weapons development programs in parallel at a cost of \$4.7 billion in FY 2023 alone, which is up from \$3.8 billion in FY 2022 (Sayler, 2023).

- **Immutability of the Source:**  
The source of asymmetry in hypersonic weapons is technology—specifically the propulsion, flight control, and thermal management systems that enable controlled



flight at these speeds. As previously discussed, technology is a fleeting source of advantage because it is constantly changing and evolving.

- **Ability to Copy or Counter:**

As with any technology, hypersonic capabilities can be replicated by other nations, as the United States is currently attempting to do. However, the resources and technical expertise required can create significant barriers for many other nations. A key asymmetric aspect of hypersonic weapons is the fact that they are more difficult to counter than cruise missiles and ballistic missiles. While ballistic missile defense is commonly compared to hitting a bullet with a bullet, defending against hypersonic weapons is more like trying to hit a highly maneuverable bullet with a bullet.<sup>1</sup> The FY 2024 budget request projects it will take 17 years of development before the DoD can begin fielding a new hypersonic defense system (Missile Defense Agency, 2023, p. 613).

- **Level of Effects and Applicability Across the Spectrum:**

Hypersonic weapons provide a tactical-level capability. While they are designed for use in conflict, they can also be applicable in competition and crisis. The ability of hypersonic weapons to hold targets at risk that other weapons may not be as effective at striking can increase deterrence in competition and potentially deter escalation in conflict if their use is deemed credible.

- **Synergies with Other Asymmetries:**

Hypersonic weapons work synergistically with geography. Because of their range and speed, they provide a greater advantage for Blue when attempting to strike highly defended Red targets from a distance, as compared to cruise missiles and ballistic missiles. A cruise missile takes much longer to reach targets over long distances (hours of flight time versus minutes at hypersonic speeds), and a maneuverable hypersonic missile is harder to defend against than a traditional ballistic missile. For Russia and China, hypersonic weapons provide an ability to target U.S. and allied bases and capital assets (such as aircraft carriers) from the relative sanctuary of their mainland. For the United States, these weapons provide the ability to strike highly defended and time-sensitive targets deep within another nation from standoff range.

- **Ability to Scale:**

With current technology, hypersonic missiles are more expensive than their alternatives, making them more difficult to field at scale. The Congressional Budget Office estimates that “hypersonic missiles would cost roughly one-third more than ballistic missiles with maneuverable warheads that had the same range and accuracy and traveled at similar speeds” (Kramer et al., 2023). Moreover, CBO notes that hypersonic weapons are only needed for a relatively small number of potential targets that are well-defended and time-sensitive, meaning they would add a diminishing incremental advantage when fielded in larger quantities because there would be fewer targets that require them.

## **Commercial Innovation**

The DoD has made a deliberate effort in recent years to improve its ability to leverage commercial innovation. The 2022 National Defense Strategy explicitly states that the DoD “will be a fast-follower where market forces are driving commercialization of military-relevant capabilities,” and that the DoD will increase collaboration with the private sector to leverage “its technological advancements and entrepreneurial spirit to enable new capabilities” (DoD, 2022, pp. 19–20). The

---

<sup>1</sup> For more on the challenges of defending against hypersonic missiles, see Karako and Dahlgren’s (2022) *Complex Air Defense: Countering the Hypersonic Missile Threat*.



democratization of technology means that more people and companies in the private sector have access to the design tools and other enabling technologies that make rapid innovation possible. According to Organization for Economic Cooperation and Development data, U.S. government and business R&D spending were roughly equal in 1981, but by 2020 R&D spending by U.S. businesses had grown to be 3.3 times that of the U.S. government.<sup>2</sup> The asymmetry for the United States is not a set of specific commercial technologies with military applications. Rather, the asymmetry is the free market economic system and access to capital that enables commercial companies to produce innovative technology. The U.S. commercial sector is widely considered more vibrant and innovative than that of its competitors, namely Russia and China, and it has a deeply rooted culture of entrepreneurship that encourages risk-taking and innovation (Hill et al., 2023).

- **Immutability of the Source:**

The source of commercial innovation is the economic, social, and cultural systems that enable it. While the economic system of a nation can change, these changes usually occur over many decades, and a nation's social and culture systems change even more gradually. The United States and many of its allies and partners have had (largely) free market economic systems and open societies for many decades, if not centuries, and this is unlikely to change in the foreseeable future.

- **Ability to Copy or Counter:**

The technology and capabilities that result from Blue commercial innovation can be copied by Red, as is evident by the extensive efforts both China and Russia have made to steal U.S. commercial technology (Editorial Board, 2022). Red could also attempt to counter each technology as it emerges. But this puts Red at a perpetual disadvantage—it will always be attempting to catch up to the latest Blue commercial innovations. The most effective long-term counter would be for Red to develop a commercial innovation base of its own that is comparable to Blue's.<sup>3</sup> This may not be a viable option for countries like China, Russia, and other authoritarian regimes because the conditions necessary for a vibrant commercial sector—namely an open society and free market economic system—would erode the foundation upon which their regimes are based. In other words, they would have to become more like the United States to counter the asymmetry commercial innovation provides.

- **Level of Effects and Applicability Across the Spectrum:**

Commercial innovation stems from the fundamental economic and governance system of a nation, anchoring it at the national level. It is most applicable in cooperation and competition, but commercial innovation can play an important role in crisis and conflict by augmenting military capabilities and enabling the ability to scale production of key items. In this situation, the relationship between the commercial sector and the military is likely to be substantially different than it is in peacetime.<sup>4</sup>

- **Synergies with Other Asymmetries:**

As previously discussed, commercial innovation is more advantageous when it is combined with an asymmetry between Blue and Red's forms of government and types of society (Blue's being free and open and Red's being authoritarian and closed). It

---

<sup>2</sup> Author's analysis of OECD data: <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm>

<sup>3</sup> Russia attempted to create a Silicon Valley of its own, known as the Skolkovo Innovation Center, beginning in 2010 under former Russian President Dmitry Medvedev. But after more than a decade of operating under an authoritarian regime and with multiple incidents of corruption, the effort has "failed to produce a single unicorn [or] even a company that has become a household name" (Hlebanov, 2022).

<sup>4</sup> For example, in World War II the U.S. government took unprecedented steps to control the commercial sector by rationing materials and fuel and by turning automotive factories into aircraft factories (Automobile Manufacturers Association, 1950).



also pairs well with a preexisting economic asymmetry where Blue has a much larger economy than Red.

- **Ability to Scale:**

When the government scales its use of commercial innovation, either by using commercial approaches more in existing areas or expanding the use of commercial approaches into new areas, it is leveraging a much larger private investment, creating a multiplier effect for every dollar the DoD spends. Moreover, the effect becomes exponential when considering the speed at which innovation occurs in the commercial sector, where the time between new generations of technology is often measured in months rather than decades and the rate of change is ever increasing.

## **Final Thoughts**

Asymmetries have historically been a powerful source of military advantage, but as this paper has shown, not all asymmetries are created equal. Asymmetries vary significantly in their ability to endure, the degree to which they maximize leverage, and their potential to scale effects exponentially. The framework presented in this paper is intended to serve as a basis by which current and future asymmetries can be evaluated and compared. The three example asymmetries analyzed (ubiquitous ISR, hypersonic weapons, and commercial innovation) serve to demonstrate how the framework can be applied across a variety of areas and are not intended to be a comprehensive listing of potential asymmetries. The goal of the framework is to help decision makers identify asymmetries that have the greatest potential.

The framework also highlights some of the key weaknesses in previous offset strategies—weaknesses that the U.S. military risks repeating in its current attempts to identify a new offset. The First Offset was subject to being copied, which allowed the Soviets to restore a level of symmetry in the competition. The Second Offset, in addition to being readily countered and copied, was based on a mercurial source—technology—that provided only a fleeting advantage. Moreover, the Third Offset strategy called on a laundry list of potential technologies, such as AI and autonomous systems, that are not likely to produce an enduring advantage for similar reasons as the first two offsets.

The 2022 National Defense Strategy rightly specifies that the asymmetric advantages it wants to pursue are those that will endure. This corresponds to asymmetries that are based on relatively more immutable sources and that are relatively more difficult to copy or counter. The strategy also focuses on the concept of integrated deterrence—shifting the competition into areas where the United States can leverage all areas of national power and influence in a coordinated manner. This calls for asymmetries that anchor at the national level of effects, that are applicable in the cooperation and competition phases of operations (i.e., left of crisis), and that leverage other underlying asymmetries beyond just the military aspect of the competition. The ultimate objective is not to fight and win wars—the goal is to win without fighting. Finding and pursuing the right asymmetric advantages is the key to establishing a stable and enduring deterrence posture for the future.

## **Metrea Strategic Insights**

Metrea Strategic Insights (MSI) specializes in pathfinding studies that look beyond the typical five-year planning horizon at long-term trends, threats, opportunities, discontinuities, and asymmetries in national security. MSI is led by Managing Director Todd Harrison and is supported by a team of experts with a variety of experience in government, industry, think tanks, and academic institutions. MSI does not take institutional positions, and any views expressed in this publication are solely those of the author(s).

MSI's parent company, Metrea, provides effects-as-a-service to national security partners in four domains and over a dozen mission-centric solution areas, including airborne ISR, aerial refueling, electronic warfare,



communications, space-based ISR, and advanced simulation. Metrea leverages commercial business models to unleash innovation cycles that anticipate emerging threats. Metrea is headquartered in Washington, DC with facilities across the United States, the United Kingdom, and the EU.

## References

- Amoroso, P., & Solis, M. (2019, May28). Improvised explosive devices, a near perfect asymmetric weapon system of necessity rather than a weapon of choice. *Small Wars Journal*.  
<https://smallwarsjournal.com/jrnl/art/improvised-explosive-devices-near-perfect-asymmetric-weapon-system-necessity-rather-weapon>
- Automobile Manufacturers Association. (1950). *Freedom's arsenal: The story of the automotive council for war production*. The Association.
- Biden, J. (2021, February 4). *Remarks by President Biden on America's place in the world*. The White House.  
<https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/02/04/remarks-by-president-biden-on-americas-place-in-the-world/>
- Bitzinger, R. A. (1989). *Assessing the conventional balance in Europe, 1945–1975*. RAND.  
<https://www.rand.org/content/dam/rand/pubs/notes/2007/N2859.pdf>
- CIA. (n.d.). *The world factbook*. Retrieved March 23, 2023, from <https://www.cia.gov/the-world-factbook/>
- Close, F. (2000). *Lucifer's legacy: The meaning of asymmetry*. Oxford University Press.
- Comptroller General of the United States. (1978). *Report to the Congress: Additional cost of the all-volunteer force*. U.S. Government Accountability Office. <https://www.gao.gov/assets/fpcd-78-11.pdf>
- Cronk, T. M. (2021, May 3). Defense official says hypersonics are vital to modernization strategy, battlefield dominance. *U.S. Department of Defense News*. <https://www.defense.gov/News/News-Stories/Article/Article/2593029/defense-official-says-hypersonics-are-vital-to-modernization-strategy-battlefie/>
- Darwin, C. (1872). *The origin of species*. P F Collier & Son.  
<https://rauterberg.employee.id.tue.nl/lecturenotes/DDM110%20CAS/Darwin-1859%20Origin%20of%20Species.pdf>
- Diamond, J. (1999). *Guns, germs, and steel: The fates of human societies*. W.W. Norton & Co.
- DoD. (2018). *2018 national defense strategy*. <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>
- DoD. (2022). *2022 national defense strategy of the United States of America*.  
<https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.PDF>
- Eddington, A. S. (1948). *The nature of the physical world*. Cambridge University Press.  
<https://henry.pha.jhu.edu/Eddington.2008.pdf>
- Editorial Board. (2022, April 18). America is struggling to counter China's intellectual property theft. *Financial Times*. <https://www.ft.com/content/1d13ab71-bffd-4d63-a0bf-9e9bdfc33c39>
- Eisenhower, D. D. (1952, October 21). *Text of the address by Dwight D. Eisenhower, republican nominee for president, delivered at Detroit, Michigan*.  
<https://www.eisenhowerlibrary.gov/sites/default/files/research/online-documents/korean-war/i-shall-go-to-korea-1952-10-24.pdf>
- Eisenhower, D. (1956, October 17). *Address at a rally in the civic auditorium*. Seattle, WA.  
<https://www.eisenhowerlibrary.gov/eisenhowers/quotes#War>
- Farley, R. (2012, November 6). Asymmetry. *Information Dissemination*.  
<http://www.informationdissemination.net/2012/11/asymmetry.html>
- Gardels, N. (2010, February 5). China is aiming at America's soft underbelly: The internet. *The Christian Science Monitor*. <https://www.csmonitor.com/Commentary/Global-Viewpoint/2010/0205/China-is-aiming-at-America-s-soft-underbelly-the-Internet>
- Gentile, G., Shurkin, M., Evans, A. T., Gris , M., Hvizda, M., & Jensen, R. (2021). *A history of the third offset, 2014–2018*, 7–40. RAND Corporation. [https://www.rand.org/pubs/research\\_reports/RRA454-1.html](https://www.rand.org/pubs/research_reports/RRA454-1.html)
- Goulding, V. J., Jr. (2000). Back to the future with asymmetric warfare. *Parameters*, 30(4).  
<https://press.armywarcollege.edu/cgi/viewcontent.cgi?article=2005&context=parameters>
- Harris, J. F., & Graham, B. (1999, June 3). Clinton is reassessing sufficiency of air war. *Washington Post*.  
<https://www.washingtonpost.com/archive/politics/1999/06/03/clinton-is-reassessing-sufficiency-of-air-war/eaf709c3-89fa-4664-8dba-8d2e5b018fb2/>
- Harrison, T., Johnson, K., & Young, M. (2021). *Defense against the dark arts in space: Protecting space systems from counterspace weapons*. Center for Strategic and International Studies. [https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/210225\\_Harrison\\_Defense\\_Space.pdf](https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/210225_Harrison_Defense_Space.pdf)



- Hill, S., Ionescu-Somers, A., Coduras, A., Guerrero, M., Menipaz, E. E., Boutaleb, F., Zbierowski, P., Schött, T., Sahasranamam, S., & Shay, J. (2023). 2022/3 global report: Adapting to a “new normal”. *Global Entrepreneurship Monitor*. <https://www.gemconsortium.org/reports/latest-global-report>
- Hlebanov, S. (2022, April 22). Skolkovo: The story of Russia’s failed attempt to build its own Silicon Valley. *The Business of Business*. <https://www.businessofbusiness.com/articles/skolkovo-russias-failed-silicon-valley-tech-putin/>
- Johnson, K., Harrison, T., Young, M., Wood, N., & Goessler, A. (2022). *Space threat assessment 2022*. Center for Strategic and International Studies. <https://aerospace.csis.org/space-threat-assessment-2022/>
- Jones, M. J., Evans, A. J., Johnson, B. C., Weller, M. B., Andrews-Hana, J. C., Tikoo, S. M., & Keane, J. T. (2022). A south pole–Aitken impact origin of the lunar compositional asymmetry. *Science Advances*, 8(14). <https://doi.org/10.1126/sciadv.abm8475>
- Karako, T., & Dahlgren, M. (2022, February 7). *Complex air defense: Countering the hypersonic missile threat*. Center for Strategic and International Studies. <https://www.csis.org/analysis/complex-air-defense-countering-hypersonic-missile-threat>
- Kimball, D. (2022). U.S.-Russian nuclear arms control agreements at a glance. *Arms Control Association*. <https://www.armscontrol.org/factsheets/USRussiaNuclearAgreements>
- Krajewski, P. (2012, November 15). Symmetric warfare—The return to symmetry. *CIMSEC*. <https://cimsec.org/symmetric-warfare-back-to-symmetry/>
- Kramer, C., Mosher, D., & Keating, E. G. (2023, January). *U.S. hypersonic weapons and alternatives*. Congressional Budget Office. <https://www.cbo.gov/publication/58924#:~:text=CBO%20estimates%20that%20hypersonic%20missiles,for%20them%20is%20well%20developed>
- Kysela, D. T., Brown, P. J. B., Huang, K. C., & Brun, Y. V. (2013). Biological consequences and advantages of asymmetric bacterial growth. *Annual Review of Microbiology*, 67, 417–435. <https://doi.org/10.1146/annurev-micro-092412-155622>
- Lepingwell, J. W. R. (1989). Soviet strategic air defense and the stealth challenge. *International Security*, 14(2), 64–100. <https://doi.org/10.2307/2538855>
- Livio, M. (2012). Why symmetry matters. *Nature*, 490, 472–473. <https://doi.org/10.1038/490472a>
- McGilchrist, I. (2012). *The master and his emissary: The divided brain and the making of the western world*. Yale University Press.
- Metz, S., & Johnson II, D. V. (2001). *Asymmetry and U.S. military strategy: Definition, background, and strategic concepts*. U.S. Army War College. <https://apps.dtic.mil/sti/pdfs/ADA387381.pdf>
- Missile Defense Agency (2023, March). Department of Defense fiscal year (FY) 2024 budget estimates. *Defense-Wide Justification Book 2a*. [https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2024/budget\\_justification/pdfs/03\\_RDTandE/RDTEVol2MDARDTEPB24JustificationBook.pdf](https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2024/budget_justification/pdfs/03_RDTandE/RDTEVol2MDARDTEPB24JustificationBook.pdf)
- Office of the Undersecretary of Defense for Research and Engineering. (2022). *Technology vision for an era of competition*. Department of Defense. [https://www.cto.mil/wp-content/uploads/2022/02/usdre\\_strategic\\_vision\\_critical\\_tech\\_areas.pdf](https://www.cto.mil/wp-content/uploads/2022/02/usdre_strategic_vision_critical_tech_areas.pdf)
- Roser, M. (2023, February 22). Technology over the long run: Zoom out to see how dramatically the world can change within a lifetime. *Our World in Data*. <https://ourworldindata.org/technology-long-run>
- Russell, C. B. (2012, August 1). *Counter-improvised explosive devices: Multiple DoD organizations are developing numerous initiatives*. U.S. Government Accountability Office. <https://www.gao.gov/assets/gao-12-861r.pdf>
- Salam, A. (1990). The role of chirality in the origin of life. *International Centre for Theoretical Physics*. <https://inis.iaea.org/collection/NCLCollectionStore/Public/22/052/22052504.pdf>
- Sayler, K. M. (2023, February 13). *Hypersonic weapons: Background and issues for Congress*. Congressional Research Services. <https://sgp.fas.org/crs/weapons/R45811.pdf>
- Shell, J. (2017, May 1). How the IED won: Dispelling the myth of tactical success and innovation. *War on the Rocks*. <https://warontherocks.com/2017/05/how-the-ied-won-dispelling-the-myth-of-tactical-success-and-innovation/>
- Skeptic’s Play. (2009, February 12). *Darwin’s flatfish flounder*. <https://skepticsplay.blogspot.com/2009/02/darwins-flatfish-flounder.html>
- Sullivan, M. J. (2009, October 8). *Testimony before the house armed services committee, defense acquisition reform panel: Rapid acquisition of MRAP vehicles*. U.S. Government Accountability Office. <https://www.gao.gov/assets/gao-10-155t.pdf>
- Ziegler, C. E., & Menon, R. (2014). Neomercantilism and great-power energy competition in Central Asia and the Caspian. *Strategic Studies Quarterly*, 8(2), 17. <https://www.jstor.org/stable/26270802>



## PANEL 21. CALCULATING RETURN ON INVESTMENT

Thursday, May 11, 2023

2:15 p.m. –  
3:30 p.m.

**Chair: John Terence Blake**, VADM USN (Ret.), Professor of the Practice,  
Conrad Program Chair, Naval Postgraduate School

***Is it Ready? Quantifying the Maturity of Emerging Technologies***

Charles Rea, George Washington University  
John Kamp, George Washington University

***Management and Business Knowledge Representation for Decision Making:  
Applying AI Machine Learning, Data Science, and Advanced Quantitative  
Decision Analytics for Making Better-Informed Decisions***

Johnathan Mun, Naval Postgraduate School

***You Can't Wait for ROI to Justify Model Based Design and Analysis for  
Cyber Physical Systems' Embedded Computing Resources***

Alfred Schenker, Carnegie Mellon University  
Alfred Schenker, Carnegie Mellon University

**John Terence Blake, VADM USN (Ret.)**—served more than 37 years in the United States Navy before retiring in February 2013. He most recently served as the Deputy Chief of Naval Operations for Integration of Capabilities and Resources (OPNAV N8). As the Navy's Chief Financial Officer, he was charged with planning, programming and executing the Navy's Budget. He also served in numerous positions in the Pentagon including Deputy Assistant Secretary of the Navy (Financial Management & Comptroller) and Deputy Director Resources and Acquisition (Joint Staff J8). His sea commands included a Destroyer, an AEGIS Cruiser, and a Carrier Strike Group. Since retiring from the Navy, VADM Blake has served as an independent consultant. He graduated from the U.S. Naval Academy with a Bachelor of Science degree in Political Science, a Master of Science degree in Finance from the Naval Postgraduate School and a Master of Science degree in National Security from the National War College. Additionally, he completed the Seminar XXI program in International Relations from MIT.



# Is it Ready? Quantifying the Maturity of Emerging Technologies

**Charles Rea**—received a Doctor of Engineering degree in engineering management from the George Washington University and a Master of Science degree in electrical engineering and Bachelor of Science degree in electrical engineering and computer science from West Virginia University. Rea is a NAVAIR Associate Fellow and member of various professional organizations. [Charles.A.Rea.CIV@us.navy.mil]

**John Kamp**—received a Doctor of Engineering degree in engineering management from the George Washington University, a Master of Engineering degree in nuclear engineering from Iowa State University, and a Bachelor of Arts degree in mathematics and French from the University of Nebraska-Lincoln. Kamp joined the George Washington University in 2019 and currently teaches graduate courses and advises doctoral candidates in the School of Engineering and Applied Sciences. He is a retired naval submarine officer with extensive experience in research and development and program management. His research interests include engineering management, maritime systems, and acquisition system research. Kamp is a Fellow in the Royal Institution of Naval Architects and a member of several professional associations. [jckamp2018@gwu.edu]

## Abstract

The Department of Defense uses advanced technology to provide U.S. weapons systems superior operational capabilities. Technology Readiness Assessments establish the technological maturity level of emergent technologies. However, these assessments often rely upon subjective evaluations that depend upon measures indirectly associated with the actual readiness of a technology for use in a specific end-use application. The challenge of measuring the readiness of an emerging advanced technology for use in a new system remains subjective and a source of early program cost and schedule risk.

Prior bibliometric-based methods are sensitive to the search logic, keywords, and the specific corpus used. Visualization tools and larger datasets provide insights into the overall body of work and identify new patterns and associations. However, such methods have not been validated against independent assessments of actual maturity.

This paper presents novel methods, strategies, and results of using publicly available publication data to identify when specific technologies were mature enough to be used in programs approaching Milestone B. The method is calibrated using declarations from authoritative sources such as Selected Acquisition Reports and correlated against independent assessments from the Government Accountability Office.

**Results statements:** The method is predictive for the analyzed technologies and is shown to be appropriate for use in pre-Milestone B activities such as source selection and Milestone decision support.

**Keywords:** Acquisition, bibliometrics, technology maturity levels

## Introduction

Independent technical risk assessments (ITRAs) are required by law and require either identification of critical technologies and manufacturing processes that need to be matured prior to program initiation (Milestone A) have not been successfully demonstrated in a relevant environment prior to start of engineering development (Milestone B; 10 U.S.C. § 4272, 2016). Their content is codified by regulation and guidance (Under Secretary of Defense for Research and Engineering [USD(R&E)], 2020b) and is intended to provide (as named) an independent assessment of technical risk.

Despite a mandated ITRA process, and mandatory demonstration of all technologies entering the engineering and manufacturing phase of acquisition be mature, the U.S.





Department of Defense (DoD), in 2021, accumulated over \$615.4 Billion (52%) in total cost growth since program start while simultaneously slipping schedule by 35% with an average delay of over 2 years (Oakley, 2021). Some of the blame for the cost and schedule growth can be firmly placed on lack of consistent knowledge-based acquisition practices, specifically on maturing critical technologies and conducting appropriate design reviews prior to starting product development (Sager, 2021).

This paper summarizes a novel way to judge a technology's technical maturity based upon simple measurements of publication volumes. The results were calibrated to independent maturity assessments.

## Literature Review

### Technology Readiness

The DoD provides specific guidance on the technical risk assessment process that requires *subjective* evaluation of achievement of specific criteria (USD[R&E], 2020a). This method makes sense when the evaluators (experts) are familiar with the technology or are active in the technology development. However, for emergent or rapidly changing technologies, the assessment may be biased or incomplete and not capture the actual technical risk associated with trying to apply an emergent technology to a new use.

There is a subtle difference between technical risk and technology readiness. Technical risk is defined by NASA as "... the risk ... affecting the level of *performance necessary to meet the stakeholder expectations and technical requirements*...." (NASA, 2022). Technology readiness characterizes whether a system (product) *performs as intended* (Persons & Sullivan, 2016). In this paper, the use of the terms "independent technical risk assessment," "technology risk assessment," and "technology readiness assessment" are treated as equivalent, consistent with current usage in the DoD.

There have been several methods developed to provide a simple answer to the program manager's question, "Is it ready for ...?" Technology Readiness Levels (TRLs) are a common example. They are an ordinal scale, placing basic research discoveries in the lower levels (1 and 2), and in-use systems at the highest level (TRL 9; Mankins, 2009). They were created to help characterize the relative readiness of a component or system for use in a particular application (Olechowski et al., 2015). Azizian et al. (2009) noted that measuring "... *technology and system maturity is a multi-dimensional process that cannot be performed comprehensively by a one-dimensional metric*...." The point is that TRLs are by themselves insufficient to support technical readiness decisions. For example, technical maturity is defined as achieving TRL 6, when a model or prototype is demonstrated in a relevant environment (Persons et al., 2020). The problem is that TRLs are ordinal and are assigned based upon a subjective decision as to whether specific attributes related to a given TRL level are satisfied.

Bearden (1999) showed how for a complex system,<sup>1</sup> insertion of various technologies, although well understood at the component level, affect system cost and schedule due to unrecognized or underappreciated interdependencies. This is similar to the concept of architectural technical debt (Soliman et al., 2021).

There are other examples of how to measure technical maturity. Bailey et al. (2014) used a ratio of immature to total critical technologies<sup>2</sup> as a measure of technology maturity. Using attributes from a standards-based definition of software quality (ISO, 2001), Azizian et al.

---

<sup>1</sup> Bearden's work was specific for small satellites.

<sup>2</sup> Bailey et al. (2014) used the 2012 GAO weapon system assessment report for the count of immature and total critical technologies (Sullivan, 2012) to build their maturity estimates.



(2011) found that engineering activities related to improving system quality were strongly correlated with technology readiness and program performance.

Scant literature addresses the time to progress between TRLs. El-Khoury and Kenley (2014) and Peisen and Schulz (1999) developed independent methods to forecast technology transition times between TRLs using existing program datasets. Reinhart and Schindler (2010) added a velocity measure to their estimation model. Ramirez-Marquez and Sauser (2009) developed System Readiness Levels to address the shortcomings of TRLs in addressing integration complexities, evidence that TRLs are insufficient to assess system maturity. Bailey et al. (2014) found TRAs have a positive return on investment for Major Defense Acquisition Programs. Peters et al. (2017) derived the confidence intervals for technology assessments and their effect on program execution. Olechowski et al. (2020) identified system complexity, planning and review, and validity of assessment as continuing challenges.

The Government Accountability Office (GAO) takes a different approach—they identified key knowledge practices and associated them with three key decision points—to invest in development, to demonstrate and test prototypes, and to proceed to production (Dodaro, 2021). They also created a technology readiness assessments (TRA) guide describing best practices to determine a technology’s readiness (Persons et al., 2020). The GAO defines technical maturity when a system prototype or operational system is demonstrated in a relevant environment.<sup>3</sup>

### Quantitative Technology Readiness Indicators

Commercial products typically will consider *commercialization potential*, which includes assessments of the market,<sup>4</sup> regulatory, legal, and intellectual property assessments, in addition to technical maturity (Mazurkiewicz et al., 2015). Radpour et al. (2021) found market penetration models based on *subjective* technology maturity estimates and market survey could be unreliable.

Scientometrics and bibliometrics are quantitative methods associated with publications and are commonly used to identify important research, using tools such as citation clustering, network analysis, and visualization (Chen, 2006). Early bibliometric methods used frequency charts, as shown in Figure 1.

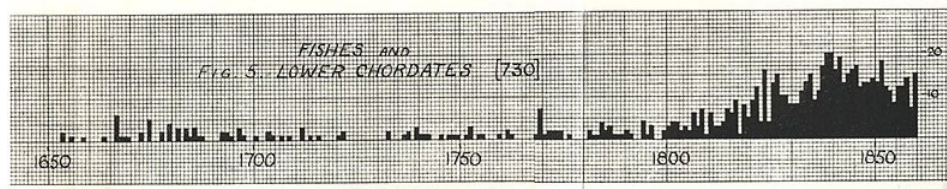


Figure 11. Example Publication Count Over Time (Public Domain)

Figure 1 shows a growth in publications between 1825 and 1850. This indicates a surge of research activity in this domain and is related to the concept of a *research front*.<sup>5</sup>

Wiesner and Ladyman (2021) described complex systems as having 10 properties, with four as conditions for complexity,<sup>6</sup> and six are the results<sup>7</sup> of those first four. Modern systems fit this complex system description. Following Wiesner and Ladyman, we consider an electric vehicle as satisfying their definition of a complex system, as they have extensive sensing for

<sup>3</sup> This is a common definition of TRL 6.

<sup>4</sup> This may include market and economic assessments.

<sup>5</sup> This is a qualitative or quantitative visualization of a research field’s state-of-the-art thinking.

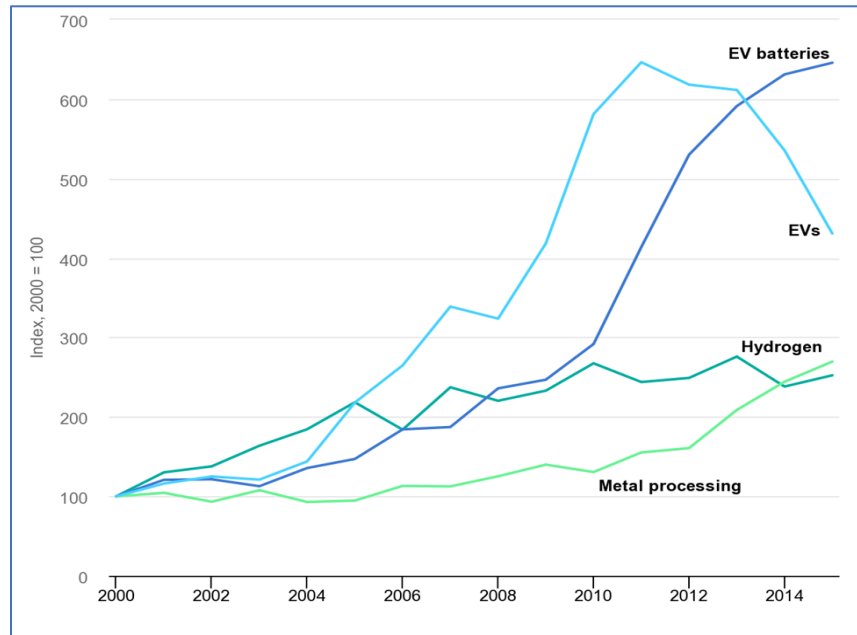
<sup>6</sup> Specifically *numerosity, disorder and diversity, feedback, and non-equilibrium*.

<sup>7</sup> Wiesner and Ladyman (2021) argue a system is complex if it has one or more of the following properties: spontaneous order and self-organization, nonlinearity, robustness, nested structure and modularity, history and memory, and adaptive behavior.



electrical system function (numerosity), uncontrolled and unpredicted interactions between the vehicle and world (disorder and diversity), and dynamic vehicle response while driving (feedback and non-equilibrium).

Following Radpour et al. (2021), we assert that electric vehicles are of a system maturity allowing market introduction and early market acceptance. According to Bloomberg, electric vehicles are about 4% of the U.S. automotive market (Stock, 2022). A common indicator of technology maturity is patent issuance. Figure 2 shows electric vehicle patent activity.



▪ **Figure 12. Electric Vehicle Related Patent Activity (IEA, 2022)**

One could argue from a qualitative perspective that the vehicle technology is mature on or after 2010, as patent issuances reduced. However, an electric vehicle is a system; what is to be made of the related patent activity? Is it correct to say that battery technology continued to mature and hydrogen power was behind? More importantly, these qualitative assessments do not answer when the various technologies were mature.

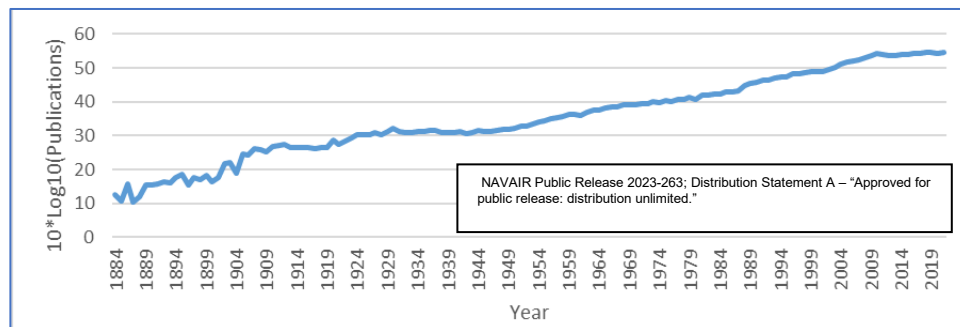
### Define Terms

This will compare a topic's research volume to that of all research conducted within online repositories. This is introduced as a relative research volume. Technologies for investigation are identified through declaration of technical maturity from authoritative government sources for Department of Defense Major Acquisition Programs. Investigation into the features of RRV, and changes in RRV (known in this praxis as  $\Delta RRV$ ), will be conducted in the Method chapter and proven through statistical testing in the Results chapter to find an objective measure of maturity that may be continuously updated and evaluated throughout a program's lifecycle. Any changes to the measure may indicate to the engineering manager to investigate or relook at specific maturation plans.

## Method

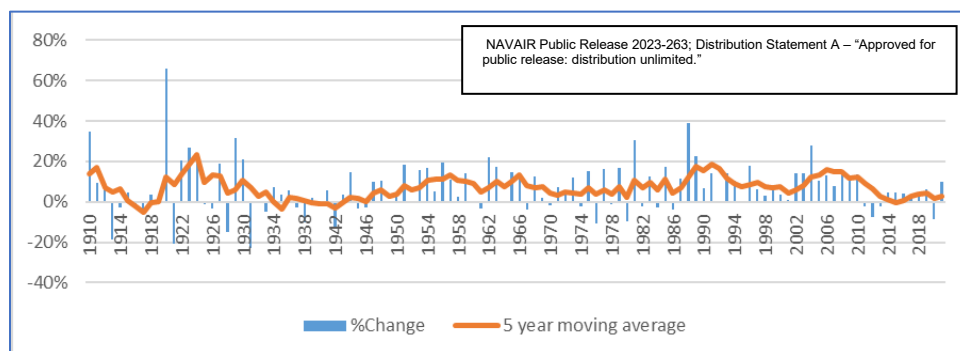
### Overview

The research used online repositories of publications such as IEEE, Wiley, SPIE, and arXiv. Search terms (keywords) are related to specific subject areas such as artificial intelligence (AI). In the case of the IEEE repository, an initial search with no keywords or search terms returns the total number of publications and breakdown of publications in six categories (Conference Papers, Journals, Magazines, Books, Standards, and Courses). An example search of IEEE Explore is shown in Figure 3.



**Figure 13. IEEE Publications  
(Rea, 2022)**

Figure 3 shows the increasing publication trend. A 10 dB rise in publications between any 2 years represents a tenfold increase in publications between years. However, the year-to-year change is noisy and is denoised by a moving average. The trend for a 5-year moving publication average is shown in Figure 4.

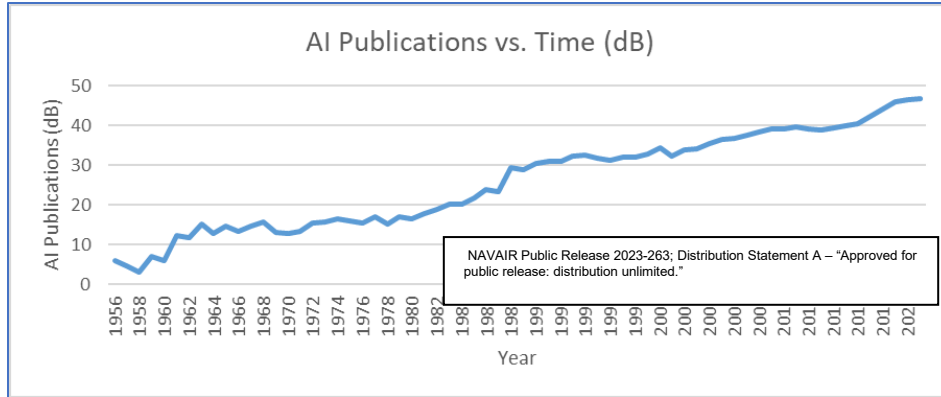


▪ **Figure 14. Percent Publication Volume 5 Year Moving Average  
(Rea, 2022)**

Figure 4 shows that trends in publication volumes over time reflect significant world events.<sup>8</sup> To demonstrate technology publication trends, we collected data from IEEE eXplore using the search term “Artificial Intelligence” and recoded total returned results were recorded for singular years, year over year, from 1956 to 2021 and shown in Figure 5 on a dB scale.

<sup>8</sup> For example, World War I, the Great Depression, World War II, the Gulf War, and the 2008 Global Recession. During SARS-COV-2 there was a -9% change in publication volume from 2019 to 2020.





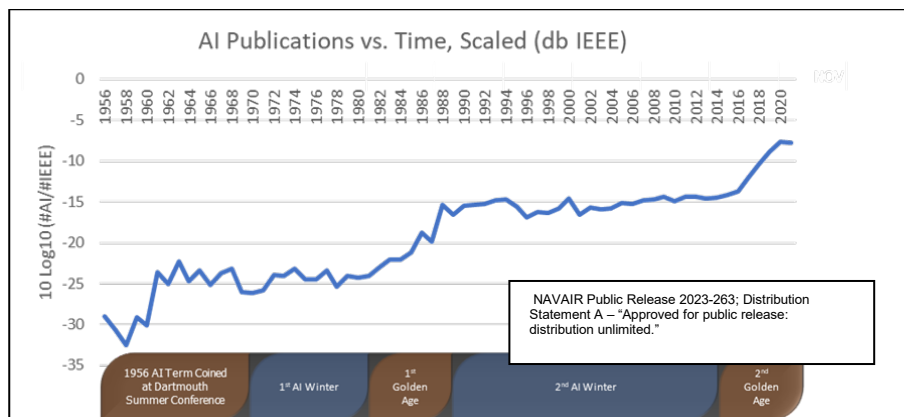
**Figure 15. AI Publications Over Time, dB  
(Rea, 2022)**

Figure 5 shows increasing AI research volume. We define relative research volume ( $RRV_n$ ) as the log of the fraction of all publications in a given repository in year  $n$ . For example, if  $AI_n$  is the total research publications containing the term “Artificial Intelligence” in year  $n$  in the IEEE eXplore repository and  $IEEE_n$  is the total publications in the IEEE eXplore repository in year  $n$ , then  $RRV_n$  is

$$RRV_n = 10 \log_{10} \left( \frac{AI_n}{IEEE_n} \right) \quad (1)$$

In this equation, a value of 0 means all research within a given year, or 100% of publications, contained the search keywords. A value of -10 dB means that the search term was found in 10% of all research for the given year.<sup>9</sup>

Artificial intelligence went through several growth periods from initial discovery; growth periods are recognizable by the upward trend in publication volume of AI relative to all publication data on IEEE 1956–1962, 1978–1987, and 2013–current. Plateaus are areas where there were few publications, such as during the “AI winter” during 1992–2012. This growth and fallback of relative research volume (RRV) is shown in Figure 6.

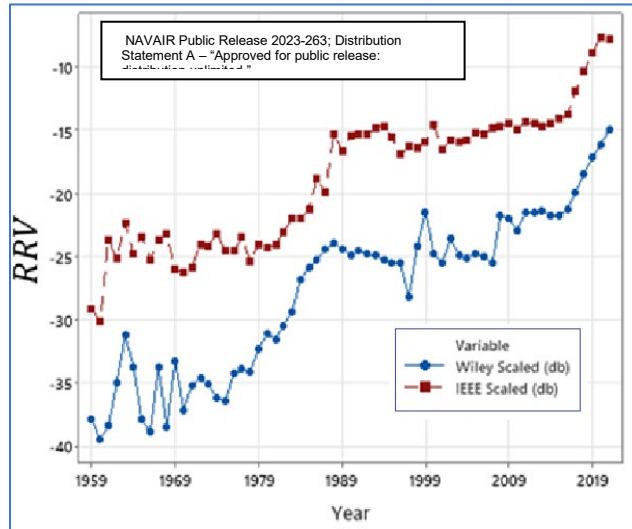


**Figure 16. AI Publications RRV  
(Rea, 2022)**

<sup>9</sup> About 17.5% of all published research in IEEE eXplore was related to AI as of 2021.



This methodology was repeated with a second technical repository (Wiley); the results are shown in Figure 7.



**Figure 17. Comparison of Wiley and IEEE RRVs (Rea, 2022)**

There is over a 94% correlation between these two results.<sup>10</sup> In summary, researchers may create RRV data sets using the following steps:

- Identify technologies of interest.
- Identify related Repository to investigate technology publications over time.
- Formulate search term in accordance with best practices to return relevant publications.
- Record the volume of publications per year for time window of interest.
- Record the total volume of publications in the repository for the same time window.
- Calculate Relative Research Volume according to Equation 5.
- Plot the results and identify trends.

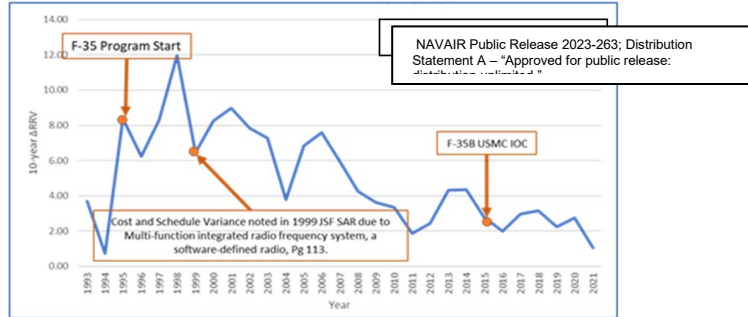
### Change in Relative Research Volume ( $\Delta RRV$ )

We define the change in relative research volume over time as

$$\Delta RRV_n = RRV_x - RRV_{x-n-1} \quad (2)$$

where x is the year of interest and n is the length of the moving time average window. Figure 8 is an annotated example of  $\Delta RRV_{10}$  for software defined radio technology.

<sup>10</sup> Using a zero lag cross correlation.



**Figure 18. Example  $\Delta$ RRV10 (Rea, 2022)**

### Method Calibration

Relative research volumes over time for specific technologies were calibrated using independent technology maturity declarations. GAO *Defense Acquisitions: Assessments of Selected Weapon Programs* reports from 2003 to 2021 were reviewed to identify technical, design, and manufacturing risks to programs of record in the DoD. The year column in the following table represents either a significant event such as program start, the first report year of maturity, or Initial Operational Capability (IOC). All *No* declarations were researched for either program IOC, a GAO declaration of *Yes*, or technology was replacement with a mature substitute.<sup>11</sup> There are a total of 60 data points—31 declared mature and 29 not mature. Table 1 summarizes the systems, critical technologies and binary<sup>12</sup> technology maturity declarations (Rea, 2022).

<sup>11</sup> Such as the Multi-Mission Maritime Aircraft Data Fusion (MMA, later known as P-8)

<sup>12</sup> If there is a clear “Mature” declaration, the technology is treated as mature. Any other adjectives, such as not mature, near (-ly, -ing) maturity are treated as not mature.



**Table 9. Technology Maturity Declarations  
(Rea, 2022)**

#	Technology	Year	Mature?	Notes
1	F-35 Software Defined Radio	1995	No	Program Start,
		1999	No	1999 JSF SAR Pg 113
		2015	Yes	USMC IOC
2	F-35 Sensor and Data Fusion	1995	No	Program Start
		2015	Yes	USMC IOC
3	F-35 Organic Light Emitting Diode	2019	Yes	Gen III HMDS Fielding
4	F-35 Agile Engineering	2017	Yes	C2D2 (Block 4) start
5	2004 DARPA Grand Challenge - Autonomous Driving	2002	No	Announced in 2002
	Autonomous Driving	2021	No	Current research volume
6	FCS - Network Intrusion Detection	2008	No	GAO Report on FCS
7	FCS - Mobile Ad Hoc Networking	2008	No	GAO Report on FCS
8	FCS - Distributed Fusion Management	2008	No	GAO Report on FCS
9	GAO - Quantum Cryptography	2021	No	GAO Report on Quantum Technol
10	GAO - Quantum Communication	2021	No	GAO Report on Quantum Technol
11	GAO - Quantum Key Distribution	2021	No	GAO Report on Quantum Technol
12	GAO - Quantum Computing	2021	No	GAO Report on Quantum Technol
13	GAO - Quantum Random Number Generation	2021	No	GAO Report on Quantum Technol
14	USPS - Optical Character Recognition	2021	Yes	USPS Deployed OCR
15	CVN 78 Dual Band Radar System	2001	No	2004 GAO Report, CVN-21 Progr
	CVN 78 Dual Band Radar System	2021	Yes	CVN 78 IOC
16	E-2D AHE Space Time Adaptive Processing Algorithms	1999	No	2004 GAO Defense Report
		2008	Yes	2008 GAO Defense Report
17	E-2D AHE SiC Power Transistor	2001	No	2004 GAO Report - 2001 program
		2007	Yes	2009 GAO Report
18	GAO - Gait Recognition	2002	No	2002 GAO Report on Biometrics T
	Commercial Gait Recognition	2018	Yes	1st Commercial Availability
19	Space qualified atomic frequency standards	2008	Yes	2008 GAO Page 153
20	MMA Data Fusion	2008	No	2008 GAO Page 157
21	Space Radar - SAR Moving Target Indication	2008	No	2008 GAO Page 167
22	TSAT Program - Dynamic Bandwidth and Resource Allocation	2008	Yes	2008 GAO Page 172
23	VH-71 Voice-over Internet Protocol Security	2008	No	2008 GAO Page 177
24	WGS - Phased Array Radar	2000	Yes	2008 GAO Page 181
25	AMDR - Digital Beamforming	2017	No	2017 GAO Page 98
26	G/ATOR Program - Gallium Nitride Power Amplifier	2016	Yes	2017 GAO Page 108
27	F-22 Geolocation Algorithm	2017	Yes	2017 GAO Page 150
28	F-22 Open Systems Architecture	2020	Yes	2021 GAP Page 130
29	MGUE Anti-Spoof	2017	No	2017 GAO Page 158
30	WSF-M Polarimetric Receiver	2017	No	2017 GAO Page 162
		2019	Yes	2021 GAP Page 114
31	ITEP Additive Manufacturing	2019	No	2019 GAO Page 97





The Table 1 dataset represents systems or technologies or systems with known technologies corresponding to search terms available in the IEEE or Wiley online repositories. There may be cases where specific technology search terms for military specific technologies *do not* match terms used in publicly available literature.

## Results

### Difference Between $\Delta$ RRV of Mature and Not Mature Technologies

Figure 9 shows histograms of 5-, 7-, and 10-year  $\Delta$ RRVs of Table 1 data overlaid with 3-parameter Weibull distributions.

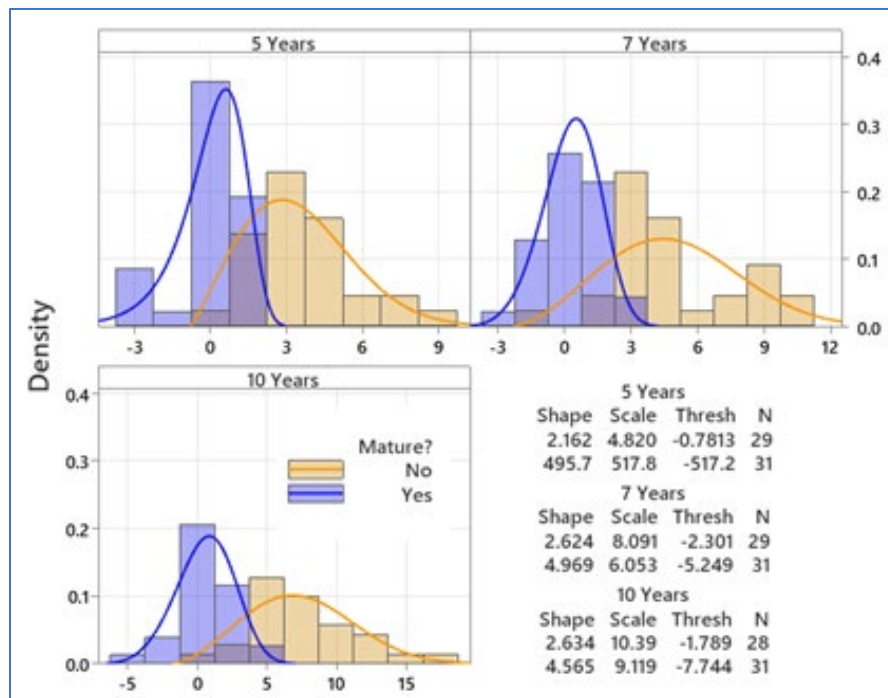
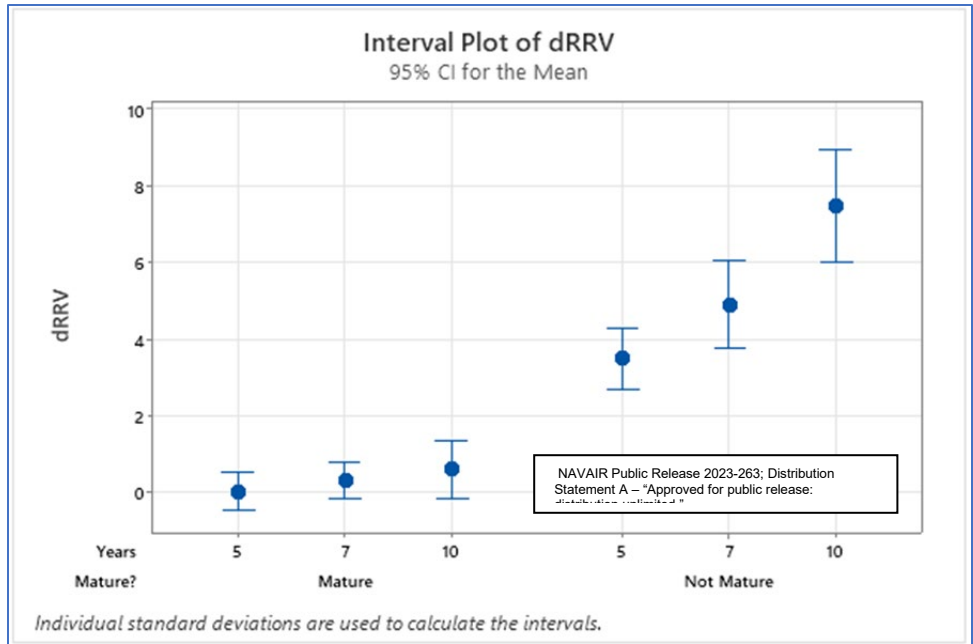


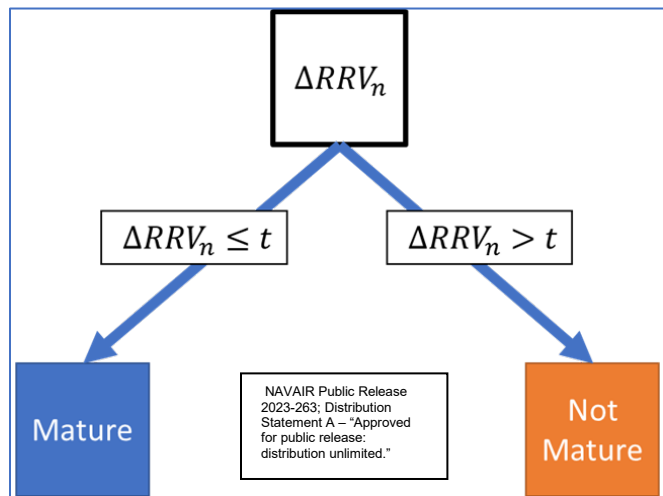
Figure 19. Histograms of Mature vs. Not Mature Declarations for  $\Delta$ RRV (Rea, 2022)

Figure 9 clearly shows separation between the mature and not mature declarations. To prove this, an analysis of variance and means is used. For 5-, 7-, and 10-year  $\Delta$ RRVs ( $n=\{5,7,10\}$ ),  $p=0.000$  for the ANOVA between mature and not mature populations; the mature and not mature samples come from statistically significant different populations. Figure 10 visually summarizes a one-way ANOVA for  $n = \{5,7,10\}$  between mature and not mature declarations.



▪ **Figure 20. Interval Plot of  $\Delta RRV$  Mature/Not Mature Declarations (Rea, 2022)**

Figure 10 clearly shows the difference between mature and not mature declarations and the existence of a threshold  $\Delta RRV$  of approximately 2 between mature and not mature declarations. This is shown visually in Figure 11, where  $t = 2$ .



▪ **Figure 21. Example  $\Delta RRV$  Maturity Classification (Rea, 2022)**

Given the separation between  $\Delta RRV_n$  for  $n=\{5,7,10\}$  does not support the data being from separate populations, decision tree depth is a maximum of 1.

### Conclusions

$\Delta RRV_n$  is shown to be a relevant measure of technology maturity. This will greatly enhance a program manager’s ability to get quick looks at potential solutions’ technology risk,



increase ability for non-experts to have insight into critical technology identification, and empower engineering managers to utilize the technology maturation planning process across a broader range of technologies assessed.

The calculation of  $\Delta RRV$  is simple, and a program or engineering manager could use it as a quick first look at potential critical technology element candidates before subject matter experts have independently reviewed proposals for new or novel uses of technologies. The data for any given technology takes minutes to generate across multiple public repositories.

### Research Limitations/Implications

This research used publicly available data from budget submissions, program-related reporting, contractor annual reports, and contemporaneous press releases. The findings are specific to the analyzed technologies and programs.

### Funding Support and Disclaimer

The views expressed in written materials or publications, and/or made by speakers, moderators, and presenters, do not necessarily reflect the official policies of the Department of Defense, nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. government.

### References

- Azizian, N., Mazzuchi, T., Sarkani, S., & Rico, D. F. (2011). A framework for evaluating technology readiness, system quality, and program performance of U.S. DoD acquisitions. *Systems Engineering*, 14(4), 410–426. <https://doi.org/10.1002/sys.20186>
- Azizian, N., Sarkani, S., & Mazzuchi, T. (2009, October 20). A comprehensive review and analysis of maturity assessment approaches for improved decision support to achieve efficient defense acquisition. *Proceedings of the World Congress on Engineering and Computer Science 2009 Vol. II*. World Congress on Engineering and Computer Science, San Francisco, CA. [www.iaeng.org/publication/WCECS2009/WCECS2009\\_pp1150-1157.pdf](http://www.iaeng.org/publication/WCECS2009/WCECS2009_pp1150-1157.pdf)
- Bailey, R. U., Mazzuchi, T. A., Sarkani, S., & Rico, D. F. (2014). A comparative analysis of the value of technology readiness assessments. *Defense Acquisition Research Journal: A Publication of the Defense Acquisition University*, 21(4), 825–850. a9h.
- Bearden, D. A. (1999). *A methodology for spacecraft technology insertion analysis balancing benefit, cost and risk* [PhD]. University of Southern California.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377. <https://doi.org/10.1002/asi.20317>
- Dodaro, G. L. (2021). *Weapon systems annual assessment updated program oversight approach needed* (GAO-21-222). <https://www.gao.gov/assets/gao-21-222.pdf>
- El-Khoury, B., & Kenley, C. R. (2014). An assumptions-based framework for TRL-based cost and schedule models. *Journal of Cost Analysis and Parametrics*, 7(3), 160–179. <https://doi.org/10.1080/1941658X.2014.982232>
- IEA. (2022). *Counts of global patents for related applications of electric mobility, 2000-2015 IEA, Paris. License CC-BY-4.0*. IEA.Org. <https://www.iea.org/data-and-statistics/charts/counts-of-global-patents-for-related-applications-of-electric-mobility-2000-2015>
- ISO. (2001). *Quality model*. International Organization for Standardization. <https://www.iso.org/standard/22749.html>
- Mankins, J. C. (2009). Technology readiness assessments: A retrospective. *Acta Astronautica*, 65(9), 1216–1223. <https://doi.org/10.1016/j.actaastro.2009.03.058>
- NASA. (2022, January). *Systems engineering handbook*. <https://www.nasa.gov/seh/>
- Oakley, S. S. (2021). *DOD acquisition reform: Increased focus on knowledge needed to achieve intended performance and innovation outcomes* (GAO-21-511T). GAO. <https://www.gao.gov/products/gao-21-511t>



- Olechowski, A. L., Eppinger, S. D., & Joglekar, N. (2015). Technology readiness levels at 40: A study of state-of-the-art use, challenges, and opportunities. *SSRN Electronic Journal*.  
<https://doi.org/10.2139/ssrn.2588524>
- 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>
- Peisen, D. J., & Schulz, C. L. (1999). *Task order 221 case studies: Time required to mature aeronautic technologies to operational*.
- Persons, T. M., Mackin, M., Bothjwell, B., Schwenn, R., & Ortiz, J., Jr. (2020). *Technology readiness assessment guide: Best practices for evaluating the readiness of technology for use in acquisition programs and projects [Reissued with revisions on Feb. 11, 2020.]* (GAO-20-48G). GAO.  
<https://www.gao.gov/products/gao-20-48g>
- Persons, T. M., & Sullivan, M. J. (2016). *Technology readiness assessment guide: Best practices for evaluating the readiness of technology for use in acquisition programs and projects*.
- Radpour, S., Mondal, M. A. H., Paramashivan, D., & Kumar, A. (2021). The development of a novel framework based on a review of market penetration models for energy technologies. *Energy Strategy Reviews*, 38, 100704. <https://doi.org/10.1016/j.esr.2021.100704>
- Ramirez-Marquez, J. E., & Sauser, B. J. (2009). System development planning via system maturity optimization. *IEEE Transactions on Engineering Management*, 56(3), 533–548.  
<https://doi.org/10.1109/TEM.2009.2013830>
- Rea, C. A. (2022). *Using relative research volume as a measure of technology maturity*. ProQuest Dissertations Publishing.
- Sager, M. (2021). *High-risk series: Dedicated leadership needed to address limited progress in most high-risk areas* (GAO-21-119SP). GAO. <https://www.gao.gov/products/GAO-21-119SP>
- Soliman, M., Avgeriou, P., & Li, Y. (2021). Architectural design decisions that incur technical debt—An industrial case study. *Information and Software Technology*, 139.  
<https://doi.org/10.1016/j.infsof.2021.106669>
- Stock, K. (2022, September 27). *Just 4% of North American car production is electric*. Bloomberg.  
<https://www.bloomberg.com/news/articles/2022-09-27/just-4-of-north-american-car-production-is-electric#xj4y7vzkg>
- Sullivan, M. J. (2012). *2012 defense acquisitions: Assessments of selected weapon programs*. 10 U.S.C. § 4272, Pub. L. No. 114–283 (2016). <https://www.law.cornell.edu/uscode/text/10/4272>
- Under Secretary of Defense for Research and Engineering. (2020a). *Defense technical risk assessment methodology (DTRAM)-0-1*. DoD. <https://ac.cto.mil/wp-content/uploads/2021/01/DTRAM-0-1.pdf>
- Under Secretary of Defense for Research and Engineering. (2020b). *Department of Defense independent technical risk assessment guidance*. DoD. <https://ac.cto.mil/wp-content/uploads/2021/01/DTRAM-0-1.pdf>
- Wiesner, K., & Ladyman, J. (2021). Complex systems are always correlated but rarely information processing. *Journal of Physics: Complexity*, 2(4). <https://doi.org/10.1088/2632-072X/ac371c>



# **Management and Business Knowledge Representation for Decision Making: Applying Artificial Intelligence, Machine Learning, Data Science, and Advanced Quantitative Decision Analytics for Making Better-Informed Decisions**

**Dr. Johnathan Mun**— is a Professor of Research at the Naval Postgraduate School in Monterey, CA, and is a specialist in advanced decision analytics, quantitative risk modeling, Monte Carlo simulation, strategic flexibility real options, predictive modeling, and portfolio optimization. He has authored over 32 books, holds 22 patents and patents pending, and has written over 200 technical reports, white papers, analytical notes, and academic journal articles. His prior positions include being Vice President of Analytics at Oracle/Crystal Ball and a senior manager at KPMG Consulting. He has developed multiple software applications that run Monte Carlo risk simulations, predictive modeling, portfolio selection and optimization, strategic real options, analysis of alternatives, total ownership cost, and others. Mun has performed over 100 consulting projects in private industry as well as with the Department of Defense, in areas including decision analytics, return on investment analysis, acquisitions analysis, analysis of alternatives, enterprise risk management, program and project selection and optimization, cost and schedule risk modeling, and others. His latest technical reports with the Naval Research Program and Naval Acquisitions Research Program deal with the return on investment on military education and research, as well as total cost ownership modeling and return on investment on DARPA's 5G protected waveform open-source initiative. He holds a PhD in finance and economics from Lehigh University with specializations in advanced risk analytics, econometric modeling, financial analytics, decision analysis, and strategic real options. He also holds an MBA in management and a BS in physics and biology, as well as multiple other designations, such as the Certified in Quantitative Risk Management (CQRM), Certified in Financial Risk Management (FRM), and Chartered Financial Consultant (CFC), among others. [jcmun@nps.edu]

## **Abstract**

How were the decisions made in the past, and what were the drivers, strategies, or rationale? The old adage holds true on how organizations should learn from the past to help make better decisions in the future. This current first-phase research looks at how the Department of Defense (DoD) can inculcate institutional and corporate memory. Specifically, the research tests and develops recommendations about how a transparent Decisions Options Register (DOR) integrated intelligent database system can be developed, where the DOR helps capture all historical decisions (assumptions, data inputs, constraints, limitations, competing objectives, and decision rules) for programs within the DoD. Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to enable making decisions under uncertainty while leaning on past best practices and allowing senior leadership to make defensible and practical decisions. The current first phase of research uses stylized data and examples to illustrate the recommended methodologies.

This research implements industry best-in-class decision analytics using advanced quantitative modeling methods (stochastic simulation, portfolio optimization) coupled with Artificial Intelligence (AI) and Machine Learning (ML) algorithms (data scraping, text mining, sentiment analysis) and Enterprise Risk Management (ERM) procedures. The DOR will be partially based on ERM methods of using risk registers, where different risk elements are subdivided into different GOPAD groups, or Goals (military capability, cost savings, novel technology, future weapons capability, public safety, government priorities, command preference, etc.), Organization (Air Force, Army, Navy, Marines), Programs (acquisition, commercial-off-the-shelf, joint-industry,



hybrid, etc.), Activity (inventory, replacement, new development, research and development, and so forth), and Domain (air, sea, cyber, etc.) categories.

Multiple competing stakeholders (e.g., the Office of the Secretary of Defense, Office of the Chief of Naval Operations, the U.S. Congress, and the civilian population) have their specific objectives (e.g., capability, efficiency, cost-effectiveness, competitiveness, and lethality, as well as alternatives and trade-offs), constraints (e.g., time, budget, schedule, and manpower), and mission-based domain requirements (e.g., balancing the needs of digital transformation in cybersecurity, cyber-counterterrorism, anti-submarine warfare, anti-aircraft warfare, or missile defense).

This research takes a multidisciplinary approach where methods from advanced analytics, artificial intelligence, computer science, decision analytics, defense acquisitions, economics, engineering and physics, finance, options theory, project and program management, simulation with stochastic modeling, applied mathematics, and statistics are applied. The ultimate goals are to provide decision-makers actionable intelligence and visibility into future decision options or flexible real options, complete with the assumptions that led to certain comparable decisions.

The recommended approaches include the use of supervised and unsupervised AI/ML sentiment text analysis, AI/ML natural language text processing, and AI/ML logistic classification and support vector machine (SVM) algorithms, coupled with more traditional advanced analytics and data science methods such as Monte Carlo simulation, stochastic portfolio optimization and project selection, capital budgeting using financial and economic metrics, and lexicographic rank approaches like PROMETHEE and ELECTRE.

Example case applications, code snippets, and mock-up DORs are presented, complete with stylized data to illustrate their capabilities. The current research outcome will provide a stepping stone to the next phase's multiyear research, where prototypes can be built and actual data can be run through the prescribed analytical engines.

## Introduction

The purpose of this proposed research is to generate a transparent Decisions Options Register (DOR) integrated intelligent database system that helps to capture all historical decisions going forward, including their assumptions, data inputs, constraints, limitations, competing objectives, and decision rules for the Department of Defense (DoD). Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to implement and enable making decisions under uncertainty while leaning on past best practices and allowing senior leadership to make defensible and practical decisions.

The DOR is based on Enterprise Risk Management (ERM) practices in private industry, which typically lists risks and lessons learned from past, current, and proposed future projects. The creation of a documentation database of decision history is critical. There is no learning curve if there is no curve, and you cannot have a curve without any data or information. With the recommended DOR and associated methodologies in this current research, we can compute probabilities of the success and failures of a new program by looking at its characteristics and using historical data as a reference to predict the outcomes. Of course, there will be a need to operationalize and define success versus failure. Just because a program is under budget, on time, requires little rework, and hits all the required specifications and technology release levels, does it mean it is successful? What other metrics might we use to determine definite success or definite failure, and what about all the other levels in between? We need to identify available data as well as the gaps to get us a solid DOR. What are some statistically significant predictors of success and failures as we have operationally defined them? The other issue is risk mitigation and strategic flexibility.



This research will showcase industry best-in-class decision analytics and ERM procedures. The DOR will be partially based on ERM methods of using risk registers, where different risk elements are subdivided into different GOPAD groups, or *Goals* (military capability, cost savings, novel technology, future weapons capability, public safety, government priorities, command preference, etc.), *Organization* (Air Force, Army, Navy, Marines), *Programs* (acquisition, commercial-off-the-shelf, joint-industry, hybrid, etc.), *Activity* (inventory, replacement, new development, research and development, and so forth), and *Domain* (air, sea, cyber, etc.) categories.

Multiple competing stakeholders (e.g., the Office of the Secretary of Defense, Office of the Chief of Naval Operations, the U.S. Congress, and the civilian population) have their specific objectives (e.g., capability, efficiency, cost-effectiveness, competitiveness, and lethality, as well as alternatives and trade-offs), constraints (e.g., time, budget, schedule, and manpower), and mission-based domain requirements (e.g., balancing the needs of digital transformation in cybersecurity, cyber-counterterrorism, anti-submarine warfare, anti-aircraft warfare, or missile defense). These elements are critical when new decisions are to be considered. A DOR database that preserves institutional knowledge and memory will assist in such endeavors and instill trust in the decisions.

This research will take on a multidisciplinary approach where we will be applying methods from advanced analytics, artificial intelligence, computer science, decision analytics, defense acquisitions, economics, engineering and physics, finance, options theory, project and program management, simulation with stochastic modeling, applied mathematics, and statistics. The ultimate goals are to provide decision-makers actionable intelligence and visibility into future decision options or flexible real options, complete with the assumptions that led to certain comparable decisions.

### **Research Current State-of-the-Art**

In a legal dispute, courts use precedents when deciding the outcomes of cases. The use of precedence has been in practice for over 200 years, often to appeal or overturn previous judgments. However, precedent-based decision-making is something that industries and governments have not yet fully embraced. Organizations, including the DoD, tend to have a short memory due to the fluctuations and outflows of human capital and the loss of institutional knowledge when employees leave or are reassigned elsewhere. The current research is intended to include an examination of how related research into the state of the art of precedent-based decision-making is performed today, what might be considered state of the art, and what its current limitations are.

### **Research Approach**

The research applies multiple novel approaches to enhance its success in generating a powerful and searchable DOR database. The recommendations will include key parameters, assumptions, input data, saved models and computations, decisions made, leadership inputs and overrides, constraints and limitations, end goals, and other pertinent information, which can then be mined using *Sentiment Analysis with Machine Learning*, coupled with *Scraping Algorithms and Text Mining with Custom Lexicographic sets*. Users of the system will be able to apply precedent-based insights into their current and future programs. In addition, whenever possible, predictive values will be complemented by actual values captured over time. This allows postmortem analysis of previous programs and provides for lessons learned along the way. Capturing the history of key decisions will help senior leadership make more credible and defensible decisions, which may eventually lead to legal and regulatory changes for the DoD.

The proposed methodologies will allow the collection of data that can be applied in a variety of areas, including, but not limited to, Integrated Risk Management<sup>®</sup> approaches where



*stochastic analyses* like *Monte Carlo simulations*, *stochastic portfolio optimization*, and advanced *data analytical* approaches, *artificial intelligence*, and *data science* methods can be run. Over time, lookback analyses can be applied to update the DOR, making it more closely aligned with the needs of the DoD. The system should be able to collect different types of economic data (total lifecycle cost, total ownership cost, acquisition cost, cost deferred, and schedule and risk costs), logistics data (e.g., inherent availability, effective availability, mission reliability, operational dependability, mean downtime, mean maintenance time, logistics delay time, achieved availability, operational availability, mission availability, fielded capabilities, and Likert levels of creative and novel technology, as well as other metrics), qualitative subject matter expert estimates (strategic value, value to society, command priorities, legal and regulatory impact scores, etc.), and market comparables to operationalize various elements of DoD benefit. At appropriate time intervals, backfitting analyses such as *nonlinear discriminant analysis*, *neural networks*, *distributional fitting*, *limited dependent variables*, *path-dependent partial least squares*, and others can be applied to tease out the *critical success factors* that lead to the success or failure of certain decisions within a program or acquisition.

### **Research Application**

The current research is important because it will create a significant difference in the DoD's decision-making process. The DoD is continually looking for better theoretically justifiable and quantitatively rigorous analytical methods for decision analysis, capital budgeting, and portfolio optimization. The specific interest lies in how to identify and quantify the value of each program to the military and optimally select the correct mix of programs, systems, and capabilities that maximizes some military value (strategic, operational, or economic) while subject to budgetary, cost, schedule, and risk constraints. This research applies private-sector and industry best practices coupled with advanced analytical methods and models to help create these methodologies to do so. However, the uniqueness of the DoD requires that additional work be done to determine the concept of value to the military while considering competing stakeholders' needs. The DoD requires defensible and quantitatively robust concepts of military value in its return on investment for making optimal funding decisions such as where, how much, and how long to invest. These decision options (strategic sequential compound real options, optimal timing options, growth options, and other options to expand, contract, and abandon) are critical when performing an analysis of alternatives and balancing cost-benefit trade-offs in a non-economic DoD environment. The DOR will provide historically preserved insights into the various alternate futures assumed, the alternatives modeled, and why certain decisions were made.

### **Artificial Intelligence and Data Science**

Artificial Neural Network (NN) is a data-driven, distribution-free nonparametric family of methods that can be used for nonlinear pattern recognition, prediction modeling, and forecasting. NN is often used to refer to a combinatorial network circuit of biological neurons. The modern usage of the term often also refers to "artificial neural networks," comprising artificial neurons, or nodes, recreated within a software environment. Such artificial networks attempt to mimic the neurons or neuronal nodes in the human brain in terms of the way humans think, identify patterns, and, in our situation, identify patterns for forecasting time-series data. NN methods can be used in well-behaved time series as well as chaotic physical systems. When used in Big Data (BD) and in conjunction with Machine Learning (ML) approaches, it can be considered as a cross-over to a semi-supervised Artificial Intelligence (AI) system. NN is still considered semi-supervised, as neural networks require a multilayered training process as part of the activation function. For instance, the neural node weights and interactive convolution can be run autonomously once the activation is triggered in the system. In multilayered neuronal nodes, the results from the first node layer will become the inputs into subsequent layers of nodes.





This paper proposes the addition of an internal optimization process to be iteratively run to continually train the nodes to minimize a series of error measurements, such as the standardized sums of squares of errors while balancing and constraining the Akaike Information Criterion, Bayes Criterion, and Hannan-Quinn Criterion. In addition, the proposal here is to add a Combinatorial Fuzzy Logic methodology to the mix to generate the best possible forecast. The term *fuzzy logic* is derived from fuzzy set theory to deal with reasoning that is approximate rather than accurate. As opposed to crisp logic, where binary sets have binary logic, fuzzy logic variables may have a truth value that ranges between 0 and 1 and is not constrained to the two truth values of classic propositional logic. This fuzzy weighting schema is used together with a combinatorial method to yield time-series forecast results.

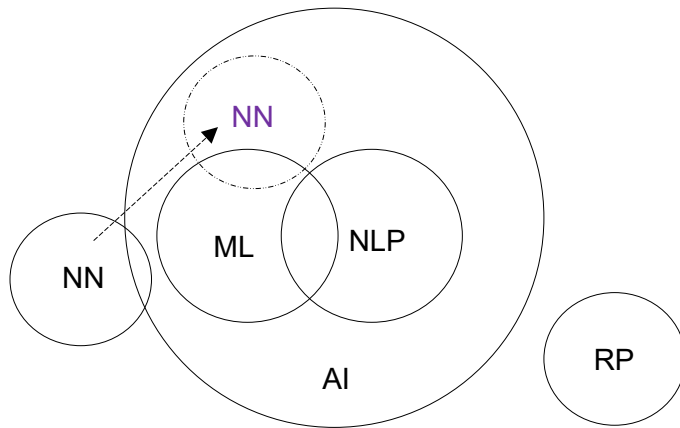
Augur (2016) provides a good summary of the history of data science. According to his research, the term “data science” first appeared as early as 1974, when Peter Naur published his article entitled “Concise Survey of Computer Methods” and defined it as “the science of dealing with data, once they have been established, while the relation of the data to what they represent is delegated to other fields and sciences.” The term took a while to catch on, having not fully integrated the vernacular until 2010. The term “data scientist” is often attributed to Jeff Hammerbacher and D. J. Patil, of Facebook and LinkedIn, in 2008. Between 2011 and 2012, “data scientist” job listings increased by 15,000%, with an emphasis on working with Big Data. By 2016, data science started to become entrenched in the fields of Artificial Intelligence, specifically in the subfields of Machine Learning and Deep Learning.

## Literature Review

Artificial intelligence (AI) is a broad term that refers to a variety of technologies. It’s a catch-all term for a group of inorganic computer science technologies that are used to simulate intelligence. The word AI is often associated with the hazy notion of machine learning, which is a subset of AI in which a computer system is trained to recognize and categorize external real-world data. It is “The ability of machines to perform tasks that normally require human intelligence—for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action—whether digitally or as the smart software behind autonomous physical systems,” according to the DoD’s (2019) AI strategy. The DoD (2019) is particularly interested in these expanded automation capabilities since prospective future near-peer enemies such as Russia and China are investing extensively in this field for military purposes. Given the vast AI field of study, this study focuses on the AI processes deemed most ideal for procurement, such as Machine Learning (ML), Natural Language Processing (NLP), and Robotic Process Automation (RPA), as illustrated in Figure 1 (modified from Sievo [2019]). The image depicts AI as a combination of AI sciences such as machine learning and natural language processing, and while RPA benefits from AI applications, it is not a simulation of human intelligence, but rather a mimic of skills.

The science of AI was established in 1956 to determine whether inorganic robots could execute human-level intelligence capabilities (Denning, 2019). It went through various hype cycles, mostly as a result of sensationalizing what it could do, with numerous disappointments (Figure 2). Significant interest in AI resurfaced about the same time as Big Data computer capacity became more widely available to researchers and businesses, allowing them to apply the science to a variety of practical applications (Haenlein & Kaplan, 2019). Manufacturing robots, smart assistants, proactive healthcare management, illness mapping, automated financial investing, virtual travel booking agents, social media monitoring, conversational marketing bots, NLP tools, and contract management are all examples of commercially feasible AI applications (Daley, 2019).





**Artificial Intelligence (AI):**  
algorithms exhibiting “smart” behavior

**Machine Learning (ML):**  
algorithms that detect patterns and use them for prediction and decision making

**Natural Language Processing (NLP):**  
Algorithms that can interpret, predict, transform, and generate human language

**Robotic Process Automation (RPA):**  
Algorithms that mimic human actions to reduce simple but repetitive tasks

Figure 22. Types of Artificial Intelligence

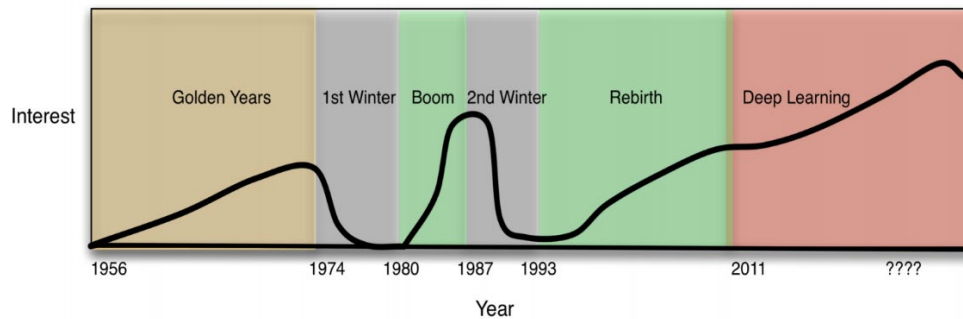


Figure 23. The Timeline of Interest in AI During Different Phases of Its Development (Denning, 2019)

### A Brief History of Data Science

Several good explanations of the history of data science can be found in Press (2013) and Augur (2016). The timeline of data science development is summarized here. We can see how mathematical statistics has evolved into applied statistics, data science, artificial intelligence, and machine learning.

**1962: John Tukey wrote “The Future of Data Analysis,” and as a mathematical statistician, he considered his critical expertise as one able to analyze data.**

1974: Peter Naur published the “Concise Survey of Computer Methods,” where he coined the term data science. He defined it as “the science of dealing with data, once they have been established, while the relation of the data to what they represent is delegated to other fields and sciences.” This term took a while to catch on.

1977: The International Association for Statistical Computing (IASC) was founded. Its main goal was to “link traditional statistical methodology, modern computer technology, and the knowledge of domain experts to convert data into information and knowledge.”

1994: The early forms of modern marketing began to appear, with the main emphasis on Database Marketing.

1996: The term Data Science appeared for the first time at the International Federation of Classification Societies in Japan. The inaugural topic was entitled “Data Science, Classification, and Related Methods.”



**1997: Jeff Wu gave an inaugural lecture titled simply “Statistics = Data Science?”**

2001: William Cleveland published “Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics.” He put forward the notion that data science was an independent discipline and named six areas in which he believed data scientists should be educated: multidisciplinary investigations, models and methods for data, computing with data, pedagogy, tool evaluation, and theory.

**2008: The term “data scientist” is often attributed to Jeff Hammerbacher and DJ Patil of Facebook and LinkedIn.**

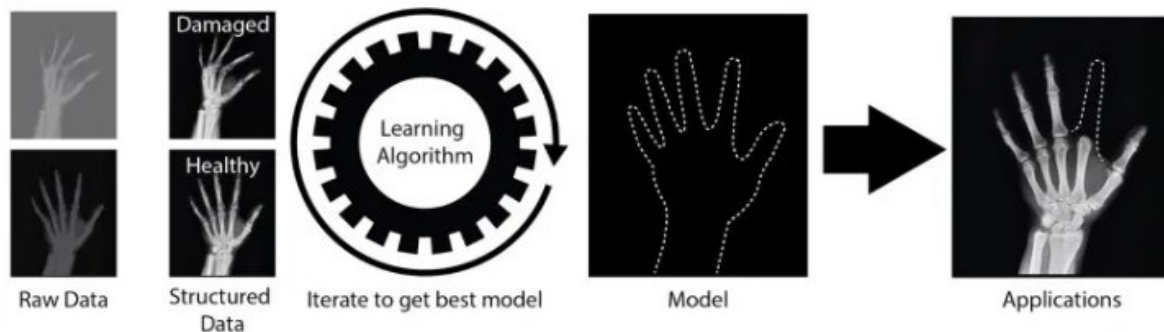
**2010: The term “data science” has fully infiltrated the vernacular. Between 2011 and 2012, “data scientist” job listings increased by 15,000%.**

2016: Data science started to be entrenched in Machine Learning and Deep Learning.

## Machine Learning

Intelligence is the ability to process a specific sort of data, allowing a processor to solve significant problems (Gardner, 1993). Beyond the traditional idea of a person’s intelligence quotient (IQ), which can often simply evaluate how well someone performs on an IQ test rather than their natural talents, psychologists have postulated multiple categories of intelligence. Howard Gardner (1993) proposed a theory of multiple intelligence, which suggests that traditional psychometric views of intelligence are too narrow. Intelligence should be expanded to include more categories in which certain processors, in this case, people, are better at making sense of different stimuli than others. Visual-spatial, linguistic-verbal, interpersonal, intrapersonal, logical-mathematical, musical, body-kinesthetic, and naturalistic intelligence are some of the categories of intelligence (Gardner, 1993). A counter-argument would be that these categories essentially represent learned and disciplined habits that people adopt over their lives as a result of their personality and surroundings. Regardless, both definitions of intelligence (traditional and many) are relevant to the stages involved in creating an artificial intelligence machine.

