ACQUISITION INNOVATION RESEARCH CENTER

Digital Data Management and Analytic Pilots

EXECUTIVE SUMMARY AND REPORT

REVISED: AUGUST 2024

PRINCIPAL INVESTIGATOR
Philip S. Antón, *Stevens Institute of Technology*

CO-PRINCIPAL INVESTIGATOR
Laura Freeman, *Virginia Tech NSI*

SPONSOR
Mr. Mark E. Krzysko, SES, Principal Deputy Director, Enterprise Information, Acquisition Data and Analytics, Office of the Under Secretary of Defense for Acquisition and Sustainment

DISCLAIMER

The Acquisition Innovation Research Center (AIRC) is a multi-university partnership led by the Stevens Institute of Technology and sponsored by the U.S. Department of Defense (DoD) through the Systems Engineering Research Center (SERC)—a DoD University-Affiliated Research Center (UARC).

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, Task Order #0480.

The views, findings, conclusions, and recommendations expressed in this material are solely those of the authors and do not necessarily reflect the views or positions of the United States Government (including the Department of Defense (DoD) and any government personnel), the Stevens Institute of Technology, Virginia Tech, or VT-ARC.

No Warranty.

This Material is furnished on an "as-is" basis. The Stevens Institute of Technology, Virginia Tech, and VT-ARC make 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.

The Stevens Institute of Technology, Virginia Tech, and VT-ARC do not make any warranty of any kind with respect to freedom from patent, trademark, or copyright infringement.

SYSTEMS ENGINEERING RESEARCH CENTER

AIRC ACQUISITION INNOVATION RESEARCH CENTER

STEVENS INSTITUTE OF TECHNOLOGY 1870

TABLE OF CONTENTS

LIST OF FIGURES

RESEARCH TEAM

Name	Organization	Labor Category
Philip S. Antón	Stevens Institute of Technology	Principal Investigator
Laura Freeman	Virginia Tech	Co-Principal Investigator
Maegen Nix	VT-ARC	Director ISAD
Christina Houfek	VT-ARC	Lead Project Manager
Orlando Florez	Virginia Tech	Assistant Director for Program Management
Grant Beanblossom	VT-ARC	Technical Contributor
Kobie Marsh	VT-ARC	Technical Contributor
James Brown	VT-ARC	Technical Contributor
Stoney Trent	Virginia Tech	Principal Advisor, R&I
Daniel Wolodkin	VT-ARC	Technical Contributor

ACRONYMS AND ABBREVIATIONS

A&S	Acquisitions and Sustainment
ACAT	Acquisition Category
AI	Artificial Intelligence
AIRC	Acquisition Innovation Research Center
API	Applications Programming Interface
BLI	Budget Line Item
BLIN	Budget Line Item Number
CAC	Common Access Card
CDAO	Chief Digital and Artificial Intelligence Office
CIO	Chief Information Officer
CISA	Cybersecurity and Infrastructure Security Agency
CJCS	Chairman of the Joint Chiefs of Staff
COCOM	Combatant Commander
CONUS	Continental United States
CUI	Controlled Unclassified Information
DARCIE	Defense Acquisition Research, Collaboration and Innovation Environment
DASD	Deputy Assistant Secretary of Defense
DCA	Defensive Counterair
DEVOPS	Development Operations
DFAS	Defense Finance and Accounting Service
DMAG	Deputy's Management Action Group
DoD	Department of Defense
DOT&E	Director of Operational Test and Evaluation
DOE	Department of Energy
EIE	Energy, Installations and Environment
EMD	Engineering and Manufacturing Development
FFRDC	Federally Funded Research and Development Center
FY	Fiscal Year
GFI	Government Furnished Information
GHG	Greenhouse Gas

HR	Human Resources
IAPR	Integrated Acquisition Portfolio Review
IDEAS	Innovative Data-Enabled Acquisition Strategy
IT	Information Technology
JCIDS	Joint Capabilities Integration and Development System
JCTD	Joint Capability Technology Demonstration
JOA	Joint Operational Area
JSA	Joint Security Area
JWICS	Joint Worldwide Intelligence Communication System
MDA	Milestone Decision Authority
MDAP	Major Defense Acquisition Program
MES	Manufacturing Execution System
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MoDSAS	Model and Data Storage and Access System
MS	Milestone
MSCO	Modeling and Simulation Coordination Office
MTA	Middle Tier of Acquisition
NAVAIR	Naval Air Systems Command
NDS	National Defense Strategy
NIIN	National Item Identification Number
NIPRnet	Non-classified Internet Protocol Router Network
NSI	National Security Institute
OESRC	Office of Export and Secure Research Compliance
OUSD(A&S)	Office of the Under Secretary of Defense for Acquisitions and Sustainment
OUSD(R&E)	Office of the Under Secretary of Defense for Research and Engineering
P_{Detect}	Probability of Detection
PDF	Portable Document Format
PE	Program Element
PEO	Program Executive Office
PM	Program Manager
POM	Program Objective Memorandum

PPBE	Programming, Planning, Budgeting, and Execution
R&D	Research and Development
R&E	Research and Engineering
REPI	Readiness and Environmental Protection Integration
RFI	Request for Information
SBOM	Software Bill of Materials
SCM	Supply Chain Management
SERC	Systems Engineering Research Center
SES	Senior Executive Service
SIPRnet	Secret Internet Protocol Router Network
SQL	Structured Query Language
UI/UX	User Interface / User Experience
UJTL	Universal Joint Task List
US	United States
VT	Virginia Tech
VT-ARC	Virginia Tech – Applied Research Corporation

EXECUTIVE SUMMARY

Background. This technical report for WRT-1057.6 describes the follow-on efforts for WRT-1047p, Digital Data Management and Analytic Pilots, which focused on establishing use cases for innovative analyses of acquisition data, and establishing a data science environment for analysis of sensitive acquisition data known as DARCIE: Defense Acquisition Research, Collaboration and Innovation Environment. These efforts are consistent with the initiatives outlined in WRT-1047s IDEAS: Innovative, Data-Enabled Acquisition Strategy.

Use Cases. The previous research task had several recommendations focused on developing high-value use cases for analysis of acquisition data to drive innovation and improve decision-making. Working with the sponsor, the research team developed materials for eight potential use cases, which were used to solicit support from potential sponsors of acquisition data analysis. In order to ensure a commitment toward using the results, AIRC recommends that potential sponsors for the use cases provide some form of commitment to either support the research or adopt solutions. Three of the use cases represent opportunities to improve management of a portfolio of acquisitions. The eight use cases were:

- Portfolio Management
  - » Tradespace Analysis in Integrated Acquisition Portfolio Reviews (IAPRs)
  - » Total Lifecycle Portfolio Cost
  - » Alignment of Major Defense Acquisition Programs (MDAPs) and Universal Joint Task Lists (UJTLs)
- Digital Data Storage Management
- Rapid, Serial Incrementalism
- Securing the Software Supply Chain
- Quantitative Risk Assessment
- Climate Crisis

Of the eight, the two use cases that attracted the most attention were data analyses of software bills of materials to strengthen the software supply chain, and methods for quantitative risk management.

DARCIE. Building on the successful establishment of the pilot DARCIE 1.0 environment in the previous research task, this research task identified the requirements to develop the next phase of DARCIE. Two primary goals were to expand access and improve the user experience. The research team held conversations with three Federally Funded Research and Development Centers (FFRDCs) to understand the lessons that they learned in establishing their enclaves for processing controlled unclassified information (CUI). The researchers also met with Snowflake, Palantir and Scale AI to discuss their capabilities and ways they could potentially contribute to the DARCIE environment. This information served as inputs to the research team, which then developed a baseline set of requirements for further development of DARCIE (see Appendix A).

The research team found that the need for DARCIE continues to be strong as the pilot DARCIE environment was used to support analysis of Department of Defense (DoD) obligation data by the Planning, Programming, Budgeting, and Execution (PPBE) Reform Commission. The U.S. Marine Corps also piloted the use of DARCIE for analysis of CUI for the Digital Data Grand Prix, providing a model for how CUI data competitions could take place within the AIRC. Additionally, the research team met several times with representatives from the U.S. Air Force Research Institute for Tactical Autonomy to figure out the best methods to support their needs.

Recommendations.

Use Cases. The previous WRT-1047p research task recommended conducting a workshop to build a consensus around a set of high-value use cases for innovative analyses of acquisition data. Overall, the research team recommended following the precepts of Kotter's model of organizational change[1] to break down the institutional and cultural barriers surrounding innovative sharing and use of acquisition data:

- Establish a sense of urgency
- Create a guiding coalition
- Develop a vision and strategy
- Communicate the change vision
- Empower employees for broad-based action
- Generate short-term wins
- Consolidate gains & produce more change
- Anchor new approaches in the culture

In the future such a workshop could provide the impetus needed to obtain the leadership support for the sharing and analysis of acquisition data, help to create a guiding coalition, and successfully executing the analyses.

In order to start making the changes needed for innovative use of acquisition data, AIRC recommended that A&S seek out like-minded partners to identify high-impact opportunities. These partners should be empowered to represent their organizations and make the commitments necessary to share the data and conduct the necessary analyses. The focus should be on generating short-term wins on issues of importance to the Department.

DARCIE. The research team recommends continued DARCIE development as the need remains strong. The resulting infrastructure will help achieve the intent of the statute that establish AIRC[2], which included provisions to establish data repositories and develop analytical capabilities to enable researchers and acquisition professionals to access and analyze historical data sets to support research and develop new policies and practices.

[1] Kotter, John P. *Leading Change*. Harvard Business Review Press, 2012.

[2] "National Defense Authorization Act for Fiscal Year 2020, Section 835." Congress.gov, 20 Dec. 2019.

BACKGROUND

The research efforts under WRT-1057.6 represent a key initiative in the Acquisition Innovation Research Center's (AIRC's) overall strategy, illustrated in Figure 1.

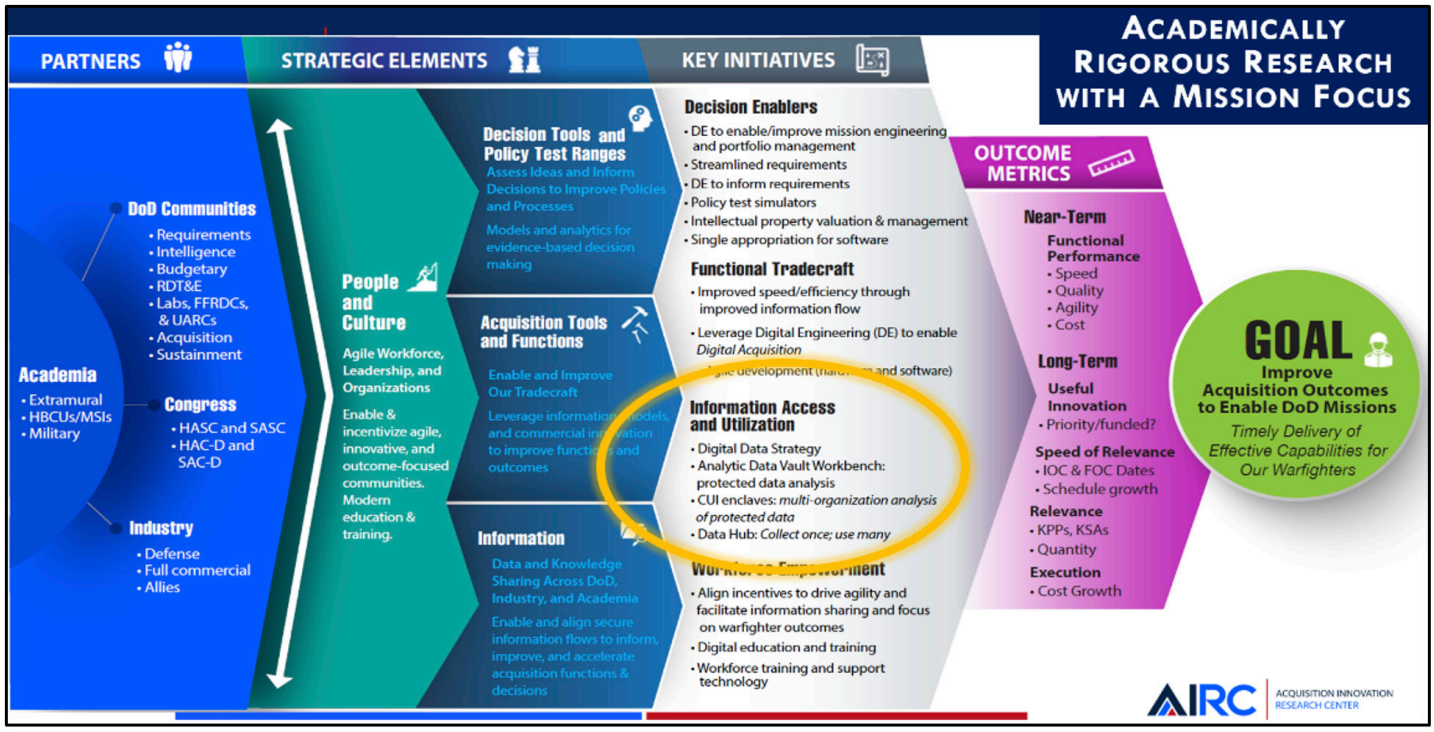

Figure 1. Focus of WRT-1057.6 Research

Efforts to establish DARCIE directly support the "Information Access and Utilization" aspects of the AIRC strategy, while the acquisition use cases directly support the Acquisition Tools and Functions aspect of the AIRC strategy.

USE CASES

Working with the sponsor, AIRC identified eight potential use cases for innovative analysis of acquisition data. The eight use cases were:

- Portfolio Management
  - » Tradespace Analysis in Integrated Acquisition Portfolio Reviews (IAPRs)
  - » Total Lifecycle Portfolio Cost
  - » Alignment of Major Defense Acquisition Programs (MDAPs) and Universal Joint Task Lists (UJTLs)
- Digital Data Storage Management
- Rapid, Serial Incrementalism
- Securing the Software Supply Chain
- Quantitative Risk Assessment
- Climate Crisis

Working with the sponsor's staff, the researchers developed materials solicit interest in pursuing the use case, specifically:

- A description of the opportunity
- An implementation strategy
- A data epic
- Milestone schedule
- Resource requirements

The technical report contains a summary of each use case below. Appendix 1 in the DARCIE Requirements Generation Report includes the detailed materials to solicit interest in the use cases.

Use Case 1: Portfolio Management. An Integrated Acquisition Portfolio Review (IAPR) is a formal process to strengthen synchronization of warfighting concepts, requirements, technologies, and program execution across portfolios of systems. The process develops enterprise acquisition portfolio roadmaps and identifies risks and interdependencies throughout the acquisition lifecycle.

> **Use Case 1a: Tradespace Analysis in IAPRs.** In this use case, data will be developed that show how acquisition decisions made earlier in the fiscal year impact the funding available for decisions that come later. Currently, the acquisition oversight process does not include a cap on the total addition to future cost for the entire tranche of MDAPs that receive Milestone (MS) B approval to enter Engineering and Manufacturing Development during a given year. Programs are reviewed one-by-one throughout the year, and programs considered later in the year are not constrained by the costs of programs approved earlier in the year, even when those cost projections increase. By addressing this shortfall, this use case will improve the effectiveness of the MS B affordability analysis.
>
> **Use Case 1b: Total Lifecycle Portfolio Cost.** Currently, research and development (R&D) funds are tracked and managed separately from sustainment funding, making it difficult to determine the overall costs associated with each system. Decisions made during development can have a significant impact on the funding needed during sustainment. For example, accepting platforms with lower than planned reliability can mean that more of those platforms are needed to achieve the same military capability. This use case will formally establish the funding relationships between R&D and sustainment to enable full lifecycle analysis of these tradeoff decisions.
>
> **Use Case 1c: Alignment of MDAPs and UJTLs.** Currently, IAPRs do not fully consider the projected combat capability to be delivered by each MDAP. The UJTL is used by all Combatant Commanders, Combat Support Agencies, and military Services to determine and report mission readiness, as well as to develop and assess joint training. UJTLs could be used to inform this important aspect of IAPR analyses.

Use Case 2: Digital Data Storage Management. This use case will accelerate the transition to digital acquisition through policy and repositories to share models and data, through development of MoDSAS—DoD's Model and Data Storage and Access System—a central site that enables enterprise access models and data through either a federated or a centralized repository.

Use Case 3: Rapid, Serial Incrementalism. This use case is designed to leverage the commercial sector's nonlinear, risk-taking approach to technology innovation. Specifically, this data product enables the rapid incrementalism, also known as agile development or development operations (DevOps). The focus is on developing data to enable Program Executive Offices (PEOs) to delay, perhaps indefinitely, the decision about whether particular aspects of performance are better delivered by hardware or software in order to take advantage of the latest products available in the market. This approach gives the PEO the ability to swap out any component or system for a better performing component at will.

Use Case 4: Securing the Software Supply Chain. This use case provides visibility to software used on a MDAP platform, as well as the software engineering tools and software used in repair and maintenance through analysis of software bills of materials (SBOMs). SBOMs will maintain accurate and up-to-date data on the origin of software code or components, place controls on internal and third-party software components, tools, and services present in software development processes, and allow for audits and enforcement of these controls on a recurring basis. SBOMs create opportunities for the DoD to enhance the acquisition process through development of appropriate controls.

Use Case 5: Quantitative Risk Assessment. This use case involves developing a quantitative risk assessment tool for managing MDAP risks at both PEO and portfolio levels. This effort integrates risk assessments into business case templates and portfolio management processes to foster faster, more predictable, and more successful acquisition programs. This effort will provide an aggregate view of the sources and potential impacts of risks within an acquisition portfolio and across the Department as a whole.

Use Case 6: Climate Crisis. This use case supports analysis of climate considerations when developing DoD's policies, strategies, partner engagements, war-gaming, modeling, and simulation in order to bolster mission resilience and reduce DoD's carbon footprint. This data product will provide analytical insights to develop strategy and answer mission-critical questions regarding DoD-managed properties, infrastructure, and installation operations.

Results

- **Use Case 1: Portfolio Management.** Methods for conducting IAPRs are somewhat ad hoc, and could benefit from a systematic review of successful practices, as well as challenges and opportunities to improve their effectiveness. The three use cases would be helpful, though a broader review may prove more useful.

- **Use Case 2: Digital Data Storage Management.** This initiative was perceived to be too similar to previous efforts for cataloguing the Department's models and simulations, which proved difficult to implement. A federated solution that leverages the growing digital engineering community may be the most viable approach to pursue.

- **Use Case 3: Rapid, Serial Incrementalism.** Though this use case offers significant opportunity for improved performance, particularly with respect to programs using the software acquisition pathway, the research team did not identify a co-sponsor to develop this use case.

- **Use Case 4: Securing the Software Supply Chain.** There is an active engagement across OSD in the area of SBOM. While no direct co-sponsorship was identified in exploratory meetings, there was enthusiasm for AIRC engagement in the ongoing activities. This use case will be explore for future data strategy engagement.

- **Use Case 5: Quantitative Risk Assessment.** The research team found that the Naval Postgraduate School (NPS) has developed methods and tools involving real options, decisions trees and quantifying knowledge value added which could support future efforts.

- **Use Case 6: Climate Crisis.** The research team met with representatives responsible for managing DoD's installations and infrastructure, but did not identify a specific use case to develop.

DARCIE

Under this task the research team developed a set of requirements for future development of DARCIE. The primary objectives were expanding access and improving the user experience. The research team developed a set of requirements for DARCIE focused in four areas: Security, Technical, System and Data (Figure 2).

Security (access, storage)	**Technical** (SW, apps, tools)	**System** (operating system, teaming, storage)	**Data** (types, formats, sources)
• CUI Compliance • User access • Sensitive data storage, sharing	• SAS • R / Rstudio • Jupyter Notebook • GitLab • Access to Linux, MAC	• Virtual desktop • Windows • Microsoft 365 w/ Teams	• Cost • System models • Test & evaluation

Figure 2. DARCIE Requirement Categories

One of the more surprising findings came from the meeting with the Rand Corporation, in discussing the tradeoffs between hosting the computational resources on-premises vs. a cloud solution. In setting up their CUI environment with a well-known cloud service provider, they found that two-factor authentication, and other user authentication methods, were taking place in different parts of the country, which led to some significant response time delays for many users. An on-premises solution would have allowed for more control to avoid such issues.

Although the research team met with Snowflake, Palantir and other analytic tool vendors, they found that open-source analytic tools would suffice for most applications for the current state of DARCIE and its data repositories. There does need to be a process, however, for research teams to be able to load licensed software for specific applications. Collaboration tools are necessary for effective work in a CUI environment.

The research team found that the demand for CUI environment remains strong. Several AIRC projects used the DARCIE environment during the period of performance. Furthermore, the research team met with representatives from Howard University, the lead for the Air Force Research Institute for Tactical Autonomy. They understand the difficulties in setting up a multi-university environment for conducting research involving CUI data. The research team shared lessons learned, and established accounts to access DARCIE.

RECOMMENDATIONS

Use Cases. The previous WRT-1047p research task recommended conducting a workshop to build a consensus around a set of high-value use cases for innovative analyses of acquisition data. The research team recommended following the precepts of Kotter's model of organizational change to break down the institutional and cultural barriers surrounding innovative sharing and use of acquisition data:

- Establish a sense of urgency
- Create a guiding coalition
- Develop a vision and strategy
- Communicate the change vision
- Empower employees for broad-based action
- Generate short-term wins
- Consolidate gains & produce more change
- Anchor new approaches in the culture

AIRC recommends that future efforts engage leadership support for the sharing and analysis of acquisition data, to create a guiding coalition, and execute data analyses.

In order to start making the changes needed for innovative use of acquisition data, AIRC recommends that the sponsor seek out like-minded co-sponsors to conduct such a workshop and/or to identify high-impact opportunities. The focus should be on generating short-term wins on issues of importance to the Department.

DARCIE. The research team recommends continued DARCIE development as the need remains strong. The resulting infrastructure will help achieve the intent of the statute that establish AIRC, which included provisions to establish data repositories and develop analytical capabilities to enable researchers and acquisition professionals to access and analyze historical data sets to support research and develop new policies and practices.

It is also recommended that DARCIE be provided as an option for all AIRC projects involving use of CUI data.

REFERENCES

Kotter, John P. *Leading Change*. Harvard Business Review Press, 2012.

"National Defense Authorization Act for Fiscal Year 2020, Section 835." Congress.gov, 20 Dec. 2019.

APPENDIX A – DARCIE REQUIREMENTS GENERATION REPORT

DISCLAIMER

The Acquisition Innovation Research Center is a multi-university partnership led by the Stevens Institute of Technology and sponsored by the U.S. Department of Defense (DoD) through the Systems Engineering Research Center (SERC)—a DoD University-Affiliated Research Center (UARC).

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 HQ0034-13-D-0004, TO#0304.

The views, findings, conclusions, and recommendations expressed in this material are solely those of the authors and do not necessarily reflect the views or positions of the United States Government (including the Department of Defense (DoD) and any government personnel), the Stevens Institute of Technology, Virginia Tech, or Virginia Tech Applied Research Corporation.

No Warranty.

This Material is furnished on an "as-is" basis. The Stevens Institute of Technology, Virginia Tech, and Virginia Tech Applied Research Corporation Virginia Tech, and Virginia Tech Applied Research Corporation make 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.

The Stevens Institute of Technology Virginia Tech, and Virginia Tech Applied Research Corporation do not make any warranty of any kind with respect to freedom from patent, trademark, or copyright infringement.

SYSTEMS ENGINEERING RESEARCH CENTER

AIRC ACQUISITION INNOVATION RESEARCH CENTER

STEVENS INSTITUTE OF TECHNOLOGY 1870

TABLE OF CONTENTS

LIST OF FIGURES

RESEARCH TEAM

Name	Organization	Labor Category
Grant Beanblossom	VT-ARC	Technical Contributor
James Brown	VT-ARC	Technical Contributor
Christina Houfek	VT-ARC	Principal Investigator
Kobie Marsh	VT-ARC	Technical Contributor
Daniel Wolodkin	VT-ARC	Technical Contributor

DARCIE STAKEHOLDER

Name	Organization	Labor Category
Stephen Adams	VT	Project Manager/Technical User
Grant Beanblossom	VT-ARC	Technical User
James Brown	VT-ARC	Project Manager
Timothy Crone	VT-ARC	Program Manager
William Headley	VT	Technical User
James McClure	VT	Technical User
Michael Rudzinski	University of Arizona	System Administrator
Dhatri Saamak	University of Maryland	Technical User
Stoney Trent	VT	Principal Investigator
Daniel Wolodkin	VT-ARC	Technical User
Justin Xu	University of Maryland	Technical User
Yabing Yang	University of Maryland	Technical User

BACKGROUND

The Defense Acquisition Research, Collaboration and Innovation Environment (DARCIE) pilot program focused on five development areas: processing, data, access, privacy, and analysis. This was done through three overlapping pieces: the DARCIE Vault, DARCIE Harness, and DARCIE Hub.

The DARCIE Vault houses the components of the DARCIE system and provides access CUI data. The Vault is maintained by Virginia Tech's Office of Export Security and Research Compliance (OESRC) who ensures DARCIE is compliant with DoD CUI requirements.