A computer is capable of doing computations and returning a response based on the data provided. It can be programmed and configured to repeat particular stages or algorithms and even change its conclusions based on previously calculated results using error-correcting techniques. The underlying principle of machine learning is a combination of these two phases. A computer system is fed data that is structured in such a way that the algorithm can identify it, deduce patterns from it, and make assumptions about any unstructured data that is presented later (Greenfield, 2019). In an x-ray learning method, Figure 3 explains how this works.



The image shows the steps an AI algorithm goes through to make a recommendation to a physician on where a missing body part should be. It takes in structured data and develops its understanding of what “right” looks like. When given unstructured data, it compares the image against previously trained models and identifies the abnormality with a recommendation on where to apply a fix, such as a prosthetic.

**Figure 24. AI Training Algorithm (Greenfield, 2019)**



## Supervised Learning

An algorithm is taught the patterns using past data and then detects them automatically in new data. Supervision comes in the form of correct answers that humans provide to train the algorithm to seek out patterns in data. This is commonly used within procurement areas such as spend classification (Sievo, 2019).

## Unsupervised Learning

The algorithm is programmed to identify and potentially detect patterns in new data. Without any human supervision, the algorithm is not expected to surface specific correct answers; instead, it looks for logical patterns within raw data. This is rarely used within critical procurement functions (Sievo, 2019).

## Reinforcement Learning

The algorithm helps to make decisions on how to act in certain situations, and the behavior is rewarded or penalized depending on the consequences. This is largely theoretical in the procurement context (Sievo, 2019).

## Deep Learning

Deep learning is an advanced class of machine learning inspired by the human brain where artificial neural networks progressively improve their ability to perform a task. This is an emerging opportunity in procurement functions (Sievo, 2019).

## Natural Language Processing

Anyone who has used devices that appear to be able to understand and act on written or spoken words, such as translation apps or personal assistants like Amazon's Alexa, is already familiar with NLP-enabled AI. NLP is a set of algorithms for interpreting, transforming, and generating human language in a way that people can understand (Sammalkorpi & Teppala, 2019). Speech soundwaves are converted into computer code that the algorithms understand. The code then translates that meaning into a human-readable, precise response that can be applied to normal human cognition. This is performed using semantic parsing, which maps a passage's language to categorize each word and, using machine learning, creates associations to represent not just the definition of a term but the meaning within a specific context (Raghaven & Mooney, 2013). Figure 4 depicts this categorization and analysis procedure in the context of a DoD procurement contract.

## NATURAL LANGUAGE PROCESSING IN PROCUREMENT

Identifying parts of a text and their grammatical roles through text parsing.

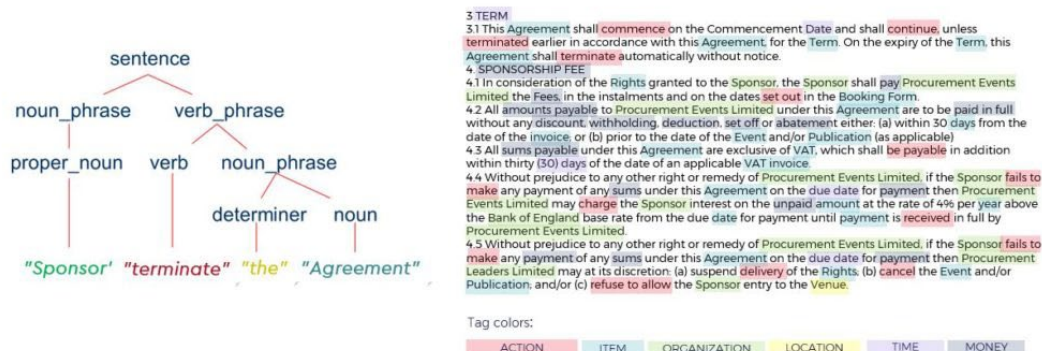


Figure 25. Semantic Parsing in Procurement (Sievo, 2019)



## **Robotic Process Automation**

RPA is not AI, as previously stated; rather, it is an existing process that has been augmented by AI. RPA is defined as “the use of technology by employees in a firm to set up computer software or a robot to capture and interpret current applications for processing transactions, altering data, triggering reactions, and communicating with other digital systems” (Institute for Robotic Process Automation & Artificial Intelligence, 2019). When used correctly, robotic automation offers numerous benefits because it is not constrained by human limitations such as weariness, morale, discipline, or survival requirements. Robots, unlike their human creators, have no ambitions. Working harder will not get you more money or get you promoted, and being permanently turned off will have no effect because robotic automation just duplicates the practical parts of the human intellect, not the underlying nature of mankind (Zarkadakis, 2019). (Note, however, that machine learning relies on an incentive system to make judgments about positive or negative reactions.) A future AI-enabled RPA option is for a machine to learn how to control the source of positive reinforcement fully independent of the rules required to achieve its aim. Things that survive evolve to stay alive because of positive reinforcement from their surroundings and the fact that they continue to act in a way that is regarded as survivable. This should be taken into account in any future AI efforts, especially in the case of why a human must always be present when final judgments are made. Regardless of whether or not AI systems have a perfect track record, they should not be entirely trusted.

## **Proposed Artificial Intelligence, Machine Learning, and Advanced Quantitative Methodologies**

### **Decision Options Database**

As discussed, the purpose of this research is to create a transparent Decisions Options Register (DOR) integrated intelligent database system that helps to capture all historical decisions going forward, including their assumptions, data inputs, constraints, limitations, competing objectives, and decision rules for the DoD. Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to implement and make decisions under uncertainty while leaning on past best practices and allows senior leadership to make defensible and practical decisions.

### **AI/ML Data Reduction and Classification and Logistic Predictive Modeling**

The dataset comprises textual information as well as whether the project was successful (completed) or failed (the program was rejected or canceled). Using the quantitative variables, AI/ML classification routines can be applied to determine the probability that a potential or future program will also be successful or fail.

The classification routine we will use applies in the situation where the dependent variable contains data that are limited in scope and range, such as binary responses (0 or 1 for failures/successes), truncated, ordered, or censored data. For instance, given a set of independent variables (e.g., age, income, education level of credit card or mortgage loan holders), we can model the probability of defaulting on mortgage payments using maximum likelihood estimation (MLE). The response or dependent variable  $Y$  is binary. That is, it can have only two possible outcomes that we denote as 1 and 0 (e.g.,  $Y$  may represent the presence/absence of a certain condition, defaulted/not defaulted on previous loans, success/failure of some device, answer yes/no on a survey, etc.) and we also have a vector of independent variable regressors  $X$ , which are assumed to influence the outcome  $Y$ . A typical ordinary least squares regression approach is invalid because the regression errors are heteroskedastic and non-normal, and the resulting estimated probability estimates will return nonsensical values of above 1 or below 0.



MLE analysis handles these problems using an iterative optimization routine to maximize a log-likelihood function when the dependent variables are limited.

A Logit or Logistic regression is used for predicting the probability of occurrence of an event by fitting data to a logistic curve. It is a generalized linear model used for binomial regression, and like many forms of regression analysis, it makes use of several predictor variables that may be either numerical or categorical. MLE applied in a binary multivariate logistic analysis is used to model the dependent variable to determine the expected probability of success of belonging to a certain group. The estimated coefficients for the Logit model are the logarithmic odds ratios, and they cannot be interpreted directly as probabilities. A quick computation is first required, and the approach is simple.

Specifically, the Logit model is specified as  $Estimated Y = LN[P_i/(1 - P_i)]$  or, conversely,  $P_i = EXP(Estimated Y)/(1 + EXP(Estimated Y))$ , and the coefficients  $\beta_i$  are the log odds ratios. So, taking the antilog or  $EXP(\beta_i)$ , we obtain the odds ratio of  $P_i/(1 - P_i)$ . This means that with an increase in a unit of  $\beta_i$ , the log odds ratio increases by this amount. Finally, the rate of change in the probability  $dP/dX = \beta_i P_i(1 - P_i)$ . To estimate the probability of success of belonging to a certain group (e.g., predicting if a program will develop issues and eventually fail given a certain combination of lifecycle cost, ROI, FTE requirements, length of time, strategic value, etc.), we simply compute the *Estimated Y* value using the MLE coefficients and convert it into the inverse antilog of the odds ratio as discussed previously. Next, we can take the statistically significant variables and apply them to a Gaussian Support Vector Machine (SVM) to classify the programs into high probabilities of approval or rejection categories.

### First Step

Model Inputs:

VAR1

VAR2; VAR3; VAR4; VAR5; VAR6; VAR7; VAR8; VAR9

Status (D)

Monthly FTE, Complexity Level, Strategic Value, Value to Command, Length in Months, Program Cost, Overrun Ratio, Annual Cost Savings

Generalized Linear Model (Logit with Binary Outcomes)

	Coefficient	Std. Error	Wald Test	P-value	Exp(B)	Lower	Upper
Intercept	-1.634198	0.754434	4.692098	0.030302	0.195109	0.000000	0.000000
VAR1	0.028625	0.020496	1.950585	0.162524	1.029039	0.988520	1.071218
VAR2	0.076812	0.144371	0.283071	0.594695	1.079839	0.813711	1.433004
VAR3	-0.262500	0.040630	41.7411	<b>0.000000</b>	0.769127	0.710254	0.832879
VAR4	-0.096195	0.027419	12.3083	<b>0.000451</b>	0.908287	0.860764	0.958434
VAR5	0.000823	0.012687	0.004210	0.948266	1.000824	0.976243	1.026022
VAR6	0.074324	0.039911	3.467833	<b>0.062573</b>	1.077155	0.996106	1.164799
VAR7	0.564136	0.134590	17.5689	<b>0.000028</b>	1.757929	1.350325	2.288569
VAR8	0.049994	0.101851	0.240943	0.623526	1.051265	0.861027	1.283535

Log-Likelihood -199.9830

Restricted Log-Likelihood -285.4773

McFadden's R-Squared 0.299479

Cox and Snell's R-Squared 0.289636

Nagelkerke's R-Squared 0.425440

Raw Akaike Info. Criterion 417.9659

Raw Bayes Criterion 455.8974

Log-Likelihood -199.9830



Restricted Log-Likelihood -285.4773  
 Chi-Square 170.9886  
 Degrees of Freedom 8  
 P-value 0.000000

**Second Step**

Model Inputs:

VAR1

VAR4; VAR5; VAR7; VAR8

Status (D)

Strategic Value, Value to Command, Program Cost, Overrun Ratio

Generalized Linear Model (Logit with Binary Outcomes)

	Coefficient	Std. Error	Wald Test	P-value	Exp(B)	Lower	Upper
Intercept	-0.781188	0.305330	6.545958	0.010512	0.457862	0.000000	0.000000
VAR1	-0.239706	0.033215	52.0818	<b>0.000000</b>	0.786859	0.737266	0.839788
VAR2	-0.074519	0.023632	9.942889	<b>0.001615</b>	0.928190	0.886178	0.972194
VAR3	0.082202	0.022767	13.0359	<b>0.000306</b>	1.085675	1.038294	1.135218
VAR4	0.588673	0.108123	29.6424	<b>0.000000</b>	1.801597	1.457549	2.226855

Log-Likelihood -201.7171  
 Restricted Log-Likelihood -285.4773  
 McFadden's R-Squared 0.293404  
 Cox and Snell's R-Squared 0.284691  
 Nagelkerke R-Squared **0.418177**  
 Raw Akaike Info. Criterion 413.4342  
 Raw Bayes Criterion 434.5072

Log-Likelihood -201.7171  
 Restricted Log-Likelihood -285.4773  
 Chi-Square 167.5204  
 Degrees of Freedom 4  
 P-value 0.000000

**Third Step**

Model Inputs:

Status (D)

Strategic Value, Value to Command, Program Cost, Overrun Ratio

Sigma, Lambda, Omega, Calibration Level: 1.00, 1.00, 0.40, 1.00

**AI Machine Learning: Classification with Gaussian SVM (Supervised)**

Relax: 8.218332

Accuracy 68.20% 67.40% 68.20% **69.40%** 67.80% 67.80% 66.60% 65.00% 63.40% 62.20%  
 Omega 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00

Forecast	Group
1.118101	1.00
0.971805	0.00
...	..
...	..
0.971828	1.00



## Stochastic Simulation and Probabilistic Analysis

Another recommended approach is to perform stochastic distributional fitting; that is, how do the collected historical data fit known probability distributions? These fitted distributions can be used as the variable's input parameters (e.g., a Fréchet or Weibull distribution with shape and scale parameters of 0.5 and 1.2). Figure 5 illustrates an example where historical program costs were fitted to determine its distributional properties. With the fitted distribution, these can be used as inputs into a Monte Carlo simulation model to forecast and predict a new program's chances of success.

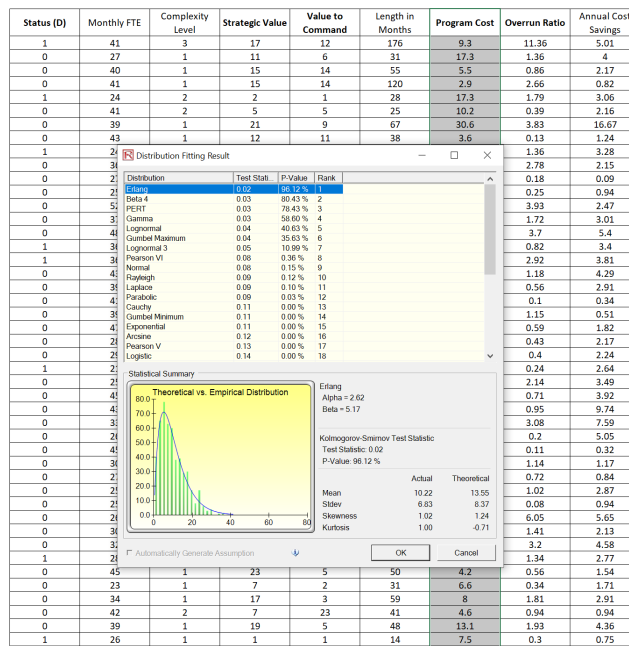


Figure 26. Distributional Fitting for Probabilistic Analysis and Stochastic Simulation

Predicted	Actual	Correct	Datapoints
		72.4%	
negative	negative	0	a few months before the release of star wars episode 1 , the phantom menace , 20th century fox decides to release another space film
negative	negative	1	bad movies described as " a swift descent into sinful pleasure , decay , and debauchery " are hard to watch . bad 2000 ' s movies that n
positive	negative	0	bruce willis needs to stay away from straightforward action pictures . mercury rising adds to a growing list ( including such stinkers as his
negative	negative	1	capsule : godawful " comedy " that ' s amazingly shabby and cut - rate , and rather bereft of laughs . i was having a bad week in my life
negative	negative	1	battlefield earth is the worst film of 2000 , and i guarantee you that nothing else this year will even come close . in fact , i ' ll be surprised
negative	negative	1	jet li busted onto the american action movie scene , when he stole the show in 1998 ' s lethal weapon 4 , with his wicked looks , his nast
negative	negative	1	one night , during a torrential downpour that flooded the streets , we went to see -- what else -- hard rain . " so , are we all going to die
negative	negative	1	it was with a huge lack of something to do that i decided to watch this on good old upn on sunday afternoon , when the only good things
negative	negative	1	vampires starts out almost in the style of a spaghetti western with an attack on a small homestead in new mexico . the house has a nest
positive	negative	0	when considering david fincher ' s latest film , " the game " , four words come to mind . " don ' t believe the hype . " this michael douglas
positive	negative	0	weighed down by tired plot lines and spielberg ' s reliance on formulas , _saving private ryan_ is a mediocre film which nods in the direct
positive	negative	0	if you ' re going to make a two - hour hollywood in - joke , why bother releasing it to the general public ? if you ' re going to create a film
positive	negative	0	it used to be that not just anyone could become a vampire . usually , you had to be an aristocrat - a count such as dracula or karstein .
positive	negative	0	hav plenty , as we are told in the beginning and reminded during the film , is a true story . life itself is a series of true stories , but most a
positive	negative	0	it seems that i ' ve stopped enjoying movies that should be fun to watch . take payback , for example . a movie that most people seem to
negative	negative	1	note : some may consider portions of the following text to be spoilers . be forewarned . " quick , robin ! the anti - shark repellent ! " - add
negative	negative	1	everything about this ninth trek movie seems on the cheap , from the roger carman - grade special effects to its highly derivative and ugly
negative	negative	1	this is one of the worst big - screen film experiences i ' ve had for a while . with this film , plus ' showgirls ' and ' basic instinct ' , paul ver
negative	negative	1	phew , what a mess ! for his fifth collaboration with director rich - ard donner ( lethal weapon i - iii , maverick ) , mel gibson plays a moto
negative	negative	1	in french , the phrase " film noir " literally means " black film . " webster defines it as " a type of crime film featuring cynical malevolent ch
positive	negative	0	plot : lara croft is british , rich and kicks a lot of ass . she also likes to raid tombs but when the illuminata discover that all nine planets ar
negative	negative	1	it happens every year -- the days get longer , the weather gets warmer and the studios start releasing their big - budget blockbusters . i
negative	negative	1	i heard actor skeet ulrich discussing this film in a couple of interviews , and in both instances , he felt the strange compulsion to compare
negative	negative	1	it is with some sad irony that i screened fright night part 2 on the day that one of it ' s stars , roddy mcdownal passed away at the age of
positive	negative	0	sometimes a stellar cast can compensate for a lot of things , and " pushing tin " certainly features some name stars who are going place
negative	negative	1	the most depressing thing about the depressingly pedestrian james bond film " the world is not enough " is its final frame : white letters o
negative	negative	1	when wait disney pictures announced a live - action feature based on the ' 60s cartoon series of " mr . magoo , " special interests group:
positive	negative	0	david spade has a snide , sarcastic sense of humor that works perfectly on the tv sitcom just shoot me . it also served as a good showc:
positive	negative	0	9 : its pathetic attempt at " improving " on a shakespeare classic . 8 : its just another piece of teen fluff . 7 : kids in high school are not th
negative	negative	1	sometimes i wonder just what the censors are thinking . take this film , " naked killer " , among it ' s ingredients are heavy doses of violer
positive	negative	0	after seeing blaze and driving miss daisy , i was ready for some mindless fun -- oh , maybe something like tango & cash . maybe not ! m
negative	negative	1	in " the 13th warrior , " arab poet ahmed ibn fahdian ( antonio banderas ) finds himself kicked out of baghdad for feeling up the king ' s o
positive	negative	0	there isn ' t much good about this movie . not much i can say about the acting , directing , or writing that would make you consider seeing
positive	negative	0	" varsity blues " is the best film of 1999 thus far . unfortunately , it is also the first film i have seen from 1999 . it is another one of those s
negative	negative	1	for a film touted as exploring relationships and black sexuality , troy is surprisingly tame . despite it ' s lurid subject matter and it ' s pass
positive	negative	0	" mission to mars " is one of those annoying movies where , in the middle of the movie , you get the sneaking suspicion that the reason th
negative	negative	1	my friend here in film school just made a two minute - long film for one of his classes that includes a staged anal rape scene , done by tv
positive	negative	0	remember back in the mid 1990s when crime and macabre movies were all the rage ? " pulp fiction " and " fargo " both managed to get i
positive	negative	0	_dirty_work_ has a premise of deliciously mean - spirited potential . mitch weaver ( norm macdonald ) and his lifelong best friend sam m
positive	negative	0	america ' s favorite homicidal plaything takes a wicked wife in " bride of chucky , " and their unholy matrimony is something old , nothing i

Figure 27. Text Scraping and Sentiment Analysis Dataset and AI Classification Results





### Lessons Learned in the Prototype Application

The prototype was very insightful in that it provided a myriad of lessons learned. For instance, the issue of hypernyms and hyponyms can be developed to create a hierarchical structure of a custom dictionary, whereby using text scraping methodologies, we can complement the learning algorithm with our custom lexicon. Sayings, proverbs, adages, and other types of word structures will also need to be considered, as will concise wordings or mixed negatives (e.g., “no good” is a negative connotation as opposed to a positive “good” implication despite the fact that the word exists in the context). The impact scores of certain words and their frequencies can also be used to generate word clouds and help create visuals of the most frequent and impactful comments from past programs.

### Neural Network Pattern Recognition Prediction Methods

Using the Box-Jenkins method of forward-looking predictive steps, we have

$$\hat{x}_{t+1} = f(x_t, x_{t-1}, \dots, x_{t-n})$$

$$\hat{x}_{t+2} = f(x_{t+1}, x_t, \dots, x_{t-n+1})$$

...

$$\hat{x}_{t+k} = f(x_{t+k-1}, x_{t+k-2}, \dots, x_{t-n+k-1})$$

Where  $x_t$  is the observation of  $x$  at time  $t$ . This means that if we use a  $k$  step ahead predictive model, we have

$$x_{t+1} = f_1(x_t, x_{t-1}, \dots, x_{t-n})$$

$$x_{t+2} = f_2(x_t, x_{t-1}, \dots, x_{t-n})$$

...

$$x_{t+k} = f_k(x_t, x_{t-1}, \dots, x_{t-n})$$

Here, we see that  $f_i$  are computed in the neural network paradigm.

### Activation Transfer Functions for Neural Networks

Logistic sigmoidal function:  $f(x) = (1 + e^{-x})^{-1}$

Hyperbolic tangent function:  $f(x) = (e^x - e^{-x})(e^x + e^{-x})^{-1}$

Sine and cosine function:  $f(x) = \sin(x)$  or  $f(x) = \cos(x)$

Linear function:  $f(x) = x$

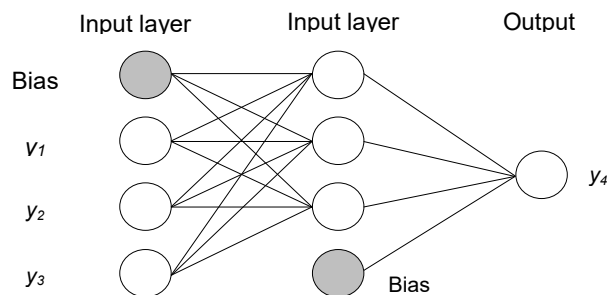


Figure 28. A Multiple Layered Perceptron Neural Network

The neural mapping assumes that  $y_4$  is the dependent variable, whereas  $y_1, y_2, y_3$  and a constant term is the set of independent variables. The neural network has an input layer, a hidden layer, and an output layer. There are three inputs in the input layer, a neuron for the biases, four neurons in the hidden layer, and one neuron in the output layer.

### Error Measurements and Error Correction for Parameter Calibration

Total Variables (Dependent and Independent):  $v$

$$\text{Mean Absolute Deviation: } MAD = \frac{\sum |e_t|}{n}$$

$$\text{Root Mean Squared Error: } RMSE = \sqrt{\frac{\sum (e_t)^2}{n}}$$

$$\text{Sums of Squared Errors: } SSE = \sum (e_t)^2$$

$$\text{Maximum Log-Likelihood: } MLL = \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \frac{SSE}{2} - SSE \left[ \frac{n}{2SSE} \right]$$

$$\text{Akaike Information Criterion: } AIC = \frac{-2MLL}{n} + \frac{2k}{n}$$

$$\text{Bayes Information Criterion (BIC): } BC = AIC + \frac{2(v+2)(k+3)}{n-k-3}$$

Pesaran-Timmermann Test:  $PT = \frac{p(xf) - p'}{\sqrt{v-w}}$  where  $v = \frac{p'(1-p')}{n}$ ,  $p' = f^+x^+ + (1-f^+)(1-x^+)$ , and where  $w = \frac{(2f^+-1)^2x^+(1-x^+)}{n} + \frac{(2x^+-1)^2f^+(1-f^+)}{n} + \frac{(4x^+f^+)(1-x^+)(1-f^+)}{n^2}$ ,  $x^+$  is the proportion of positives on the data and  $f^+$  is the proportion of positives on the forecast

$$\text{Hannan-Quinn Information Loss Criterion: } HC = \frac{-2MLE}{n} + \frac{2k \ln(\ln(n))}{n}$$

### Training Algorithms

For model 1:

Variables used in Training set	Predicted value
$y_1, y_2, y_3$	$\hat{y}_4$
$y_2, y_3, y_4$	$\hat{y}_5$
...	...
$y_{400}, y_{401}, y_{402}$	$\hat{y}_{403}$

Using the coefficients obtained from the training set, we do the following on the testing set:

Variables used in Testing set	Predicted value
$y_{401}, y_{402}, y_{403}$	$\hat{y}_{404}$
$y_{402}, y_{403}, y_{404}$	$\hat{y}_{405}$
...	...
$y_{420}, y_{421}, y_{422}$	$\hat{y}_{423}$

When forecasting, we use

Independent Variables	Predicted value
$y_{401}, y_{402}, y_{403}$	$\hat{y}_{404}$
$y_{402}, y_{403}, \hat{y}_{404}$	$\hat{y}_{405}$



For model 2:

Variables used in Training set	Predicted value
$y_1, y_2, y_3, y_4, y_5, y_6$	$\hat{y}_7$
$y_2, y_3, y_4, y_5, y_6, y_7$	$\hat{y}_8$
...	...
$y_{300}, y_{301}, y_{302}, y_{303}, y_{304}, y_{305}$	$\hat{y}_{306}$

Using the coefficients obtained from the training set, we do the following on the testing set:

Variables used in Testing set	Predicted value
$y_{301}, y_{302}, y_{303}, y_{304}, y_{305}, y_{306}$	$\hat{y}_{307}$
$y_{302}, y_{303}, y_{304}, y_{305}, y_{306}, y_{307}$	$\hat{y}_{308}$
...	...
$y_{330}, y_{331}, y_{332}, y_{333}, y_{334}, y_{335}$	$\hat{y}_{336}$

When forecasting, we use

Independent Variables	Predicted value
$y_{301}, y_{302}, y_{303}, y_{304}, y_{305}, y_{306}$	$\hat{y}_{307}$
$y_{302}, y_{303}, y_{304}, y_{305}, y_{306}, \hat{y}_{307}$	$\hat{y}_{308}$

Figure 8 shows a neural network for time series forecasting. The functions  $\phi_0$  and  $\phi_h$  are called activation functions, and the logistic function and linear function are usually chosen. One of the input nodes is sometimes called the data bias node.

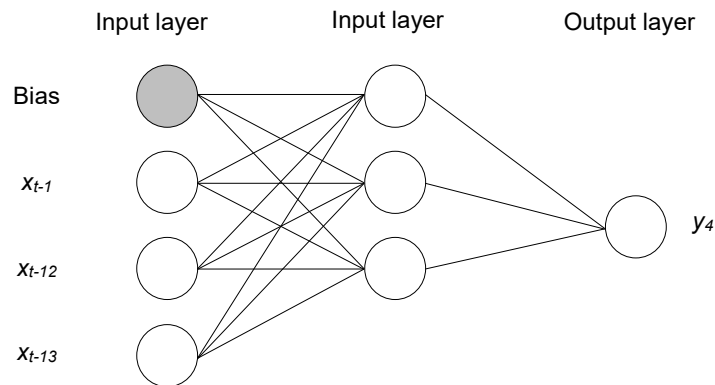


Figure 29. Example of Multiple Layered Perceptrons

The functional form to be modeled looks like this:

$$\hat{x}_t = \phi_0 \left( w_{co} + \sum_h w_{ho} \phi_h \left( w_{ch} + \sum_h w_{ih} x_{t-j_i} \right) \right)$$

$w_{ch}$  is the weight for the connections between the constant input and the hidden neurons,  $w_{co}$  is the weight of the direct connection between the constant input and output, and  $w_{ih}$  and  $w_{ho}$  are the weights for the other connections between the inputs and hidden neurons.

The model is built in the following steps:

- Obtain periodic time-series data.
- Calculate log-returns  $\log \frac{x_t}{x_{t-1}}$ .
- Transform the log returns to the interval [0, 1]. The reason is that we use logistic functions in both the hidden layer and the output layer. The output of the logistic function lies between [0, 1].
- Decide which part of the data are as a training set, i.e., to train the neural network to obtain the weights  $w_{ih}$ ,  $w_{ch}$ ,  $w_{ho}$ ,  $w_{co}$  in the figure, and which part of the data is used as the testing set. That is, after we train the neural network, we use those weights for forecasting and comparing with the data in the testing set.
- Decide on the inputs:  $x_{t-1}$ ,  $x_{t-2}$ ,  $x_{t-3}$  are used to predict  $x_t$ . Then,  $x_{t-1}$ ,  $x_{t-2}$ ,  $x_{t-3}$ ,  $x_{t-4}$ ,  $x_{t-5}$ ,  $x_{t-6}$  are used to predict  $x_t$ .
- Build the model to calculate the output.
- Use the data table to calculate the output by neural network for the data in the training set. Calculate the sum of squares of errors. Minimize this term using an internal optimization routine. We obtain the trained neural network, and the weights can be used for forecasting.
- Use the data table to calculate the output by neural network for the data in the testing set. Compare with true value.

## Decision Analytics Methodology

### A Combined Lexicographic Average Rank Approach for Evaluating Uncertain Multi-Indicator Matrices with Risk Metrics<sup>1</sup>

In many situations, projects are characterized by several criteria or attributes that can be assessed from multiple perspectives (financial, economic, etc.). Each criterion is quantified via performance values (PV), which can either be numerical or categorical. This information is typically structured in a multi-indicator matrix  $\mathbf{Q}$ . A typical problem faced by a decision-maker is to define an aggregate quality (AQ) able to synthesize the global characteristics of each project and then derive the rankings from the best to the worst base-case ranking (Mun et al., 2016).

Ranking techniques can be classified as parametric and nonparametric. A parametric technique requires information about decision-maker preferences (e.g., criterion weights). According to Dorini, Kapelan, and Azapagic (2011), some examples of parametric techniques include the ELECTRE methods (Roy, 1968) and Preference Ranking Organization Methods for Enrichment Evaluations (PROMETHEE; Brans & Vincke, 1985). Nonparametric techniques, such as Partial Order Ranking (Bruggemann et al., 1999) and Copeland Scores (Al-Sharrah, 2010), do not require information from the decision-maker. In general, all of these techniques can produce a ranking of the alternatives from the best to the worst.

Therefore, given a matrix  $\mathbf{Q}$ , the selected procedure generates a ranking defined as the base-case rank (BCR). As a result of this assessment, for each alternative, a specific rank  $R_i$  that considers the multiple perspectives defined by the decision-maker is obtained. The set of  $R_i$  corresponds to the global evaluation under the first synthetic attribute, defined and named as *base ranking* and capable of characterizing the alternatives in the base case.

---

<sup>1</sup> Some of the material discussed in this section is based on previous work by the author, Dr. Johnathan Mun, and his team, Dr. Elvis Hernández-Perdomo and Dr. Claudio M. Rocco. Their work has been published as a chapter, "A Combined Lexicographic Average Rank Approach for Evaluating Uncertain Multi-Indicator Matrices with Risk Metrics," in *Partial Order Concepts in Applied Sciences*, M. Fattore and R. Bruggemann (eds.), Springer International Publishing (2017).



However, in real-life situations, each performance value could be affected by uncertain factors. Several approaches have been presented for analyzing how the uncertainty in the performance values (the input) affects the ranking of the objects (the output; Rocco & Tarantola, 2014; Corrente et al., 2014; Hyde et al., 2004; Hyde & Maier, 2006; Yu et al., 2012). The approaches, based on Monte Carlo simulation, consider each uncertain factor as a random variable with known probability density functions. As a result, the AQ of each alternative and, therefore, the ranking also become random variables with approximated probability distributions. In such situations, the decision-maker could perform probability distribution evaluations. For example, the decision-maker could be interested in determining not only the worst rank of a specific alternative, but also its probability and volatility (risk evaluation).

In the standard approach, the probability of an alternative being ranked as in the BCR is selected as the synthetic attribute *probability* able to characterize the alternatives under uncertainty.

The stochastic nature of the AQ of each alternative could be further assessed to reflect the risk evaluation induced by uncertainty. In this case, it is required to compare several random variables synthesized through their percentiles and statistical moments. Several approaches have been proposed to this end, such as a simple comparison of the expected value, the expected utility (Von Neumann & Morgenstern, 1947), the use of low-order moments (Markowitz, 1952), risk measures (Jorion, 2007; Mansini et al., 2007; Rockafellar & Uryasev, 2000), the Partitioned Multiobjective Risk Method (PMRM; Asbeck & Haimes, 1984; Haimes 2009), and the stochastic dominance theory (Levy, 2006), among others.

To consider the risk evaluation induced by uncertainty, each alternative is represented by the third synthetic attribute: *compliance*. This new attribute is based on a simultaneous assessment of several risk measures and some moments of each AQ distribution (Mun et al., 2016).

At this point, each alternative is assessed from three different angles:

1. Multiple decision-making perspectives that include several aspects such as economic, financial, technical, and social (*base ranking*)
2. Uncertainty propagation on performance values (*probability*)
3. A risk evaluation based on the generated probability distribution (*compliance*)

These perspectives are then used for defining a new multi-indicator matrix  $\mathbf{Q}_1$  correlated to projects and synthesized using a ranking technique. However, in some situations, decision-makers need to select projects following their most preferred criteria successively. For this reason, an aggregation ranking technique that allows compensation is useless.

Therefore, the final assessment is derived using a combined approach based on a *nonparametric aggregation rule* (using the concept of average rank) for attributes 1 and 2; a simple procedure for score assignment for attribute 3; and a *lexicographic rule*. In addition, a preliminary analysis of the alternatives is performed using a Hasse diagram (Bruggemann et al., 1999). To the best of the researcher's knowledge, this type of combined assessment has not been reported in the literature.

### Average Rank Approach

Let  $P$  define a set of  $n$  objects (e.g., alternatives) to be analyzed and let the descriptors  $q_1, q_2, \dots, q_m$  define  $m$  different attributes or criteria selected to assess the objects in  $P$  (e.g., cost, availability, environmental impact). Attributes must be defined to reflect, for example, that a low value indicates low rankings (best positions), while a high value indicates high ranking (worst positions; Restrepo et al., 2008). However, for a given problem or case study, this convention could be reversed.



If only one descriptor is used to rank the objects, then it is possible to define a total order in  $P$ . In general, given  $x, y \in P$ , if  $q_i(x) \leq q_i(y) \forall i$ , then  $x$  and  $y$  are said to be comparable. However, if two descriptors are used simultaneously, the following could happen:  $q_1(x) \leq q_1(y)$  and  $q_2(x) > q_2(y)$ . In such a case,  $x$  and  $y$  are said to be incomparable (denoted by  $x||y$ ). If several objects are mutually incomparable, set  $P$  is called a partially ordered set or *poset*. Note that since comparisons are made for each criterion, no normalization is required.

The objects in a poset can be represented by a directed acyclic graph whose vertices are the objects  $\in P$ , and there is an edge between two objects only if they are comparable and one covers the other, that is, when no other element is in between the two. Such a chart is termed a Hasse diagram (Bruggemann et al., 1995).

A Hasse diagram is, then, a nonparametric ranking technique and can perform ranking decisions from the available information without using any aggregation criterion. However, while it cannot always provide a total order of objects, it does provide an interesting overall picture of the relationships among objects.

A useful approach to producing a ranking is based on the concept of the average rank of each object in the set of linear extensions of a poset (De Loof et al., 2011). Since the algorithms suggested for calculating such average ranks are exponential (De Loof et al., 2011), special approximations have been developed, such as the Local Partial Order Model (LPOM; Bruggemann et al., 2004), the extended LPOM (LPOMext; Bruggemann & Carlsen, 2011), or the approximation suggested by De Loof et al. (2011).

From the Hasse diagram, several sets can be derived (Bruggemann & Carlsen, 2011). If  $x \in P$ ,

1.  $U(x)$ , the set of objects incomparable with  $x$ :  $U(x) := \{y \in P: x||y\}$
2.  $O(x)$ , the *down* section:  $O(x) := \{y \in P: y \leq x\}$
3.  $S(x)$ , the successor section:  $S(x) := O(x) - \{x\}$
4.  $F(x)$ , the *up*:  $F(x) := \{y \in P: x \leq y\}$

Then, the following average rank indexes are defined:

$$a) LPOM(x) = (|S(x)| + 1) \times (n + 1) \div (n + 1 - |U(x)|)$$

$$b) LPOMext(x) = |O(x)| + \sum_{y \in U(x)} \frac{P_y^<}{P_y^< + P_y^>}$$

where  $n$  is the number of objects,

$|V|$  defines the cardinality of the set  $V$ ,

$$P_y^< = |O(x) \cap U(y)|, P_y^> = |F(x) \cap U(y)|, \text{ and } y \in U(x)$$

### Lexicographic Approach

A lexicographic technique enables decision-makers to develop choice rules in which they select more items based on their most important criteria. When two objects have the same influence on the most preferred criteria, decision-makers prefer the one with the biggest impact on the second most preferred criteria, and so on, according to Saban and Sethuraman (2014). This lexicographic form simulates situations in which decision-makers have a strong preference for one criterion over another or are in charge of non-compensatory aggregation (Yaman et al., 2011; Pulido et al., 2014).

Finally, decision-makers can model their strong preferences for the criteria chosen since, after additional investigation of the situation, they are neither indifferent nor uncertain about their preferences for the criteria considered. In other words, they will always favor one criterion over another, regardless of criterion weights.



## Risk Metrics and Compliance

Risk metrics are statistical indicators or measurements that enable decision-makers to assess the dispersion (volatility) of specific events or outcomes. As a result, a random variable can be evaluated using statistical moments (e.g., mean, variance, skewness, kurtosis), or risk metrics, such as Value at Risk (VaR) and Conditional VaR, can be used to investigate extreme values (Bodie et al., 2009; Fabozzi, 2010; Matos, 2007; Mun, 2015).

Risk metrics are used to analyze the volatility or stability of a set of options or a portfolio of alternatives in decision problems, such as financial risk management (Chong, 2004), portfolio risk management (Bodie et al., 2009), enterprise risk management (Scarlat et al., 2012), and a variety of other areas (Fabozzi, 2010).

A compliance strategy, or the establishment of a set of rules to guide decision-makers, is used to evaluate how risky an object is and its interaction with other objects (Hopkins, 2011). For assessing compliance, several methodologies have been presented. Barrett and Donald (2003), for example, propose a stochastic dominance analysis to compare probability distributions before establishing a hierarchy; Boucher, Danielsson, Kouontchou, and Maillet (2014) use risk metrics and forecasting to adjust models based on historical performance; and Zanolli, Gambelli, Solfanelli, and Padel (2014) investigate the effects of risk factors on non-compliance in UK agriculture.

Because it permits evaluating whether an item performs according to decision-makers' preferences and overstated risk measures, the compliance approach is more user-friendly for decision-making. The main concept is to divide the risk spectrum into two categories (Hopkins, 2011). As a result, the higher the compliance with a stated risk metric, the closer the decision-makers' preferences are aligned. Scarlat et al. (2012) and Tarantino (2008) examine similar approaches based on important risk indicators.

## PROMETHEE and ELECTRE

Another layer of complexity emerges when decision-makers must integrate potentially conflicting decision criteria (quantitative or qualitative, monetary and nonmonetary) into project management, such as legal (taxes, compliance, social responsibility, etc.), environmental (level of pollution, noise, watershed issues, etc.), and economic (level of economic growth, monetary and nonmonetary). Furthermore, the relative significance (RI) or weights of those criteria may differ. The phrases in BP's (2015) sustainability report that businesses "must earn and keep society's support" and "must take action to assist to conserve the environment for future generations" may imply that certain decision-makers value profit over social responsibility or vice versa. As a result, it is critical to factor those variances into the decision-making process (Mun et al., 2017).

To solve this issue, multicriteria analysis (MCA) has emerged as an effective tool for dealing with multi-dimensional problems and obtaining an Aggregate Quality (AQ) that may be used to support a final decision (Bouyssou et al., 2006; Brito et al., 2010). MCA is a set of strategies, techniques, and tools that aid individuals in solving choice issues (description, grouping, ranking, and selection) by considering multiple objectives or criteria at the same time (Roy, 1996; Ghafghazi et al., 2010; Kaya & Kahraman, 2011; Afsordegan et al., 2016).

The authors propose PROMETHEE (Goumas & Lygerou, 2000; Brans & Mareschal, 2005; Behzadian et al., 2010; Tavana et al., 2013) as an appropriate MCA technique. Outranking the connection  $S$  is the basis of PROMETHEE techniques. This notion defines whether "the alternative is at least as good as the alternative  $b$ ," rather than determining whether the relationship between two alternatives  $a$  and  $b$  is a strong preference (" $a P b$ "), a weak preference (" $a Q b$ "), or indifference (" $a | b$ "; Brans & Mareschal, 2005).



Because of their theoretical and practical merits, PROMETHEE procedures are appropriate. They can, for example, assign an AQ index to each project that maximizes the available information in terms of decision-makers' preferences for the criteria chosen, as well as the intensity of those preferences among alternatives and the nature of each criterion (Bouyssou et al., 2006). Many energy-related studies have used PROMETHEE methods, including sustainable energy planning (Pohekar & Ramachandran, 2004; Cavallaro, 2005); renewable energy alternatives (Georgopoulou et al., 1997); heating system options (Ghafghazi et al., 2010); and oil and gas pipeline planning (Tavana et al., 2013); and oil and gas pipeline planning (Behzadian et al., 2010).

There are other approaches, such as the ELECTRE methodologies (Bouyssou et al., 2006), the Analytical Hierarchy Process (AHP; Desai et al., 2012; Saaty, 2013), MACBETH (Cliville et al., 2007; Costa et al., 2012), and TOPSIS (Kaya & Kahrama, 2011). These alternative approaches, on the other hand, do not clearly describe the aforementioned benefits, and the AQ is harder to read.

Although some studies have attempted to incorporate real options (RO) into MCA (Cavallaro, 2005; Angelou & Economides, 2008; Tolga & Kahraman, 2008; Zandi & Tavana, 2010; Tolga, 2011, 2012), there is little evidence of an integrated RO-MCA methodology for ranking a portfolio of projects in state-owned energy companies that pursue nonfinancial objectives.

According to the author, while RO values and assesses flexibility and uncertainty for PM, MCA allows for the inclusion of additional factors such as GDP and employment in strategic planning criteria to produce an AQ for picking the best projects.

## References

- Asbeck, E., & Haimes, Y. Y. (1984). The partitioned multiobjective risk method. *Large Scale Systems*, 6(1), 13–38.
- Barrett, G. F., & Donald, S. G. (2003). Consistent tests for stochastic dominance. *Econometrica*, 71(1), 71–104.
- Behzadian, M., Kazemzadeh, R. B., Albadvi, A., & Aghdasi, M. (2010). PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, 200(1), 198–215.
- Betts, K. D., & Jaep, K. R. (2017). The dawn of fully automated contract drafting: Machine learning breathes new life into decades-old promises. *Duke Law and Technology Review*, 15(1), 216–233.
- Boucher, C. M., Danielsson, J., Kouontchou, P. S., & Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking and Finance*, 44, 72–92.
- Brans, J.-P., & Mareschal, B. (2005). Multicriteria decision aid. The PROMETHEE-GAIA solution in multiple criteria decision analysis: State of the art surveys. *International Series in Operations Research & Management Science*, 78, 163–186.
- Brans, J. P., & Vincke, P. H. (1985). A preference ranking organization method: The PROMETHEE method for multiple criteria decision making. *Management Science*, 31(6), 647–656.
- Bruggemann, R., Bücherl, C., Pudenz, S., & Steinberg, C. (1999). Application of the concept of partial order on a comparative evaluation of environmental chemicals. *Acta hydrochimica et hydrobiologica*, 27, 170–178.
- Cliville, V., Berrah, L., & Mauris, G. (2007). Quantitative expression and aggregation of performance measurements based on the MACBETH multicriteria method. *International Journal of Production Economics*, 105(1), 171–189.
- Daley, S. (2019, September 24). 19 examples of artificial intelligence shaking up business as usual. <https://builtin.com/artificial-intelligence/examples-ai-in-industry>
- Darken, R. (2019, October 21). Human-machine teaming AI. Naval Postgraduate School.
- Defense Advanced Research Projects Agency. (2019). AI next campaign. <https://www.darpa.mil/work-with-us/ai-next-campaign>
- De Loof, K., De Baets, B., & De Meyer, H. (2011). Approximation of average ranks in posets. *MATCH Communications in Mathematical and in Computer Chemistry*, 66, 219–229.
- Denning, P. (2019, September). Harnessing artificial intelligence. Naval Postgraduate School.





- DoD. (2019, February 12). Summary of the 2018 Department of Defense artificial intelligence strategy: Harnessing AI to advance our security and prosperity. <https://media.defense.gov/2019/Feb/12/2002088963-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF>
- Desai, S., Bidanda, B., & Lovell, M. R. (2012). Material and process selection in product design using decision-making technique (AHP). *European Journal of Industrial Engineering*, 6(3), 322–346.
- Dorini, G., Kapelan, Z., & Azapagic, A. (2011). Managing uncertainty in multiple-criteria decision-making related to sustainability assessment. *Clean Technologies and Environmental Policy*, 13, 133–139.
- Eykholt, K., Evitimov, I., Fernandes, E., Li, B., Rahmati, A., Xia, C., & Song, D. (2018). Robust physical-world attacks on deep learning visual classification. Cornell University.
- Gardner, H. (1993). *Multiple intelligences*. Basic Books.
- Ghafghazi, S., Sowlati, T., Sokhansanj, S., & Melin, S. (2010). A multicriteria approach to evaluate district heating system options. *Applied Energy*, 87(4), 1134–1140.
- Golstein, B. (2018, October 10). A brief taxonomy of AI. *Sharper AI*. <https://www.sharper.ai/taxonomy-ai/>
- Greenfield, D. (2019, June 19). Artificial intelligence in medicine: Applications, implications, and limitations. Harvard University. <http://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/>
- Goumas, M., & Lygerou, V. (2000). An extension of the PROMETHEE method for decision making in fuzzy environment: Ranking of alternative energy exploitation projects. *European Journal of Operational Research*, 123(3), 606–613.
- Gunning, D. (2017, November). Explainable artificial intelligence. DARPA. <https://www.darpa.mil/attachments/xaiprogramupdate.pdf>
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On past, present, and future of AI. *California Management Review*, 61(4), 5–14.
- Haimes, Y. Y. (2009). *Risk modeling, assessment, and management* (3rd ed.). John Wiley & Sons.
- Icertis. (2019a). Icertis customer profile: Mindtree. <https://www.icertis.com/customer/mindtree/>
- Icertis. (2019b, April 2). Microsoft streamlined its contract management. <https://www.icertis.com/customers/microsoft-information-exchange-agreements-case-study/>
- Institute for Robotic Process Automation & Artificial Intelligence. (2019, November 16). What is robotic process automation? <https://irpaai.com/what-is-robotic-process-automation/>
- Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
- Kaya, T., & Kahraman, C. (2011). Multicriteria decision making in energy planning using a modified fuzzy TOPSIS methodology. *Expert Systems with Applications*, 38(6), 6577–6585.
- King, A. D. (2019, November 16). Talk to a transformer. <https://talktotransformer.com/>
- Knight, W. (2017, April 11). The dark secret at the heart of AI. *Technology Review*. <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>
- Lawgeex. (2018, February). Comparing the performance of artificial intelligence to human lawyers in the review of standard business contracts. *Law.com*. <https://images.law.com/contrib/content/uploads/documents/397/5408/lawgeex.pdf>
- Leviathan, Y. (2018, May 8). Google Duplex: An AI system for accomplishing real-world tasks over the phone. *Google AI Blog*. <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html>
- Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, 77–91.
- Mun, J. & Anderson, M. (2021). Technology trust: System information impact on autonomous systems adoption in high-risk applications. *Defense Acquisition Research Journal*, 28(1), 2–39. <https://doi.org/10.22594/10.22594/dau.19-841.28.01>
- Mun, J. (2015a). *Modeling Risk: Applying Monte Carlo risk simulation, strategic real options, stochastic forecasting, portfolio optimization, data analytics, business intelligence, and decision modeling* (3rd ed.). Thomson-Shore and ROV Press.
- Mun, J. (2015b). *Readings in certified quantitative risk management (CQRM)* (3rd ed.). Thomson-Shore and ROV Press.
- Mun, J. (2016b). *Real options analysis: Tools and techniques for valuing strategic investments and decisions with integrated risk management and advanced quantitative decision analytics* (3rd ed.). Thomson-Shore and ROV Press.
- Mun, J. (2016a). *Real options analysis* (3rd ed.). Thomson-Shore and ROV Press.
- Mun, J. (2016c, October). Empirical cost estimation tool [Paper presentation]. Naval Acquisition Research Conference, Monterey, CA.
- Mun, J. (2019). Empirical cost estimation for U.S. Navy ships. *Universal Journal of Management*, 7, 152–176.
- Mun, J., & Housel, T. (2010). A primer on applying Monte Carlo simulation, real options analysis, knowledge value-added, forecasting, and portfolio optimization. Calhoun.



- Mun, J., Ford, D., & Housel, T. (2012, October 10). Naval ship maintenance: An analysis of the Dutch shipbuilding industry using the Knowledge Value Added, Systems Dynamics, and Integrated Risk Management methodologies (NPS-AM-12-204). <https://calhoun.nps.edu/bitstream/handle/10945/33851/NPS-AM-12-204.pdf?sequence=1>
- Mun, J., George, K., & Ledbetter, E. (2020). Total ownership with lifecycle cost model under uncertainty for surface ships' electro-optical-infrared-sensors [Unpublished manuscript].
- Mun, J., Hernandez, E., & Rocco, C. (2016). A combined lexicographic average rank approach for evaluating uncertain multi-indicator matrices with risk metrics. In M. Fattore & R. Brüggemann (Eds.), *Partial order concepts in applied sciences*. <http://www.springer.com/la/book/9783319454191>. eBook ISBN (978-3319454214)
- Mun, J., Housel, T., & Wessman, M. D. (2010). PEO-IWS ACB insertion portfolio optimization. *Proceedings of the Seventh Annual Acquisition Research Symposium*, 2 (No. NPS-AM-10-069-VOL-2). Naval Postgraduate School. <https://my.nps.edu/documents/105938399/108624025/NPS-AM-10-069.pdf/c71c6830-853a-448b-beac-242bea4c8b?version=1.0>
- Nayak, P. (2019, October 25). Understanding searches better than ever before. Google. <https://blog.google/products/search/search-language-understanding-bert>
- Oppy, G., & Dowe, D. (2016, February 8). The Turing test. In E. N. Zalta (ed.), *Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/entries/turing-test/>
- Parloff, R. (2016, September 28). From 2016: Why deep learning is suddenly changing your life. *Fortune*. <https://fortune.com/longform/ai-artificial-intelligence-deep-machine-learning/>
- Pulido, F. J., Mandow, L., & de la Cruz, J. L. P. (2014). Multiobjective shortest path problems with lexicographic goal-based preferences. *European Journal of Operational Research*, 239(1), 89–101.
- Raghaven, S., & Mooney, R. J. (2013). Online inference-rule learning from natural-language extractions. The University of Texas.
- Restrepo, G., Brüggemann, R., Weckert, M., Gerstmann, S., & Frank, H. (2008). Ranking patterns, an application to refrigerants. *MATCH Communications in Mathematical and in Computer Chemistry*, 59, 555–584.
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21–41.
- Roy, B. (1968). Classement et choix en presence de points devue multiples (la methode ELECTRE). *Revue d'Informatique et de recherche opérationelle*, 6(8), 57–75.
- Saban, D., & Sethuraman, J. (2014). A note on object allocation under lexicographic preferences. *Journal of Mathematical Economics*, 50, 283–289.
- Sammalkorpi, S., & Teppala, J. P. (2019). AI in procurement. Sievo Oy.
- Scarlat, E., Chirita, N., & Bradea, I. A. (2012). Indicators and metrics used in enterprise risk management (ERM). *Economic Computation and Economic Cybernetics Studies and Research*, 46(4), 5–18.
- Shanahan, P. (2018). DOD cloud strategy. Department of Defense.
- Shaw, M. (2019, October 15). Why Google is the best search engine (and why businesses should care). *Tower Marketing*. <https://www.towermarketing.net/blog/google-best-search-engine/>
- Sievo. (2019, November 16). AI in procurement. <https://sievo.com/resources/ai-in-procurement>
- Tarantino, A. (2008). *Governance, risk, and compliance handbook: Technology, finance, environmental, and international guidance and best practices*. John Wiley & Sons.
- Tavana, M., Behzadian, M., Pirdashti, M., & Pirdashti, H. (2013). A PROMETHEE-GDSS for oil and gas pipeline planning in the Caspian Sea basin. *Energy Economics*, 36, 716–728.
- Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton University Press.
- Yaman, F., Walsh, T. J., Littman, M. L., & Desjardins, M. (2011). Democratic approximation of lexicographic preference models. *Artificial Intelligence*, 175(78), 1290–1307.
- Zanoli, R., Gambelli, D., Solfanelli, F., & Padel, S. (2014). Assessing the risk of non-compliance in UK organic agriculture. *British Food Journal*, 116(8), 1369–1382.
- Zarkadakis, G. (2019, September 11). The rise of the conscious machines: How far should we take AI? *Science Focus*. <https://www.sciencefocus.com/future-technology/the-rise-of-the-conscious-machines-how-far-should-we-take-ai/>



# You Can't Wait for ROI to Justify Model-Based Design and Analysis for Cyber Physical Systems' Embedded Computing Resources

**Fred Schenker**—works in the SEI's Software Solutions Division and has worked there for over 20 years. He works to improve software acquisition and product development practices throughout the armed services and other organizations. He has actively worked in software process, architecture, model-based systems engineering, and metrics. Before joining the SEI, Schenker spent over 20 years in industry as an active contributor in all phases of product development. Schenker is also an inventor and has obtained patents for a pressure switch (used in automotive airbag applications) and for a manufacturing process to seal gas inside a vessel. [ars@sei.cmu.edu]

**Jérôme Hugues**—is a Senior Researcher at the Carnegie Mellon University/Software Engineering Institute in the Assuring Cyber-Physical Systems team. He holds a Habilitation à Diriger les Recherches (HDR, 2017), a PhD (2005) and an engineering degree (2002) from Telecom ParisTech. His research interests focus on the design of software-based real-time and embedded systems and tools to support it. More specifically, he concentrates on software architecture to support the design of complex software-based real-time and embedded systems, and programming languages and artifacts to support them. He is a member of the SAE AS-2C committee working on the AADL since 2005. Prior to joining the CMU/SEI, he was a Professor at the Department of Engineering of Complex Systems of the Institute for Space and Aeronautics Engineering (ISAE), in charge of teaching curriculum on systems engineering, safety-critical systems, and real-time systems. [jhugues@sei.cmu.edu]

## Abstract

The practical, pragmatic benefits of building early architectural models of the embedded computing resources for Cyber Physical Systems (CPS) have been documented and demonstrated. However, the rate of adoption of this practice by the contractor community has been slow. Empirically, we have observed skepticism with respect to the increased cost of building these models, as being of sufficient value to justify their expense. This paper elaborates the reasons why using traditional methods, such as return on investment (ROI), to justify the increased expense (of building and maintaining these virtual models) is inadequate. Alternate ways to quantify and rationalize the benefits are discussed, but ultimately the decision to adopt may require a leap of faith.

We begin by describing the problem space and advancements in the design and implementation of the embedded computing resources for CPS. We discuss the proposed process change we seek: using model-based methods to reduce integration and test risk. We discuss the potential effects of that change on CPS, as well as our thoughts on ROI and the issues that can arise when using ROI. Finally, we recommend how organizations can move forward with a model-based approach in the absence of solid ROI data.

## Introduction

Technological advancements have been suggested, supported, funded, and incorporated into our lives since the earliest times, and we continue to identify opportunities to push our technological envelope. Frequently, technological innovations are identified to produce incremental improvements; less frequently the improvement forces a change to the “way we do business.” This paper focuses on a particular technological improvement, one that will have a significant impact on the “performance” of a project (i.e., cost and schedule) to build, test, and sustain a cyber-physical system (CPS).

The improvement we are advocating is that projects build virtual architectural models early in the system development lifecycle of the CPS they are developing. These projects must explicitly define the embedded computing resources within these models and use the models as



early as possible to validate the systems' requirements. In particular, the models should be used to identify embedded computing system constraints so the project can manage the constraint as a risk. To contrast this with our observations of current CPS development projects, we find that the constraints are not identified as risks, requiring mitigation, until much later in the lifecycle, typically when the components are being integrated in a lab. The late validation of the constraint usually causes big problems, e.g., significant cost overruns, schedule delays, and/or compromises in capability.

In the development environment of the future, early architectural models of the CPS, coupled with model-based analysis methods are applied iteratively and recursively from requirements analysis, through product design, and virtual integration and test. As design decisions are made, the architectural model fidelity is increased, enabling more accurate estimates of computing resource performance. Eventually, the practical application of the model is replaced by physical hardware and software in a laboratory environment, i.e., a Systems Integration Laboratory (SIL), but the architectural models are kept up to date as issues are found and resolved. After the CPS is completed, the models are maintained and are used to assess the impact of potential changes, possibly as part of a system upgrade.

By the time you reach the end of the paper, you should have a good idea of the practical, pragmatic benefits of building early architectural models of the embedded computing resources for CPS. You should also understand the challenges that arise when trying to use ROI to justify the increased expense of building and maintaining these virtual models. The paper elaborates the reasons for this assertion.

## The Problem Space

It is apparent that CPS are getting more and more complex. The software that is incorporated within the system gets to be a bigger part of the overall technical solution, and the number of physical parameters that are monitored and controlled by the system contribute to this complexity. This increased complexity results in potentially hard-to-predict behavior, as it is very difficult to understand the suitability of a proposed system solution without a deep understanding of how its computing resources are going to be used (as part of the overall technical solution). We might think a solution is adequate, we start the implementation, and then we find issues. Typically, these issues are surfaced at the end of product development, as the components are being integrated, and the system is being tested. As time goes on (i.e., after deployment) the complexity only gets worse, making it more and more difficult to predict the impact of a change made as part of incremental updates or modernization efforts.

At the same time, these CPS development organizations have been slow to adopt a key potential process improvement, to develop a virtual architectural model with associated analysis tools that would help them deal with this very problem. This observation is evident from the cost and schedule issues that arise late in the development process (e.g., integration and test) for virtually every Department of Defense (DoD) cyber-physical system built in the last 30 years. We continue to claim that we “do the best we can,” but in the end our efforts fall short, and system after system fails to meet its expectations. We are not doing the “best we can.” We are doing the same thing we have been doing for the last 30 years...and we can do better.

One extremely common reason for not adopting newer methods (initially) is that there is a perceived need to prove that the new way is better than the old way. We have all heard people say things such as, “Better the devil you know...,” or “The grass is always greener...” The decision makers that need to advocate to move forward with a process improvement can always find ways to delay, “Bring me a rock. No. Bring me a different one.” The patterns of behavior for individuals and organizations are well-established. Everett Rogers' (2003) theory, *Diffusion of Innovations*, is a widely accepted model for characterizing this behavior (Rogers,



2003). Terms such as *early adopters* and *laggards* are well known and used by many in industry.

In other domains, model-based design and analysis has been employed by engineers for centuries. For example, improvements in structural analysis were facilitated by work done by da Vinci, Newton, Euler, among many others, leading ultimately to the development of the *finite element method* for predicting stress in structural components in the mid-1950s. The model-based theories were incorporated into software tools, starting in the 1960s and continuing to this day. It is common practice now for mechanical engineers to use finite element models to help improve the quality of their designs, and to provide an element of verification not available prior to the development of the finite element models and analysis methods. The modeling tools are used iteratively as part of the design process to optimize and to reduce elements of design risk.

Bridges collapsing or rockets exploding on the launch pad are graphic images of design failure, and the need to prevent future occurrences can rise to the top of a nation's agenda. As such, it is unlikely that a need to demonstrate financial ROI existed back when we were developing the finite element tools to prevent these very public failures. In these cases, elimination of the public disgrace was the benefit, and the community was fine with that.

To contrast this with embedded computing resources for CPS, the design failures are invisible until the physical devices are connected. As mentioned above, this may be due to increased system complexity. Even when the devices are connected, and the system is not working the way it was supposed to, it may take quite a while to investigate and determine the root cause of the problem. Because these problems are surfaced in the lab, the cost to deal with them can be up to 80x more than what it might have been if caught during design. Models representing the CPS and specifically the CPS' embedded computing resources can be used during early lifecycle stages to predict these issues (or constraints) and can also be used to evaluate alternate designs that might mitigate the future problems.

We do not live in a perfect world. Models can be incomplete, assumptions can be made that are incorrect, boundary conditions may not be represented accurately. All these things might affect the quality of the models, analyses, and resulting designs. If the model-based methods produce an incorrect answer, for whatever reason, the development may go off the rails. However, we do not design systems assuming that the designers and engineers are going to make mistakes. We trust our teams and their processes to produce high quality designs. When we improve our design process, it is usually because we have identified new methods that enhance our understanding of the problem space. Who would argue that properly using models and analytical methods to provide higher fidelity design verification could possibly be more expensive than not doing so? The work to create, verify/validate, and apply the models will cost more than not doing so, but when problems and computing resource constraints are found early in the development process, we will avoid more expensive rework of the system, and possibly provide enhanced capability for new implementations. Better design will cost more.

## Cyber-Physical Systems

Cyber-Physical Systems (CPS) are pervasive in DoD systems; their capability is extended by their embedded systems. Embedded systems are made of software and hardware sub-systems and integrated into a larger system, e.g., in charge of monitoring and controlling a physical process such as the trajectory of a vehicle. They are often associated with real-time or safety non-functional requirements: providing a function that must be completed under time constraints, (e.g., respect of deadline, periodicity, etc.) while ensuring safety invariants, e.g., avoiding unsafe situations that would create an unbearable risk to the system or its environment. CPS adds extra complexity to the system because of the greater degrees of



coupling between computations and physical processes and were first recognized by the NSF as part of an emerging field of research in late 2006 (NSF, 2010).

Because of the interleaving between physics and computer sciences concerns, there is hardly a single state-of-practice for engineering CPS: understanding the system concept of operations and high-level requirements is key to narrowing down the engineering body of knowledge. For instance, controlling a swarm of UAVs will rely on control theory and flight dynamics in addition to wireless communication stacks and distributed algorithms, whereas the definition of a robot operating along with human operators will rely on mechatronics, inverse kinematics along with stringent design methods for real-time safety-critical systems.

Hence, CPS engineering is deeply connected with both Systems Engineering approaches for capturing concept of operations and high-level requirements, methods for architecting systems and specific analysis methods. Industry standards have been developed to facilitate the development of CPS such as simulation techniques to validate a system or Digital Twin to monitor a system as it is being deployed (Bickford et al., 2020).

Although these approaches support the engineering of CPS, they do not address the diversity of analysis methods required. As a response, Model-Based Design and Analysis has been suggested as a discipline of its own to support the broad need to address performance, safety, security, or behavioral analyses of a system.

## **Model-Based Design and Analysis**

### **Model-Based Design**

Model-based design, or model-based systems engineering (MBSE) is a key aspect of the DoD's press for digital engineering. Models have been used for centuries to provide an environment to predict the performance of products under development. We do not need to look back very far to see evidence of how models have helped improve our lives.

The evolution of the rail, used for railroads, has seen significant improvement from model-based design and analysis. Early rails were developed for horse-drawn wagons, and they were built using wooden rails, eventually replaced by cast iron. When the steam engine was introduced to power the transport, the increased weight of the locomotive caused significant breakage; the rails were too weak and brittle. This led to huge costs for maintaining the rails. Without having models to represent the material properties of different metals, we may never have gotten to the current standard for steel. Engineers were able to (1) increase carbon content to improve tensile strength and reduce ductility, (2) increase manganese content to reduce abrasion, and (3) reduce phosphorus and sulfur impurities to improve the brittleness of the rail. All these improvements led to the creation of standards for the material content of the rail, specified by the American Railway Engineering Association (Walsh, 1909).

Without the material science, leading to innovations in the actual rail material, many of the advances of the industrial revolution would have been attenuated. This is because one of the key enablers of the industrial revolution was the means to efficiently transport these products. Necessity was the mother of this development, and it involved many different innovations. However, individuals and companies sponsored the innovations. There was a competitive nature that spurred on the innovators. The innovators did not know for sure, when new rail material was laid for the first time, whether this composition would work better than the previous, but as time went on, the innovators were able to get more and more precise with the chemical composition of the steel, the processes used to produce it, and the methods by which the rails are joined. Now we have standards and predictable results.



An important observation to be made is that even though the models (i.e., the science) predicted the performance, the rail companies had to verify and validate the rail's performance before committing to large scale production. In practice, this not only provides confidence in the models, but it also provides the environment to understand why the models do not work as expected. This feedback loop improves the model as it incorporates the lessons learned from the evaluation. This is true for all model-based development, and we do not learn about these things if we do not build the environment to develop the models and test them.

Looking back at the evolution of the rail, what became of the companies that did not invest in the efforts to build better rails? No doubt that there were early adopters that benefitted greatly from the material science advancements and built these practices into their development process. There were others that did not fare so well.

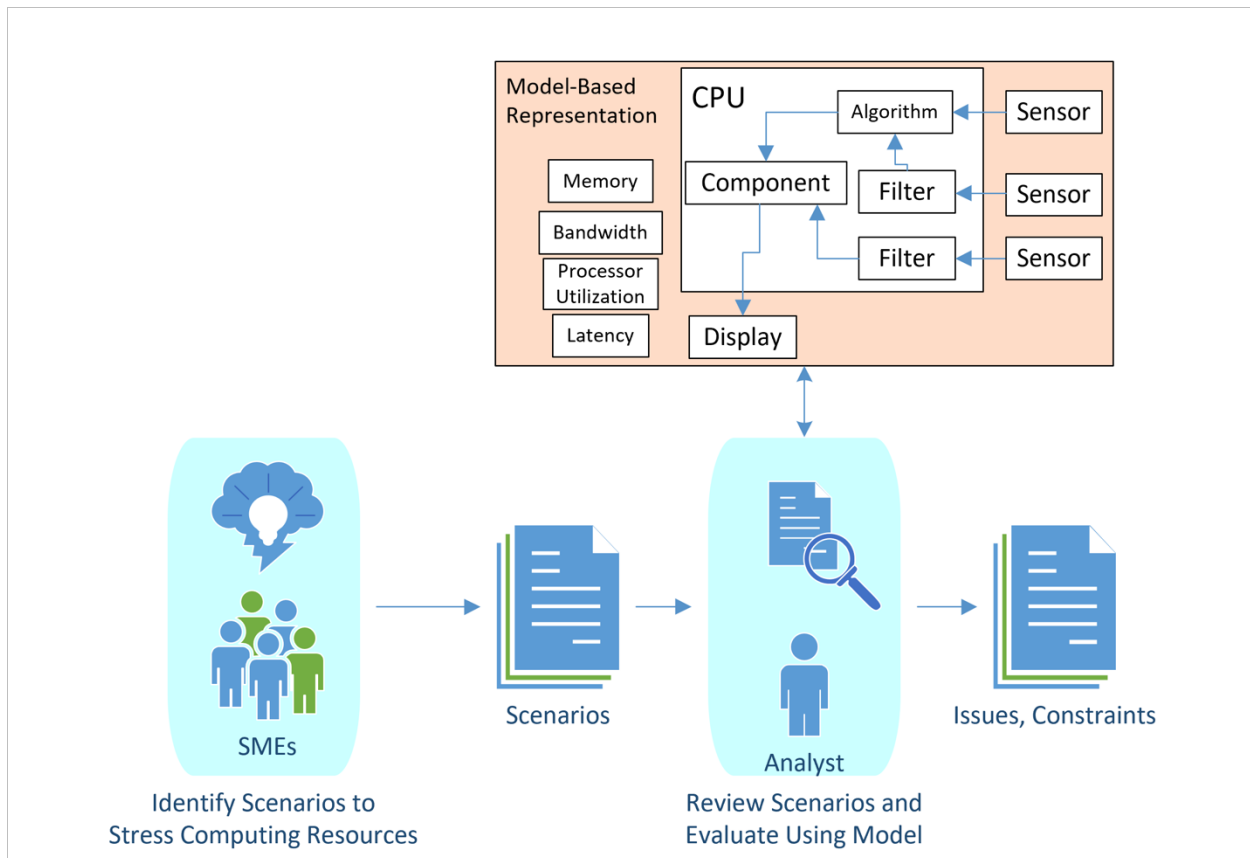
In our current context, models are widely used for all sorts of different applications. In some engineering domains, they are the trusted and authoritative sources of truth for the design. The introduction of the model-based methods in such fields as mechanics, thermodynamics, electromagnetic spectrum, electrical engineering, logistics, maintenance, process optimization, and manufacturing have had a transformative effect on the "way we do business." They are used to document critical design decisions, illustrating graphically the choices that were available to the designers, the design selection criteria, and the rationale for selecting the winning design. In some of these domains, the digital engineering environment is used to transfer model-based designs into analysis environments, and use the results to verify performance characteristics, or conduct model-based what-if analyses. A good example of this is mechanical design, using a model-based 3D computer-aided design (CAD) tool, and analyzing stresses and other aspects using finite element analysis (FEA).

In DoD CPS design, the move towards MBSE is progressing in the right direction. However, there is room for improvement. Many organizations are not working natively in the MBSE tools. They do their work outside the MBSE environment, then "document" the resulting design in the MBSE environment. To unlock the true potential of MBSE, we need to build the system models and the associated analysis environment, as has been done in the other domains, and use the digital environment organically to test design ideas and build quality in.

## **Model-Based Analysis**

As stated in the section above, model-based analysis is how we leverage the investment in model-based design. Without analysis methods, the properties of new rail material compositions would not be evaluated until the rail was laid. Imagine the world where we didn't do analysis as an integral part of the design process.





**Figure 30. Notional Model-Based Analysis Process**

As shown in Figure 1, in a mature model-based design environment, subject matter experts (SMEs) are used to identify design stressors to uncover elements of weakness in the design and then to build an environment to evaluate designs as they evolve. A simple example is the use of a wind tunnel used to assess air drag in the design of a performance car. Using the analysis environment, the time needed to create a final design can be significantly reduced. With more experience and validation of the analysis methods and tools, the designers and engineers learn to rely on them to provide the early performance prediction needed for design verification.

Once constraints have been identified, they are managed by the design team. Having an environment available to evaluate scenarios that stress the constraints is an essential element to predict product performance. Many times, it is possible to identify unintended consequences of design decisions by using the analysis tools and facilities. These issues, if identified later in the design process (because of the lack of analysis capability or poor implementation of the design constraint) typically have a much more significant impact on the project. The economic case for investing in processes based on early defect detection was argued by Feiler et al. (2013) and Hansson et al. (2018). In short, they present evidence that in the domain of embedded safety-critical systems, 35% of the errors are introduced in requirements, and 35% of the errors are injected in architectural design. Nonetheless, 80% of all errors are not discovered until system integration or later. Figure 2 is a graphic depiction of the impact of the late discovery.



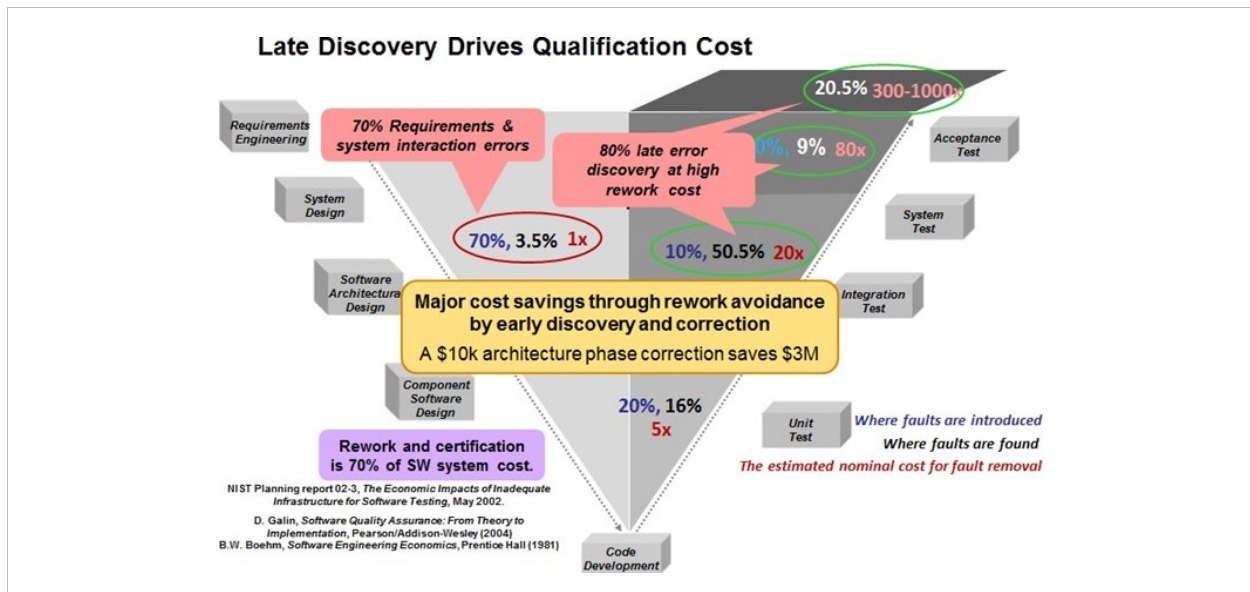


Figure 31. Gap Between Defect Origin and Discovery (Feiler et al., 2013)

Normally, there will be opportunities to trade off critical constraint allocations to meet overall system requirements. There may even be options to increase the capacity of the constrained resource if the impact of the constrained resource will be too severe. The point is that the analysis capability provides the design team with early insight into the product performance, allowing the design team to improve their management of the technical risk.

For DoD CPS, the physical aspect of the equipment constrains the overall analysis. In addition, the DoD acquisition timeline adds additional turmoil. To contrast, non-DoD CPS (e.g., automotive manufacturers) generally come out with new models every year, so last year's analytical tooling needs only minor modification to work for this year's model. It is also generally the case that last year's models functioned properly, constraints are known and planned for. The DoD acquisition timeline and the systems engineering process (methodical and rigorous, but also following a waterfall approach) mean that by the time requirements have been allocated to components, it may be too late to make some of the changes that would be needed as part of the management of a technical constraint such as an emerging need. It is important for the development team to have an analysis capability throughout the systems engineering processes to help with all those critical systems engineering decisions.

Analysis can be performed to support multiple domains (e.g., performance, safety, security). The different analyses could be performed using similar (or the same) models. In some cases, the analysis performed in support of one domain may conflict with a different domain. Managing the design and development using analytical tools should provide higher levels of design assurance and fewer issues during integration and test. In this context, the Software Engineering Institute (SEI) has been the technical lead of the Architecture Analysis and Design Language (AADL) SAE-AS-5506 standardization effort. AADL provides the foundations for the precise analysis of safety-critical CPS (SEI, 2023). A pilot study by the Aerospace industry consortium Aerospace Vehicle Systems Institute at Texas A&M under the System Architecture Virtual Integration (SAVI) in 2008 chose AADL as primary candidate to address the Embedded Software System affordability problem (Feiler et al., 2009). DARPA HACMS (2012-17) used AADL as part of their MBSE toolkit to use formal methods to build embedded computing systems that are resilient against cyber-attack because they have been proven not to have typical security vulnerabilities.



## DoD Digital Engineering Strategies

In recent years, Congress mandated that the DoD adopt a Modular Open Systems Approach (MOSA) to systems development, directing procurement officials to pursue modularity in CPSs to reduce costs across families of systems.<sup>1</sup> The Air Force published an agile software strategy to reduce software integration costs across platforms (Roper, 2020). The Army spearheaded development of the Architecture Centric Virtual Integration Process (ACVIP) to find cyber-physical integration errors early through virtual integration (Boydston et al., 2019).

Additionally, the DoD has heavily invested in DE. The DoD provided a framework for DE in the 2018 *Digital Engineering Strategy*, which relates five expected benefits from DE:

1. Informed decision making/greater insight through increased transparency
2. Enhanced communication
3. Increased understanding for greater flexibility/adaptability in design
4. Increased confidence that the capability will perform as expected
5. Increased efficiency in engineering and acquisition practices

The Digital Engineering (DE) Strategy (DES) calls for practitioners to “Establish accountability to measure, foster, demonstrate, and improve tangible results across programs and enterprise” (DoD DES, 2018). Recent efforts such as the Joint Multi-Role (JMR) Comprehensive Architecture Strategy (CAS) provide a framework for measurement by formalizing relationships between Key Business Drivers (KBDs), Key Architecture Drivers (KADs), and Quality Attributes (QAs) for a cyber-physical system (CCDC, 2018). Schenker et al. (2022) called for practical measures to support achieving these DE benefits (Schenker, Smith, & Nichols, 2022). Without methods to measure DE effectiveness, particularly model-based analysis, it will be difficult for new programs (e.g., the Army’s Future Vertical Lift efforts) to gauge whether they are on track to reap the benefits of DE.

## Return on Investment

The essential concept to understand regarding return on investment (ROI) is that when an organization commits to invest in something, there ought to be a financial justification for the investment (i.e., the investment should improve some aspect of the product). Sometimes the investment is long term (e.g., research, and it is not clear how the research will be applied). More frequently the investment is tactical, with a specific target in mind. There are generally many opportunities for investment, and one of the most important criteria in ranking the opportunities is the perceived value of the benefit. Built into the ROI calculus, there must be a discrete thing that the organization is trying to improve (e.g., time to market, cost to produce, quality). The improvement may not result in a direct financial benefit, but the investing organization will recognize that the improvement is desirable for the business. For example, reducing the time to bring a new product to market may enable greater market share. The key points are:

- The benefit may not be easily dollarized because the financial return is indirect.
- Imputing a return introduces *subjectivity* into the ROI calculation.
- The benefits that are not *directly* financial (e.g., lead time, market share, efficiency) are not made explicit.

Investments do not always pan out, or, more likely, do not meet the original goals of the investment proposal. There are predictable reasons for this, A change in a process may require

---

<sup>1</sup> See NDAA 2021, section 804.B.iii.



training, and the organization may experience a “learning curve” (i.e., a dip in productivity that eventually is overcome as the staff learns the new methods). A change in technology may be short lived, as there may be portions of the underlying technology that continue to evolve, requiring further investments and possibly eroding the potential benefits.

Some organizations adopt a different approach when it comes to investments in process or people, in which the investments are made incrementally, or continuously. These types of situations typically require instrumentation to measure the impact of the change. The culture of continuous process improvement is therefore more data-driven and requires a fair amount of consistency in the type of task being performed. In these situations, the actual ROI for each increment may not be measured, but rather the entire program may be evaluated over time. The organizations that adopt this approach also need to establish the culture that supports this type of methodology. Individuals need to be willing to change and adapt as the methods evolve.

In the context of DoD CPSs, we have observed regularly that systems fail or are constrained unexpectedly when entering integration and test. One goal that ought to be non-controversial is that an ROI goal for developers of DoD CPSs is to mitigate the impact of this inevitable occurrence. This could be achieved in several ways:

- Earlier identification of the constraints would allow for mitigation strategies to be planned and executed.
- Early identification and correction of defects/issues would improve the overall quality of the system, reducing the likelihood of significant amounts of defects to be found at integration and test.

What is the magnitude of this benefit? Prior research has cited that cost overruns, schedule delays, and technical compromises have a significant negative impact on these CPS programs. Even with the extra investment that we make to finish the programs, we often find that we wind up with *good enough* instead of *what we wanted*. Then, because the requirements have been *paid for*, we must accept that all the requirement implementations that exist are what we wanted, no matter how poorly they are implemented. When future changes are proposed to make the requirement what we want, the objection is that what you got was good enough, and we don't want the taxpayer to pay twice for the same capability.

## The Opportunity for Cyber Physical Systems

As cited in the prior paragraph, our experience is that our inability to discover issues and constraints until we perform integration and test, coupled with the correction of the issues found during these activities, is almost certain to cause program delays and cost overruns. We could just accept this “meta-physical certainty” and adjust our budgets and schedules to account for this. However, we don't. Time and time again, we claim that (1) we know what the issues are, (2) we have the best people on the job, and (3) we have learned from prior experience. We go in with rose-colored glasses. Then, it happens again, and we find ourselves in the middle of another acquisition nightmare. Note that this occurs both with new acquisitions and with upgrades of existing systems. It could easily be argued that the magnitude of the impact on the legacy system upgrade is more significant than new system acquisitions... we all agree that it costs much more than it should to upgrade our legacy systems.

It is important to note types of issues we find during integration and test (we do not claim to be an authoritative source for all systems). We find that:

- There are basic incompatibilities between the components that comprise the system, usually connected through the infrastructure of the system. These are most frequently caused by incomplete or inadequate interface descriptions. It is not that we have not



reviewed the interfaces for completeness, it is much more the case that a critical element needed to achieve a process cycle time requirement cannot be achieved because the timing of the element was not (or was incorrectly) specified.

- Something unexpected happens when we connect the components together. These types of issues are difficult to predict. Systems are becoming more and more complicated with data being needed in many different areas of the systems.
- There are computing resource constraints that limit the system capability, especially when the system is under load. These will typically present themselves as latencies (i.e., operations will take longer to perform than normal). Occasionally, memory or data bus issues will also constrain system performance or cause unexpected system behavior.

Concurrently, our experience is that most DoD contractors (that we have observed) are not taking advantage of model-based methods to address the root causes of these late-breaking issues. Specifically, we are not finding models of the computing resources being used to assess the adequacy of the planned computational, memory, and bandwidth loading. We do see that the analyses are being performed, but either they are not model-based, or the models are not kept up to date as the system evolves. The use of an architectural model to provide early assessments of computing resource performance issues ought to be at the top of the list of DoD CPS process improvements. Model-based methods are well established in many other engineering domains for similar purpose, but not in this one.

There are many possible reasons for their reluctance to embrace new technology. The most common objection we have experienced is that the modeling and analysis effort is somehow redundant and not necessarily as effective as more traditional methods. The detractors seek conclusive data that demonstrates the ROI. This is very difficult to provide, currently.