The DARCIE Harness consists of the tools that allow for acquisition data analysis. These tools give researchers access to different languages used for data science, digital engineering, as well as AI/ML. The harness also includes an intuitive User Interface team to query data appropriate for user permissions.

The DARCIE Hub consists of the available data housed within the vault and accessed via the harness. The DARCIE Hub is set up in a Hub and Spoke data model, wherein one central hub of data is attached via spokes to newly integrated data. This model will allow for a quicker ingest of data as the system grows. The hub data was scraped from the SAM.gov databank. There are currently two data spokes attached: FedBizOpps Data also coming from the SAM databank and CUI Earned Value Management data. The DARCIE Hub is designed to allow users to bring their own data or to leverage existing data as appropriate.

The goal of the DARCIE Pilot was to demonstrate an initial capability for extramural research into complex acquisition questions. An as-is model of this pilot can be seen below (Figure 1). Following the proof-of-concept pilot, the DARCIE team was tasked to scale the capability for extramural research across AIRC partners and other university as identified by the sponsor. As the potential research in support of data-driven decisions is varied, the DARCIE team recognized the existing capabilities may not meet the diverse research requirements. Prior to initiating the effort to scale DARCIE, the team focused on identifying a comprehensive requirement list. To identify the initial set of DARCIE 2.0 requirements, the team met with a number of both technical and non-technical stakeholders. As the pace of innovation in both technology and data science methods continues to accelerate, DARCIE requirements will need to be continually updated.

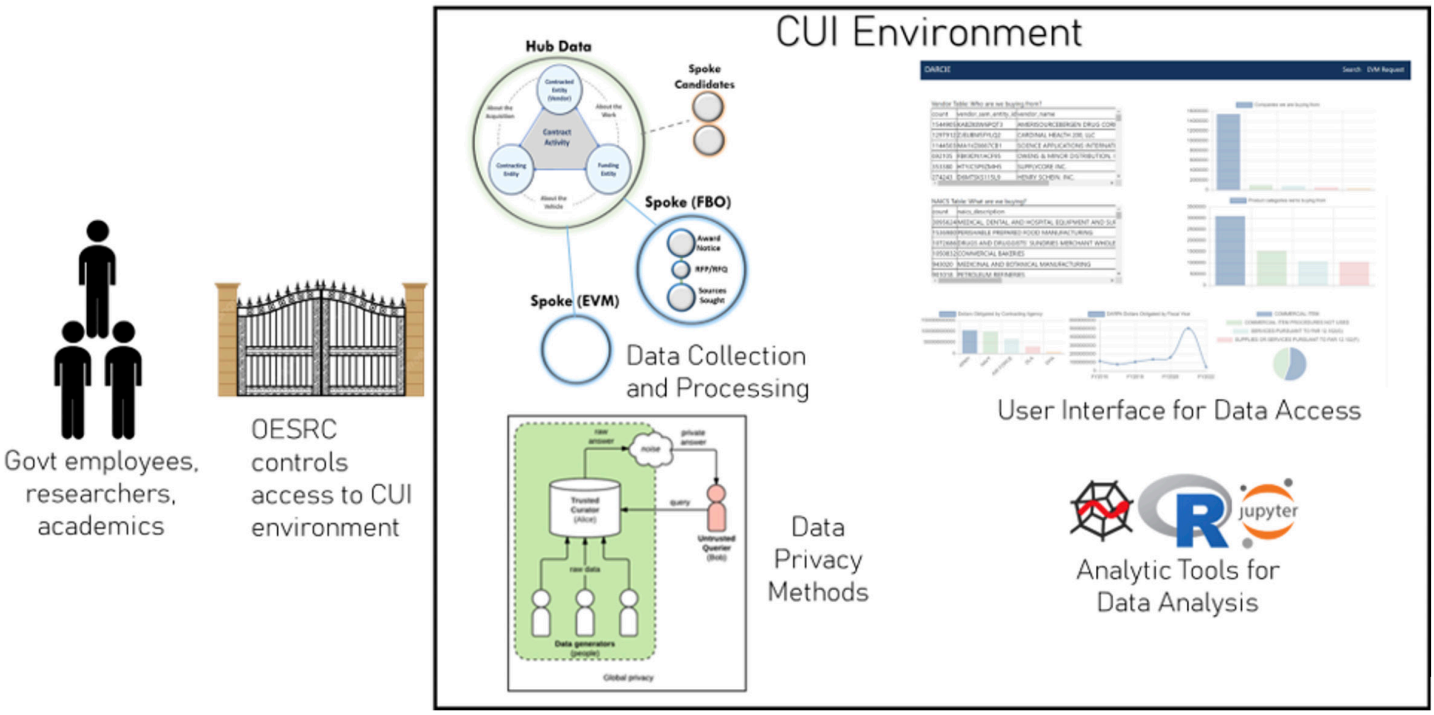

Figure 1. DARCIE Phase 1 as-is model

STUDY DESIGN

The study design created a requirements structure that splits requirements into four categories: security, technical, system, and data. Examples of each of these requirement types can be seen in figure 2 below.

Security (access, storage)	**Technical** (SW, apps, tools)	**System** (operating system, teaming, storage)	**Data** (types, formats, sources)
• CUI Compliance • User access • Sensitive data storage, sharing	• SAS • R / Rstudio • Jupyter Notebook • GitLab • Access to Linux, MAC	• Virtual desktop • Windows • Microsoft 365 w/ Teams	• Cost • System models • Test & evaluation

Figure 2. DARCIE Requirement Categories

The questions used to elicit DARCIE 2.0 requirements can be found in Appendix. Stakeholders were structured to ensure the diverse user group would identify a wide range of requirements across each of the categories based on varied perspectives and technical knowledge. The team identified a range of user types which was narrowed down to four categories: sponsors (government partners), system administrators, project managers, and technical users (e.g., data scientists, statisticians, and students). The team met with Individuals representing all four types of users. In total, 12 individuals contributed to the requirements generation. An additional meeting with academic technical users from Virginia tech with experience using the OESRC-maintained CUI environment provided an understanding of existing challenges and potential solutions. Finally, the team reviewed the as-is pilot DARCIE model and identified known, or existing, requirements identified during phase I development.

REPORT FINDINGS

The body of this report includes an overview of the findings from the requirements generation. These requirements are split into four different categories, specifying four different types DARCIE Users. These types of people include sponsors, system administrators, project managers, and researchers. Because users range from highly technical to non-technical, with diverse objectives or reasons for use, the team approached requirements identification Each type of user poses unique needs and challenges to a system such as this, and so they should each be approached with unique tailoring to make sure this system can fit everyone's needs.

SPONSORS

Sponsors present a unique challenge to any CUI system such as this. These sponsors will be external users to the system, and so a series of protocols must be in place to enable their access while ensuring the security of the rest of the system. Additionally, it is important to show the sponsors the strength and flexibility of the system, as they are where the money flows from. Sponsors need to be able to gain access to what they want quickly and easily, and there shouldn't be a long list of hoops that these sponsors need to jump through in order to get there.

The first set of requirements relating to sponsors is how they will get onto the system. One very important piece of this is a codified policy on how external users will access the system. This policy needs to be well-defined and written out, accessible for people in the projects that want to add the sponsors to the system. Not having this spelled out can cause serious delays, that could be easily circumvented if the information was readily available. This policy can involve a series of forms, or questions, that need to be filled out or answered by the sponsor, which should also be located in an easily accessible place.

A limitation that currently exists relating to getting sponsors onto our system relates to an ability to share code repositories easily with external sponsors. Git and Gitlab exist on the system, and allow for collaboration by the technical contributors, however external sharing is difficult. Having a place where these repositories can be shared is essential, so government sponsors are able to see code that's being written in support of their projects. Having a process with Gitlab to add these external sponsors would suffice, as long as it is easy to follow.

Once the sponsors are added to the system, we need to ensure that the system is easy to access. This set of requirements will be shared across each of our different types of people, but this poses special importance with the sponsors. If they are infrequently accessing this drive, whether that be to view a report or pass along data, they will be much less familiar with the process of getting onto the system, and so ease of access is important. An item that was heavily stressed by stakeholders was that the login should be instantaneous, taking just a matter of a couple of seconds once the information has been submitted. Ideally, this should be done by just clicking a desktop shortcut. The current CUI system allows for this, and there should be no trouble with something like this. One concern as well is the amount of times that two factor authentication is required. In the current CUI environment, people may have to use two factor authentication up to three times. One requirement mentioned is that once you are in the CUI environment, you shouldn't have to use two factor authentication again, just to initially get into the environment initially.

One thing mentioned a few times was the desire for the ability to log into the system by using a CAC card. This is a lower priority item, as integration might be difficult, however, this is a potentially secure way of allowing users to log in very quickly. Given the time and long list of additional requirements, however, this is something that can be pushed to a future phase when the system is scaled.

A concern held amongst many of the sponsors was the security of communications. There are many levels to this, whether that be from communicating with internal teams, communicating to sponsors, or publishing findings. The process for that is up in the air, and not entirely clear to everyone what can be discussed outside of the CUI environment. The next set of requirements, when satisfied, will help to alleviate those concerns.

The easiest way to establish clarity is to have access to government Microsoft Teams. This version of Microsoft Teams would be cleared to communicate CUI information, and so this would allow for team members to have no concerns about if they are communicating CUI material in the appropriate way. Being able to then allow sponsors or individuals external to the team access to conversations and communications in a secure way, where there is no concern about the security of the communications.

SYSTEM ADMINISTRATORS

System administrators play a key role in ensuring technical workers have an ability to perform their work within the system, while also making sure the system is working securely, and running without issues. The system administrator role is the backbone of the entire system, and the requirements that system administrators need are essential, as it dictates a lot of how the system will run itself.

Many of the requirements discussed for sponsors similarly apply to system administrators. Safe, secure, and easy login is essential for system security, which is a core component of the role of system administrator's. Having a process for adding external users to a system is key, as system administrators will be the ones to process that and add new users to systems.

However, there are requirements that uniquely apply to the role of system administrators. There needs to be tools on the system that allow system administrators to manage infrastructure and user accounts. This should allow them to monitor and change permissions, view resource usage, and check to see the infrastructure of the system is running effectively with no bugs. This should also allow the system administrators to monitor the activity of users logging in and out, making sure that there are no bad actors permeating the system, or that there is no unusual activity from the users that are on the system. Software such as the Powershell console is an example, but there are many options to explore that allow system administrators to perform these functions.

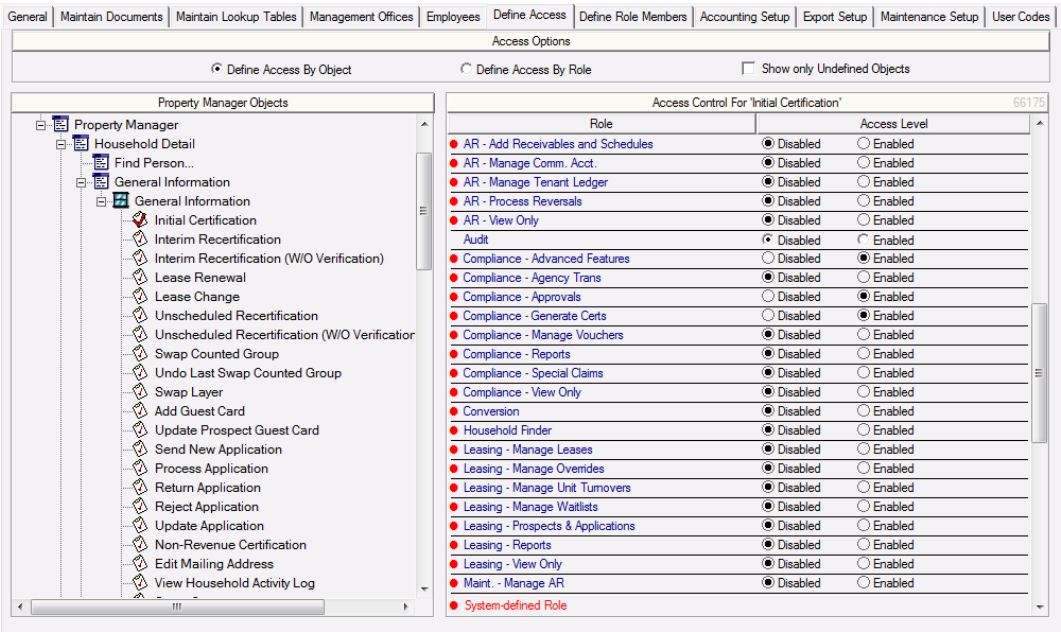

Figure 3. Access control is a key component of the job of system administrators

One requirement noted from the current CUI environment is issues with software not being regularly updated. There will be issues with files or programs no longer running because dependencies are outdated. Some process for regularly updating software needs to be a key component of this system. One way to handle that could involve Powershell scripts that automatically update software as updates are needed. Another option could be simply having an Excel sheet with the full list of software and schedule updates for each through that spreadsheet.

In addition to updating the software that currently exists on the system, there needs to be a clear process for requesting new software to be added to the environment. Within the CUI environment, there will need to be audits of software to make sure that the software is compliant with any DoD requirements, and not on any sort of blacklist. This will have to be performed by the system administrator, but before it gets to that check, there is no clear way of requesting software currently. A process needs to be outlined for this, so that users know what they need to do to request software, and system administrators know what to do once software is requested.

One issue that came up was the current system of folders within the CUI environment. For security purposes, users are concerned about the ability to see the names of every folder in the environment, whether or not you have access to that folder. There is the obvious security implication, but this also makes it difficult for users to navigate as the list of folders is extremely long, rather than just one or two folders that the user has access to.

One last key requirement for system administrators is a neatly outlined process for data backup and recovery. There needs to be a clearly defined schedule for data backup. This makes sure that there is no risk of data being lost due to system shutdowns or any damage to the system. On top of having this data backup process, there needs to be an accompanying data recovery plan, so that if data are lost from the main environment, anyone can access information about how that data can be recovered, so there isn't any delays in working progress.

PROJECT MANAGERS

Project Managers (PMs) have to worry about more than just the technical work. They also must worry about the logistical problems such as the number of users on a project, the deliverables necessary to produce a viable product, and what resources must be utilized in project development. To accommodate those needs, some requirements were sourced from project managers who might have to interact with the CUI system.

Project managers will need support within the system for project closeout. One thing highlighted was that project closeout within CUI environments usually is lacking, as there are a lot of resources available outside of the CUI environment that are not available to the project managers. Replicating these resources within the CUI environment will make the job easier for project managers at crunch time whenever they are finishing out.

In addition to this support, there needs to be an interface for project managers to assess resources and provision resources for users. This will work along with the system administrators, to make sure that the system is not getting overloaded.

Other requirements that apply to other types of people also apply to project managers and were mentioned. The ease of logging into the system also applies, as well as streamlining processes. One that is very important for project managers is the process for adding new users to the project, having that as a streamlined process that is easy to follow will be very key for project managers, as they have to add new members to the team as the project goes along.

USERS

As a user of any system, it seems that it would be a simple task to improve your experience. If a button is not responsive – just make it more so, or if navigation is stalled by having to reenter information – just make accounts that store your information for ease of use. Most, if not all, users in a system can think of one or two ways of improving it. That, however, is where a glaring issue makes itself known. If every user can think of some way to improve their experience, what do you do when these improvements contradict one another, sacrifice security for ease of access, or are otherwise infeasible to implement. This problem can be remedied by creating a set of requirements that are both sourced by users and curated by the developers of the CUI environment.

The report identifies key findings for the development of a CUI system based on requirements gathered from the users. The first set of findings revolve around the idea that in a CUI environment, the best possible outcome for users is to give them the ability to perfectly replicate their workflows. Specifically, there was an emphasis on the need for a convenient login process, seamless integration of specific software tools (VS Code, Tableau, Git) that facilitate data visualization and manipulation, and a seamless way to export results. A convenient and uncomplicated login process, that does not sacrifice security, will allow users a low stress environment and help them devote more time and energy into completing their necessary research goals and tasks. Having the software tools normally available to them also available inside the CUI environment will empower the users to provide deliverables efficiently. Lastly an intuitive way to export their non-CUI insights will allow users to deliver their findings to their project managers, and spend less time devoted to managing an unruly system and more time focused on their research.

The second set of requirements relates to the ease of traversing different data sources, since finding data relevant to the user's work must be as seamless as possible. Specifically, users requested that there is accurate metadata attached to each dataset, an intuitive search functionality, and tools for data filtering. The metadata attached to each variable in the dataset will help the user determine what the data are. It can also help with search functionality since it maybe be possible to search by tags currently existing in a dataset. An intuitive search functionality might be difficult to implement, since currently search is based on a SQL Select query. Creating a data filtering capability inherent to the can allow inexperienced users to limit the amount of data they are ingesting into their work, allowing them to focus on only relevant data.

The last set of requirements relates to accessibility and support when it comes to the environment. Specifically, users requested that they would appreciate some form 24/7 IT Support (whether that'd a FAQ or some sort of troubleshooting tool), access to qualitative data, access to both Windows and Linux environments, and the ability to SSH into other environments. The IT support option does not have to be someone physically available 24/7, since that might be difficult to do. However, having some sort of tool to help those facing issues within the environment at any time would be a boon to any of the affected users. Accessing qualitative data can also be helpful to users running language processing models or even compiling findings for reports. It is never a problem to house more data, and including qualitative as well as quantitative data will increase the usefulness of the environment. The ability to use multiple environments and OS within the CUI system can also help users who primarily develop on Linux vs Windows and vice versa, giving users the option allows them to keep their workflows. Finally, allowing users to ssh into other environments allow users to use different hardware specifications (HPC or other remote desktops with different specifications) that suit their unique needs while also having a secure way to transfer data between two environments.

CONCLUSION

Over the course of the requirements generation, the research team uncovered a lot of similarities in requirements across the different types of individuals interacting with the system. The point that resonated the strongest, was the ease of login. It cannot be stressed enough how essential it is for this system to be easy, intuitive, and fast to log in to. The appeal of a system that is complicated to login, or has issues anytime you login to it, or is slow, will dissuade users from using the environment entirely.

In addition to ease of login, one thing that was stressed was the importance of having processes and policies clearly defined, and easily accessible. Many of the problems with the current CUI environment, whether that be adding new users to COMPASS or adding new software to the system, boil down to there not being clearly defined policies and procedures. This leads to time being lost due to information exchange through emails, delaying projects, and ultimately posing a risk to the project's ability to complete work on time.

Users stressed that the draw of this system would be the data, and having data that they can only use within this environment. Many of them emphasized that they have great setups outside of this environment, and so a combination of the data and an environment that is competitive with their current environment is needed. This means that our system needs to have the software that developers would use to work on projects immediately accessible, including coding programs, visualization tools, and data analysis tools at their disposal. In addition to that, the program needs to have computation ability that makes users able to run large programs with ease. Access to the high powered computing server would be a big draw, as currently exists on the CUI environment.

Ultimately, this work highlighted that AIRC should consider every aspect of an environment like this and what would need to be in place to attract users. Having clearly defined requirements will be essential in making sure the project can progress smoothly, with a deliverable that fits the sponsors requirements at the end of it.

ANNEX A: SAMPLE REQUIREMENTS MATRIX TEMPLATE

Project Requirements

Project Name:	DARCIE	PM	Tim C.
Date:	6/14/2023	Team Lead	Chris H.
Verified Date		Co-Lead	Grant

Requirements List					Requirements Analysis				
	Requirements Description	Requested By	Category	Priority	Acceptance Criteria	Complexity	In Scope?	Test/Verification	Phase
	Feature/Function/Capability that is required for the project	Stakeholder	Functional, Technical, operational, KIPI…	High Medium Low Convert to SLA - when underway.	Requirement that must be met by Stakeholder	Not Possible High Medium Low	In Scope or out of Scope for the system	How will we test/verify	1 2 3 4

ANNEX B: STAKEHOLDER QUESTIONS

1. Why do you need access to a CUI data analysis environment?
    a. What types of projects will you be doing in DARCIE?
    b. What types of users will be on your team (e.g., government sponsors, data scientists, project managers)
        i. What would be their home organization (i.e., VT, other universities, government, or business)
        ii. Will they be US Citizens?
        iii. How will you verify/validate citizenship?
    c. How will study outcomes be evaluated?
        i. Who will do the evaluations?
        ii. What will they require to conduct the evaluations?
2. How would you like to use a CUI environment?
    a. What features or functions will be helpful for you to access and use the environment?
    b. What accessibility accommodations should be offered?
        i. 508, ADA accessibility
3. What are the system components you imagine being the most useful in DARCIE?
    a. Data
        i. What types of data should be accessible in DARCIE, and what should that data look like?
        ii. How much data do you anticipate being able to access via DARCIE?
        iii. How much storage will you require for your project?
        iv. What are your data transfer requirements?
        v. Are there any data sets that would be useful for DARCIE to have available?

b. Compute

i. What are the hardware specifications that you expect for this project?

ii. How many users/research project members do you anticipate?

iii. Are there any time requirements for input/output/data sizes/data types/CUI?

iv. Will any of your projects require the use of specialty hardware?

v. Will you need to install a licensed application?

vi. Will you need to use a specific operating system?

c. Storage

i. What data storage best practices should DARCIE feature?

ii. How would you like to store your data during the project?

iii. How will data be archived or destroyed at the conclusion of the project?

d. Communication

i. How will you communicate your findings to your government sponsor?

ii. Do you have any project collaboration requirements?

1) Virtual meeting space

2) Email

3) Instant messaging

4) Project management

5) Code sharing

iii. What type of file sharing capabilities do you need?

1) Interface with systems?

2) Exchange with systems?

3) Data upon request?

4. What are the specific security considerations do you need in place of using DARCIE?
    a. Exchange/subscription/feed/interface?
    b. How will security considerations change upon project completion and transferred to open source?
    c. What does this transfer look like?
5. What features will make DARCIE useful in future projects/proposals?
    a. Assuming the system specs will need to be communicated, what aspects of the DARCIE system should be included in the communication package?
    b. What do we need to communicate to ensure your sponsors have confidence in DARCIE security?

APPENDIX 1 – DATA PRODUCTS TO SUPPORT DEVELOPMENT OF USE CASES

DATA PRODUCT 1A: IAPR TRADESPACE ANALYSIS

Creating IAPR tradespace through transparent competition for MDAP funding increases

SUMMARY

The data product proposed in this use case will improve the effectiveness of affordability analysis in the current Milestone B (MS B) review process. The data product will enable reviewers to see how the costs of MDAPs approved earlier in the fiscal year impact Acquisition funding. It will enable a program's costs to be viewed by either Program Element, Budget Line Item (BLI), or both, thereby creating a shared view of projected costs. It will also enable Acquisition and Sustainment (A&S) to enforce consistent definitions and values for mathematical constructs like: slopes assumed for cost progress curves; overhead rates, the rate of reliability improvement; learning curves (associated with a given technology).

OPPORTUNITY SPACE

Currently, the Acquisition oversight process does not include a cap on the total addition to future cost for the entire tranche of Major Defense Acquisition Programs (MDAPs) that receive MS B approval during a given year. Programs are reviewed one-by-one throughout the year, and programs considered later in the year are not constrained by the costs of programs approved earlier in the year, even when those cost projections increase. This reduces the effectiveness of the "affordability analysis" in the current MS B review process. Acquisition funding is not adjusted in the background to accommodate the costs of the MDAPs (previously) approved by the Milestone Decision Authority (MDA), presumably by finding any additional funds required elsewhere in the budget. In fact, funding limits are present in the Budgeting activity of PPBE, not in the Acquisition process.

STRATEGY FOR IMPROVEMENT

We recommend improving the process by creating transparent competition for funding between MDAPs at MS B. Given that:

1. A program is a cost projection organized in Program Element (PE) categories (strategic forces, mobility forces, and so on) intended to show what capability the funding would support.

2. A budget is a cost projection instead organized in terms of input categories (e.g., military pay, operations, and maintenance).

The recommendation here is to create a master dataset of cost projections that can be viewed and analyzed either by: program categories; budget categories; or both simultaneously. In other words, create a shared view – 'common ground' – of projected costs such that Program decisions are visible in real time to Budgeters, and Budget decisions are visible in real time to Programmers.

Properly designed the dataset can enforce consistent definitions and values for mathematical constructs like:

- slopes assumed for cost progress curves
- overhead rates
- the rate of reliability improvement
- learning curves (in a given technology)
- speed of the savings due to obtaining (post Engineering and Manufacturing Development (EMD)) competition on subassemblies
- projected savings from the adoption of some innovation, e.g., computer aided design
- savings from a given cost reduction initiative
- speed of maturation of a given innovation (for example image recognition)

DAY-IN-THE-LIFE:

Pat Smith sat down with the day's first cup of coffee and logged in to look at the morning schedule. Pat thinks…

> "Three hours for a Milestone B review of an MDAP? <groan>. Oh well, three hours is way better than the bad old days before XYZ. Jeez. Two-to three months of data calls trying to get information. Two to three more months trying to merge spreadsheets with completely random formats. Plus trying to pull information out of PowerPoints, Word documents, reports PDFs. Cutting and pasting like a Neanderthal. Never feeling 100% confident, I was actually looking at solid evidence. Researching useful figures for overhead rates, rate of reliability improvement, learning curves. Oh yeah, those were the bad old days.
>
> Now at least the data are all in one place and I can spend my time on analysis. I can view the cost projections by force category and by budget input category. I can be confident my peers are using the same figures for overhead rates, rate of reliability improvement, learning curves. And no more arguments with the PEO about why my numbers are different than the budget numbers!
>
> OK, what does the afternoon schedule look like?"