In the development environment of the future, we envision that integration and test engineers build a virtual environment to assess the state of development from Day 1, refining and elaborating the model(s) (as the designs are matured), but always able to answer fundamental questions about the system performance, safety, security, modularity, or any other relevant quality attribute. Perhaps the initial models are primitive and incomplete; however, the virtual environment will still be able to provide an early verification & validation (V&V) check on the systems engineering processes: requirements analysis, functional design, and allocation. The systems engineers would either use the environment themselves, or they would reach out to the integration and test engineers to conduct what-if analyses. The results of the analyses would get documented in the system design.

## **The “Problem” With ROI for Model-Based Analysis**

Changing the way we do business is difficult. Changes are disruptive and require commitment. Commitment is needed from both the management and the technical staff. Prior research has shown the types of barriers that exist for situations such as this, and strategies have been developed to manage change (e.g., *The Lippitt-Knoster Model for Managing Complex Change* (Lippitt & Knoster, 1987)). The management must commit to this change wholeheartedly. They need to accept the responsibility for the change, create the environment that would make the change successful, and not back down in the face of opposition. The technical staff need to commit to the new way of producing work products. They need to be flexible with their personal process, and participate actively as a process performer, suggesting changes as appropriate to make the process better.

When deciding whether to employ model-based methods as part of the organization’s culture, there are several ways to rationalize the change. All of these (described below) will



have potential plusses and minuses. At the end of the day, management will need to decide the path forward. It has become customary to assess change using some data, such as ROI, to reinforce the decision to move forward. We don't think that it is possible to use ROI to justify the change to incorporate model-based methods and will explain our rationale in this section. *What cannot be ignored, through all the discussion that follows, is that model-based methods have been successfully employed in virtually every domain where they have been introduced.*

## Creating an ROI Experiment

Wouldn't it be nice if there was a documented study that showed how to use the model-based methods to improve our process? It's not so easy. The following discussion (summarized below) makes several points:

- The DoD acquisition lifecycle is so long. By the time we get to integration and test we can't remember what we found during requirements analysis or other early reviews.
- Teams of developers will not have the same skill sets. Trying to set up an experiment to compare "apples with apples" will be challenging.
- We need to acknowledge that the organization would still be learning how to apply the new technology while it was conducting the study.
- Determining what to measure may vary by organization. Different organizations will characterize benefits in different ways.

A realistic way to evaluate a proposed process against an established process would be to pilot the new one, iterate to establish a reasonably repeatable process, and then do some sort of side-by-side comparison of the two approaches. This is straightforward for normal process improvement, where the cycle time might be measured in weeks or months. In the context of DoD CPS, the time between early lifecycle work and integration can be five years. Five years, for the duration of an experiment, is a long amount of time. There would be so many opportunities to create legitimate anomalies over such a time interval (that could wind up invalidating the results) that it seems like the experiment would quickly be dismissed, either as a success or as a failure, without the data to support the decision. Even if the scope of the experiment were sufficiently small to enable a quick result, how likely would it be for the process evaluators to feel confident about scaling the result for a large-scale system? One of the real problems with CPS is that software is everywhere within the system, which leads to more and more complexity within the solution space. One of the most important benefits of the model-based approach is that the model will maintain relationships between the software components and predict behavior that otherwise would be very difficult to predict. How is this measured? The duration of the experiment is an issue.

In a side-by-side experiment, it is important to try to limit the variable to the process itself. We would like to try not to introduce variation, but when humans are involved, and the processes are not performed routinely, it is easy to see that the human element would be an easy way for the experiment's results to be ruled invalid. It's not just the new processes that are not repeatable. In many DoD contractor settings, processes are performed at specific times during the project lifecycle. For example, early in lifecycle, there is a need to review and validate requirements, eventually leading to the systems engineering Systems Requirements Review (SRR). The activity to review and validate the requirements is intensely done for a short amount of time (as compared to the overall project). If we were introducing a new process to use model-based methods to review, analyze, and validate requirements, a variable might be how the contractor manages their staff to keep them proficient in the process. Humans that perform the same process more frequently become more proficient and more repeatable, and the resulting process performance is easier to baseline. According to a 1993 report published by Ericson et



al., experts in a particular domain have more knowledge and experience than novices, which allows them to perform tasks more quickly and accurately (Ericsson, Krampe, & Tesch-Römer, 1993). This is because experts have more knowledge and experience that allow them to anticipate and recognize patterns that are critical for performing the task.

This leads to a discussion of a team's proficiency with a new process. It is expected that the introduction of a new type of artifact (e.g., the MBSE model and the resulting analysis capability) would take some time to stabilize the process that takes advantage of these tools. We typically would call this a *learning curve*. How long would it take for a team to become proficient with the new process? It's one thing for management to decree that "you shall use MBSE and model-based analysis." It's quite different to have figured out how to use these tools within the context of the "way we do business." We would also expect that an organization that is learning how to apply a new process would continue to tweak things over a series of pilots. Management ought to expect this and encourage incorporation of lessons learned towards the determination of how best to incorporate the new tools, although they should also expect that eventually the process will stabilize.

The means for calculating the ROI benefit will vary from organization to organization, possibly from project to project. Is the on-time delivery of capability to the warfighter more valuable than avoiding a \$500 million cost overrun and two-year delay in schedule? Our experience is that the model-based methods will support either goal, but the development organization must decide what their benefit is going to be. Market share, for example, contributes to top line, increased revenue. ROI is a more complicated bottom line calculation. Other benefits might include flexibility and agility or aid of reuse on other products.

We need to develop objective criteria each time we try to apply the model-based methods. We should prepare to accept that the criteria may change from group to group, although there should be some commonality. For example, the main benefit to the project from adopting model-based methods ought to be that there is significantly less rework required during integration and test. How this is measured, either by effort or schedule or number of issues found (or some combination), ought to be expected. What else they might measure, possibly as leading indicators, is left to the process improvement teams.

## How Can You Count Defects That Aren't There?

The primary benefit we have been describing in this paper is that using models and analysis methods earlier in lifecycle will lead to fewer issues later in lifecycle. A conundrum for the calculation of this benefit is that if the model-based methods were so effective at identifying constraints and issues, many of the expected defects would not be present during integration and test. We wouldn't know, except by making assumptions, what the effect of the model-based methods were. Although we have evidence of the magnitude of the problem (from observation of prior project performance), it could be claimed that "we'll do better this time."

An interesting term from psychology that is useful for thinking about this is "counterfactual," or contrary to the facts. In our context, this might refer to the number of defects, issues, and constraints we find at system integration and system test. The counterfactual thinking would be to ask, "*If only we had been using a model-based approach from the beginning...?*" Daniel Kahneman and Amos Tversky (1982) pioneered the study of counterfactual thought, showing that people tend to think "if only" more often about exceptional events than about normal events. It may be the case that the acceptance of the types of issues (that we find every time we build a CPS) is actually normal, and that it is hard for us to ask the if-only question because it is not exceptional.



Using counterfactual thinking, we would envision the world where we had employed model-based methods as an integral part of the design process and use the outcomes to justify the investment in the model-based methods. This approach would be facilitated by some of the work suggested in the following paragraph.

### Post-Mortem Analysis

A different way to arrive at a model-based analysis justification for an organization is by using prior project data to illustrate the things that could have been avoided if only we had applied model-based methods. By itself, this method would only identify the opportunities. The organization would then need to figure out how to incorporate model-based analysis into their process in such a way as to have discovered the issue earlier in lifecycle. This is a useful method for organizations to identify process improvement opportunities.

Regarding post-mortem analysis, it should be relatively straight-forward for an organization to develop a process:

1. Identify a set of projects to review.
2. Examine the defect database, pareto the defects by amount of time to correct the issue.
3. For each of the defects in the top 80%, determine how a model-based method could have been employed to prevent the issue from occurring, along with an assessment of how practical it would have been to have done this.
4. Summarize the effort that would have been saved, realistically, by using the model-based methods, and use this as the potential benefit for the investment.

Using this approach, a root cause analysis would assess where the issue could have been identified, had model-based methods been used. Note that this practice typically already exists for many organizations, where a defect found late in lifecycle is characterized as an *escape*, and that this type of data is used to improve the quality of design reviews. The model-based methods enhance the ability to critically review the system and component designs as they evolve, and an escaped issue could be considered a failure of the model-based review.

This process could be applied at the end of a project, or it could be done iteratively and recursively as the work progresses. Schenker et al. (2022) suggested that the feedback on a set of processes, coupled with mature root cause analysis practices, would rapidly evolve the adoption of model-based methods by an organization (Schenker, Smith, & Nichols, 2022).

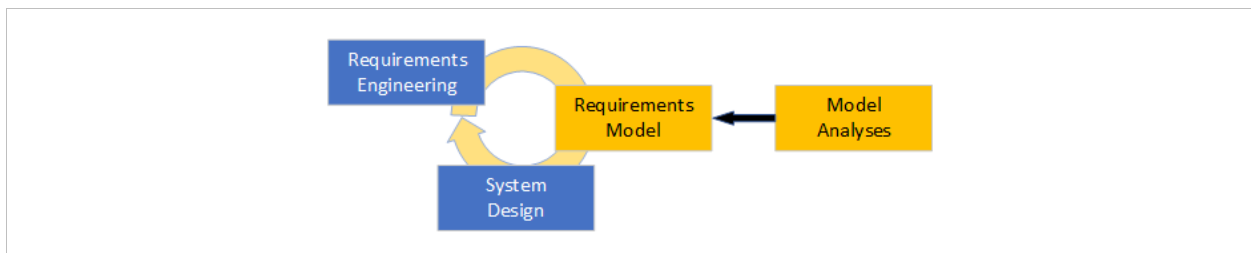


Figure 32. Feedback Loop Incorporating Model-Based Methods

Figure 3 illustrates a way that a model could be used to iterate different technical solutions as requirements are elaborated, and a system design is emerging. Model analyses could be applied to the model as it evolves to predict system characteristics, such as performance, safety, and security.

In the context of this diagram, there would be similar use of model-based methods at the next level of the design, i.e., system/software architecture. A mature development practice

would perform root cause analysis of any issues found downstream, with the goal of understanding whether this issue could have been found in the prior step, i.e., requirements analysis. The goal is to use the models to identify as many of the issues and constraints as are practical to identify, accepting that some of the issues will not be practical to identify.

A critical review at the end of a project, addressing all the issues found during the project, would be of significant value to an organization trying to implement such a continuous process improvement practice.

### **Acceptance by Analogy**

Yet another way to rationalize the decision to move forward is by examining the experiences of similar applications of model-based methods in other domains. This ought to be a fast review. As stated above, our experience is that over and over again, the introduction of the model-based methods in such fields as mechanics, thermodynamics, electromagnetic spectrum, electrical engineering, logistics, maintenance, process optimization, and manufacturing have had a transformative effect on the “way we do business.”

This approach accepts that the underlying benefit from applying a model-based approach will occur analogously in embedded computing resources for CPS as it does in the other domains.

### **Recommendations and Possible Path Forward**

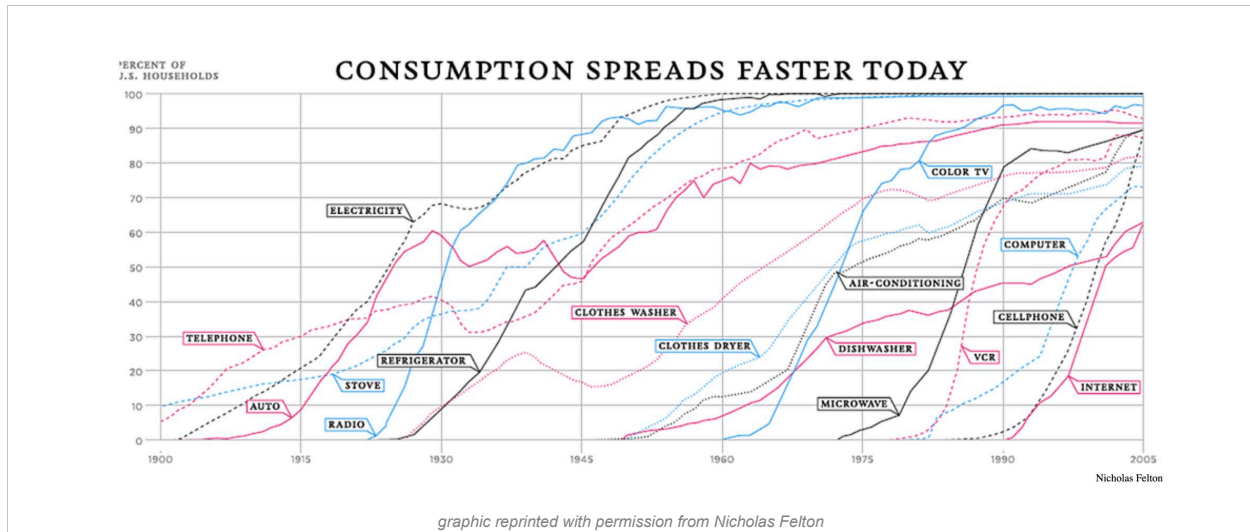
Throughout this paper, we have been asserting that there are real problems with using traditional methods, such as ROI, to justify an investment in better engineering practices, such as model-based design and analysis. The goal of this paper is to point out these problems while providing some alternative means of justification. Our opinion is that once the organization decides to move forward with model-based methods, the process definition experts will figure out how best to apply the model-based tools within the context of their organization’s product development lifecycle. Once a critical mass of practitioners has adopted the methods, then the methods will evolve naturally towards the goal of higher product quality and a lower number of issues discovered late in lifecycle.

We recognize that using model-based methods for embedded computing resources is still in its early days. There has not been enough practical experience with the tooling and the methods, so it is very difficult to make claims about ROI. We’d be cautious of anyone that did make such claims. Comparing the maturity of this model-based technology to something like CAD and FEA is simply not fair.

Investigating other instances of technology adoption, it is not clear what occurred to make the new technology ubiquitous. The chart below, created by Nicholas Felton of *The New York Times*, shows the technology adoption trends for U.S. households for a variety of different technologies.







**Figure 33. Adoption Curve for Various Technological Innovations**

If we were to examine the curve for the microwave oven, for example, we see that there is a significant change in the adoption of microwaves that occurred in the mid-1980s. It is interesting that the first full year of sales of microwave popcorn took place in 1983. The 1987 *New York Times* article, “Microwave Key to Popcorn War,” describes the rapid growth of microwave popcorn sales over the next few years, increasing in revenue from \$53 million in 1983 to \$250 million in 1986 (New York Times, 1987) (“Microwave Key,” 1987). That had to be accompanied by a rapid growth in microwave oven sales, which is supported by this chart. Microwaves had been around since the late 1960s, with a very slow rate of consumer adoption. After the introduction of microwave popcorn, the number of households jumps dramatically from less than 10% in the early 1980s to about 80% of U.S. households by 1990. It is hard to find a household now that does not have a microwave oven.

It's important to note that while the sales of microwaves were struggling through the late 1960s and 1970s, the appliance manufacturers did not back off their commitments to provide this new technology to the American household. Investment in new technology, advertising, and manufacturing capability continued to occur despite the poor sales. Then, something happened, and there was widespread adoption of this technology.

We believe that a similar trigger will occur as the model-based methods become more widely used. One day we believe that we will wake up and find that companies that do not perform this type of work are the exception. In effect, we will have leveled up our CPS development capability, leading to more predictable budgets and schedules, and more buying power for the DoD. We don't know what the trigger will be, but we do expect that there will be a trigger. We think that CPS developers should accelerate their investment in this technology or prepare to be left behind.

### Recommendations for Acquirers

When establishing a new discipline within an acquisition organization, it's necessary to focus not just on the specific practices that should be established, but also on the care and feeding for the practices. Model-based analysis for embedded computing systems is not different. Acquisition programs reside within agencies or PEO structures, and there needs to be support for the practice both at the program level and at the higher echelon level.



1. Continue to set expectations with contractors that model-based design and analysis will be required for current and future acquisitions. Use this as a driver to spur investment in model-based methods. This might have several different elements to it:
  - Language in SEPs, SOWs, and other acquisition documents.
  - RFIs asking contractors to describe their practices for integrating model-based methods into their development practices.
  - Award fees and other contract incentives for successful application of model-based methods.
2. Train staff on how to use the tooling to be able to effectively review, verify, and validate contractor model-based deliverables.
3. Build an enterprise-level competency for model-based methods to establish consistency across programs, and collect lessons learned for future process enhancement.
4. Build the supporting infrastructure (digital engineering environment) to provide the capability to collect and analyze contractor deliverables.

### Recommendations for Contractors

The challenge for contractors will initially be cultural, because using model-based methods as we have described will have a significant change to the way the contractor does business. Winning over the hearts and minds of all the practitioners, from managers to engineers, will be extremely challenging. In particular, the effort required to build a predictive architectural virtual integration model early in lifecycle will be viewed by management as an unnecessary expense, because the model-based methods have not been demonstrated with ROI, and the shifting left of the effort means that there will be less effort available when the real hardware and software show up in the SIL. We must get past this.

Once the culture has been established, and the team has accepted that the model-based methods will improve our likelihood of success, they will need to determine how to apply the methods to improve the existing development process. Then, they will need to establish the root cause analysis practice when defect escapes are found downstream to try to improve the model-based processes.

1. Establish the culture to enable the model-based methods to thrive and add value.
2. Establish how the model-based methods are to be implemented. This will naturally involve some form of digital engineering, in identifying the tools, and configuring the tools into some kind of toolchain.
3. Train staff on how to use the tools to perform the new practices.
4. Develop a strategy for model management when working with heterogenous teams of contractors. Don't assume that it's *my way or the highway*. Elements such as where the authoritative source of design information resides need to be established and communicated with all stakeholders.
5. Take a critical look at the defect resolution process. Examine the criteria for when root cause analyses are performed. Use the results of the root cause analyses to spur innovation with the model-based development methods.
6. Establish a project post-mortem process.
7. Establish a plan for how to account for the added costs and measuring the value received from applying model-based methods to the existing process.



## Conclusion

In this paper, the authors make the case that ROI is not a useful way to assess the viability of adopting model-based systems engineering practices, especially for architecting and evaluating the embedded computing resources of CPS. Instead, we propose alternative ways, such as post-mortem analysis, analogy, or just a leap of faith to justify the increased usage of MBSE techniques to support these CPS projects. These findings are informing some of the current engagements performed by the SEI. The SEI is supporting multiple DoD projects in their adoption towards MBSE and will transfer some of these recommendations into practice as part of our transition work.

## Acknowledgements

Copyright 2023 Carnegie Mellon University

This material is based upon work funded and supported by the U.S. Army Combat Capabilities Development Command Aviation & Missile Center 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.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of U.S. Army Combat Capabilities Development Command Aviation & Missile Center.

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-U.S. government use and distribution.

Internal use:\* Permission to reproduce this material and to prepare derivative works from this material for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.

External use:\* This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other external and/or commercial use. Requests for permission should be directed to the Software Engineering Institute at [permission@sei.cmu.edu](mailto:permission@sei.cmu.edu).

## References

- Bickford, J., Van Bossuyt, D. L., Beery, P., & Pollman, A. (2020, November). Operationalizing digital twins through model-based systems engineering methods. *Systems Engineering*, 23(6), 724–50. <https://doi.org/10.1002/sys.21559>
- Boydston, A., Feiler, P., Vestal, S., & Lewis, B. (2019). Architecture centric virtual integration process (ACVIP): A key component of the DoD digital engineering strategy. *22nd Annual Systems and Mission Engineering Conference*. <https://www.adventiumlabs.com/publication/architecture-centric-virtual-integration-process-acvip-key-component-dod-digital>
- CCDC. (2018). *Comprehensive architecture strategy (CAS) version 4.0*. Redstone Arsenal: U.S. Army Combat Capabilities Development Command, Aviation and Missile Center, Redstone Arsenal. <https://apps.dtic.mil/sti/pdfs/AD1103295.pdf>
- DoD DES. (2018). *Digital engineering strategy (DES)*. Directorate of Defense Research and Engineering for Advanced Capabilities. [https://ac.cto.mil/digital\\_engineering/](https://ac.cto.mil/digital_engineering/)



- Ericsson, A. K., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406.  
<https://doi.org/10.1037/0033-295X.100.3.363>
- Feiler, P. H., Goodenough, J. B., Gurfinkel, A., Weinstock, C. B., & Wrage, L. (2013). *Four pillars for improving the quality of safety-critical software-reliant systems* [White paper]. Software Engineering Institute. <https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=47791>
- Feiler, P., Hansson, J., de Niz, D., & Wrage, L. (2009). *System architecture virtual integration: An industrial case study*. Software Engineering Institute.
- Hansson, J., Helton, S., & Feiler, P. (2018). *ROI analysis of the system architecture virtual integration initiative* [Technical report]. Software Engineering Institute. doi:10.1184/R1/12363080.v1
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, T. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty* (pp. 201–208). Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511809477>
- Lippitt, M., & Knoster, T. (1987). *The Lippitt-Knoster model for managing complex change*. Microwave key to popcorn war. (1987, June 22). *The New York Times*.  
<https://www.nytimes.com/1987/06/22/business/microwave-key-to-popcorn-war.html>
- NSF. (2010). *Cyber-physical systems (CPS)*. National Science Foundation.  
[https://www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf11516](https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf11516)
- Rogers, E. (2003). *Diffusion of innovations* (5th ed.). Free Press Publishers.  
<https://www.simonandschuster.com/books/Diffusion-of-Innovations-5th-Edition/Everett-M-Rogers/9780743258234>
- Roper, W. (2020). *There is no spoon: The new digital acquisition reality*. U.S. Air Force.  
<https://software.af.mil/wp-content/uploads/2020/10/There-Is-No-Spoon-Digital-Acquisition-7-Oct-2020-digital-version.pdf>
- Schenker, A., Smith, T., & Nichols, W. (2022). *Digital engineering effectiveness* [White paper]. Software Engineering Institute & Adventium Labs. <https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=884463>
- SEI. (2023). *Architecture analysis and design language (AADL)*. Software Engineering Institute.  
[https://www.sei.cmu.edu/our-work/projects/display.cfm?customel\\_datapageid\\_4050=191439,191439](https://www.sei.cmu.edu/our-work/projects/display.cfm?customel_datapageid_4050=191439,191439)
- Walsh, W. J. (1909). *The history of steel railway rails: Thesis*. Forgotten Books.



## PANEL 22. ACQUISITION THROUGH MODELING & SIMULATION

Thursday, May 11, 2023	
3:45 p.m. – 5:00 p.m.	<p><b>Chair: Dr. Dennis L. Lester</b>, Associate Provost, Graduate Education, Naval Postgraduate School</p> <p><b><i>The Design and Development of a Defense Acquisition Workforce Virtual Environments for Asynchronous Collaboration (VEAC)</i></b> Domonique Hittner, Naval Postgraduate School</p> <p><b><i>A Reference Architecture for a Policy Test Laboratory</i></b> Alejandro Salado, University of Arizona Hanumanthrao “Rao” Kannan, University of Alabama Zoe Szajnfarber, George Washington University William B. Rouse, Georgetown University Young-Jun Son, Purdue University Nirav Merchant, University of Arizona</p> <p><b><i>Towards an Enterprise All-Domain M&amp;S Environment for T&amp;E: Overcoming M&amp;S Challenges within the DOD</i></b> Jeremy Werner, OSD/DOT&amp;E</p>

**Dennis L. Lester**— joined the Naval Postgraduate School (NPS) team in December 2018 and reports to the Provost and Academic Dean. He has served in a variety of leadership, management, technical, and instructional positions in government, industry, and academia. He started his professional career in the United States Air Force. During his military career, he was an instructor pilot, flight examiner, and commander. He led programs that developed leading-edge technologies supporting research, system development, training, testing, and experimentation. After completing his Air Force career, Denny was employed by The Johns Hopkins University Applied Physics Laboratory, Scientific Research Corporation, Modern Technology Solutions, Inc., and Clemson University. In his most recent roles at Clemson, he was assigned as Associate Director for Science and Technology at the Watt Family Innovation Center, Research Professor, and Interim Director of Clemson Online. In these positions, he was responsible for managing research projects; overseeing building construction and operations; developing partnerships with industry; integrating new learning technologies into the Watt Center, and advancing distance learning programs. Additionally, as a faculty member at Clemson, he taught the capstone course in business strategy to graduating seniors in the College of Business. Prior to his assignment at Clemson, he taught graduate and undergraduate courses in team building, strategic management, and instructional use of computer simulation at the University of New Mexico. At NPS, he serves as the Associate Provost for Graduate Education and Director of Graduate Education Advancement and teaches courses in strategic management and strategic marketing for the Graduate School of Defense Management.



# **The Design and Development of a Defense Acquisition Workforce Virtual Environments for Asynchronous Collaboration (VEAC)**

**Domoniqué Hittner**—was commissioned as a Second Lieutenant in the Field Artillery Branch in 2010 after graduating from Officer Candidate School at Fort Benning, Georgia. Upon successful completion of battery command, she was assessed into the Army Acquisition Corps. Throughout her career, she deployed and served in support of Operation Enduring Freedom (2014), Operation Inherent Resolve, and Operation Freedom Sentinel (2018 and 2019). Hittner earned a Bachelor of Science degree from Florida State University. She went on to earn a Master of Healthcare Administration and Management from the University of Arkansas-Grantham, graduating with distinction. As an acquisition professional, she served as an assistant product manager and then as an executive officer to the PEOC3T, where she was accepted into the Acquisition Technical Expert Program and is currently a PhD candidate in Computer Science-Modeling of Virtual Environments and Simulation (MOVES) at the Naval Postgraduate School. Her dissertation topic is “Virtual Environments as Tools for Asynchronous Collaboration of Diverse Communities of Practice” and she will graduate in June 2023. [domoniquehittner@gmail.com]

## **Abstract**

Asynchronous collaboration is an inevitable part of the global workforce. However, there is a gap in asynchronous collaborative research and solutions that integrates diverse groups engaged in a shared objective. This research demonstrates the process of designing and developing a novel virtual environment (VE) system interface to augment the defense acquisition community’s asynchronous collaborative work. Using an ethnographic approach, the researcher methodologically conducts task analysis, comparative analysis, case studies, and a usability study to derive the best practices to implement in VE. The resulting insights inform the design of a prototype 3D user interface for asynchronous collaboration. In this 3D non-immersive user interface, a set of analytical tools—such as user-generated and in-session system guidance—that support the participants’ asynchronous collaborative tasks is implemented. Based on these studies, the resulting VE is tested for its usability and the extent to which it brings value to the collaborative team in this research’s next phase. This prototype user interface collaborative environment has the potential to be beneficial for a range of communities of practice including the Defense Acquisition workforce, industry, the medical domain, and educational domain.

## **Purpose**

Though many 2D synchronous and asynchronous collaborative solutions exist—e.g., Slack, Skype, Microsoft Teams, or Miro—currently, there are no guiding principles on the type of system features needed to support asynchronous collaborative work in the VE systems. Similarly, there are no documented best practices with which to implement asynchronous virtual collaboration channels that enable diverse communities to conduct their daily work and to do that effectively and efficiently. This unexplored domain space creates a capability gap that hinders businesses with hybrid workers and industries in product development from implementing best practices in their asynchronous work. Currently, no existing framework informs diverse communities of practice, including the defense and industry acquisition community, on the criteria for, and instances in which to use, asynchronous VEs for collaboration, evaluation, testing, optimization, and prioritization of the work on innovative systems and critical solutions slated for the end-users.

The DoD acquisition community is an example of the large problem space and a diverse community of practice needing a superior 3D asynchronous collaborative tool. The acquisition community is a segment of professionals within the industry and the defense domains who design, develop, produce, and procure systems or solutions for users. Though the defense acquisition workforce shares a common mission, they are often operating towards their



objectives on different regulatory tracks. The lack of fidelity, horizontal cohesion, and collaboration in the decision space caused by using less-than-optimal collaborative tools undermines the adoption of various novel systems such as tactical vehicles, cyber operations tools, satellite communication systems, personal protective gear, and VEs over their life cycle (Argyris & Schön, 1992, 1974; Mankin et al., 2004; Thomas et al., 2007).

The study fills this gap by providing a novel theoretical and conceptual framework for asynchronous collaboration in VEs. This research identifies the instances where VEs have great potential in supporting various daily asynchronous collaborative tasks among heterogeneous communities of practice that need to operate with 3D artifacts. The research then uses a methodological approach to develop the guidelines for a 3D asynchronous collaborative tool to augment the defense acquisition workforce's daily work using the author's Theoretical Framework for Asynchronous Collaboration-Virtual Environment (TFRAC-VE).

## Methods

This study uses a multistep developmental design approach to determine the design and developmental features to implement in the prototype VEAC user interface. First, an acquisition domain task analysis study was conducted to determine the purpose, goals, steps, standards of practice, experiences, tools, biases, challenges, and team structure to achieve asynchronous collaborative tasks in a sample population of acquisition professionals engaged in asynchronous collaborative work. Since the acquisition workforce comprises fourteen diverse communities that must work synchronously and asynchronously to deliver timely innovative solutions to support the warfighter, this research examined a sample of career fields—program managers, contracting officers, logisticians, and test and evaluation professionals—that may benefit from the enhanced capability of VR technology (Rendon & Snider, 2008, 2019). Within this step, an acquisition regulation and literature task analysis study was conducted to compare the user findings to domain standards of practice. Second, the resulting insights—along with an extensive literature review in asynchronous collaboration, acquisition research, and virtual environments—were gathered to form a theoretical framework for asynchronous collaboration in human activity. Third, a comparative analysis study was conducted to generate the necessary lessons to apply the domain and technology-agnostic theoretical framework to VE. Fourth, the new elements of the resulting VE framework and an additional comparative analysis study using acquisition research were conducted to demonstrate and implement the VE framework for the defense acquisition workforce. The resulting insights about the framework as well as the themes were derived from the comparative analysis studies. A subset of features that would produce the best results in 3D was implemented in the prototype user interface.

## Resulting Design Features

The cumulative results of the user and comparative analysis studies generated the output processes to support asynchronous collaboration and the best tools to be implemented in 3D. A subset of those processes and tools are implemented in the VEAC as design features to improve the usability and likelihood of adoption. The resulting primary design features are aggregated into user-generated and in-session guidance or system-generated features. The first output of the task and regulation analysis studies produced the user-generated features. The second out of the comparative analysis studies and theoretical framework produced the in-session guidance or system-generated features. The subset of primary features implemented in VE provides a novel mode of asynchronous collaboration for the defense acquisition workforce and supports VE domain scholars' standards of designing systems with adoption in mind (Sadagic et al., 2019).

Figure 1 depicts the primary features of a Defense Acquisition VEAC.



User-generated Features	In session guidance and system features
<ul style="list-style-type: none"> <li>• Performance metrics over the artifact's development</li> <li>• Annotations in objective with the source of changes</li> <li>• Pop-up notifications of the last work completed on objective</li> <li>• Task or objective hierarchy visual</li> <li>• Import new data function</li> <li>• <i>Immersive plug-in option (as requested)</i></li> </ul>	<ul style="list-style-type: none"> <li>• Audio and Video record and playback</li> <li>• 3D models (create, edit, manipulate)</li> <li>• Author tracking with <ul style="list-style-type: none"> <li>➢ Collaborator roles and functions</li> <li>➢ Breadcrumbs of last user work (different color)</li> <li>➢ Save progress or changes (when collaborators step away from VE).</li> </ul> </li> <li>• "Last seen" user location tracker</li> <li>• Catalog of inventory pop-up (reference materials, organizational property book, and commercially funded inventory)</li> <li>• Selection tools: <ul style="list-style-type: none"> <li>➢ Click for the most frequently used inputs</li> </ul> </li> <li>• 3D Object manipulation: <ul style="list-style-type: none"> <li>➢ A transparent object (if in progress or development)</li> </ul> </li> <li>• 3D Objective/Task Features <ul style="list-style-type: none"> <li>➢ System feedback (when an error is made, or task is incomplete prior to selecting the next prompt)</li> <li>➢ Visual task progression bar</li> <li>➢ Input/Output recognition embedded in objective</li> <li>➢ Sign out or task complete feature</li> </ul> </li> <li>• 30-sec video recording (for clarity and quick explanation) <ul style="list-style-type: none"> <li>➢ Closed captioning for video-enabled</li> </ul> </li> <li>• Mute function (to disable audio, excess visuals, and haptics depending on the user's work/learning preference)</li> <li>• Catch-up reel feature or focus feature (receive push notifications during set hours to prevent disturbances)</li> <li>• Teleportation: Navigation <ul style="list-style-type: none"> <li>• Enter a location: navigation</li> <li>• Select and change the background environment</li> </ul> </li> </ul>

Figure 1. Primary Features of a Defense Acquisition VEAC

## Conclusion

The prototype design provides a support mechanism to augment asynchronous collaborative work such as providing clarity to participants and stakeholders during preliminary and critical design review, automating and providing enhanced visuals of provider capabilities during source selection, and aiding stakeholders in finding novel solutions in market research. This capability provides an enhancement to current analog and 2D processes within the defense acquisition workforce and provides program managers, contracting officers, logisticians, and any other of the career fields who may want or need to view or manipulate 3D artifacts a new method of exploring the usability, sustainability, form, fit, and function of new systems or solutions.

## References

- Argyris, C., & Schon, D. A. (1994). *Theory in practice—increasing professional effectiveness*. Jossey-Bass Inc., San Francisco, California, 1992/1974, 224 pages. ISBN-55542-446-5 (paperback). *Behavioral Science*, 39(3), 254–256. <https://doi.org/10.1002/bs.3830390308>





- Mankin, D., Cohen, S., & Fitzgerald, S. P. (2004). Developing complex collaborations: Basic principles to guide design and implementation. In *Complex Collaboration: Building the Capabilities for Working Across Boundaries* (Vol. 10, pp. 1–26). Emerald Group Publishing Limited.  
[https://doi.org/10.1016/S1572-0977\(04\)10001-0](https://doi.org/10.1016/S1572-0977(04)10001-0)
- Rendon, R. G., Snider, K. F., & Allen, N. (2008). *Management of defense acquisition projects*. American Institute of Aeronautics and Astronautics.
- Rendon, R. G., & Snider, K. F. (2019). *Management of defense acquisition projects* (Second edition). American Institute of Aeronautics and Astronautics, Inc.
- Sadagic, A., Attig, J., Gibson, J., Rashid, F., Arthur, N., Yates, F., & Tackett, C. (2019). Designing VR and AR systems with large Scale adoption in mind. In G. Bebis, R. Boyle, B. Parvin, D. Koracin, D. Ushizima, S. Chai, S. Sueda, X. Lin, A. Lu, D. Thalmann, C. Wang, & P. Xu (Eds.), *Advances in Visual Computing* (Vol. 11845, pp. 117–128). Springer International Publishing.  
[https://doi.org/10.1007/978-3-030-33723-0\\_10](https://doi.org/10.1007/978-3-030-33723-0_10)
- Thomas, G. F., Jansen, E., Hocevar, S. P., & Rendon, R. G. (2007). *Field validation of collaborative capacity audit*. Naval Postgraduate School.



## A Reference Architecture for a Policy Test Laboratory

**Alejandro Salado**—is an Associate Professor of systems engineering with the Department of Systems and Industrial Engineering at the University of Arizona, and consults in areas related to enterprise transformation, cultural change of technical teams, systems engineering, and engineering strategy. Alejandro conducts research in problem formulation, design of verification and validation strategies, model-based systems engineering, and engineering education. Before joining academia, he held positions as systems engineer, chief architect, and chief systems engineer in manned and unmanned space systems of up to \$1 billion in development cost. He obtained his PhD in Systems Engineering from the Stevens Institute of Technology. [alejandrosalado@arizona.edu]

**Hanumanthrao “Rao” Kannan**—is an Assistant Professor in the ISEEM department at the University of Alabama in Huntsville. Prior to this, he was an Assistant Professor in the ISE department at Virginia Tech since 2018. Dr. Kannan received a BE in Aeronautical Engineering from Anna University, India, in 2010, an MS in Aeronautical Engineering from the University of Southern California in 2011, and a PhD in Aerospace Engineering from Iowa State in 2015. He then worked as a Postdoctoral Research Associate at Iowa State and Virginia Tech. His research focuses on formalization of Systems Engineering by leveraging disciplines including Decision Analysis, formal philosophy, and engineering design. [hk0049@uah.edu]

**Zoe Szajnfarder**—is the Professor and Chair of the Engineering Management and Systems Engineering Department at the George Washington University. She studies the design and development of complex systems, primarily in the aerospace and defense sectors. Dr. Szajnfarder holds a PhD in Engineering Systems and dual SM degrees in Aeronautics & Astronautics and Technology Policy, all from Massachusetts Institute of Technology, and a BAsc degree in Engineering Science from the University of Toronto. Outside of Academia, she has worked as a systems engineer at Dynacon and MDA Space Systems, and a technology and innovation policy advisor at European Space Agency and NASA. [zszajnfa@gwu.edu]

**William B. Rouse**—is Research Professor in the McCourt School of Public Policy at Georgetown University, and Professor Emeritus and former Chair of the School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research focuses on mathematical and computational modeling for policy design and analysis in complex public–private systems, with particular emphasis on health care, higher education, transportation, and national security. He is a member of the National Academy of Engineering and fellow of IEEE, INCOSE, INFORMS, and HFES. Rouse received his BS from the University of Rhode Island, and his SM and PhD from MIT. [wr268@georgetown.edu]

**Young-Jun Son**—is the Head and Professor of School of Industrial Engineering at Purdue University. Prior to this position, he was the Department Head and Professor of Systems and Industrial Engineering at University of Arizona. His research focuses on a data-driven, multi-scale, simulation and decision model for various applications, including manufacturing enterprise, homeland security, and social network. He has authored over 110 journal papers and 100 conference papers. He is a Fellow of IISE and has received the SME 2004 Outstanding Young ME Award, the IISE 2005 Outstanding Young IE Award, and the IISE Annual Meeting Best Paper Awards. [yjson@purdue.edu]

**Nirav Merchant**—received a BS in industrial engineering from the University of Pune and an MS in systems and industrial engineering from The University of Arizona. He is the Co-PI for NSF CyVerse and NSF Jetstream. Over the last two decades, his research has been directed toward developing scalable computational platforms for supporting open science and open innovation, with emphasis on improving research productivity for geographically distributed interdisciplinary teams. His research interests include data science literacy, large-scale data management platforms, data delivery technologies, managed sensor and mobile platforms for health interventions, workforce development, and project-based learning. [nirav@arizona.edu]

### Abstract

The government has identified several obstacles to inform effective and efficient acquisition policies. Effective modeling, simulation, and analysis of acquisition policies require a multi-domain, multi-scale approach. However, existing research in acquisition policy analysis has primarily



remained siloed. Policy researchers lack a platform that enables sharing, reusing, or integrating the methods, models, and data developed and/or generated by different research teams in different projects. Government envisions a Policy Test Laboratory (PTL) as a potential solution to this need. The PTL is conceived as a service where a domain model developed in a project can be used and/or integrated with another model of a different domain developed in a different project. This paper presents a reference architecture for the PTL, defined as a set of guidelines and constraints that will enable (1) the sharing and use across acquisition research projects of data, models, and tools, and (2) the construction and composition of multi-disciplinary models of government acquisition, that addresses both technical and governing aspects.

## Introduction

The government has identified several obstacles to inform effective and efficient acquisition policies. The defense budget serves many purposes, with many stakeholders. This could lead to inherent conflicting objectives. For example, socioeconomic objectives, including free and fair competition for taxpayer money, can be at odds with the most expedient means to achieve military objectives. We suggest that this complex context results in NDAA, statutes, requirements, etc., that are driven by an overreliance on process metrics because of an inability to define outcome metrics.

In our experience, effective modeling, simulation, and analysis of acquisition policies require a multi-domain, multi-scale approach. Among others, informing a policy decision requires understanding not only financial implications, market reactions, supply chain availability, resulting technical capabilities, societal impacts, and effects on national security, which requires assessing how they relate to each other. However, existing research in acquisition policy analysis has primarily remained siloed to the best of our knowledge. Policy researchers lack a platform that enables sharing, reusing, or integrating the methods, models, and data developed and/or generated by different research teams in different projects.

The Acquisition Innovation Research Center (AIRC) has envisioned a Policy Test Laboratory (PTL) as a potential solution to this need. The PTL is conceived as a service where a domain model developed in a project can be used and/or integrated with another model of a different domain developed in a different project. In this sense, the PTL is not necessarily a unique simulator or aggregated model. While it could be implemented in such a way, non-monolithic implementations are also considered.

This paper presents an initial reference architecture to support the development of the PTL. The reference architecture defines a set of guidelines and constraints that enable (1) the sharing and use across acquisition research projects of data, models, and tools, and (2) the construction and composition of multi-disciplinary models relevant to government acquisition policy research questions. In essence, the PTL's reference architecture is intended to guide research teams in developing models, gathering data, and performing simulations in different domains so that they can be reused and integrated by others.

## Background

This section provides the results of an initial assessment of the characteristics, scope, drivers, and main capabilities of existing efforts in other domains that have attempted or are attempting to integrate models and data across disciplinary boundaries. The effort allocated to identify and assess existing architectures and/or frameworks was timeboxed. This section is not aimed at being comprehensive but rather exploratory.

Identification was performed by aggregating frameworks and architectures already known to the researchers, as well as by a quick online search. Assessment was performed



based on publicly available documentation and/or conversations with some of the people involved in the architecture or framework being assessed. In line with the exploratory spirit, the activity was not intended to necessarily achieve an accurate characterization of existing work. Therefore, it is recognized that there may be some inaccuracies in the information provided in this section. Nevertheless, the information is still considered relevant and useful for the purpose of informing the developing of the initial reference architecture for the PTL.

Nine frameworks were assessed: the MIT Joint Program on the Science and Policy of Global Change,<sup>1</sup> the IEEE Std 1516-2010 (IEEE, 2010), the Multi-level Modeling framework (Rouse, 2019, 2022), CyVerse,<sup>2</sup> the National Socio-Environmental Synthesis Center (SESYNC),<sup>3</sup> the Simulation Framework (CSF; (Haynes et al., 2003; Singh & Mathirajan, 2014), the One Semiautomated Force (OneSAF; Parsons et al., 2005), the Modeling Architecture for Technology, Research, and Experimentation (MATREX; Hurt et al., 2006), and the OpenGMS.<sup>4</sup>

Each framework was assessed in the following attributes, noting that for some frameworks some of this information might have not been available or identified during the activity: (1) Background, goal, and maturity (or state of the development) of the framework/architecture; (2) Types of research questions it is intended to support, including application domains it serves; (3) Kinds of disciplinary models, data, and tools it is intended to support, including integration capabilities (i.e., connecting across models, data, tools...); (4) Architectural aspects, such as layers, components, integration, relationship between parts, services it provides, etc.; (5) Technical governance, including maintenance and, if possible, rough estimate of effort; and (6) Organizational governance, including maintenance and, if possible, rough estimate of effort.

Existing frameworks display a wide variety of approaches to establish frameworks that enable the integration of models across disciplinary boundaries. There seem to be three main trends in establishing these frameworks:

*Structural frameworks:* These frameworks provide structure and guidelines that enable reuse and integration of models but do not provide any integrated model. These are independent of research question. Details of how integration occurs are left open for the different modeling actors to define. These frameworks are generally established through working groups or standards bodies.

*Top-down frameworks:* These frameworks are constructed around a research question. An integrated overarching model is constructed, even if not at once. Using and contributing the model requires evaluation and approval of a governing body that oversees the growth of the model. Answering research questions requires interacting with all or part of the integrated model. As a result, deployment requires a substantive portion of the integrated model to be constructed before it can be used. This, together with the extensive oversight required to maintain the model, leads to high upfront and sustainment investments.

*Bottom-up frameworks:* Similar to structural frameworks in the sense that a structure to enable integration is provided, but additional guidance and infrastructure are provided to integrate models around a class of research questions. These frameworks often rely on open source and open access artifacts, as well as a decentralized contribution from researchers, which reduces the investment needs to deploy and sustain the resulting models.

---

<sup>1</sup> <https://globalchange.mit.edu/>

<sup>2</sup> <https://cyverse.org/>

<sup>3</sup> <https://www.sesync.org/>

<sup>4</sup> <https://geomodeling.njnu.edu.cn/>



## **MIT Joint Program on Science and Policy of Global Change**

The MIT Joint Program on the Science and Policy of Global Change has the mission of “advancing a sustainable, prosperous world through actionable, scientific analysis of the complex interactions among co-evolving, interconnected global systems.” Founded in 1991, the “program” has pursued research that enables decision-makers to answer policy questions related to sustainability. Specifically: environmental protection, economic viability, and social equity. It has always been a collaboration mainly between earth scientists and economists and specializes in “integrated assessments.” It is a research program in that it houses a mix of faculty, research scientists, and graduate assistants (at multiple levels, but weighted the technology and policy program masters students). It received anchor funding from the Department of Energy and also works with a consortium of sponsors. Over a history, it has done a mix of inquiry-driven development vs. infrastructure development; that balance has shifted over time.

The program was designed to provide relatively quick comprehensive analysis to support decision-making on global and climate relevant policy questions. Their work in seven focus areas. The most relevant to this project is the policy scenarios.

Most of the Joint Program’s work leverages the Integrated Global System Modeling (IGSM). It has two interacting components: (1) The Economic Projection and Policy Analysis (EPPA) model (a computable general equilibrium model from economics) and (2) the MIT Earth System model (MESM; from atmospheric science). Both include discipline-specific models of the “physics” of the relevant system. EPPA draws on trade data that was curated over decades.

A version of each of these models existed at the time when the program was founded. Since then, most of the new work has focused on building additional resolution in segments of the economy of the earth system when they are needed to answer a specific policy question. For some specific purposes, new models are developed that use different data sets or aggregate sectors differently.

For the first 20 years of the program, technical development was led by one key research scientist. He worked with every student contributing a module and retained authority to include a new module into the live EPPA instance. Most new technical tasks focus on “building-out” a specific relevant module. Before it is integrated into the overall model, would take responsibility for V&V. As the program has grown, there are a few more technical leads, but the group is still small, and technical governance is centralized. Their approach has been quite centralized too, in that there has generally been a Director/PI for each of the economic and earth systems sides, with a few senior research staff and a lot of student research assistants. They collaborate through weekly lunch tag ups where the RAs got to watch the discussions of the senior folks about what work to prioritize. Even though the effort is highly problem-driven making external stakeholders were important, the team retains a strong emphasis on the overall goal of developing “this global modeling competency.” This has led to a lot of co-creating of the intersection of model extensions to support groups of pressing questions.

## **IEEE Std 1516-2010**

The IEEE Std 1516-2010 describes the framework and rules of the High Level Architecture (HLA), which is an integrated approach to provide a common architecture for federated simulations. HLA was initially developed under the leadership of the U.S. Department of Defense in the mid 1990s. In 1998, the Defense Modeling and Simulation Office (DMSO) released HLA 1.3, an official document of HLA. The second version (HLA-2000) and third version (HLA-2010) were then further refined and published by IEEE. The latest version (HLA 1516-20XX; HLA 4) is currently developed by Simulation Interoperability Standards Organization (SISO).



The goal of HLA was to assure interoperability and reusability of defense models and simulations (training, analysis, and control) (original goal), which was extended to broader applications (e.g., manufacturing/supply chain management, health care, infrastructure, and more). It enables us to connect simulations running on different computers, locally or widely distributed, independent of their operating system and implementation language, into one federation. Maturity: Run-time Infrastructure (RTI) is major component of HLA, and a software that provides a standardized set of services, as specified in the HLA interface specification. In the past two decades, multiple RTIs have been developed as an open source (e.g., <http://porticoproject.org>), by a commercial sector (e.g., MAK Technologies), or research team projects (web services).

HLA has been used to address interoperability and reusability of defense models (e.g. DoD projects), development of supply chain network simulation, integrating geographically dispersed member simulations (e.g., National Institute of Standards and Technology and Boeing), and a city-level traffic simulation (the Federal Highway Administration). In these projects, researchers and practitioners used HLA to integrate a mix of the following elements: (1) system dynamics (aggregate level), (2) discrete event (process flows), (3) agent based (decision making, communications), (4) dynamic systems or physics-based game engine, (5) hardware (e.g., robots, machines, drones; simulations running in real-time), and (6) human (simulations running in real-time).

Following a publish and subscribe architecture, HLA can be applicable to various types of operating systems, software, applications, and languages. For example, it allows integration of wide ranges of software: AnyLogic, Simio, Arena, ProModel, Repast, DynusT (traffic simulator), hardware (robots, machines, drones), Unity (game engine), and more.

To maintain or govern models, an open source Portico (<http://porticoproject.org>) or a commercial RTI (e.g., MAK Technologies) can be used, and efforts are needed to develop technical governance. In addition, a governance structure and agreement need to be established among sponsors and users.

### **Multi-Level Modeling**

The Multi-Level Modeling approach to modeling represents an enterprise or an ecosystem at four levels of abstraction: people, processes, organizations, and society. The levels are typically represented by agent-based models (people) discrete event or network process flow models (processes), microeconomic models of decision making (organizations), and macroeconomic models of policies (society). This framework has been in use, and continually refined, for over 10 years, addressing research questions related to economics of scaling clinical trials to broader use (Emory, Indiana, Penn, Vanderbilt), likely impacts and efficacy of health policies (ACA, CMS), and impacts of incentives on consumer energy behaviors (Accenture, GM).

The engagement of sponsors and subject matter experts is central to this approach. Such dependency makes scheduling and conducting meetings a challenge. The approach is not very adaptable, and models are difficult to update once the sponsor's questions have been answered. As such, the models are not necessarily maintained or governed. Instead, each new question demands the development of new models, which require an investment in order of \$200,000–300,000 for familiar domains and \$500,000–1,000,000 for new domains.

### **CyVerse**

CyVerse provides scientists with powerful platform to handle huge datasets and complex analyses, thus enabling data-driven discovery. CyVerse offers extensible platforms that provide data storage, bioinformatics tools, data visualization, interactive analyses, cloud services, and



APIs, among others, with the purpose of transforming science through data-driven discovery. It is conceived as a federated platform for enabling diverse teams to collaboratively develop and share solutions for data driven questions, and support analyses that need domain specific models and machine learning workloads. Current applications range from astronomy to Earth sciences to hydrology, traffic engineering, and life sciences.

CyVerse is built on a layered architecture that abstracts data storage and execution environments. Data management is driven by metadata, remaining agnostic of the physical storage provider (which can be on the cloud, private premises, etc.). Access to the execution environment is secure with federated sign on. Layers are connected through automation using APIs and the end user facing applications are customized for specific purposes through web interfaces. This allows for developing methods and securing sharing underlying tools/pipelines and data without needing software installation on client side.

Operationally, CyVerse has a public deployment and the capability to be deployed privately at different organizations. The public deployment is maintained by the University of Arizona, and it can be integrated with private infrastructure.

## **SESYNC**

The SESYNC, established in 2011, brings together the science of the natural world with the science of social systems and decision making to solve problems at the human-environment interface. SESYNC has accelerated research and learning that seeks to understand the structure, functioning, and sustainability of coupled social and environmental systems. This is achieved by enhancing teams' and individual participants' capacities and skills to bridge varying epistemologies, methods, and approaches. SESYNC has supported over 340 projects, engaging over 4,700 researchers in over 70 countries. Its research output accounts for over 750 peer-reviewed publications.

SESYNC research relies on many different forms of information (data collected by quantifying an event or outcome, running a computer simulation, collecting photographs, transcribing interviews, or capturing social media activity), highly heterogenous data, and synthesis and analysis methods (systematic literature reviews, meta-analyses, expert elicitation, statistical and spatially explicit modeling, system dynamics, and agent-based modeling).

The SESYNC builds upon a decentralized infrastructure of several dedicated software and tools. In terms of organizational governance, all products developed under SESYNC-sponsored activities are made accessible with no restrictions for use and dissemination through FTP or code repository services.

## **CSF**

The CSF was initiated by the U.S. Army Aviation and Missile Research, Development, and Engineering Center (AMRDEC) in 1999 as a standardized structure to support dynamic simulations. The original intent for the framework was to be domain agnostic, but it evolved as a specific toolkit to support modeling and simulation of tactical missile systems. It supports both discrete event simulation and differential equations. It supports simulation of missile deployment, 6 DOF Propulsion Aerodynamics Controls and Kinematics module, and hardware-in-the-loop testing, with both real-time and non-real-time capabilities. The framework is flexible and supports various models, data, and tools, with a common library approach and C++ implementation. It has a GUI for model composition and allows for plug-ins.

## **OneSAF**

The OneSAF is intended to foster interoperability and reuse across modeling and simulation communities of the Army. The framework supports the development of advanced concepts for doctrine and tactics, training of unit commanders and staffs, development of new



weapon systems, and production of data as input to other simulations. Its applications include testing algorithms for real C4I systems, modeling WWII tank combat, creating a “cyber range” for cyber warfare analysis and training, and virtual training on operating construction equipment. The framework supports physics-based models for platforms, soldiers, equipment, logistical supplies, communications systems and networks, emerging threats, and aviation assets. It has a layered approach with components linked into a common executable, and data exchange occurs through method calls. However, the framework has no inherent mechanism to enforce assumptions and dependencies of a component if used in a different context, and the validity of the composed system is up to the developer.

## **MATREX**

The purpose of MATREX is to develop a composable Battle Command-centric modeling and simulation MS environment consisting of multi-fidelity models, simulations and tools that are integrated and mapped to a Future ForceBlended Force architecture for use across the acquisition spectrum, specifically integrating live, virtual, and constructive models at the entity or engineering level. MATREX is not limited to a specific application and can be extended, serving as a support system for various types of research questions and application domains, including modeling command and control, communications actions and effects, and network centric warfare systems. The framework supports the integration of different disciplinary models, data, and tools, including OneSAF, Aviation Mobility Server, Countermine Server, Missile Server, and others. The architecture of MATREX includes a layered approach with three layers - Federates, Middleware, and Distributed execution infrastructure. Technical governance and organizational governance have not been assessed.

## **OpenGMS**

OpenGMS supports open web-distributed sharing of modeling and simulation resources for geographic applications by providing a virtual community for collaboration among researchers from various domains. The models are heterogeneous, both in terms of domain of application and scale.

OpenGMS uses a layered architecture with four layers: Model repository, Data repository, Models as a service, and Thematic center. The model repository collects model resources to build a dictionary where all models (also include related tools, algorithms, etc.) are organized in a formal way. Users can find a model with its detailed information, conceptual and logical descriptions, computable resources, developing history, and applications. This repository publishes model resources under the permission of the author. The data repository collects data resources to build a community where users can explore modeling-related data through a universal center. Users can share their data resources to the data repository. Various data sharing sites can be also linked to support users so that they do not visit individual sites. This data repository publishes data and their related information under the permission of the author. This platform provides model, data, and computing resources as corresponding services in an open web environment. Users can setup input data and run a model via a web client, and the related model will be executed in a remote computer node. Users can invoke a model service before boarding and obtain results when get off. A set of alternative solutions are available to convert original models as reusable services, to publish data files as reusable services and to share computers as available services. Several thematic centers are constructed to help researchers collect models, data and other related resources. Topic-related or problem-related resources could be easily discovered within a thematic center.





## Reference Architecture

### Use Cases

Three main use cases were used to inform the development of the reference architecture:

- 1) The government has a policy question that could be answered using the PTL. Example: *How should investments in acquisition supply chains be managed across mission areas with highly uncertain demands?*
- 2) A (policy) researcher wants to leverage the PTL. Example: *How can a STEM policy found to be successful in a pilot study in one state, best be scaled to provide benefits to all states?*
- 3) A researcher wants to integrate their work in the PTL. Example: *How can a large data set on technology innovations in sensors and semiconductors be imported to the PTL for access and use by other researchers?*

### Major Drivers

Success of the PTL requires two key contributions: researchers that use and contribute to the PTL, and a sponsor that trusts the results generated with the PTL. Having in mind that the needs of the sponsor will change over time, as well as the science, models, methods, and tools used by researchers, the ability to seamlessly evolve the PTL is likely to be instrumental for its own sustainment. Therefore, the major drivers that informed the development of the reference architecture were sponsor trust, researcher adoption, and evolutionary needs.

Researcher adoption is likely to be driven by two questions:

- (1) As a researcher, why should I use the PTL?
- (2) As a researcher, why should I make an extra effort to make my models, data, and methods reusable by other researchers and interoperable with other models, data, and methods that I do not plan to make use of?

Addressing these may require incorporating provisions for establishing incentives in the reference architecture.

Gaining the trust of the sponsors to use the results provided by the PTL will likely depend on several factors. There is abundant literature on this topic, but it was not possible to explore it in detail as part of the sponsored project due to time constraints. Instead, the team started off their own experience in working with sponsors in the context of modeling. A summary of factors leading to trusting different aspects of the modeling effort are summarized in *Table 1*. In addition, it is noted that trust between the modeler and the stakeholders takes time to build and the path to build such trust depends on the type of relationship between them.

**Table 10. Informal model of sponsor trust**

Who/What am I trusting?	Modeler (track record of interacting with stakeholder or reputation)	Model (previously used/accepted or careful V&V in this context)	Inputs (provenance, e.g., censuses, vs. careful look at representativeness for this application)
Validity (solve my problem)	Gut of senior stakeholder	Classic model V*V Depends on generation (block 1 different than n)	Good data vs. right data for this application
Acceptability (in ways I prefer)	Comfort/confidence in understanding (and the way they talk to me)	Type of models used (understand representation, e.g., pde vs. econometrics) Explainability	Support credibility of the data (available in community and has been vetted by experts)



Viability (worth my time to learn how to do)	Cost-effort to work with expert vs. use their tool (either learning to work together or learning the tool)	Effort to develop my own comfort (learning curve)	Effort to compile/clean. Proprietary/classified/expensive?
--	--	---	--

Trust in the concept of the PTL adds an additional dimension; that of trusting the integration of models and data. Stakeholders do not only need to trust the individual components that form the integrated models, but also the integration process of the models and the resulting integrated model. Furthermore, whereas some models may have been considered valid, acceptable, and/or viable on their own by their dedicated stakeholders, these assessments may need to be revisited in the context of the integrated model and the new stakeholders. Assuming a bottom up PTL, as discussed earlier, stakeholders may not even have access to the modelers that modeled some of the components of the integrated model, which further hinders trust. Transparency and clarity on model usage may likely be a key aspect the reference architecture must facilitate.

Facilitating the evolution of the PTL with respect to research questions, scientific discoveries, modeling frameworks, novel methods, etc., results in some development challenges. While the reference architecture can provide flexibility, evolution cannot be unbounded. In fact, guidelines and bounded actions may be necessary to guarantee that existing models and data do not inadvertently become not usable due to the evolution. In other words, it is likely that the reference architecture does not simply facilitate evolution but that it guides it to maintain relevance, validity, and acceptability of the artifacts it possesses at the time of the evolution.

The ability to compose, in varying combinations, simulation components (e.g., models, applications, etc.) into simulation systems to satisfy specific user requirements. The defining characteristic of composability is that different simulation systems can be composed in a variety of ways, each suited to some distinct purpose, and the different possible compositions will be usefully valid. Composability is more than just the ability to assemble simulations from parts; it is the ability to combine and recombine, to configure and reconfigure, sets of components into different simulation systems to meet different needs (Petty & Weisel, 2019).

Furthermore, the different artifacts in the PTL must allow for model composability to enable integrating heterogeneous models. Model composability can have two interpretations (although both are needed for a valid composition): (1) Syntactic composability and (2) Semantic composability. Syntactic composability deals with the actual implementation aspects of model composition, where the focus is on the implementation details such as parameter passing mechanisms, external data accesses, and timing assumptions. This strives to ensure that the composed models are compatible for all of the different configurations that might be composed. In contrast, semantic composability is a question of whether the models that make up the composed simulation system can be meaningfully composed (i.e., if their combined computation is semantically valid). Even if the components can be composed syntactically, the models may or may not be composable semantically. Since one of the critical attributes of a simulation system is the degree of reorganizability, to answer a wide range of questions, semantic composability is a more appropriate notion. Note that syntactic composability is a necessary but insufficient condition for semantic composability.

Model composability requires metadata associated to each model and may be facilitated by certain tenets of the framework in which composability occurs. Desired model metadata that are required to enable composability include, among others (Petty & Weisel, 2019):



- **Nature:** assumptions, spatial and temporal resolution, boundary conditions, range of validity, inputs and outputs, details about model interpretations, etc.
- **Tools/Technology:** software, implementation language, operating system, compiler version, tools, etc.
- **Interfaces:** syntax, data definitions, standards
- **Applications:** run modes, performance, intended uses
- **Provenance:** developers, prior uses, validation history

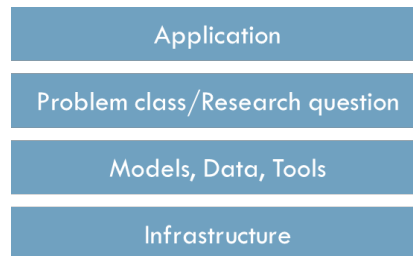
Useful characteristics for a framework that can facilitate implementation of composability include, among others (Petty & Weisel, 2019):

- Dynamic model registration and discovery, supported by a directory (or directories) of registered models and repositories.
- Semantic query, search, and reasoning capabilities for model selection, supported by model specifications (i.e., metadata).
- Distributed processing across multiple platforms, systems, services, and domains.
- Support for intelligent and polymorphic proxies for models.
- Automated composition processes to combine models.
- Virtual repositories that include version control.
- Ability to save compositions and composition templates.
- Compliance with relevant standards.
- Software authentication and information exchange services.

## Architecture

The reference architecture for the PTL is defined as a set of guidelines and constraints that will enable (1) the sharing and use across acquisition research projects of data, models, and tools, and (2) the construction and composition of multi-disciplinary models of government acquisition, that addresses both technical and governing aspects.

A layered reference architecture is proposed (Figure 1). The Application layer handles aspects related to how organizations and infrastructure engage (e.g., security aspects or UI/UX). The Problem class/Research question layer handles aspects related to assessing if a given task can be supported by the PTL (as an integrated assessment tool). The Models, Data, Tools layer handles the actual research artifacts indicated by their names. The Infrastructure layer handles all aspects related to hosting, storing, and exchanging the research artifacts with the PTL consumers.



**Figure 34. Reference architecture**

This layered architecture allows for PTL designs that can embed the useful characteristics to facilitate model composability listed earlier. For example:

- Dynamic model registration and discovery, supported by a directory (or directories) of registered models and repositories. → Through the *Application Layer*, a user can query



the *Problem Class/Research Question Layer*, which accesses the directory of models in the *Models, Data, and Tools Layer*.

- Semantic query, search, and reasoning capabilities for model selection, supported by model specifications (i.e., metadata). → Through the *Application Layer*, a user can semantically query the *Problem Class/Research Question Layer*, which accesses the directory of models and their metadata in the *Models, Data, and Tools Layer*.
- Distributed processing across multiple platforms, systems, services, and domains. → The *Infrastructure Layer* can be implemented as a distributed platform.

Specific choices of what characteristics to implement are left to specific PTL designs.

The next subsections provide further details and discussions for each layer in the reference architecture. Each layer is elaborated with a different depth, based on prioritizations made by the research team in the scope of the sponsored project leading to this paper.

### Application Layer

The Application layers provide a framework for the engagement of the different actors and the PTL. Three actors have been identified: researcher, sponsor, and the AIRC. Anticipated engagements are depicted in Figure 3. Note that, while the application layer is defined in the context of the tasks performed by the different actors, the application layer does not include the tasks but provides the means to the different actors to interact with the PTL to execute those tasks.

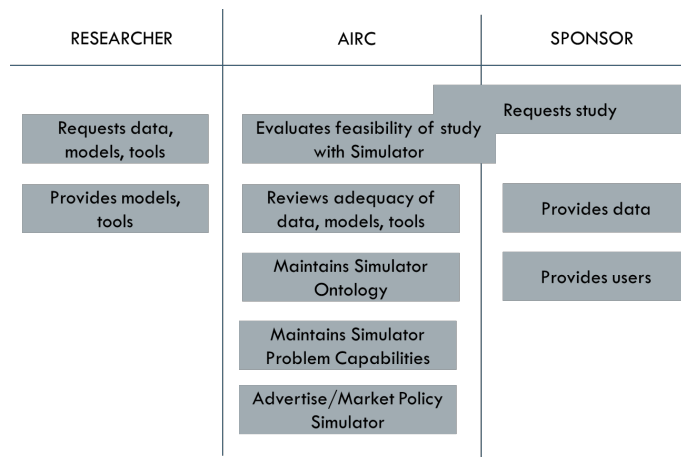


Figure 35. Engagements Between PTL actors

Researchers are anticipated to contribute with their models, data, and tools to build the PTL. Essentially, models, data, and tools resulting from their research projects will be fed into the PTL. At the same time, researchers are anticipated to be consumers or users of the data, models, and tools already available in the PTL. This is, in fact, the purpose of the PTL: a researcher can make use of models, data, and tools already in the PTL to conduct cross-disciplinary research. The application layer handles the exchange of requests and data, models, and tool exchanges between the researcher and the PTL, including aspects related to UX/UI, security, and access restrictions, among others. The extent to which the reference architecture should constraint these aspects needs to be addressed in future work.

Sponsors are anticipated to feed the PTL with data to support research and be the consumers and users of the results generated by the PTL. Furthermore, sponsors are also expected to initiate most of the research supported by the PTL. This is expected to be done in tandem with the AIRC team, which will possess the detailed knowledge of what research



questions could the PTL support. As for researchers, the application layer handles these exchanges between the sponsors and the PTL, including aspects related to UX/UI security, and access restrictions, among others. The extent to which the reference architecture should constraint these aspects needs to be addressed in future work.

The AIRC is anticipated to act as the governing body of the PTL, undertaking the activities associated with sustaining the PTL and supporting its operations. On the technical front, because, as described earlier, the PTL is expected to be developed bottom-up, the adequacy of the models, data, and tools injected by researchers and/or suppliers into the PTL should be assessed for conformance to guarantee future integration efforts. Furthermore, since some of these models, data, and tools may incorporate new aspects not previously addressed by the PTL, its ontology and evolving capabilities will also need to be maintained. On the programmatic front, the AIRC is anticipated to jointly work and support its sponsors in assessing the feasibility and adequacy of the PTL to support desired research questions, as well as to advertise the PTL's capabilities to reach to a wide variety of researchers and sponsors that could benefit from them. As for researchers and sponsors, the application layer handles these exchanges between the AIRC and the PTL, including aspects related to UX/UI. The extent to which the reference architecture should constraint these aspects needs to be addressed in future work.

### ***Problem Class/Research Question Layer***

The Problem Class/Research Question layer provides the necessary services to characterize the artifacts in the PTL in the context of trust. Particularly, these services evaluate the information contained in the different PTL's artifacts to determine the questions or problems that the PTL can support and the level of confidence to be expected in such support. This can be thought of as the identification of capabilities enabled by the models, data, and tools in the PTL; this includes those already existing and those that may be created during the research.

While a more in-depth assessment is necessary, this layer handles taxonomical aspects important to trust such as:

- *Scale*: it indicates the context in which the model, data, and/or tools have been used (e.g., from a successful prototype to a large-scale application).
- *Projection of tipping points*: it indicates the likelihood of achieving change (e.g., confidence on organizational or social change)
- *Risk assessment*: it indicates risks associated with using the different artifacts in the PTL for a particular problem class or research question.
- *Control mechanisms*: it indicates the interoperability of models, data, and/or tools with respect to a specific problem class or research question.

While traditional concepts, methods, and tools for model verification and validation are likely to be adopted in this layer, novel methods to forecast and execute verification and validation of integrated (heterogeneous) models may need to be developed. These will refine the constraints imposed in the metadata to be provided with every model, dataset, and method that is fed to and/or used together with the PTL.

### ***Models, Data, and Tools Layer***

This layer represents the models, data, and tools in the PTL. It encompasses the artifacts and their associated metadata or ancillary information, which include the information necessary to (1) integrate each model, data, and/or tool with other models, data, and/or tools, and (2) assess the confidence level or trust in the artifacts. The layer implements control



mechanisms to guarantee that every artifact in the PTL conforms with pre-defined requirements for such metadata, ancillary information, and confidence characterization.

Several model metadata standards and/or protocols to enable model integrability are in use in other fields. Two examples are presented below, the Open Modelling Interface (OpenMI) (Moore & Tindall, 2005) and the ODD (Overview, Design concepts, Details) Protocol (Grimm et al., 2010).

In the field of geospatial information modeling, the OpenMI was defined with the goal to “bring about interoperability between independently developed modelling components, where those components may originate from any discipline or supplier” (Moore & Tindall, 2005). The standard is over 100 pages long and formally defined through schemas in UML. Its coverage is comprehensive, including requirements on model elements, interfaces, values, and linking capabilities and/or protocols. UX/UI are also covered with templates to document the metamodel and its conformance to the standard. The standard is very specific to geospatial modeling, so it cannot be directly reused for the PTL. However, it provides a good indication of the kind of effort that goes into defining a modeling interface for acquisition research.

The ODD Protocol has a narrower scope, focusing on fully defining agent-based models (Grimm et al., 2010). It requires every model to incorporate details of general nature (purpose, entities, state variables, scales, and process overview and scheduling), design concepts (basic principles, emergence, adaptation, objectives, learning, prediction, sensing, interaction, stochasticity, collectives, observation), and details of the model (initialization, input data, and sub-models).

The same applies to metadata standards. An example is the FAIR Data Standard (Wilkinson et al., 2016). FAIR provides a set of principles that have the goal to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets, with an emphasis on machine-actionability:

- Findable:
  - F1. (Meta)data are assigned a globally unique and persistent identifier.
  - F2. Data are described with rich metadata (ref. to R1 below).
  - F3. Metadata clearly and explicitly include the identifier of the data they describe.
  - F4. (Meta)data are registered or indexed in a searchable resource.
- Accessible
  - A1. (Meta)data are retrievable by their identifier using a standardized communications protocol.
    - A1.1 The protocol is open, free, and universally implementable.
    - A1.2 The protocol allows for an authentication and authorization procedure where necessary.
  - A2. Metadata are accessible, even when the data are no longer available.
- Interoperable
  - I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
  - I2. (Meta)data use vocabularies that follow FAIR principles.
  - I3. (Meta)data include qualified references to other (meta)data.
- Reusable
  - R1. (Meta)data are richly described with a plurality of accurate and relevant attributes.



- R1.1 (Meta)data are released with a clear and accessible data usage license.
- R1.2 (Meta)data are associated with detailed provenance.
- R1.3 (Meta)data meet domain-relevant community standards.

Detailed descriptions of the principles and several implementation examples are publicly available. There are also tools available to support the generation of metadata that guarantee abiding to some of the FAIR principles.

Furthermore, this layer also incorporates two mechanisms that provide an underlying structure to foster internal consistency between the artifacts in the PTL:

- (1) An acquisition ontology. In line with some of the requirements and principles identified earlier, an ontology will establish common understanding and interpretation of acquisition concepts, avoiding terminological and conceptual conflicts between models, as well as redundancies.
- (2) (Potentially) A hetero-functional graph (Schoonenberg et al., 2019). Building upon the ontology, a hetero-functional graph provides a mathematical structure to integrate heterogeneous models. While this still needs some investigation, hetero-functional graphs may provide valuable capabilities to assess confidence and trust resulting from such integrations.

### **Infrastructure Layer**

The Infrastructure Layer hosts all the artifacts of the PTL. It can be thought of as a repository containing models, datasets, and tools, as well as the tools that implement the different layers of the PTL.

Three major alternatives have been identified:

- Use an *existing infrastructure*, such as CyVerse. This alternative usually requires the minimum upfront development effort but might provide insufficient security protection for certain datasets.
- Use a *custom, centralized infrastructure*. In this case, AIRC would develop and maintain the repository. This option offers the maximum flexibility to satisfy sponsors hosting needs but likely requires a significant upfront development effort.
- Use a *decentralized approach*, where each researcher must host the artifacts that they develop and provide PTL users with access to them, both within requirements set forth by the AIRC.

The reference architecture does not need to constraint the implementation of the PTL to any particular alternative. The selection may be done in the context of the PTL design.

### **Conclusions**

An initial reference architecture to support the development of a PTL has been presented. The reference architecture consists of four layers that are aimed at enabling the sharing and use across acquisition research projects of data, models, and tools, and the construction and composition of multidisciplinary models of government acquisition, that addresses both technical and governing aspects.

A PTL could be purposed to support a suite of activities aimed at answering a wide diversity of policy questions or to center on a type of policy problem and focus on building test



range infrastructure over time. In the first option, the extent of reuse is mostly data and generic modeling strategies, standards, and best practices. In the second option, reuse goes beyond data and standards; it requires a core set of models that can be quickly customized for specific questions. The proposed reference architecture is intended to support both kinds of developments, particularly promoting the organic growth of the PTL as data, models, and results from different projects become available and injected into the PTL.

In fact, given a lack of clarity on the existence of commonality or agreements related to modeling in acquisition-related research, a bottom-up implementation approach is suggested initially. The basic idea consists of, first, not constraining the work of acquisition policy researchers to specific models, tools, or modeling approaches. Instead, researchers are requested to deliver a set of artifacts (and metadata) associated with the models, datasets, and tools they generate during their project. These are then consolidated and aggregated by a dedicated team, resulting in a PTL that will grow larger, more mature, and more capable with every new acquisition research project. As the PTL matures, sponsors could incorporate additional constraints to be met by researchers to facilitate integrability with the PTL. It is anticipated that this implementation plan requires minimal upfront effort, which will organically increase as the maturity and capabilities of the PTL increase.

## Acknowledgment

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 0309.

## References

- Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, 221(23), 2760–2768. <https://doi.org/10.1016/j.ecolmodel.2010.08.019>
- Haynes, B., Carroll, T., Tollison, D., Kendrick, W., & Salter, A. (2003). *Development of the Missile Component Simulation Library (MCLib) for tactical missile simulation*. Paper presented at the Huntsville Simulation Conference, Huntsville, AL.
- Hurt, T., McDonnell, J., & McKelvy, T. (2006, December 3–6). *The modeling architecture for technology, research, and experimentation*. Paper presented at the Proceedings of the 2006 Winter Simulation Conference.
- IEEE. (2010). IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA)– Framework and Rules. *IEEE Std 1516-2010 (Revision of IEEE Std 1516-2000)*, 1-38. <https://doi.org/10.1109/IEEESTD.2010.5553440>
- Moore, R. V., & Tindall, C. I. (2005). An overview of the open modelling interface and environment (the OpenMI). *Environmental Science & Policy*, 8(3), 279–286. <https://doi.org/10.1016/j.envsci.2005.03.009>
- Parsons, D., Surdu, J., & Jordan, B. (2005). *OneSAF: A next generation simulation modeling the contemporary operating environment*. Paper presented at the European Simulation Interoperability Workshop, Toulouse, France.
- Petty, M. D., & Weisel, E. W. (2019). Chapter 4 - Model composition and reuse. In L. Zhang, B. P. Zeigler, & Y. Laili (Eds.), *Model Engineering for Simulation* (pp. 57–85).
- Rouse, W. B. (2019). *Computing possible futures: Model based explorations of "What if?"* Oxford University Press.
- Rouse, W. B. (2022). *Transforming public-private ecosystems: Understanding and enabling innovation in complex systems*. Oxford University Press.
- Schoonenberg, W. C., Khayal, I. S., & Farid, A. M. (2019). *A hetero-functional graph theory for modeling interdependent smart city infrastructure*. Springer.
- Singh, R., & Mathirajan, M. (2014, December 9–12). *A conceptual simulation framework for the performance assessment of lot release policies*. Paper presented at the 2014 IEEE International Conference on Industrial Engineering and Engineering Management.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>





# **Towards an Enterprise All-Domain M&S Environment for T&E: Overcoming M&S Challenges Within the DoD**

**Jeremy Werner**—was appointed DOT&E’s Chief Scientist in December 2021 after initially starting at DOT&E as an Action Officer for Naval Warfare in August 2021. Before then, Werner founded a data science-oriented military operations research team at JHU/APL that transformed the analytics of an ongoing military mission. Werner previously served at IDA supporting DOT&E in the assessment of a variety of systems. Werner earned a PhD in physics from Princeton University where he was an integral contributor to the Compact Muon Solenoid collaboration in the experimental discovery of the Higgs boson at the Large Hadron Collider, CERN, Geneva, Switzerland. [jeremy.s.werner.civ@mail.mil]

## **Abstract**

Our earlier paper Data Driven Modeling and Simulation to Test the Internet of War Things described how (Werner, 2023):

The Director, Operational Test and Evaluation’s Strategic Initiatives, Policy, and Emerging Technologies division (DOT&E SIPET) is shaping the test and evaluation (T&E) of future multi-domain warfighting.

and how:

Comprehensive Live testing of multi-domain capabilities currently under development is not possible due to environmental, fiscal, safety, classification, and ethical constraints, and so our evaluations will become more dependent on modeling and simulation (M&S) to test the efficacy and interoperability of our systems.

That paper was written for a general audience but still explored the following technical challenges:

- Architecting M&S and live tests to engender a “predict, live test, refine” feedback loop to improve M&S accuracy over systems’ life cycles
- Ensuring integration across all warfighting domains and digital capabilities
- Fielding M&S as a service so that the skillset required to operate it and understand its outputs mirrors the skills required of warfighters in the real world
- Implementing an environment with real-time analysis and accurate results for T&E and operational decisions

The present paper complements the earlier one by targeting a DoD technical audience and addressing several more challenges we see within the DoD:

- Using M&S to credibly extrapolate outside of the operational envelope covered in live test
- Rigorous life cycle approaches to V&V that are centered around quantitative estimates of uncertainty
- Accelerating M&S processing times
- The risk that we do all of this rigorous work and our models still turn out to be wrong
- Providing policy, guidance, best practices, executable examples, and training an M&S V&V/Uncertainty Quantification (UQ) workforce within the DoD.

that we must overcome to outpace our adversaries’ capabilities.

## **Introduction**

The need to maximally leverage modern M&S solutions for the T&E of multi-domain warfighting capabilities is manifest, as comprehensive live testing of these capabilities is not possible due to environmental, fiscal, safety, classification, and ethical constraints. Nevertheless, we must overcome a multitude of challenges to most effectively utilize M&S for



the T&E of these capabilities. Our earlier article *Data Driven Modeling and Simulation to Test the Internet of War Things* discussed means for overcoming several of these challenges while being addressed to a general audience; the article at hand complements the earlier one by targeting a DoD technical audience to overcome several more challenges we see within our department:

- Providing policy, guidance, best practices, executable examples, and training an M&S V&V workforce within the DoD
- Using M&S to credibly extrapolate outside of the operational envelope covered in live test
- Rigorous life cycle approaches to V&V that are centered around quantitative estimates of uncertainty
- Accelerating M&S processing times
- The risk that we do all of this rigorous work and our models still turn out to be wrong
- Providing policy, guidance, best practices, executable examples, and training an M&S V&V/Uncertainty Quantification (UQ) workforce within the DoD.

### **Using M&S to Credibly Extrapolate Outside of the Operational Envelope Covered in Live Test**

Deducing models that accurately characterize phenomena far beyond and more generally than just those limited input observations from which they were originally derived has been a cornerstone of science, engineering, and technology for millennia. The Antikythera mechanism built by the ancient Greeks is the earliest known mechanical computer and was able to accurately predict astronomical positions and eclipses decades in advance. Isaac Newton was able to glean the mechanics underlying the motion of all bodies in the universe from the limited number of terrestrial-based observations available to him (that is, at least, until one looks at either the quantum or relativistic regimes—more on that later).

In general, the fact that a small handful of physical laws can be used to accurately characterize and predict phenomena across a vast—even infinite—set of input conditions is a crowning achievement of science. And one with strong implications for T&E since it means:

It is possible to apply firmly-established physical laws joined to a limited number of live test observations to credibly assess system performance in regions of the operational envelope not directly covered in test.

### **National Aeronautics and Space Administration (NASA) Example<sup>1</sup>**

NASA Langley's 14'×22' subsonic wind tunnel provides an excellent example. The wind tunnel is used to assess conventional performance for low-speed tests of powered and unpowered models of various fixed- and rotary-wing civil and military aircraft over a wide range of takeoff, landing, cruise, and high-angle-of-attack conditions. The 14'×22' wind tunnel is ideally suited for low-speed tests to determine high-lift stability and control, aerodynamic performance, rotorcraft acoustics, turboprop performance, and basic-wake and flow-field surveys (NASA Aeronautics Test Program, 2009). Small-scale models of aircraft are tested in the tunnel and the results then scaled to the full-size platforms using a thoroughly vetted and continuously-validated computational fluid dynamics (CFD) model.

The ongoing feedback loop between live data and the CFD model is the key to success: Wind tunnel data is used to both calibrate and validate the model, while the model can then be used to provide accurate results scaled to the full-size platform as well as help identify important

---

<sup>1</sup> Special thanks to James Warner, NASA, for his help in formulating this example.



design constraints, air flows, and the like which in turn can then be tested in the wind tunnel. Of course, once the full-scale platform is eventually built and tested, then it's measured performance characteristics can also be compared to the design predictions and used to refine the prediction process as needed.

### **National Nuclear Security Administration (NNSA) Example<sup>2</sup>**

The Department of Energy's (DOE) NNSA Office of Defense Projects mission is to maintain a safe, secure, and effective nuclear weapons stockpile for our nation; its Advanced Simulation and Computing (ASC) program is not only vital to this mission but provides another excellent example for the DoD to follow. The ASC has its origins in the nation's ongoing need to maintain and assess the readiness of our stockpile following the discontinuation of live, underground nuclear explosion tests in the 1990s. Over the span of nearly three decades, the ASC has successfully developed and validated advanced simulation capabilities based on well-known physics of magnetohydrodynamics, inertial confinement fusion, structural dynamics, etc.—and then calibrated and validated them against current experiments and historical data from live underground testing—to credibly assure our nuclear weapon stockpile in the absence of live nuclear explosion tests.