DATA EPIC SCOPE:

The initial data epic implementation will be scoped to support Integrated Acquisition Portfolio Reviews within a given Portfolio.

Data Story - In my role as the MS B decision maker for an MDAP, when I conduct an MS B review, I want to have ready access to a robust set of program data, budget data, and standard guidance. This includes:

- a repo that provides authoritative data on the (specific) MDAP Program, capable of being displayed according to in force categories (strategic forces, mobility forces, and so on) showing what capability the funding would support, and also capable of being displayed according to budget input categories (e.g., military pay, operations and maintenance).
- in the same repo, authoritative data on program/budget decisions made with respect to the costs of programs approved earlier in the fiscal year, so I can maximize the effectiveness of my "affordability analysis" in the current MS B review
- approved, consistent definitions and values for mathematical constructs like:
  - slopes assumed for cost progress curves
  - overhead rates
  - the rate of reliability improvement
  - learning curves (in a given technology)
  - speed of the savings due to obtaining (post EMD) competition on subassemblies
  - projected savings from the adoption of some innovation, e.g., computer aided design
  - savings from a given cost reduction initiative
  - speed of maturation of a given innovation (for example image recognition)

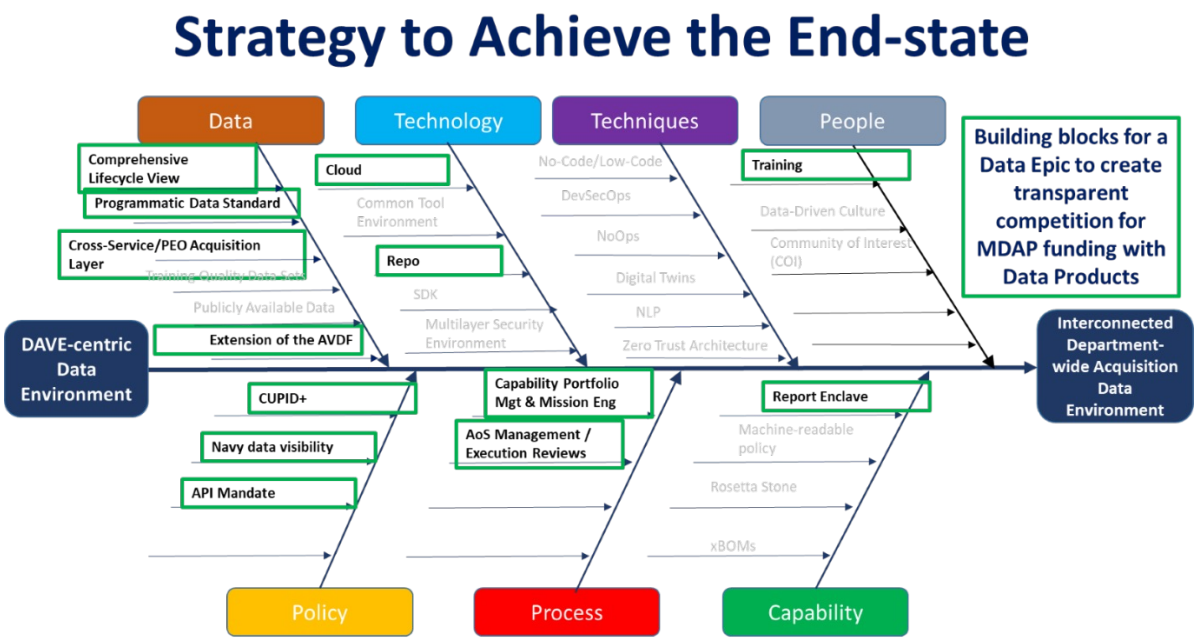

Figure 4. Strategy for IAPR Tradespace Use Case

The Acquisition Digital Data Strategy itemizes several building blocks (Figure 4) that will be combined to prosecute this Data Epic. This data epic can be developed using an Agile approach. Once the capability has been pressure-tested in IAPRs within a portfolio, the scope will be enhanced to support trade-space across the full set of Acquisition Portfolios.

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. Executing the tasks under the Enabling Objective will move the acquisition community further towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 1B: TOTAL LIFECYCLE PORTFOLIO COST

The keys to optimizing total lifecycle cost of MDAP portfolios

SUMMARY

This data product will deliver a complete view of the risks – including budget, technology, and supply chain risks – throughout the complete lifecycle of each weapon platform about to be put into operation. That risk profile includes sustainment, which can last for decades and generate 70-80% of the lifecycle cost of the MDAP. This data product addresses analytical workloads like: "How much did we spend on platform XYZ in the last twelve months?" – AND – "If I had another million dollars to spend, which of these two platforms in my Acquisition portfolio would deliver the greatest combat capability for that spend?" – AND – "Which (30%) of my development & procurements decisions (for this MDAP) will have the greatest impact on projected lifecycle costs?"

OPPORTUNITY SPACE

Today we ask Services to accept risk with each weapon platforms, especially with new weapon platforms about to be put into production. But we don't give the Services a complete view of the risks. For example, a platform may involve commitments to multiple, named suppliers in multiple Congressional districts. That sourcing profile implies supply chain risks, and budgetary risks, in sustainment which can last for decades and generate 70-80% of the lifecycle cost of the MDAP. That sourcing profile also implies obsolescence risks that are difficult to objectively quantify today. Further, decisions made within the first 30% of development and procurement determine 70% of the lifecycle cost. Even when a platform is operating at scale, we struggle to answer simple questions like:

> "How much did we spend on platform *XYZ* in the last twelve months?"
>
> "If I had another million dollars to spend, which of these two platforms in my Acquisition portfolio would deliver the greatest combat capability for that spend?"
>
> "If I had another million dollars to spend, which of the platforms in this particular kill chain should get that money to deliver the greatest improvement for the system of systems?"
>
> "Which (30%) of my development & procurements decisions (for this MDAP) will have the greatest impact on projected lifecycle costs?"

STRATEGY FOR IMPROVEMENT

The root cause of the problem is the lack of common data keys across the various systems holding the data. While CUPID will standardize the platform identifier for Acquisition IT systems Department-wide, there is no equivalent data key (for platforms) in accounting standards, program reporting standards, personnel standards, or accounting standards. (See Figure 5 below.) Worse, in some of the authoritative sources the information is in an unstructured text field or a blob text field.

Figure 5. Budget relationship difficulties

The way forward is to harness the power of Machine Learning (ML) and Artificial Intelligence (AI) to build a "Rosetta Stone" capability that identifies, extracts, and interpolates the appropriate information. There are three important categories of AI algorithms to train and leverage in this data epic:

1. Natural Language Processing
2. Image Recognition
3. Generative AI

The Acquisition Digital Data Strategy itemizes several building blocks that will be combined to prosecute this Data Epic. This data epic can be developed using an Agile approach. Once the capability has been pressure-tested in IAPRs within a portfolio, the scope will be enhanced to support trade-space across the full set of Acquisition Portfolios.

DAY-IN-THE-LIFE:

Acquisition PEO professional Riley Jones has been working hard to avoid the valley of death. They are guiding an MDAP for a new tactical communications network through the Middle Tier of Acquisition (MTA) pathways. They have actively collaborated with the Army, Marine Corps and Air Force, all of whom express commitment to deploying the new weapon platform when it's mature. The next conversation is today. Riley knows the Service is concerned about the risk they are accepting, particularly with respect to the 40-year sustainment tail.

To prepare for the Teams meeting, Riley logs into the Acquisition Data Environment then clicks on the Total Lifecycle View button. The screen displays the full system Bill of Materials, with links to the entire supply chain displayed on a Continental United States (CONUS) map. With the next click Riley gets a profile of the financial viability of each supplier. The same window shows annual sustainment costs (with a dial to raise or lower the planned quantities). Projected obsolescence is shown on the timeline with the costs. Project alternative technologies, based on Venture Capital funding of product development in Silicon Valley, are shown on the same timeline. Riley thinks, "OK, I'm ready for the Teams call."

DATA EPIC SCOPE:

Initial scope is the Aegis. As an Acquisition Portfolio Manager, when I review my portfolio, I want a data product that enables me to answer the following questions:

- How much did the DoD spend on this Acquisition Category (ACAT) 1 weapon platform in the last 12 months? What is our mathematical confidence in the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- How much combat capability did we get for that spend? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- If I had another $1 Million to spend, how much more combat capability could I generate? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?

Expand scope to include all ACAT1 assets in the portfolio (for example, Long-range Precision Fires). As an Acquisition Portfolio Manager, when I review my portfolio I want a data product that enables me to answer the following questions:

- How much did the DoD spend on each platform in the last 12 months? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- How much combat capability did we get for that spend? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- If I had another $1 Million to spend, where (in the portfolio) should I spend it & how much more combat capability could I generate? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?

Expand scope to include the rest of the assets in the Aegis (primary) kill-chain for the mission threads most often prosecuted by Combatant Commanders today (or those mission threads prioritized by Joint Capabilities Integration and Development (JCIDS) for tomorrow). As an Acquisition Portfolio Manager, when I and other Acquisition Portfolio Managers discuss our portfolios with the USD A&S, I want a data product that enables me to answer the following questions:

- How much did the DoD spend on each asset in the last 12 months? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- How much combat capability did we get for that spend? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- If I had another $1 Million to spend, where (in the portfolio) should I spend it & how much more combat capability could I generate? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?

Expand scope to include the rest of the assets in the other Long-Range Precision Fires (primary) kill-chains. As an Acquisition Portfolio Manager, when I and other Acquisition Portfolio Managers discuss our portfolios with the USD A&S, I want a data product that enables me to answer the following questions:

- How much did the DoD spend on each asset in the last 12 months? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- How much combat capability did we get for that spend? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?
- If I had another $1 Million to spend, where (in the portfolio) should I spend it & how much more combat capability could I generate? What is our mathematical confidence on the completeness of this answer? What steps do we take to improve our confidence? Is the juice worth the squeeze?

CONCLUSION

Completion of this use case will produce a new capability / function that enables AE to better accomplish the mission objectives outlined in the DoD Data Strategy. Prosecuting the use case will move the acquisition community further together towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 1C: ALIGN MDAPS WITH UJTL MEASURES

Expanding IAPR trade-space by integrating Universal Joint Task List (UJTL) measures of projected combat capability delivered

SUMMARY

This data product will enable IAPRs to incorporate the use of an effects/kill chain framework (the UJTL) to calibrate each MDAP's output of combat capabilities. This will permit A&S Portfolio leaders to assess the impact of their portfolio on future mission performance. Published by JS J7 on behalf of the Chairman of the Joint Chiefs of Staff (CJCS), the UJTL is used by all Combatant Commanders, Combat Support Agencies, and Services to determine and report mission readiness and performance, as well as to develop and assess joint training. In other words, this data product enables A&S Portfolio leaders and Combatant Commands to 'speak the same language'.

OPPORTUNITY SPACE

DoDD 5000.01 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 13 Capability Portfolios with the "mission is to use Capability Portfolio Management to analyze, manage, and inform acquisition and resourcing decisions in platform and weapon portfolios …" Recent changes to support capability portfolio management, such as the IAPR efforts, are working to produce actionable advice for senior leaders. Currently, MDAP program performance is measured in terms of cost, schedule and risks. But those performance measures do not include any objective measures of results – the combat capability delivered by the program. Nor can those performance measures be expressed in terms of contribution to mission. Multiple capabilities can map to multiple missions. Missions are rapidly evolving and a comprehensive mapping of capabilities to missions does not exist. Therefore, it is difficult for Portfolio leaders to assess the impacts of capability portfolio performance on mission performance currently.

STRATEGY FOR IMPROVEMENT

Meanwhile, the DoD does have objective measures of combat capability, which measures are designed to be mapped to missions. Published by JS J7 on behalf of CJCS, the UJTL is used by all Combatant Commanders, Combat Support Agencies, and Services to determine and report mission readiness, as well as develop and assess joint training. Each Universal Joint Task has an explicit description with a list of sample measures (See Figure 6 below). These measures are the basis for describing varying levels of joint task performance. When a given commander's approved measures are combined with these criteria for performance, they yield a task standard which describes how well a joint organization or force must perform a joint task under a specific set of conditions in the operational environment.

Consider the F-16. As an example, UJTL OP 6.1.4 Conduct Defensive Counterair (DCA) might apply. 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 for assessing the capability.

#	Unit	Description
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.

Figure 6. Defensive Counterair UJTL

The DoD uses results-based performance measures (e.g., military recruiting success, readiness ratings, tracking obligations) to measure the impact of its spending. But today the Acquisition community lacks results-based performance measures for MDAPs. This data epic lays out steps to strengthen MDAP management by establishing results (i.e., combat capability delivered) based performance measures that apply the UJTL to each MDAP. This will not only give both portfolio executives & PEOs a more robust set of performance objectives, but it will also align those objectives with both JCIDS objectives and the National Defense Strategy (NDS).

The Acquisition Digital Data Strategy itemizes several building blocks that will be combined to prosecute this Data Epic. This data epic can be developed using an Agile approach. Once the capability has been pressure-tested in IAPRs within a portfolio, the scope will be enhanced to support trade-space across the full set of Acquisition Portfolios.

DAY-IN-THE-LIFE:

Halsing Verma is an A&S Portfolio Executive preparing for the latest IAPR. All of the MDAPs show green on the dashboard, but there's a "Look at me!" flag turned on for the fourth one in the list. Expecting bad news, Halsing obliges by double-clicking on the icon. Halsing thinks 'outloud'…

> "What's the problem then?
>
> The schedule is fine. ✓·
>
> Costs are within budget. ✓·
>
> We've obligated all our funds. ✓·
>
> The risks haven't changed. ✓·
>
> So why the flag?
>
> Oh, there it is. The PEO says that if we let their schedule slip by six months they think they can double their goal for *UJTL OP 6.1.4 measure M4*. What the heck is that? Oh, M4 is 'Percentage of enemy air attacks detected early enough to allow engagement.' And the PEO says they're going to double it? That's a bold claim. What was the original target? Ah it was 75%. They're going to deliver 150% Yeah, no. I need to talk to the PEO."
>
> <click> There goes an Outlook invite to a Teams meeting with the PEO.

DATA EPIC SCOPE:

The initial implementation will be scoped to support Integrated Acquisition Portfolio Reviews within a given Portfolio. The actions are:

1. set up an automated feed of the master task list from the UJTL portal
2. produce a data product that enables the portfolio manager to assign UJTL goals to each MDAP in the portfolio, and specific targets for each performance measure under the UJTL goal
3. apply the new data product during the next round of IAPRs. Use the IAPRs to confirm the selection of UJTL goals, and to assign performance targets for all measures in the chosen goals
4. add the UJTL goals and performance targets to the data product feeding Advana

Data Story – As an Acquisition Portfolio Manager, when I conduct an IAPR I want a data product that enables me to answer the following questions.

- What impact will a schedule change have on the combat capability delivered?
  - » In other words if I field the capability 6 months earlier what level of performance will I get?
  - » Is it better to get 50% of the expected performance in 6 months, than to wait 24 months to field 100% of the expected performance?
- If I hold constant the results we need, what is the trade-space between funding and risk?
- If I hold constant the results we need, what is the trade-space between schedule and funding?
- What is my mathematical confidence in the validity & accuracy of these answers?

Subsequently expand scope to include the funding profile for each ACAT 1. Currently there is no standardize way to even identify which Budget/Program Objective Memorandum (POM) lines across the appropriations are funding the programs. Further, a platform could be funding its modification completely on its own or systems/ subsystems could be funded outside the platform with no apparent rule to drive consistency. The actions are:

- pull funding data from the data products produced by the data epic named “Develop the keys to optimizing total lifecycle cost of MDAP portfolios.”

Subsequently expand scope to include the rest of the assets in each ACAT 1 kill-chain for each mission thread. The actions are:

5. create an additional data product that assembles performance goals across the ACAT 1 kill chains; engage Combatant Commanders (COCOMs) in a review of the performance profile for the ACAT 1 kill chain MDAPs in typical (or prioritized mission threads). Adjust performance targets accordingly.
6. track the adjustments and reflect their source in the data product feeding Advana
7. Extend approach to ACATs 2-4

Data Story – As an Acquisition Portfolio Manager when I discuss my portfolio of MDAPs with Combatant Commanders, I want a data product that enables me to answer the following questions.

- When will each MDAP deliver new combat capability?
- How much will that new capability improve the performance of the COCOMs' kill chains in typical (or prioritized mission threads)?
- What is my mathematical confidence in the validity & accuracy of these answers?

CONCLUSION

Completion of this use case will produce a new capability / function that enables AE to better accomplish the mission objectives outlined in the DoD Data Strategy. Prosecuting the use case will move the acquisition community further together towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 2: DIGITAL DATA STORAGE MANAGEMENT

Accelerate transition to digital acquisition with repositories to share models and data

SUMMARY

This data product will accelerate the improvement of acquisition outcomes using models as an improved information base for acquisition and sustainment decisions and functions. It facilitates archiving, sharing, and management of acquired capability models and data, leveraging commercial sector's best practices for managed repositories. This use case calls for development of MoDSAS—DoD's Model and Data Storage and Access System—a central site that enables enterprise access either through a federated or centralized repository. This effort develops policy and technical approaches, exploring the use of existing sites such as the Chief Digital and AI Office's (CDAO) git repositories on Non-classified Internet Protocol Router Network (NIPRnet), Secret Internet Protocol Router Network (SIPRnet), and Joint Worldwide Intelligence Communication System (JWICS). It will explore the tradeoffs between federated and centralized implementations. A federated approach allows organizations to establish their own repositories, but a centralized approach may reduce redundancies and improve efficiency.

OPPORTUNITY SPACE

A nascent digital acquisition community is getting established in several DoD organizations—primarily in the systems engineering community. The Army's PEO Aviation is providing Government-Furnished Information (GFI) using common models across multiple acquisition programs. The Naval Air Systems Command is investigating the use of models to accelerate air worthiness qualification. Although some sharing of data and models is occurring within organizations, there are currently no enterprise policies or repositories to support these activities.

Beyond the use of models at the program level for engineering, they have great promise for improving other acquisition functions, such as test and evaluation (T&E), cost estimating, maintenance, logistics, and upgrades. Moreover, the use of models across multiple systems can improve analysis and decision making of portfolios of capabilities. Both expanded uses hinge (in part) on the archiving, management, and availability of models and associated data beyond program offices. Thus, establishing the requirement and providing the facilities for an enterprise model repository will accelerate the Department's transition to digital acquisition.

STRATEGY FOR IMPROVEMENT

Working with the CDAO and OUSD(R&E)—with insights from the military departments—we will develop a policy and approach for sharing acquisition models and data through a MoDSAS. MoDSAS provides central access to a federated or centralized storage system modeled on industry best practices for open-source development (e.g., git repositories or federated databases). Organizations would be able to leverage the CDAO's NIPRnet, SIPRnet and JWICS git repositories or develop their own repository. MoDSAS developers will take advantage of lessons learned from the model repository/catalog efforts of the OUSD(R&E) Modeling and Simulation Coordination Office (MSCO) as well as Massachusetts Institute of Technology's (MIT) model curation research.

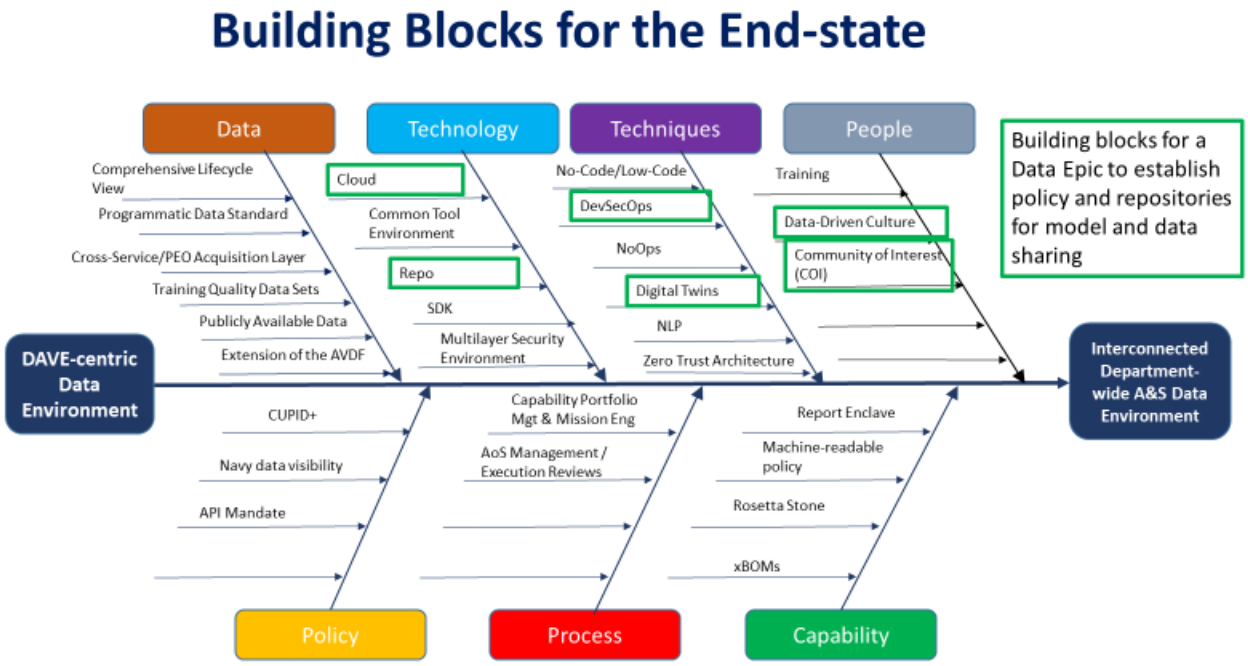

Figure 7. MoDSAS Use Case Strategy

DAY-IN-THE-LIFE:

Sally Citizen is an acquisition professional in a program management office putting together a solicitation for an improved radar system for an ACAT 1D platform. The improved system will support multiple platforms, which will reduce sustainment costs. Processor and receiver-transmitter technologies have been evolving rapidly, so she is putting together a request for information (RFI) to help define achievable requirements for the solicitation. Fortunately, the acquisition platforms have put together models and interface specifications which will be included with the RFI as GFI. The current DoD acquisition policies and guidance will also be embedded in the models as GFI as well. Sally is checking to be sure that the latest policies and guidance are included in the GFI.

> "Hmm, let's see if there is anything new to include in the GFI." She does a search on MoDSAS, and uses her CAC to access the site. "I heard that there is some new cyber survivability guidance from the Joint Staff. Or perhaps it was policy. Wasn't NAVAIR tracking that?" She types 'cyber survivability' into the MoDSAS search bar, which returns a link to a NAVAIR git repository. She uses her common access card (CAC) to authenticate into the NAVAIR git repository. "It looks like there is new guidance, but the updated process model is still in beta. I'll include the new guidance and keep the current process model." She downloads the new guidance. "What were the other things in that list?" She moves on to her next item to check.

Pedro Patriot is reviewing recent funding decisions for Joint Capability and Technology Demonstration (JCTD) projects to determine if upcoming training exercises would demonstrate a different mix of capabilities to defeat a new hypersonic cruise missile threat, which was recently projected to become available in the next five years.

> "It looks like that new hyper-spectral imager got funded, and will launch in time to participate in Northern Edge 27. Let's see if they've updated the operational architecture for this mission thread in the exercise." Pedro authenticates with his SIPRnet token and does a search on 'Northern Edge 27' in MoDSAS. MoDSAS returns a link to the Joint Staff git, which is hosted by the CDAO. Pedro searches for the mission thread for cruise missile defeat, and looks to see if the JCTD is included. Unfortunately, it hasn't been updated yet, but he could run the mission thread analysis with the updated Probability of Detection (Pdetect) figures for the JCTD. Pedro starts reviewing the other systems involved in the mission thread to see if there are any other planned updates to the exercise that he was unaware of. As he reviews the mission thread, he concludes that the updated mission thread analysis might have an impact on the upcoming missile defense portfolio review.

DATA EPIC SCOPE:

As an Acquisition professional working with models and data, I want straightforward methods to identify and access authoritative models relevant to my needs. MoDSAS helps in providing a single location for searching the federated set of git repositories.

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. Executing the tasks under the Enabling Objective will move the acquisition community further towards a data-driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 3: RAPID SERIAL INCREMENTALISM

Increasing the agility of acquisition programs through data that facilitate rapid, serial incrementalism

SUMMARY

This data product will enable acquisition professionals to leverage the commercial sector's nonlinear, risk-taking approach to technology innovation. Specifically, this data product enables the digital transformation discipline of rapid incrementalism (in the vernacular, Agile development and DevOps) combined with improving decision-making based on enhancing and aligning data around desired strategic outcomes. This data product is a strong complement to Earned Value. MTA Pathways already push authority, tools, and training to the 'tactical edge' of acquisition. This data product will push data (on outcomes and learning) to the tactical edge of acquisition. This will enable the PEO to delay, perhaps indefinitely, the decision about whether a particular performance parameter is delivered by hardware or software in order to take advantage of the latest products available in the market. This gives the PEO the ability to swap out any component or system for a better performing component any time one emerges.