We have provided both tactical (NASA) and strategic (NNSA) examples from our partner government agencies outside of the DoD showing that it is possible to apply firmly-established physical laws joined to a limited number of live test observations to credibly assess system performance in regions of the operational envelope not directly covered in test. DOT&E will seek collaboration and knowledge exchange with NASA and the NNSA, including national labs such as Sandia, Los Alamos, and Lawrence Livermore, so that we may adapt their M&S V&V methods to our mission.

### **Rigorous Life Cycle Approaches to V&V that are Centered Around Quantitative Estimates of Uncertainty**

Our previous article discussed architecting M&S and live tests to engender a “predict, live test, refine” feedback loop to improve M&S accuracy over systems' life cycles. This is aligned with DoD policy; in particular DODI 5001.61 (DoD, 2009) describes how it is DoD policy that, “Models, simulations, and associated data used to support DoD processes, products, and decisions shall undergo verification and validation (V&V) throughout their life cycles.”

Furthermore, the National Academies 2012 report *Assessing the Reliability of Complex Models* (National Academies, 2012) recognizes “the ubiquity of uncertainty in computational estimates of reality and the necessity for its quantification.”

That report also contains a treasure trove of examples and methods for our DoD M&S community to learn from and apply.

We agree with the National Academies on the criticality of uncertainty quantification (UQ) and view UQ as the principal essence of rigorous V&V. After all, even simple measurements are subject to imperfections caused by stochastic variation, calibration tolerances, etc., and so all measurement results must be associated with an uncertainty estimating these effects.

---

<sup>2</sup> Special thanks to Thuc Hoang, Director, Office of Advanced Simulation and Computing, NNSA for reviewing this example and all other references to NNSA throughout this article.



## DOT&E Example

The 2022 DOT&E article *Uncertainty Analysis Demonstration: A Missile Case Study* (Werner, 2022) provides the reader a primer on uncertainty, including a succinct description of the different types of uncertainty and their causes brought to life using the measurement of the mean weight of basketballs approved for NBA games as an example. From there, the article presents a case study of a notional missile performance analysis to demonstrate how statistical uncertainty in live test data—due to the limited number of samples from which it was generated—can be reinterpreted as a systematic uncertainty in the simulation. The article is available on DOT&E’s website and includes a link to download the executable code to reproduce all of the charts and findings it contains with a single command; in this way, the study is packaged so that it can be easily used and adapted to our community’s applications.

## NNSA Example

The National Nuclear Security Administration’s (NNSA) 2022 Advanced Simulation and Computing (ASC) Simulation Strategy (Etim, 2022) delves into the ASC’s mature uncertainty quantification capabilities and describes their uncertainty and margins framework. Central to the ASC’s strategy is “addressing the demand of uncertainty quantification efforts being performed routinely and more quickly by making them more user friendly and more easily integrated into daily practice;” the need for the DoD to do the same is apparent.

It is DoD policy that Models, simulations, and associated data used to support DoD processes, products, and decisions shall undergo V&V throughout their life cycles. The National Academies’ 2012 report *Assessing the Reliability of Complex Models* recognizes the ubiquity of uncertainty in computational estimates and the necessity for its quantification, while providing a treasure trove examples and methods for our DoD M&S community to learn from and apply. DOT&E’s 2022 article *Uncertainty Analysis Demonstration: A Missile Case Study* provides a primer on uncertainty and a case study demonstrating uncertainty quantification that is packaged with the executable code for our community to use and adapt. The NNSA’s 2022 Advanced Simulation and Computing Simulation Strategy delves into their mature uncertainty quantification capabilities and framework. That strategy recognizes the need to integrate uncertainty quantification into their analysis workflows in a more routine and user friendlier way; although our community’s uncertainty quantification capabilities are less mature than the NNSA’s, the need for us to do the same is apparent.

## Accelerating M&S Processing Times at the Hardware and Tactical Performance Levels

### Achieving Machine Speeds Using Hardware

The fact that custom tactical hardware and their associated integrated circuit boards such as Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) are much faster for their dedicated purposes than more generalized computer processors such as Central Processing Units (CPU) and Graphical Processing Units (GPU) presents a technical challenge to the development of an enterprise all-domain M&S environment for T&E. This can be easily understood by highlighting that the duration between input data and output response from a given logical algorithm can be shorter than 1 microsecond when implemented on an FPGA but may take 50 microseconds or more on a CPU (van der Ploeg, 2018).



These speed constraints are fundamental to the device architectures; no amount of CPU/GPU parallelization or clever software engineering can overcome them. (After all, that's why our tactical platforms have integrated circuit boards in the first place.) Nevertheless, the need for achieving hardware speed in M&S is paramount to synchronously include operators in the loop and other complicating, real-world tactical effects. Therefore, achieving performant M&S at the low levels and high fidelity of hardware necessitates that CPU/GPU-based M&S capabilities be augmented with tactical hardware in the loop (HWIL). From there, these HWIL-integrated capabilities can be used to feed machine learning and other advanced methods to generate reduced high-level tactical performance models that can run in real time.

### **Scaling Across the Enterprise Using Parallelized and Distributed Computing**

Fortunately, modern CPU- and GPU-based computing architectures do not present any fundamental limitations to the development of an enterprise all-domain M&S solution beyond the low level of hardware. Alternatively, modern computing architectures provide a wealth of solutions for us to exploit. CPUs, of course, are the primary engines of computer processing and modern hardware and software solutions enable CPU computations to be parallelized and scaled across vast numbers of cores; the large computational workloads and high throughput required of M&S can be accelerated by sharing the workload across a large number of CPU cores operating in parallel. Additionally, GPUs are natively massively parallelized and are faster and more energy efficient than CPUs for a variety of computational workloads; many industrial data centers are currently shifting their infrastructures to include more GPUs for these reasons. GPUs are particularly efficient at the rendering used in a variety of M&S solutions, and have enabled real time 3D ray tracing in video games and the production of movies that entirely use photo-realistic Computer Generated Imagery such as Disney's 2019 *The Lion King*. Finally, both CPU- and GPU-based parallelized computing capabilities can be distributed across large geographical areas and have disparate, asynchronous workflows integrated and brought into harmony using modern, enterprise software architectural solutions such as RESTful APIs (Amazon, n.d.).

Achieving performant M&S at the low levels and high fidelity of hardware necessitates that CPU/GPU-based M&S capabilities be augmented with tactical hardware in the loop (HWIL). From there, these HWIL-integrated capabilities can be used to feed machine learning and other advanced methods to generate reduced high-level tactical performance models that can run in real time. Beyond the low level of hardware, the large computational workloads and high throughput required of M&S can be accelerated by sharing the workload across a large number of CPU and GPU cores operating in parallel. Furthermore, both CPU- and GPU-based parallelized computing capabilities can be distributed across large geographical areas and have disparate, asynchronous workflows integrated and brought into harmony using modern, enterprise software architectural solutions such as RESTful APIs.

### **The Risk That We Do All of this Rigorous Work and Our Models Still Turn Out to be Wrong**

It took humanity until the 1680s to consolidate all of our observations dating from antiquity into three fundamental laws that describe the mechanics underlying the motion of all bodies in the universe. Or so we thought. Around 1890 hints of Newton's laws breaking down towards the speed of light began to appear. Then around 1900 more hints of Newton's laws breaking down—this time at microscopic distance scales—began to appear. But by then the scientific method had been fully institutionalized. Einstein's theory of relativity solved the speed of light problem in 1905, effectively extending our foundational laws of mechanics to their light



speed limit. By the mid-1920s quantum mechanics had been fully formulated by Niels Bohr, Erwin Schrodinger, and a cast of many others to extend our understanding of mechanics to the microscopic scale.

The institutionalization of the scientific method meant that it only took a couple of decades to mend cracks in physical laws that took millennia to formulate in the first place. And then of course it took just a couple of more decades to realize the weaponization of these new physics through the advent of nuclear weapons and propulsion.

Now let's return to the topic at hand: Using M&S for the rigorous T&E of our military systems and future joint warfighting concepts, with the ultimate goal of reducing the risk posed to our warfighters. Warfare is wrought with risk and uncertainty; real-world combat data may prove our models wrong despite our best efforts. But by design—by institutionalizing the rigor, mechanisms, and processes discussed in this article, its preceding one, and many related efforts to follow—we will have most thoroughly prepared ourselves for this exact eventuality. Just as the scientists of the early 20th century were well-prepared to reformulate physics for entirely new and unexpected regimes at an incredibly rapid pace, so too must our community be prepared to quickly ingest data from all venues—including real-world combat operations—to rapidly adapt our models. But to truly be prepared for these critical risks, we need to move out with agility now to build and stress this enterprise all-domain M&S plumbing. The discoveries we make along the way could be surprising and profound for warfare; after all, they have been before.

Warfare is wrought with risk and uncertainty; real-world combat data may prove our models wrong despite our best efforts. The institutionalization of the scientific method meant that it only took a couple of decades to mend cracks in physical laws that took millennia to formulate in the first place, and this eventually led to nuclear weapons and propulsion. By institutionalizing the rigor, mechanisms, and processes discussed in this article, its preceding one, and many related efforts to follow, our community will be best prepared to quickly ingest data from all venues—including real-world combat operations—to rapidly adapt our models as needed. To truly be prepared for these critical risks, we need to move out with agility now to build and stress this enterprise all-domain M&S environment. The discoveries we make along the way could be surprising and profound for warfare; after all, they have been before.

## **Providing Policy, Guidance, Best Practices, Examples, and Training an M&S V&V/Uncertainty Quantification (UQ) Workforce Within the DoD**

DOT&E is taking action to provide updated policy, guidance, best practices, and executable examples for M&S V&V and UQ:

- We are in the process of updating our 2016 and 2017 policy memoranda (DOT&E, 2016, 2017) on M&S V&V and consolidating them into a single, expanded DoD Manual (DODM) that will be released later this year.
- Later this year, we will start developing a M&S V&V UQ Companion Guide that will include a wide array of best practices and examples; these examples will each be provided along with the executable code to fully reproduce them with a single command, inclusive of all charts, tables, and numerical results; in this way, each example will be packaged in such a way that it can be easily used and adapted to our community's applications. Many of these best practices and examples can already be found in IDA's *2019 Handbook on Statistical Design & Analysis Techniques for Modeling & Simulation*



*Validation*, which also describes a rigorous methodology for planning tests that utilize both M&S and live test data for evaluation<sup>3</sup>.

We further recognize the need to train a dedicated M&S V&V UQ workforce and to learn from V&V UQ communities outside of our department—NASA and the DOE NNSA enterprise for nuclear weapons stockpile assurance, in particular:

- DOT&E will seek collaboration and knowledge exchange with NASA and the NNSA, including national labs such as Sandia, Los Alamos, and Lawrence Livermore, so that we may adapt their M&S V&V UQ methods to our mission.
- DOT&E will pursue a training curriculum targeted at creating a dedicated M&S V&V UQ workforce for the T&E enterprise in conjunction with the Under Secretary of Defense for Research and Engineering and the Services.

Preparing our nation for the next generation of T&E capabilities means providing updated M&S policy, guidance, best practices, executable examples, and the training of a dedicated M&S V&V UQ workforce; DOT&E is taking a number of actions to fulfill these needs and will pursue a training curriculum to create a dedicated M&S V&V UQ workforce.

## Conclusion

DOT&E's SIPET division—Strategic Initiatives, Policy, and Emerging Technologies—is shaping the T&E of future multi-domain warfighting. We understand that comprehensive Live testing of multi-domain capabilities currently under development is not possible due to environmental, fiscal, safety, classification, and ethical constraints and so our evaluations will become more dependent on modeling and simulation (M&S) to test the efficacy and interoperability of our systems. Our earlier article *Data Driven Modeling and Simulation to Test the Internet of War Things* discussed means for overcoming the following M&S challenges while being addressed to a general audience:

- Architecting M&S and live tests to engender a “predict, live test, refine” feedback loop to improve M&S accuracy over systems’ life cycles
- Ensuring integration across all warfighting domains and digital capabilities
- Fielding M&S as a service so that the skillset required to operate it and understand its outputs mirrors the skills required of warfighters in the real world
- Implementing an environment with real-time analysis and accurate results for T&E and operational decisions.

The article at hand complemented the earlier one by targeting a DoD technical audience to overcome several more challenges we see within our department:

- Using M&S to credibly extrapolate outside of the operational envelope covered in live test
- Rigorous life cycle approaches to V&V that are centered around quantitative estimates of uncertainty
- Accelerating M&S processing times
- The risk that we do all of this rigorous work and our models still turn out to be wrong
- Providing policy, guidance, best practices, executable examples, and training an M&S V&V/UQ workforce within the DoD.

---

<sup>3</sup> Many of these best practices and examples can already be found in Wojton (2019).



The exposition of these challenges was not academic; alternatively, the discussion was pragmatic and centered around already-mature or rapidly maturing technologies, advanced methods, and real-world use cases pertinent to the M&S required for the T&E of future joint warfighting concepts. DOT&E and our partners will soon have multiple R&D projects underway to advance our M&S and V&V/UQ capabilities for T&E and position us to meet this critical challenge; we have released our S&T strategy to the public (<https://www.dote.osd.mil/News/News-Display/Article/3118739/dote-strategy-update-2022/>) as well as our detailed implementation plan.

We will do our part by transforming M&S and its V&V/UQ for T&E to enable delivery of the world's most advanced warfighting capabilities at the speed of need. We seek your proposals to collaborate with us.

## References

- Amazon. (n.d.). *What is a RESTful API?*. <https://aws.amazon.com/what-is/restful-api/#:~:text=RESTful%20API%20is%20an%20interface,applications%20to%20perform%20various%20tasks>
- DoD. (2009). *DODI 5000.61, incorporating a change from 2018, DoD modeling and simulation (M&S) verification, validation, and accreditation (VV&A)*.
- DOT&E. (2016). *Guidance on the validation of models and simulation used in operational test and live fire assessments*. [https://www.dote.osd.mil/Portals/97/pub/policies/2016/20140314\\_Guidance\\_on\\_Valid\\_of\\_Mod\\_Sim\\_used\\_in\\_OT\\_and\\_LF\\_Assess\\_\(10601\).pdf?ver=2019-08-19-144201-107](https://www.dote.osd.mil/Portals/97/pub/policies/2016/20140314_Guidance_on_Valid_of_Mod_Sim_used_in_OT_and_LF_Assess_(10601).pdf?ver=2019-08-19-144201-107)
- DOT&E, 2017, *Clarifications on Guidance on the Validation of Models and Simulation used in Operational Test and Live Fire Assessments*, [https://www.dote.osd.mil/Portals/97/pub/policies/2017/20170117\\_Clarification\\_on\\_Guidance\\_on\\_the\\_Validation\\_of\\_ModSim\\_used\\_in\\_OT\\_and\\_LF\\_Assess\(15520\).pdf?ver=2019-08-19-144121-123](https://www.dote.osd.mil/Portals/97/pub/policies/2017/20170117_Clarification_on_Guidance_on_the_Validation_of_ModSim_used_in_OT_and_LF_Assess(15520).pdf?ver=2019-08-19-144121-123)
- Etim. (2022). *Advanced computing and simulation, NNSA, LLNL, simulation strategy*. <https://asc.llnl.gov/file-download/download/public/2731>
- National Academies. (2012). *Assessing the reliability of complex models*. <https://nap.nationalacademies.org/catalog/13395/assessing-the-reliability-of-complex-models-mathematical-and-statistical-foundations>
- NASA Aeronautics Test Program. (n.d.). *14- by 22-Foot Subsonic Tunnel*. [https://www.nasa.gov/sites/default/files/atoms/files/m187003\\_14\\_22print\\_508.pdf](https://www.nasa.gov/sites/default/files/atoms/files/m187003_14_22print_508.pdf)
- Stevens. (2022). *25 years of accomplishments*. Advanced Simulation Program, National Nuclear Security Administration. <https://www.energy.gov/sites/default/files/2022-10/20221013%20ASC%2025-year%20Accomplishments%20Report.pdf>
- van der Ploeg. (2018). *Why use an FPGA instead of a CPU or GPU?*. <https://blog.esciencecenter.nl/why-use-an-fpga-instead-of-a-cpu-or-gpu-b234cd4f309c>
- Werner. (2022). *Uncertainty analysis demonstration: A missile case study*. DOT&E. <https://www.dote.osd.mil/News/What-DOT-Es-Saying/Saying-Display/Article/3254156/uncertainty-analysis-demonstration-a-missile-case-study/>
- Werner. (2023). *Data driven modeling and simulation to test the internet of war things*. *MODSIM World Proceedings*.
- Wojton. (2019). *Handbook on statistical design & analysis techniques for modeling & simulation validation*. <https://apps.dtic.mil/sti/pdfs/AD1122387.pdf>





## PANEL 23. NEXT GENERATION PRIMES - MOVING FROM INNOVATION TO FIELDING

Thursday, May 11, 2023	
3:45 p.m. – 5:00 p.m.	<p><b>Chair: L. Neil Thurgood, LTG USA (Ret.)</b></p> <p><b><i>The Innovation Paradox – Merging Process with Disruptive Thinking to Accelerate Capability Transition to the War Fighter Through the Educational Innovation Capstone Process</i></b></p> <p style="text-align: center;">Raymond Jones, Naval Postgraduate School</p> <p><b><i>Assessing the Effectiveness of Defense-Sponsored Innovation Programs as a Means of Accelerating the Adoption of Innovation Forcewide</i></b></p> <p style="text-align: center;">Amanda Bresler, PW Communications Alex Bresler, PW Communications</p> <p><b><i>Leverage AI to Learn, Optimize, and Wargame (LAILOW) for Strategic Laydown and Dispersal (SLD) of the Operating Forces of the U.S. Navy</i></b></p> <p style="text-align: center;">Ying Zhao, Naval Postgraduate School Doug MacKinnon, Naval Postgraduate School</p>

**L. Neil Thurgood, LTG USA (Ret.)**—is a retired United States Army lieutenant general who last served as the director of Hypersonics, Directed Energy, Space, and Rapid Acquisition of the Office of the Assistant Secretary of the Army for Acquisition, Logistics, and Technology. He previously served as the Special Assistant to the Assistant Secretary of the Army for Acquisition, Logistics, and Technology.



# **The Innovation Paradox—Merging Process with Disruptive Thinking to Accelerate Capability Transition to the War Fighter Through the Educational Innovation Capstone Process**

**Raymond Jones, COL USA (Ret)**—retired as a Colonel from the U.S. Army in 2012 and is a Professor of Practice within the Graduate School of Defense Management at the U.S. Naval Postgraduate School in Monterey, CA. He also serves as a Guest Lecturer for the IDARM Program within the Institute for Security Governance (ISG), Defense Security Cooperation Agency (DSCA). 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. [rdjone1@nps.edu]

## **Abstract**

Innovation is the process of creating something new or improving an existing product, service, or process. In the national security environment, it is critical to ensuring operational and strategic overmatch against one's' adversaries. Without innovation in ideas and capabilities, nations lose their ability to outmaneuver their competitors and begin their ultimate decline into irrelevance on the world's stage. Innovation can take many forms. It can be the development of a new product or service intended to meet the needs of the end user or customer and It can also be the implementation of a new process that improves efficiency and productivity in an organization. Innovation can be incremental, such as small improvements to existing capabilities or services, or they can be disruptive, completely transforming an organization. Disruptive innovation tends to change the nature of warfare and are marked by paradigm shifts known as revolutions in military affairs.

Innovation is not without its challenges. It can be difficult to come up with new ideas, and even harder to turn those ideas into a successful product or service. It can also be challenging to manage the risks associated with innovation, such as the cost of research and development and the potential for failure. True innovation requires a strategy to transition the innovative idea into a usable capability that has a measurable impact of intended purpose. Therein lies the paradox of innovation. To realize true innovation, the curse of bureaucracy is necessary to allow the innovative thought and concept to move from an idea to an actionable capability. In effect, to transition an idea from concept across the "valley of death," a deliberate and sometimes slow and structured process is necessary to align all the competing interests that might otherwise crush the new idea, much like the immune system attacks a foreign object in one's body. Despite these challenges, innovation remains a critical element of progress and growth in society, business, and the military. It is the driving force behind many of the world's most successful institutions and has been responsible for some of the most significant technological advancements in human history.

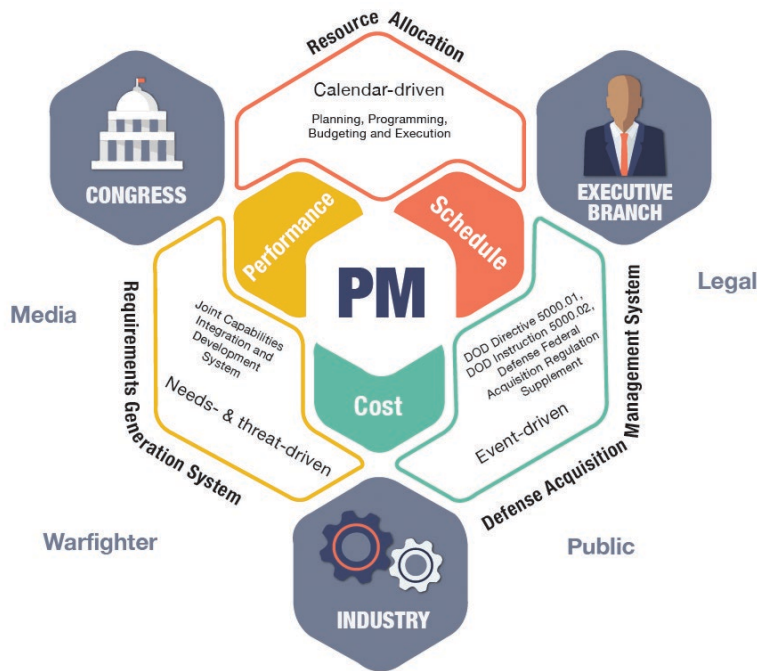
This paper will address the fundamental problem that most new ideas have regarding transitioning from a "good idea" to becoming a viable capability in the hands of the user. The problem most militaries have is that the process of capabilities development tends to take too long, is too costly, and lacks the agility to allow for innovative and disruptive ideas to gain a fold hold, once the acquisition process has started for specific needs of the warfighter. Additionally, many of the critical and disruptive ideas born in the "foxhole" tend to die in place for lack of a



clear pathway out of the foxhole. I will seek to define these challenges and present a new pathway to successfully cross the valley of death, beyond the traditional six pathways defined in the Department of Defense 5000 instruction. While these pathways appear to provide a well-defined and deliberate approach to technology maturation and innovation, they lack the opportunity to tap into disruptive innovation rapidly and in a way that supports both government and industry. In essence the current methods of transitioning innovative ideas is simply not robust enough for the rapidly changing dynamics of the future national security environment and it is time to change the paradigm and embrace the innovation paradox.

## The Valley of Death

Defense acquisition is the process through which the United States Department of Defense (DoD) acquires goods and services, including weapons, equipment, and technology, to support the nation’s security and military operations. However, the defense acquisition process is often fraught with challenges that can impede its effectiveness and efficiency. A significant challenge facing defense acquisition is the sheer complexity of the process. The defense acquisition process involves a vast number of stakeholders, including the DoD, industry partners, Congress, and the public.



**Figure 1. Defense Acquisition Stakeholder Environment**

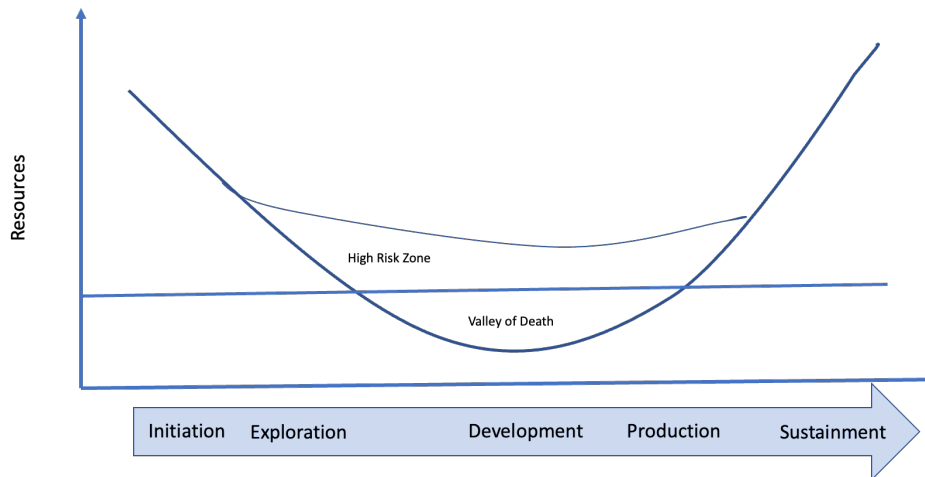
Army ALT Magazine (<https://asc.army.mil/web/news-big-a-acquisition-a-primer/>)

Each of these stakeholders has their own set of priorities, interests, and requirements, making it difficult to align everyone’s goals and objectives. Furthermore, the defense acquisition process involves numerous regulations, procedures, and documentation requirements that can be time-consuming and burdensome.

The valley of death is the gap between the development of new technologies and their successful delivery to the end user or customer. This gap is often referred to as the “valley”



because it represents a challenging period of uncertainty and risk for technology developers and investors. New ideas can enter the high-risk zone, as depicted in Figure 2, and quickly spiral below the point of no return and fall into the “valley of death.” There needs to be an organized effort and well-defined path for the new technology or concept to stay above the point of no return and be able to transition into production and resource investment. The valley of death is particularly relevant in technology development initiatives where the pace of innovation is rapid and the time it takes for new ideas to gain traction can be slow. The gap between the initial development of a new technology and its successful operationalization can be vast, with many new technologies failing to make it through this difficult phase.



**Figure 2. Technology Transition and the “Valley of Death”**

There are several reasons why technology transitions fail to cross the Valley of Death. One reason is that new technologies are often untested, with limited data available to demonstrate their effectiveness or safety. This lack of data can make it difficult to make informed decisions about whether to continue to invest in a new technology. From a defense acquisition perspective, the failure to invest often reflects the lack of willingness of programs of record to recognize the value of emerging technologies as they relate to specified requirements. In essence program managers suffer from requirements myopia by failing to see how new innovative ideas align to existing requirements within their portfolios. Without a resource sponsor or investor, the technology idea falls into the infamous valley of death and fails to realize the full potential of innovation.

Current defense acquisition pathways do little to encourage innovation once a program is established and contracts are awarded. In fact, innovation is discouraged if it strays from the well-defined acquisition strategy and contract agreements established with industry. Deviation from the “plan” is seen as a distraction that violates a prescribed performance baseline, particularly if the deviation comes from outside of the agreed upon contract relationships between the government and industry provider. In effect, once a strategy is approved, the baseline is agreed upon, and contracts are awarded, innovation stops. The plan is executed as prescribed with little introspective assessment of new ideas. New ideas bring risk and drive baseline variance with is a sure path to program failure. Therein lies the paradox. While innovation is born in an unconstrained environment of risk and experimentation, for new ideas to mature and evolve into real capability able to cross the valley of death, the same complex process of rules is the process that is necessary to facilitate the successful transition of innovative ideas that ultimately lead to innovation.



## Change at the Margins

The defense acquisition process is often fraught with inefficiencies, delays, and cost overruns, which can impact the readiness and effectiveness of the military. Typically, the defense department seeks new ways to improve the process, and while the department has a significant success record in developing new and capable systems, many more potential opportunities are lost due to the arduousness of the process. Typical approaches that are used to improve a perceived process problem include:

1. **Streamline the procurement process.** The procurement process is often complicated and bureaucratic, leading to delays and increased costs. Streamlining the process can reduce the time and resources required to procure equipment and services, resulting in cost savings and improved readiness. This can be achieved by simplifying the procurement requirements, consolidating procurement efforts, and reducing paperwork and administrative burdens.
2. **Increase competition.** Competition is essential in any procurement process, and the defense acquisition process is no exception. Increased competition can lead to lower costs, improved quality, and better innovation. To increase competition, the government can promote the participation of small businesses and minority-owned businesses and encourage collaboration between industry and academia to foster innovation.
3. **Foster collaboration and communication.** Collaboration and communication between government and industry are crucial to the success of the acquisition process. By working together, they can identify potential risks, mitigate them early on, and find innovative solutions to procurement challenges. Regular communication and collaboration between government and industry can also help identify best practices, reduce redundancies, and streamline the procurement process.
4. **Invest in technology and data analytics.** Technology and data analytics can improve the acquisition process by providing real-time data and insights that can inform decision-making. This can help identify cost savings, reduce inefficiencies, and improve the quality of equipment and services procured. By investing in technology and data analytics, the government can also increase transparency and accountability in the procurement process.
5. **Implement performance-based contracting.** Performance-based contracting is a procurement approach that focuses on the outcome rather than the process. It allows the government to specify the desired results and leaves it up to the contractor to determine how best to achieve those results. This approach can incentivize contractors to find innovative solutions and reduce costs, resulting in improved quality and efficiency.

While these are excellent process improvement techniques, none of this address the root issue of how to enhance innovation. These typical solutions, address the symptoms and not the root cause that new and disruptive ideas simply have a difficult time of entering the process and finding a “sponsor” that can accelerate disruptive ideas and technologies. The DoD has attempted to address the overall process by redefining the acquisition process. A new approach to describing the fundamental process of meeting a “customer’s” needs was drafted and marketed as a different way to speed up technology transition. This process, referred to as the Adaptive Acquisition Framework (AAF) provides clear pathways to a system that has always allowed for agility and tailoring as appropriate to meet the needs of the user.



## Adaptive Acquisition

The defense acquisition process requires an adaptive approach that is agile and responsive enough to change with the evolving threat and speed of technology. Over the years, the DoD has been criticized for being slow, bureaucratic, and inefficient. In response to these criticisms, the DoD has implemented a number of reforms, including the Adaptive Acquisition Framework (AAF), which is designed to make the acquisition process more efficient and effective. The AAF is a set of guidelines and procedures that are designed to make the DoD's acquisition process more agile and responsive to changing requirements. The AAF was introduced in 2018, and it is based on the principles of the DoD's Better Buying Power initiative. The AAF is intended to be a flexible framework that can be adapted to different acquisition programs and situations.

The AAF is divided into three phases and six pathways: the Explore and Engage phase, the Assess and Approve phase, and the Execute and Deliver phase. Each of these phases includes a number of steps and activities that are designed to ensure that the acquisition process is efficient, effective, and responsive to changing requirements. The Explore and Engage phase is focused on identifying the user's needs and requirements, as well as identifying potential solutions and vendors. This phase includes activities such as market research, engagement with industry partners, and development of the acquisition strategy. The Assess and Approve phase is focused on evaluating potential solutions and vendors and selecting the best option. This phase includes activities such as requirements development, solicitation of proposals, and source selection. The Execute and Deliver phase is focused on implementing and delivering the chosen solution. This phase includes activities such as contract management, testing and evaluation, and delivery and sustainment.

There are six distinct pathways (Figure 3) in the AAF that are designed to align to a potential system of processes maturity level. These pathways represent derivatives of the Major Defense Acquisition process and are designed to mitigate potential inertia program managers could encounter. Each pathway addresses the maturity and type of capability being developed or procured. A common misconception of the AAF, is that choosing an alternative to the Major Capability Acquisition process, allows programs to avoid some regulatory and statutory requirements. While the pathways help to structure a program acquisition strategy relative to its maturity, all specified regulations and statutes are still required to be met or justified. The AAF is a convenient way to show the relationship between system maturity, urgency of need, and time.



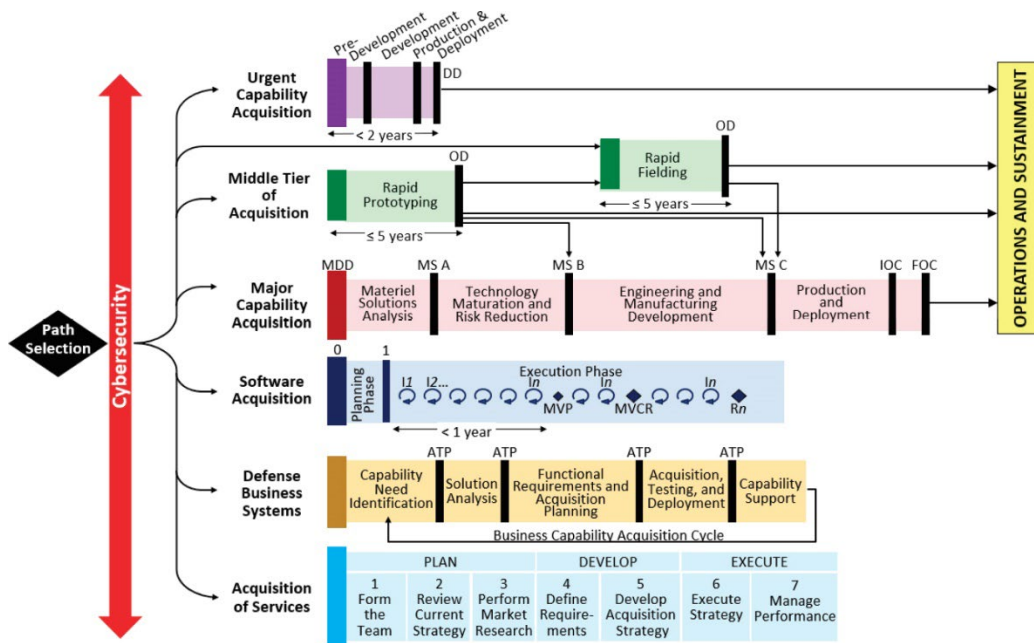


Figure 3. Adaptive Acquisition Framework (AAF) Pathways

One of the benefits of the AAF is that it provides a more flexible and agile acquisition process, allowing the DoD to be more responsive to changing requirements and to adapt to new technologies and threats. While this is accurate, the ability to do this has always been available to acquisition professionals and leaders. The AAF helps the acquisition leader align their technologies to a prescribed strategy, but it does not allow any more agility to rapidly introduce new innovative technologies into existing programs of record. Additionally, the AAF does not address the most critical challenge of how to leverage small innovative companies that have limited resources to compete in an environment that is not suited to speed of thought and innovation. The AAF simply provides more fidelity and definition to the status quo.

Providing more definition to the process seems like a good idea, but has ultimately created a more risk adverse environment focused on compliance centered leaders that are focused on the management of programs at the expense of adaptive and creative leadership of programs. One simply needs to dissect the many program failures to see that part of the root cause of failure lies in program teams managing through compliance rather than making the case for talking calculated and informed risk. Data suggests that the leading cause affecting PM decision-making is the restrictions imposed through processes and oversight within the acquisition environment. While years of acquisition reform, such as AAF, have aided in process maturity within the DoD, these reform efforts to improve oversight, reduce risk, and aid cost control may drive negative, unintended consequences (Neterer & Petrone, 2018).

Compliance centered leadership is preventing critical technologies from crossing the “valley of death” because of inflexible and misinterpretation of the fundamental purpose to the technology mature nation phase of the product development. Requirements are too quickly tied to specific technologies under contract leaving little room for new ideas for fear of violating a “rule” that funds can only be spent of specific, well defined, requirements. As a result, once a contractor with their specific solution is selected innovation in new ideas stops. Project managers are now bound by statutory baselines that are tied to specific technologies. Additionally, smaller companies are not able to compete in this system that is bias towards compliance and risk aversion rather than innovation, agility, and adoption.



The Adaptive Acquisition Framework is an important initiative that is designed to improve the efficiency and effectiveness of the DoD's acquisition process, it still falls short on providing effective opportunities for innovative disruptive ideas and technologies to make their way into programs that are already under way or that do not fit into one of the prescriptive pathways of AAF. Innovation requires disruptive ideas and solutions to be able to gain a foot hold into more structured developmental efforts to optimize capability at the right time and place. DoD developmental programs are inherently designed to seek stability and are even punished for varying from the "approved" performance and technical baseline. The AAF represents a refinement of the status quo and simply helps the acquisition process categorize technologies and processes based upon their maturity levels. It does not allow for the rapid insertion of change once a program is established, leaving small business innovative thinkers left to struggle with finding a "transition" agent to help them move their ideas into more mature and producible capabilities.

### The Innovation Paradox

When organizations and individuals recognize the need for innovation yet struggle to implement change and progress effectively, they are often experiencing the tension between the rigidity of the process and unstructured critical thinking. Paradoxically, to innovate both are necessary. The deliberate and often slow linear thinking of process is necessary to ensure disruptive ideas are shaped in ways that are not threats to the system. To be effective in the business of warfighting, disruptive thinking that moves inside the OODA loop of the adversary requires a deep understanding of the complexity of the business battlespace. Failure to navigate this space leads to missed steps and opportunities in the quest for resources and advocacy across the instruments of national power.

When individuals or organizations try something new, there is always a chance that it will not work out as planned, which can lead to wasted time, money, and effort. Risk aversion can make it difficult to take the necessary steps to innovate and can cause organizations to give up and fail to deliver critical capability.

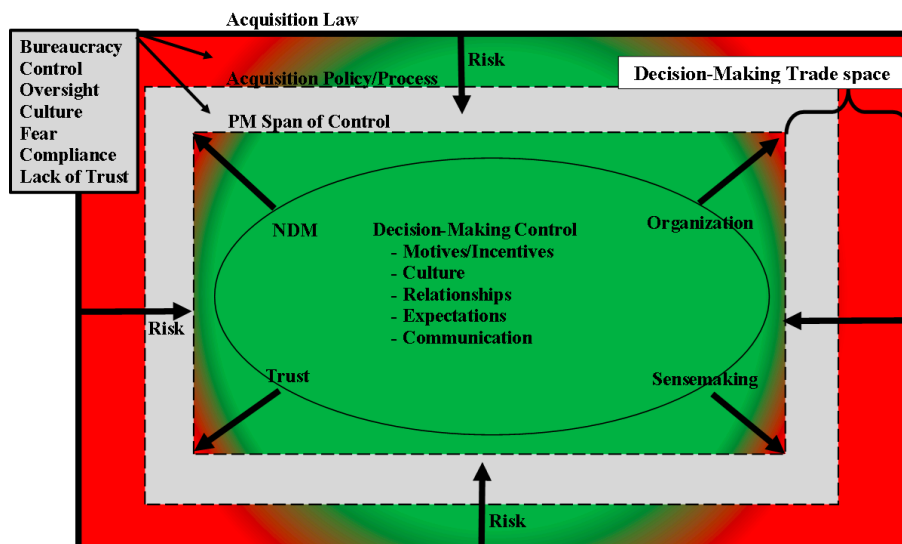


Figure 4. Risk Relationship to Decision Making in the Defense Acquisition Environment. (Neterer & Petrone, 2018).





Risk aversion is often linked to a lack of understanding or confidence in one's knowledge of the process. Those with insight of the technical vision, often lack the insight and clarity in the business process necessary to navigate through the vast sea of bureaucrats that are necessary to move an idea from inception to realization. The process is an essential part of the innovation process that allows the innovator to succeed. Figure 4 shows the relationship between culture, bias, insight and awareness, and risk suggesting that the more perceived risk the more likely a program team will become more process myopic and less innovative. In essence, the system is not structured to introduce new ideas that are not in the technology roadmap, limiting the significant opportunity and technology transition of new ideas.

Another reason why the innovation paradox exists is that innovation requires a certain degree of creativity and outside-the-box thinking. Unfortunately, many individuals and organizations are not naturally inclined towards these qualities and may struggle to come up with truly innovative ideas. Additionally, innovation often requires a willingness to challenge the status quo, which can be difficult for individuals and organizations that have become comfortable with their current methods and practices. Systemic in this is the basic lack of detailed knowledge of the process. All too often the process is blamed as the reason for program failures and the lack of visionary thinking that leads to effective innovation. A detailed and integrative understanding of the process is required to be able to think "out of the box" within a very structured process. The tendency is to allow the "system" to take over.

Bold leadership and innovative process adaptation is necessary to drive new concepts into a very deliberate and structured environment. The procurement process will adapt to new ideas for those that are able to take advantage of the inherent mechanisms within the process. In order to successfully do this however, requires in depth and critical understanding of the process and a leadership culture that encourages informed risk rather than a compliance management mindset. Managers keep the trains on time and leaders keep them going in the right direction. The process is a necessary component of innovation and leaders with an intense understanding of the tools are needed to keep the creativity alive throughout the process. To effectively navigate the innovation paradox, it is important to strike a balance between the need for creativity and the need for structure and stability. Ultimately, the innovation paradox is a complex and multi-faceted issue, but one that must be addressed if individuals and organizations hope to remain competitive and relevant in the modern world. By recognizing the challenges associated with innovation and taking steps to overcome them, it is possible to harness the power of innovation and drive meaningful progress and growth.

## **Educational Innovation Pathway**

The first step in creating a successful innovation strategy is to identify and understand the strategic and operational needs of the customer. In the defense environment, a market analysis is considered the requirements analysis, which seeks to identify gaps in capability needed to meet national security objectives. This involves conducting research and analyzing operational user or customer feedback to identify areas where improvements can be made or where new capability or services are needed. By understanding the strategic and operational needs, organizations can create capabilities and services that are not only relevant but also meet the needs of their target audience.

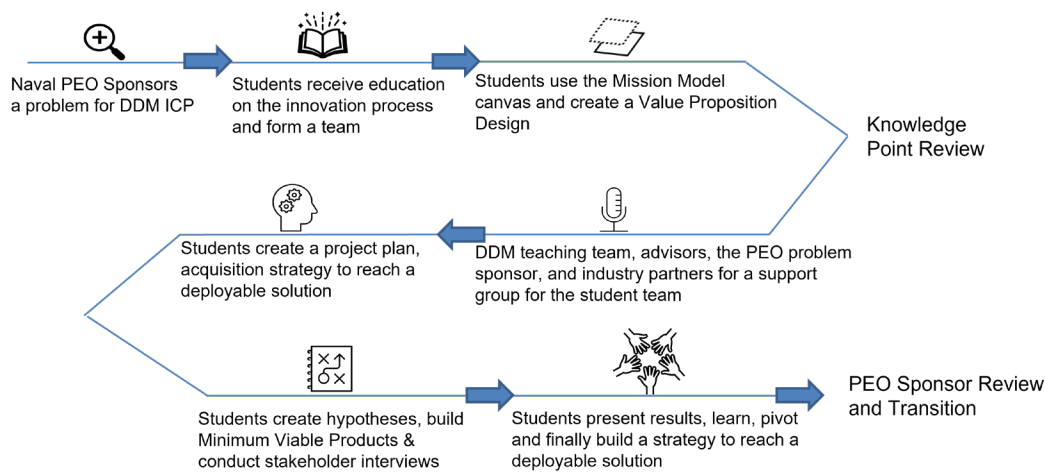
Once an organization has identified the market and customer needs, the next step is to create a culture of innovation within the organization. This involves encouraging the sharing of ideas and providing them with the necessary resources and support to pursue those ideas. Fostering an environment of innovation involves creating an environment that fosters creativity and innovation, such as encouraging cross-functional collaboration. Cross functional collaboration should be integrated into the entire life cycle process for as to encourage



disruptive thinking and potential for novel insight. A successful innovation strategy also involves a willingness to take risks and experiment with new ideas anywhere is the life cycle of a system under development. This means accepting that not all ideas will succeed and being willing to learn from failures and pivot when necessary. It also means being open to feedback and continually iterating and improving upon ideas.

The Innovation Capstone Project (ICP) established by the Department of Defense Management (DDM) at the Naval Postgraduate School (NPS) seeks to integrate the principles of innovation into an educational pathway that merges student/faculty teams with industry and relevant DoD Program Executive Offices (PEOs) to rapidly shape potential solutions for critical operational requirements. Bringing these three institutional entities together early seeks to develop a capability strategy early with an eye toward integration into approved programs of record. A key enabler in this ICP process is the early creation and development of an acquisition strategy for technologies that are very early in their technology readiness levels. By thinking about how a system will transition into a production and sustainment phase, thought must be given to how the process will be influenced relative to the technology to gain adoption. Stakeholders from each phase of the life cycle must be considered and integrated into the plan early and often, not when the new technology is “ready to go.”

The process begins with a requirement from the operational forces or PEO. As requirements are vetted by the DDM ICP Program Manager, student cross functional teams (CFT) are formed and assigned to faculty mentors that will help guide them through the process. As the CFTs iterate the problem they create a proposed recommendation and acquisition strategy that defines the total life cycle strategy and impact of their recommendation. At that point the CFT plan and strategy is presented to a board of subject matter experts who assess the viability of the plan and strategy with the intent to proceed into the concept development phase of the ICP. CFTs that successfully meet the standards of the knowledge point review are then married up with relevant industry partners through agreement such as Partnership Intermediary Agreements (PIA) or Cooperative Research and Development Agreements (CRADA) and a PEO with the portfolio requirements appropriate to the technology being considered. The integration of industry and PEO at the first knowledge point is critical in that the industry partners provide the focused technical expertise, and the PEO provides the transition engine for the technology being developed. The CRT acquisition strategy provides the PEO with the specified business plan by which they can adopt the new technology, provided it can be shown to achieve a relevant technology readiness level. Figure 5 summarizes the Educational Innovation Pathway.



**Figure 5. Innovation Capstone Program Pathway**



The value proposition for this Educational Pathway is the early development of the acquisition strategy and integration of industry and PEO in the process during the process of exploration. This allows small innovative companies to gain traction with government title 10 program offices early and it allows the PEOs to begin to shape the business strategy for adoption into existing programs of record. From an educational perspective, students at the Naval Postgraduate School begin to work with organizations that are responsible for developing and delivering technology and start to build the scaffolding of creative thought that will follow them into operational positions as they progress through their careers.

Leveraging the educational process for both learning and technology evolution is foundational to experiential learning and provides valuable insight to the institutions responsible for ensuring that the national security posture of the United States stays far ahead of any current and future adversary. The tangible products are the technologies that transition. The less obvious products are the students that learn to think more creatively and deeply about complex problems as well as the relationships that are developed between the DoD and small competent companies that find dealing with the DoD challenging at best.

Developing successful innovation strategies is critical for organizations looking to stay competitive and relevant in today's marketplace. This requires identifying and understanding market and customer needs, creating a culture of innovation, using technology and data analytics, taking risks and experimenting with new ideas, and having a strong leadership team committed to innovation. By incorporating these key components into their innovation strategy, organizations can achieve their goals and create products and services that meet the needs of their target audience.

## References

DoDI 5000.02. (2020). *Operation of the adaptive acquisition framework*.

Neterer J., & Petrone, M. (2018). *Why DO programs fail? An Analysis of defense program manager decision making in complex and chaotic environments*. Naval Postgraduate School.



# Assessing the Effectiveness of Defense-Sponsored Innovation Programs as a Means of Accelerating the Adoption of Innovation Force Wide

**Amanda Bresler**—serves as Chief Strategy Officer for PW Communications Inc. She runs the company’s Future Capabilities Practice, focused on helping the Department of Defense improve its ability to identify, access, and retain innovative solutions providers. Prior to joining PW Communications, she served 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; Chatham International Inc.; and AlmaLinks. 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. Bresler is a data-driven technologist, investor, and advisor to world-class high-tech companies. He is an expert 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

The Department of Defense (DoD) invests billions of dollars into innovation programs every year. One primary objective of these programs is to accelerate the adoption of critical new technologies force wide. This paper assesses the extent to which companies funded through defense-sponsored innovation programs (“program participants”), specifically the DoD Small Business Innovation Research (SBIR) program, subsequently deliver their capabilities to the warfighter. By analyzing millions of contracting and subcontracting actions associated with thousands of program participants, we demonstrate that the DoD awards most SBIR funding to a small subset of program participants. Furthermore, companies in receipt of the greatest share of overall program funding are among the least likely to transition their technologies to the warfighter. We analyzed the structure of DoD SBIR to identify potential causes for this poor rate of inter-government technology transition. We determined that this outcome results from misaligned incentives, antiquated policies and regulations, anticompetitive solicitation processes, and the absence of thoughtful, standardized metrics for defining and measuring programmatic success. In conclusion, we offer a series of concrete recommendations to address these issues and position DoD SBIR to more effectively deliver capabilities to the warfighter.

## Introduction

The Department of Defense (DoD) invests billions of dollars annually into innovation programs with the stated objective of enabling and/or accelerating the adoption of cutting-edge technologies. However, the DoD does not consistently track how companies engaged in these innovation programs (program participants) perform in the defense market, subsequent to program completion. Our research aimed to fill this gap by evaluating the extent to which program participants’ capabilities were subsequently procured by the DoD, either directly or indirectly.

While the DoD funds dozens of innovation programs, we focused our research on the Small Business Innovation Research/Small Business Technology Transfer (SBIR/STTR) program for several reasons. One primary goal of the SBIR/STTR program is to “support scientific excellence and technological innovation through the investment of Federal research funds in critical American priorities to build a strong national economy and accelerate capabilities to the warfighter” (*DoD Small Business Innovation Research / Small Business Technology Transfer, n.d.*). Other program objectives include investing in research and development (R&D) that has the potential for commercialization and encouraging “participation



in innovation and entrepreneurship by women and socially or economically disadvantaged persons” (*About, n.d.*). The DoD receives more than 50% of the entire more than \$4 billion SBIR/STTR budget annually, making it the largest DoD innovation initiative.

**Transitioning state-of-the-art capabilities to the warfighter must be the priority of the DoD SBIR/STTR program.** In decades past, the DoD was at the forefront of technological innovation and exported *its* technologies to the commercial sector. Today, companies outside of the traditional defense industrial base are driving advancements in areas critical to our national defense. The DoD must identify, engage, and retain these suppliers. Furthermore, as noted by former Secretary of Defense James Mattis in the 2018 National Defense Strategy, “Success no longer goes to the country that develops a new technology first, but rather to the one that better integrates it and adapts its way of fighting” (DoD, 2018). It is not enough for the DoD to simply invest in cutting-edge capabilities; it must integrate these capabilities force wide as quickly and seamlessly as possible.

From a research perspective, SBIR/STTR awards are explicitly identified in procurement data, enabling us to quantitatively analyze a wide range of information about program participants—including subsequent defense business—in great detail. By comparison, no consistent, publicly-available data exists to indicate whether a company participated in other DoD innovation programs.

### Analyzing the SBIR//STTR Program

To assess the extent to which investments in the SBIR/STTR program have resulted in new capabilities reaching the warfighter, we first needed to isolate a data set of entities that won defense-funded SBIR/STTR awards for analysis (SBIR companies, DoD SBIR companies). To do so, we aggregated SBIR/STTR 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 include defense-funded Phase I/Phase II SBIR awards from fiscal year (FY) 2012 through FY 2021.

We determined that there were 52,746 defense-funded Phase I/Phase II SBIR/STTR awards totaling approximately \$13.1 billion from FY 2012–FY 2021. Table 1 provides a breakdown of the count and total dollar value of DoD-funded Phase I/Phase II awards annually.

**Table 1. Sizing the DoD SBIR/STTR Program Annually**

FY	Count of Distinct DoD-Funded PI/PII SBIR/STTR Awards	Total DoD-Funded PI/PII SBIR/STTR Funding
2012	4973	\$1,090,143,968.02
2013	4901	\$988,818,482.23
2014	4796	\$1,082,209,915.19
2015	4832	\$1,040,778,157.84
2016	4971	\$1,105,200,418.39
2017	5190	\$1,260,999,327.89
2018	5251	\$1,240,980,063.70
2019	5796	\$1,691,062,982.31
2020	6046	\$1,905,575,032.16
2021	5990	\$1,711,005,800.94
<b>Total</b>	<b>52,746</b>	<b>\$13,116,774,148.67</b>



Next, we filtered the award data by Unique Entity Identifier (UEI) in each year to calculate the unique number of recipients of Phase I/Phase II awards annually.<sup>1</sup> We excluded entities with less than \$50,000 in DoD-funded Phase I/Phase II awards. Doing so produced a cleaner data set that eliminated potential administrative errors or otherwise unexplainable data—namely, companies associated with SBIR/STTR funding below the standard \$50,000 minimum award value.

We determined that the 52,746 DoD-funded Phase I/PII awards in our data set were distributed across 4,703 unique entities. These 4,703 companies (SBIR companies, DoD SBIR companies) became our analysis data. Table 2 breaks-down the number of unique companies that received DoD funded Phase I/II SBIR/STTR awards each year.

**Table 2. Unique SBIR Companies Annually**

<b>FY</b>	<b>Count of Distinct DoD-Funded PI/PII SBIR/STTR Awards</b>	<b>Count of Distinct Entities in Receipt of DoD-Funded PI/PII SBIR/STTR Awards</b>
2012	4973	1584
2013	4901	1627
2014	4796	1609
2015	4832	1648
2016	4971	1625
2017	5190	1695
2018	5251	1660
2019	5796	1999
2020	6046	2276
2021	5990	2190

### Multiple Award Winners

The count of distinct contract awards is significantly higher than the number of distinct SBIR companies, indicating that certain SBIR companies receive multiple awards in each year (“multiple award winners [MAWs]”). This finding aligns with earlier research we published, highlighting the fact that certain companies not only win multiple SBIR/STTR awards annually, but also participate in the program year over year. For instance, we determined that from FY 2010–FY 2019, 90% of DoD Phase I funds were awarded to existing DoD vendors. Over that same period, the top 5% of DoD SBIR companies with the most in DoD Phase I/Phase II awards received 51% of all DoD SBIR/STTR Phase I/Phase II funding (Bresler & Bresler, 2020). There is a major difference between a DoD SBIR company with decades of program experience and tens of millions of dollars or more in Phase I/Phase II funding, and a company new to the program with one or two awards. Given the share of SBIR/STTR funding awarded to MAWs, and given that they are well versed in navigating the government ecosystem, transition rates among MAWs should differ from less experienced SBIR companies. To evaluate this, our analyses considered transition rates not only SBIR-wide, but also among MAWs as a group.

<sup>1</sup> In April 2022, UEIs replaced DUNS numbers as the identifier provided by the System for Award Management (SAM)



## Defining Transition

Neither the SBIR/STTR program nor the DoD offer a standard taxonomy or set of metrics to define intragovernmental transition rate. Thus, to conduct this research we first needed to establish a working definition for transition rate along with a set of corresponding quantitative metrics. We define “transition” as a company developing a set of capabilities through Phases I and/or II of the SBIR program and subsequently delivering these capabilities to end-users in the DoD.

We focused on three metrics as a proxy for transition:

- 1) **Phase III awards attributed to a DoD-funded SBIR/STTR company.** The SBIR/STTR program is divided into three phases. Whereas Phases I and II provide funding for companies to conduct research and development (R&D), Phase III awards are contracts for work that “derives from, extends, or completes an effort made under prior SBIR funding agreements, but is funded by sources other than the SBIR Program” (Boyer, 2017). Phase III awards are identified in FPDS, making them the most explicit indicator that a company’s SBIR/STTR innovations were subsequently procured by the DoD.
- 2) **Non-Phase III DoD-funded procurement awards attributed to a DoD-funded SBIR/STTR company.** Some procurement contracts awarded to SBIR/STTR companies should be marked as Phase III in the data but are not. While there is no way to know if a subsequent contract related to a company’s SBIR/STTR research if it was not coded as a Phase III, for the purposes of our analyses we wanted to consider the possibility that the SBIR/STTR program delivers capabilities to the warfighter via non-Phase III contracts. For each company, we considered non-SBIR DoD procurement awarded subsequent to the first DoD-funded SBIR/STTR Phase I/Phase II award during our analysis period.
- 3) **DoD-funded subcontract awards attributed to a DoD-funded SBIR/STTR company.** Given the complexity and costs associated with pursuing government contracts, some SBIR/STTR companies have neither the ability nor the desire to contract with the DoD directly. Instead, they may deliver their capabilities to the DoD by subcontracting to a prime contractor. While there is no way to know if a subsequent DoD-funded subcontract award related to a company’s SBIR/STTR research, we wanted to consider the possibility that some SBIR companies transition their capabilities to the warfighter through a prime. As such, we considered DoD-funded subcontract awards attributed to each SBIR company, subsequent to the first DoD-funded SBIR/STTR Phase I/Phase II award they received during our analysis period.

## Research Limitations and Future Research

It is possible that a SBIR/STTR company was wholly or partially acquired by a prime contractor, and that the prime contractor then integrated the SBIR company’s capabilities into a DoD contract that it held. We did not have access to a reliable set of commercial acquisitions data, so we were unable to consider this metric. To the extent this information can be compiled in the future, it would be valuable to incorporate into subsequent research.

Our most significant research limitation was the fact that we could not distinguish between which non-SBIR DoD procurement contracts and DoD subcontracts related to a SBIR company’s SBIR/STTR work, and which did not. Because we treated all subsequent non-SBIR DoD procurement contracts and DoD subcontracts as indicators that SBIR/STTR capabilities transitioned to the warfighter, we gave the program more than its due credit. Future research



could leverage advanced text analysis to compare a company’s SBIR/STTR project description with text describing a subsequent contract/subcontract award, to evaluate the possibility that the two are related. Subsequent contracts/subcontracts that appear unrelated to SBIR/STTR work could be excluded. However, the most effective way to reduce false attributions would be to require more comprehensive reporting for Phase III contracts and thereby eliminate the need to analyze non–Phase III procurement contracts entirely. Similarly, the government should establish formal criteria for “Phase III subcontract awards,” create a code in USASpending to denote Phase III subcontract awards, and require that they be reported by relevant stakeholders from government and industry. Doing so would make it easier to track when SBIR/STTR capabilities transition to the warfighter indirectly. In light of these limitations, conclusions drawn from this research should place a greater emphasis on coded Phase III transition metrics because of their significantly higher efficacy.

### Calculating Transition by Metric

To analyze transition rate across our three metrics, we leveraged procurement data from FPDS and subcontracting data from USASpending. First, we aggregated procurement data from FPDS and filtered it for FY 2012–FY 2021. Next, we isolated DoD Phase III awards attributed to the 4703 SBIR companies in our analysis group. We repeated this process for non-SBIR DoD procurement contracts and subsequent DoD-funded SBIR/STTR Phase I/Phase II awards during the analysis period.

To identify DoD-funded subcontracts awarded to the SBIR companies, we aggregated subcontract award data from USASpending. We filtered the data for DoD-funded subcontracts awarded to the 4703 SBIR companies in our analysis group from FY 2012–FY 2021. We then isolated DoD-funded subcontracts subsequent to their first SBIR/STTR Phase I/Phase II award.

Table 3 provides a breakdown of the total funding amount and number of SBIR companies that transitioned capabilities to the warfighter, based on three increasingly broad ways of measuring transition:

- 1) Companies that transitioned via Phase III award(s)
- 2) Companies that transitioned via Phase III award(s) and/or subsequent non-SBIR DoD Procurement contracts
- 3) Companies that transitioned via Phase III award(s) and/or subsequent non-SBIR DoD Procurement contracts, and/or subsequent DoD-funded subcontracts

**Table 3. DoD SBIR/STTR Program Transition Rate by Metric**

Total SBIR Companies	Total DoD PIII Funding to SBIR Companies	SBIR Companies w/ PIII Award(s)	% SBIR Companies w/ PIII Award(s)	Total DoD PIII + Non-SBIR DoD Procurement to SBIR Companies	SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement	% SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement	Total DoD PIII + Non-SBIR DoD Procurement + DoD Subcontracts to SBIR Companies	SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement and/or DoD Subcontracts	% SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement and/or DoD Subcontracts
4703	\$10,276,728,376	748	16%	\$60,004,772,641	2731	58%	\$118,726,886,820	2949	63%

Over the last decade, only 16% of DoD SBIR companies won Phase III awards. The transition rate noticeably improved when considering non-SBIR procurement and subcontracts, but as discussed previously it is difficult to draw conclusions about the nature of these awards.

### Assessing the Distribution of Transition Funding

Next, we were interested in assessing the distribution of transition funding across the SBIR companies. Specifically, for the SBIR companies that transitioned, we wanted to determine the extent to which they generated more in subsequent defense revenue across





these three metrics, relative to the amount of non-dilutive Phase I/Phase II funding they were awarded. For each SBIR company that transitioned, we compared the total amount of funding they received in DoD Phase I/Phase II awards against the total amount of revenue they generated across these three metrics:

- 1) Phase III awards
- 2) Phase III awards and/or subsequent non-SBIR DoD procurement contracts
- 3) Phase III awards and/or subsequent non-SBIR DoD procurement contracts, and/or subsequent DoD-funded subcontracts

### Ratio of Phase I/Phase II Funding to Phase III Awards

As shown in Table 4, of the 748 SBIR companies that transitioned via Phase III awards, only 39% generated more in Phase III contract dollars than they were awarded in Phase I/Phase II non-dilutive funding. Taken as a percentage of the overall DoD SBIR program, just 6% of all SBIR companies generated more in Phase III contracts than they were awarded in Phase I/Phase II funding.

Table 4. Phase III Funding vs. Phase I/Phase II Funding

DoD SBIR Companies	SBIR Companies w/ PIII Award(s)	SBIR Companies w/ more in PIII than DoD PI/PII Funding	% PIII Companies with More in PIII Funding than PI/PII Funding	% All SBIR Companies w/ More in PIII Funding than PI/PII Funding
4703	748	293	39%	6%

### Ratio of Phase I/Phase II Funding to (Phase III Awards + non-SBIR Procurement)

As shown in Table 5, nearly half of companies that transitioned via Phase III and/or non-SBIR procurement contracts consumed more in Phase I/Phase II funding than they generated in transition revenue. Taken as a percentage of the overall DoD SBIR program, just 29% of all SBIR companies generated more in Phase III funding and/or non-SBIR procurement contracts than they were awarded in Phase I/Phase II funding.

Table 5. (Phase III + Non-SBIR Procurement Funding) vs. Phase I/Phase II Funding

DoD SBIR Companies	SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement	SBIR Companies w/ more in PIII and/or non-SBIR procurement than DoD PI/PII Funding	% PIII + Procurement Companies with More in PIII and/or Procurement Funding than PI/PII Funding	% All SBIR Companies w/ More in PIII and/or Procurement Funding than PI/PII Funding
4703	2731	1382	51%	29%

### Ratio of Phase I/Phase II Funding to (Phase III Awards + non-SBIR Procurement + DoD Subcontract Awards)

Adding DoD-funded subcontract awards to the calculation, 42% of companies that transitioned via one or more transition metric consumed more in Phase I/Phase II funding than they generated in subsequent transition revenue, as shown in Table 6. Taken as a percentage of the overall DoD SBIR program, just 36% of all SBIR companies generated more in Phase III awards and/or non-SBIR procurement contracts, and/or DoD-funded subcontracts than they were awarded in Phase I/Phase II funding.



**Table 6. Phase III + Non-SBIR Procurement Funding + DoD Subcontract Funding vs. Phase I/Phase II Funding**

DoD SBIR Companies	SBIR Companies w/ PIII Award(s) and/or non-SBIR DoD Procurement and/or DoD Subcontracts	SBIR Companies w/ more in PIII and/or non-SBIR procurement and/or DoD Subcontracts than DoD PI/PII Funding	% PIII + Procurement Companies with More in PIII and/or Procurement Funding and/or DoD Subcontracts than PI/PII Funding	% All SBIR Companies w/ More in PIII and/or Procurement and/or DoD Subcontracts than PI/PII Funding
4703	2949	1705	58%	36%

### Grading Transition Rate: The Jury is Out

Our analysis revealed that a substantial portion of DoD SBIR companies failed to transition their capabilities to the warfighter by any metric, and nearly all that transitioned still consumed more in Phase I/Phase II funding than what they generated in subsequent non-SBIR defense revenue. However, we could not draw conclusions about the success or failure of the SBIR program based on these metrics alone. Investing in early stage R&D means, to some extent, investing in ideas that will fail. If all Phase I/Phase II projects produced usable capabilities, it would signal that the DoD SBIR program was too risk averse. One could also argue that it is worth investing billions into companies that failed to transition if that investment also produced even a small number of capabilities that truly transformed the warfighter.

Additionally, these metrics alone offered no insight into specific factors inhibiting transition rate. Lawmakers and DoD officials often use the term “valley of death” to “[refer] to how many defense technologies fail to transition from prototypes into actual products for the military,” citing “the Pentagon’s bureaucracy”—the complexity of pursuing and winning DoD contracts—as its primary cause (Luckenbaugh, n.d.). However, our data shows that a subset of DoD SBIR companies won tens of millions of dollars or more in Phase I/Phase II awards annually. If the valley of death is caused primarily by companies lacking resources or expertise, there should be noticeable differences between the transition rates among these MAWs, relative to DoD SBIR companies with less experience. To draw more insightful conclusions about the DoD SBIR/STTR program as a means of delivering capabilities to the warfighter and to begin to understand why certain participants may fail to transition, we coupled our analysis of transition rates across program participants in general with an analysis of transition rates among MAWS specifically.

### Assessing the Top SBIR Companies

Our data set includes hundreds of MAWs. For example, the top 5% of DoD SBIR companies in our analysis group with the most in Phase I/Phase II awards—about 235 companies—collectively received 49% of all Phase I/Phase II funding. However, to meaningfully analyze the features and transition rates of MAWs at an individual company level, we focused on a smaller data set. Specifically, we isolated the 25 DoD SBIR companies in receipt of the most Phase I/Phase II funding during our analysis period. As shown in Table 7, the top 25 SBIR companies cumulatively received 18% of all DoD Phase I/Phase II funding—more than \$2.3 billion—from FY 2012–FY 2021.



**Table 7. Top 25 DoD SBIR Companies' Phase I/Phase II Funding Totals, FY 2012–FY 2021**

Company	Total DoD PI/PII Funding, FY 2012–FY 2021	% of Total DoD PI/PII Funding, FY 2012–FY 2021
PHYSICAL OPTICS CORPORATION	\$198,222,973	1.51%
INTELLIGENT AUTOMATION INC	\$172,174,305	1.31%
PHYSICAL SCIENCES INC	\$168,520,875	1.28%
CREARE INCORPORATED	\$158,034,669	1.20%
CHARLES RIVER ANALYTICS INC	\$153,639,314	1.17%
TRITON SYSTEMS INC	\$121,816,610	0.93%
LUNA INNOVATIONS INCORPORATED	\$115,727,487	0.88%
CFD RESEARCH CORPORATION	\$103,029,444	0.79%
LYNNTECH INC	\$95,715,220	0.73%
TOYON RESEARCH CORPORATION	\$92,398,212	0.70%
ARETE ASSOCIATES	\$86,856,904	0.66%
PROGENY SYSTEMS CORPORATION	\$76,422,839	0.58%
SA PHOTONICS INC	\$75,002,150	0.57%
MAINSTREAM ENGINEERING CORPORATION	\$70,653,705	0.54%
APTIMA INC	\$70,561,859	0.54%
CORVID TECHNOLOGIES LLC	\$64,965,146	0.50%
SOAR TECHNOLOGY INC	\$67,302,292	0.51%
CORNERSTONE RESEARCH GROUP INCORPORATED	\$59,984,693	0.46%
ENGINEERING AND SOFTWARE SYSTEM SOLUTIONS INC	\$57,145,087	0.44%
TDA RESEARCH INC	\$56,439,024	0.43%
INTELLISENSE SYSTEMS INC	\$55,685,545	0.42%
MAXENTRIC TECHNOLOGIES LLC	\$55,054,742	0.42%
OCEANIT LABORATORIES INC	\$54,091,626	0.41%
FIRST RF CORPORATION	\$53,791,669	0.41%
SYSTEMS TECHNOLOGY RESEARCH LLC	\$52,631,563	0.40%
<b>Total</b>	<b>\$2,335,867,952</b>	<b>18%</b>

To capture a more complete picture of the Phase I/Phase II funding attributed to MAWs, for each of these top 25 companies we expanded the analysis time frame to calculate their total DoD Phase I/Phase II funding, from their first DoD Phase I/Phase II award through the date we ran the analysis (March 29, 2023). Table 8 shows the total amount of DoD Phase I/Phase II funding each company received over its lifetime.



**Table 8 Lifetime DoD Phase I/Phase II Funding—Top 25 DoD SBIR Companies**

Company	FY of Initial DoD PhI/PhII Award	FY of Most Recent DoD PhI/PhII Award	Lifetime Total DoD PI/PII Award Funding
PHYSICAL OPTICS CORPORATION	1997	2023	\$359,325,897
PHYSICAL SCIENCES INC	1997	2023	\$321,023,208
CREARE INCORPORATED	1997	2023	\$274,156,442
INTELLIGENT AUTOMATION INC	1997	2023	\$269,444,012
CHARLES RIVER ANALYTICS INC	1997	2023	\$260,141,162
TRITON SYSTEMS INC	1997	2023	\$243,888,188
CFD RESEARCH CORPORATION	1997	2023	\$213,364,011
LUNA INNOVATIONS INCORPORATED	1997	2023	\$199,301,561
LYNNTECH INC	1997	2023	\$158,497,089
TOYON RESEARCH CORPORATION	1997	2023	\$153,759,374
APTIMA INC	1997	2023	\$152,596,850
ARETE ASSOCIATES	1997	2023	\$139,482,615
PROGENY SYSTEMS CORPORATION	1997	2023	\$133,489,054
TDA RESEARCH INC	1997	2023	\$106,391,125
CORNERSTONE RESEARCH GROUP INCORPORATED	1998	2023	\$105,438,088
MAINSTREAM ENGINEERING CORPORATION	1997	2023	\$102,005,756
SOAR TECHNOLOGY INC	2000	2023	\$101,166,814
SA PHOTONICS INC	2003	2023	\$98,359,670
INTELLISENSE SYSTEMS INC	2018	2023	\$84,704,547
FIRST RF CORPORATION	2003	2023	\$84,536,933
CORVID TECHNOLOGIES LLC	2005	2023	\$80,279,823
OCEANIT LABORATORIES INC	1997	2023	\$76,722,560
ENGINEERING AND SOFTWARE SYSTEM SOLUTIONS INC	2007	2023	\$75,206,735
MAXENTRIC TECHNOLOGIES LLC	2005	2023	\$71,623,153
SYSTEMS TECHNOLOGY RESEARCH LLC	2011	2022	\$53,419,184
<b>TOTAL</b>			<b>\$3,558,997,955</b>

All but one of the top 25 companies have received DoD Phase I/Phase II SBIR/STTR awards for more than 10 years, and 20 of the top 25 companies have been awarded DoD Phase I/Phase II funding for more than 20 years.

### Transition Rate Among MAWs

For each of the top 25 DoD SBIR companies, we calculated the total amount of Phase III, non-SBIR DoD Procurement, and DoD subcontract revenue generated between FY 2012–FY 2021. We then compared each metric to the company’s total Phase I/Phase II funding during



the analysis period to generate a ratio of transition revenue to total Phase I/Phase II funding. As shown in Table 9, only four of the top 25 DoD SBIR companies generated more in DoD Phase III contracts than they received in non-dilutive Phase I/Phase II awards.

Adding non-SBIR DoD procurement to the transition calculation, the majority of the top 25 DoD SBIR companies still received more in Phase I/Phase II funding than they generated in subsequent Phase III and/or non-SBIR DoD contracts. By the most liberal transition metric—subsequent DoD Phase III funding, and/or non-SBIR DoD procurement, and/or DoD-funded subcontract awards—just over half of the top 25 DoD SBIR companies generated more in transition revenue than they were awarded in Phase I/Phase IIs.

**Table 9. Transition Metrics for FY 2012–FY 2021, Top 25 DoD SBIR Companies**

Company	Total DoD PI/PII \$	Total PIII \$	Ratio PIII \$ vs. PI/PII \$	Total PIII + non-SBIR \$	Ratio PIII + non-SBIR \$ vs. PI/PII \$	Total PIII + non-SBIR + DoD Subcontract \$	Ratio PIII + non-SBIR + DoD Subcontract \$ vs. PI/PII \$
PHYSICAL OPTICS CORPORATION	\$198,222,973	\$296,550,639	150%	\$506,752,621	256%	\$543,835,766	274%
INTELLIGENT AUTOMATION INC	\$172,174,305	\$14,607,362	8%	\$68,236,490	40%	\$86,709,123	50%
PHYSICAL SCIENCES INC	\$168,520,875	\$10,303,411	6%	\$74,941,384	44%	\$101,913,061	60%
CREARE INCORPORATED	\$158,034,669	\$53,366,123	34%	\$85,743,425	54%	\$88,505,471	56%
CHARLES RIVER ANALYTICS INC	\$153,639,314	\$15,930,109	10%	\$206,213,710	134%	\$241,430,984	157%
TRITON SYSTEMS INC	\$121,816,610	\$6,430,752	5%	\$35,544,912	29%	\$36,091,069	30%
LUNA INNOVATIONS INCORPORATED	\$115,727,487	\$3,616,872	3%	\$32,884,666	28%	\$36,422,619	31%
CFD RESEARCH CORPORATION	\$103,029,444	\$450,378	0%	\$21,122,072	21%	\$53,267,339	52%
LYNNTECH INC	\$95,715,220	\$3,849,136	4%	\$20,586,029	22%	\$20,742,065	22%
TOYON RESEARCH CORPORATION	\$92,398,212	\$19,174,422	21%	\$129,289,686	140%	\$228,169,816	247%
ARETE ASSOCIATES	\$86,856,904	\$125,140,457	144%	\$179,414,186	207%	\$231,727,064	267%
PROGENY SYSTEMS CORPORATION	\$76,422,839	\$875,436,015	1146%	\$1,326,867,356	1736%	\$2,068,581,929	2707%
SA PHOTONICS INC	\$75,002,150	\$11,267,031	15%	\$82,407,497	110%	\$205,665,144	274%
MAINSTREAM ENGINEERING CORPORATION	\$70,653,705	\$143,565	0%	\$26,159,461	37%	\$51,320,790	73%
APTIMA INC	\$70,561,859	\$82,468,290	117%	\$193,482,868	274%	\$276,564,268	392%
CORVID TECHNOLOGIES LLC	\$64,965,146	\$26,602,284	41%	\$112,915,222	174%	\$201,785,024	311%
SOAR TECHNOLOGY INC	\$67,302,292	\$5,760,555	9%	\$104,177,240	155%	\$213,942,061	318%
CORNERSTONE RESEARCH GROUP INCORPORATED	\$59,984,693	\$4,820,260	8%	\$20,992,906	35%	\$27,303,828	46%
ENGINEERING AND SOFTWARE SYSTEM SOLUTIONS INC	\$57,145,087	\$66,924,136	117%	\$177,492,020	311%	\$178,879,990	313%
TDA RESEARCH INC	\$56,439,024	\$610,100	1%	\$17,383,352	31%	\$18,439,670	33%
INTELLISENSE SYSTEMS INC	\$55,685,545	\$15,624,644	28%	\$31,418,599	56%	\$58,408,779	105%
MAXENTRIC TECHNOLOGIES LLC	\$55,054,742	\$6,290,024	11%	\$22,033,549	40%	\$27,717,398	50%
OCEANIT LABORATORIES INC	\$54,091,626	\$22,630,526	42%	\$52,124,554	96%	\$53,565,949	99%
FIRST RF CORPORATION	\$53,791,669	\$33,006,900	61%	\$70,982,752	132%	\$468,983,023	872%
SYSTEMS TECHNOLOGY RESEARCH LLC	\$52,631,563	\$49,937,790	95%	\$594,811,635	1130%	\$677,348,738	1287%

We were interested in seeing how these top 25 companies ranked in terms of the amount of Phase III contract dollars they received, compared to the other companies in our



analysis group that received Phase IIIs. We ranked the 748 companies from our analysis group that received Phase III awards, where “1” denoted the company with the most in Phase III funding and “748” denoted the company with the least in Phase III funding. Table 10 shows where each of the top 25 DoD SBIR companies ranked. Only nine of the top 25 companies fell in the top 10% of DoD SBIR companies receiving the most Phase III contract dollars.

**Table 10. Ranking of Top 25 SBIR Companies, Based on Phase III Funding Amount**

Company	Company Ranking, Based on Total DoD Phase III Funding
PHYSICAL OPTICS CORPORATION	6
INTELLIGENT AUTOMATION INC	112
PHYSICAL SCIENCES INC	147
CREARE INCORPORATED	40
CHARLES RIVER ANALYTICS INC	104
TRITON SYSTEMS INC	207
LUNA INNOVATIONS INCORPORATED	273
CFD RESEARCH CORPORATION	585
LYNNTECH INC	266
TOYON RESEARCH CORPORATION	91
ARETE ASSOCIATES	18
PROGENY SYSTEMS CORPORATION	1
SA PHOTONICS INC	136
MAINSTREAM ENGINEERING CORPORATION	653
APTIMA INC	27
CORVID TECHNOLOGIES LLC	68
SOAR TECHNOLOGY INC	223
CORNERSTONE RESEARCH GROUP INCORPORATED	247
ENGINEERING AND SOFTWARE SYSTEM SOLUTIONS INC	30
TDA RESEARCH INC	555
INTELISENSE SYSTEMS INC	106
MAXENTRIC TECHNOLOGIES LLC	210
OCEANIT LABORATORIES INC	80
FIRST RF CORPORATION	56
SYSTEMS TECHNOLOGY RESEARCH LLC	42

The data revealed no consistent relationship between the amount of Phase I/Phase II funding a company received and the extent to which it delivered capabilities to the warfighter. In fact, some MAWs continued to receive a disproportionate share of overall DoD Phase I/Phase II funding, yet had below average rates of transition. Their inability to transition cannot be attributed to a lack of resources or wherewithal—after all, they have decades of experience in the defense market and tens of millions in non-dilutive contract awards. Instead, the inconsistent and often poor transition rates among MAWs revealed a disconnect between both the stated



objectives of the program and the role the program should play, in light of today's threat environment; and how the program functions in actuality.

The DoD SBIR program awards a disproportionate share of Phase I/Phase II funding to a set of companies that, based on extensive past performance data, are unlikely to deliver capabilities to defense end-users. That the most active DoD SBIR companies are not necessarily those with the greatest potential for transition indicates that they are selected for Phase I/Phase II awards based on other, unrelated criteria. As such, "the valley of death" is not simply the result of companies struggling to navigate the bureaucracy associated with transitioning from R&D into a DoD program of record. By continuing to disproportionately fund companies that, based on their extensive past performance, will not transition, the DoD SBIR program effectively guarantees the existence of a "valley of death."

### **Small By What Standards?**

The data related to MAWs brought to light another fundamental issue related to the SBIR program. While the SBIR/STTR program was established to serve small businesses, companies can win tens of millions of dollars or more annually in non-dilutive R&D grants and still qualify by program standards as small. In fact, Phase I/Phase II awards represent only a snapshot of MAWs' overall revenue—many generate tens of millions of dollars or more in government revenue annually from other sources, as demonstrated in Table 9; in addition to commercial revenue. Some, like Luna Innovations, are publicly-traded.

Companies can qualify as "small" by SBIR/STTR size standards irrespective of how much revenue they generate, as long as they have fewer than 500 employees (*DOD Small Business Innovation Research / Small Business Technology Transfer, n.d.*). A significant share of Phase I/Phase II funds are not simply awarded to companies unlikely to transition their capabilities to the warfighter; they are awarded to companies that most reasonable Americans would never consider to be "small businesses."

Additionally, MAWs win Phase I/Phase II awards for projects that span a wide range of unrelated topics. We searched a subset of the top 25 companies by name on the SBIR Award Database website, <https://www.sbir.gov/sbirsearch/award/all>, to better understand the nature of some of their DoD Phase I/Phase II awards. We found that Charles River Analytics received Phase I/Phase II funding for projects including, but not limited to, data analytics for ship maintenance, decision support systems to assist Army soldiers with career planning, wearable sensors for Navy divers, algorithms to enhance robotic caregivers, the development of "smart fabrics" that incorporate sensors and communication networks, and more. Physical Optics received Phase I/Phase II funding to develop artificial intelligence for unmanned systems, coatings for missiles, cyber detection and attack tools, remote unmanned refueling systems, night vision cameras and more. Progeny won Phase I/Phase II awards to develop cyber security for unmanned aerial systems, self-serve kiosks to display human performance information, platforms to manage food service on Navy ships, augmented reality displays for submarine command teams, and more.

Furthermore, from our earlier research we know that most MAWs not only win DoD Phase I/Phase II awards, but also participate in the SBIR/STTR program across multiple non-defense agencies. To capture a picture of their experience in other agencies' SBIR programs, we linked all Phase I/Phase II SBIR/STTR award data associated with each of the top 25 companies from FPDS and USASpending, irrespective of funding agency. As shown in Table 11, all but one of the top 25 DoD SBIR/STTR companies generated Phase I/Phase II funding from non-DoD sources.



**Table 11. Top 25 DoD SBIR/STTR Companies' Lifetime Phase I/Phase II Funding, DoD and non-DoD Sources**

Company	Lifetime Total DoD PI/PII Funding	Lifetime Total PI/PII SBIR/STTR Funding
PHYSICAL OPTICS CORPORATION	\$359,325,897	\$384,534,627
PHYSICAL SCIENCES INC	\$321,023,208	\$355,985,614
CREARE INCORPORATED	\$274,156,442	\$330,887,539
INTELLIGENT AUTOMATION INC	\$269,444,012	\$313,815,023
CHARLES RIVER ANALYTICS INC	\$260,141,162	\$281,737,900
TRITON SYSTEMS INC	\$243,888,188	\$249,656,762
CFD RESEARCH CORPORATION	\$213,364,011	\$240,851,455
LUNA INNOVATIONS INCORPORATED	\$199,301,561	\$238,795,534
LYNNTECH INC	\$158,497,089	\$176,441,321
TOYON RESEARCH CORPORATION	\$153,759,374	\$165,561,850
APTIMA INC	\$152,596,850	\$156,214,311
ARETE ASSOCIATES	\$139,482,615	\$141,259,857
PROGENY SYSTEMS CORPORATION	\$133,489,054	\$136,432,764
TDA RESEARCH INC	\$106,391,125	\$129,951,953
CORNERSTONE RESEARCH GROUP INCORPORATED	\$105,438,088	\$124,861,304
MAINSTREAM ENGINEERING CORPORATION	\$102,005,756	\$113,875,803
SOAR TECHNOLOGY INC	\$101,166,814	\$103,579,056
SA PHOTONICS INC	\$98,359,670	\$99,259,498
INTELLISENSE SYSTEMS INC	\$84,704,547	\$88,161,845
FIRST RF CORPORATION	\$84,536,933	\$85,129,445
CORVID TECHNOLOGIES LLC	\$80,279,823	\$80,653,711
OCEANIT LABORATORIES INC	\$76,722,560	\$78,712,745
ENGINEERING AND SOFTWARE SYSTEM SOLUTIONS INC	\$75,206,735	\$76,722,560
MAXENTRIC TECHNOLOGIES LLC	\$71,623,153	\$73,821,632
SYSTEMS TECHNOLOGY RESEARCH LLC	\$53,419,184	\$53,419,184

It is hard to imagine how any company, let alone a small business, can be at the cutting-edge of innovation in dozens of unrelated fields. Rather, these companies are experts in navigating the SBIR program. Despite the stated objectives of the program, DoD SBIR program managers are primarily measured by whether or not they award the requisite amount of total funding to eligible firms every year; and whether or not these recipient firms comply with program rules over the course of their projects. Based on this criterion, companies with expertise submitting SBIR proposals, rather than companies with the best ideas, are the likely recipients of Phase I/Phase II funding. The sheer amount of SBIR/STTR funding attributed to MAWs across the entirety of the program further underscores that poor transition rates cannot be attributed exclusively to a lack of resources. Simply allocating more money to SBIR companies does not address the “valley of death.” SBIR program managers must begin to evaluate a company’s potential for transition as the primary criterion for award.