OPPORTUNITY SPACE

Today's Acquisition professionals are constrained to linear, serial processes. These processes require them to develop detailed, specific requirements category by category. This approach inhibits considerations of the entire system, let alone the system of systems in which the MDAP will operate. It also inhibits agility in the life of the program. Meanwhile the commercial sector has convincingly demonstrated that the organization which adopts the latest technological advances is the best poised to dominate in future competition. The commercial sector has also demonstrated that today's nonlinear, risk-taking approach to technology innovation requires agility, creativity, and speed. Meanwhile formal acquisition programs don't have access to the data that would enable increased responsiveness and investment decision space.

STRATEGY FOR IMPROVEMENT

In the commercial sector, digital transformation disciplines focus on rapid incrementalism (in the vernacular, Agile development and DevOps) combined with improving decision-making based on enhancing and aligning data around desired strategic outcomes.

Step 1 - deliver a capability as early as possible.

Step 2 - use user feedback to prioritize future development, thereby focusing more on realized value. In this approach there are two key questions: What new functionality did we actually just deliver? And what did we just learn about what works vs. what doesn't? This approach requires flexible, bottom-up data frameworks built to reveal tangible outcomes and learning. By contrast, practices like Earned Value measure inputs and activities, not outcomes or learning. In other words, this data product is a strong complement to Earned Value.

MTA Pathways already push authority, tools, and training to the 'tactical edge' of Acquisition. The next step is to push data (on outcomes and learning) to the tactical edge as well. This can enable the PEO to delay, perhaps indefinitely, the decision about whether a particular performance parameter is delivered by hardware or software in order to take advantage of the latest products available in the market. In the commercial sector, digital transformation leaders refer to this as a 'plug-and-play' approach to technical architectures. It gives them the ability to swap out any component or system for a better performing component any time one emerges. This Data Epic aims to give Acquisition professionals the same flexibility throughout the lifecycle of the weapon platform. DoD creates a wealth of data that can inform decision making, including exercise results, test and evaluation data, modeling and simulation data, digital twin outputs, and business operations execution data.

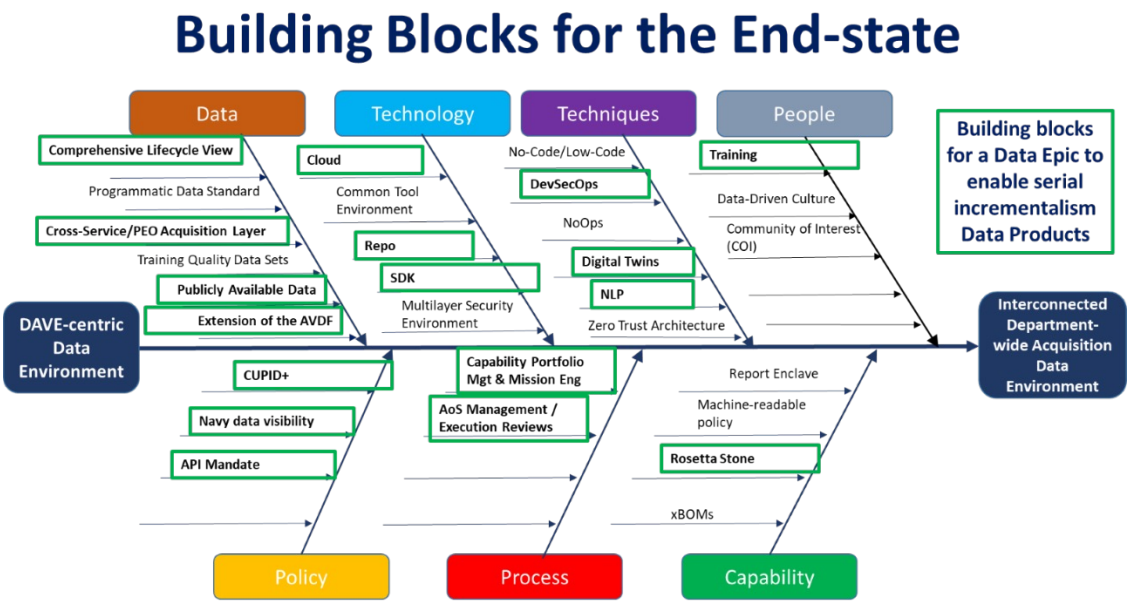

Figure 8. Serial incrementatlism use case strategy

DAY-IN-THE-LIFE:

Chris Jane is an Acquisition professional doing a portfolio review on MDAP in the MTA pathways. The PEO has been experimenting with a new approach to introducing additional competition in the design phase for a new targeting sensor. Multiple Silicon Valley start-ups are piloting their preferred mixes of custom microchips and custom image recognition algorithms for the targeting sensor. They all know what the performance test parameters are and are free to put their best engineering talent to work. The test ranges are automated and produce highly calibrated measures of performance. The PEO's own engineering just uploaded a new set of test parameters into the test range. Time to see the results. Hayden opens the A&S data repo, clicks on the Data Products, and shares the screen in a Teams meeting.

> "There are the results from the new test. Wow, look how that flipped the performance ratings. Group B jumped from last place to first. Group C slipped from first place to second, but only just. And Group F slid to last place. Huh.
>
> So now let's switch the way we're looking at this. If we map the test parameters to the conditions we expect to face against pacing adversaries, the Group C solution is perfect for a high intensity fight inside the First Island Chain. Group B is perfect for a bit of push-and-shove in Eastern Europe. And both supplier Groups can produce the quantities we need for the next phase of development."

DATA EPIC SCOPE:

As an Acquisition professional managing MDAPs in MTA Pathways, I want data products that enable me to delay, perhaps indefinitely, the decision about whether a particular performance parameter is delivered by hardware or software in order to take advantage of the latest products available in the market. The data products will do this by answering questions like:

1. How do these alternative mixes compare in exercise results?
2. What happens to performance when the engineers tune test and evaluation data to different operational parameters?
3. Where are the limits of our modeling and simulation data?
4. How good are our digital twin at representing the real world?

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. (See Reference section below for implementation plan template.) Executing the tasks under the Enabling Objective will move the acquisition community further towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 4: SECURE THE SOFTWARE SUPPLY CHAIN

Securing the software supply chain for MDAP portfolios

SUMMARY

The data product proposed in this use case will provide visibility to software in an MDAP platform, as well as to the supporting software engineering tools and software in repair and maintenance equipment for those platforms. It will deliver an automatically updated 'bill of materials' for its software supply chain starting with the embedded software and will map these software 'bills of materials' to databases of known software vulnerabilities like malware. It will generate alerts when new vulnerabilities are identified, to enable remediation of those vulnerabilities.

OPPORTUNITY SPACE

President Biden issued E.O. 14028 to improve the nation's cybersecurity. In Section 4, it calls for enhancing software supply chain security by software manufacturers producing Software Bill of Materials (SBOM) for each product, and software customers maintaining accurate and up-to-date data, provenance (i.e., origin) of software code or components, and controls on internal and third-party software components, tools, and services present in software development processes, and performing audits and enforcement of these controls on a recurring basis. The software manufacturing industry is broadly mature with respect to producing SBOMs. This creates the opportunity for the DoD to enhance the Acquisition process by building out appropriate controls.

STRATEGY FOR IMPROVEMENT

We recommend beginning with a focus on the crown jewels of software in an MDAP platform: embedded software, including Basic Input/Output System and Unified Extensible Firmware Interface. This software is most difficult to inspect for vulnerabilities and most difficult to remediate when vulnerabilities are discovered. Accordingly, prevention (of software defects) is critical. We would also like to conduct a parallel pilot looking at an alternative software associated with a semi-autonomous drone. The building blocks for both Data Epics are shown in Figure 4 and would remain consistent if a different ACAT platform or SBOM exemplar is selected. In addition, because we are part of a larger Interagency organization involved in SBOM/XBOM work, we will integrate/adapt our activity across efforts in CDAO, A&S, Chief Information Officer (CIO), R&E, Director, Operational Test and Evaluation (DOT&E), Department of Energy (DOE), and Cybersecurity and Infrastructure Security Agency (CISA). Some pilots and communities exist that might benefit or benefit from our Acquisition efforts. Developing a community of interest in which we are a contributing player is key to understanding how best to execute this data product.

The benefit of conducting two parallel efforts lies in their comparability. It also allows us to consider the impact on the acquisition process of different types of acquisition paths – and how SBOMs and their associated toolsets would impact decision making. For example, how will we automate data feeds so that people can see the most current data with the least effort. How will we present the data to support actionable decision-making? How will we manage and model the data as it changes for the greatest insights? How will we enable a balance between uniformity of insight while permitting the Services the flexibility they require? The acquisition process is unique in the sense that modifications and updates are not enacted of the lifecycle of a product. So, once it leaves the acquisition sphere of influence and moves into sustainment, it begins to travel across the force and tracking SBOM material can become increasingly complex. Having an origination point for capabilities will be critical to future tracking endeavors.

DAY-IN-THE-LIFE:

Ashley True is preparing for the latest IAPR. The first step? Review the progress made on managing software vulnerabilities. Ashley opens the SBOM control tower in the DoD Department-wide Acquisition Data Environment.

> "Hmm. I had four platforms with improvement targets. Three of the four have now reached their intermediate goal of 97% visibility of SBOMs, with 85% confidence levels. That's great. I'll make a note to praise the improvement. The fourth platform beat the goal. They're at 99% visibility with 95% confidence. Now to set the next round of targets…
>
> Now to rotate the view from platforms to suppliers. Let me see. The first 5 suppliers are at the targets we agreed. Same for seven through ten. But supplier number six is just not cooperating. Time to get tough with them.
>
> <click> There goes a note to Contracting Offices dealing with that contractor.
>
> <click> There goes a note to the PEOs buying from that contractor.
>
> <click> There goes a note to our lawyers."

DATA EPIC SCOPE:

Build and automate updates to a repo of embedded software for an MDAP platform as the data in the source systems evolves. Figure 9 below illustrates a sequence of data gathering and automation that delivers insights of value with each data story. It emphasizes statistical calculations of completeness and confidence. It includes discovery for the full range of software vulnerabilities.

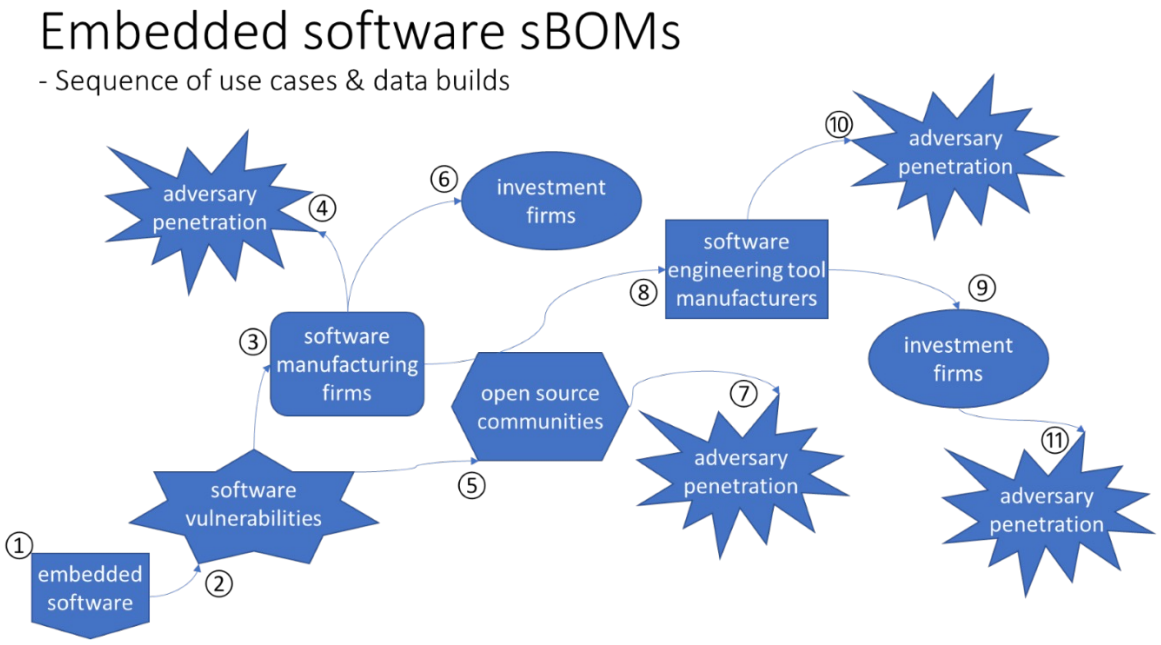

Figure 9. SBOM sequence of use cases and data builds

As an Acquisition Portfolio Manager, when I conduct an IAPR I want a data product that enables me to answer the following questions.