For decades, MAWs have comfortably won tens of millions of dollars or more in non-dilutive R&D funding, year in and year out. In spite of the stated objectives of the program and that now more than ever it is critical for the military to harness innovations stemming from the private sector, neither the DoD SBIR program managers nor the participating companies are held accountable for ensuring these investments benefit the warfighter. When making award decisions in relation to MAWs, SBIR program managers must be required to factor the ratio of previous Phase I/Phase IIs awarded to a company, compared to the subsequent Phase III/Phase III subcontracts generated. Additionally, Congress must establish clear Phase III transition requirements for DoD SBIR program offices—specifically, a formal goal for the minimum number of companies awarded Phase III contracts and/or Phase III subcontracts annually. Doing so will direct more SBIR resources to non-MAWs, and/or will force the most active participants in the DoD SBIR program to focus on delivering capabilities to DoD end-users.

### **Transition Challenges for Smaller SBIR Companies**

Clearly, large-scale improvements to the transition rate among DoD SBIR companies will require creating new incentives, changing the eligibility criteria for participants, and changing the metrics for evaluating DoD program offices. That said, we also wanted to consider the unique challenges smaller DoD SBIR companies face when navigating the defense market. Unlike MAWs, smaller companies with less experience in the DoD market often pursue the SBIR/STTR program with the expectation that, if they perform well, it will lead to follow-on defense business. However, the DoD SBIR program rarely positions them for success in the broader defense market for a variety of reasons.

We have interacted with and surveyed dozens of DoD SBIR companies and DoD SBIR program offices over the last five years, both in conjunction with earlier research papers published through the Naval Postgraduate School and as part of work we have undertaken—with Phase I/Phase II funding from the Navy, the Air Force, and the Defense Technical Information Center—to develop solutions to improve defense stakeholders’ ability to leverage capabilities funded and fielded through innovation programs.

Through this qualitative research, we identified several specific factors keeping DoD SBIR companies from serving the needs of the warfighter subsequent to program completion (Bresler & Bresler, 2021):

- SBIR companies are not educated on how or where to identify DoD opportunities, and they are unlikely to succeed if and when they attempt to bid on them.
  - The design and archaic search functionality of the website where DoD solicitations are marketed (SAM.gov) make it near impossible for companies to find relevant opportunities.
  - If a company manages to identify a relevant opportunity, the submission deadline makes it nearly impossible to prepare and submit a bid. Our analysis of more than 1 million DoD solicitations from 2002 through 2020 found that 70% required companies to respond within 21 days of when they were posted, and 30% required responses within 10 days or less.
  - DoD solicitations are not written clearly. Evaluating the readability of the description fields associated with more than 1 million DoD solicitations using the Flesch-Kincaid Readability and Grade Level scores, we found that fewer than 3% of solicitation descriptions were written in “plain English.”



- Government stakeholders do not coordinate their requirements, despite often seeking the same capabilities. For instance, on a single day in October 2020, there were 132 open requirements related to UAVs. Small companies new to the defense market cannot reasonably respond to dozens or hundreds of potentially relevant opportunities, and they lack the insider knowledge to effectively prioritize them.
- The individual that oversees Phase I/Phase II contract work typically does not have the authority and/or resources to fund a follow-on contract/program of record directly. And he/she may not have knowledge of or access to those who do. As a result, in the absence of investing in lobbyists or business development consultants, companies have no way of connecting with their potential DoD customers—regardless of their Phase I/Phase II performance.
- The DoD SBIR program offers no clear instructions to companies regarding internal systems (accounting, cybersecurity, etc.) that may be required to qualify for follow-on contracts. Small companies not only walk away from the defense market because they cannot justify the investment, but also because they simply cannot get clear information on what the required level of investment will be.
- The DoD SBIR program does not effectively market participants' capabilities to the broader armed services community. One of the most frequent comments from DoD stakeholders over the last five years was that they received very few briefings on the projects funded by their own branch, and almost never received information on capabilities funded by other branches. As a result, rather than leveraging existing investments made through the DoD SBIR program, DoD stakeholders either continuously invest in redundant market research or fail to modernize altogether.

## Conclusion and Recommendations

On the whole, the DoD SBIR program has failed to incubate capabilities that go on to serve DoD end-users. This poor rate of transition can be attributed to multiple factors. SBIR program managers are not held accountable for funding companies with the greatest promise for transition. Instead, they have directed the majority of Phase I/Phase II funds to companies that have cultivated an expertise in how to navigate the SBIR program. Regardless of SBA size standards, these MAWs look and act like large businesses. They effectively submit winning proposals and deliver compliant milestones. Their institutional knowledge of processes is more relevant than the innovativeness of their solutions. Because these companies can win tens of millions of dollars annually in non-dilutive funding, they have no incentive to transition. In fact, they are incentivized to continue to focus their resources and attention on pursuing *more* SBIR awards, rather than undertake the complex process of pursuing non-SBIR contracts.

Additionally, companies that participate in the DoD SBIR program with the goal of continuing to support the DoD thereafter are not well-positioned to do so. The SBIR program fails to educate them on the requirements associated with pursuing traditional defense contracts. While the SBIR program affords participants with sole-source justification within scope, it does not facilitate connections between SBIR companies and viable DoD customers. To successfully capture defense business after Phase I/Phase II project completion requires a significant investment. Many small, innovative companies—particularly those with viable commercial revenue streams—choose to abandon the defense market altogether, rather than “pay to play.”

To address these issues and position the SBIR program to more effectively deliver capabilities to the warfighter, we offer the following recommendations:



- Require more comprehensive reporting for Phase III contracts to eliminate the need to analyze non-Phase III procurement contracts when measuring transition.
- Create a code in USASpending specifically for Phase III subcontract awards, to denote when a subcontract award relates to a company's SBIR work.
- Overhaul SBA size standards so that the SBIR program benefits *truly* small businesses.
- Establish clear Phase III transition requirements for DoD SBIR/STTR program offices. Specifically, there should be a formal goal for the minimum number of companies awarded Phase III contracts and/or Phase III subcontracts annually.
- When a company submits a Phase I/Phase II SBIR proposal, the ratio of its total Phase I/Phase II funding relative to the amount of revenue it has generated in Phase III contract and subcontract awards should be an important evaluation criteria. Firms with \$10 million or more in cumulative Phase I/Phase II DoD SBIR awards must meet minimum Phase III transition thresholds in order to remain eligible for additional Phase I/Phase II funding.
- Incentivize DoD stakeholders to integrate capabilities funded and fielded through the SBIR/STTR program. The incentive program can mirror existing set-aside programs that require DoD stakeholders to award a certain percentage of contract awards to various company types (woman-owned small business, 8a, etc.). There should be goals for awarding a percentage of contracts annually as Phase IIIs or Phase III subcontract awards, to encourage the DoD to leverage investments made through the SBIR/STTR program. Additional credit should be given when a DoD stakeholder awards a Phase III contract or subcontract to a company funded and fielded by a different agency.
- Incentivize prime contractors to integrate capabilities funded and fielded through the SBIR/STTR program. Much like prime contractors have goals for awarding a certain share of subcontracting business to various set-aside companies, they should receive additional credit—beyond what would count towards their small business set-aside goals—when subcontracting for capabilities funded and fielded through SBIR/STTR.
- Make it easier for companies to identify and bid on DoD solicitations. Specifically, redesign SAM.beta to improve search functionality; require solicitations to have a response time of more than 30 days unless a justification is provided; require solicitation descriptions to be written in plain English; and require government stakeholders with similar requirements to coordinate their outreach and communication efforts.

It is essential for our national security that the DoD have access to the most promising new technologies. As the largest and most long-standing defense innovation initiative, the DoD SBIR program must adapt with this imperative in mind. With strong leadership and a thoughtful restructuring of resources and incentives, the DoD SBIR program has the potential to channel its multibillion-dollar budget into solutions that could revolutionize the military.

## References

- About.* (n.d.). About | SBIR.gov. Retrieved March 23, 2023, from <https://www.sbir.gov/about>
- Boyer, L. A. (2017). *SBIR/STTR phase III contracting—What you need to know.*
- Bresler, A., & Bresler, A. (2020). The effect of defense-sponsored innovation programs on the military's industrial base. *The Acquisition Research Program of the Graduate School of Defense Management at the Naval Postgraduate School.*



<https://calhoun.nps.edu/bitstream/handle/10945/64763/SYM-AM-20-059.pdf?sequence=1&isAllowed=y>

Bresler, A., & Bresler, A. (2021). Why marketing matters: Strengthening the defense supplier base through better communication with industry. *The Acquisition Research Program of the Graduate School of Defense Management at the Naval Postgraduate School*.

<https://dair.nps.edu/bitstream/123456789/4404/1/SYM-AM-21-097.pdf>

DoD. (2018). *Summary of the national defense strategy of the United States of America 2018*.

<https://www.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-StrategySummary.pdf>

DoD Small Business Innovation Research / Small Business Technology Transfer. (n.d.). Retrieved March 22, 2023, from <https://www.defensesbirsttr.mil/>

Luckenbaugh, J. (n.d.). //ITSEC NEWS: Valley of death isn't "that big of a problem," military official says. Retrieved March 30, 2023, from

<https://www.nationaldefensemagazine.org/articles/2022/12/1/valley-of-death-isnt-that-big-of-a-problem-official-says>



# Leverage AI to Learn, Optimize, and Wargame (LAILOW) for Strategic Laydown and Dispersal (SLD) of the Operating Forces of the U.S. Navy

**Dr. Ying Zhao**—is a Research Professor at the Naval Postgraduate School (NPS). Her research focuses on data sciences, machine learning, and artificial intelligence methods for defense military applications such as semantic and social networks, common tactical air pictures, combat identification, logistics, and mission planning. Since joining NPS, Zhao has been a principal investigator (PI) on many projects awarded for Department of Defense (DoD) research projects. She received her PhD in mathematics from MIT and is the co-founder of Quantum Intelligence, Inc. [yzhao@nps.edu]

**Dr. Doug MacKinnon**—is a Research Associate Professor at the Naval Postgraduate School (NPS). He teaches graduate Operations Research (OR) courses in probability and statistics, wargaming, and simulation. He is also heavily involved in leading data mining research projects that explore domains ranging from Navy acquisition to recruiting and has performed as principal investigator (PI) on many projects awarded for DoD research and has advised 100+ doctoral and master's research studies. He earned his PhD in civil engineering from Stanford University and his undergraduate degree from the U.S. Naval Academy. [djmackin@nps.edu]

## Abstract

The Secretary of the Navy disperses Navy forces in a deliberate manner to support Department of Defense (DoD) guidance, policy, and budget. The current strategic, laydown, and dispersal (SLD) process is labor intensive, time intensive, and less capable of becoming agile for considering competing alternative plans. SLD could benefit from the implementation of artificial intelligence. We introduced a relatively new methodology to address these questions which was recently derived from an earlier Office of Naval Research funded project that combined deep analytics of machine learning, optimization, and wargames. This methodology is entitled LAILOW which encompasses Leverage AI to Learn, Optimize, and Wargame (LAILOW). In this paper, we developed a stand-alone set of pseudo data that mimicked the actual, classified data so that experimental excursions could be performed safely. We show LAILOW produces a score from a wargame-like scenario for every available ship that might be moved. The score for each ship increases as fewer resources (e.g., lower cost) are required to fulfill an SLD plan requirement to move that ship to a new homeport. This produces a mathematical model that enables the immediate comparison between competing or alternate ship movement scenarios that might be chosen instead. We envision a more integrated, coherent, and large-scale deep analytics effort leveraging methods that link to existing real data sources to more easily enable the direct comparisons of potential scenarios of platform movement considered through the SLD process. The resulting product could facilitate decision makers' ability to learn, document, and track the reasons for complex decision making of each SLD process and identify potential improvements and efficiencies for force development and force generation.

**Keywords:** artificial intelligence, machine learning, optimization, strategic laydown, and dispersal, SLD, data mining

## Introduction

The Secretary of the Navy (SECNAV) disperses Navy forces in a deliberate manner to support Department of Defense (DoD) guidance, policy and budget. The current strategic, laydown, and dispersal (SLD) process is labor intensive and may be benefited by digitalization, automation and application of AI.

The laydown and dispersal of U.S. Naval forces requires manual manipulation of data via weekly Working Groups, which is manpower intensive, and only presents one option to the Chief of Naval Operations (CNO) and SECNAV for consideration. The current SLD process



takes one full year to develop and is not responsive to changes in the operating environment or strategic guidance. For example, there is no mechanism to leverage existing data resources to monitor plan execution and track progress toward completion. The 10 years of projected force laydown optimization problem can be overwhelming. The SLD plan needs more than just simple process revision—it needs a modernization with a holistic design leveraging digitalization, automation, and application of AI.

The objective is to digitalize the SLD process with more automation using a cloud-based SLD database, deep analytics, ML/AI to aid decision making, and reduce manpower requirements to focus on the strategic basis and integration of the SLD Plan for improved efficiency and better-informed decision making.

## Literature Review

More specifically, based on a memo from RDML T. R. Williams, former director for Plans, Policy, and Integration (N5) for the Deputy Chief of Naval Operations for Operations, Plans, and Strategy (N3/N5; Williams, 2021), N52 is teaming with industry and academia to modernize the SLD process; the challenges are described in the following phases.

### Descriptive Phase

What decisions were made? This phase is focused on developing a new database utilizing modern data analytics to display information in a shareable website. The current SLD database exists on a standalone computer with a single user's access in the Pentagon requiring manual updates. This phase's end state is a cloud-based SLD database accessible to the SLD working group that offers permission controls and features improved analysis and display functions.

### Predictive Phase

How are we making decisions? What happens if I make a different decision? This phase's end state is an Excursion Modeling Tool. The goal is to develop a decision support tool that uses existing authoritative data and models SLD excursions to assist in rapid decision making with increased accuracy.

### Prescriptive Phase

Are we making the right decisions? This phase's end state will utilize deep analytics including AI to take the SLD calculations and other inputs to evaluate the SLD plan and create an optimized plan by including global and theater posture and time-phased force and deployment data (TPFDD) into the calculations.

## Methods

This paper details the methods related to the research questions and the prescriptive phase. We apply a mathematical model (i.e., Leverage AI to Learn, Optimize, and Wargame [LAILOW] model) to address deep analytic aspects of the research. LAILOW was derived from an ONR-funded project that focuses on deep analytics of machine learning, optimization, and wargame, essentially Leveraging AI, and consists of the following steps:

**Learn:** D data, data mining, machine learning, and predictive algorithms are used to learn the patterns from historical data on what and how decisions were made. Data derived from competing demands refer to the excursion proposals and requirements from fleet commanders, national leaders, and assessment data done in various function areas in different installation locations. The current manual process focuses on balancing the budget of unit moving costs with the known demands. Moving cost is developed from permanent change of station (PCS)



orders of manpower and readiness of infrastructure. The data are in the form of structured databases and unstructured data such as PowerPoint slides and .pdf documents.

**Optimize:** Patterns from learning are represented as Soar reinforcement learning (Soar-RL) rules or AGI Transformer models used to optimize future SLD plans. An SLD plan includes a complete gain or loss of naval assets at each installation, homeport, home base, hub, and shore posture location (Fd) and staff (Fg). The optimization can be overwhelming considering numerous combinations. LAILOW instead uses integrated Soar-RL and coevolutionary algorithms to map a total SLD plan to individual units mentioned in an excursion proposal, assessment report, and other what-if analyses.

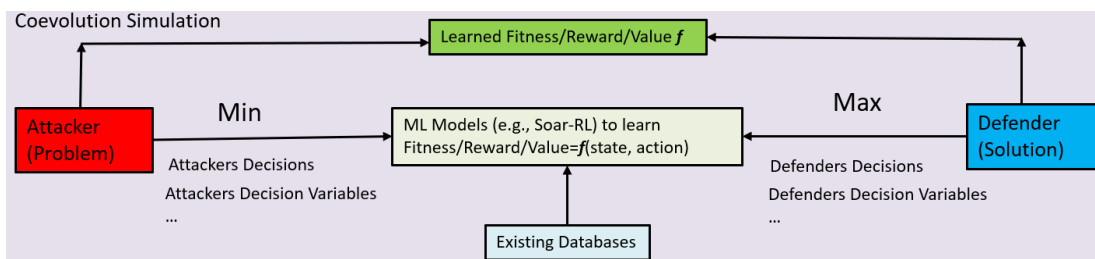
**Wargame:** There might be no or rare data for new warfighting requirements and capabilities. This motivates wargame simulations. An SLD plan can include state variables or problems (e.g., future global and theater posture, threat characteristics), which is only observed, sensed, and cannot be changed. Control variables are solutions (e.g., an SLD plan). LAILOW sets up a wargame between state and control variables. Problems and solutions coevolve based on evolutionary principles of selection, mutation, and crossover.

A LAILOW framework can be set up as a multi-segment wargame played by a self-player and the opponent, as shown in Figure 1. The self-player or defender is the SLD enterprise. The opponent or attacker is the environment including competing demands. When applying LAILOW, we first divide the processes into state variables and decision variables as follows:

**State variables:** These variables and data can be sensed, observed, and estimated, however, cannot be decided or changed by the self-player. They are the input variables, or problems that the self-player must consider. They are also called *tests* or *attacks* for the SLD enterprise.

**Decision variables:** These variables are needed to solve the problem using optimization algorithms. In LAILOW, the optimization of the decision variables is achieved by the integration of Soar-RL and coevolutionary search and optimization algorithms (Back, 1996; O’Reilly et al., 2020).

Both opponent (tests) and self-player (solutions) evolve and compete as in a wargame. LAILOW is like a Monte Carlo simulation but guided by ML/AI learned patterns with optimization algorithms. In the wargame, the opponent generates large-scale what-if tests to challenge the self-player to come up with better solutions, e.g., SLD configurations to answer the questions such as “what happens if I choose a different decision?” in a systematic simulation.



**Figure 1. LAILOW viewed in a Coevolutionary wargame simulation; ML algorithms (i.e., Soar-RL) are used to model the fitness or utility functions for both players.**

Each “learn, optimize, wargame” cycle dynamically iterates in each stage and across all the value areas with the analytic components and algorithms detailed as follows.



In a LAILOW framework, the “learn” component usually employs supervised ML algorithms such as classification, regression, and predictive algorithms. For example, one can apply a wide range of state-of-the-art supervised ML algorithms from the *scikit-learn* python such as logistic regression, decision trees, *naïve Bayes*, random forest, k-nearest neighbors, and neural networks. Deep learning or AGI Transformers can also be placed in this category where the input data is diversified. An AGI framework typically contains large-scale machine learning models (e.g., billions of parameters in the ChatGPT model; OpenAI, 2023) to learn and recognize patterns from multimodality data.

Supervised ML algorithms can be used to learn the state variables and assessment measures in the function areas for potential SLD and excursion plans such as speed, quality, and fitness of deployment and execution, balance of competing demands and constraints (e.g., avoidance of unacceptable reduction of capability), along with Fd and Fg measures.

In LAILOW, we use Soar-RL to learn two fitness functions separately for the self-player and opponent. In reinforcement learning, an agent takes an action and generates a new state, based on its current state and on the expected value it estimates from its internal model (Sutton & Barto, 2014). It also learns from the reward data from the environment by modifying its internal models. Soar-RL can scalably integrate a rule-based AI system with many other capabilities, including short- and long-term memory (Laird, 2012). Soar-RL carries the following advantages for the military applications, as it

- Can include existing knowledge (e.g., rules of engagements of SLD) and also modify and discover new rules from data
- Learns in an online, real-time, incremental fashion and thus does not require batch processing of (potentially big) data
- Provides the advantage of explainable AI because discovered patterns are represented as rules
- Links to causal learning since it fits the pillars of causal learning (e.g., associations, intervention, and counterfactuals; Pearl & Mackenzie, 2018) by generating the desired effect data using intervention (Wager & Athey, 2018).

The “learn” component can also apply unsupervised learning algorithms. The self-player performs unsupervised machine learning algorithms such as k-means, principle component analysis (PCA), and lexical link analysis (LLA; Zhao & Stevens, 2020; Zhao et al., 2016) for discovering links.

An SLD process needs to perform what-if analysis, as this motivates wargame simulations. An SLD plan can include state variables or problems (e.g., future global and theater posture, threat characteristics, fleet demands to handle these threats), which is only observed, sensed, and cannot be changed. Control variables are solutions (e.g., an SLD plan). LAILOW sets up a wargame between state and control variables. Problems and solutions coevolve based on evolutionary principles of selection, mutation, and crossover.

The number of state and decision variables for an SLD plan and excursion models can be extremely large. Coevolutionary algorithms can simulate dynamic configurations of future warfighting requirements, threats, and global environment and future capabilities, and other competing factors in a wargame simulation. As shown earlier in Figure 1, competitive coevolutionary algorithms are used to solve minmax-problems like those encountered by generative adversarial networks (GANs; Goodfellow et al., 2014; Arora et al., 2017). Adversarial engagements of players can be computationally modeled. Competitive coevolutionary algorithms take a population-based approach to iterate adversarial engagement and can explore a different behavioral space. The use case tests (adversarial attacker population) are





actively or passively thwarting the effectiveness of the problem solution (defender). The coevolutionary algorithms are used to identify successful, novel, as well as the most effective means of solutions (defenses) against various tests (attacks). In this competitive game, the test (attacker) and solution (defender) strategies can lead to an arms race between the adversaries, both adapting or evolving while pursuing conflicting objectives.

A basic coevolutionary algorithm evolves two populations with a tournament selection and for variation uses such as crossover and mutation. One population comprises tests (attacks) and the other solutions (defenses). In each generation, engagements are formed by pairing attack and defense. The populations are evolved in alternating steps: First, the test population is selected, varied, updated and evaluated against the solutions, and then the solution's population is selected, varied, updated, and evaluated against the tests. Each test--solution pair is dispatched to the engagement component, and the result is used as a part of the fitness for each of them. Fitness is calculated overall from an adversary's engagements.

Each SLD configuration possesses a fitness value which is related to measures that need to optimize, such as force development (Fd) and force generation (Fg) efficiency. Patterns from "learn" are used to optimize future SLD plans with the measures of the following:

- Cost of an SLD: for a ship to move from one location to another
  - How much does it cost to move personnel: PCS cost per person x # of billets?
  - How much does it cost to prepare requisite infrastructure (matched assessments) to support that ship move?
- Fd/Fg Efficiency: How many excursions or demands are met (matched)?

The optimization can be overwhelming. LAILOW uses integrated Soar-RL and coevolutionary algorithms and simplifies the optimization process.

LAILOW has been used in wargames in DMO and EABO (Zhao, 2021), discover vulnerability and resilience for the logistics operations for Navy ships and Marine's maintenance and supply chain (Zhao & Mata, 2020), and over-the-horizon strike mission planning (Zhao et al., 2020; Zhao & Nagy, 2020).

## Use Case

To illustrate the process, we first designed and developed a mock unclassified data set to reflect the SLD process. We began by customizing LAILOW to the SLD process in a high level, as shown in Figure 2. This involved defining self-player variables and opponent variables in the SLD process. Self-player variables are also called defender, control, decision, action, or solution variables. The opponent variables are also called attacker, state, problem, or test variables. Opponent variables include profile variables for a ship such as age, maintenance status, decommission schedule, current installation location, capabilities required by fleet commanders as reflected in the excursion plans, and assessments reflected in a collection of warfighting function areas; these variables are considered pre-determined and known information for a ship and cannot be easily changed for decision makers (defenders) at the time of the SLD process. Attacker variables are the state variables for the defenders to address. Decision variables include move (to what location) or stay, cost, manpower, and maintenance readiness, and are also known as defender variables. Both the defenders and attackers evolve and coevolve, and both are guided by their own fitness functions that reflect the self-player and opponent's competing objectives.



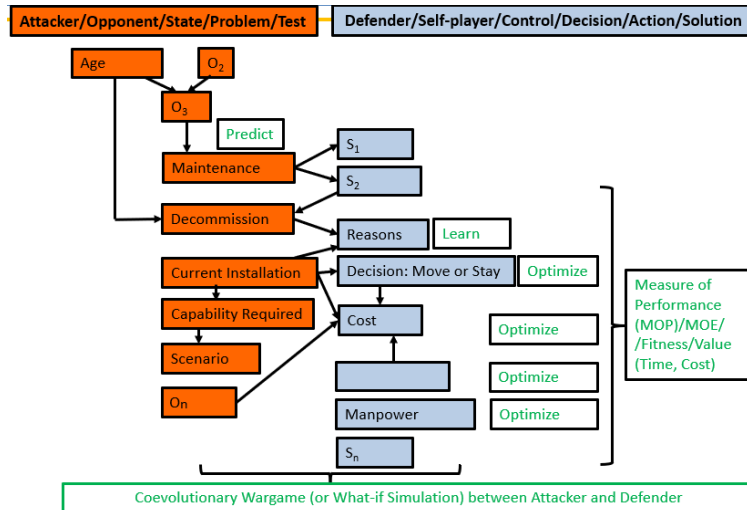


Figure 2. The LAILLOW is tailored to the SLD process in a high level to reflect the what-if decision process used by decision makers in the process.

Figure 3 shows the unclassified mock data set to reflect the understanding of the SLD process in Figure 2.

Name_I	(O)Hull	(O)CurrentInstallationGeolocation	(O)Reason	(S)Decision	(S)NextInstallationGeolocation	(O)Billets_I	(O)DistanceCost_I	(O)Age_N	TotalCost	DecisionCostLow
Wittos	DDG-275	RotaES	OCONUS_PACOMScenario	MOVE	GuamUS	855	7000	15	7835	0
Bismarck	DDG-25	SigonellaIT	COMM	MOVE	MaineUS	692	7000	1	7692	0
Banks	DDG-24	SoudaBayGR	DECOMM	MOVE	NorfolkUS	591	7000	30	7591	0
Banks	DDG-24	SoudaBayGR	DECOMM	MOVE	NorfolkUS	591	7000	30	7591	0
Banks	DDG-24	SoudaBayGR	MAINT	MOVE	SanDiegoUS	591	7000	11	7591	0
Windsor	DDG-245	SoudaBayGR	OCONUS_EUCOM	MOVE	ChinhaeKR	591	7000	15	7591	0
Baldwin	DDG-117	BahrainBH	OCONUS_AFRICOMScenario	MOVE	GuantanamoBay	495	7000	11	7495	0
EarlIsilver	DDG-124	RotaES	OCONUS_CENTCOMScenario	MOVE	BahrainBH	495	7000	11	7495	0
Cameo	DDG-116	NorfolkUS	COMM	MOVE	SigonellaIT	494	7000	1	7494	0
Yorky	DDG-123	ChinhaeKR	OCONUS_EUCOMScenario	MOVE	SigonellaIT	494	7000	11	7494	0
Hokuto	DDG-115	GuantanamoBayCU	OCONUS_EUCOMScenario	MOVE	SoudaBayGR	493	7000	11	7493	0
Rome	DDG-122	GuantanamoBayCU	OCONUS_PACOMScenario	MOVE	KanedaAB	493	7000	11	7493	0
Wild Chrisp	DDG-121	YokosukaJA	OCONUS_PACOMScenario	MOVE	RotaES	492	7000	11	7492	0
Jonathan	DDG-113	GuamUS	OCONUS_PACOMScenario	MOVE	BarkingSandsUS	491	7000	11	7491	0
Avajitlja	DDG-23	NorfolkUS	COMM	STAY	n/a	490	7000	1	7490	0
Godfrey	AS-29	SaseboJA	DECOMM	MOVE	NorfolkUS	420	7000	30	7420	0
Abram	AS-39	SaseboJA	DECOMM	MOVE	NorfolkUS	420	7000	30	7420	0
Nyack	AS-18	MaineUS	COMM	MOVE	SigonellaIT	338	7000	1	7338	0
Hampus	AS-28	MaineUS	COMM	MOVE	GuamUS	338	7000	1	7338	0
Shockley	AS-48	MaineUS	COMM	MOVE	GuamUS	338	7000	1	7338	0
Apollo	DDG-22	NorfolkUS	COMM	MOVE	GuantanamoBayCU	286	7000	1	7286	0
Acheson	AS-37	YokosukaJA	DECOMM	MOVE	NorfolkUS	149	7000	30	7149	0
Lodi	DDG-114	YokosukaJA	OCONUS_PACOMScenario	MOVE	SaseboJA	492	1000	11	1492	0
Ultra Gold	DDG-120	GuamUS	OCONUS_PACOMScenario	MOVE	ChinhaeKR	491	1000	11	1491	1
Fuji	DDG-112	SaseboJA	OCONUS_PACOMScenario	MOVE	YokosukaJA	490	1000	11	1490	1
Suncrip	DDG-119	KanedaAB	OCONUS_PACOMScenario	MOVE	YokosukaJA	490	1000	11	1490	1
Metzger	DDG-311	SaseboJA	OCONUS_PACOMScenario	MOVE	ChinhaeKR	205	1000	2	1205	1
Goldspur	AS-27	YokosukaJA	MAINT	MOVE	HawaiiUS	149	1000	11	1149	1
Adzarnovka	DDG-19	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	5	1080	1
Herma	DDG-191	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	2	1080	1
Orin	DDG-192	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	5	1080	1
Shoesmith	DDG-193	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	5	1080	1
Tinmoth	DDG-194	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	10	1080	1
Wedge	DDG-195	BahrainBH	OCONUS_CENTCOM	STAY	n/a	1080	0	15	1080	1

Figure 3. An unclassified data set designed and developed to reflect the understanding in Figure 2

## Results

We input the mock data into the LAILLOW software and simulate a large number of alternative configurations of Navy assets using the mock data set. Figure 4 shows LAILLOW solutions as heatmaps (solutions). Each cell in each iteration (i.e., generation in the coevolution algorithm), e.g., circled as 1, 2, and 3, represents a potential SLD plan (Defender) against an



environmental test (Attacker), is produced. The heat color shows the fitness for the solution. Clicking on the heatmap cell shows the detail of the corresponding solution configuration.



Figure 4. LAILOW solutions as heatmaps (solutions)

We drill down details of the LAILOW simulation in Figure 4. As shown in Figure 5. The LAILOW software illustrates that better decision configurations (6) than ones in the historical databases (4 and 5) can be discovered using the LAILOW software.

Sequence	Variable	Variable Name	Mean	Soar-RL_1_1	Soar-RL_0_1	Soar-RL_1_0	Soar-RL_0_0	Defender's Reward
0	F0	(O)Age_bt_02_08	0.467889908	0.000605917	0.000137289	-1.90E-06	1.27E-06	1
9	F9	(O)CurrentInstallationGeolocation_KanedaAB	0.009174312	1.20E-05	0.000731227	0	-6.33E-07	1
41	F41	(O)Hull_DDG-119	0.009174312	1.20E-05	0.000731227	0	-6.33E-07	1
128	F128	(O)Reason_OCONUS_PACOMScenario	0.018348624	-1.36E-06	0.000744571	0	-6.33E-07	1
138	F138	(S)Decision_MOVE	0.256880734	-0.00019516	0.00093837	1.27E-06	-1.90E-06	1
155	F155	(S)NextInstallationGeolocation_YokosukaJA	0.018348624	2.40E-05	0.000719252	0	-6.33E-07	1
								0.108772543

Sequence	Variable	Variable Name	Mean	Soar-RL_1_1	Soar-RL_0_1	Soar-RL_1_0	Soar-RL_0_0	Defender's Reward
2	F2	(O)Age_it_02	0.091743119	-5.75E-05	0.000800722	5.93E-11	-6.33E-07	1
10	F10	(O)CurrentInstallationGeolocation_MaineUS	0.055045872	-4.09E-06	0.000747293	5.93E-11	-6.33E-07	1
41	F41	(O)Hull_DDG-119	0.009174312	1.20E-05	0.000731227	0	-6.33E-07	1
123	F123	(O)Reason_COMM	0.073394495	-8.14E-05	0.000824637	5.93E-11	-6.33E-07	1
138	F138	(S)Decision_MOVE	0.256880734	-0.00019516	0.00093837	1.27E-06	-1.90E-06	1
142	F142	(S)NextInstallationGeolocation_ChinhaeKR	0.027522936	1.05E-05	0.000732682	0	-6.33E-07	1
								0.107222755

Sequence	Variable	Variable Name	Mean	Soar-RL_1_1	Soar-RL_0_1	Soar-RL_1_0	Soar-RL_0_0	Defender's Reward
0	F0	(O)Age_bt_02_08	0.467889908	0.000605917	0.000137289	-1.90E-06	1.27E-06	1
6	F6	(O)CurrentInstallationGeolocation_GuamUS	0.073394495	7.00E-05	0.000673199	0	-6.33E-07	1
41	F41	(O)Hull_DDG-119	0.009174312	1.20E-05	0.000731227	0	-6.33E-07	1
128	F128	(O)Reason_OCONUS_PACOMScenario	0.018348624	-1.36E-06	0.000744571	0	-6.33E-07	1
138	F138	(S)Decision_MOVE	0.256880734	-0.00019516	0.00093837	1.27E-06	-1.90E-06	1
142	F142	(S)NextInstallationGeolocation_ChinhaeKR	0.027522936	1.05E-05	0.000732682	0	-6.33E-07	1
								0.108861738

4 is the original in the database, 6 is better than 4 and 5 (in terms of lower cost)

Figure 5. Better decision configurations (6) than ones in the historical databases (4 and 5) can be discovered using the LAILOW software. This shows the potential to discover alternative SLD plans for Naval assets. A defender is an SLD plan for a ship.



## Discussions

In reality, the Navy may need to consider many more variables, such as

- Availability of maintenance, pier space, required training schools, etc.
- The policy that requires each ship overseas to return to the continental United States within 10 years
- How each unit fulfills tactical and strategic requirements that must be maintained
- Unseen political pressures that can outweigh numerically based resource requirements

We anticipate our findings to guide the way forward toward further exploration in this area through our suggested methodology. This would likely save time and energy of the decision makers and offer otherwise undiscovered potential alternative solutions toward the development of future SLD plans. In consideration of future efforts, we envision a more integrated, coherent, and large-scale deep analytics effort leveraging methods that link to existing data sources to enable direct comparisons of potential scenarios of platform movement considered through the SLD process. The resulting product could facilitate decision makers' ability to learn, document, and track the reasons for complex decision making of each SLD process and identify potential improvements and efficiencies.

## Conclusions

We demonstrate the feasibility of the methodologies of leveraging AI to learn, optimize, and wargame (LAILOW) using mock data. LAILOW produces a score derived from a wargame-like scenario for every available ship that might be moved to a new homeport. The score for each ship increases as fewer resources (i.e., lower cost) are required to fulfill an SLD plan requirement to move that ship to a new homeport. This produces a mathematically optimal response and enables an immediate comparison between competing or alternate ship movement scenarios that might also be chosen, thus improving the automation, consistency, and efficiency of the SLD process.

## Acknowledgment

We thank the Naval Postgraduate School (NPS) Naval Research Program (NRP) and the Office of Naval Research (ONR) Naval Enterprise Partnership Teaming with the Universities for National Excellence (NEPTUNE 2.0) program for supporting this research. We would also like to thank the OPNAV N3/N5 team for sponsoring the research topic and providing insightful discussions.

## Disclaimer

The views presented are those of the authors and do not necessarily represent the views of the U.S. Government, DoD, or their Components.

## References

- Arora, et al. (2017). Generalization and equilibrium in generative adversarial nets (GANs). In *International Conference on Machine Learning*, PMLR, 224–232.
- Back, T. (1996). *Evolutionary algorithms in theory and practice: Evolution strategies, evolutionary programming, genetic algorithms*. Oxford University Press.
- Goodfellow, et al. (2014). *Generative adversarial networks*. arXiv preprint arXiv:1406.2661.
- Laird, J. E. (2012). *The Soar cognitive architecture*. MIT Press.
- OpenAI. (2023). <https://openai.com>
- O'Reilly, U., et al. (2020). Adversarial genetic programming for cyber security: A rising application domain where GP matters. *Genetic Programming and Evolvable Machines*, 21(1), 219–250.



- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*.
- Sutton, R. S., & Barto, A. G. (2014). *Reinforcement learning: An introduction*. MIT Press.
- Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228–1242.
- Williams, T. R. (2021). *Modernizing the strategic lay down and dispersal process* (N3/N5, unclassified). [Email communication with the N3/N5 team on 5/14/2021].
- Zhao, Y. (2021). Developing a threat and capability coevolutionary matrix (TCCW) - Application to shaping flexible C2 organizational structure for distributed maritime operations (DMO). In *The 18th Annual Acquisition Research Symposium*.
- Zhao, Y., MacKinnon, D. J., Gallup, S. P., & Billingsley, J. L. (2016). Leveraging Lexical Link Analysis (LLA) to discover new knowledge. *Military Cyber Affairs*, 2(1), 3.
- Zhao, Y., & Mata, G. (2020). Leverage artificial intelligence to learn, optimize, and win (LAILOW) for the marine maintenance and supply complex system. In *The 2020 International Symposium on Foundations and Applications of Big Data Analytics* (FAB 2020) in conjunction with the IEEE/ACM ASONAM. <https://ieeexplore.ieee.org/document/9381319>
- Zhao, Y., & Nagy, B. (2020). Modeling a multi-segment war game leveraging machine intelligence with EVE structures. In *Proc. SPIE 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, 114131V. <https://doi.org/10.1117/12.2561855>
- Zhao, Y., Nagy, B., Kendall, T., & Schwamm, R. (2020). Causal learning in modeling multi-segment war game leveraging machine intelligence with EVE structures. In *Proceedings of AAAI Symposium on the 2nd Workshop on Deep Models and Artificial Intelligence for Defense Applications: Potentials, Theories, Practices, Tools, and Risks*. <http://ceur-ws.org/Vol-2819/session3paper3.pdf>
- Zhao, Y., & Stevens, E. (2020). Using lexical link analysis (LLA) as a tool to analyze a complex system and improve sustainment. In *Unifying themes in complex systems X*. Springer.



## PANEL 24. DIGITAL ENGINEERING IN TEST AND EVALUATION

Thursday, May 11, 2023	
3:45 p.m. – 5:00 p.m.	<p><b>Chair: Sandra Hobson</b>, Deputy Director, Operational Test and Evaluation</p> <p><b><i>Proven Warfighting Capabilities Delivered at the Speed of Need</i></b>            Sandra Hobson, DOT&amp;E            Jeremy Werner, DOT&amp;E            Kristen Alexander, DOT&amp;E            Nilo Thomas, DOT&amp;E</p> <p><b><i>Shifting Left: Opportunities to Reduce Defense Acquisition Cycle Time by Fully Integrating Test and Evaluation in Model Based System Engineering</i></b>            Craig Arndt, Georgia Tech Research Institute            Awele Anyahun, Georgia Tech Research Institute            Jeremy Werner, DOT&amp;E</p> <p><b><i>Using Digital Twins to Tame the Testing of AI / ML and Adaptive Systems</i></b>            David Zurn, Georgia Tech Research Institute            Craig Arndt, Georgia Tech Research Institute            Jeremy Werner, DOT&amp;E</p>

**Sandra Hobson, Ph.D.**—was appointed to the Senior Executive Service in 2014 as a Deputy Director in the Office of the Director, Operational Test and Evaluation (DOT&E), within the Office of the Secretary of Defense. She is currently responsible for defining and executing strategic initiatives, and supporting the development of policy and guidance to meet the test and evaluation demands of the future as the complexity of Department of Defense weapons systems and multi-domain operational environments evolve. Dr. Hobson was selected to perform the duties of the DOT&E Principal Deputy Director from January 20 to December 19, 2021.

Prior to this appointment, Dr. Hobson supported DOT&E in two other capacities: first as a Research Staff Member at the Institute for Defense Analyses and then as an Aircraft Systems and Weapons Staff Specialist. During these tenures, she provided technical oversight of test and evaluation programs to enable adequate assessments of the survivability and lethality of a subset of Department of Defense aircraft and weapons acquisition programs.

Dr. Hobson earned her Bachelor of Science degree in Aerospace Engineering from the United States Naval Academy and her Doctor of Philosophy degree in Aerospace Engineering from A. James Clark School of Engineering at University of Maryland. Dr. Hobson is a recipient of the National Fellowship for Exceptional Researcher awarded by the United Nations Educational, Scientific and Cultural Organization (UNESCO), and two Secretary of Defense Medals for Meritorious Civilian Service. She currently resides in Virginia with her husband and two German Shepherd dogs.



## Proven Warfighting Capabilities Delivered at the Speed of Need

**Sandra Hobson**—was appointed to the Senior Executive Service in 2014 as a Deputy Director in the Office of the Director, Operational Test and Evaluation (DOT&E), within the Office of the Secretary of Defense. She is currently responsible for defining and executing strategic initiatives, and supporting the development of policy and guidance to meet the test and evaluation demands of the future as the complexity of Department of Defense weapons systems and multi-domain operational environments evolve. Hobson was selected to perform the duties of the DOT&E Principal Deputy Director from January 20 to December 19, 2021. [sandra.hobson3.civ@mail.mil]

**Jeremy Werner**—was appointed DOT&E's Chief Scientist in December 2021 after initially starting at DOT&E as an Action Officer for Naval Warfare in August 2021. Before then, Jeremy founded a data science-oriented military operations research team at JHU/APL that transformed the analytics of an ongoing military mission. Jeremy previously served at IDA supporting DOT&E in the assessment of a variety of systems. Jeremy earned a PhD in physics from Princeton University where he was an integral contributor to the Compact Muon Solenoid collaboration in the experimental discovery of the Higgs boson at the Large Hadron Collider, CERN, Geneva, Switzerland. [jeremy.s.werner.civ@mail.mil]

**Kristen Alexander**—is the Chief Learning and Artificial Intelligence Officer for DOT&E. Alexander previously served as the Technical Advisor for the Deputy DOT&E for Land & Expeditionary Warfare. As Technical Advisor, she provided technical and analytical advice for test design, evaluation, and reporting on over 50 Army and Marine Corps programs. Alexander has been involved with operational test and evaluation since 1999 as a Research Staff Member at the Institute for Defense Analyses. She holds BS and PhD degrees in chemical engineering from University of Rochester and Carnegie Mellon University, respectively. [kristen.l.alexander5.civ@mail.mil]

**Nilo Thomas**—graduated from New Mexico State University in 2013 with a BS in aerospace engineering. He worked for the Air Force's 47th Cyberspace Test Squadron for 8 years, where he was a Test Manager leading the largest test programs in the test squadron, Unified Platform and Joint Cyber Command and Control, the DoD's premier cyberspace weapon systems. In 2021, he started working for DOT&E and works as the organization's Software and Cyber Advisor, managing the diverse portfolio of DOT&E software and cyber initiatives on behalf of operational test and evaluation. [nilo.a.thomas.civ@mail.mil]

### Abstract

The enduring mission of the Department of Defense (DoD) is to provide the military forces to deter war and ensure the nation's security. Test and Evaluation (T&E) is critical to the DoD mission success: It enables delivery of the proven, combat-ready systems needed to enable the lethality, suitability, resilience, survivability, agility, and responsiveness of the future Joint Force. With the complexity of the multi-domain operating environment increasing at more dynamic rates, the T&E tools, processes, infrastructure, and workforce must capitalize on the latest advances in science and technology to transform T&E strategies, stay ahead of the adversary, and continue to inspire trust and confidence in the nation's warfighting capabilities while responding to the adaptive acquisition framework to deliver those capabilities at the speed of need. This paper focuses on identifying the transformational changes that the T&E enterprise needs to implement to enable accurate characterization of the operational performance and limitations of the DoD to prevail in conflict and defend the homeland. This paper summarizes the desired end state and preliminary actions to motivate a call for action across government, industry, and academia to define the right measures of performance and accelerate the proposed transformation.

**Keywords:** test and evaluation, mission threads, data management, software-reliant systems, continuous test and evaluation, adaptive acquisition framework, multi-domain operating environment, speed-to-field, culture.



## Strategic Drivers

We have identified seven disruptors, there may be more, that are driving us to rethink the way we do T&E.

### Engineering of Software and Software Reliant Systems.

The DoD Software Modernization Strategy lists an array of challenges that the acquisition and T&E communities need to overcome to deliver the next generation Department of Defense (DoD) software and software-reliant systems, at the speed of need. Modern software development techniques introduce one of the greatest departures from traditional Test and Evaluation (T&E) approaches—a need for a truly integrated, iterative, yet still independent T&E, from code conception to software deployment on weapon systems.

Traditional operational testing and evaluation concepts that focus on one large test in support of a full-rate production decision are not suitable for modern software practices of rapidly deploying capability upgrades. Instead, the adequacy of software T&E design, execution, and reporting depends on their integration into the software pipeline and systems engineering process, while also ensuring continuous user engagement and operationally representative conditions. Contractors' supply chain risk management, program protection, cloud services, software factories, and data rights represent critical factors in the evaluation of software operational effectiveness, suitability, survivability, and lethality (as applicable). The execution of such T&E within existing organizational structures, laws, and policies presents a challenge. It is not yet, for example, fully defined how government T&E could interface with software development teams while still maintaining their independence. It is also not clear how to invoke flexibility to keep pace with the software development cadence while still meeting all documentation requirements. As we consider these challenges for software, we might also consider them for hardware-intensive systems, because there may be some lessons learned here for all of T&E.

### Artificial Intelligence and Machine Learning

Related to engineering of software and software-reliant systems, artificial intelligence (AI) and machine learning (ML) are transformational technologies that introduce yet another layer of complexity to T&E. Rigorous, defensible testing of AI-enabled systems requires additional research to evaluate AI algorithms and ML models, and to ensure AI assurance and to certify that AI algorithms operate as intended and are free of vulnerabilities either from faulty design or from maliciously inserted data or code. T&E also needs to consider the uniquely contextual operational and responsible performance of AI/ML capabilities, especially as they learn and change during real operational use. Given that, where does AI T&E begin and where does it end? Or should it end at all? What tools and processes do we need to put in place to enable continuous evaluation of AI-enabled systems as they get exposed to new and different operational environments? Some form of T&E during operations and sustainment might be our new reality, and we may have to use AI itself and other digital tools to enable such T&E and the related scaling challenges.

### Joint All Domain Operations

Joint all domain operations bring into question the entire concept of testing one system at a time, as has been done historically. Is that approach still suitable in an environment where operational effectiveness, suitability, survivability, and lethality may depend on multiple sensors and shooters, joint targeting, joint kinetic and non-kinetic precision fires, and the like? This machine-speed warfare, integrated across all combat domains, requires us to focus intently on testing the mission threads that make up the system-of-systems environment, including the entire potential attack surface and the persistence we expect from our adversaries. The sheer volume of systems, and their extensive reliance on each other to form effective kill webs, will





require tools and infrastructure that facilitate continuous and automated T&E. It will also require a physical and virtual infrastructure that can adequately represent the operating environment that is changing in both space and time. Some initial thoughts are summarized in Figure 1.

T&E may also need to be more closely tied to training and large force exercises to leverage the various test and training events and corresponding infrastructure. In addition to the infrastructure and instrumentation to support such events, there are several other items that would have to be fully thought through before leveraging training and large force exercises a regular part of our T&E practice: the appropriate range/training capability requirements, mission thread requirements, a joint T&E organization(s) to plan and execute such events, and the associated T&E concepts and cost.

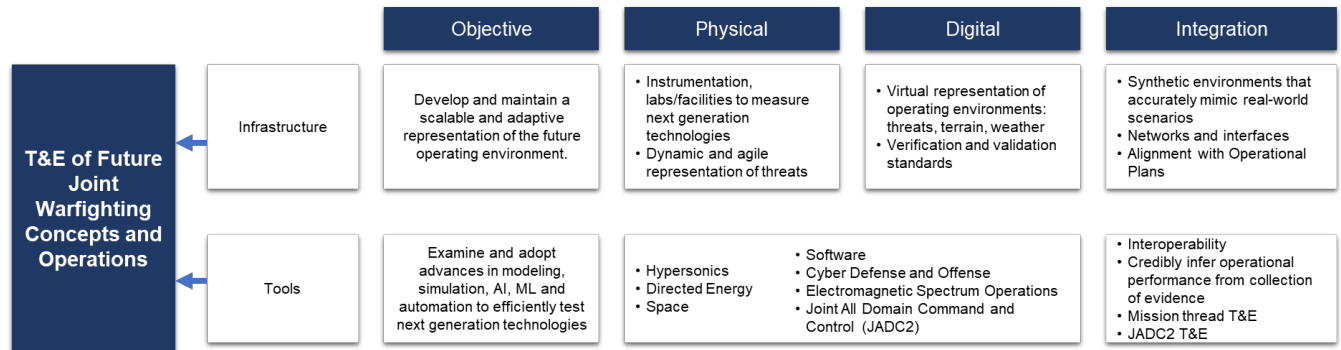


Figure 1. Crude Summary of the Physical and Virtual Infrastructure and Tools Needed to Support T&E of Joint Warfighting Concepts and Operations

## Data

Related to joint all domain operations is the challenge of data. The effectiveness of joint all domain operations will hinge on the ability of relevant systems to deliver information advantage at the speed of relevance: the ability to ingest, sense, analyze, predict, decide, act, and secure data across the entire Joint Force, at every echelon, from the strategic level to the tactical edge, at machine speeds. This will require innovative T&E data management tools to measure and evaluate data-oriented operational performance, especially as data and elements of the kill webs change over time. Let's pause to consider the complexity of the interoperability of such concepts and let's then imagine them in contested space, electromagnetic spectrum (EMS), and cyber environments. This requires an array of different T&E changes—some of which we may have not even considered yet.

At a minimum, the T&E community needs T&E data management plans, standards, repositories, and a large-scale industrial computing infrastructure deployed across the T&E enterprise to enable credible, data-driven decisions on both the complexity level and rapid timescale of future joint warfighting operations. In other words, the collection, analysis, and high-level aggregation of test data must be networked and automated. Moreover, this infrastructure must be integrated with a unified data-driven, all-domain modeling and simulation (M&S) environment—all the way from the tactical edge to the C-suite—to complement live testing in conditions not possible in live tests due to environmental, fiscal, safety, classification, and ethical constraints. *At the edge*, we likely would need raw streaming data from the platform, onboard data reduction, and distributed raw binary data stores. *Across the enterprise*, we may need data post-processing to machine-readable formats; integration into data-backed M&S; “online” system performance analyses using an open, general-purpose automated data analysis environment; and ad hoc “offline” manual, analyst-performed experimentation and development.



At the C-suite, we should leverage the DoD's Advana<sup>1</sup> platform for big data-aggregated advanced analytics using a data mart<sup>2</sup> composed of high-level system performance analysis artifacts.

The rigor behind the automated data and analysis infrastructure described above will enable not only faster analysis but also the complex, tiered analysis necessary to glean high-level and synergistic mission effects: live data collected in simpler scenarios will be “remixed” and fused with M&S to create “digital arenas” for evaluating more complex warfighting scenarios and their related emergent behaviors, using a variety of advanced analysis techniques including AI and Bayesian networks. We look forward to this future.

## Speed to Field

Speed to field is another challenge for the T&E community brought to us by an increasingly capable adversary with access to new science and technologies. Accelerating warfighting capability to the field while not compromising quality and while also confronting the complexity of the technologies themselves requires innovative, if not revolutionary, T&E approaches.

The DoD's Digital Engineering Strategy states that *current acquisition processes and engineering methods hinder meeting the demands of exponential technology growth, complexity, and access to information*. Digital engineering offers an opportunity to enable the acquisition and T&E forces to automate and accelerate workflows and processes. For example, the development of interoperable digital data about systems under development and test could accelerate acquisition by enabling the coherent management of large numbers of systems and mission sets.

Adaptive test planning and execution also offer opportunities to enable faster and more efficient acquisition: the test adequacy for any phase of test is dependent on the outcome of prior testing. For example, the underperformance of a system that falls well below its reliability requirement can be determined with fewer trials than the over-performance of a system that falls slightly above its reliability requirement. Moreover, the development status of the system under test will mature over time (or acquisition milestones); often we will not have a production-representative test asset until initial operational test and evaluation, and yet this test asset will likely be strongly related to its prototype predecessors used in earlier phases of test. In addition, the capabilities and conditions tested in initial operational test and evaluation will often be more stressing than those examined in earlier test phases and subject to different limitations; though, again, the conditions and limitations will typically be related. These dependencies in time among prior observed test results, test assets, and tested capabilities/conditions/limitations suggest the creation of a test design that is both sequential and adaptive to support optimal but adequate T&E in a way that most efficiently leverages data collected across the test life cycle. Figuring this must be a priority.

## Culture

There are several factors that affect the T&E culture that may have to be reconsidered to enable the transformation of T&E processes and outputs:

- Communication is key to success, no matter the situation. Disparate test communities, sometimes spread across different corners of the country, different appreciations for test as compared to M&S, different priorities, and misunderstandings can get in a way of progress, even though all acquisition and T&E stakeholders have the same objective—

---

<sup>1</sup> Advana is the DoD's big data platform for advanced analytics, supported by the Office of the Assistant Secretary of Defense for Acquisition. It is accessible via DoD network only.

<sup>2</sup> A data mart is a structure/access pattern specific to data warehouse environments, used to retrieve client-facing data.



that is, to support the warfighter. Common data and tool repositories, digital engineering, secured and available networks, and infrastructure may improve communication within T&E working-level integrated product teams or integrated test teams and within the T&E enterprise as a whole to identify, prioritize, and track T&E capability requirements.

- The value of T&E can be difficult to appreciate but not if it is characterized as an approach to quantify risk to the program and the warfighter. Programs would benefit from including T&E stakeholders in the development of requirements to evaluate their scope and testability. Similarly, acquisition contracts should also be informed by T&E to confirm access to needed data, artifacts, tools, support to government T&E, and similar.
- Change can be difficult to accept and even more difficult to implement, but continuous change will be the T&E reality as we transition into a digital ecosystem. Application of design thinking principles, for example, could help the T&E working-level integrated product teams solve problems in more structured and efficient ways. Continuous learning and training might become an inherent part of the acquisition and T&E career progression to help all stakeholders successfully maneuver in this new space.

### **Talent Management**






This leads us to our last identified disruptor—talent management. As already alluded to, emerging technologies and digital transformation concepts will require unique skills that the existing workforce may not have. To complicate things, some fields such as software engineering, AI, and cyber science are changing rapidly, requiring continuous training and professional development. To further complicate things, the T&E community has been absorbing responsibilities to keep pace with emerging threats, leaving minimal to no room for training, learning, mentorships, or the like. In addition, high-demand skill sets are also sought by the commercial sector, making it even more challenging for the government to acquire and retain the right T&E workforce. The T&E community is left with the challenge to craft a new approach to recruitment, training, education, and long-term management of the talent pipeline.

### **Strategy**

To respond to these disruptors, the Director, Operational Test and Evaluation, Nickolas Guertin, challenged us to work with him on a strategy that will help transform T&E and enable the delivery of the world's most capable warfighting capability. Under his leadership, we defined the desired end state as five strategic pillars, depicted in Table 1.



**Table 1. Strategic Pillars of T&E Transformation and Capability Delivery**

	Pillars	Key Actions	Desired End State
	<b>Test the Way We Fight</b>	Architect T&E around validated mission threads and demonstrate the operational performance of the Joint Force in multi-domain operations	<ul style="list-style-type: none"> <li>• Accurate representation of the Joint, multi-domain operating environment in test</li> <li>• Adequate evaluation of Joint warfighting capabilities and mission threads (kill webs, system-of-systems performance)</li> </ul>
	<b>Accelerate the Delivery of Weapons That Work</b>	Embrace digital technologies to deliver high-quality systems at more dynamic rates	<ul style="list-style-type: none"> <li>• Discoverable, accessible, and secure data repositories</li> <li>• Near-real-time test data analysis and assessment</li> <li>• Established tools and processes that optimize integrated T&amp;E</li> <li>• Digital documentation and tracking of T&amp;E strategies and plans</li> </ul>
	<b>Improve the Survivability of the DoD in a Contested Environment</b>	Identify, assess, and act on cyber, EMS, space, and other risks to the DoD mission – at scale and speed	<ul style="list-style-type: none"> <li>• Minimized mission-critical vulnerabilities in a contested environment</li> <li>• Timely tracking and response to mission-critical vulnerabilities as systems and threats evolve</li> </ul>
	<b>Pioneer the T&amp;E of Weapon Systems Built to Change Over Time</b>	Implement fluid and iterative T&E across the entire system life cycle to help assure continued combat credibility as the system evolves to meet warfighter needs	<ul style="list-style-type: none"> <li>• Standardized and increased use of credible digital twins in T&amp;E</li> <li>• Adequate assessment of operational and ethical performance of AI-enabled systems</li> <li>• Effective tracking of any degradation of operational performance of DoD systems in theater</li> </ul>
	<b>Foster an Agile and Enduring T&amp;E Enterprise Workforce</b>	Centralize and leverage efforts to access, curate, and engage T&E talent to quicken the pace of innovation across the T&E enterprise	<ul style="list-style-type: none"> <li>• Highly skilled T&amp;E workforce prepared to meet the toughest challenges</li> <li>• Effective continuous learning program and a robust recruitment/retention plan</li> <li>• Agile and innovative workforce operating model</li> </ul>



## **Pillar 1: Test the Way We Fight**

This pillar is focused on accurate evaluation of next-generation warfighting capabilities and requires an adequate representation of the theater operating environment during test, training, and mission rehearsal. It also requires equipment, both physical and digital, that can adequately measure technical and operational performance of emerging or fielded warfighting capabilities in that environment. The DoD has an array of test and training ranges and capabilities managed, funded, and operated by different stakeholders. To enable efficient and structured modernization and sustainment of existing range capabilities while also transforming the ranges to meet the demands of the future, it is important to have an accurate and common picture of the existing and required future range capabilities and requirements. It will be equally important to ensure this common picture is accurate, digitized, and transparent to key T&E stakeholders to enable collaboration in developing joint/interoperable solutions, avoiding redundancies, and increasing capability delivery and efficiencies.

This pillar also considers real-world mission scenarios that involve the use of multiple systems of varying complexities and pedigrees working together to achieve the desired lethal effect. The emergence of highly network-centric concepts, greater dependency on connectivity, and the use of large amounts of data from a wide array of shooters and sensors across multiple domains, at machine speeds, warrants a review of our T&E processes within individual acquisition programs. Evaluating warfighting capability is further challenged by asynchronous updates and continuous evolution of the various components that comprise these system-of-systems operations. These evolutions demonstrate an inherent need to continually characterize the interoperability and effectiveness of such systems as they would be employed by the combatant commands. With the emergence of joint all-domain command-and-control solutions and the concept of kill webs, it is important to define the process and the required T&E tools that would effectively measure the success rates of mission threads, concepts, and solutions.

## **Pillar 2: Accelerate the Delivery of Weapons That Work**

Data are strategic assets that fuel automation and algorithms designed to alleviate our workload, speed up our processes, help us achieve new insights, and achieve T&E at scale and speed. As data-driven and complex systems continue to proliferate, it is important to develop T&E data and interface standards, stores, and platforms to ensure that the data are credible, trustworthy, available, and secure across the T&E enterprise. The T&E community must demonstrate its compliance with and contribution to implementing the DoD Data Management Strategy and enable data collection, storage, visibility, sharing, accessibility, ingestion, and security across commercial and government stakeholders to expedite data analysis, optimize T&E planning and execution, and enable more automated T&E. This compliance translates to availability of data stores, knowledge management tools, and data fusion/analytics tools that will enable the new fluid and iterative nature of T&E demanded by software- and data-reliant systems. Data are also critical to verify and validate digital tools. Lastly, all data (contractor, developmental, operational, and live fire) must be effectively leveraged to adapt, inform, and optimize T&E plans—no data should be left behind.

Having a data management plan will also enable modern model-based engineering and adaptive inference processes that offer integrated, holistic approaches to generating and managing knowledge of system performance throughout its life cycle. Early test data from system components, for example, can be integrated into a larger system model to predict mission-level performance early in development. Advanced performance inference techniques (e.g., Bayesian) can be used to carry forward data from early prototypes through evaluation of production-representative systems. Moreover, model-based engineering can eliminate manual workflows through automation that enables generation and distribution of up-to-date dynamic reports on systems and their status in the acquisition life cycle.



This is further emphasized in Section 223 of the Public Law 117-81. The T&E community will have to leverage heavily model-based systems engineering and other digital tools and technologies to enable full-spectrum survivability (and lethality) evaluations. Full-spectrum survivability evaluations are intended to enable the survivability of DoD systems and services in a multi-domain operational environment, accounting for both kinetic and non-kinetic threats—such as cyber; directed energy weapons; EMS fires; chemical, biological, radiological, and nuclear threats; and any combination thereof. Moreover, full-spectrum survivability evaluation is intended to leverage digital technologies required to enable such evaluation throughout the life cycle of the acquisition program, as both the fielded system and the threat(s) evolve over time at more dynamic rates.

### **Pillar 3: Improve the Survivability of the DoD in a Contested Environment**

Seamless integration of various systems and technologies working together across multiple domains introduces the potential for vulnerabilities that cannot be evaluated one system or one threat at a time. As discussed under the first pillar, testing must consider the mission thread, specifically the composition of weapon systems, networks, critical infrastructure, equipment, and tactics, techniques, and procedures. A rapid and accurate mission-based survivability assessment would define specific steps to enhance mission assurance and identify the defenses required against threats to those missions.

To add to this complexity, since weapon systems of today and the future are defined by both software and hardware, battle networks are central to the kill web, and information technology is at the heart of cyber, space, and EMS warfare. The complex interactions between software and hardware can sometimes be difficult to predict or evaluate. Our challenge is to evaluate cyber–physical systems against advanced cyber and EMS threats at scale and speed. Attack surfaces are growing exponentially, reaching into supply chains, software pipelines and factories, the EMS, and an array of cloud solutions. We therefore must aggressively pursue verified and validated digital tools and transformative technologies to manage cyber, EMS, and advanced kinetic threat survivability T&E and assess the effectiveness of countermeasures and other self-defense solutions.

In addition, space is increasingly congested and highly contested, with a broad array of rapidly evolving threats warranting its own focus. Reliance on space-based capabilities has sharpened the DoD's—and our adversaries'—focus on deploying both offensive and defensive weapons in space. Because the DoD must operate in this contested environment, the T&E enterprise must be ready to accurately evaluate space-based and space-dependent systems' operational performance, including survivability against current and anticipated threats.

### **Pillar 4: Pioneer T&E of Weapon Systems Built to Change Over Time**

The fourth pillar is focused on addressing the T&E challenges associated with complex, largely software-reliant systems, the operational performance of which could be affected by incremental and frequent software upgrades and/or frequent and dynamic changes to the operating environment. Related to software-reliant systems, this pillar also focuses on the challenges brought by AI and ML capabilities. All elements of this pillar are counting on advances of the digital ecosystems, which start with the development of credible digital twins—high-fidelity digital representations of physical objects.

In addition, the combination of new domains and operational constraints makes verified and validated digital technologies the necessary, practical approach for development and T&E of certain systems where live T&E is not possible or practical. For example, digital twins that can be subjected to repeated cyberattacks—as the system itself, the threats it will face, and adversary tactics, techniques, and procedures change over time—will help developers and program managers improve system cyber survivability at an increased pace. These types of



models allow us to find out how real-world objects might behave under different conditions or requirements. The defining feature of a digital twin is the ongoing data integration between the digital model and its physical unit counterpart. Digital twins have demonstrated the capability to incorporate transmission of real-time data sensed by the real-world object. These new, higher-resolution sensor data allow the digital twin to reason about future behaviors, then transmit feedback to the physical object. This ability could be particularly useful in enabling continuous monitoring of operational performance of systems as they evolve over time. Unfortunately, while digital twins create new opportunities for T&E to determine the performance of continuously evolving systems, they also create new verification, validation, and accreditation challenges.

On a related note, AI-based systems have accelerated the need to re-engineer T&E to enable continuous assessment once fielded. The T&E enterprise must be positioned to monitor and evaluate the drift in deployed AI models' behavior, which could occur when real-world data deviate from the training data used to create the model. Testing also must demonstrate with confidence that AI-based systems are responsible, ethical, equitable, traceable, reliable, and governable. Ethical and safe use of AI is necessary to reduce risks to U.S. strategic initiatives, reputation, operations, legal standing, and privacy issues. Due to their reliance on ever-changing data, however, AI-based systems are uncertain by nature. Emerging approaches that have the potential to address such uncertainty propagation deserve further investigation. Additional research is also needed to re-envision the T&E process with increased AI and automation tools to support T&E professionals and identify opportunities where AI can assist them—relieving them of tedious tasks so they can better focus on tasks that require the creativity and innovation that only humans can provide.

Lastly, and as discussed in the Strategic Drivers section, modern warfighting systems are increasingly software-reliant. They are developed through complex software pipelines filled with a myriad of tools intended to ensure automatically that the product is effective and secure. However, developers frequently use open-source and third-party software, which raises the risk from the security and sustainability perspectives. It is important to identify new approaches to address change propagation within software-reliant systems. For example, the survivability T&E community needs to influence and measure the development and cyber defense of software pipelines up front with accredited tools, techniques, and procedures. Automated testing should be embraced at every level, and a rigorous standard of testing should continue to be implemented at the speed of relevance.

## **Pillar 5: Foster an Agile and Enduring T&E Workforce**

A structured approach for the collective development and sustainment of the T&E enterprise workforce will enhance workforce agility and response to emerging T&E requirements. Dedicated T&E skill codes and qualifiers to track T&E professionals' knowledge, skills, and abilities would improve the DoD's awareness of the T&E workforce's overall health and development. An infrastructure to make data-driven workforce planning decisions would enable the T&E enterprise to forecast, track, and address gaps in the T&E workforce's collective capabilities. It would also enable unified development of the T&E enterprise workforce, as well as its agility to move among the requirements developers, technology developers, buyers, and across the service T&E communities.

T&E professionals of the future require access, bandwidth, and clear requirements to engage in continuous learning opportunities. Providing these opportunities will better prepare them for the advances in T&E operational and technical capabilities needed to perform their duties. It is important to establish enterprise-wide baseline education and training needs and the ability to identify all T&E-related course offerings to strengthen workforce capabilities. The T&E learning apparatus should change as quickly as the T&E operating environment, with easily



adaptable courses, content, and training based on current workforce demands. It is also important to establish a continuum of cutting-edge learning opportunities that can tie training and education to specific job and career outcomes across the enterprise; this would improve and incentivize T&E learning and workforce retention. To compete with private sector organizations for top-tier talent and promote retention, the T&E enterprise will need to invest in workforce experiences and the SkillBridge program to appeal to a diverse T&E workforce in terms of skill development, rotational opportunities, and leadership roles.

## **Implementation Plan**

Requirements, intelligence, and the acquisition pathways are the foundation of the T&E process. Changes in capabilities, such as kill webs, complex all-domain environments, and gaps newly identified by intelligence reports, will steer acquisition decisions and commensurate T&E responses. Based on the requirements, intelligence, and mandates sourced from the six acquisition pathways, the T&E community must collaborate to identify and develop the T&E capabilities necessary to test and evaluate systems in the acquisition pipeline. These T&E activities will realize the goals of the five strategic pillars that will in turn inform T&E policy and guidance with the potential to inform operational and system requirements, system development, and acquisition contracts. It is imperative to work together and promote a pioneering spirit, as well as a culture of continuous learning, agility, transparency, and co-ownership, to use our combined talents most effectively and accelerate research and development needed to transform T&E tools, processes, infrastructure, and human capital.

## **Conclusion**

The DoD faces a shifting threat landscape and the need to swiftly leverage advanced technologies to increase the lethality, suitability, resiliency, survivability, agility, and responsiveness of our future Joint Force. To continue to deliver credible warfighting capability at the speed of need, the acquisition and T&E enterprise must rethink traditional approaches. We must respond with agility, efficiency, and effectiveness to adequately account for the technology disruptors as we face an inflection point in the scope, scalability, and capabilities of our infrastructure, tools, processes, and workforce. It is our responsibility to set the framework to leverage ongoing government-based activities, the best practices of industry, academia, and our allies to develop a future-ready T&E enterprise.

The T&E enterprise of the future must be agile, motivated by mission thread approaches, joint warfighting concepts, and the power of digital tools and technologies. It must be strengthened by the effects of these changes on our ability to support the warfighter. It must be empowered by continuous learning and supported by unbound access to state-of-the-art skills and technologies to be better positioned to stay ahead of the adversary and continue to advocate for the warfighter and its mission as defined by the National Defense Strategy 2022





# Shifting Left: Opportunities to Reduce Defense Acquisition Cycle Time by Fully Integrating Test and Evaluation in Model Based Systems Engineering

**Craig Arndt**—currently serves as a principal research engineer on the research faculty of the George Tech Research Institute (GTRI) in the System Engineering Research division of the Electronic Systems Lab. Arndt is a licensed Professional Engineer (PE), and has over 40 years of professional engineering and leadership experience. Arndt holds engineering degrees in electrical engineering, systems engineering, and human factors engineering and a Masters of Arts in strategic studies from the U.S. Naval War college. He 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. In industry he has been an engineering manager, director, vice president, and CTO of several major defense companies he is also a retired naval officer. [Craig.Arndt@gtri.gatech.edu]

**Awele I. Anyanhun**—is a senior research engineer in the MBSE Research and Application Branch at Georgia Tech Research Institute (GTRI). She is an INCOSE-Certified Systems Engineering Professional (CSEP) and Senior Member of IEEE with over 13 years' professional experience in architecting complex automotive, space, and defense system architectures. In her current role as a project director and systems architect, she provides thought leadership and systems engineering expertise on DoD-sponsored DE and MBSE projects. Anyanhun is an OMG-Certified Systems and Software Modeling Professional, SysML-MBA, OCUP 2-MBA, and a UL-Certified Functional Safety Professional (UL-CFSP). Anyanhun has authored multiple conference and journal publications, and holds a PhD in Electrical Engineering with a concentration in systems engineering. [Awele.Anyanhun@gtri.gatech.edu]

**Jeremy Werner, PhD, ST**—was appointed DOT&E's Chief Scientist in December 2021 after initially starting at DOT&E as an Action Officer for Naval Warfare in August 2021. Before then, Werner was at Johns Hopkins University Applied Physics Laboratory (JHU/APL), where he founded a data science-oriented military operations research team that transformed the analytics of an ongoing military mission. Werner previously served as a Research Staff Member at the Institute for Defense Analyses where he supported DOT&E in the rigorous assessment of a variety of systems/platforms. Werner received a PhD in physics from Princeton University where he was an integral contributor to the Compact Muon Solenoid collaboration in the experimental discovery of the Higgs boson at the Large Hadron Collider at CERN, the European Organization for Nuclear Research in Geneva, Switzerland. Werner is a native Californian and received a bachelor's degree in physics from the University of California, Los Angeles where he was the recipient of the E. Lee Kinsey Prize (most outstanding graduating senior in physics). [jeremy.s.werner.civ@mail.mil]

## Abstract

The reduction in cycle time for acquisition programs, or “Shift Left” is important to realizing the benefits of digital engineering (DE), as specifically addressed in the DOT&E Strategy update in 2022. Although DE has long held the promise of making programs faster, and achieving goals and priorities more efficiently, its effect on reduced acquisition cycle time is still difficult to identify and quantify. Furthermore, problem discovery during testing and evaluation (T&E) has been identified as a critical driver in the time it takes to develop systems and is said to have significant impact on the acquisition cycle time. Hence, a reduction in acquisition cycle time can be achieved through a systemic approach that positively impacts the time required to test systems while maintaining or reducing risk. Therefore, expanding the use of DE and model-based systems engineering (MBSE) to include test capability models creates the opportunity to improve testing and development of defense systems as well as reduce the defense acquisition life cycle time. To this end, this paper will present the quantitative results of a project that expands the use of MBSE within the test and evaluation space through the creation of a model-based test integration prototype. The results will show where and how test modeling can be used to impact acquisition decision making and reduce overall program schedule.

**Keywords:** Digital Engineering, MBSE, Test planning, Shift left



## Introduction

The transformation from the historical, document-based acquisition system to digital engineering (DE) is resulting in some of the most significant changes to the way the DoD has engineered and developed weapon systems in decades. The shift to the use of DE will not only impact the DoD but the entire military-industrial complex. Coined by President Eisenhower in a 1961 address to the American people, the “military-industrial complex” includes the contractors that develop and manufacture our nation’s combat systems (History.com Editors, 2009).

In some ways, the transition to DE is the DoD’s reaction to the larger endeavor in the engineering community to reduce development time and cost by using digital data management technologies across development and manufacturing enterprises. In the DoD’s “Digital Engineering Strategy,” the DoD states that “current acquisition processes and engineering methods hinder meeting the demands of exponential technology growth, complexity, and access to information” (DoD, 2018). DoD leadership believes that DE will enable the DoD to meet the current and upcoming challenges to delivering new capabilities to the warfighters in support of the DoD’s numerous complex missions. To accomplish this, it is crucial to have a realistic DE strategy in place that can be implemented with new DE technologies while maintaining compliance with current acquisition policy and best practices.

In balancing these constraints against the opportunities of DE, several key goals and needs of the DoD must be considered. First the goal of the department acquisition activities is to deliver to the warfighters the best possible systems in a timely, cost-effective manner in order to maximize lethality and survivability. Second the different acquisition activities need to use and create data, information, and knowledge in a manner that improves critical processes already in the acquisition system and allows programs to be managed based on their risk profile and their impact to the existing and future operational use in concert with other operational systems and in the presence of future threats. In order to maximize the positive impact of DE and modeling we need to implement DE in a manner that specifically addresses speed, risk, and quality of decision making, across portfolios, in a manner responsive to relevant missions.

## Background

The adoption of additional technologies or methodologies is often accompanied by a myriad of questions regarding the scope of adoption or degree of utilization of the introduced concept. Digital engineering is no different. Many of the original implementations of DE, and more specifically model-based systems engineering (MBSE), have focused on the development of models of requirements and the design of systems to meet these requirements. In many instances they lack a meaningful set of modes of the test process. This is problematic for a number of different reasons. First, because test and evaluation (T&E) are critical parts of the development life cycle accelerating the development of systems (a key goal for the DoD) cannot be properly addressed without detailed modeling of the T&E process. Additionally, the information and data that is collected during different parts of the T&E process (including developmental and operational test) is critical to making good decision about every aspect of a given program. In the *DOT&E Strategy Update 2022*, Nickolas Guertin, the Director of DoD Operational Test and Evaluation states the second strategy pillar of DOT&E strategy is “acceleration the delivery of weapons that work” and he points out that MBSE is needed to achieve this shift-left (DOT&E, 2022).

At the beginning of the development process essential mission requirements are identified as Key Performance Parameters (KPPs), key acceptance criteria, or Measures of Effectiveness (MOEs), organizations typically use one or more of these terms for their essential mission requirements. All key stakeholders must agree to these KPPs or MOEs early in the life



of the project, because these are the select few, critical, and non-negotiable criteria that the solution must satisfy to be acceptable. They represent the absolutely critical subset of measurable and observable capabilities and characteristics that the solution must meet. Developing KPPs, MOEs, or key acceptance criteria is a joint effort between the systems engineers and the key stakeholders. The DoD defines KPPs in the initial capabilities document and validates them in the capability description document. Defining KPPs often takes the collaboration of multiple stakeholders, but it is critical to providing focus and emphasis on complex multi-year multi-agency development programs.

Modeling of the testing capability, in this case the test ranges, is a key part of integrated program modeling. In modeling the system, the KPPs and the MOE define the test cases needed for the testing program. In order to demonstrate the modeling, we developed models for the test cases and for the testing capabilities of a generic electronic warfare system. The testing capabilities are modeled in the form of a test range model. For this project we developed a partial model of the Eglin range. The range model was used to capture specific capabilities of the range to be used in constructing the test use cases. The test range models the requirements, system design, and test cases. A risk model is then linked to the other models. In this way the risk model gets data from the other models, and can be used to aggregate the risks based on differences between the available test resources and the requirements for those resources in test execution.

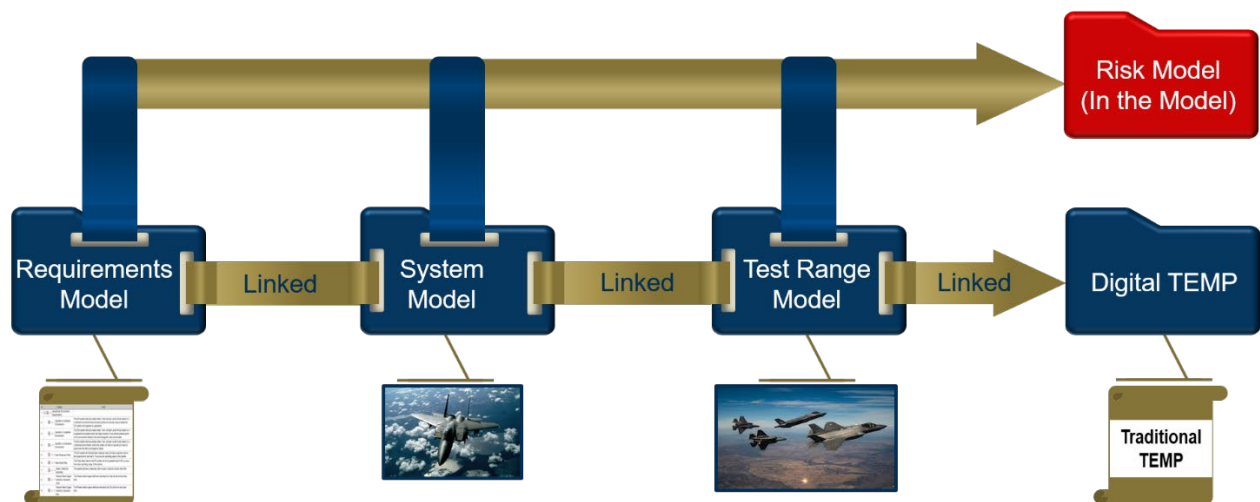


Figure 1. Life Cycle Modeling Structure

In addition to the need for greater use of modeling in T&E and the linking of these models to the larger set of requirements and systems models, there is a great need to look at the way risk is modeled as a function of the models that exist in DE processes. Risk in this context is inherently a function of different aspects of the program's ability to create and deliver a useful product to the field. Traditionally risk functions have focused only on the risk to the program of developing the end item product. The traditional risk approach also developed risks based on specific design risks. This approach has several shortcomings. First, this approach does not guarantee a comprehensive coverage of all possible risks. Second, the traditional approach has no specific way of addressing risk created in the testing program independent of the acquisition risk of different parts of the system. Third, the traditional approach to risk does not have a direct means of aggregating risk from the system's operational mission or operational environments. As a result, these three specific areas form the requirements set for the development of a new risk approach.

1. Develop a risk function that is comprehensive across all areas of the program.
2. It is critical that the risk function capture risks that are inherent to the testing of systems.
3. A risk function needs the ability to aggregate risk across different aspects of the program, specifically aggregate risks across mission areas and operational environments.

The development of more robust T&E and risk modeling generate the data, viability and insights needed to make decisions to accelerate acquisition programs.

## **Shifting Left: Applying an MBSE Approach to Test Planning, System Testing, and Evaluation**

The power of system modeling to address the challenge of accelerating the acquisition process and reduce cycle time can be best demonstrated by a specific use case. In this use case we developed a representative set of models that captures requirements, system architecture, test range architecture, test cases, and risk functions. The use case shown in this section will demonstrate how test organizations can integrate and link together requirements models, system architecture/design models, test ranges, and system under test (SUT) models in order to reduce the time it takes to test systems while maintaining or reducing risk. In order to develop the system and test models in a form that would serve as a good example for future acquisition programs as well as to prototype the process of developing and linking the models, a simplified version (notional) of a real electric warfare (EW) system, the AN/ALQ-161A (Angry Kitten) electronic countermeasures system design for the B-1B bomber aircraft was chosen. The integrated model is composed of several independent models holding requirements, system architectures and test artifacts in self-contained modules that can be updated and augmented independently when new data is available. The EW use case was selected because of our ability to abstract it to more general defense acquisition and because there are several well understood missions that can demonstrate different risk profiles.

### **Specifying Requirements Modeling for Test**

Requirements engineering is a vital part of the (model-based) systems engineering (SE) process because it defines the problem scope and links all subsequent system development, system testing, and risk analysis information to it (Dick, 2017). A set of exemplar requirements that capture the requirements of the system to be developed and tested, requirements of the test range(s) required for performing system tests, and testing requirements are captured in a requirements model and are further refined in the system architecture and test range infrastructure models. Highlighted in this section are the requirements sets for a specific capability of an EW electronic countermeasures system, the requirements sets for a test range that would be used to test the EW system capability, and the testing requirements. Requirements engineering constitutes the branch of SE that bridges the gap between the informal world of stakeholder needs—which in this context is representative of the test community and program office—and the formal world of a reduced cycle time for defense acquisition.

**Specify the System of Interest (Sol) Requirements.** The desired mission capabilities of the EW countermeasures system are first specified and modeled as system requirements. The system-level requirements describe the functions and quality attributes (non-functional requirements) the system must fulfill in order to satisfy the program office's needs. Functionally, the EW system is expected to operate optimally within several operational environments based on specific user-defined missions. A couple of operational environment requirements levied on the EW system are shown in Figure 2. For this exemplar model, the main mission capability



expected of the EW countermeasures system is the ability to *provide situational awareness* during missions.

As highlighted in Figure 3, the high-level requirement *EW Situational Awareness* is decomposed into two main requirements: *EW Operationally Effective* and *EW Situationally Effective*, which are further decomposed into atomic requirements that can be tested. Also captured in the requirements diagram are *Derived Requirements—Computed Correct ID Performance*, *Computed Incorrect ID Performance*, and *Computed Missed ID Performance* which are derived from the *Computed Identification (ID) Performance Requirement*. These mission requirements represent Key Performance Parameters (KPPs) or Measures of Effectiveness (MOEs) that are critical for the success of the mission and which a test range will need to have the capability to test.

#	△ Name	Text
1	Operational Environment Requirements	
2	62.1 Operate in Contested Environment	The EW system shall accurately detect, track and jam, active threat radars in a contented environment where threat systems are actively trying to defeat the EW system and degrade its capabilities.
3	62.2 Operate in Congested Environment	The EW system shall accurately detect, track, and jam, active threat radars in a congested environment where are large numbers of non-threat systems active in the environment adding to the electromagnetic noise environment.
4	62.3 Operate in Constrained Environment	The EW system shall accurately detect, track, and jam, active threat radars in a constrained environment, where the system will need to operate are reduces power and not emit in all frequency bands.
5	62.4 False Response Rate	The EW system will minimize false response rates, the false response rate for the system will be less that 0.1% across the operating range of the system.
6	62.5 False Alarm Rate	The False alarm rate for the EW system will be not greater than 0.02%, across the entire operating range of the system.
7	62.10 Mode 2 Detection Sensitivity	The system will have a detection rate in mode 2 shall be not less than 90%.
8	62.11 Passive Detect Signal Detection Sensitivity - Omni	The Passive detect signal detection sensitivity for Omni will be not less than 90%.
9	62.12 Passive Detect Signal Detection Sensitivity - ESA	The Passive detect signal detection sensitivity for ESA will be not less than 95%.

Figure 2. Operational Environment Requirements View

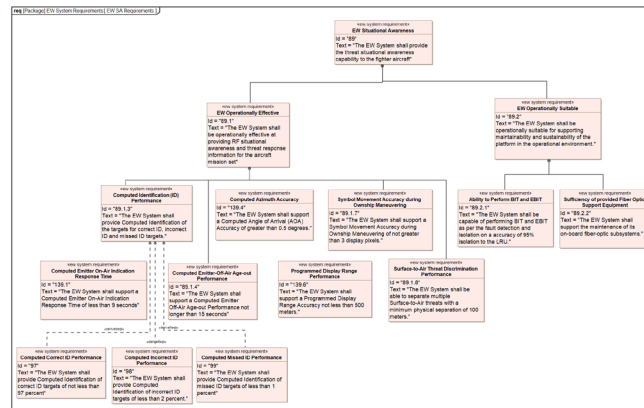


Figure 3. EW System SA Requirements

**Specify Test Range Requirements Views.** For the purpose of this exemplar, the requirements captured as Test Range Requirements are limited to a specific test range capability as shown in Figure 4. To leverage MBSE as a means of positively impacting the time it takes to test systems, development of test range models is vital. Therefore, test range capability requirements are captured within the requirement model which enables traceability to both the test range infrastructure and the Sol requirements. For example, the *Probability of Target Identification Test Range Requirement* specifies that the test range of interest must be capable of testing whether an EW system correctly identifies target/threat systems with a confidence greater than or equal to 90%. The benefit of specifying test range requirements in



model form is that it allows for a real-time gap and impact analysis of a test range's capability to test certain system capabilities and enables better test planning by organizations. Having such information readily available can help reduce the time needed to identify the appropriate test ranges needed to test specific systems/system capabilities.

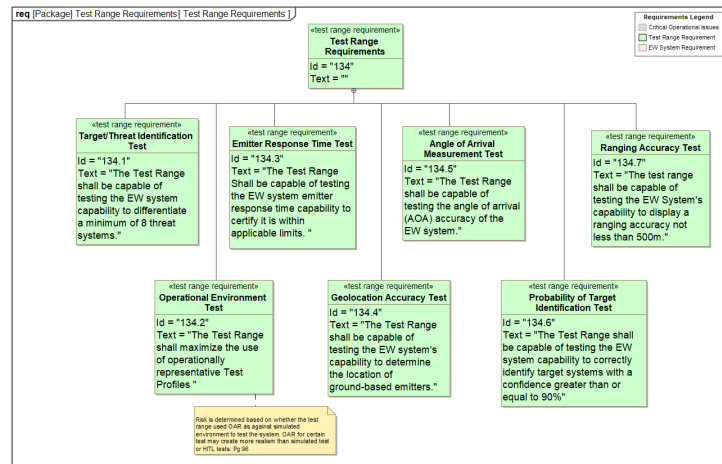


Figure 4. Notional Test Range Requirements View

**Define Test Requirements and Test Objectives.** Also captured as part of the requirements model are testing requirements and test objectives as shown in Figure 5 and Figure 6. Testing requirements also called critical operational issues (COI) outline the issues that are examined during testing and evaluation to determine the system's capability to perform the mission. An example of a testing requirement for the notional system is stated as *“Does the EW system provide effective situational awareness to the aircrew?”* It follows then that COIs represent the requirements by which the suitability of the system under test (SUT) will be assessed from a mission perspective. Capturing Test Objectives in a model-based format facilitate tracing of test objectives to system models and test range infrastructure, enabling test personnel to make key decisions in a timely manner.

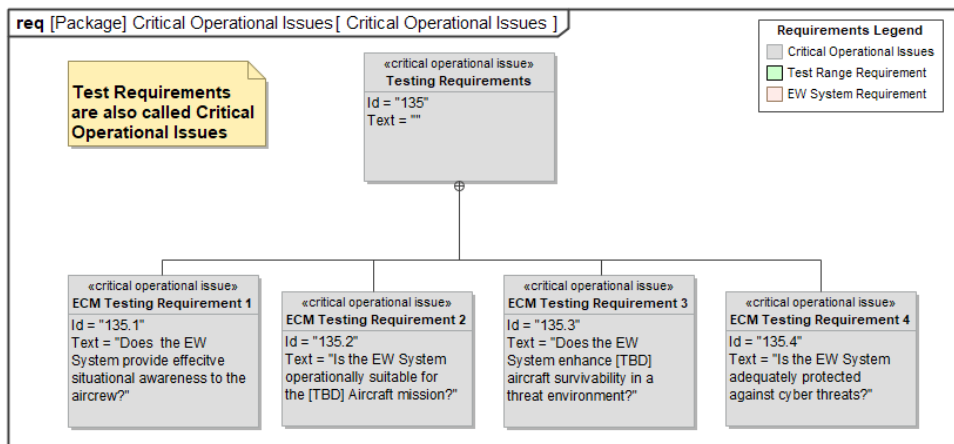


Figure 5. EW System Testing Requirement View



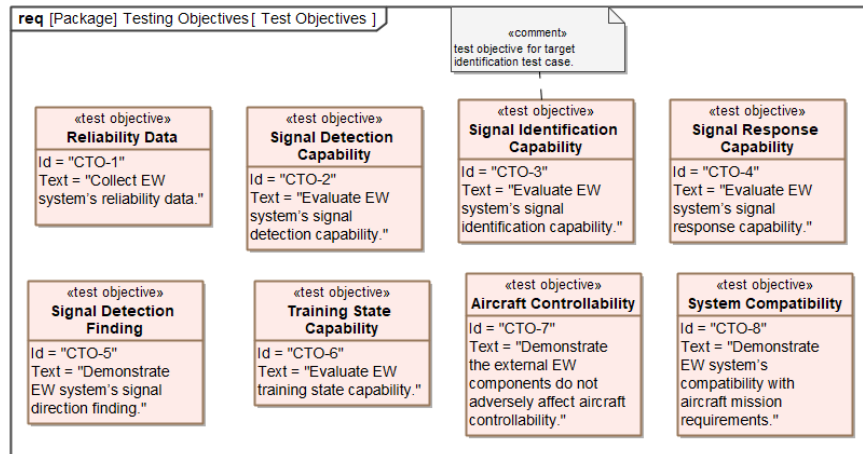


Figure 6. EW System Test Objectives View

## EW Countermeasures System Architecture Definition

The approach taken in the development of the EW countermeasures notional architecture was to first define the conceptual or black box architectural view of the system which entailed defining the EW system as a black box, and capturing its operational domain. This approach enabled the identification of external interfaces and specification of high level (black box) test cases. Once the conceptual view of the system had been defined, the next step taken was to develop the logical or white box architectural view of the system. This section highlights some of the architectural views created as part of the system architecture definition. The views which comprise both behavioral and structural depictions of the system architecture facilitate the development of test cases for the EW countermeasures system. In order to simplify the modeling and make the process more generalizable for multiple programs, the decision was made to use only unclassified information.

**Identify SOI Capabilities and Specify Conceptual Architecture.** Two main capabilities of an EW countermeasures system include its ability to *provide situational awareness* to the pilot and *execute self-protection*. The situational awareness capability is the focus for this exemplar model, and hence, most architectural views and artifacts presented here are skewed to this capability with all other aspects either abstracted out or simplified. Figure 7 depicts the *Perform RF Source ECM* capability as a use case of the RF Electronic Countermeasures System. The combined behavior of the use cases, *provide situational awareness* and *execute self-protection*, represent the overall behavior of the *Perform RF Source ECM* capability. Identified primary actors include the *pilot* and *aircraft* while secondary actors (systems) have been identified as EW Threats and Enemy EW Systems. A significant benefit of an MBSE approach is the ability to determine very early in the system acquisition cycle which test resources are required to test a certain system/capability. In the case of the EW countermeasures system, it is apparent that in order to perform a live test of the *provide situational awareness* capability, the test range used should have the capabilities that represent an Enemy EW system. The high-level *perform RF Source ECM* scenario view depicted in Figure 8 highlights the interactions between the Sol and external systems (i.e., enemy threat and radar systems) within its operational domain, while Figure 9 highlights a structural view of the EW countermeasures system domain. Also specified at this level are MOEs of the Sol.



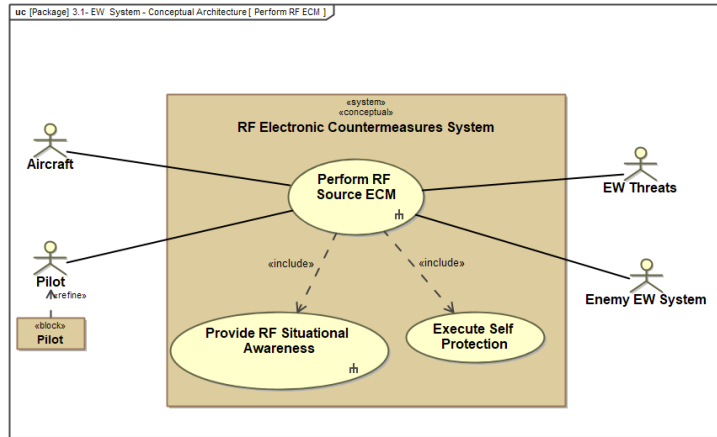


Figure 7. Perform RF Source ECM Capability View

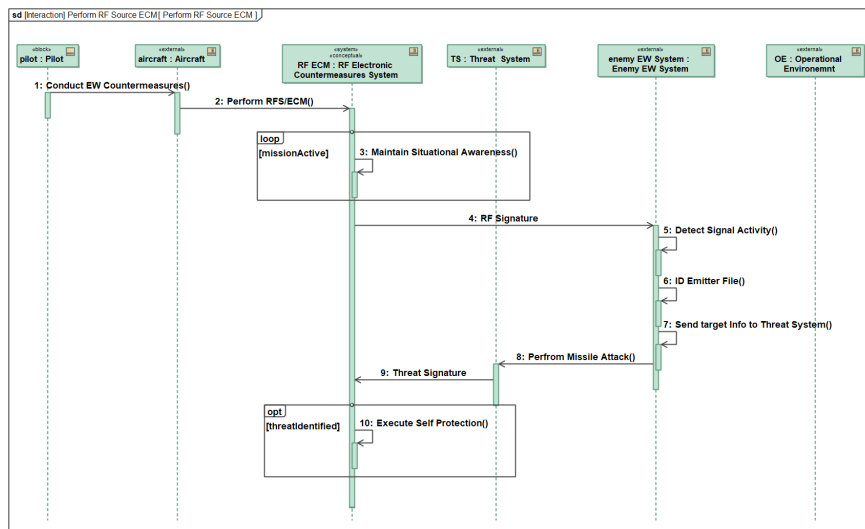


Figure 8. Perform RF Source ECM Scenario View

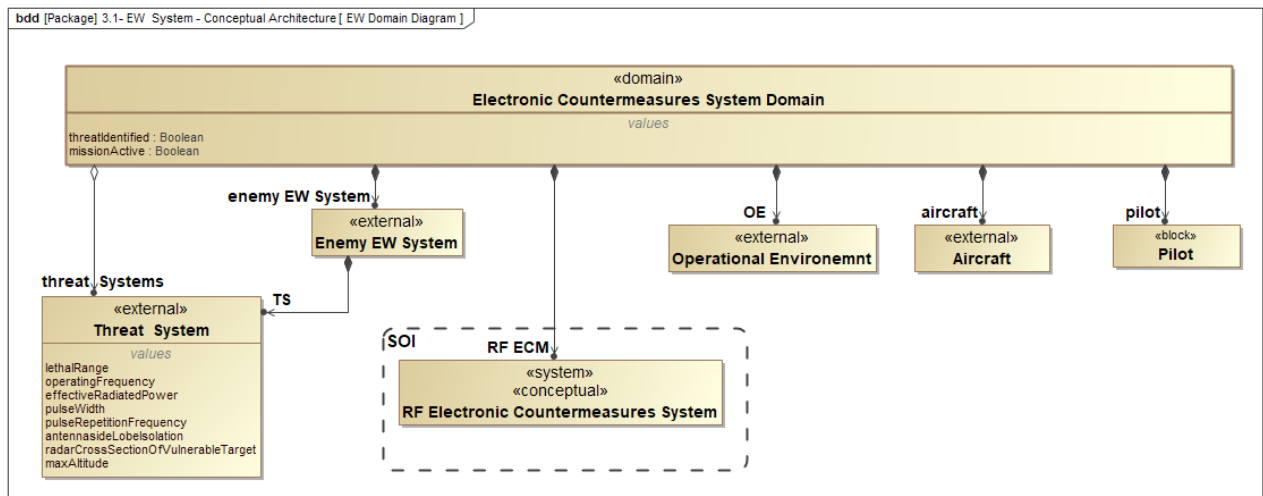
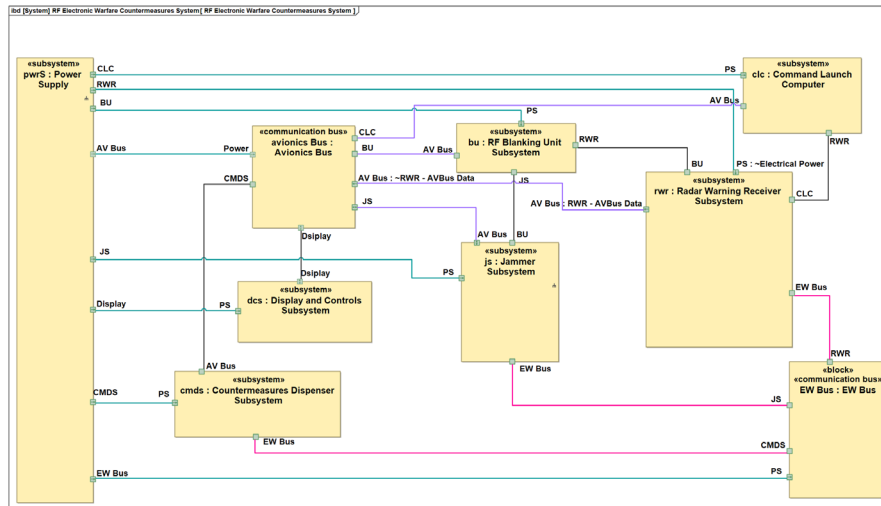


Figure 9. Electronic Countermeasures System Domain View

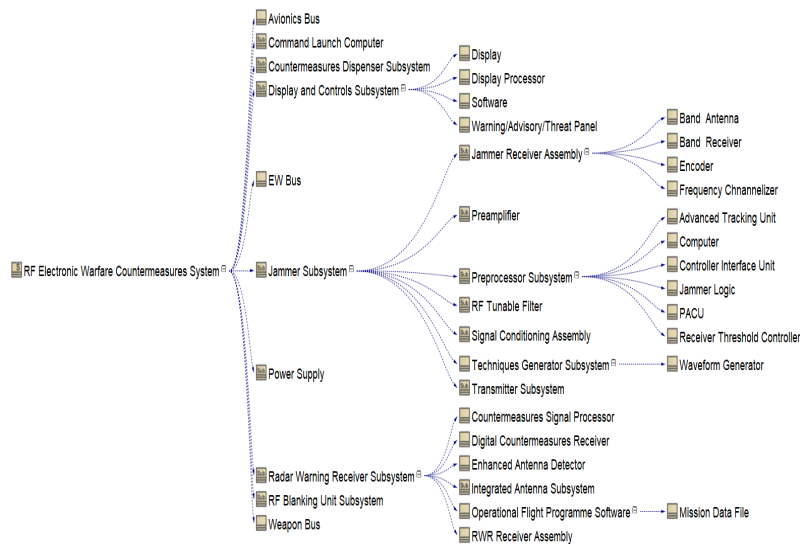




**Develop Logical Architecture Views.** The process of developing the logical architecture for the EW system begins with defining the functional/behavior architectural view, identifying logical subsystems, developing a configuration view, and finally allocating the functions (actions) to logical/structural subsystems. Portrayed in Figures 10 and 11 are the logical configuration view portraying the interconnections between subsystems and structural decomposition (hierarchy) view of the exemplar EW countermeasures system. Additionally, the logical architecture view shown in Figure 12 portrays the allocation of system functions to logical subsystems using swimlanes. The EW system’s functional hierarchy / decomposition view is shown in Figure 13.



**Figure 10. EW System Logical Configuration View**



**Figure 11. EW System Structural Decomposition View**



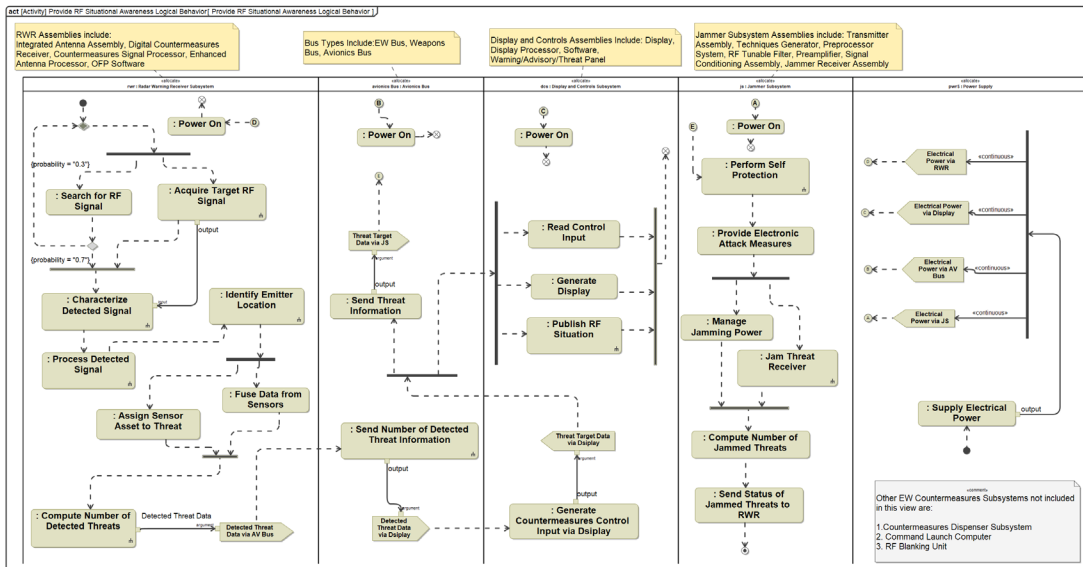


Figure 12. EW System Logical Architecture View

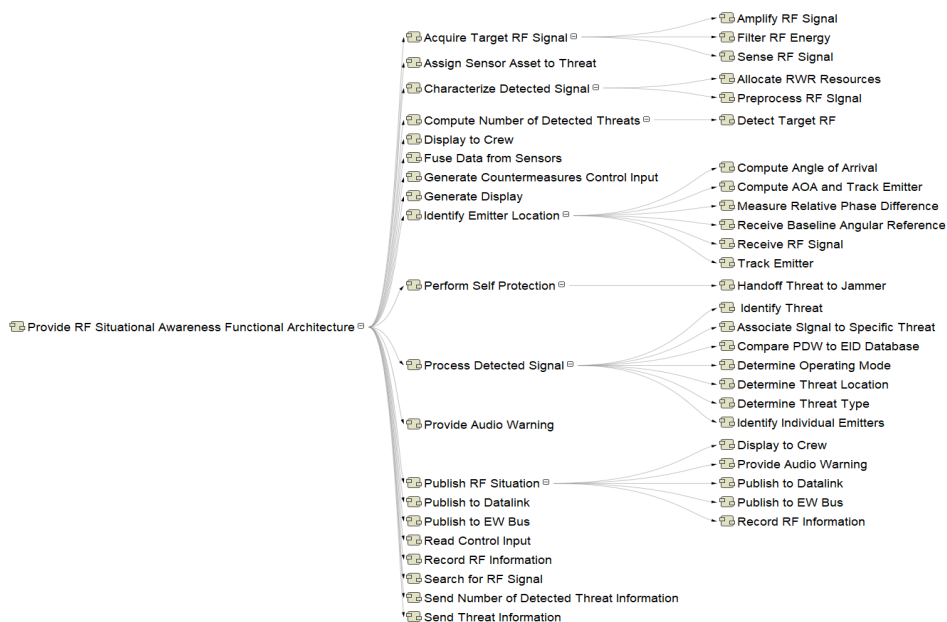


Figure 13. EW System Functional Decomposition View

Furthermore, shown in Figures 14 and 15 are simplifications of the behavior of two EW system functions: *Identify Emitter Location* and *Compute Number of Detected Threats*. They represent key functionality of the EW system needed to *provide situational awareness* to the pilot and other subsystems onboard the aircraft. Identification and modeling of these system capabilities inform the program office and test planners during the early phases of system development of tests that would need to be performed on the system and can enable early testing via simulation of the system model before it is actually built. This approach greatly limits the tendency to develop systems that do not satisfy stated requirements, thereby reducing the overall acquisition cycle time in the event that design rework is required.

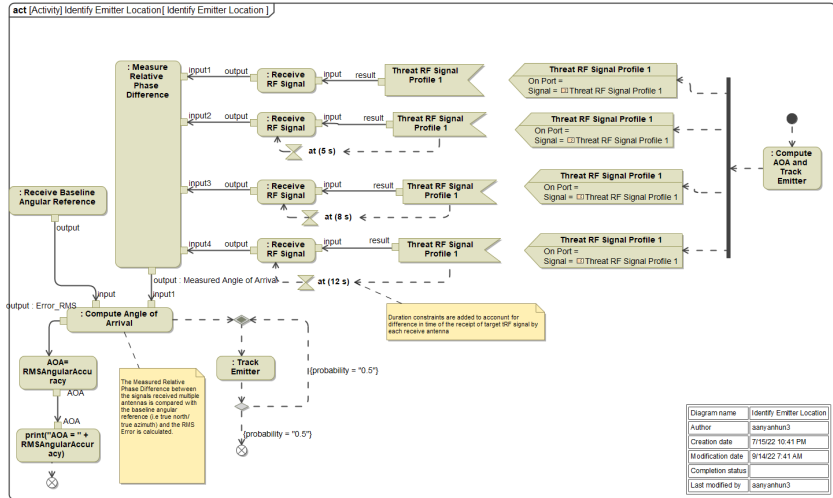


Figure 14. Identify Emitter Behavior View

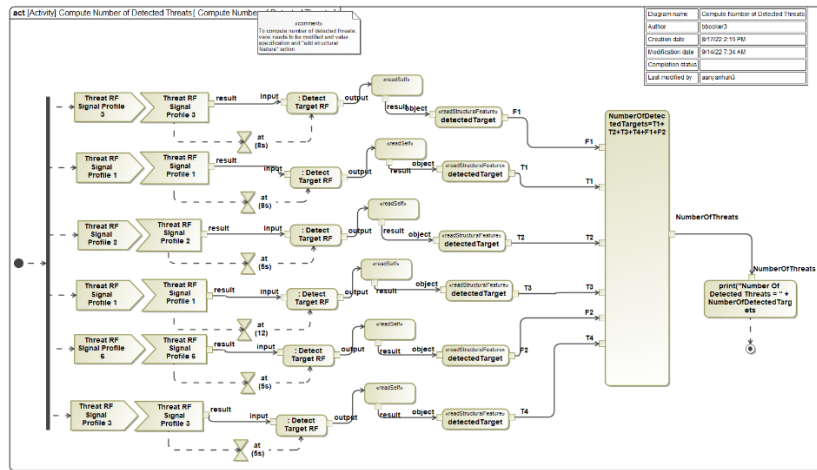


Figure 15. Compute No. of Detected Threats Behavior View

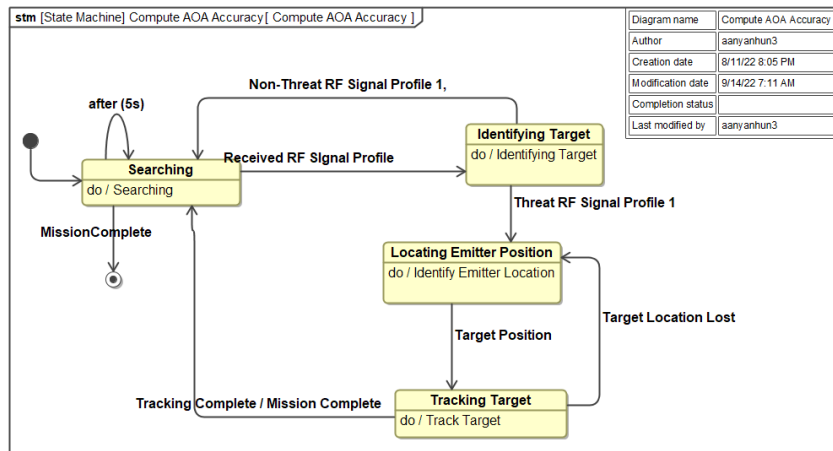


Figure 16. EW System Behavior View



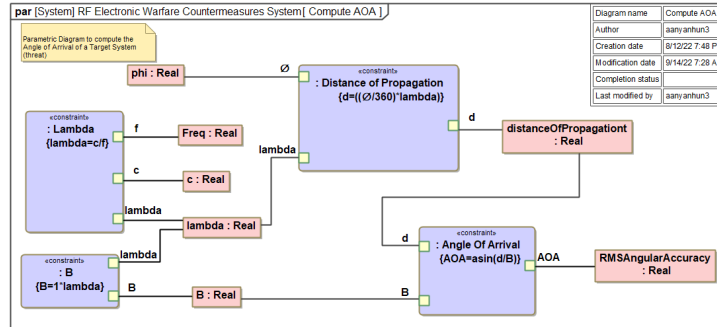


Figure 17. Angle of Arrival Parametric View

In addition, developed within the system architecture model are the views highlighted in Figure 16 which portrays the system behavior using a *state machine* and a system performance analysis view created to compute the angle of arrival (AOA) which is shown in Figure 17. The AOA is a system property that was specified as a measure of effectiveness in the *conceptual* architecture. Development of these system views are critical since they are used during model-based testing activities.

**Curate Key Traceability Views.** During the model-based development of the EW countermeasures system architecture, several model views and artifacts are created which describe the system from multiple perspectives such as structural, behavioral, interfaces, data, etc., and varying levels of fidelity. These views explicitly portray existing relationships between model elements, however, there also exist implicit relationships between some model elements that are not captured in these model views but which are crucial to understanding the system. Identifying and capturing these implicit relationships enable the performance of impact analysis, regression analysis, and promote the understanding of how these relationships impact system behavior and by extension test results. As reported by Konigs et al., “traceability allows changes to be propagated efficiently while implications can be detected easily based on relations between multiple artifacts” (2012). Highlighted in Table 1 are traceability views which portray existing explicit/implicit relationships between model data.

Table 1. EW System Traceability View Mapping Structure to Function

#	Name	Owned Assembly	Function List	Subsystem Interface
1	Display and Controls Subsystem	<ul style="list-style-type: none"> <li>Display Processor</li> <li>Software</li> <li>Display</li> <li>Warning/Advisory/Threat Panel</li> </ul>	<ul style="list-style-type: none"> <li>Read Control Input</li> <li>Generate Display</li> <li>Publish RF Situation</li> <li>Generate Countermeasures Control In</li> <li>Power On</li> <li>Forward Threat Data(context Display e</li> </ul>	<ul style="list-style-type: none"> <li>In PS : ~Electrical Power</li> <li>Inout Display : Display Data</li> </ul>
2	Radar Warning Receiver Subsystem	<ul style="list-style-type: none"> <li>Digital Countermeasures Receive</li> <li>RWR Receiver Assembly</li> <li>Countermeasures Signal Process</li> <li>Enhanced Antenna Detector</li> <li>Integrated Antenna Subsystem</li> <li>Operational Flight Programme So</li> </ul>	<ul style="list-style-type: none"> <li>Search for RF Signal(context Radar W</li> <li>Acquire Target RF Signal</li> <li>Characterize Detected Signal</li> <li>Process Detected Signal</li> <li>Identify Emitter Location</li> <li>Assign Sensor Asset to Threat</li> <li>Fuse Data from Sensors</li> <li>Compute Number of Detected Threats</li> <li>Power On</li> </ul>	<ul style="list-style-type: none"> <li>In PS : ~Electrical Power</li> <li>EW Bus</li> <li>BU</li> <li>CLC</li> <li>out AV Bus : RWR - AVBus Data</li> </ul>
3	Jammer Subsystem	<ul style="list-style-type: none"> <li>Preprocessor Subsystem</li> <li>Transmitter Subsystem</li> <li>Techniques Generator Subsystem</li> <li>Jammer Receiver Assembly</li> <li>Preamplifier</li> <li>RF Tunable Filter</li> <li>Signal Conditioning Assembly</li> </ul>	<ul style="list-style-type: none"> <li>Perform Self Protection</li> <li>Send Status of Jammed Threats to RW</li> <li>Compute Number of Jammed Threats</li> <li>Manage Jamming Power(context Jamme</li> <li>Jam Threat Receiver(context Jammer s</li> <li>Provide Electronic Attack Measures(cc</li> <li>Power On</li> </ul>	<ul style="list-style-type: none"> <li>EW Bus</li> <li>BU</li> <li>In PS : ~Electrical Power</li> <li>Inout AV Bus : ~Jammer - AVBus</li> </ul>



## Develop Test Range Capability Architecture

A test capability model can be described as a model-based representation of all test resources required to enable the testing of a given set of systems/capabilities. The test range infrastructure is a key part of the model-based integrated test prototype and consists of the testing capabilities required to test an EW countermeasures system. In this section, aspects of a notional test range model based on the Eglin test range (Eglin Customer Guide, 2021) will be presented. The test range model captures the system capabilities the test range is capable of testing, test range test resources, its structural composition, and test operational environments. Development of the test range model begins with the identification and definition of the capabilities of the test range.

**Identify Test Range Capabilities.** The first step taken in the development of the model-based test range was to identify the test range capabilities. Test range capabilities in this work refer to the types/kinds/categories/forms of tests a test range is capable of executing. The Eglin Test and Training Complex (ETTC) has a total of 45 test capabilities (Eglin Customer Guide, 2021), some of which are shown in Figure 18. Specifically, the test capabilities of interest for this model prototype shown in the use case diagram highlighted in Figure 19 is the *Perform EW Countermeasures Test* capability and the *Perform RF Source Countermeasures Test*. Once capabilities have been identified and defined as part of the test range model, next steps entail the development of the test range infrastructure's architecture.

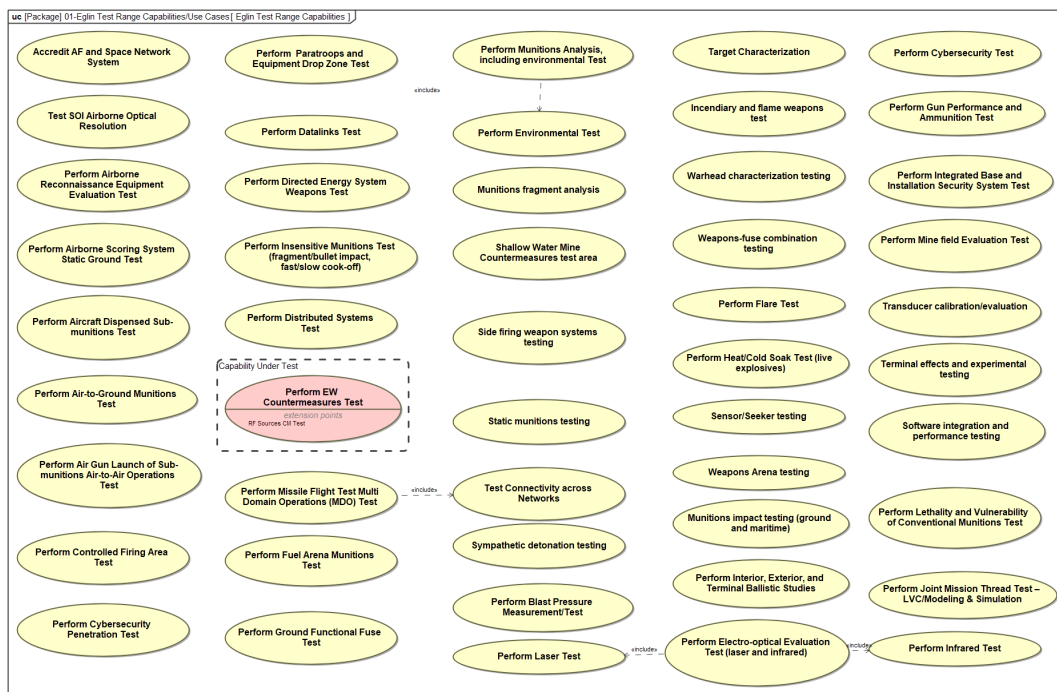


Figure 18. List of Eglin Test Center Test Capabilities

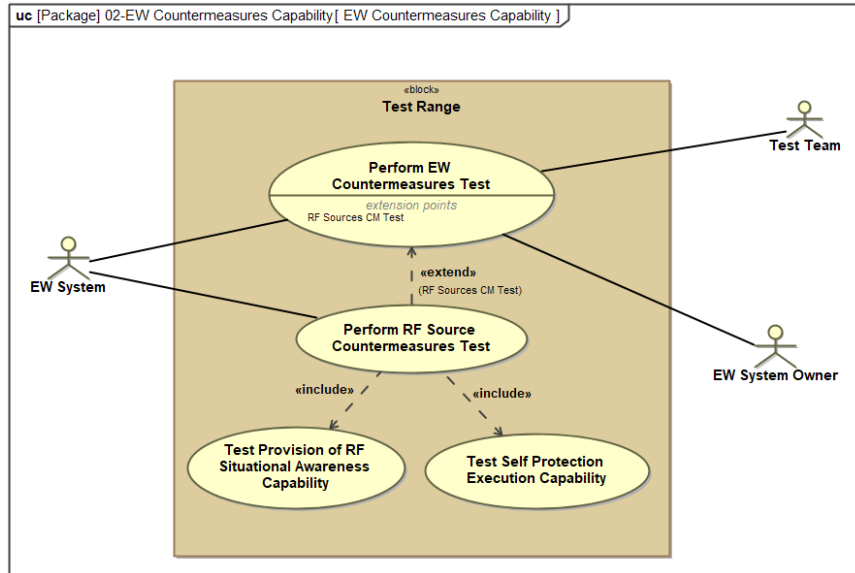


Figure 19. Perform EW Countermeasures Test

**Specify Test Range Infrastructure.** The test range model infrastructure shown in Figure 20 specifies specific aspects of a test range required for performing an EW Countermeasures Test. Range infrastructure include *Test Instrumentation*, *EW Threat Systems*, *EW Non-Threat Systems*, and *Air Threat Defense Systems*. Additionally, the EW threat systems view portrayed in Figure 21 highlights several types of radar threat systems that form part of the test environment configuration while test resources captured in Figure 22 are also used as part of the test configuration. The capture of these test range resources and their properties within the test range model enables the construction of holistic and integrated test case configurations.

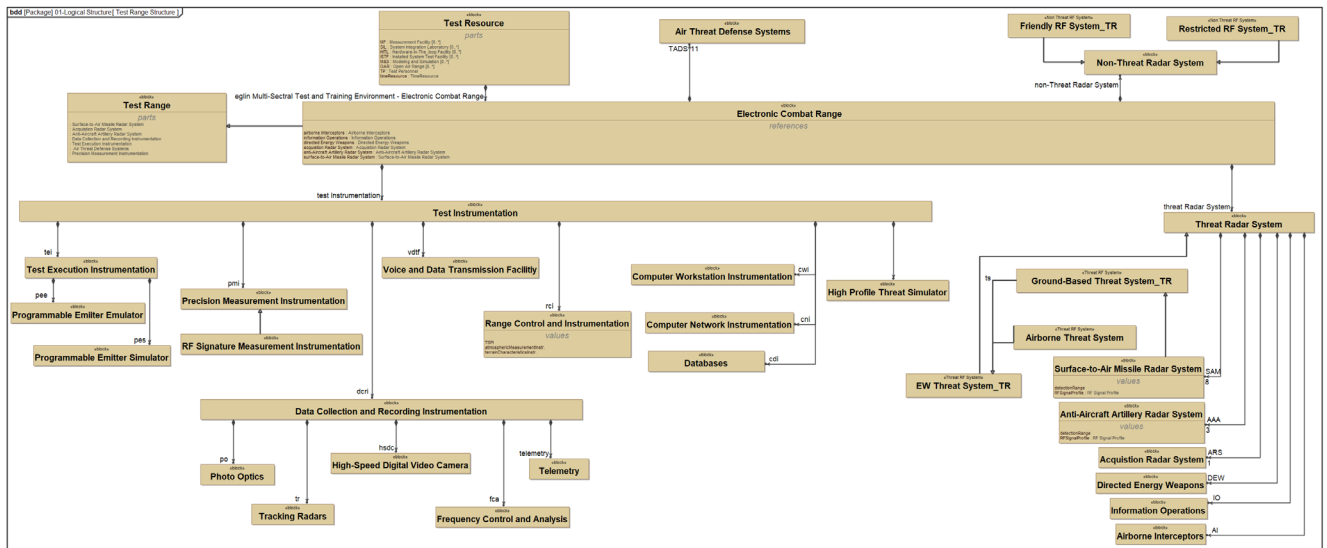


Figure 20. Notional Test Range Infrastructure View

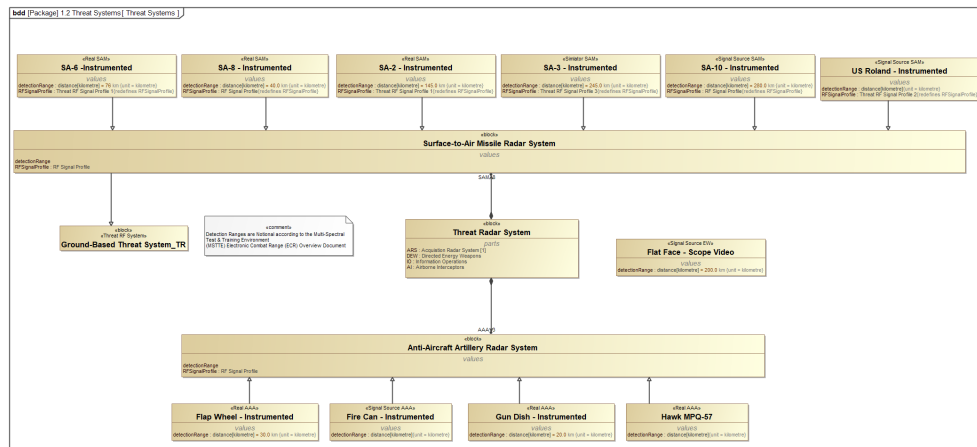


Figure 21. Test Range Threat Radar Systems View

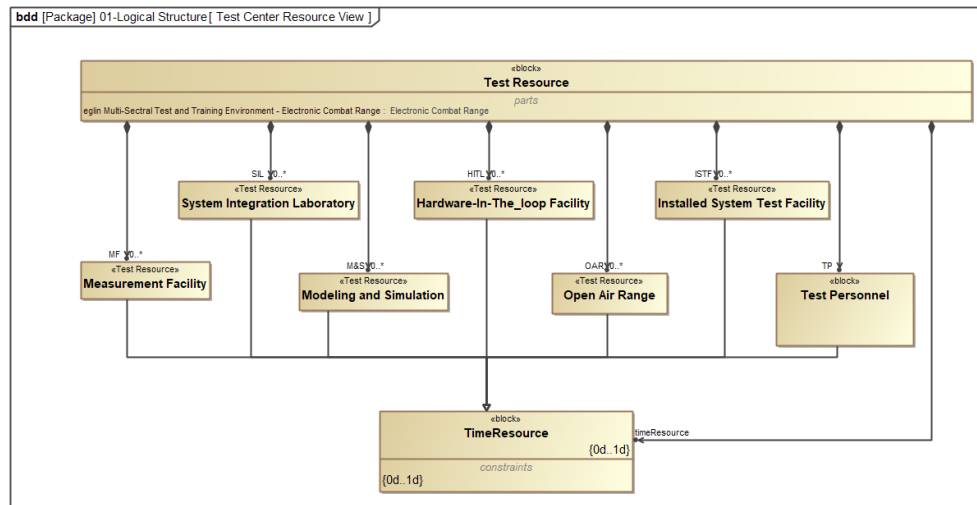


Figure 22. Test Range Test Resources View

Consequently, the capture and definition of test range artifacts in model form plays a crucial role in *test planning* by providing information on available test range resources, in *testing* by enabling model execution of test cases early in the system development life cycle, and in identifying *test risks* relating to test range resource availability.

**Test Range Operational Environment Definition.** A very important aspect of test range model definition involves specifying the various operational environments required for performing specific types of EW system tests. Shown in Figure 23 is a high-level structural view of the Test Range Operational Environment. As shown, the test range operational environment is grouped into two main categories: *Test Range Electromagnetic Operational Environment* (EMOE) and *Test Range Geophysical Environment*. Of importance for this exemplar however, are the EW countermeasures system operational environments, namely, congested environment, contested environment, and constrained environment types respectively. Defining these environments as part of the test range test capability are necessary to enable testing of the EW system to verify that the EW system requirements can be satisfied as well as enable the mapping of risk to the specific system operational environments.

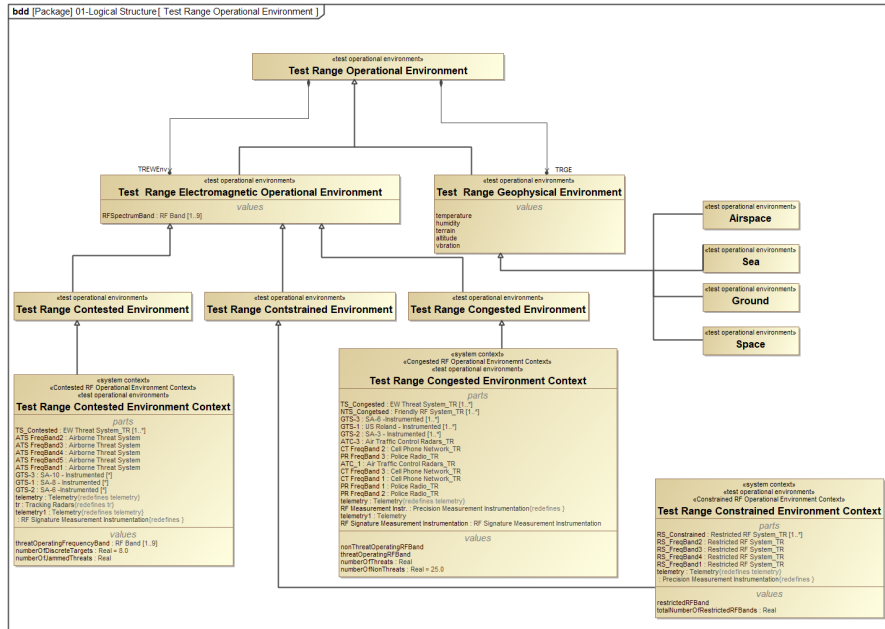


Figure 23. High-Level Test Range Operational Environment View

**Test Range Infrastructure Traceability.** Creating traceability views of test range resources such as the test instrumentation infrastructure, threat systems, and air defense systems serve as crucial model data views which enable test planning and testing activities using the *model-based test-integrated system* prototype. Figure 24 highlights several explicitly and implicitly identified relationships among test range infrastructure artifacts.

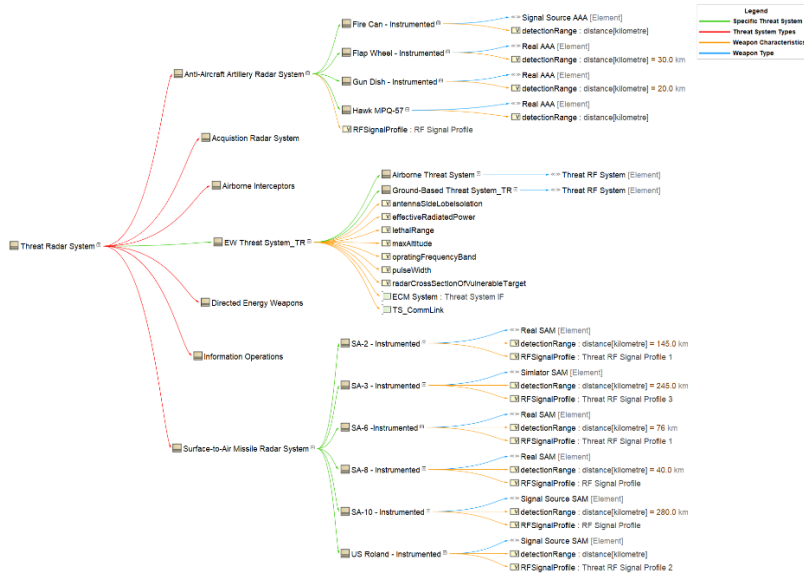


Figure 24. Traceability View of Test Range Threat Radar Systems

## System Under Test (SUT) and Test Case Modeling

A key aspect of the *model-based test-integrated system* prototype is the development and modeling of test cases and test case configurations for the system under test (SUT). In order to perform model-based testing within the MBSE environment, test cases and the test





scenario configuration need to be defined for the SUT. In the context of this work, the test configuration describes the testing context for the SUT and comprises the SUT, test resources, test personnel, test case, and the system requirements that need to be satisfied. The test community and program offices can use this model-based test configuration to inform decision making regarding availability of test range resources and the system requirements that need to be satisfied by the SUT per (mission) test case. The test configuration pattern shown in Figure 25 is an abstract representation of a test context and depicts the components required for test case execution and the relationships between them.

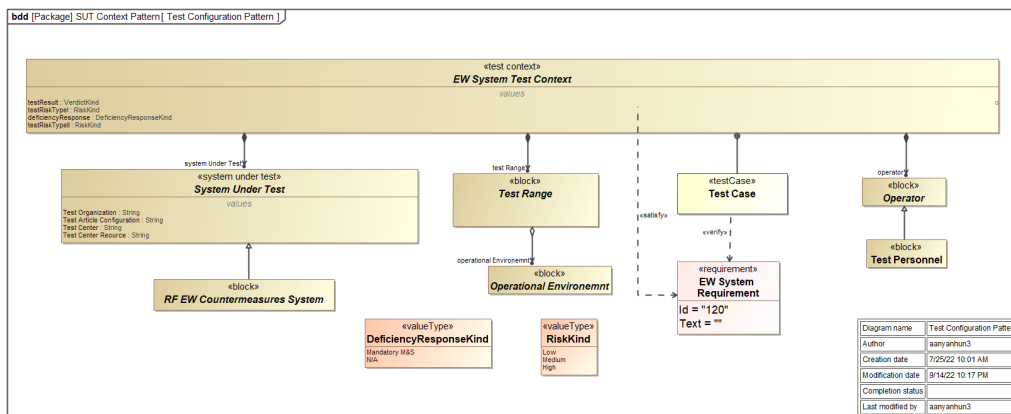


Figure 25. Model-Based Test Configuration Pattern

**Execute Test Case Model and Capture Results.** The model view portrayed in Figure 26 is an example of an implementation of the test configuration pattern shown in Figure 25. The model view shown represents the test context for the EW countermeasures system *Angle of Arrival Test Case*. It can be noted that the operational environment in which the system is being tested is designated as a specific contested operational environment. Test range resources listed as part of the contested environment include multiple threat radar systems, telemetry, and RF signature measurement instrumentation. Also captured as test scenario participants are test personnel, the requirement being tested, the SUT, and the test case artifact. Results gotten from the execution of the AOA test case context are captured in the test instance specification table shown in Table. 2.

The SUT requirements traceability view shown in Figure 27, highlighting implicit relationships that may exist between the artifacts of the integrated test model prototype. Moreover, such traceability views allow planned or unplanned change implications to be quantified and assessed.



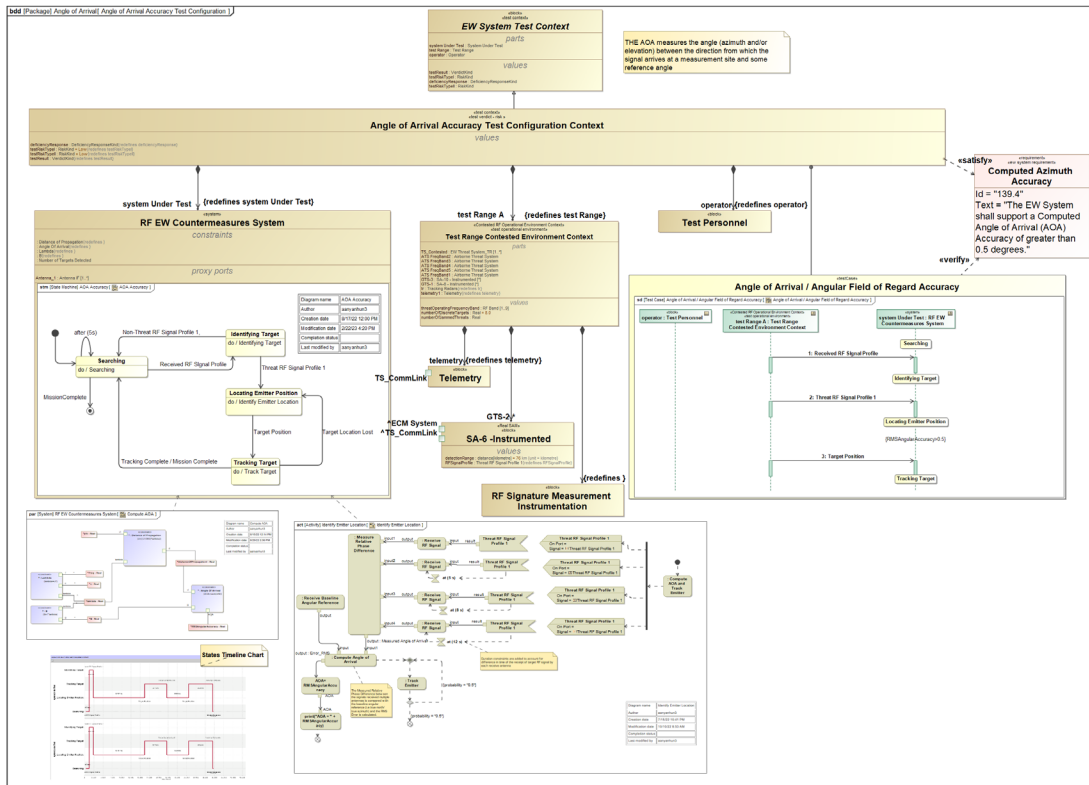


Figure 26. EW System Angle of Arrival Accuracy Testing Context

Table 2. Angle of Arrival Accuracy Test Execution Results for the EW System

#	Name	system Under Test.RMSAngularAccuracy: Real	testResult: VerdictKind	deficiencyResponse: DeficiencyResponseK	testRiskType: RiskKind	testRiskType: RiskKind
1	angle of Arrival Accuracy Test Configuration Context at 2022.10.25 23.13	0.3994	fail	Mandatory M&S	Medium	Low
2	angle of Arrival Accuracy Test Configuration Context at 2022.10.25 23.12	0.9851	pass	Mandatory M&S	Low	Medium
3	angle of Arrival Accuracy Test Configuration Context at 2022.10.26 10.38	0.294	fail	Mandatory M&S	Low	Low
4	angle of Arrival Accuracy Test Configuration Context at 2023.02.22 16.03	0.4251	fail	Mandatory M&S	Low	Low
5	angle of Arrival Accuracy Test Configuration Context at 2023.02.23 09.19	0.6842	pass	Mandatory M&S	Low	Low
6	angle of Arrival Accuracy Test Configuration Context at 2023.02.23 09.21	0.9851	pass	Mandatory M&S	Low	High

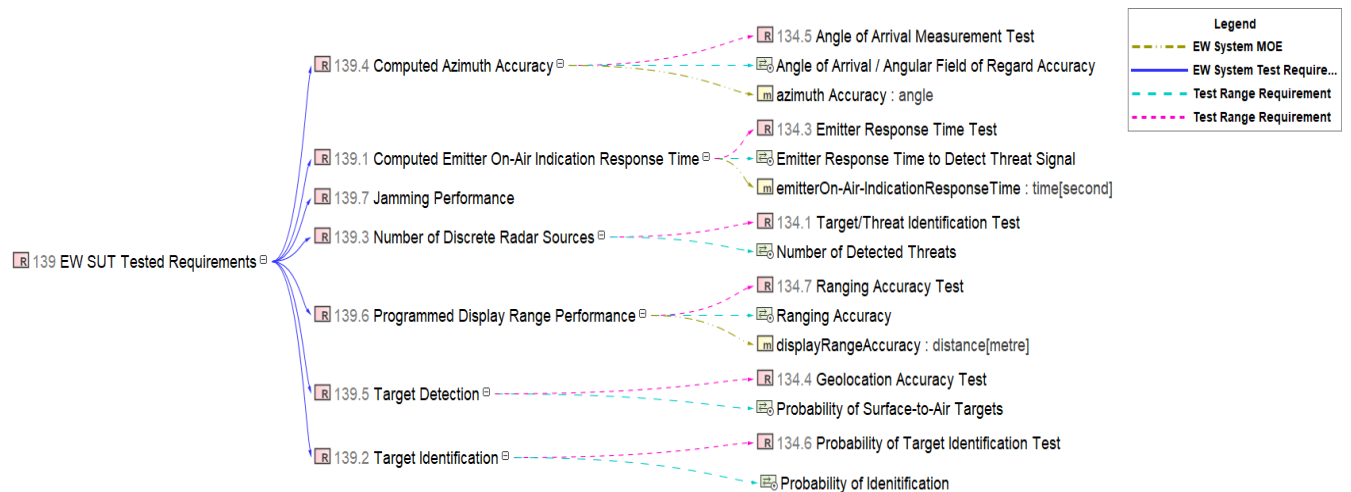


Figure 27. Test-Related Artifacts Traceability View



A key benefit of the model-based test integrated system prototype concept is that it enables the execution and analysis of multiple user defined missions and operational environment test configurations in a minimum amount of time. As a consequence, this testing prototype provides quantifiable value to program offices and the test community through its ability to impact the defense acquisition cycle time positively by enabling program offices to make informed decisions regarding system tests and associated risk in a timely manner.

### **Development and Modeling of the Risk Function**

Program risk in the context of this project is inherently a function of different aspects of the program's ability to create and deliver a useful product to the field. Traditionally risk functions have focused only on the risk to the program of developing the end item product. The traditional risk approach also developed risks based on specific design risks. This approach has several shortcomings. First, this approach does not guarantee a comprehensive coverage of all possible risks. Second, the traditional approach has no specific way of addressing risk created in the testing program independent of the acquisition risk of different parts of the system. Third, the traditional approach to risk does not have a direct means of aggregating risk from the system's operational mission or operational environments. As a result, these three specific areas form the requirements set for the development of a new risk approach.

1. Develop a risk function that is comprehensive across all areas of the program.
2. It is critical that the risk function capture risks that are inherent to the testing of systems.
3. A risk function needs the ability to aggregate risk across different aspects of the program, specifically aggregate risks across mission areas and operation environments.

Additionally, the risk function must be compatible with the system modeling functions.

### **Elements of the Risk Function**

In developing the comprehensive risk function, it was determined assessing risk on each requirement represented a means of assessing risk on all parts of the system in a comprehensive way. Undoubtedly, linking risk to the system requirements addressed several different issues. First, the requirements are modeled as a part of the MBSE process and are therefore part of the integrated engineering model of the system. Second, by linking the risk to the requirements, it is assured that the evaluation of all possible risks in the system is done in a comprehensive way. The requirements, if specified correctly and unambiguously, describe all the different aspects of the system and its operations including its different missions and operating environments.

Therefore, for a given mission and operational environment, a program's risk can be aggregated and evaluated by combining the system requirements specified for the system operations based on a given mission and operational environment. Functionally, this approach allows for different weighting of factors to the individual risks so that the overall risk profile for a given mission can reflect the different priorities of the mission. To effectively model testing risks, a test-based risk function was developed which addresses the risks inherent in the testing infrastructure's ability to completely test system requirements, and the risk of the testing infrastructure's ability to replicate future operational environments during system test. Specifically, the three risk categories defined in the test-based risk function include:

- Type 1 Test Risk: the ability to test to the requirements.
- Type 2 Test Risk: the reliability of the testing to validate the operational environment (confidence in test).
- Implementation Risk: the risk of being able to design and build the system to meet its requirements. This could be viewed as the traditional risk function (cost, schedule, and performance risks).



## Use of the Risk Function to Determine a Mission-Based Risk Profile

Operational environments and missions of interest contribute to the risk of a given system not being able to perform as designed, or as needed during operations. The risk function provides the ability to roll up the risk for the different test cases. In particular, the test risks allow the program office and test community to assess whether the testing resources available at a given test range can effectively test the system requirements to the levels expected in the operational environment for different missions of interest.

Mapping of the risk function characteristics to specific operational environments and user defined missions can be accomplished in one of two ways. The first involves using the system requirements to describe a specific mission. *Given that the complete set of requirements contain all requirements necessary for a given system to perform all required missions under all specified operating conditions, it follows that a subset of the requirements can be selected which describe a specific mission and operating environment.* While the second involves the use of the system's operational testing requirements. *In this method, the testing requirements and test cases for a specific mission are linked to the risk model to capture risks from the mission of interest.* The risks can then be aggregated to form a mission-based risk profile.

A demonstration of the mission-based risk profile developed for the EW countermeasures system and created within the *model-based test-integrated system* prototype is portrayed in Table 3, Table 4, and Table 5. In this example, the mission-based risk profile is created by mapping the set of requirements of the EW countermeasures system needed to perform a user-defined mission within specific contested, congested, and constrained environments. As shown in the congested, contested, and constrained risk function tables, values for the *Likelihood* and *Consequence* of each risk type are entered into the tables following which the value for each risk type is then automatically computed and given the necessary risk color based on the computed value.

**Table 3. Contested Environment Mission-Based Risk Function for the EW Countermeasures System**

Risk Type I: High Low Moderate Risk Type II: High Low Moderate Implementation Risk: High Low Moderate												
#	Id	Name	Text	Ability To Test - Likelihood	Ability To Test - Consequence	Ability To Test - Risk	Confidence In Test - Consequence	Confidence In Test - Likelihood	Confidence In Test - Risk	Implementation - Consequence	Implementation - Likelihood	Implementation - Risk
1	153	Target identification	The EW System shall correctly identify target system not less than 95% of the time with a confidence of or greater than 90%.	5	1	5	5	1	5	1	1	1
2	151	Cluttered EMI Environment	The B-1 band 8 replacement system shall be able to meet its performance requirements in the presence of high levels of commercial EM transitions as modeled by XX simulation.	3	1	3			0			0
3	154	Number of Discrete Radar Sources	The EW System shall be able to detect greater than 8 target systems at the same time.	3	3	9	1	3	3	1	1	1
4	156	Jamming Performance	The EW System shall be able to meet performance requirements in the presents of jamming at the level of X.	3	5	15	1	1	1	1	1	1
5	157	Operate in Contested Environment	The EW system shall accurately detect, track and jam, active threat radars in a contented environment where threat systems are actively trying to defeat the EW system and degrade its capabilities.	3	3	9	5	3	15	3	3	0
6	152	Congested Civilian EMI Environment	The B-1 Band 8 replacement system shall meet all its performance requirement in the present of multiple cell phone networks (4 or more), police radios (15 or more transmitters, Civilian radar systems, (6 or more (ATG, weather, or other radars) operating in the same or adjacent frequencies as the Band 8 replacement system.	1	5	5			0			0



**Table 4. Congested Environment Mission-Based Risk Function for the EW Countermeasures System**

Risk Type I: High Low Moderate		Risk Type II: High Low Moderate		Implementation Risk: High Low Moderate								
#	Id	Name	Text	Ability To Test - Likelihood	Ability To Test - Consequence	Ability To Test Risk	Confidence In Test - Consequence	Confidence In Test - Likelihood	Confidence In Test Risk	Implementation - Consequence	Implementation - Likelihood	Implementation Risk
1	140	Jamming Performance	The EW System shall be able to meet performance requirements in the presents of jamming at the level of X.	3	3	9	1	1	1	1	1	1
2	141	Number of Discrete Radar Sources	The EW System shall be able to detect greater than 8 target systems at the same time.	1	3	3	1	3	3	1	1	1
3	143	Target Detection	The EW System shall detect surface-to-air radar targets correctly 99% of the time at- a 90% confidence.	1	1	1	3	1	3	1	1	1
4	142	Target Identification	The EW System shall correctly identify target system not less than 95% of the time with a confidence of or greater than 90%.	3	5	15	5	1	5	1	1	1
5	144	Congested EMI Environment	The EW System shall meet all of the performance requirement in the presence of a high level of red and blue force EM system operating in close proximity of the systems location.	1	3	3	3	3	9	1	1	1

**Table 5. Constrained Environment Mission-Based Risk Function for the EW Countermeasures System**

Risk Type I: High Low Moderate		Risk Type II: High Low Moderate		Implementation Risk: High Low Moderate								
#	Id	Name	Text	Ability To Test - Consequence	Ability To Test - Likelihood	Ability To Test Risk	Confidence In Test - Consequence	Confidence In Test - Likelihood	Confidence In Test Risk	Implementation - Consequence	Implementation - Likelihood	Implementation Risk
1	145	Congested Civilian EMI Environment	The B-1 Band 8 replacement system shall meet all its performance requirement in the present of multiple cell phone networks (4 or more), police radios (15 or more transmitters, Civilian radar systems, (6 or more (ATC, weather, or other radars) operating in the same or adjacent frequencies as the Band 8 replacement system.	1	1	1	1	1	1	1	1	1
2	146	Constrained EMI Environment - Max Power	The B-1 Band 8 replacement system shall meet all its performance requirements while operating a maximum of 50% normal output power.	3	3	9	1	1	9	1	1	1
3	147	Target Identification	The EW System shall correctly identify target system not less than 95% of the time with a confidence of or greater than 90%.	1	3	3	5	1	5	1	1	1
4	150	Jamming Performance	The EW System shall be able to meet performance requirements in the presents of jamming at the level of X.	1	1	1	1	1	1	1	1	1
5	149	Target Detection	The EW System shall detect surface-to-air radar targets correctly 99% of the time at- a 90% confidence.	3	1	3	3	3	9	1	1	1
6	148	Number of Discrete Radar Sources	The EW System shall be able to detect greater than 8 target systems at the same time.	5	3	15	3	5	15	1	1	1

**Discussion**

As many parts of the DoD are moving to SysML-based MBSE and digital engineering to manage their programs, there is significant opportunity to leverage the power of these tools for test and evaluation. The test program for any DoD system is vital to ensure the future performance of the system. However, testing can be very costly and time consuming on projects and may not produce high confidence. Modeling will increase the program and test organizations ability to more effectively plan and manage the test program and to ensure that all data collected on systems during contract or test, developmental test, and operational test is captured and used to its best advantage. The current system level risk approach does not adequately capture test risk or how changes to the test program and the requirements will impact overall system risk. More robust risk analysis will positively impact test planning and acquisition outcomes.

This work has demonstrated that test risk can be effectively modeled within a MBSE model and directly related to requirements and the design of the system. In addition, this work has proposed a risk function that addresses the DoD’s need for a risk function that can be focused on modeling directly as a function of mission profile and an operating environment. The development of integrated system modeling to include the full acquisition life cycle, particularly the testing of systems, will be a major advancement in the development of the practice of model-based systems engineering and is critical to the use of MBSE in the acquisition community going forward. Results of this work demonstrate the ability to directly link the program requirements and design directly to the ability to test and test planning and develop risk functions dependent on both the system and the ability to effectively test the system.

In order to get the maximum benefits for the use of MBSE in the development of systems for the DoD we investigated the creation of an advanced risk function to include traditional risk functions (cost, schedule and performance, likelihood, and consequence) as well



as linking risk to testing and requirements. In addition, the model-based risk functions were designed as a function of requirements in order to allow for defining specific missions (based on a set of requirements) and looking at the risks as a function of the mission and operation profile (environment and threats) for that defined mission.

## Conclusion

The model-based test risk function is a new development that will give the program offices and test organizations better visibility into the critical aspects of program performance during the development and testing life cycle of the program. By expanding the use of model-based systems engineering and digital engineering to include more of the program life cycle, the DoD can gain better visibility into the management of these programs. The use of these digital models also provides the means necessary to better look across portfolios of developmental programs and existing systems for portfolio management, mission and threat analysis, and long-term campaign planning.

The Expansion of MBSE and DE in DoD acquisition to fully include the different aspects of T&E and risk management creates several significant advantages in managing programs and portfolios. Greater knowledge of risk and the data needed to inform decision making all along the acquisition life cycle will allow for the acceleration of DoD programs in a manner consistent with reasonable risk taking and data driven decision making that will result in more rapid fielding the highly capable systems.

## References

- Dick, J., Hull, E., Jackson, K., Dick, J., Hull, E., & Jackson, K. (2017). Requirements engineering in the problem domain. *Requirements Engineering*, 113–134.
- DoD. (2018). *Digital Engineering Strategy*. Office of the Deputy Assistant Secretary of Defense for Systems Engineering.
- DOT&E. (2022). *Office of the Director, Operational Test and Evaluation strategy update—Strategic pillars. FINAL DOTE 2022 Strategy Update 20220613.pdf (osd.mil)*
- History.com Editors. (2009). *President Eisenhower warns of Military-Industrial Complex*. History.com, A&E Television Networks. <https://www.history.com/this-day-in-history/eisenhower-warns-of-military-industrial-complex>
- Königs, S. F., Beier, G., Figge, A., & Stark, R. (2012). Traceability in systems engineering—Review of industrial practices, state-of-the-art technologies and new research solutions. *Advanced Engineering Informatics*, 26(4), 924–940.
- Range Operations & Sustainment 96T/XPO. (2021). *96<sup>th</sup> Test Wing Customer Guide*. <https://www.eglin.af.mil/Portals/56/documents/Customer%20Guide%202021.pdf>



# Using Digital Twins to Tame the Testing of AI/ML Systems

**David G. Zurn**—received his Bachelor and Master of Science degrees in Electrical Engineering from the Georgia Institute of Technology in 1985 and 1990 respectively. Since joining GTRI's Electronic Systems Laboratory (ELSYS) in 2003, he has worked on a variety of EW-related research efforts including Radar Warning Receiver hardware and software development and test, Missile Warning System hardware and software test, and development of Hardware in the Loop (HITL) test solutions tailored to EW applications. Zurn is currently serving as the Division Chief of the Test Engineering Division within ELSYS. Zurn is a lecturer for the RWR Design short course offered through GT's Professional Education program. His recent research interest areas are Cognitive EW T&E and Space EW T&E. [David.zurn@gtri.gatech.edu]

**Dr. Craig Arndt**—currently serves as a principal research engineer on the research faculty of the George Tech Research Institute (GTRI) in the System Engineering Research division of the Electronic Systems Lab. Dr. Arndt is a licensed Professional Engineer (PE), and has over 40 years of professional engineering and leadership experience. 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. He 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. In industry he has been an engineering manager, director, vice president, and CTO of several major defense companies. He is also a retired naval officer. [Craig.Arndt@gtri.gatech.edu]

**Jeremy Werner**—Ph.D., ST was appointed DOT&E's Chief Scientist in December 2021 after initially starting at DOT&E as an Action Officer for Naval Warfare in August 2021. Before then, Jeremy was at Johns Hopkins University Applied Physics Laboratory (JHU/APL), where he founded a data science-oriented military operations research team that transformed the analytics of an ongoing military mission. Jeremy previously served as a Research Staff Member at the Institute for Defense Analyses where he supported DOT&E in the rigorous assessment of a variety of systems/platforms. Jeremy received a PhD in physics from Princeton University where he was an integral contributor to the Compact Muon Solenoid collaboration in the experimental discovery of the Higgs boson at the Large Hadron Collider at CERN, the European Organization for Nuclear Research in Geneva, Switzerland. Jeremy is a native Californian and received a bachelor's degree in physics from the University of California, Los Angeles, where he was the recipient of the E. Lee Kinsey Prize (most outstanding graduating senior in physics). [jeremy.s.werner.civ@mail.mil]

## Abstract

Program test managers and test engineers should carefully consider Digital Twinning approaches for addressing training and testing challenges for Artificial Intelligence/Machine Learning (AI/ML) systems. A hybrid Hardware in the Loop (HITL) and Digital Twin (DT) architecture is discussed for a notional Cognitive EW system. This architecture may provide effective training and testing for complex AI/ML systems that incorporate extensive Cyber-Physical interactions. Considerations for generating realistic RF test environments for Cognitive EW systems are also considered.

**Keywords:** Digital Twin, AI/ML, Cognitive EW, HITL

## Executive Summary

This research investigates the challenges associated with testing and training of AI/ML systems in the Electronic Warfare (EW) domain and how these challenges can be addressed using Digital Twins. The specific AI/ML testing and training challenges were identified during a Cognitive EW T&E working group conducted by GTRI while under contract to DOT&E. Several key DT capabilities are identified for addressing AI/ML training and testing challenges –

1. Simulation of the system and its operational environment with sufficient realism
2. Ability of the DT to create training and testing data



3. Ability to efficiently virtualize hardware models, system firmware, and software components into the Digital Twin, allowing for efficient Continuous Integration/Continuous Delivery (CI/CD)

To better understand whether a DT can provide these capabilities, a specific detailed Cognitive EW receiver use case is developed. A high-level hybrid HITL DT architecture for this use case is discussed along with specific functional use cases, such as training and testing data set generation and validation, AI/ML component training and DT validation. Using lessons learned from the Cognitive EW Receiver use case, considerations and limitations for using DT for the Cognitive EW Receiver are discussed.

## Background

Weapons systems augmented with Artificial Intelligence/Machine Learning (AI/ML) capabilities are a new reality and driven by several trends. The modern battlefield is becoming dependent on connected kill-webs and the Joint All Domain Operations (JADO) environment, which is driving the emergence of AI/ML weapons systems on the Blue and Red side (NASEM, 2021). Indeed, strategic competitors, such as China and Russia, are making significant investments in AI for national security purposes (GAO, 2022a). The rapid explosion of AI/ML in the commercial sector is also enabling the adoption of AI/ML in weapons systems (USAF Chief of Staff, 2020).

According to the GAO, AI/ML is expected to transform all sectors of society, including, according to the Department of Defense (DoD), the very character of war. The failure to adopt and effectively integrate AI technology could hinder national security. As a result, the DoD is investing billions of dollars and making organizational changes to integrate AI into their warfighting plans. A total of almost 700 separate AI/ML programs were identified across the services either funded through R&D or procurement. This does not include classified programs or programs funded through O&M, which would inflate that total (GAO, 2022b). According to a recent National Defense Strategy, “The Department will invest broadly in military application of autonomy, artificial intelligence, and machine learning, including rapid application of commercial breakthroughs, to gain competitive military advantages” (DOD, 2018).

Historically, one of the more significant areas of DOD investment in AI/ML has been in the EW domain. GTRI has been involved in multiple efforts to develop, evaluate, and implement AI/ML algorithms on multiple RF EW systems. EW systems sample the RF environment and benefit from AI/ML capabilities designed to infer the behavior and intent of threat Radar waveforms in adversarial conditions. The remainder of this paper will consider AI/ML efforts specifically in that arena.

## AI/ML EW T&E Challenges

GTRI, under contract to DOT&E Test and Evaluation Threat Resource Activity (TETRA), conducted a five session Cognitive EW T&E Working group in 2020–2021 to explore AI/ML T&E challenges for Cognitive EW systems. A variety of stakeholders from the AI/ML research community, the DOD T&E community, and acquisition and sustainment community gathered to identify Cognitive EW T&E challenges, gaps, and potential solutions. The working group findings relating to T&E challenges are summarized in Figure 1.

AI/ML systems present a unique set of test challenges. The massive coverage space and wide range of potential behaviors are difficult to address via legacy test methods. Major AI/ML T&E challenges are summarized as follows.

- A. **Massive Coverage Space** - Extensive analysis has been done for the autonomous driving use case, specifically looking at testing for AI/ML techniques such as Deep





Neural Nets (DNN). For complex systems these DNNs can be very high order non-linear functions. Common test issues arising from these functions are massive, multi-dimensional input–output coverage spaces. This creates issues such as how to optimize/efficiently explore these spaces during test, how to efficiently create test data, and whether it is possible to create a test oracle to determine whether the test has passed or failed (Tian, 2018).

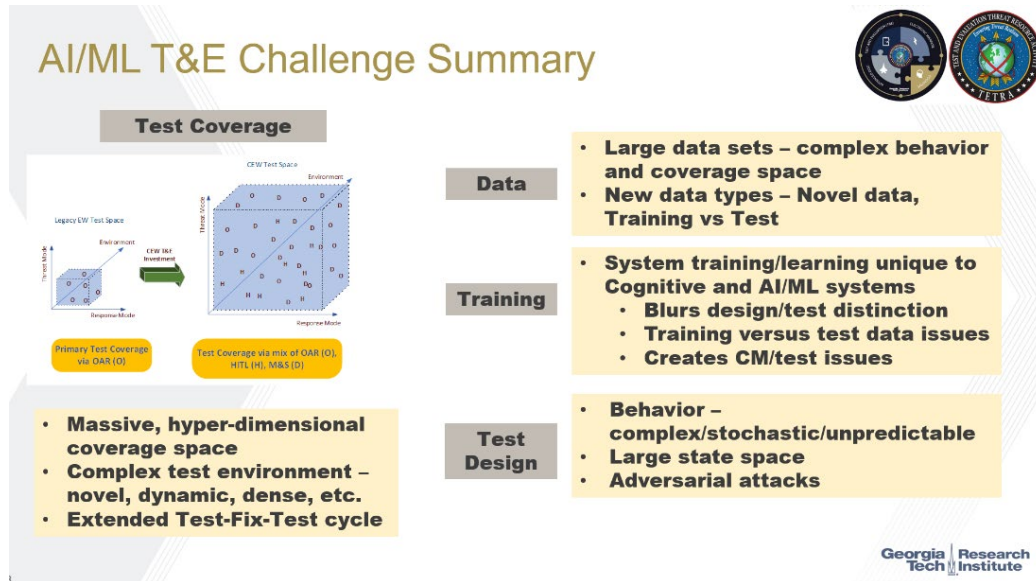


Figure 36 - AI/ML T&E Challenge Summary

- B. **Unique quality parameters** - AI/ML systems, particularly AI/ML software, present new quality parameters or measures of performance associated with learning such as correctness, accuracy, explainability, system stability, timeliness, and robustness that are not typically considered. Rigorous definitions and processes relevant to DoD test systems need to be developed to address these new parameters (Chuanqi Tao, 2019).
- C. **Adversarial exploits** - AI/ML systems require adversarial testing approaches to ensure that during operations adversarial manipulation of data does not affect the system in unpredictable ways (Prokhorov, 2019). Extensive research is underway to generate adversarial exploit data for improved system testing (Anthony Ortiz, 2018).
- D. **Assurance case testing** - For AI/ML enabled autonomous systems, testing to assure safe operation can become an issue. According to a RAND study, the current state of T&E for AI technologies cannot ensure the performance and safety of AI systems, especially those that are safety-critical. Assurance case testing is required for these types of systems (RAND, 2021). An assurance case is “a structured argument that the system is sufficiently dependable to permit fielding in a specific operational context” (Tate, 2019). AI/ML systems exhibiting autonomy require assurance cases because they are unpredictable due to the following attributes:
1. State space explosion
  2. Non-smooth or fractal response
  3. Lack of transparency
  4. Changing system behaviors over time
  5. Emergent behaviors

- E. **Continuous test** - Another unique challenge relates to the fact that AI/ML systems require train-test cycles throughout the system life cycle. This is due to the need to continuously train AI/ML components to cope with environmental and threat changes. Indeed, this is a feature - the system can learn from changes in the environment, but learning must be followed by testing as part of a continuous cycle. These cycles are short in duration and potentially continue through the fielded life cycle of the system. The legacy waterfall and distinct separation of coding/testing/fielding phases are not adequate for AI/ML systems. According to the Defense Science Board (DSB) Summer Autonomy study, to address the train-test cycle challenge, the DoD should look to commercial practices like Agile for developing autonomous, AI/ML based systems. Agile and DevSecOps development practices provide an incremental development approach enabling tight train-test cycles (DSB, 2016).
- F. **Data Generation** - Data generation for training and testing of AI/ML algorithms presents a significant T&E challenge. It's been estimated that 80% of the effort required to implement AI/ML systems is involved in data generation, tagging, and curation (Antonio Nieto-Rodriguez, 2023). The difficulty of procuring data depends on the AI/ML application area. For EW-related AI/ML applications, which this paper addresses, data is a significant challenge: collected and recorded raw high-fidelity data is often not tagged and cannot always be correlated with Blue (U.S.), Red (Adversary), and Gray (Commercial) RF sources. Synthetic data can be generated but replicating real-world environmental and propagation effects can be difficult.

These T&E challenges are exacerbated for AI/ML systems involving extensive interaction with the physical world (Autonomous vehicles, Industrial systems, RF systems). The Cyber–Physical interaction via sensors and effectors and the system interaction with the environment are often difficult or impractical to create in the real-world for test purposes. Testing in a real-world operational environment is ideal from a fidelity perspective, but testers face significant challenges generating sufficiently wide test coverage, creating edge cases and assuring the repeatability of complex test scenarios. Synthetic digital environments and DTs are often created to mitigate these challenges.

## Digital Twin Overview

According to the Digital Twin Consortium, “A Digital Twin is a virtual representation of a real-world system. A digital twin is synchronized with the physical twin at a specific fidelity and frequency” (Digital Twin Consortium, 2020). The National Institute of Standards and Technology (NIST) definition is “A digital twin is the electronic representation—the digital representation—of a real-world entity, concept, or notion, either physical or perceived’ (NIST, 2021). The application and usage of the DT concept varies widely across commercial industry and the DoD. A DOT&E memo assessing the usage of DT in DoD testing shows some progress in the adoption of DT, but it also sharply illustrates how far the DoD has to go:

- Approximately 7% of programs under DOT&E oversight have built or are planning to build a DT.
- Most of the programs that report usage of DTs are applying them for contractor-level testing in support of Engineering Manufacturing Development (EMD) and none have been used DT for operational testing (DOT&E, 2022).

The DoD recognizes the need to accelerate the adoption of DTs. The increasing use of AI/ML introduces “never-before-seen capabilities and vulnerabilities that change at never-before-seen dynamic rates.” The DOT&E 2022 strategy defines five strategic pillars to transform T&E, two of which support the use of DTs for testing AI systems – **Accelerate the delivery of**



**weapons** by embracing digital technologies as a key action and **Pioneer the T&E of weapons systems built to change over time** where enabling adequate assessment of AI-enabled weapons systems is one of the desired end states (Sandra Hobson, 2022).

Researchers in the Advanced Driver Assistance Systems (ADASs), Autonomous Vehicles (AVs), and other industries are taking up the usage of DTs and have explored the use of Digital Twins to address AI/ML training and test challenges. Recognizing the challenges of the complex Cyber-Physical interactions involved in these systems, the use of Hybrid DT systems has been considered (Jörn Thieling, 2021; Kirill Semenov, 2020).

In the AI/ML training and testing context, a Hybrid DT might consist of (1) a real hardware/software system design instantiated in a HITL testbed, (2) a set of digital models and virtualized firmware and software representing that system and the system’s operating environment and (3) a method for validating the digital model versus observed system behavior.

The Hybrid DT concept may be able to address some of the difficult AI/ML training and test challenges such as Massive Coverage Space, Continuous Test, and Data challenges outlined above. The following DT capabilities are required to address these challenges:

1. Simulation of the system and its operational environment with enough realism to support AI/ML training and testing to assure performance as expected in a real operational environment
2. Ability of the Hybrid DT operational environment simulation to create trusted training and testing data suitable for the system’s AI/ML components
3. Ability to efficiently virtualize system digital models, system firmware, and software components into the Digital Twin, allowing for efficient Continuous Integration/Continuous Delivery (CI/CD)

Next, we’ll explore a specific Cognitive EW Receiver use case to evaluate the applicability of the Hybrid DT approach.

### AI/ML Training and Test Use case - Cognitive EW Receiver

First, consider the notional EW receiver in Figure 2. The receiver system is segmented into RF input, receiver system, and Pilot Vehicle and Federated systems interfaces.

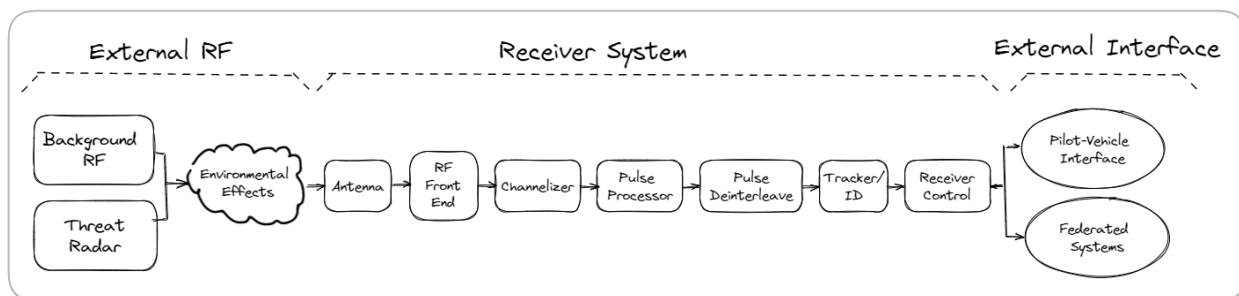
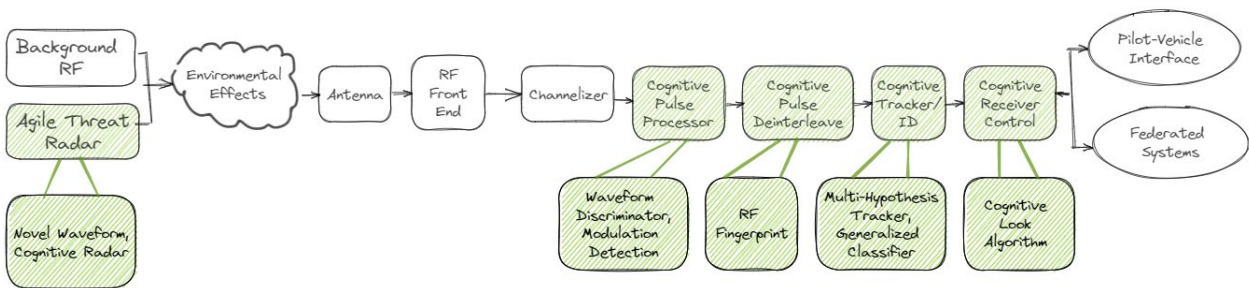


Figure 37 – Notional EW Receiver

The notional EW receiver typically processes and identifies threats using fixed lookup tables and relatively simple signal processing algorithms. This design performs well in simple threat environments where the number of threats is small and the threats produce known, predictable RF waveforms. As the number of threats increase and produce unexpected, unknown RF waveforms, the receiver performance degrades. To counter this problem, receiver designers add AI/ML algorithms to key processing components to improve their overall

performance. A notional Cognitive EW Receiver with some of these cognitive components highlighted is shown in Figure 3.

For example, the Pulse Processor Component can be supplemented with a DNN based waveform discriminator. Traditional waveform discriminators measure waveform parameters such as frequency and phase modulation, then determine waveform type using a look-up table or simple heuristic. If these waveform properties are modified by the threat radar in a way that cannot be measured accurately, or measurements fall outside of the bounds of the lookup table, the traditional discriminator will not perform well. The DNN based waveform discriminator performs similarly to a DNN used for image recognition. The DNN ingests waveforms of different modulation types and attempts classification based on observable and latent waveform features. During training, DNN weights are iteratively adjusted to minimize classification error. The DNN can potentially outperform the traditional discriminator because the DNN extends beyond general classification and is able to handle waveforms with parameters that may not match pre-programmed receiver boundaries/features.



**Figure 38 – Notional Cognitive Receiver**

Typical receiver components and their AI/ML enhancements are listed in Table 1. This is a notional list—receiver designers may create many other AI/ML enhancements depending on receiver requirements.

These AI/ML components have great potential for increasing performance but come at a price. As discussed above, each component requires training and adds an extra test burden during system development and sustainment.

**Table 11 – Cognitive EW Receiver Components**

<b>EW Receiver Component</b>	<b>AI/ML Enhancements</b>	<b>Comment</b>
Pulse Processor	Waveform Discriminator	DNN classification based on observable and latent waveform features
Pulse Deinterleaver	RF Fingerprinting	Deinterleave pulse trains using RF features unique to a given RF threat
Tracker/ID	Multi-Hypothesis Tracker (MHT), Generalized classifier	MHT uses Bayesian inference to more accurately establish and maintain tracks; generalized classifier uses DNN or mode-intent algorithms instead of lookup table to identify threat based on class
Receiver Control	Cognitive Look Algorithm	Receiver uses inference engine to optimize receiver frequency look schedule in a dense threat environment

### **Cognitive Receiver Digital Twin**

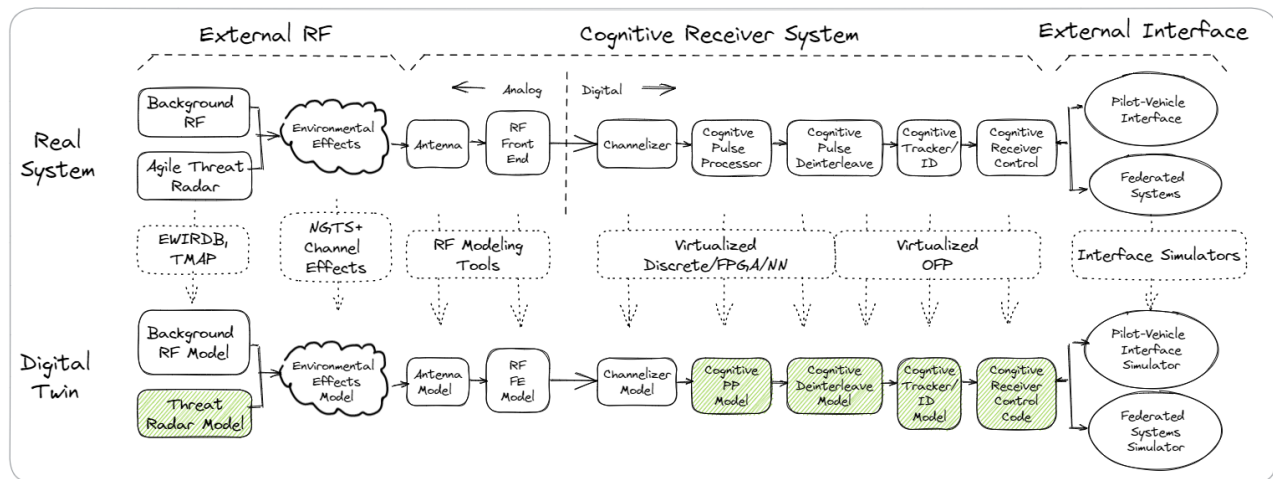
A Cognitive Receiver DT may be able to address these training and test issues. The DT provides a framework for training and testing individual AI/ML components efficiently. Without the DT, each component must be trained in a stand-alone hardware instantiation or in-situ in the receiver system. This may not seem like a problem, as the receiver developer typically implements stand-alone subcomponents for unit test. However, this is typically done only once during system development. AI/ML training is a continuous activity that needs to be done many times throughout the system’s life cycle. It is required during systems development, integration testing, developmental testing, operational testing and system sustainment. It is not practical to create and maintain stand-alone component training setups like this for the entire system life cycle. In-situ training is also impractical. Training requires the introduction of a very large set of inputs to the AI/ML component and adjustment based on component output. Generating this set of inputs through the entire system processing chain is difficult and time-consuming. Moreover, training using real hardware either stand-alone or in-situ can only be done at real-time system operation speed which could be very time-consuming for large datasets. Frequently AI/ML systems are virtualized to enable Faster than Real-Time (FTRT) training.

The DT depicted in Figure 4 is implemented by digitally instantiating each system component using either digitally hosted hardware models or through virtualization of firmware and software. Note that the DT incorporates the complete Cognitive Receiver System and External RF and External Interface elements. Considerations with the Cognitive Receiver System visualization will be discussed, followed by External elements.

The Cognitive Receiver system consists of the antenna, RF front end, the chain of processing elements and the Receiver control block. The antenna and RF front end are modeled using RF modeling tools. Depending on complexity the antenna could be an engineering model based on frequency and polarization dependent azimuth and elevation lookup tables. The RF front end is more problematic as it typically consists of a chain of complex linear and non-linear RF components – limiters, amplifiers, filters and mixers and A/D converters, that can be difficult to accurately model. These components must be accurately modeled to create a useful DT. Crude engineering-based models will not re-create the RF front end effects found in a real receiver system. These effects, such as noise, harmonics, distortion,



ringing and filtering all impact overall receiver performance. If they aren't modeled with sufficient fidelity, the DT may not accurately predict real performance. The digital subcomponents are more straightforward. The discrete logic and Field Programmable Gate Array (FPGA) firmware can be more accurately virtualized in a digital environment. The Operational Flight Program (OFP) can be rehosted on a virtual processor. Salient challenges in firmware and OFP implementation include synchronizing multiple clock domains, replicating propagation delays and accurately virtualizing embedded processors, that need to be addressed however.



**Figure 39 - Cognitive EW Receiver Digital Twin**

The External RF element, though not part of the Cognitive receiver system, is critical to implementation of the DT. An accurate replication of the RF environment must be generated to feed the receiver model with realistic inputs. Threat Radar models and Background RF models create realistic waveforms that are modified by an RF Environmental effects model that introduces the doppler, gain, delay and other effects (multi-path and other topographical effects) the waveforms will be subjected to when propagating from RF source to the receiver. The Threat Radar models typically create Waveform Descriptor Words (WDW) or Pulse Descriptor Words (PDW). Higher fidelity models may output Digital I/Q waveforms. There are a range of techniques for creating RF environmental effects from high fidelity Complex Electromagnetics (CEM) to engineering models using simpler RF propagation formulas. The fidelity of the Threat models and RF Environmental models should be matched in the DT. For the notional DT, we are assuming Digital I/Q for the Threat and Background RF models and RF propagation formulas for environmental effects.

The External Interface shown connected to the Receiver Control components represents the receiver connection to the platform Pilot Vehicle Interface (PVI), which consists of the display and control used to operate the system. The Federated Systems are the avionics and other EW systems the receiver may be connected to. An ideal DT requires the modeling of these devices, with accurate interfaces connecting them to the DT Cognitive receiver control block. Note that for simplicity, we have omitted these interfaces from the discussion that follows.



## Cognitive Receiver Digital Twin Use Cases

The DT is a valuable tool for complex systems development, training, test and sustainment. Following is a partial list of potential DT use cases in the Development and Sustainment life cycle of the Cognitive Receiver:

- System Development
  - Early algorithm development, 1st order AI/ML training
  - Verifying initial hardware design, unit test
  - Refined AI/ML Training/Testing
  - System/Integrated Test
- Formal testing - Developmental Test/Operational Test
- Sustainment
  - AI/ML training, firmware/software updates
  - Regression testing

The following discussion will focus on uses of the DT for AI/ML training/testing for development and sustainment functions for the Cognitive Receiver and will discuss the Hybrid DT concept in detail for a Cognitive Receiver.

## DT applied to AI/ML Training and Testing

As discussed earlier, a DT is a valuable tool for addressing AI/ML training and the unique challenges associated with AI/ML testing. Specific DT benefits for the Cognitive Receiver use case are:

- It is very difficult to create the needed complex RF training and test environment efficiently either in the lab or on the Open-Air Range. The DT has the potential for doing this for the RF Threat, RF Background, and Systems interfaces needed.
- The DT provides test scalability to traverse the training and testing coverage space more quickly for regression training/testing.
  - FTRT training and testing is likely needed, which may be possible in a DT.
  - The DT can virtualize multiple instantiations of AI/ML algorithms to provide accelerated training.
- The use of a DT enables early Modeling and Simulation (M&S) for the design cycle, which is critical for AI/ML systems.
  - The DT supports AI/ML Algorithm development/design/training.

However, there are practical limits to the realism that can be achieved simulating complex systems in complex environments. Specifically, the Cognitive receiver RF and analog components may be difficult to model accurately. A Hybrid DT combining real and simulated components may be able to address this limitation.

## Hybrid DT Architecture

The basic Cognitive Receiver hardware and its DT in the Hybrid DT architecture is redrawn in Figure 5. The Hybrid DT architecture supports RF stimulus (a primary component of the training/testing dataset) from one of three sources: Recorded RF, HITL RF and Digital RF. The selected RF stimulus feeds an Environmental effects generator to provide realistic RF that



changes throughout a dynamic scenario. The Environmental Effects block can feed either real Cognitive receiver hardware in a HITL setup or a DT implementation of the Cognitive receiver. For simplicity, the receiver processing chain blocks have been broken into Pre-Processing, AI/ML Component, and Post-Processing blocks. The AI/ML Component could be any of the AI/ML component blocks included in the Notional Cognitive EW receiver shown in Figure 3.

In summary, the Hybrid DT setup provides stimulus from recorded, real, and digital RF sources, feeding real or DT hardware. This flexibility is useful for conducting both AI/ML training and testing.

We'll discuss three major aspects of the Hybrid DT architecture – AI/ML training and testing dataset generation, AI/ML component isolation testing, and a specific process for using the Hybrid DT testbed during training and testing.

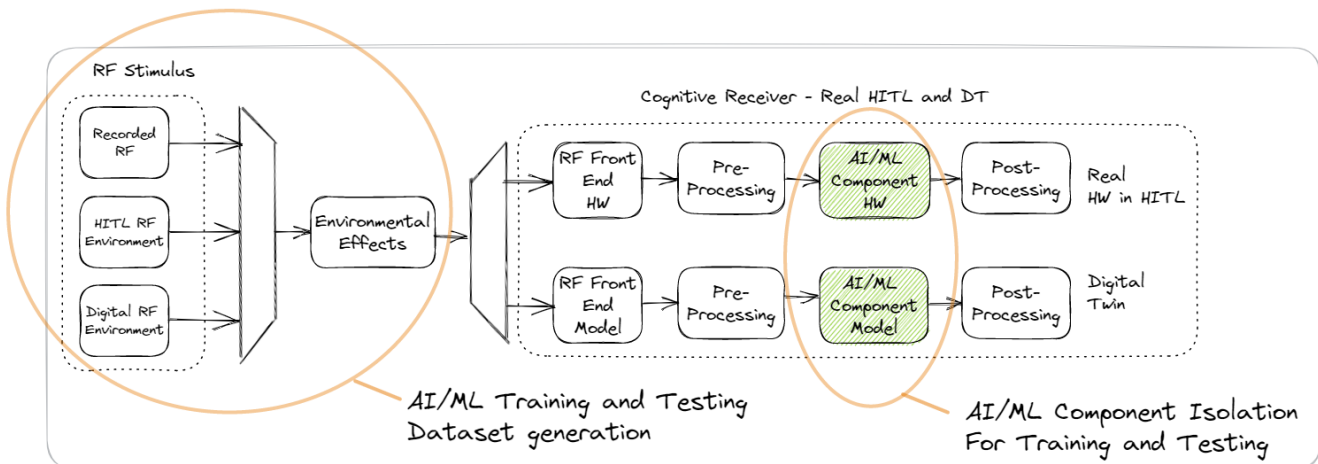
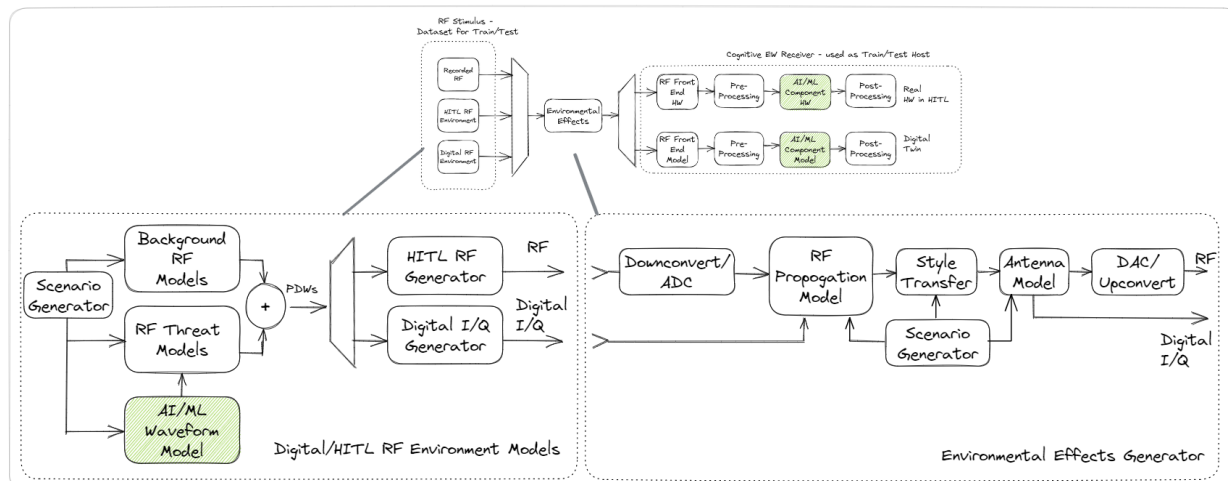


Figure 40 – Hybrid DT Architecture

## AI/ML System Training and Testing Dataset Generation

A Cognitive receiver training/testing dataset is required to train and test the system. It consists of either RF or digitized I/Q data generated by the Digital RF Environment and Environmental Effects blocks shown in Figure 6. Data elements are tagged so they can be correlated with RF emitter activity, RF band, and environmental effects used to create it.





**Figure 41 – Generating Training/Testing Dataset**

The Scenario Generator (SG) simulates the path of the Cognitive receiver platform through an operational RF threat environment. It is scripted with the expected RF threat laydown, background RF sources, PVI and Federated system interaction and simulated flight path for the Cognitive receiver platform. It executes the script, generating messages that feed the Background RF model, RF threat model, Environmental Effects model and External interface models (not shown in diagram), allowing them to generate synchronized, scenario-representative interactions with the receiver. Note that the scenario script, SG and RF and Environmental Effects generator all work together to create the training and test data spaces for the receiver. The test designer needs a complete understanding of how these elements work together to create these spaces. Given the complexity of the problem a test coverage tool should be created to assess how well scenario scripts are covering the overall space.

Based on SG inputs, the background RF model and RF threat models continuously generate PDWs that define discrete RF pulses. Note that the RF threat model is also driven by an AI/ML waveform model that generates “Novel” waveforms outside of the parameter space of known RF threats. This model is intended to enable generalized classification training for relevant AI/ML receiver components.

The PDWs feed a digital I/Q generator that streams wide-band digital I/Q, providing realistic threat data for the receiver system. The I/Q data feeds an RF Propagation model that adds doppler, gain, delay, clutter and multipath effects that would be induced on the RF as the scenario executes.

The I/Q data is then fed through a Style Transfer block where additional effects can be applied. These effects might be additional RF threat or environmental effects that are added for realism.

Note that the receiver antenna model block is incorporated into the Environmental effects block, assuming that the HITL version of the receiver will not incorporate an antenna.

The system training/testing dataset, generated via the Digital I/Q generator and Environmental Effects generator, needs to be validated prior to usage for training and test. Validation should consist of comparison of individual model performance with real data and validation of end-to-end performance versus real data. Note that validation in this context does not refer to the formal, rigorous model validation required for operational test. Several end-to-end validation methods are briefly considered below. The most straightforward end-to-end validation is done by generating an equivalent dataset using the HITL RF generator to generate



real RF, then comparing the HITL and digital I/Q generated datasets. Recorded RF data could also be injected into the Environmental effects generator and compared to digital I/Q data as a further validation step. In this case the digital I/Q data would need to be driven with a scenario script matching the real scenario used when recording the RF data.

Recall that efficient data generation is a significant issue for training and test of AI/ML systems. This data generation method should mitigate the issue to some degree. It's important to note that real data is still required to verify the synthetic data.

### **AI/ML Component Isolation**

The Hybrid DT is designed to provide individual AI/ML isolation, which allows for direct injection of inputs and direct access to outputs of a given component. This feature is required for training components in a complex system. Direct injection of component inputs allows for efficient input data generation and a higher degree of control for taking the AI/ML component input data through the complete coverage space. Direct access to outputs provides greater transparency when doing early training and testing – the tester can directly observe whether a component is performing as expected.

Component isolation can be done readily with a digital environment, but is more of a challenge with real hardware in a HITL environment. Isolation is enabled by ensuring that algorithms implemented in software and hardware conform to interface standards that are transparent to an ecosystem of potential algorithm developers. This minimizes the level of effort required to insert these algorithms into program of record (POR) systems.

Note that component isolation is very similar to AI/ML component training/testing in a standalone environment. The difference is that initial standalone development of AI/ML components provides a first order approximation of real inputs, meaning that the AI/ML algorithm at that point will not be adequate for usage in a real system. The next step for training should be done using AI/ML component isolation.

### **Hybrid DT Training and Testing process**

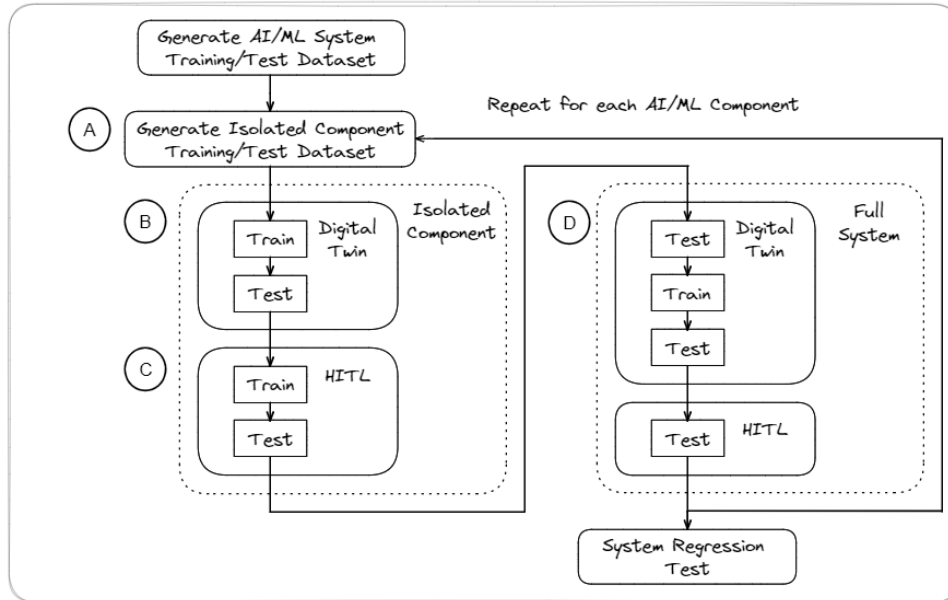
To illustrate how this setup for AI/ML component training/testing can be used, the process is broken down into major process steps in Figure 7.

The first process step is generating the AI/ML system training and testing dataset, which was discussed in detail above. The next step consists of a loop of train-test cycles performed for each isolated AI/ML component. Each loop is comprised of the following major processes –

- A. Generate isolated component training/testing Dataset
- B. Train and test Isolated component on DT
- C. Train and test Isolated component on HITL
- D. Test, train and test component on full system DT and HITL

Once all components are trained and tested, the entire system is regression tested using the DT and HITL HW. Each of the major process steps is discussed below.





**Figure 42 - AI/ML Training/Testing Process**

### Generate Isolated Component Training/Testing dataset

A dataset that is injected into an isolated component can be generated by stimulating the system with the system training and test dataset and recording inputs from components that feed the isolated component, as shown in Figure 8. These inputs are the direct injection data that will be used to train and test that component. The direct injection inputs are correlated with RF stimulus and Environmental effects that produced the data, along with event tags. The component designer will need to associate injected inputs with expected outputs for the AI/ML component model. These associations will form the truth data used for training and testing.

It would be beneficial to perform a coverage space analysis to determine how much of the AI/ML component model input space is actually covered by the generated dataset. If large parts of the direct inject dataset are uncovered, the designer may need to determine if there are issues with the SG scripting for the RF Stimulus or Environmental Effects blocks, or issues with the way these blocks are functioning.

There are challenges associated with this approach. Note in Figure 8 that Digital I/Q data is directly fed into the digital pre-processing model of the Cognitive Receiver, bypassing the RF front end model. This is done to simplify the creation of the DT. It may not be feasible to create a high-fidelity RF front end model. Additionally, feeding digitized RF into the front end may not be feasible either. The penalty paid by bypassing the RF front end model, is that the digital I/Q will have the RF front end effects absent, which could affect performance of downstream AI/ML processing in the real system. This may be mitigated by introducing RF front end effects in the Style Transfer Block (refer to Figure 6) during system training and testing dataset generation.

Care needs to be taken with the order of isolated AI/ML components for which direct inject data is generated. AI/ML component blocks that precede the chosen component must be trained before a given downstream component is addressed. If multiple components feed a given component, or there is component data feedback, then this process could become difficult and iterative training with multiple components may need to be done. This process works well for the Cognitive receiver example due to its straightforward signal processing pipeline. It may

not work as well for more complex systems; indeed, the feasibility and potential success of this approach heavily depend on the specific system architecture.

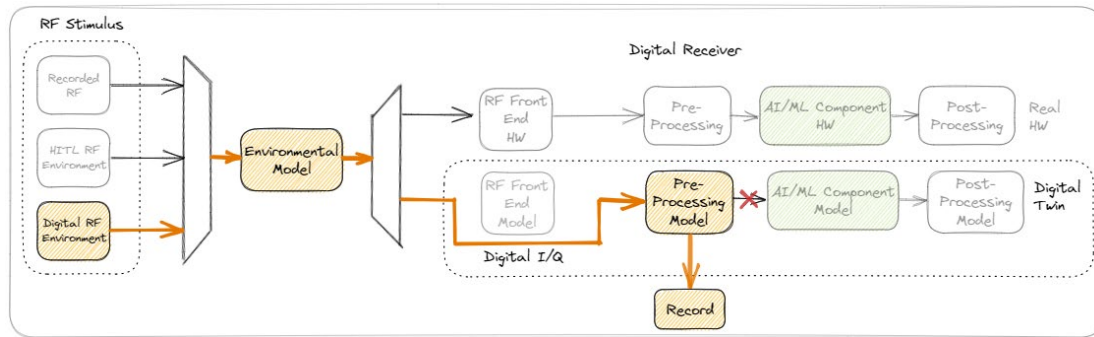


Figure 43 - Generating Isolated Component Dataset

### Train and test Isolated component on DT

In this step, the Direct Inject dataset generated for the specific component is played back into the AI/ML component in the DT to train the component as shown in Figure 9. Conducting the training may involve very large data sets and require multiple training cycles. Doing the training on the DT using the recorded dataset allows the training to potentially be conducted faster than real time. This is useful in a system like a Cognitive Receiver that likely requires continuous training in sustainment to adapt to a changing RF and threat environment.

Note that it is crucial that the digital training data used in this step be as realistic as possible, reflecting real threat, environment and system front end effects. If not, the AI/ML algorithm probably won't handle these effects properly in an operational environment. It was noted above that the RF front end would likely need to be bypassed with Digital I/Q data generation for the DT. This may be mitigated through the use of Style Transfer in the Environmental Effects generator, but it is anticipated that this will present its own set of challenges. Generally, training is conducted in training-validation-test cycles. This architecture should support these functions.

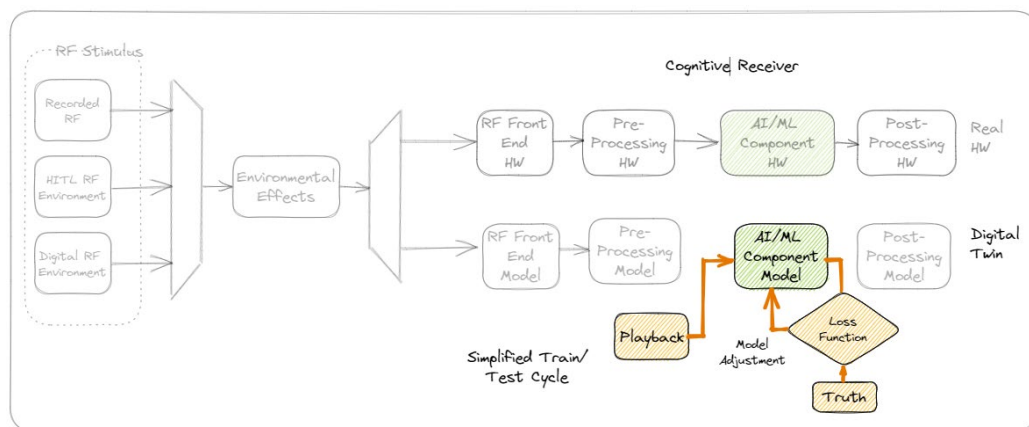
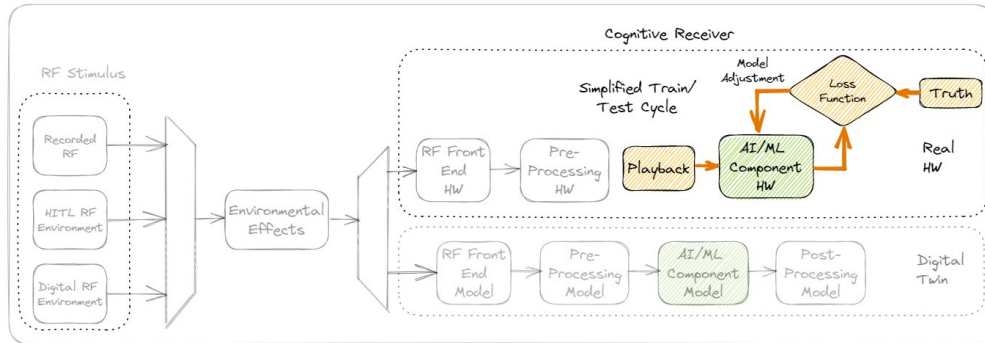


Figure 44 – Training and Testing Isolated Component on DT

### Train and test Isolated component on HITL

In this step, training and testing is conducted on isolated AI/ML components on the HITL (refer to Figure 10). The HITL will provide a higher level of realism; it uses real RF sources and

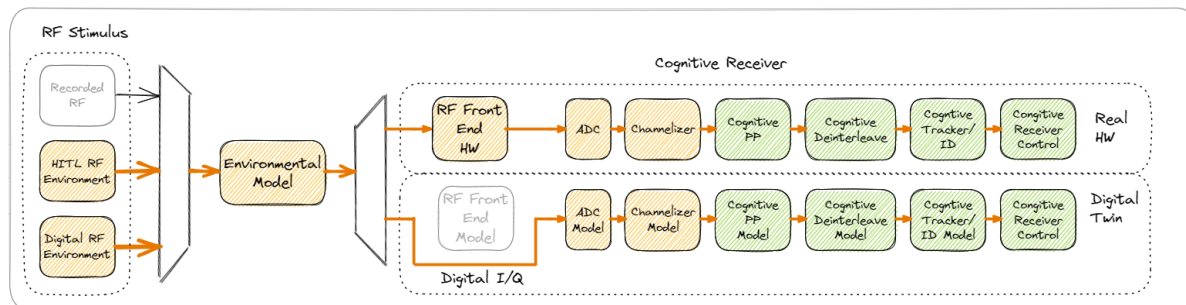
incorporates the RF front end signal path. The downside is that it must be run in real time, limiting extensive training cycles. It may be feasible to do fine tuning of AI/ML algorithms if the real time limitation does not create unacceptably long training cycles. It should be feasible, however, to use the HITL path for regression testing, which could be critical for verifying AI/ML performance after training on the DT.



**Figure 45 – Training and Testing Isolated Component on HITL**

### Test, Train and test component on Full System DT and HITL

After each component is individually trained and tested, it must be evaluated in the context of the overall system (refer to Figure 11). There may be interactions and dependencies that impact performance of the component that would only be seen in the full system environment.



**Figure 46 – Training and Testing Full System on HITL and DT**

The recommended approach is to initially test the component on the full system using the DT to determine whether the component has an acceptable level of performance. The intent is early identification of component performance issues and root cause analysis. If causes are related to training data variances, then it may be necessary to adjust the isolated training dataset and re-train and re-test the component in the isolated environment. If the component performs acceptably, then a train–test cycle is initiated to further refine training.

Next the component is tested on the HITL setup. If performance is acceptable, the full process is repeated for the next AI/ML component. Root cause analysis is conducted if the component fails, which may result in adjustment of training data and re-training and re-testing in the isolated component mode.

Depending on system complexity, there may be confounding interdependencies among the AI/ML components that prevent complete training and testing of a given component. For example, it may not be feasible to completely train/test component A, then completely train/test

component B, etc., given component interdependencies. An iterative capability approach will likely be required: train/test component A with initial capability, train/test component B with initial capability, etc., iterating through the process repeated times, layering on additional capabilities for components in the chain.

A full system regression test will be run on the DT and HITL once component training and testing is completed.

### Hybrid DT Model Validation

Initial Hybrid DT validation is required as soon as real system operational data can be collected. Prior to or during Developmental Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E), it may be feasible to collect data using an Installed System Test Facility (ISTF) like an anechoic chamber. During DT&E and OT&E Open Air Range (OAR) testing, real operational data will also be available for validation. Validation will also continue over the life of the system, if data can be collected during operational usage.

The type of data collected during these events will dictate how it is used for Hybrid DT validation. Ideally, a data recorder would collect RF data at the receiver faceplate in an operational environment with ground and airborne threat simulators. Truth data such as aircraft Time, Space, Position Information (TSPI), threat state data and range RF instrumentation for other emitters would also be required. The RF data would be played back as indicated in Figure 12. Using collected truth data, the Hybrid DT performance can be verified against actual Cognitive receiver performance.

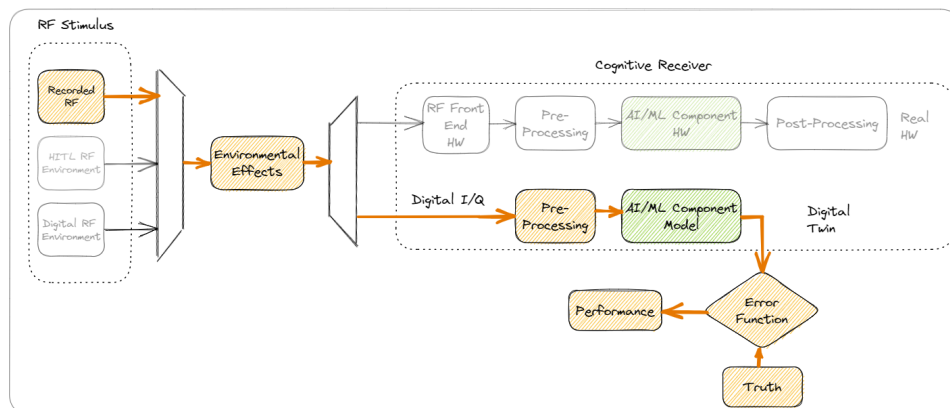


Figure 47 – Hybrid DT Model Validation - Component

It would also be useful to record and playback internal Cognitive receiver instrumentation data such as digital I/Q, PDW buffers, or emitter track file buffers to verify individual receiver component performance.

During Hybrid DT validation, it is likely that there will be some variance between the real system performance and the Hybrid DT performance. This should be viewed as an opportunity to further refine the Hybrid DT by verifying recorded playback, RF Stimulus block, Environmental effects block, RF front end assumptions, and correct virtualization of the AI/ML components and OFP, or other potential sources of variance. As sources of variance are found and fixed, further confidence will be established for the Hybrid DT.

## Considerations

The general advantages of using DTs is clearly understood and was stated above. For this discussion we'll consider the advantages of a Hybrid DT versus a purely digital DT. The primary advantage of the Hybrid DT architecture is improved realism. Systems with complex Cyber-Physical interaction and heavy sensor-dependencies are difficult to implement with purely digital DT. Implementing high-fidelity digital models for sensors and complex RF and analog signal processing can be a significant challenge. The use of hardware in the Hybrid DT to replicate these behaviors, serving as an adjunct to the DT can improve realism for those elements.

There is a basic trade-off of simulation realism versus simulation time that can be balanced with the Hybrid DT. The hardware components have increased realism but also have increased simulation time (they cannot be run faster than real time). The digital components will have less realism but can potentially be run faster than real time. The Hybrid DT uses hardware implementation only for the components that can't be simulated on a digital environment with sufficient realism. Fortunately, in the case of the Cognitive receiver, the AI/ML and software components can be virtualized with reasonably high fidelity because they are already in the digital domain.

However, the Hybrid DT will require significant effort to develop and maintain. Some in the EW T&E community argue that resources should be dedicated to improve operational testing of systems instead of DT and Digital M&S. Certainly, digital M&S has been overhyped in the past, leading to false perceptions of feasibility, accuracy, and utility. Several of the anticipated challenges implementing the Hybrid DT are as follows:

- There is no one-size-fits-all solution. The specific implementation and training/test process will vary depending on the system. The Cognitive receiver example is essentially an open loop system. More complex systems such as RF jammers will present additional difficulties.
- It is essential to verify the Hybrid DT with real OAR data collected during DT&E/OT&E and to continue validation over the system life cycle.
- For Cognitive EW applications, a critical part of the DT is the RF and threat environment. Great care needs to be taken to ensure that this environment is accurately replicated. Other AI/ML applications such as autonomous driving have similar challenges simulating realistic environments.

## Conclusion

The Hybrid DT approach demonstrated above is a promising approach for providing the improved training and test capabilities required for complex AI/ML systems.

Through targeted usage of real hardware, coupled with digital simulations, the Hybrid DT should be able to simulate the system and its operational environment with sufficient realism. If the RF operational environmental simulation is built with scenario generation capability and environmental effects simulators, much of the required training and test data may be able to be generated. Finally, the Continuous Integration/Continuous Deployment (CI/CD) process required for AI/ML systems can be supported if the system is constructed with a development pipeline that supports efficient virtualization of AI/ML components and firmware/software components.

Program test managers should carefully consider Digital Twinning and Digital model approaches, and adapt test constructs that are best suited for their system, considering system Cyber-Physical interactions and system complexities. This is particularly true for AI/ML based



systems, like Cognitive EW systems. Test constructs should also be chosen in the context of the complete system life cycle, including design, implementation, DT&E/OT&E and sustainment.

## References

- Anthony Ortiz, O. F. (2018). On the defense against adversarial examples beyond the visible spectrum. *Milcom 2018 Track 5 - Big Data and Machine Learning*, (pp. 553–558). Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8599763>
- Antonio Nieto-Rodriguez, R. V. (2023). How AI will transform project management. *Harvard Business Review*.
- Chuanqi Tao, J. G. (2019, August 23). Testing and quality validation for AI software—Perspectives, issues, and practices. *IEEEAccess*, 120164–120175. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8811507>
- Collins, C. (2023). *Test and evaluation as a continuum*. OUSD (R&E).
- Digital Twin Consortium. (2020, December 3). *Definition of a digital twin*. <https://www.digitaltwinconsortium.org/initiatives/the-definition-of-a-digital-twin/>
- DoD. (2018). *National defense strategy*. [https://nssarchive.us/wp-content/uploads/2020/04/2018\\_NDS.pdf](https://nssarchive.us/wp-content/uploads/2020/04/2018_NDS.pdf)
- DOT&E. (2022). *Digital twin assessment, Agile verification processes, and visualization technology*.
- DSB. (2016). *Summer study on autonomy*.
- GAO. (2022a). *How artificial intelligence is transforming national security*. <https://www.gao.gov/blog/how-artificial-intelligence-transforming-national-security>
- GAO. (2022b). *AI - Status of developing and acquiring capabilities for weapons systems*. <https://www.gao.gov/products/gao-22-104765>
- Jörn Thieling, J. R. (2021). Scalable sensor models and simulation methods for seamless transitions within system development: From first digital prototype to final real system. *IEEE SYSTEMS JOURNAL*, 3273–3282.
- Kirill Semenov, V. P. (2020). Verification of large scale control systems with hybrid digital models and digital twins. *2020 International Russian Automation Conference (RusAutoCon)*, 325–329.
- NASEM. (2021). *Necessary DoD range capabilities to ensure operational superiority of U.S. defense systems: Testing for the future fight*. The National Academies Press.
- NIST. (2021). *Considerations for digital twin technology and emerging standards*. <https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8356-draft.pdf>
- Prokhorov, D. (2019). Toward next generation of autonomous systems with AI. *IJCNN 2019. International Joint Conference on Neural Networks*, 1–5. Budapest, Hungary. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8851867>
- RAND. (2021). *The Department of Defense's posture for artificial intelligence*. [https://www.rand.org/pubs/research\\_briefs/RB10145.html](https://www.rand.org/pubs/research_briefs/RB10145.html)
- Sandra Hobson, P. L. (2022). DOT&E strategy update 2022: Transforming T&E to enable delivery of the world's most advanced warfighting capabilities at the speed of need. *Aircraft Survivability Program - JASP Online*.
- Tate, D. (2019). *What counts as progress in the T&E of autonomy*. Institute For Defense Analyses.
- Tian, Y. (2018). DeepTest: Automated testing of deep-neural-network-driven autonomous cars. *2018 ACM/IEEE 40th International Conference on Software Engineering*, 303–314. <https://doi.org/10.1145/3180155.3180220>
- USAF Chief of Staff. (2020). *Accelerate change or lose*. [https://www.af.mil/Portals/1/documents/csaf/CSAF\\_22/CSAF\\_22\\_Strategic\\_Approach\\_Accelerate\\_Change\\_or\\_Lose\\_31\\_Aug\\_2020.pdf](https://www.af.mil/Portals/1/documents/csaf/CSAF_22/CSAF_22_Strategic_Approach_Accelerate_Change_or_Lose_31_Aug_2020.pdf)





## PAPERS ONLY

---

	<p><b><i>Guiding the Hands of Time: Toward Reliable Schedule Estimates</i></b></p> <p>Charles Pickar, Naval Postgraduate School Raymond “Chip” Franck, Naval Postgraduate School</p> <p><b><i>Training an Agile Acquisition Workforce to Combat Emerging Threats</i></b></p> <p>Amanda Swanson Goff</p> <p><b><i>Through the Looking Glass: Why EVM Is An Essential Risk Mitigation Measure for Decision Makers and Program Managers</i></b></p> <p>Symantha “Sam” Loflin</p>
--	---



# Guiding the Hands of Time: Toward Reliable Schedule Estimates

**Charles Pickar**—has recently retired from the Naval Postgraduate School (NPS) faculty, where he taught project management, defense acquisition, and systems engineering. Before joining NPS, he led the Applied Systems Engineering Program Area at the Johns Hopkins University Applied Physics Laboratory. He is also a retired Army officer with extensive experience in the U.S. defense industry, including Director and VP levels at Lockheed Martin, Northrop Grumman, and SAIC. He has also served as Chair of the Systems Education Technical Committee of the IEEE Systems Council. His research and published work focus on applying systems engineering and system dynamics analytical approaches to defense acquisition problems. [ckpickar@gmail.com]

**Raymond Franck**—retired from the NPS faculty in 2012. He retired from the Air Force in 2000 in the grade of Brigadier General. His active-duty career included several operational tours, staff positions, and head of the Department of Economics and Geography, United States Air Force Academy. His published work includes many journal articles and research reports on military innovation and defense acquisition management. [cfranck215@aol.com]

## Abstract

This paper continues our research agenda concerning advancing the state of the art for estimating defense acquisition program schedules. Accurate schedule estimates provide valuable benchmarks for program managers and reliable dates for the availability of new systems for warfighters. However, most schedule estimates are not correct, with actual time to complete significantly greater initial estimates. This happens for several reasons. One of them is the inherent complexity of modern defense acquisition programs. Another is the generally unfavorable influence of factors outside the program (and program management control). While achieving improved estimates is worthwhile, we also conclude that accurate estimates are generally unobtainable. However, we remain convinced that improvements are possible, which benefit all concerned.

## Introduction

*Why do weapon system developments always take longer than planned, and why are we always surprised when they do? Why indeed?* These two questions have been the core of our schedules research agenda for the past 7 years. Estimating, developing, and executing weapon systems development schedules is rife with challenges, from the schedule estimating process, the complexity of schedules, and the impact of system dynamics on estimation and execution to intangibles associated with our astonishment at the inaccuracies of our schedules. A system development schedule is a promise between the acquisition organization and the customer—and the customer takes us at our word. It is essential to get better.

Substantial resources have been spent over several decades on improving data and forecasting models. Nevertheless, this has had no effect on the accuracy of forecasts. ... This indicates that something other than poor data and models is at play in generating inaccurate forecasts. (Flyvbjerg, 2006, p. 6)

Schedules provide both planned sequencing and duration of weapons system development events. When accurate, the schedule offers the warfighter a reliable date the systems will be available. However, most projects overrun the schedule. This research aims to identify ways to predict the duration of defense acquisition schedules more accurately. This effort is part of a continuing research agenda started by Franck et al. (2016). The latest paper in this line of research (Pickar & Franck, 2022) discussed the issues of commonalities versus individual differences in estimating program schedules, complexities in the processes that



determine schedule duration, and program management decisions (which are hard to observe and generally impossible to predict reliably).

It is common knowledge that defense acquisition programs (especially major defense acquisition programs [MDAPs]) experience “deviations” from planned schedules (generally delays). These typically result from untoward events, such as continuing resolution constraints on program progress, or discoveries encountered during program execution. Delays come in many forms, such as deliveries, tests, inadequate system performance, and misspecified tasks (leading perhaps to rework).

However, perhaps the most crucial aspect of this problem is that an adverse development in one aspect of system development can cause undesirable effects on other parts of the program. Thus, for example, the 2004 discovery of weight increases in the F-35 aircraft led to a comprehensive program to pare weight from the aircraft. This included changing methods of assembling the fuselage from snap-together panels to assembly with “traditional fasteners.” This solution to one problem caused other problems, such as manufacturing costs, schedule delays, and lessened F-35 performance (due to losing “good weight”; Pappalardo, 2006).

Also costly (in financial and operational metrics) are delayed retirements of systems that were earmarked to be replaced. As Sweetman (2012) put it regarding the F-35, “The failure of the so-called fifth-generation fighters . . . to arrive on time and cost is having cascading effects throughout U.S. and allied fighter forces.”

Therefore, our overarching research objective is to identify delay-causing events by both examining why the delays happen and proposing tools and processes to better estimate schedule duration at various phases of a given program:

- **Ex ante:** Identify propensity toward untoward developments. Our past work has taken a broad approach to (a) summarize and characterize key elements of the literature on program duration, (b) add to that literature through case study methods, and (c) formulate a more sophisticated model of acquisition program trade-offs. We still think Schedule Estimating Relationships are promising—albeit enriched with modern data science techniques.
- **In medias res:** When the almost-inevitable adverse events occur, it’s helpful to estimate the effects of those events and (ideally) mitigate them. Methods such as computer-enabled content analysis (or text mining) have shown promise in schedule estimation and significant research—including work by acquisition professionals (e.g., Joseph & Sconion, 2020). We also explore prediction markets as another means of crowdsourcing useful information about programs’ progress.
- **Ex post:** What can experience gleaned from past defense acquisition programs enhance our understanding of the basic (albeit complex) processes in play? How might that experience enhance the arts of schedule estimation and program management? We discuss the art of Root Cause Analysis—which the DoD has in place for cost outcomes. We also highlight issues associated with cutting-edge data-gathering and analysis methods.

Our central research question is: *How can the schedule estimating process be improved to reflect the data-identified causes of schedule duration?* Doing that can make program managers more effective. We are also committed to multiple approaches to estimate improvement as appropriate for a significant and challenging problem.



## Agenda for Improvement: Ex Ante

Having a realistic initial schedule estimate is vital for many reasons. First is a credible plan for completing the program—within reasonable limits on time and money expended. Such estimates also mitigate management difficulties in programs experiencing challenges.

Second is a reliable plan to bring new capabilities online to operate with other forces. This also facilitates managing the remaining operational life of the “legacy” systems (Sweetman, 2012). This is likely more important in an era of capabilities portfolios (Drabkin, 2019).

Third, a credible schedule provides guardrails for programs encountering difficulties. “How did (program management) know its program (execution) was failing? By the schedule and budget slipping. . . . If those forecasts were fundamentally unrealistic, a team expected to meet them would fail no matter what they did” (Flyvbjerg & Gardner, 2023, p. 99). Without a reasonable schedule estimate, program management is, in effect, flying blind.

However, there are good reasons why ex ante schedule estimates are commonly not valid (Flyvbjerg, 2006, p. 6) and are much less accurate. First, those best qualified to provide program schedule estimates are, as a rule, optimists who are incentivized to be optimistic (Pickar & Franck, 2022, pp. 12–13).

Second, an acquisition program is a managed effort. Pickar and Franck (2021) discussed the importance of management decisions in determining actual schedule times (pp. 4–12). Since those decisions are unknowable before the actual decision, there’s significant uncertainty baked into any schedule.

Third, the difficulty is compounded dramatically by defense acquisition programs’ propensity for complexity. Deterministic project scheduling assumes complete information about the scheduling problem to be solved and a static environment within which the schedule will be executed. However, the actual project environment does not behave predictably.

As Dörner (1989) puts it, “Complicated systems . . . derive their complexity from the presence of interrelated variables. One cannot see everything one would like to see” (p. 35). Moreover, complex systems are predisposed to “emergent” behavior (Franck et al., 2012, p. 107; Complex System, n.d.). This implies the system is prone to unpredictable results (Complex System, n.d.)—especially viewed ex ante.

One approach to mitigate this problem is “anchoring” the estimate with reference to past experience: “To create a successful project estimate, you must get the anchor right” (Flyvbjerg & Gardner, 2023, p. 106). Part of “anchoring” is to identify a “reference class,” approaching the estimate as “one of a class of similar projects already done” (p. 107). One takeaway from this discussion is that formulating schedule forecasts entails much thought and preparation.<sup>1</sup> Anchoring is a practical step but may not be sufficient for the estimate to be helpful.

One method of anchoring estimates is through Schedule Estimating Relationships, which are generally derived from completed programs. These are (as a rule) ingenious quantitative studies that relate observed program characteristics to actual program schedules.

An excellent example of this approach is Light et al. (2018). The authors related actual schedules to various program characteristics, including the acquisition policy era (pp. 3–6). Using those results, a method is available to assess the plausibility of newer schedule estimates based on “program characteristics” (pp. 11–14).

---

<sup>1</sup> We find nothing inconsistent between the “anchoring” approach and the Scheduling Estimating Relationship methods in acquisition research literature. There is a significant difference in perspective: ex ante versus ex post.



Another is to assume the schedule is not predictable and hedge against that; that is, strive for robust estimates<sup>2</sup> to facilitate resilient program execution. This is no panacea, but as Flyvbjerg and Gardner (2023) put it, those who “lead a big project . . . should . . . protect themselves against overruns. The obvious way to do that is to build a buffer” (p. 9)

These approaches are not mutually exclusive. Current program management practice includes such measures, including contingency and management reserves. Contingency reserves are funds or time set aside to mitigate defined risks. At the same time, management reserves are intended for risks yet to be fully explained (e.g., Project Practical, n.d.), like “black-swan” events (Flyvbjerg & Gardner, 2023, pp. 10–11). A more expansive view would include risks outside the program’s scope and, therefore, beyond management control. As noted, original estimates (cost and schedule) are powerfully predisposed to being optimistic. If needed reserves are identified, the overall plan has a degree of resilience that is otherwise missing; management would have the wherewithal to address untoward events.

These should include reserves hedging against developments outside the program. For example, a major acquisition program schedule could consist of a schedule time contingency reserve to account for the effects of restricted funding (and program actions) due to continuing resolutions instead of appropriations.<sup>3</sup>

## In Media Res

Since estimating is an inherently uncertain craft, updating program outcomes (especially cost and schedule) is a handy capability. Even more potentially useful is identifying emerging program issues—hopefully in time for program managers to mitigate or forestall them.

## “Wise Crowds”: Crowdsourcing Acquisition Program Predictions

There is good reason to believe that the collective estimate from a group can be considerably more accurate than the judgment of an expert panel. Substantial experience supports this hypothesis. However, groups can be spectacularly wrong (e.g., financial bubbles, long-shot winners, and black swans). Yet, “even if most people in a group are not especially well-informed or rational, it can still reach a collectively wise decision” (Surowiecki, 2004, pp. xiii–xiv).<sup>4</sup>

A primary framing assumption for groups potentially being intelligent is that every member has private information. Each set of data includes insights and errors (of various kinds).<sup>5</sup> In a “proper” group setting (discussed just below), the (private) errors across the group tend to cancel each other out in the aggregation of opinions, while the private sets of information add to the quality of the collective opinion (Surowiecki, 2004, pp. 10, 41).

Fundamentally, the wisdom-of-crowds hypothesis is a proposition that the capability of the whole is greater than the sum of the individual members’ capabilities. Aggregating private knowledge improves the group’s capacity to solve problems, while the individuals’ errors largely cancel out. For example, Hayek (1945, pp. 17, 19–3) and Smith (1776, pp. 13–16) discussed the ability of a crowd of market participants to reach a sensible economic equilibrium.

Various lines of inquiry have identified characteristics of “wise crowds.”

---

<sup>2</sup> These two approaches are not, of course, mutually exclusive.

<sup>3</sup> One could make a case for including continuing resolutions (CRs) in schedule estimates. The very high probability of a CR in any given could be part of the baseline estimate—with a contingency reserve to address the unknown length of the CRs.

<sup>4</sup> Surowiecki (2004) included an extensive set of notes and citations (pp. 275–296). Due to editorial constraints, we do not delve deeply into that literature here.

<sup>5</sup> One can view each set of private information as having two components: useful knowledge and errors, without individuals being aware of how their private information is divided between those components.



- **Cognitive diversity** is formed in good part by the heterogeneity of private information. Insufficient “cognitive diversity” can lead to “groupthink” and associated pressures to conform (Surowiecki, 2004, pp. 23, 38). Groups that are too much alike find it harder to keep learning because each member is (incrementally) bringing less and less new information to the table (p. 31). Diversity adds different perspectives to the group and lessens the pressure to conform to a consensus, stated or emerging (p. 39).<sup>6</sup>
- **Independence of members** means that individuals are not influenced by other group members. (With insufficient independence, there is a tendency for “herding”). It also promotes a diversity of errors in the sets of private information, which are more likely to cancel out. It also means that each individual’s information component is more likely to be additive to the group’s information rather than the “same old data” (p. 41).
- **Decentralization** (in an organizational sense) means that information is processed throughout the organization (or outside of it) rather than through a hierarchy.<sup>7</sup> This can foreclose the tendency of hierarchies to filter out information and judgments at lower levels in a structured process to arrive at the “best” answer.

The characteristics above cancel private errors out (or tend, on average, to zero). After the error cross eliminates itself, the valuable sets of personal knowledge make the entire group capable of solving significant problems, such as “cognition” (e.g., election winners), “coordination” (e.g., the operation of a market), and “cooperation” (e.g., getting a disparate group to work together; Surowiecki, 2004, pp. xvii – xviii). This assumes some method of pooling the sets of private information to reach a desired solution. That gets us to the final characteristic.

**Aggregation** is a process (or set of processes) to bring out a collective assessment related to the entire group’s diverse, independent, and decentralized opinions to a good evaluation, forecast, or decision. Aggregating information in a traditional bureaucracy is a well-defined process of screening and assessing information through a series of filters. However, this runs valuable information can be suppressed or disregarded by a hierarchical organization (e.g., the 1986 Challenger launch accident; McCleary, 2023).

Aggregating the collective “wisdom” of a group implies another method other than hierarchical screening, such as prediction markets.

## Prediction Markets<sup>8</sup>

The advantage of prediction markets is that they can benefit from the wisdom of crowds. By collecting and weighing the predictions of a large number of traders, they can provide a market-wide forecast that is generally more reliable and balanced than any single expert opinion. (Peters, 2022)

In a prediction market, group members may place bets on defined outcomes. The event may be an election or athletic contest. The group judgment is the prevailing market “price” for the event may be expressed as the probability of a candidate winning or the margin of victory (or defeat). Thus, if Group Member X believes Candidate Y has a 25% chance of winning, he would buy a contract that would be willing to pay up to 25 cents for a contract predicting

<sup>6</sup> For example, the Bay of Pigs Invasion in 1961 was planned and executed “without ever really talking to anyone who was skeptical of the prospects for success” (Surowiecki, 2004, p. 37).

<sup>7</sup> In a hierarchical assessment of available information, perspectives of lower-level individuals generally count for much less than those at higher levels.

<sup>8</sup> Due to editorial limits, we provide only an overview of prediction market research results and a sample of the literature. However, the subject is well-recognized as suitable for in-depth research. See, for example, *The Journal of Prediction Markets* at <https://www.scienceopen.com/collection/755392c6-34de-437c-9d0a-c768f6f128bb>



Candidate Y's victory. The prevailing price is frequently expressed as a point spread in a sports betting market.

An excellent example of a prediction market was discussed in *Nature* (Mann, 2016, pp. 308–310). A “reproducibility project” initiative was formed to determine whether experimental results reported in psychology journals would return the same results with an independent replication. Because the participants “thought it would be fantastic to bet on the outcome(s),” they formed a prediction market to place those bets on whether a series of given results could be replicated. The salient results of this market were that individual experts “hadn't done much better than chance with their individual predictions. But working collectively through the markets, they correctly guessed the outcome 71% of the time” (Mann, 2016, p. 308).

Worth noting is that the group involved were experts in the relevant academic field and could be expected to have significant commonalities in outlook and opinions. Nonetheless, the group outperformed the individuals. Also noteworthy is that a prediction market initiated as an afterthought worked well.

## Forms of Prediction Markets

Prediction markets can take many forms by organizing principles and modes of operation. A non-exhaustive list appears below (Peters, 2022).

- **Continuous Double Auction** matches willing buyers and sellers of contracts at a specific market price—much like a stock exchange. The market authority records each transaction.
- **Automated Market Makers** act like a casino or parimutuel betting organization; the “house” serves as the other party to all bets (or trades) and adjusts odds (and payoffs) based on volume for each outcome.
- In **Play Money Markets**, the bets placed convey no market value. Participants are perhaps motivated by reputation or satisfaction in being right.
- A **Decentralized Prediction Market** features trades executed without any central management. “Smart contracts” can then “self-execute . . . to distribute payoffs.”

## Some Prediction Market Issues Relevant to Defense Acquisition

- **Self-Fulfilling or Self-Negating Group Predictions:** For example, a group prediction of an untoward acquisition program event can lead to management actions to prevent that event. We discussed this issue in a previous paper (Pickar & Franck, 2022, p. 24).
- **Positive and Perverse Incentives:** Prediction markets look like and can operate like betting markets. As such, there can be incentives to engineer a favorable outcome, which has happened in sports betting operations. This problem can be addressed by limiting the stakes. For example, the Iowa Electronic Market limits positions to \$500 (University of Iowa, n.d.). The Reproducibility Project gave each group member \$100 to wager (Mann, 2016, p. 308). Hence, relatively small stakes can nonetheless elicit candid assessments.<sup>9</sup> And that is good news for acquisition prediction markets.

However, motivated participation is beneficial. As one observer put it, “I can create a poll that can mimic everything about a prediction market, except markets, have a way of incentivizing you to come back at 2 a.m. and update your answer” (Mann, 2016, p. 310). Most importantly, a well-functioning prediction market provides valuable incentives for participants to reveal their judgments regarding the event at hand sincerely.

---

<sup>9</sup> There's reason to believe that simply being proved right is a useful incentive for thoughtful participation. As Kathryn Schulz (2010) put it, “The experience of being right is . . . one life's cheapest and keenest satisfactions” (p. 4).



## Prediction Market Issues

As good as prediction markets can be, there are some problems. Typically, they focus on well-defined binary outcomes (e.g., win or lose) that occur at a definite time. Results of sporting events and elections are good examples of this category.

But what happens if the outcomes are more complicated? Suppose a wise group identifies an emerging problem in an acquisition program (such as a schedule slip). Suppose also that alerted program management undertakes a remedy that averts the crisis.<sup>10</sup> How do the prediction market rules determine the winner?

One way around this problem is to have more detailed results. For example, group members could choose an outcome in perhaps two parts. Will the group identify the particular problem? If so, will management action avert the problem? While this seems a reasonable solution, even more, complex bets might arise in a well-designed prediction market for an acquisition program.

There might also be dilemmas (or trade-offs) in prediction market design. An essential assumption for prediction markets harnessing the wisdom of crowds is the “marginal trader” (Adam, 2016, p. 310), one who acts to benefit from current group misconceptions.<sup>11</sup> Incentivizing effective marginal traders might entail substantial incentives to be correct. Doing that could, in turn, constitute a significant incentive to take action (unethical or illegal) to increase the odds of winning the bet.

Another obvious issue is that acquisition programs (especially MDAPs) are lengthy and have uncertain termination or milestone dates. Defense acquisition prediction markets operating arena will likely need special care in framing the questions upon which to place bets.

The issues and problems we’ve raised are untested hypotheses but could add to the practical difficulties of organizing a functional prediction market. As such, they appear to be matters for more research and experience.

## Acquisition Data Qualitative Analysis

As part of our ongoing study this year, we apply qualitative research methods to improve on a 2018 macro-level study of factors that define schedule delays (Pickar, 2018). The 2018 analysis of Selected Acquisition Reports (SARs) used a cumbersome, manual process to code each schedule explanation text entry to convert it to structured, measurable data.

A constraint of studying the schedule process is data availability and data analysis techniques. Data for this project come from the DoD SAR. The documents to be examined are the SAR sections on the executive summary and the schedule change explanation of the SARs at the Washington Headquarters Services website (DoD, n.d.). The SARs are reports to Congress and, as such, are text or unstructured data. Text mining and Computer Assisted Qualitative Data Analysis Software (CAQDAS) are two ways to analyze unstructured data. CAQDAS allows qualitative analysis of hundreds to thousands of documents. Text mining, on the other hand, offers analysis for millions of documents. For this paper, we consider CAQDAS and text mining to be interchangeable. We are attempting

the discovery by computer of new, previously unknown information by automatically extracting information from different written resources. A key

---

<sup>10</sup> This is not so far-fetched. Miller (2012) offered a method that can yield actionable indication of an emerging problem in acquisition programs (pp. 48–49). It’s therefore reasonable to suppose that prediction markets might also provide similar warnings.

<sup>11</sup> Beaton and Cohen (2023) offered advice for would-be marginal traders in this year’s NCAA Men’s Basketball Tournament, noting that “there’s value in finding the best, least popular team and making it your national champion.” However, their championship pick, UCLA, lost in the Round of 16.





element is the linking together of the extracted information together to form new facts or new hypotheses to be explored further by more conventional means of experimentation. (Hearst, 2003, p. 1)

## Qualitative Analysis Method

This qualitative analysis uses a modified version of the “grounded theory” method. Grounded theory is “the discovery of theory from data systematically obtained and analyzed in social research” (Glaser & Strauss, 2017, p. 2). The grounded theory approach uses an iterative data collection and analysis process as central to theory development. In this case, studying the events described in the SARs could contribute to a model of schedule activities in ACAT I programs examining causes for both delay and acceleration. Schedule delays often consist of more than one root cause, and the dynamics of the delays frequently cause further difficulties in schedule and elsewhere. Grounded theory is practical when no existing theory explains the activities being observed. The process steps include collecting the data; analyzing the data; identifying/grouping the discovered concepts; and finally, identifying any relationships between categories of data.

A necessary pre-analysis step required collecting and preparing the data. Preparing the data involved resolving/removing duplicate data, fixing structural errors, and identifying outliers. Once accomplished, the qualitative analysis started with a literature review of the broad area of scheduling.

## Qualitative Analysis Literature Review

We used a qualitative, two-part process to examine the literature on schedule. The first part was a broad examination of 83 peer-reviewed papers on schedule. The literature review identified themes for use in the CAQDAS analysis for coding. Figure 1 is a treemap of the initial analysis of those papers. Of the top five themes, uncertainty (a subject we have not examined) ranked fifth in the literature. In this literature review, uncertainty included both schedule risk and schedule uncertainty. The literature ranged from uncertainty impacting the schedule development process from estimates to management to execution. In some papers, the terms are used interchangeably. We then conducted a separate literature review on uncertainty.

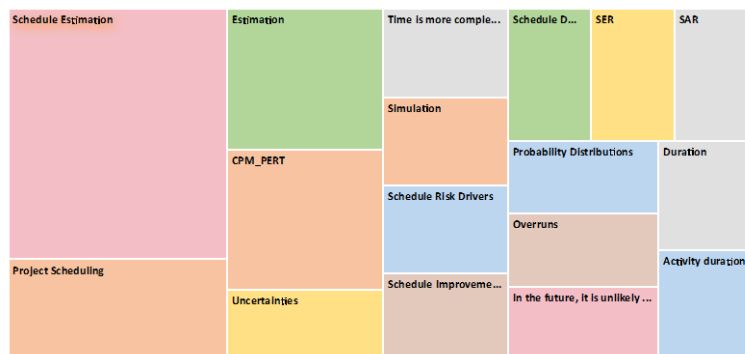


Figure 48. Treemap of the Schedule Literature

This review also provided the initial word cloud, a visual representation that allows the organization of the unstructured data to classify themes and find the relationships between those themes. Figure 2 is the word cloud for the uncertainty literature review.





Project management uncertainty is “the inability to predict future outcomes” (Shenhar & Dvir, 1996, p. 610). PMs crave certainty; however, the opposite is ever present in weapons system development projects. In fact, uncertainty in project management is generally assumed in the development of high-tech weapons. Uncertainty manifests in many other aspects of project management, including uncertainty in funding and the workforce (both skills and labor issues), to name a few.

On the other hand, risk is not only managed; it is quantified with probability (Koleczko, 2012). Uncertainty can’t be quantified. Risk is a distinct and identifiable project influence (and a practice embedded in the project management process.)

Like risk management, dealing with uncertainty depends on the program manager’s tolerance for ambiguity (Koleczko, 2012). Further, decision-making relates to certainty, risk, and uncertainty (Kerzner, 2013). Finally, too much information creates uncertainty, an idea all too familiar for combat commanders and staff. Uncertainty drives the need to gather more information, and that information inhibits decision-making in a vicious cycle—of paralysis by analysis.

## Coding & Variables

The last part of this qualitative analysis applies the codes extracted from the literature review to facilitate sensemaking. Some liken qualitative research to finding the “needle in a haystack,” an apt analogy. The SAR schedule database alone contains more than 3,000 records with over 585,000 words—an impossible task if one sets out to read, comprehend, and analyze without computer assistance.



Figure 4. Word Cloud for SAR Schedule

## Qualitative Methods Summary

This study used a qualitative analysis approach to automate the analysis of the SARs. A qualitative analysis approach potentially improves our understanding of schedule delay in ways beneficial to program managers. Part of this study used CAQDAS to perform a broad literature review of system development schedules and a more focused examination of schedule uncertainty. Both permitted us to refine the results of our 2018 research effort.

- The automation provided by the CAQDAS software significantly improves the ability to search and comprehend large amounts of unstructured data (both literature reviews and SARs data). We noticed improvements in the time necessary to do the research and the accuracy provided by the abovementioned tools. In fact, the software uncovered minor errors in the original manual study that had remained unnoticed.
- One of the difficulties in the original study was understanding the differences in terminology used by the different authors/program offices. For instance, some authors



would say, “The schedule was updated to reflect actual dates.” Others did not use that phrase but referred to the same idea of modifying the report to reflect the actual date something occurred. Still, others simply noted an update to the scheduled dates. The software provides the ability to identify these similar types of activities. Still, since they can be examined side-by-side, it gives the ability to be more accurate in assessment.

- Using uncertainty as the central theme for coding the documents provided insight into the different perceptions of program offices between uncertainty and risk. Uncertainty, as a code, occurred 27 times in the executive summary of the SAR and 26 times in the schedule change exclamation part of the SAR. A concern in the original manual analysis was the duplication of different terms. In other words, CAQDAS provided the ability to identify the existing programs and issues being discussed in the schedule context. This provided a check on the accuracy of the counts.
- The further coding of uncertainty required an understanding of the different authors’ uses of the word uncertainty and synonyms for uncertainty. For example, some authors describe situations as ambiguous in the context referring to uncertainty. As we continue our analysis across the complete SAR, understanding the actual terms being used will become one of the critical activities.
- Examination of the concept of uncertainty provided another further classification of schedule delay. We created a category to track whether the source of the delay was internal (by the PMO) or external (outside the PMO). Thirteen percent of the data records were classified as having causes external to the PMO. Examples of external change reasons included contract negotiation delays; contract completion delays; follow-on contract award delay; testing delays based on the availability of the testing unit; service changes in start dates; delays in contractor delivery, installation, and completion; and labor activities (strikes, etc.).
- This approach provides granularity in the administrative factors category. In the initial analysis, administrative reasons were more than 30% of the entire analysis. Administrative includes updates to APB, ADM changes, and changes resulting from Nunn–McCurdy processes, program restructuring, and a general category.

## Ex Post Analysis

Those who execute acquisition programs seldom clearly understand the situation’s dynamics. As Dörner (1989) put it, “A tendency, under time pressure, to apply doses of established measures. . . . An inability to . . . properly assess the side repercussions of one’s behavior” (p. 33), and “a tendency to think in terms of isolated cause-and-effect relationships” (p. 35).

Root Cause Analysis (RCA) is an ex post method intended to provide a detailed understanding of the process that led to a bad (or excellent) outcome. The intention is to discover how to remedy (or replicate) similar processes. RCAs set out to find the complete chain of events that led to the result in question (good or bad).

RCA is practiced in many fields, such as manufacturing process control, medical procedures, and accident investigations. It consists of answers to three questions.

- **What happened?** While the result is generally self-evident, the sequence of events that produced the result typically involves careful study.
- **Why did it happen?** What set of causal relationships ties those events together? The primary intent is to discover the chain of events to identify proximate and root (or underlying) causes.
- **What should we do about it, or what can we learn from it?** These “takeaways” involve measures to prevent (replicate) similar occurrences of bad (good) results.



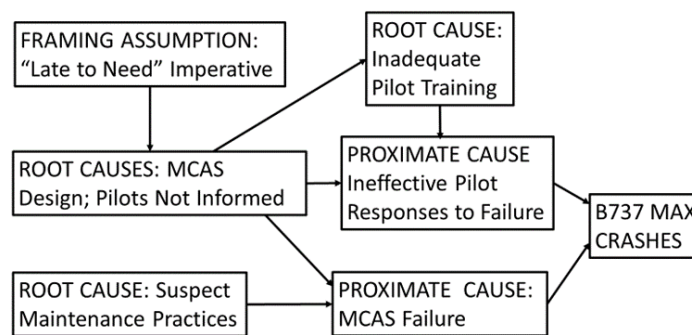
RCA of cost growth has also been practiced in defense acquisition programs, with a lead organization (e.g., OSD, PARCA) mandated by legislation.<sup>12</sup> Among its commissioned cost RCAs were studies of MDAPs undertaken by Institute for Defense Analyses (Arnold et al., 2010) and RAND (Blickstein et al., 2011).

One example from commercial aerospace is the fate of the Boeing 737 MAX airliner program. Program difficulties became evident with two catastrophic crashes (*what happened*)—in October 2018 and March 2019. Pickar and Franck (2020) took an acquisition-schedule perspective in their RCA.

*Why it happened* (from a development schedule perspective) is summarized in Figure 5. The decision (about 2010) to delay the Boeing 737 replacement program until 2020 allowed Airbus to steal a march in the form of a re-engined 320 family. The rapid commercial success of the new variant (A320neo) forced Boeing management to promise a new (more efficient) variant of the 737—promised for delivery at a highly optimistic date. Adapting the Boeing 737 airframe to accommodate new and larger turbofans on a tight schedule included (a) moving the engines forward and higher and (b) resolving the attendant aerodynamic complications with software fixes rather than airframe modifications.

Pickar and Franck (2020) traced the causal chain back to Boeing’s decision to delay the Boeing 737 replacement program: “The fundamental root cause of the current 737 MAX difficulties was a strategic miscalculation a decade in the past” (p. 11). Boeing had a severe problem in two parts: “(1) coaxing additional competitive life from a half-century-old design and (2) doing so in a manner responsive to the threat posed by the Airbus A320neo” (p. 11). In short, the 737 MAX program was begun “late to need.” Boeing could have avoided this problem by starting a 737 replacement program around 2010.

*What could also be learned* was the dangers of rigidly specifying total program times (Pickar & Franck, 2020, pp. 11–12): “aspirational” schedules, which “are driven by political and commercial processes and decisions.”<sup>13</sup> It is an example of making engineering development fit a strategy, rather than allowing the engineering (tasks) to define the time needed.



**Figure 5. A Root Cause Analysis of the Boeing 737 MAX Accidents of 2018, 2019 (Pickar & Franck, 2020, p. 9)**

<sup>12</sup> This function is now within the Office of Acquisition Data and Analytics.

<sup>13</sup> Acquisition programs starting “late to need” is not a rarity. The new US ICBM (GBSD) and Airborne Early Warning Aircraft (E-7) are contemporary examples.



## Concluding Comment

We believe current methods (even improved) acquisition schedule estimates are unlikely to solve this thorny problem completely.

We've discussed several reasons for this, including those summarized below.

- Those most knowledgeable about the program at the start are likely those most optimistic about the program. Furthermore, competitive bidding processes tend to reward that innate optimism. This can also result in “winner’s curse” outcomes in which the chosen supplier can be the most optimistic—and, therefore, the most likely to be inaccurate.
- MDAPs are systems with many moving parts that interact with each other in ways not known in advance and imperfectly known during program execution. This complexity seems inherent in defense acquisition programs.
- Program execution is a managed process in which management is expected to make decisions with trades (perhaps implicit) among schedule, cost, and performance. Each decision is generally impossible to predict and usually tricky to discern ex post.

A modest proposal: Given the innate degree of uncertainty in any program schedule, a helpful benchmark for a schedule estimate might be the answer to two questions. First, what's the schedule duration if everything goes right?<sup>14</sup> Second, how much do we wish to hedge that bet (through cost and schedule “reserves”)?<sup>15</sup>

## References

- Arnold, S. A., et al. (2010). *WSARA 2009: Joint Strike Fighter root cause analysis* (IDA Paper P-4612).
- Beaton, A., & Cohen, B. (2023, March 14). Want to win your NCAA tournament bracket pool? Pick this team. *Wall Street Journal*. <https://www.wsj.com/articles/march-madness-bracket-predictions-ucla-464d0524>
- Blickstein, I., et al. (2011). *Root cause analyses of Nunn-McCurdy breaches Volume I: Zumwalt-Class Destroyer, Joint Strike Fighter, Longbow Apache, and Wideband Global Satellite*. RAND Corporation.
- Complex system. (n.d.). In *Wikipedia*. Retrieved February 9, 2023, from [https://en.wikipedia.org/wiki/Complex\\_system](https://en.wikipedia.org/wiki/Complex_system)
- Demeulemeester, E. L., & Herroelen, W. S. (2006). *Project scheduling*. Springer Science & Business Media.
- DoD. (n.d.) *Washington Headquarters Services*. Retrieved April 6, 2023, from [https://www.esd.whs.mil/FOIA/Reading-Room/Reading-Room-List\\_2/Selected\\_Acquisition\\_Reports/](https://www.esd.whs.mil/FOIA/Reading-Room/Reading-Room-List_2/Selected_Acquisition_Reports/)
- Dörner, D. (1989). *The logic of failure* Translated by R. & R. Kimber, Pegasus Books.
- Dörner, D. (1997). *The logic of failure*. Basic Books.
- Drabkin, D. (2019). Section 809 Panel recommends reforming defense acquisition with updated structures, simplified procedures and an empowered workforce. *DAU News*, 31.
- Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. *Project Management Journal*, 37(3), 5–15.
- Flyvbjerg, B., & Gardner, D. (2023). *How big things get done*. Currency.
- Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Hayek, F.A. (1945). American economic review. In H. Townsend (Ed.), *Price theory: Penguin modern economics readings* (pp. 17–31).
- Hearst, M. (2003). *What is text mining?* University of California, Berkeley.
- Kerzner, H. R. (2013). *Project management*. John Wiley.

<sup>14</sup> The DoD tends to get these anyway.

<sup>15</sup> We view this as more “truth in advertising” than a new approach.



- Klein, G. (2004). *The power of intuition: How to use your gut feelings to make better decisions at work*. Currency.  
<https://books.google.com/books?hl=en&lr=&id=okyNDQAAQBAJ&oi=fnd&pg=PP15&dq=the+power+of+intuition&ots=eRTyAR5S3r&sig=-xqpMc0W6C17YYG4fYxW-a-ltxl>
- Knight, F. H. (1921). *Risk, uncertainty, and profit*. Houghton Mifflin Company.  
<https://cir.nii.ac.jp/crid/1573387450829321344>
- Koleczko, K. (2012). Risk and uncertainty in project management decision-making. *Public Infrastructure Bulletin*, 1(8), 1–8.  
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fbbe5a5a8da599f31b641fa78ca9a3bda44817dd>
- Light, T., et al. (2018). *Benchmarking schedules for major defense acquisition programs* (Document No. RR-2144-AF). RAND. <https://doi.org/10.7249/RR2144>
- Loxton, M. H. (2021). *Using MAXQDA for strategic foresight*. MAXQDA Press.
- Mann, A. (2016, October 20). Market forecasts: Prediction markets can be uncannily accurate. *Nature*, 308–310. <https://www.doi.org/1038/538308a.PMID277622382>
- McCleary, R. M. (2023, January 27). The man who tried to stop the Space Shuttle Challenger’s launch. *Wall Street Journal*. <https://www.wsj.com/articles/the-man-who-tried-to-stop-the-challenger-launch-space-shuttle-exploration-roger-boisjoly-moral-injury-11674857494>
- Miller, T. P. (2012). *Acquisition program problem detection using text mining methods* [Thesis, Air Force Institute of Technology].
- Pickar, C. K. (2018). *Informing DoD program planning through the examination of the causes of delays in acquisition using acquisition data*. [https://calhoun.nps.edu/bitstream/handle/10945/58762/SYM-AM-18-079-019\\_Pickar.pdf?sequence=1](https://calhoun.nps.edu/bitstream/handle/10945/58762/SYM-AM-18-079-019_Pickar.pdf?sequence=1)
- Pickar, C., & Franck, R. (2020). *It’s about time: Toward realistic acquisition schedule estimates* (SYM-AM-20-051). *Proceedings of the Seventeenth Annual Acquisition Research Symposium*. <https://event.nps.edu/conf/app/researchsymposium/unsecure/diprop646/52>
- Pickar, C., & Franck, R. (2021). *Telling time: Getting relevant data for acquisition schedule estimating relationships* (NPS-AM-22-013). Naval Postgraduate School.
- Pickar, C., & Franck, R. (2022). *Time for a change: Data-driven schedule relationships* [Paper presentation]. Western Economic Association Annual Conference.
- Project Practical. (n.d.). *Contingency reserve vs management reserve: PMP concept*. Retrieved March 1, 2023, from <https://www.projectpractical.com/contingency-reserve-vs-management-reserve/>
- Schulz, K. (2010). *Being wrong: Adventures in the margin of error*. HarperCollins.
- Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations. In E. Canaan (Ed.), *Modern library*.
- Surowiecki, J. (2004). *The wisdom of crowds*. Doubleday.
- Sweetman, B. (2012, November 26). Generation gap. *Aviation Week*, 22–23.
- Torenberg, T. (2021). *A primer on prediction markets: Ideas and musings*. <https://eriktorenberg.substack.com/p/a-primer-on-prediction-markets>
- University of Iowa. (n.d.). *What is the IEM?* Retrieved January 23, 2023, from <https://iemweb.biz.uiowa.edu/media/summary.html>



# Training an Agile Acquisition Workforce to Combat Emerging Threats

**Amanda Swanson Goff**—joined The Pulse in 2022 as the Director of Research and Analysis. Amanda received her Bachelor's in Linguistics from the University of Kansas (KU) and her Master of Public Policy from George Washington University (GWU). In May of 2024, she will graduate from Vanderbilt University with a Doctor of Education in Organizational Learning. She brings 6+ years of experience in project management, data analytics, and stakeholder engagement. Amanda is especially adept at using publicly-available federal spending data to inform policy work and identify key trends impacting the GovCon industry.

## Abstract

The Department of Defense (DoD) current source selection methods are at an increased risk of experiencing sustained bid protests. During source selections, the government frequently contradicts itself between its advertised stated order of importance for acquisition evaluation criteria (pre-award) and its actual choice behavior during source selections (Butler, 2014). This paper provides a summation of research, conducted from 2021 to 2022, that explored the following research objectives: 1) Determine the degree of disconnect between stated preferences during pre-award acquisition phase and actual choice behavior in defense acquisition source selections, 2) develop a deep understanding of quality attributes in evaluating logistics-based service acquisitions, 3) provide a Choice-Based Conjoint (CBC) framework that the DoD could utilize to enhance source selection criteria development in both logistics and further categories of government spending. The research utilized methods such as interviews and spend analysis techniques to identify quality attributes of logistics-based acquisitions that would best discriminate as evaluation factors for award. Later, these attributes were used to develop a CBC exercise that enabled us to calculate attribute utilities and relative importance for each attribute. The summarized research in this paper provides a way forward to empirically deduce the relative importance for source selection evaluation factors, potentially reducing bid protest occurrences in future source selections.

## Introduction

In FY22 alone, the Department of Defense (DoD) spent nearly \$420 billion on contracted products and services to support mission needs around the globe. Procurement professionals executed contracts for everything from janitorial services to major weapons systems. However, regardless of size and scope, the procurement process was often slow and cumbersome. Contract execution from solicitation to award is taking months, and in some cases, years. Resultantly, the United States will not be able to keep pace with evolving global threats.

Particularly as the DoD shifts its focus from the conflict in Afghanistan to near-peer threats in China and Russia, the Department's ability to quickly acquire emerging technology will be paramount to its success. This will require working with non-traditional contractors using Other Transactions Authority (OTAs), Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) grants, and lesser-known authorities like DoD Section 2373. Successful implementation of more flexible procurement authorities will require that procurement professionals across the DoD are well-trained and empowered to make decisions with due speed. Achieving this level of competency among the Department's procurement workforce necessitates a new approach to training that employs the most effective techniques for maximizing knowledge retention.

## The DoD's Acquisition Workforce

As of September 2022, the DoD employed 41,374 contracting professionals as classified under the Office of Personnel Management's (OPM) 1102 occupation code. This represents just





1.8% growth in the Department’s acquisition workforce since September 2021. Over the same period, the DoD’s contract obligations grew 7.38%.

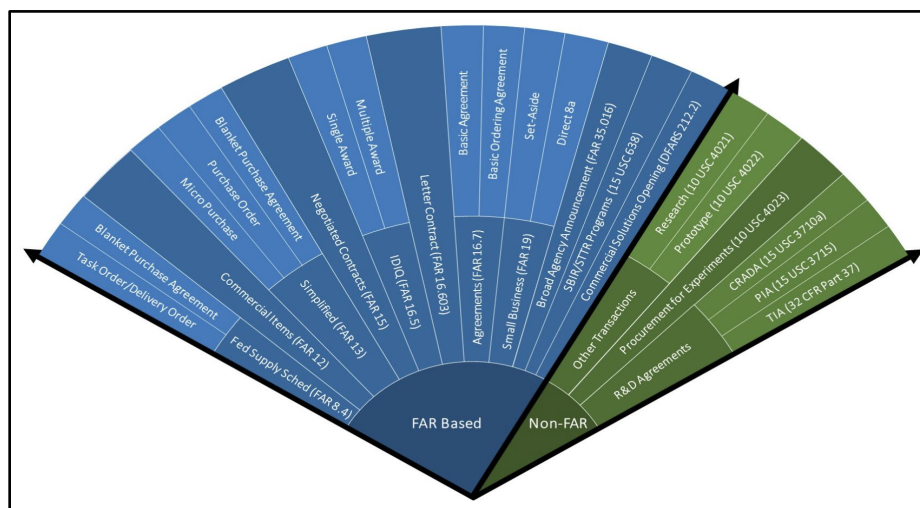
Fiscal Year	Total Contract Obligations	Obligation Growth	Total 1102 Workforce	Workforce Growth
2018	\$369,977,711,948.32	-	38,226	-
2019	\$396,698,449,057.20	7.22%	39,388	3.04%
2020	\$429,828,139,793.80	8.35%	39,904	1.31%
2021	\$390,852,687,172.09	-9.07%	40,619	1.79%
2022	\$419,693,725,077.08	7.38%	41,374	1.86%

Data over the last five fiscal years indicate a considerable gap between growth in the Department’s contract obligations and the number of contracting professionals available to manage those contracts. The growing burden on contracting professionals necessitates effective training that empowers them to be efficient and agile in their work. However, the Department continually fails to empower contracting professionals to leverage flexible contracting authorities that would improve access to critical and emerging technologies.

Given existing workforce recruitment and retention challenges across the federal enterprise, the burden on individual contracting professionals is likely to worsen. Over reliance on cumbersome, traditional contracting methodologies is unsustainable—both for the professionals executing DoD contracts and the service members who rely on the products and services being purchased.

### What Alternative Acquisition Procedures are Available to DoD Procurement Professionals?

According to the Defense Acquisition University (DAU), the DoD has 17 different statutory authorities that may be leveraged in the procurement of goods and services. Six of these authorities exist outside the constraints of the Federal Acquisition Regulation (FAR).



Well-known alternative contracting methods include SBIR and STTR grants, as well as OTAs. In FY19 alone, the DoD obligated \$1.7 billion in SBIR and STTR grants.<sup>1</sup> The Department obligated an additional \$7.4 billion through OTAs. Although these two contracting methodologies comprised only 2% of the DoD's contract obligations in FY19, they are well known in government and industry.

Commonly, SBIR/STTR grants as well as OTAs and other non-traditional contracting methods are used for research and development of emerging technologies. Often, the Department faces significant barriers to procuring these technologies once development is complete.

DoD Section 2373 (USC §4023), Procurement for Experimental Purposes, provides a crucial opportunity for the Department to procure and test the effectiveness of these new technologies. The authority provided therein enables the DoD to put the equipment directly in the hands of service members, allowing for expedited evaluation of its usefulness in meeting mission needs.

Section 2373 is just one example of how alternative contracting authorities may be leveraged to better access the latest and greatest in warfighting technology. But unfortunately, most contracting professionals across the DoD seem unaware of or unwilling to use these additional authorities. Ultimately, empowering contracting professionals to think creatively requires more frequent and higher quality training in how to use alternative authorities correctly and strategically.

## **Strategies for Designing Effective Training**

Decades of research in cognitive psychology indicate that designing effective training hinges on four things: the right length, the right timing, the right structure, and the right level of engagement. In the case that the Department continues to opt for internal training design and execution through the DAU, it is crucial that DoD officials are aware of and make use of these tenets to maximize knowledge retention and facilitate knowledge application by contracting professionals.

### **Selecting the Right Training Length**

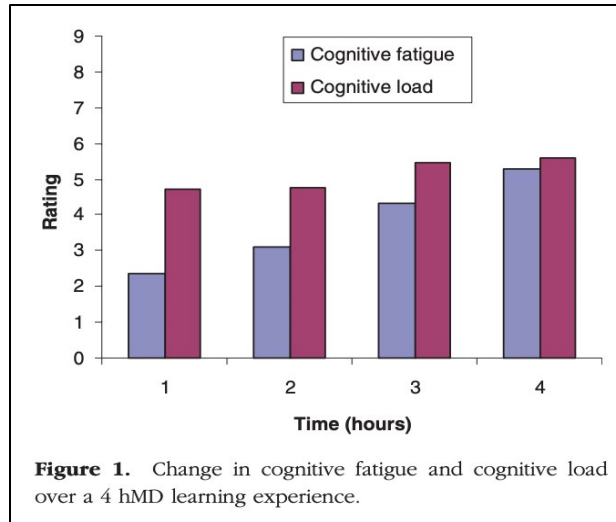
Traditionally, training is often conducted over two or three days wherein participants spend upwards of eight hours per day receiving instruction. Simply, the research does not support this.

Studies have shown that the human brain cannot effectively maintain focus for more than 60 minutes at a time. In a 2010 study (Raman et al., 2010), researchers found that “over time, fatigue increases with an increase in cognitive load.” In simple terms, as learners are exposed to more information, their brains get “tired.” This implies diminishing returns over the course of the day as participants continue receiving information that they are unable to process and retain.

---

<sup>1</sup> Retrieved from [https://www.sbir.gov/awards/annual-reports?program=STTR&abbr%5B%5D=DOD&view\\_by=Year](https://www.sbir.gov/awards/annual-reports?program=STTR&abbr%5B%5D=DOD&view_by=Year)





Effectively training DoD contracting professionals will require a fundamental shift in how the federal government thinks about human capital and professional development.

### Proper Spacing Between Training Sessions

Too often, training is conducted once annually during slower periods for the business. The first quarter of the fiscal year would represent one such period, wherein the prohibition on new starts included in continuing resolutions brings contracting activities to a halt. However, psychological research indicates a need for more frequent instruction. Ultimately, training that is conducted too close together or too far apart results in poor information retention.

Ideal minimum spacing between learning sessions covering the same kind of content is 12 hours (Kornmeier et al., 2022). Further, some studies have demonstrated better long term retention of information when learning is spaced by one month or more.

The new Federal Acquisition Certification in Contracting (FAC-C) structure released by the Office of Federal Procurement Policy (OFPP) in January 2023 is a step in the right direction (OFPP, 2023). Among other structural changes, the new certification includes an increased focus on continuous learning (CLE).

Certification holders must now complete 100 hours of additional coursework every two years, an increase from the previously required 80 hours. Even so, OFPP does not require specific subjects to be covered as part of that continuing education. Without clear guidance from DoD officials as to how contracting professionals may meet CLE requirements, the Department will miss out on a critical opportunity to ensure its acquisition workforce is up to date on the most current procurement practices and authorities available to them.

Further, OFPP guidance does not include requirements for spacing out CLE hours. As a result, there is nothing that prevents Department leadership from insisting that all 100 CLE hours be completed in a short period of time, eliminating the value of spaced learning and likely resulting in hours-long training sessions. Ultimately both results greatly diminish the value of continuing education.

### Implementing Retrieval Practice

Encouraging participants to recall information they've already learned is another crucial component of effective training design. A 2012 study (Karpicke, 2012) found that the use of



repeated retrieval practice was as much as four times more effective than studying the information one time.

Often, retrieval practice is assumed to be a “pop quiz” of sorts. However, retrieval practice simply implies an opportunity to recall information that was presented previously. Facilitators may consider Dr. Pooja Agarwal’s “Two Things” Methodology, wherein learners are asked to write down two things that come to mind in response to a prompt.<sup>2</sup>

Example prompts include:

- Write down two questions you have related to the previous session’s material.
- What are your two biggest takeaways from the previous session?
- Can you find two connections between today’s material and yesterday’s?

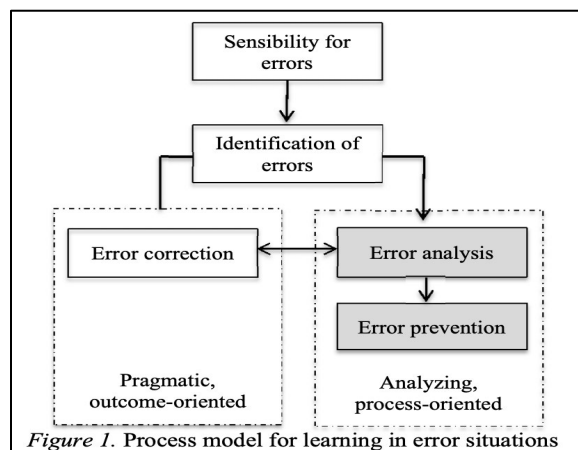
The Department may also consider applied learning strategies, wherein learners are asked to utilize information from a previous session to address a problem identified in a case study. Any such activity that prompts learners to recall and apply information will increase knowledge retention and contribute to a more ready acquisition workforce.

### Encouraging Active Participation

Finally, effective training requires actively engaged participants. This includes both trainers and trainees. Simply sitting in a classroom where information is being disseminated does not ensure comprehension and retention.

In particular, classroom studies indicate that encouraging errors in the learning process and actively responding to them in a positive but instructive way can contribute to increased learning. This is especially important for contracting professionals who are widely considered to be mistake-avoidant.

Practical applications of a session’s material may occur in the large group or in smaller breakout sessions, but should always include feedback from the facilitator. The figure included below illustrates how the feedback cycle contributes to the learning process, ultimately resulting in increased error prevention (Rach et al, 2013).



Further, all attendees should be expected to participate regardless of title or experience. Active participation from senior professionals can encourage younger, less experienced participants to engage with the group more meaningfully.

<sup>2</sup> Retrieved from <https://www.retrievalpractice.org/strategies/2018/two-things>.

As the DoD acquisition workforce continues to age, engagement from experienced professionals will become increasingly important in the preservation of institutional knowledge. The inclusion and active participation of these individuals is paramount in creating a more nuanced understanding of the procurement process among younger professionals.

## Conclusion

The Federal procurement process is slow and cumbersome, operating in stark contrast to the rapidly evolving needs of the Department of Defense. To maintain readiness and address near-peer threats from Russia and China, DoD contracting professionals must operate with greater agility to access cutting edge technologies. And in order to do so, they must be well trained and feel empowered to leverage alternative procurement authorities. That is not currently the case.

Simply, the existing training infrastructure at the DoD is not well-suited to develop an agile acquisition workforce. The Department must shift from the traditional, certification-based training system to one that focuses on shorter, more frequent training that addresses new procurement processes and lesser known contracting authorities. Ultimately, the Department's failure to leverage all available contracting methodologies will limit its ability to equip the warfighter of the future. The antiquated training infrastructure in use at the DoD is not sufficient to produce well trained contracting officers, and until corrected, the DoD's procurement workforce will be unable to acquire critical technologies with appropriate speed—thus leaving the country vulnerable.

## References

- Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, 21(3), 157–163.
- Kornmeier, J., et al. (2022). Spacing learning affects both learning and forgetting. *Trends in Neuroscience and Education*, 26.
- OFPP Memorandum. (2023, January 19). <https://www.whitehouse.gov/wp-content/uploads/2023/01/FAC-C-Modernization-Memorandum-19-Jan-2023.pdf>.
- Rach, S., et al. (2013). Learning from errors: Effects of teachers' training on students' attitudes towards and their individual use of errors. *PNA*, 8(1), 21–30.
- Raman, M., et al. (2010). Teaching in small portions dispersed over time enhances long-term knowledge retention. *Medical Teacher*, 32(3), 250–255.



# Through the Looking Glass: Why EVM Is an Essential Risk Mitigation Measure for Decision Makers and Program Managers

**Symantha “Sam” Loflin**—has an MS in program management and a 2020 certification in Advanced Acquisition Studies from the Naval Postgraduate School (NPS), where she was a contributor to the 18th and 19th Annual NPS Acquisition Research Symposium. She also holds a BS in finance from the University of Houston. Loflin is a PM/EVM SME at Manufacturing Technical Solutions (MTS)/Aeyon and is supporting NASA’s Earned Value Management Program Executive, OCFO-SID. She has over 20 years of acquisition experience supporting the DCMA, the DoD, NASA, and the military services. She recently served as an acquisition program manager on the Coronavirus Task Force that focused on building the industrial base for personal protective equipment. Loflin’s career began at NASA, supporting the Space Shuttle, ISS, and the Constellation Programs in Houston. [symanthaloflin@gmail.com]

## Abstract

The author has written this paper to defend and strengthen the use of risk mitigation measures inherent in the implementation, maintenance, and surveillance of a government contractor’s business system (CBS)<sup>1</sup>, specifically the contractual addition of the Earned Value Management System (EVMS; DFARS 252.234-7002) clause.<sup>2</sup> The contractor and the United States Government’s ability to engage the **right people, processes, and tools at the right time** is essential to effective program management policy and control (Hite, 2010. p. 23).<sup>3</sup> This course of action will provide the stakeholders with the capabilities required to plan and execute the program/project<sup>4</sup> to support **proper stewardship of taxpayer dollars**. The projects performance outcomes provide lessons learned “through the looking glass” that will ensure objective and rationale for the measurements of reliability, availability, and maintainability<sup>5</sup> of a compliant EVM business system. These risk mitigation measures objective is to **reduce project cancellations, strengthen national security** and build the **domestic industrial base** sourcing of goods and services.

## Research Issue

Why is it imperative for program managers and decision makers to use EVM as a risk mitigation measure?

## Research Results Statement

It takes a whole-of-government approach by Department of Defense (DOD), National Aeronautics and Space Administration (NASA), and General Services Administration (GSA) to defend and protect the world’s dependence on the sea, air, and space. By embracing EVM as a risk mitigation measure, through legal and regulatory processes, the federal government’s procurement of Made in America<sup>6</sup> products will promote economic opportunities to grow American small businesses that strengthen the Defense Industrial Base (DIB) and promote the national security strategy. The implementation and use of EVMS, system surveillance, and Integrated Baseline Reviews (IBRs), through the roles of the contracting officer, program

<sup>1</sup> Contractor business system and CBS are used interchangeably throughout the document.

<sup>2</sup> DFARS 252.242-7002, Earned Value Management System (2011, May). [https://farclause.com/FARregulation/Clause/DFARS252.234-7002\\_Basic-earned-value-management-system#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.234-7002_Basic-earned-value-management-system#gsc.tab=0)

<sup>3</sup> Hite, Randolph C. (2010, August). Organizational Transformation: A Framework for Assessing and Improving Enterprise Architecture Management (Version 2.0). (GAO-10-846G), Government Accountability Office.

<sup>4</sup> Program/project will be used interchangeably throughout document.

<sup>5</sup> DOD Guide for achieving Reliability, Availability, and maintainability, August 3, 2005

<sup>6</sup> White House (2021a, January 25). Executive Order 14005. Section 4. (a) *Ensuring the Future Is Made in All of America by All of America's Workers*.

<https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/25/executive-order-on-ensuring-the-future-is-made-in-all-of-america-by-all-of-americas-workers/>



manager, functional specialist, and additional stakeholders are essential to a fundamental assessment of the contractor's performance. These measures enhance risk mitigation by controlling cost, schedule visibility, and enhancing technical readiness.

Each year, the federal government increases the funding of developmental contracts as a measure to "Protect Sea, Air, and Space" (National Security Strategy, 2022, October 12).<sup>7</sup> These efforts aim to protect the United States interests in developing technologies, creating economic opportunities, and enabling climate surveillance, and to responsibly oversee the space environment. In addition, value is added to the taxpayer, federal government, and the National Security Strategy as follows:

*Biden-Harris Administration's National Security Strategy,*

---

*By enhancing our industrial capacity, investing in our people, and strengthening our democracy, we will have strengthened the foundation of our economy, bolstered our national resilience, enhanced our credibility on the world stage, and ensured our competitive advantages. (Biden-Harris Administration's National Security Strategy, 2022)*

---

**Three recommendations for consideration:** **1.** Require and provide continuous learning to all stakeholders (government and contractor). Research shows that subject matter knowledge increases the likelihood of successful problem resolutions "...experiential learning offers a way to ensure we are imparting not just rote learning and certifications but providing our people the knowledge, skills, and experience to effectively control the efforts we charge them to lead..."<sup>8</sup> **2.** Promote regular and reoccurring EVMS surveillance throughout the project's lifecycle, reveals whether the internal controls and business management practices comply to the 32 guidelines. **3.** Implement a federal government consortium between all stakeholders from Federal Agency's, Services, and Industry, as this will provide effective and efficient cross functional lessons learned.

The results are clear that the federal measures enacted by The Banking Act of 1933 and the creation of the EVM 32 guidelines criteria for industrial management systems in 1967 promote stakeholder confidence. It takes a whole-of-government approach to provide government oversight and insight to mitigate fiscal financial risk and ensure the stewardship of taxpayer resources.

## Introduction

This paper analyzes the whole-of-government approach to demonstrate that EVM, a program management best practice tool, is a key principle of risk management, budgeting, contracting, and capital asset acquisition, performance-based management requirement of federal agencies, as noted in the Office of Management and Budget (OMB) Circular A-11, Federal Acquisition Regulation (FAR), Subpart 34.2 requires that EVM systems comply with the Electronic Industries Alliance (EIA) Standard 748, and Defense Federal Acquisition Regulation Supplement (DFARS), Subpart 234.201,<sup>9</sup> which allowed for more stringent government oversight of contractors. The USA spending for all budget function for FY 2023 is \$3.7 Trillion.

---

<sup>7</sup> Biden-Harris Administration's National Security Strategy (2022, October 12). Biden-Harris Administration's National Security Strategy.pdf (whitehouse.gov)

<sup>8</sup> Pickar, Charles. (2020), Naval Postgraduate School, Acquisition Research Symposium. *Learning from Experience: Acquisition Professional Education for this Century*, SYM-AM-20-070.pdf (nps.edu)

<sup>9</sup> National Defense Industry Association. (2016, October). U.S. Federal Agency EVMS Policy Summary. Federal Agency EVM Policy Summary (ndia.org)



The total USA spending is shown by state geography infographic from the official website is in Figure 1.<sup>10</sup>

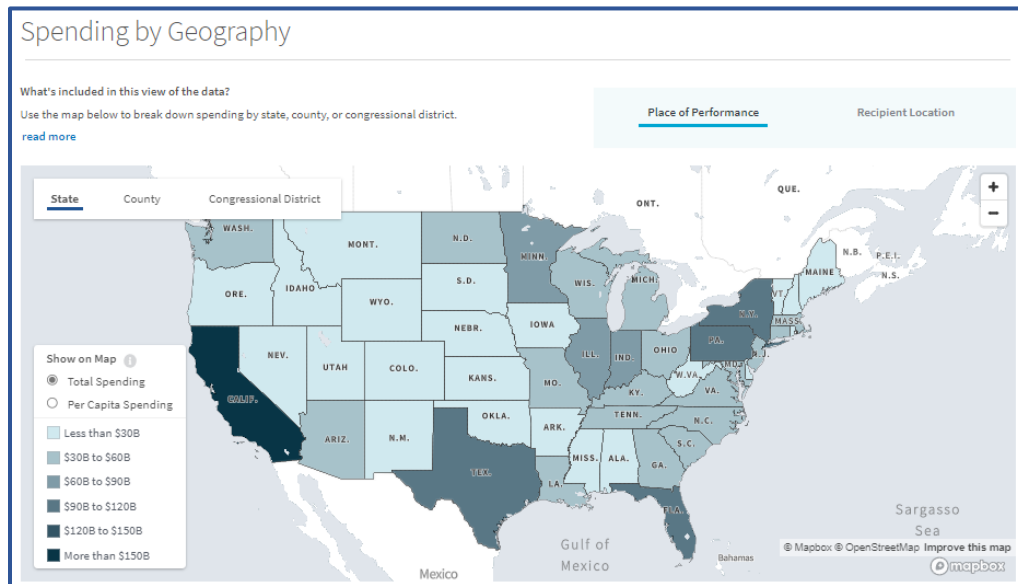


Figure 1. USA Spending by State Infographic, FY 2023

## Federal Risk Mitigation Measures: “Protect Sea, Air, and Space”

The Banking Act of 1933 and Earned Value Management (EVM) 1967

In 1929, the lessons learned from bank failures, during the Great Depression, prompted the swift action of the United States Federal Government to pass the Glass-Steagall Act “The Banking Act” of 1933, which established the Federal Deposit Insurance Corporation (FDIC), that was signed into law by President Franklin D. Roosevelt. These federal laws, regulations, and government oversight “bank examinations [surveillance]” measures restored America’s confidence in the banking system and the federal government.<sup>11</sup>

In 1967, like the government oversight laws in banking, EVM was pioneered “...by [industry] contractors and later Air Force executive A. Ernest “Ernie” Fitzgerald and Air Force Lt. Col. Hans “Whitey” Driessnack, captured industry’s best management practices [restore confidence in project cost and schedule status] and issued them in defense policy not as “how to” requirements but as 35 [now 32 guidelines] criteria for industrial management systems.<sup>12</sup>

For more than 50 years, EVM is virtually unchanged, as it captures program managements best practices, issued through defense policy, as the criteria for 32 guidelines with the DOD Instruction 7000.2 [later DoDI 5000.2], Cost/Schedule Control Systems Criteria (C/SCSC).<sup>13</sup> The responsibility and requirements of government surveillance of contractors EVM business system is a critical component of EVM. The quantitative and qualitative surveillance analysis results shows whether data is current, accurate, timely, and reliable. Surveillance reveals the effectiveness of the contractor’s implementation and sustainment of compliance laws, policies, and internal management controls.

The historical quantitative metric analysis results to the qualitative recommendations are documented and retained that revealed valuable insight into the integrity of the project’s current

<sup>10</sup> USAspending.gov. (2023, April 4). *Spending by Geography*. Federal Awards | Advanced Search | USA spending. <https://www.usaspending.gov/>

<sup>11</sup> Federal Reserve History. (1933), Banking Act of 1933 (Glass-Steagall). Banking Act of 1933 (Glass-Steagall) | Federal Reserve History

<sup>12</sup> Abba, Wayne F. (2017), Defense AT&L: March-April 2017. *The Evolution of Earned Value Management*. Defense AT & L Magazine (dau.edu)

<sup>13</sup> Abba, Wayne F. (2017), Defense AT&L: March-April 2017. *The Evolution of Earned Value Management*. Defense AT & L Magazine (dau.edu)





cost and schedule status. The analysis results are used as a predictor of future cost and schedule growth compared to the past. In addition, the results will reveal any material weaknesses in data integrity that may require the projects to provide a root cause for the corrective action(s) with a plan that includes preventive measures to reduce the risk of reoccurrence.

### **FAR Rule Part 34.201(a): EVMS, Integrated Baseline Review (IBR), and Surveillance 2006<sup>14</sup>**

On July 5, 2006, The Federal Register issued a FAR rule by the DoD, GSA, and NASA that amended FAR Part 34.201(a), Major System Acquisition, as follows:

to implement Earned Value Management System (EVMS) policy in accordance with OMB Circular A-11, Part 7 and the supplement to Part 7, the Capital Planning Guide... shall conduct an Integrated Baseline Review (IBR), ...Contracting Officer or a duly authorized representative as necessary to permit Government surveillance to ensure that the EVMS conforms, and continues to conform, with the performance criteria...

### **EVM Newsflash: Low Cost and High Value**

EVM requirements promote sound planning and effective program execution, but baseless claims that EVM is high cost and low value are the norm among non-EVM practitioners. In the *Defense AT&L* January-February 2017 publication of “EVM Systems are High Cost - FACT or FICTION?” the authors reveal the results of the Joint Space Cost Council Better EVMS Implementation Survey conducted with industry contractors, DCMA, and the DoD Office of Performance Assessment and Root Cause Analyses.

The survey results from Phase I in 2015 and Phase II in 2016 verifies and validates that **EVM is low cost and high value to program managers as shown in Figure 2**. Authors: Ivan Bembers, National Reconnaissance Office (NRO) EVM Focal Point, Earned Value Management (EVM) Center of Excellence (ECE). He is responsible for EVM system acceptance and surveillance reviews, facilitating integrated baseline reviews and supporting programs’ use of EVM across the enterprise. Ed Knox, Michelle Jones and Jeff Traczyk support EVM at the NRO.<sup>15</sup>

---

<sup>14</sup> Federal Acquisition Regulation. (2006, July 5). Earned Value Management System (EVMS) (FAR Case 2004-019). Federal Register. <https://www.govinfo.gov/content/pkg/FR-2006-07-05/pdf/06-5966.pdf>

<sup>15</sup> Bembers, Ivan, Knox, Jones, Traczyk. (2017, January-February). *Defense AT&L, EVMS System’s High Cost-Fact or Fiction?* Defense AT & L Magazine (dau.edu)



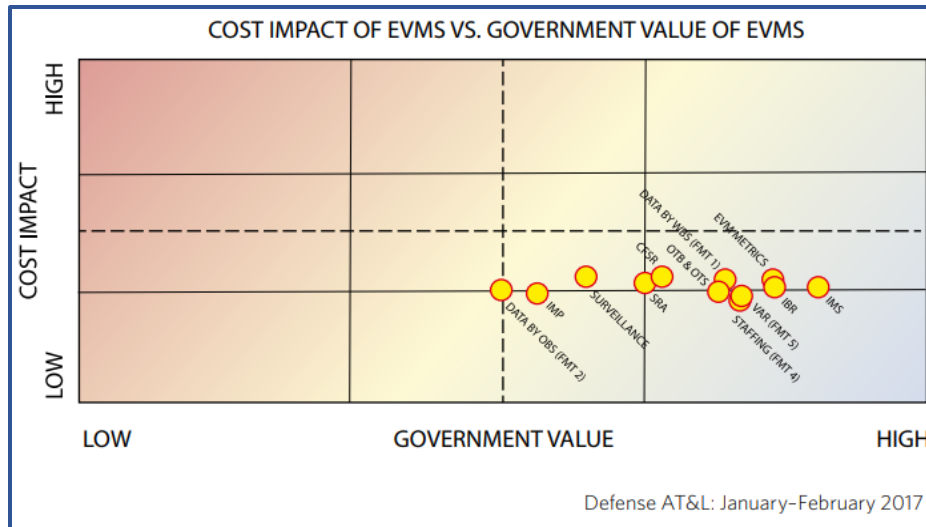


Figure 2. Cost Impact of EVMS vs. Government Value of EVM (2017)

### National Security in Space

On September 4, 2022, at the U.S. Chamber of Commerce’s Global Aerospace Summit, leaders from industry, space, and aviation sectors collectively discussed how the development of space technologies, barriers in space exploration, impact national security, sustainability, and life on earth, as noted by Alexander MacDonald, NASA’s chief economist, in Figure 3.



Figure 3. Left to Right: Scott Pace, PhD, Former Executive Secretary of the National Space Council; Roy Azevedo, President, Raytheon Intelligence and Space; and Alexander MacDonald, Chief Economist, NASA (2022)

The Outer Space Treaty, signed by President Johnson, enforced the U.S. commitment to peaceful use of outer space exploration.<sup>16</sup> In October 1967, the **Outer Space Treaty** was signed by three collective Governments (the **Russian Federation**, the **United Kingdom**, and the **United States of America**), entered into force under the United Nation Office for Outer

<sup>16</sup> United Nations Office for Outer Space Affairs. (1967, October). Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies. The Outer Space Treaty (unoosa.org)

Space Affairs. The treaty provides the basic framework on international space law with principles.

MacDonald emphasized that NASA’s return to the moon with the Artemis program is a national security measure that shows the world the capabilities of the United States. MacDonald said,

“...in order to maintain our security in a more general sense in Earth’s orbit, we need to maintain it as a place of rules, of behavior, [and] of norms,” he said. “In that regard, I think the support of the U.S. by the UN for a resolution around ASAT tests is incredibly important because that’s a commitment to a peaceful, lower orbit environment, which we all know is needed for business.”<sup>17</sup>

### NASA – 2019 Economic Impact \$64.3 Billion<sup>18</sup>

In 1958, NASA was established with an initial funding of \$330M and in FY2023 NASA’s budgetary resources are \$32.35 Billion. NASA reported that in Fiscal Year (FY) 2021, that the diverse workforce is comprised of just under 18,000 civil servants and supports more than 312,000 [prime contractors/subcontractors (large and small businesses)] jobs across the United States. In addition, the workforce provides an extensive knowledge base that includes academia and international and commercial partners, which provide a benefit to humanity.<sup>19</sup> The Agency’s Mission is to “Drive advances in science, technology, aeronautics, and space exploration to enhance knowledge, education, innovation, economic vitality, and stewardship of Earth,” and the funding details of the FY 2023 Summary of NASA’s \$32.35 Billion in budgetary resources infographic Figure 4.<sup>20</sup>

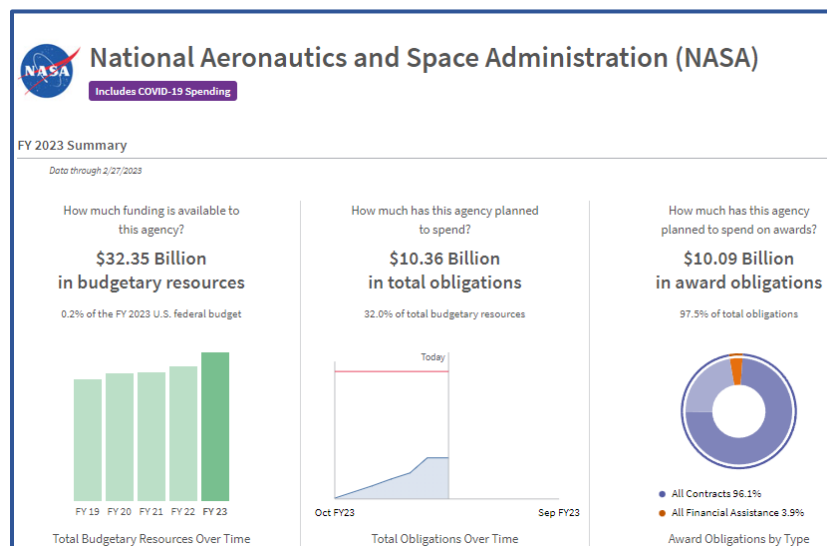


Figure 4. National Aeronautics and Space Administration (NASA) Budgetary Resources, FY 2023

<sup>17</sup> U.S. Chamber of Commerce. (2022, September 14). How the U.S. Is Establishing National Security in Space. How the U.S. Is Establishing National Security in Space | U.S. Chamber of Commerce (uschamber.com)

<sup>18</sup> NASA. (2020, September 25). Release 20-093. *NASA Report Details How Agency Significantly Benefits US Economy*. NASA Report Details How Agency Significantly Benefits US Economy | NASA

<sup>19</sup> NASA. (Fiscal Year 2019). About NASA. About NASA | NASA

<sup>20</sup> NASA. (2023, February 27). FY 2023 Summary, National Aeronautics and Space Administration (NASA) | Spending Profile | USA spending. <https://www.usaspending.gov/>



## People, Processes, and Tools

### Small Business - EVM – NASA

On November 16, 1998, NASA enhanced the federal governments oversight to mitigate fiscal financial risk, as noted above by the banking laws of 1933, by updating the "...NASA FAR Supplement relative to the application of Earned Value Management (EVM) at NASA. The proposed change would establish NASA-wide clauses and provisions compatible with those used by DoD..."<sup>21</sup> The Agency's insight cost for people, processes, and tools to support EVM compliance, IBR, and surveillance is negligible compared to the extraordinary benefits the analytical results provide to all stakeholders.

NASA's EVMS compliance and surveillance are functions that support, small business, key acquisition priorities, and are vital to effective mission management. The projects' cost and schedule status are reflected in the quantitative metric analysis to the qualitative recommendations from the EVM compliance and surveillance subject matter experts (SME)s. The **SMEs** are comprised of civil servants and employees on **small business contracts** e.g., Consolidated Program Support Services (CPSS) Program, Planning, and Control (PP&C).<sup>22</sup> The overall surveillance objective is to ensure that management control processes support the projects Program Management Baseline (PMB) compliant with the EIA-748 EVM System guidelines.<sup>23</sup>

There are 32 EVM compliance guidelines (GL) that are divided into five categories: Organization (GL 1-5); Planning, Scheduling, and Budgeting (GL 6-15); Accounting Considerations (GL 16-21); Analysis and Management Reporting (GL 22-27); and Revisions and Data Maintenance (GL 28-32). Surveillance is conducted by performing 142 DCMA EVMS Compliance Metrics (DECM) over a 3-year period. The metrics provide an objective methodology to assess the health of the EVMS. The implementation and validation of a project's EVMS along with surveillance increases cost and schedule visibility for stakeholders. The reliance on the right EVM people, process, and tools by the PMs and decisions makers is a risk mitigation measure that reduces project cancellations, strengthens national security, and builds the domestic industrial base sourcing of goods and services.

### Small Business Encore Analytics Software Tool

Government Agencies, Military Services, Academia, Corporate Industry

The EVM analysts use manual and automated software tools contracted primarily from Encore Analytics that is a verified Veteran-Owned Small Business (VOSB) by the U.S. Department of Veterans Affairs. The Empower software is aligned to the Defense Contract Management Agency's DCMA EVMS Compliance Metrics (DECM) that provides automated analytic results for 73 of the 142 metrics. Encore Analytics' current customer base is comprised of more than 49 government, federal agencies, military services, academia, and corporate industry leaders. The most notable customers include NASA, the Department of Energy (DoE), the National Reconnaissance Office, the National Security Agency, NAVAIR, the U.S. Army, the U.S. Marine Corps, Lockheed, Raytheon, Northrup Grumman, other federal agencies, Defense Acquisition University (DAU), and industry.



<sup>21</sup> Federal Register. (1998), Vol. 63, No. 220 / Monday, November 16, 1998 / Proposed Rules 63655, 98-30554.pdf (govinfo.gov)

<sup>22</sup> NASA. (2021, May 12). Contract Release C21-011. NASA has selected Manufacturing Technical Solutions Inc. of Huntsville, Alabama. NASA Selects CPSS PP&C Contractor | NASA

<sup>23</sup> NASA EVM Implementation Handbook. (2023), NASA EVM Implementation Handbooks | NASA



The **Empower** software developed, supported, and maintained by **Encore Analytics** (see the **Empower Appendix**). “Empower is the first and only browser-based analytical tool that integrates



earned value, schedule, work authorization, and other key performance data to enable proactive management of complex projects.”<sup>24</sup> The manual and automated DECM analysis culminates in a series of validity checks and balances of the monthly cost and schedule performance as well as measuring the integrity of the data. When the analytic results reveal significant deficiencies in data integrity the projects are required to provide a root cause, implement a corrective action, and provide a preventative measure to prevent reoccurrence.

## **EVM – NASA, DoD, Military Services – DCMA Delegation**

An example of Defense Contract Management Agency (DCMA) EVM delegation responsibilities is the NASA/DCMA Memorandum of Understanding (MOU) for Earned Value Management for EVM System Acceptance/EVM Project Surveillance. The MOU is used to delegate DCMA the responsibility “...for reviewing suppliers Earned Value Management (EVM) system plans and verifying initial compliance [and surveillance] with NASA and DOD Earned Value Management system criteria and conformity with ANSI Standard EIA-748, Industry Guidelines for Earned Value Management Systems.”<sup>25</sup>

This research is focused on the value of Earned Value Management: System review, program management tools, and internal controls used to ensure government compliance with the integration of the scope of work to cost, schedule, and performance. As mentioned above, the DCMA EVMS Compliance Metrics (DECM) are being utilized for manual and automated analysis by NASA, other government agencies, and industry contractors. Factors: Cost or incentive contracts subject to the EVM clause.<sup>26</sup>

There are two guides; one is for implementation, and the other is for interpretation, from the Office of “Acquisition Analytics and Policy (AAP)...is accountable for EVM Policy, oversight, and governance across the DOD” (AAP, 2019a, p. 1). The AAP Earned Value Management Implementation Guide (EVMIG) provides guidance for the EVM concepts, use, and application to contracts (AAP, 2019a, p. i). The AAP Earned Value Management System Interpretation Guide (EVMSIG) provides the “DOD interpretation of the 32 guidelines” (AAP, 2019b, p. 4). The guidance for interpretation of EVM policy pertains to several internal and regulatory requirements, for example, 32 guidelines covered in the EIA-748-C Standard, “Earned Value Management Systems,” and DFARS 252.234-7002, as well as DFARS Subpart 242.302, “Contract Administrations Functions, etc.” (AAP, 2019b, p. 90). Additionally, the DCMA-MAN 2301–01, Section 5: Earned Value Management System, provides guidance related to several internal and regulatory requirements (2019c).

The contractor must ensure they have the proper internal control tools and a formal documented process that includes standard business management practices. To this end, the contractor maintains internal controls documented in their EVM System Description (SD). As mentioned previously, the EVMS functional specialist must review and ensure that the contractor’s internal controls and business management practices comply with the 32 guidelines.

<sup>24</sup> Empower. (2023, April 5). Encore Analytics, Encore Analytics – Actionable insight for complex projects. (encore-analytics.com)

<sup>25</sup> NASA/DCMA Memorandum of Understanding for Earned Value Management (n.d.), NASA/DCMA Memorandum of Understanding for Earned Value Management | NASA

<sup>26</sup> Government Accountability Office. (2019). Contractor business systems: DOD needs better information to monitor and assess review process (GAO-19-212). <https://www.gao.gov/assets/700/696801.pdf>



## DCMA – Contract Management

In fiscal year 2021, the Defense Contract Management Agency (DCMA) celebrated more than two decades of delivering contract management support to the nation's warfighters. DCMA reports to the under secretary of defense for acquisition and sustainment and is a valued contributor to the greater national defense team, ensuring readiness and delivering 409.2 million items to the warfighter annually. According to the DCMA Insight Magazine, a yearly publication where the agency provides data related to its contract management role. DCMA's role is primarily funded by the military or federal government (e.g.,



**Department of the Army, Navy, Marines, Air Force, or NASA**). The agency's oversight role, as the first line of defense against fraud, waste, and abuse of taxpayer dollars, ensures that a contractor business system (CBS) is compliant. Since it has managed more than 232,166 contracts at 13,335 contractor facilities valued at \$4.55 trillion, there is a return on investment of 1.44 to 1 on every dollar invested.<sup>27</sup> This measure is critical in identifying fraud, waste, and abuse of taxpayer dollars.

## Role of the DCMA and DCAA

The DCMA is the cognizant federal agency responsible for the management of contracts. Additionally, these laws and timelines directly affect the DCAA and DCMA as they are the government audit agencies responsible for surveillance and auditing of DoD contractor accounting systems. A DCMA or DCAA functional specialist issues a business system report, and they make compliance recommendations to the DCMA through an audit report (FAR 42.101, 2020).

DCMA issues a business system analysis summary (BSAS) for

- Earned value management systems (EVMS; DFARS 252.234-7002)
- Contractor property management systems (DFARS 252.245-7003, 2012)
- Contractor purchasing systems (CPSR; DFARS 252.244-7001, 2014)

DCAA issues a business system report (audit report) for

- Accounting systems (DFARS 252.242-7006, 2012)
- Cost estimating systems (DFARS 252.215-7002, 2012)
- Material management and accounting systems (MMAS; DFARS 252.242-7004, 2011)

When the military or federal government (e.g., Department of the Army, Navy, Marines, Air Force, or NASA) awards a contract, if the program manager (PM) uses the DCMA to manage the contract, the procurement contracting officer (PCO) will send a delegation authority to the DCMA. The PCO can withhold all but FAR 42.302 (2020), Contract Administration Functions, to the DCMA:

The contracting officer normally delegates the following contract administration functions to a contract administration office [CAO]. The contracting officer may retain any of these functions, except those in paragraphs (a)(5), (a)(9), (a)(11) and (a)(12) of this section, unless the cognizant Federal agency (see 2.101) has designated the contracting officer to perform these functions. (FAR 42.302, 2020)

---

<sup>27</sup> Tremblay, P. (2021). Agency awarded expanded mission. *Defense Contract Management Agency Insight*, DCMA\_Insight\_2022.pdf



In addition to the clauses, DCMA directives, policies, manuals, and instructions are used to determine and mitigate risk with the contractor business systems. The DCMA provides contract oversight, surveillance, and compliance processes when performing contractor business system compliance reviews and evaluating data integrity.

It is DCMA policy to

- Ensure contractors maintain effective business systems, processes, and procedures
- Perform contractor business system reviews and determinations in a multifunctional, integrated, synchronized, and coordinated manner
- Execute this [DCMA-MAN 2301-01] manual in a safe, efficient, effective, and ethical manner<sup>28</sup>

### GSA Provides Guidance for Buying Agencies - Small Business

The focus for this research is the support of GSA to the Office of Small Business, which provides nationwide coverage for national security and building the industrial base. GSA provides workplaces, acquisition solutions, promotes management best practices for efficient government operations through the development of governmentwide policies.<sup>29</sup> Their goal is to meet and exceed statutory prime and subcontracting small business and socio-economic small business goals. GSA assists small businesses in finding federal contract opportunities. In addition, GSA's assistance with nationwide procurement opportunities, ensures small business participation, and training.<sup>30</sup>



### Small Business Improvement Acts

On February 3, 2022, Small Business Committee Passes and Recommends Five Bills to the House of Representatives that will help American small business entrepreneurs succeed, which are shown below.<sup>31</sup>



H.R. 6445, Small Business Development Centers Improvement Act of 2022 - to amend the Small Business Act to require an annual report on entrepreneurial development programs, and for other purposes. On April 26, 2022, passed House.

H.R. 6441, Women's Business Centers Improvement Act of 2022 - to amend the Small Business Act to improve the women's business center program, and for other purposes.

<sup>28</sup> Defense Contract Management Agency. (2019c, April 28). Contractor business system (DCMA-MAN 2301-01). Department of Defense. <https://www.dcms.mil/Portals/31/Documents/Policy/DCMA-MAN-2301-01.pdf?ver=2019-05-03-123122-347>

<sup>29</sup> U.S. General Services Administration. (2022). About us. About Us | GSA

<sup>30</sup> U.S. General Services Administration. (2022). FY 2024 Annual Performance Plan and FY 2022 Annual Performance Report. Office of Small and Disadvantaged Business Utilization | GSA

<sup>31</sup> House Small Business Committee Republicans (2022, February 3). Small Business Committee Passes and Recommends Five Bills to the House of Representatives.



On April 27, 2022, received; read twice and referred to the Committee on Small Business and Entrepreneurship.

H.R. 6450, SCORE for Small Business Act of 2022 - to amend the Small Business Act to reauthorize the SCORE program, and for other purposes. On April 26, 2022, passed House.

H.R. 4877, "ONE STOP SHOP FOR SMALL BUSINESS COMPLIANCE ACT OF 2021" - To amend the Small Business Act to require the Small Business and Agriculture Regulatory Enforcement Ombudsman to create a centralized website for compliance guides, and for other purposes. **On October 10, 2022, Public Law No: 117-188<sup>32</sup>**

H.R. 6454, Small Business Advocacy Improvements Act of 2022 - to clarify the primary functions and duties of the Office of Advocacy of the Small Business Administration, and for other purposes. On April 26, 2022, passed House.

### Office of Small Business Programs DoD Mentor Protégé Program

The OSBP is under the Small Business Act that established mandatory small business contracting goals and programs that apply to **DOD and all Federal agencies** (Office of Small Business Programs, 2022, April 5).<sup>33</sup> Its mission is to contribute to national security by maximizing opportunities for small businesses that provide combat supplies for our troops and economic sustainment for our nation. One of the highest responsibilities is the management of the DOD Mentor Protégé Program (MPP). The program is critical to developing high priority sectors of the **DOD Industrial Base** (OSBP, 2022). A representation of some of the DOD MPP Project Spectrum Program Partnerships is shown in Figure 5 (Diaz, 2021, pg. 6).

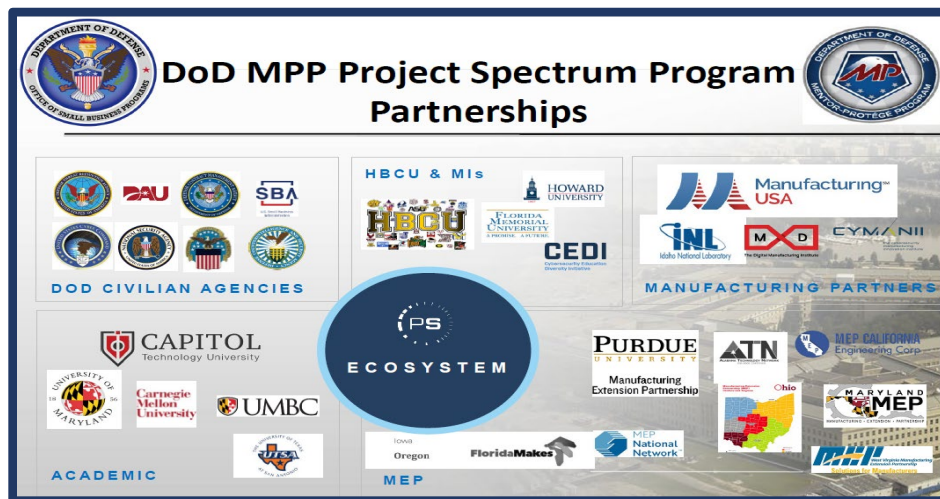


Figure 5. Extracted From DoD MPP Project Spectrum Program Partnerships (Diaz, 2021)

On November 5, 1990, H.R. 4739 – National Defense Authorization Act (NDAA) for Fiscal Year (FY) 1991, directs the Secretary of Defense to establish a Mentor-Protégé Program [MPP] to provide incentives to major DOD contractors (mentors) to help disadvantaged small

<sup>32</sup> H.R.4877 - One Stop Shop for Small Business Compliance Act of 2021, Public Law 117-188, (2022, October 10). H.R.4877 - 117th Congress (2021-2022): One Stop Shop for Small Business Compliance Act of 2021 | Congress.gov | Library of Congress

<sup>33</sup> Office of Small Business Programs (2022, April 5). Mentor-Protégé Program (MPP).





businesses (protégés) perform as subcontractors and suppliers under DOD and other government contracts.<sup>34</sup>

On October 1, 1991, the DOD MPP was the first operative federal mentor-protégé program that since its inception as a pilot program. It has received continuous funding extensions as a pilot despite the 1994-scheduled expiration. Currently, it is funded through FY2026 for reimbursement of cost incurred under existing agreements and FY2024 for the formation of new agreements. DOD's MPP is the only federal pilot program that is mandated by law and receives authorized and appropriated funds (Mentor Protégé Pilot Program, 1990).<sup>35</sup>

Historically, the DoD's Mentor-Protégé Program is a front-runner with mentors' commitment to leveraging small business protégés in successfully growing the DIB, but the Office of the Secretary of Defense (OSD) must champion support and funding for the MPP. In FY20, the MPP experienced a zeroed-out funding from the DoD in the FY2020 Defense Wide Review (DWR). The President's Budget Request (PBR) rescued funding for the MPP by adding it back in for FY2021 (Defense Business Board [DBB], 2022, p. 33).

### **Sec. 856. Codification of the Department of Defense Mentor–Protégé Program**

December 23, 2022, President Joe Biden signed the H.R.7776 - James M. Inhofe National Defense Authorization Act for Fiscal Year 2023 National Defense Authorization Act<sup>36</sup>, funding \$858 Billion for **DOD and national security programs** under the Department of Energy. Notably, since 1991 the Mentor- Protégé Program Pilot withstood the test of time for 32 years until the 856. Codification, which is no longer a pilot program. The eligibility threshold for mentors dropped \$25M in the total of defense contracts. In addition, the MPP participation duration extends to three years. To strengthen the software development DIB, Sec. 856 establishes a five-year Protégé Technical Reimbursement Program with an incentive to the protégé and mentor. The protégé firm may receive up to 25 percent of the reimbursement that is also offered to the mentor in the agreement for the engineering and software development that will be integrated with a DOD program or system. The **DCMA and NASA Mentor-Protégé point of contacts** are as follows:

---

<sup>34</sup> National Defense Authorization Act (1990, November 5). Pub. L. No. 101-510, 104 STAT. 1490, Title VIII: Acquisition Policy, Acquisition Management, and Related Matters - Part D: Miscellaneous, Sec. 831. Mentor-protégé pilot program (1990).


<sup>35</sup> Mentor-Protégé Pilot Program, Section 807 (a) of Pub. L. 102-484 (1991).

<sup>36</sup> National Defense Authorization Act (2022, December 23). Pub. L. No. 117-263, Subtitle E – Industrial Base Matters. Codification of the Department of Defense Mentor-Protégé Program. Text - H.R.7776 - 117th Congress (2021-2022): James M. Inhofe National Defense Authorization Act for Fiscal Year 2023 | Congress.gov | Library of Congress



## The DCMA and NASA Mentor Protégé Program Points of Contacts

### DCMA Mentor Protégé – Angela Dokes<sup>37</sup>

 <p><b>Angela Dokes</b> Mentor-Protégé Program Team Lead, Defense Contract Management Agency (DCMA)</p>	<p><b>Angela Dokes</b> Mentor-Protégé Program Team Lead, Defense Contract Management Agency (DCMA)</p> <p>Ms. Dokes serves as the Team Lead for the Defense Contract Management Agency (DCMA) Mentor-Protégé Program (MPP) team. She has over 14 years of experience working with DCMA, including five years in contracts. In her role as the Team Lead, Mrs. Dokes is responsible for the administration and reporting of the Department of Defense (DoD) Mentor Protégé Program (MPP) to include providing data for congressional inquiries. Ms. Dokes is Lean Six Sigma, Yellow Belt Certified, Small Business Professional Certified, DAWIA Level III Certified, and is a member of the Defense Acquisition Corps.</p> <p>Prior to her government career, Ms. Dokes worked as a teacher/ administrator in Public Education for 15 years. Ms. Dokes holds a Bachelor of Science in Math Education from the University of Arkansas at Pine Bluff and a Masters in Educational Administration from Lindenwood University.</p>
--	--



### NASA Mentor Protégé – Glen A. Delgado<sup>38</sup>

 <p><b>Glen A. Delgado</b> Associate Administrator for Small Business Programs</p>	 <p>Delgado provides executive leadership and policy direction for developing and implementing initiatives, that ensure all categories of small businesses are afforded opportunities to compete for agency contracts.</p> <p>With more than 40 years of acquisition experience, Delgado has received several awards and medals, including the Presidential Rank Award (Meritorious Service). The Congressional Black Caucus honored Delgado as a Small Business Champion Living Legends Award. He also received two NASA Outstanding Leadership Medals and the NASA Exceptional Service Medal.</p> <p>Delgado earned a Bachelor of Science degree from the University of New Hampshire and a Master of Business Administration degree from Marymount University. He is Level III-certified in the acquisition professional field of contracting and a member of the acquisition professional community.</p>
--	---



<sup>37</sup> Dokes, Angela. (2023, April 5). DCMA. Mentor Protégé Program. Small Business (dcma.mil)

<sup>38</sup> Delgado, Glenn A. (2023, April 5). NASA. Small Business Programs. Glenn A. Delgado | NASA



## Appendix. Empower, Encore Analytics



Encore Analytics LLC is a verified Veteran-Owned Small Business (VOSB) by the U.S. Department of Veterans Affairs. The customer base of more than 49 is comprised of government, federal agencies, military services, academia, and corporate industry leaders. The most notable customers include NASA, the Department of Energy (DOE), the National Reconnaissance Office, the National Security Agency, the Missile Defense Agency, NAVAIR, the U.S. Army, the U.S. Marine Corps, Lockheed Martin, Raytheon, Northrop Grumman, other federal agencies, Defense Acquisition University (DAU), and industry (customer base listed below).

The customer base uses Empower, an enterprise level web-based analytical tool to collect earned value and schedule performance data from suppliers in standardized electronic formats. The system contracts are loaded in an enterprise database, then populated and managed by a central staff. The analytical tool allows for timely dissemination of integrated cost/schedule performance data to all program stakeholders via a single web URL that contains tailored dashboards for various roles including program managers, government technical managers, schedule analysts, and cost analysts.

Empower is also capable of running standardized data quality checks in accordance with the DCMA DECM tests for compliance professionals to evaluate supplier data quality. This ability of the customer to test data quality has significantly improved with the new Integrated Program Management Data and Analysis Report (IPMDAR) delivery specifications for cost (earned value) and the integrated master schedule (IMS) data on new contracts. The IPMDAR formats are currently in use by most of Encore Analytics' customers. The formats allow the customer to receive better integrated cost and schedule data, as well as data at a more detailed level (control account or work package) with element of cost delineation. This allows the customer to quickly identify the root cause of cost issues and schedule delays for corrective action initiatives and/or estimate at completion updates.

The use of Empower, especially as a corporate enterprise program performance analytical tool, eliminates software costs, software maintenance costs, IT support costs, and training costs for all customer programs since it is centrally funded. Also, since Empower is a widely used tool, it is relatively easy to find and hire experienced resources familiar with the tool.

Empower allows for integration with external data sources as well as other web-based tools. Encore Analytics has provided integrated Empower capabilities within web-based systems such as the Program Analysis and Reporting System (PARS) for analytics. PARS includes all programmatic data, including scope of work, budget, budget execution, risk management, key program milestones, etc.

To accelerate and increase the software's adoption, dashboards can be deployed at an enterprise level and customized for each user. A series of notional dashboards are included in this paper to illustrate dashboards by role (PM, technical manager, cost/schedule analyst, compliance manager, etc.).



## Notional Program Manager (PM) Dashboards

The dashboards shown in this section include multiple projects/contracts that a PM might be responsible for. Projects can be grouped into programs or portfolios and have interactive drill-down to locate significant cost and/or schedule drivers for management-by-exception. The tool allows configuration by role to reduce menu items and options to make the tool more intuitive for senior managers who only occasionally review contract performance data. Notional dashboards are provided.

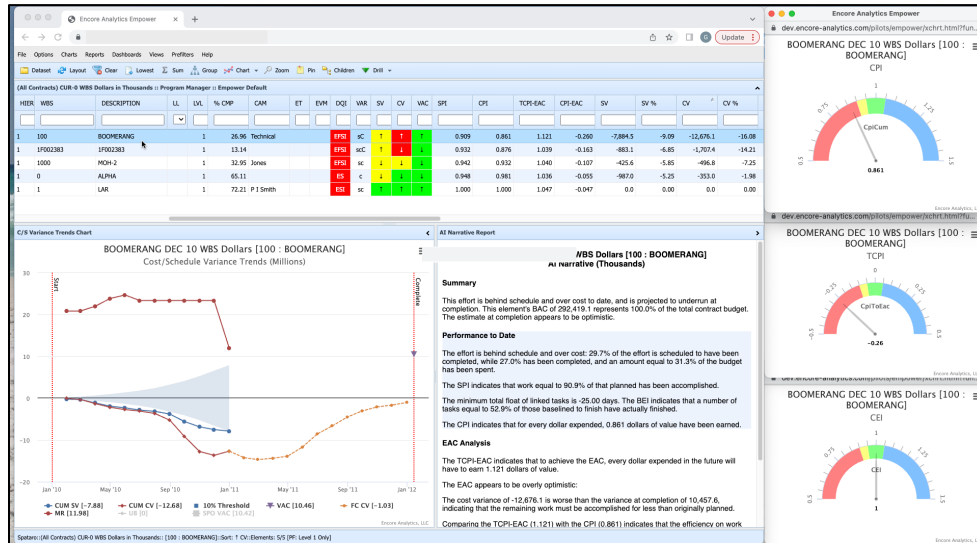


Figure 8. PM Contract Performance

The PM Contract Performance dashboard shows in the grid all contracts the PM has responsibility for with color coded performance indicators, trend arrows and key performance metrics for each contract. The PM can sort or filter on any field and/or drill down to the root cause of the variance. The charts, reports and gauges automatically update as the PM selects new projects or elements within a project. This dashboard displays a trend chart of cost/schedule variances with a light blue shaded tolerance threshold, a rule-based assisted intelligence (AI) Narrative Report that transforms the performance data to explanatory text regarding the performance, and three gauges showing key performance metrics.

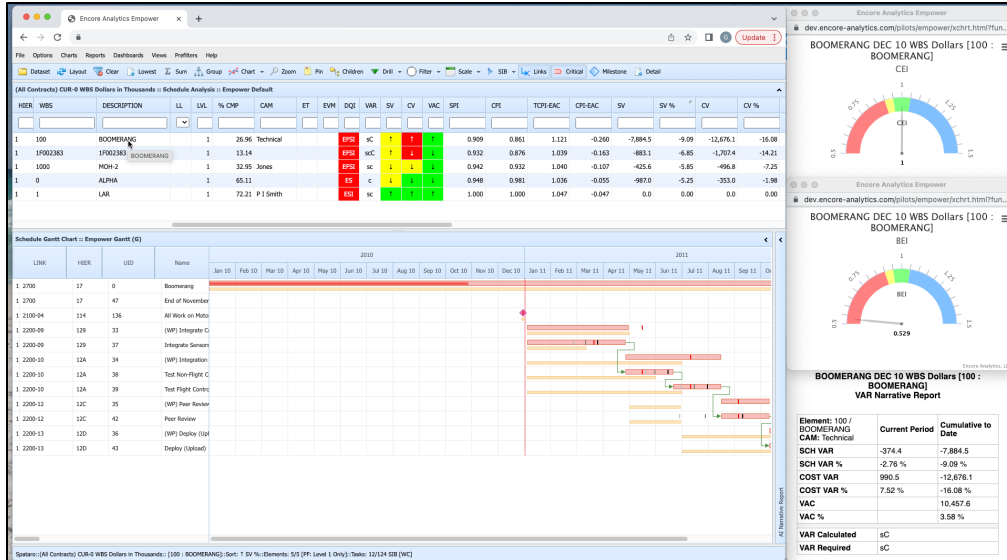


Figure 9. PM IMS Dashboard

This dashboard shows the critical path for each selected contract along with the IMS related Critical Execution Index (CEI) and Baseline Execution Index (BEI) gauges and the contractor's narrative submission regarding the performance on the contract. Note in the schedule Gantt display, it shows finish dates reported in prior schedules with the "I" bars in black and shades of gray and (I) as the drop-dead date for negative float. This allows the program manager to see slips on the critical path and if negative float exists in the schedule which is a key indicator of risk.

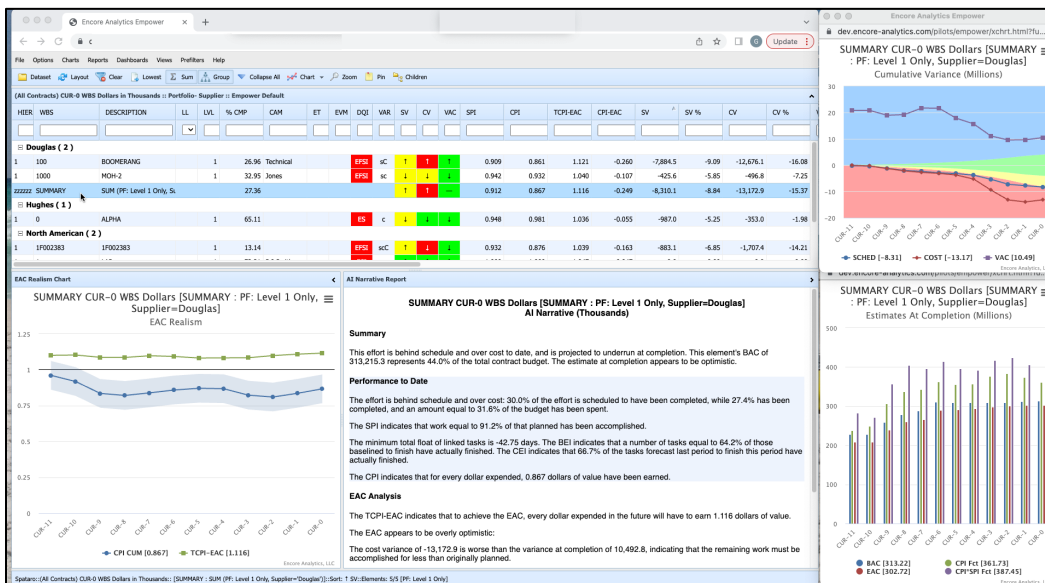


Figure 10. PM Portfolio Dashboard

This dashboard groups projects by supplier and provides for a summary line where the aggregate performance for that supplier can be analyzed. This type of view can be used to determine if performance is an issue on a single contract or might be a broader problem with a certain supplier.



# Technical Manager/Analyst Dashboards

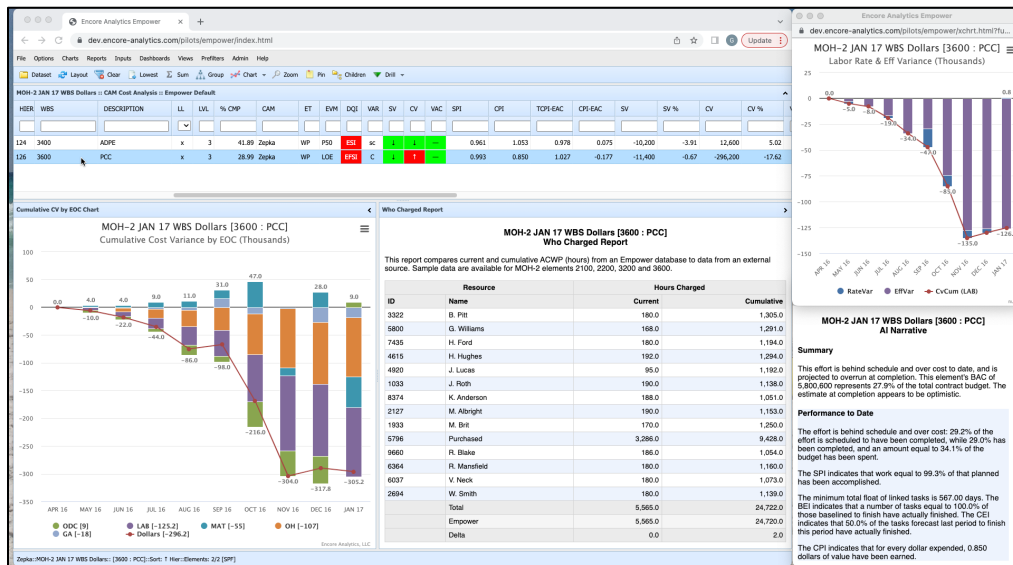


Figure 11. Detailed Cost Dashboard

This dashboard displays the two accounts assigned to government focal point “Zepka” along with an Element of Cost (EOC) chart (lower left) that shows the cost variance (CV) broken out by EOC (Labor, Material, Other Direct Charges [ODC], Overhead [OH] and General & Administrative [G&A]). Note that labor is a large contributor to the negative CV, so two more widgets are displayed as additional information regarding labor performance for this account. The first is the Who Charged Report that shows individuals or subcontractors who charged labor to this account. The second is the Labor Rate and Efficiency Variance Chart in the upper right corner. This chart clearly shows the labor issue is an efficiency problem where more hours are being expended to complete work than planned.

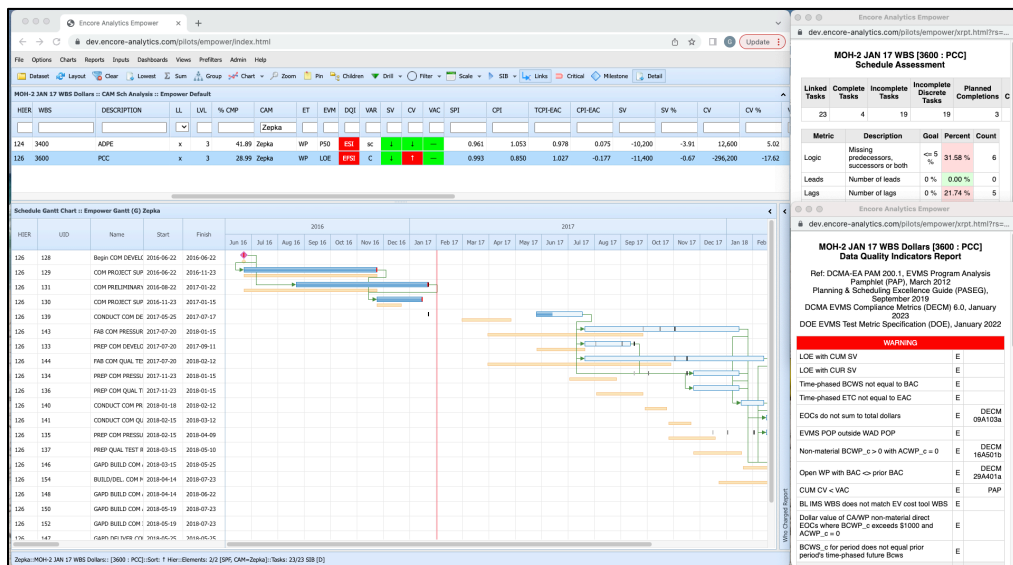


Figure 12. Detailed Schedule Dashboard

This dashboard shows schedule activities for the selected account with data quality indicator reports to the right.



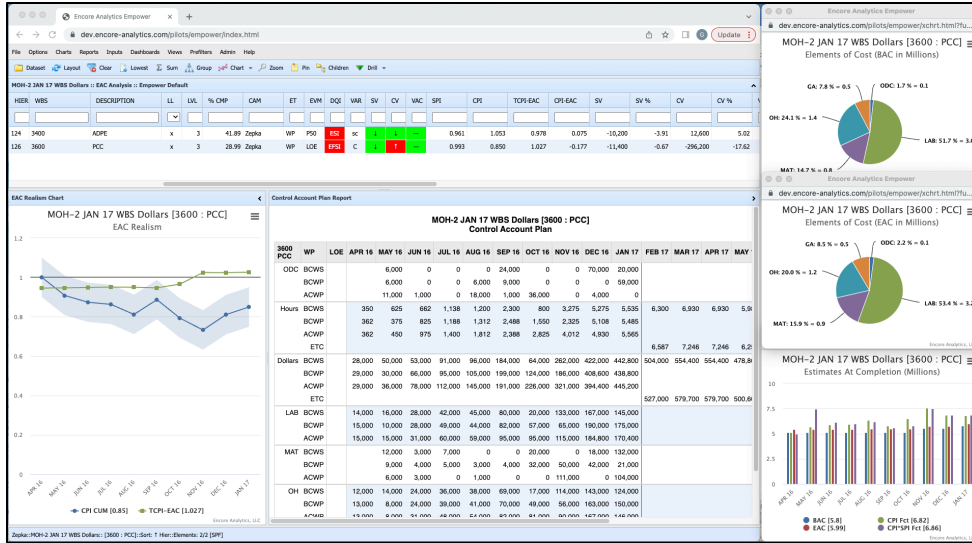


Figure 13. Estimate at Completion (EAC) Dashboard

This dashboard shows a report like a Control Account Plan (CAP) and charts related to EAC Realism, mathematically calculated EACs, and pie charts that break out EOCs by the budget and the estimate at complete.

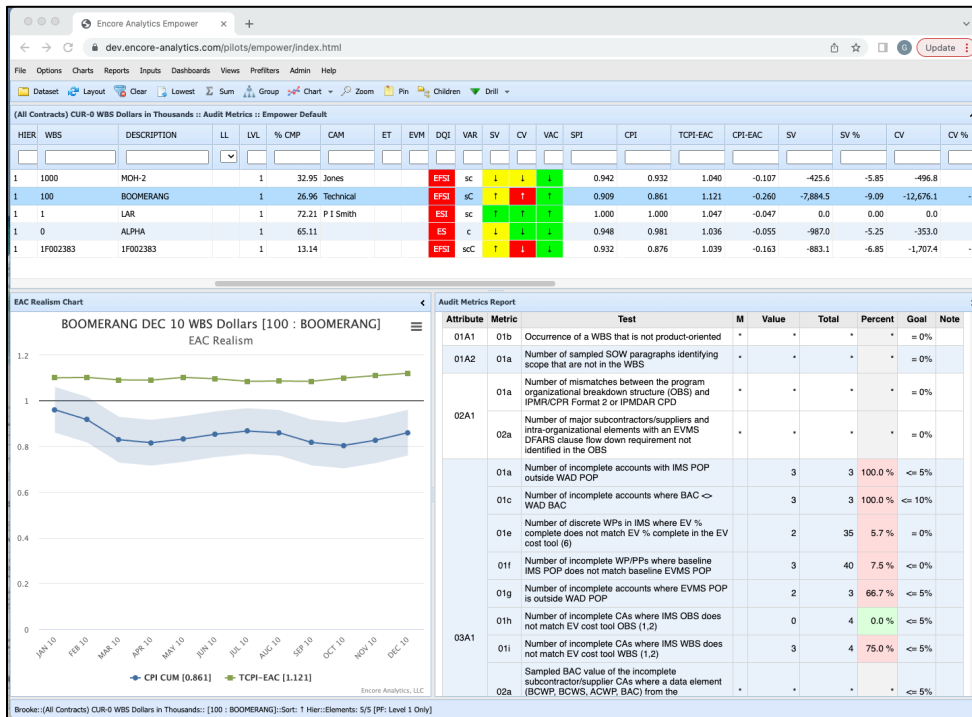



Figure 14. Compliance Manager Dashboard

The compliance manager dashboard shows the Audit Metrics Report with the numerator, denominator, and goal for each test. Tests with a “\*” in the “M” column are manual tests per DCMA instructions, and the remaining 73 tests can be automatically calculated by the Empower software from the performance data.



 	
No	Customer Base*
1	Austal
2	Betchel
3	Blue Halo
4	Blue Origin
5	Brookhaven National Laboratory
6	Cobham
7	Consolidated Nuclear Security/Y12 National Security Complex
8	Defense Acquisition University
9	Department of Energy
10	Eastern Shipbuilding
11	Fincantieri Marinette Marine
12	Four Rivers Nuclear Partnership
13	General Atomics Aeronautical
14	Honeybee Robotics
15	Honeywell FM&T KCNSC
16	Idaho National Laboratory
17	Jacobs
18	John Hopkins Applied Physics Laboratory
19	L3Harris
20	Lawrence Livermore National Laboratory
21	Lockheed Martin Aeronautics
22	Lockheed Martin Space
23	Los Alamos National Laboratory
24	Mid-America Conversion Services
25	Missile Defense Agency
26	Mission Support and Test Services, LLC
28	NASA
29	NASA Jet Propulsion Laboratory
30	National Nuclear Security Administration
31	National Reconnaissance Office
32	National Security Agency
33	NAVAIR
34	Northrop Grumman
35	Orbital/ATK
36	ORCC Oak Ridge
37	Pratt & Whitney
38	Progeny Systems
39	Raytheon
40	Rolls Royce
41	Sandia National Laboratories
42	Savannah River Remediation
43	Sierra Nevada Corporation
44	Sierra Space
45	Southwest Research Institute
46	US Army
47	US Marine Corps
48	VT Halter Marine
49	Washington River Protection Solutions

\*List is as of September 2022

## List of References

Abba, W. F. (2017). The evolution of earned value management. *Defense AT&L Magazine*. dau.edu  
 Bembers, I., Know, Jones, & Traczyk. (2017, January-February). EVMS system's high cost-Fact or fiction? *Defense AT&L Magazine*. dau.edu





*Biden-Harris administration's national security strategy.* (2022, October 12). whitehouse.gov

Defense Business Board. (2022, March 8). *FY2022 assessment of the Department of Defense mentor-protégé program.*  
[https://dbb.defense.gov/Portals/35/Documents/Meetings/2022/Feb%202022%202022/DBB%20FY22-01%20MPP%20Assessment%20Final%20Report%20w\\_Cover%20-%20Cleared.pdf?ver=s1EXfZ7pqYtqa5FAoDMwwg%3d%3d](https://dbb.defense.gov/Portals/35/Documents/Meetings/2022/Feb%202022%202022/DBB%20FY22-01%20MPP%20Assessment%20Final%20Report%20w_Cover%20-%20Cleared.pdf?ver=s1EXfZ7pqYtqa5FAoDMwwg%3d%3d)

Defense Contract Management Agency. (2019, April 28). *Contractor business system* (DCMA-MAN 2301-01). DoD. <https://www.dcma.mil/Portals/31/Documents/Policy/DCMA-MAN-2301-01.pdf?ver=2019-05-03-123122-347>

Delgado, G. A. (2023, April 5). *Small business programs.* NASA.

DFARS 252.215-7002, Cost Estimating System Requirements (2012, December).  
[https://farclause.com/FARregulation/Clause/DFARS252.215-7002\\_Basic-costestimating-system-requirements#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.215-7002_Basic-costestimating-system-requirements#gsc.tab=0)

DFARS 252.242-7002, Earned Value Management System (2011, May).  
[https://farclause.com/FARregulation/Clause/DFARS252.234-7002\\_Basic-earned-value-management-system#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.234-7002_Basic-earned-value-management-system#gsc.tab=0)

DFARS 252.242-7004, Material Management and Accounting System (2011, May).  
<https://www.acquisition.gov/dfars/part-252-%E2%80%93-clauses?&searchTerms=252-242.7004%28d%29%285%29#DFARS-252.242-7004>

DFARS 252.242-7006, Accounting System Administration (2012, February).  
[https://farclause.com/FARregulation/Clause/DFARS252.242-7006\\_Basic-accounting-system-administration#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.242-7006_Basic-accounting-system-administration#gsc.tab=0)

DFARS 252.244-7001, Contractor Purchasing System Administration (2014, April).  
[https://farclause.com/FARregulation/Clause/DFARS252.244-7001\\_Alt\\_lcontractor-purchasing-system-administration#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.244-7001_Alt_lcontractor-purchasing-system-administration#gsc.tab=0)

DFARS 252.245-7003, Contractor Property Management System Administration (2012, April).  
[https://farclause.com/FARregulation/Clause/DFARS252.245-7003\\_Basic-contractor-property-management-system-admini#gsc.tab=0](https://farclause.com/FARregulation/Clause/DFARS252.245-7003_Basic-contractor-property-management-system-admini#gsc.tab=0)

Diaz, K. (2021, May). *DoD mentor protégé program.* DoD, Office of Small Business Programs.  
[https://www.mda.mil/global/documents/pdf/Diaz,%20K%20-%20DoD%20Mentor%20Protege%20Program\\_MDA%20SB%20CONF%20\(May\\_2021\)FINAL.pdf](https://www.mda.mil/global/documents/pdf/Diaz,%20K%20-%20DoD%20Mentor%20Protege%20Program_MDA%20SB%20CONF%20(May_2021)FINAL.pdf)

DoD, Office of Small Business Programs. (2022, April 5). *Mentor-protégé program (MPP).*  
<https://business.defense.gov/Programs/Mentor-Protege-Program/>

*DOD guide for achieving reliability, availability, and maintainability.* (2005, August 3).

Dokes, A. (2023, April 5). *DCMA. Mentor protégé program. Small business* (dcma.mil).

Empower. (2023, April 5). *Encore Analytics – Actionable insight for complex projects.* encore-analytics.com

FAR 42.101, Contract Audit Responsibilities (2020). <https://www.acquisition.gov/content/part-42-contract-administration-and-audit-services#id1617MD0K0QY>

FAR 42.302, Contract Administration Functions (2020). <https://www.acquisition.gov/content/42302-contract-administration-functions>

Federal Acquisition Regulation. (2006, July 5). *Earned value management system (EVMS)* (FAR Case 2004-019). Federal Register. <https://www.govinfo.gov/content/pkg/FR-2006-07-05/pdf/06-5966.pdf>

Federal Register. (1998). Vol. 63, No. 220 / Monday, November 16, 1998 / Proposed Rules 63655. 98-30554.pdf (govinfo.gov)

Federal Reserve History. (1933). Banking Act of 1933 (Glass-Steagall).

GAO. (2019). *Contractor business systems: DOD needs better information to monitor and assess review process* (GAO-19-212). <https://www.gao.gov/assets/700/696801.pdf>

GSA. (2022a). *About us.*

GSA. (2022b). *FY 2024 annual performance plan and FY 2022 annual performance report.* Office of Small and Disadvantaged Business Utilization.

Hite, R. C. (2010, August). *Organizational transformation: A framework for assessing and improving enterprise architecture management* (Ver. 2.0; GAO-10-846G). GAO.



House Small Business Committee Republicans. (2022, February 3). *Small business committee passes and recommends five bills to the House of Representatives*.

H.R. 4877, One stop shop for Small Business Compliance Act of 2021, Pub. L. No. 117-188. (2022, October 10). H.R. 4877 - 117th Congress (2021-2022): One Stop Shop for Small Business Compliance Act of 2021. Congress.gov

Mentor-Protégé Pilot Program, Section 807(a), Pub. L. No. 102-484 (1991).

NASA. (2019). *About NASA*.

NASA. (2020, September 25). *NASA report details how agency significantly benefits US economy* (Release 20-093).

NASA. (2021, May 12). *NASA has selected Manufacturing Technical Solutions Inc. of Huntsville, Alabama. NASA selects CPSS PP&C contractor* (Contract Release C21-011).

NASA. (2023, February 27). *FY 2023 summary, National Aeronautics and Space Administration (NASA) | Spending profile | Government spending open data | USA spending*.  
<https://www.usaspending.gov/>

NASA/DCMA memorandum of understanding for earned value management. (n.d.). NASA.

NASA EVM implementation handbook. (2023). NASA.

National Defense Authorization Act, Pub. L. No. 101–510, 104 Stat. 1490 (1990, November 5). *Title VIII: Acquisition policy, acquisition management, and related matters - Part D: Miscellaneous, Sec. 831. Mentor-protégé pilot program*.

National Defense Authorization Act, Pub. L. No. 117–263 (2022, December 23). *Subtitle E – Industrial base matters*. Codification of the Department of Defense Mentor-Protégé Program. Text - H.R.7776 - 117th Congress (2021-2022): James M. Inhofe National Defense Authorization Act for Fiscal Year 2023. Congress.gov

National Defense Industry Association. (2016, October). U.S. federal agency EVMS policy summary. [ndia.org](http://ndia.org)

Office of Acquisition Analytics and Policy. (2019a, January 18). *Earned value management system implementation guide (EVMIG)*. DoD. <https://www.acq.osd.mil/evm/assets/docs/DoD%20EVMIG-01-18-2019.pdf>

Office of Acquisition Analytics and Policy. (2019b, March 14). *Earned value management system interpretation guide (EVMSIG)*. DoD. [https://www.acq.osd.mil/evm/assets/docs/DoD\\_EVMSIG\\_14MAR2019.pdf](https://www.acq.osd.mil/evm/assets/docs/DoD_EVMSIG_14MAR2019.pdf)

Office of Small Business Programs. (2022, April 5). *Mentor-protégé program (MPP)*.

Pickar, C. (2020). *Learning from experience: Acquisition professional education for this century* (SYM-AM-20-070). Naval Postgraduate School, Acquisition Research Symposium. [nps.edu](http://nps.edu)

Tremblay, P. (2021). *Agency awarded expanded mission*. Defense Contract Management Agency Insight. [DCMA\\_Insight\\_2022.pdf](https://www.dcmadocs.com/DCMA_Insight_2022.pdf)

United Nations Office for Outer Space Affairs. (1967, October). *Treaty on principles governing the activities of states in the exploration and use of outer space, including the moon and other celestial bodies*. The Outer Space Treaty. [unoosa.org](http://unoosa.org)

USAspending. (2023, April 4). *Spending by geography*. <https://www.usaspending.gov/>

U.S. Chamber of Commerce. (2022, September 14). *How the U.S. is establishing national security in space*. [uschamber.com](http://uschamber.com)

White House. (2021, January 25). *Executive order 14005, section 4(a): Ensuring the future is made in all of America by all of America's workers*. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/25/executive-order-on-ensuring-the-future-is-made-in-all-of-america-by-all-of-americas-workers/>







ACQUISITION RESEARCH PROGRAM  
DEPARTMENT OF DEFENSE MANAGEMENT  
NAVAL POSTGRADUATE SCHOOL  
555 DYER ROAD, INGERSOLL HALL  
MONTEREY, CA 93943

[WWW.ACQUISITIONRESEARCH.NET](http://WWW.ACQUISITIONRESEARCH.NET)