1. Based on an initial SBOM for embedded software, what is our mathematical confidence that we have identified the full SBOM for the embedded software in the platforms in my portfolio?
    a. What steps can I take to improve my confidence?
        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for Manufacturing Execution System (MES) or Supply Chain Management (SCM))? Is the improvement in confidence worth the cost of the data, including automation?
2. Expand scope to include software vulnerabilities - What is our mathematical confidence that we have identified the software vulnerabilities in the full SBOM for the embedded software in the platform?
    a. What steps can I take to improve my confidence?
        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?
    b. What steps can I take to improve reduce my vulnerabilities? Which mitigations provide the widest or most valuable reduction at the lowest cost?
3. Expand scope to include software manufacturing firms - What is our mathematical confidence that we have identified the companies producing the software in the full SBOM for the embedded software?
    a. What steps can I take to improve my confidence?
        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?
4. Expand scope to include open source software communities - What is our mathematical confidence that we have identified the open source communities producing the software in the full SBOM for the embedded software in receiving platforms?
    a. What steps can I take to improve my confidence?
        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

5. Expand scope to include the investment community behind software manufacturers - What is our mathematical confidence that we have identified VC/Program Element (PE) firms funding startups of the companies producing the software in the full SBOM for the embedded software in receiving platforms?

    a. What steps can I take to improve my confidence?

        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

6. Expand scope to include adversary penetration of investment community behind software manufacturers - What is our mathematical confidence that we have identified adversary-penetration (including by funding VC/PE startups) of the companies producing the software in the full SBOM for the embedded software in receiving platforms?

    a. What steps can I take to improve my confidence?

        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

7. Expand scope to include software engineering tools - What is our mathematical confidence that we have identified the Software Engineering tools used by companies producing the software in the full SBOM for the embedded software in receiving platforms?

    a. What steps can I take to improve my confidence?

        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

8. Expand scope to include firms manufacturing software engineering tools - What is our mathematical confidence that we have identified the firms manufacturing the Software Engineering tools used by companies producing the software in the full SBOM for the embedded software in receiving platforms?

    a. What steps can I take to improve my confidence?

        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx, another government agency, a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

9. Expand scope to include adversary penetration of investors in firms manufacturing Software Engineering tools -What is our mathematical confidence that we have identified adversary-penetration (including by funding VC/PE startups) of the companies producing the Software Engineering tools used be the companies producing the full SBOM for the embedded software in receiving platforms?

    a. What steps can I take to improve my confidence?

        i. Test case: Can I improve my confidence by pulling data from a new source (Open Source company *xxx; another government agency; a public software repo xxx*; supplier's private repos; suppliers enterprise applications for MES or SCM)? Is the improvement in confidence worth the cost of the data, including automation?

ALTERNATIVE / PARALLEL DATA EPICS

10. **Opportunity:** The obsolescence and capability upgrade timelines for modern electronic components are very short, and in a resource-constrained environment, the desire for cost-effective, long-lived major weapon systems is challenged by these short timelines. New defense acquisition programs must therefore consider these resource constraints and plan for design growth. Doing so involves maintaining design margins within system parameters, as well as considering future periodic component upgrades to enable long service lives for systems by incorporating the latest technology hardware obsolescence timeline, or life-cycle cost reduction.
    Way Forward: Change scope to modern electronic components used in receiving platforms – repeat steps 1-8. Include microelectronics marketplaces in new sources.

11. **Opportunity:** similar to 10 above but with hardware components.
    Way Forward: Change scope to the hardware components/systems used to support the ACAT1 weapons platforms – repeat steps 1-8. Include industrial manufacturing marketplaces in new sources.

12. **Opportunity:** similar to 10 & 11 but looking at indirect contributions.
    Way Forward: Change scope to the operational systems (like machine tools) used to support receiving platforms – repeat steps 1-8. Include industrial manufacturing marketplaces in new sources.

13. **Opportunity:** similar to 13 looking at indirect contributions.
    Way Forward: Change scope to the engineering and design systems used to support the platforms – repeat steps 1-8. Include industrial manufacturing marketplaces in new sources.

14. **Opportunity:** similar to 13 looking at indirect contributions.
    Way Forward: Change scope to the business systems (requirements, acquisition, contracting, finance) used to buy the platforms – repeat steps 1-8.

15. **Opportunity:** gradually deliver the same benefits for all weapons platforms.
    Way Forward: Change scope to alternative ACAT 2-4 platforms– repeat steps 1-12.

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. Executing the tasks under the Enabling Objective will move the acquisition community further towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 5: QUANTITATIVE RISK ASSESSMENT

Strengthening risk management in MDAPs by creating a quantitative risk assessment tool for managing MDAP risks at both PEO and portfolio levels

SUMMARY

This data product will enable A&S and the Services to apply commercial sector best practices in managing the business risks inherent in program & project management for MDAPs. This data product will enable a disciplined, orderly approach to estimating technology and schedule risks, as well as costs and cost-related risks. By enabling the integration of risk assessments into business case templates and the integration of risk register reviews into portfolio management processes, this data product will foster faster, more predictable, and more successful acquisition programs. This data product will improve uniformity in projecting costs and risks and permit rapid responses to changing economic conditions. It will also make it easy to create an aggregate view of the sources and potential impact of risks within an Acquisition portfolio and across the Department.

OPPORTUNITY SPACE

Commercial sector best practices instill discipline in program & project risk management by leveraging a combination of:

- light-weight automation and data feeds,
- training and reinforcement through performance measurement,
- integrating risk assessments into business case templates
- integrating risk register reviews into portfolio management processes.

Best practices in risk management emphasize the rule that risks are synonymous with uncertainty. Not just the uncertainty of things going wrong (downside risks), but also the uncertainty of desirable things happening (upside risks). A strong culture of risk management works to mitigate things going wrong, but also cultivate desirable outcomes.

STRATEGY FOR IMPROVEMENT

We recommend improving MDAP performance and predictability by delivering data products that enable a disciplined, orderly approach to estimating technology and schedule risks, as well as costs. These data products, and the outputs thereof, will be integrated into the MDAP business case template. They will also be incorporated into Integrated Application Portfolio Reviews.

The User Experience (UX) / User Interface (UI) will leverage the familiarity of spreadsheets for data capture, with automated feeds of standard costs, inflation rates, etc. from open source and commercial data products. This approach delivers uniformity in projecting costs and risks and permits rapid responses to changing economic conditions. It also makes it easy to create an aggregate view of the sources and potential impact of risks within an Acquisition portfolio and across the Department.

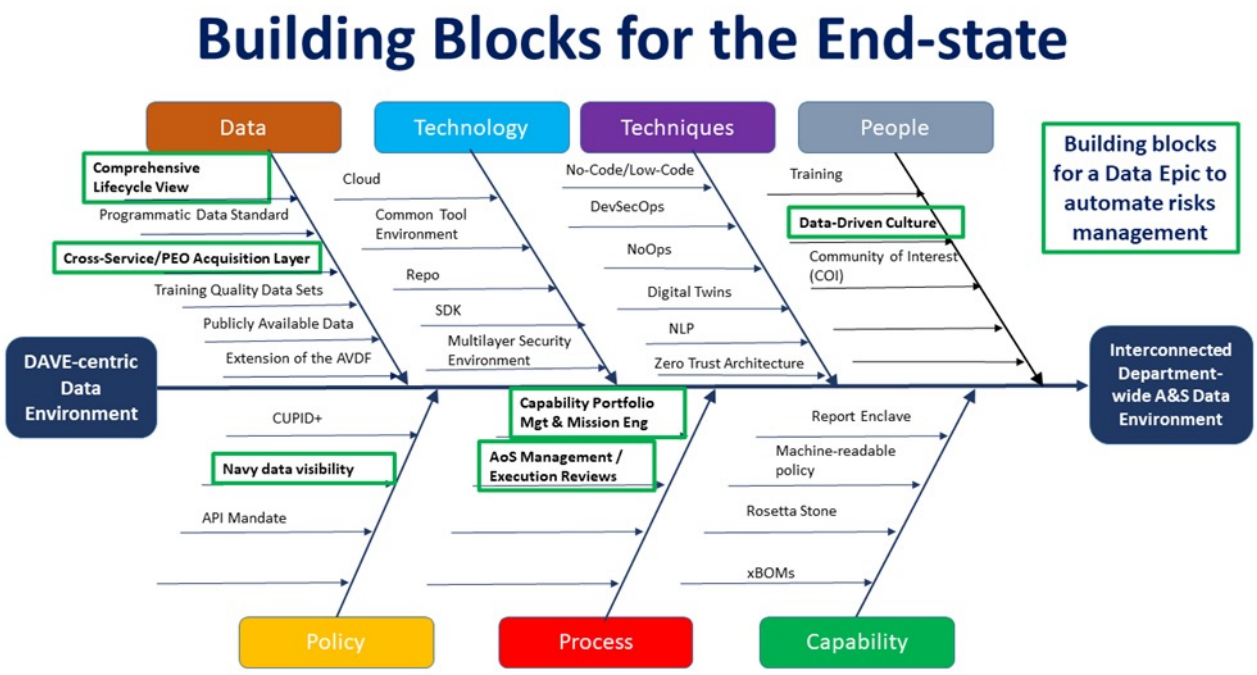

Figure 10. Strategy elements to develop quantitative risk management tool

DAY-IN-THE-LIFE:

Darsi Bando is a program manager organizing a new MDAP in the Middle Tier of Acquisition. The planned weapon platform will rely on ML algorithms, some of which are mature. But other algorithms are dual-use and undergoing rapid revolutionary advancements in both academic and consumer R&D projects. Moreover, the algorithms will be running on low-power microchips, a technology sector of the economy that has been unpredictable of late. The algorithms will be used to control dual-band quantum cascade lasers, another group of emerging technologies.

Darsi just completed the one-hour, self-paced training on the new risk management data product and is about to use it for the first time. Darsi logs into DAVE and clicks on the Risk Management icon. Darsi enters basic information about the MDAP. Then Darsi imports a draft digital engineering model for the weapons platforms. Using drop-down menus, Darsi rapidly confirms the desired performance parameters. Darsi also uses drop-down menus to select the technologies that will comprise the components of the new weapons platform. Darsi reviews and confirms the initial costs parameters, automatically pulled from open source and commercial data feeds. Darsi also reviews and confirms the automatically populated lists of assumptions and issues generated by the data product based on the nature of the weapons platform. Finally, Darsi reviews the initial cost estimates and fills a number of gaps. The results are saved automatically and routed to the PEO financial manager for a review of the costs.

DATA EPIC SCOPE:

The initial implementation will be scoped to support both PEOs and Acquisition Portfolio leaders.

Data Story 1 – As an Acquisition Project Manager or Program Manager, when I organize a new MDAP I want to create a baseline profile of the material assumptions, issues, and risks I take into account for planning purposes.

Action – Build a UI/UX that delivers the logic and functionality detailed immediately below. Incorporate DoD Acquisition guidance into the logic and the drop-down menus. Incorporate GAO guidance (see footnotes 1-7, but especially GAO-20-195G) into the logic and the drop-down menus. Integrate a Defense Finance and Accounting Service (DFAS) Applications Programming Interface (API) to provide standard BLIs, as well as National Item Identification Numbers (NIINs) and their average costs.

- an Assumption is defined as a planning parameter which might affect the successful outcome of the project/program. When an assumption is verified, it becomes an Issue to be managed. Further:

    1. Assumptions are statements accepted as true without proof. Given enough time, all assumptions will eventually be proven true or false.

    2. Assumptions and risks share two key characteristics: Uncertainty (probability) and consequence (Impact). Because of this, assumptions can benefit from qualitative analysis.

    3. Typically assumptions have a low probability of being untrue. Any assumption with a high probability of being untrue is a candidate for management as a risk.

    4. Typically assumptions have a high or medium impact to the project. Any assumption with a very high impact is a candidate for management as a risk. Any assumption with a low impact is a candidate for being ignored.

    5. Some risks with a medium or low risk ranking (according to the P-I matrix) are candidates for management as assumptions.

    6. Assumptions should be identified, qualitatively analyzed, quantified, and monitored, but they do not require in-depth quantitative analysis or response planning unless and until they are converted into Risks or Issues.

- a Risk is defined as a planning parameter that represents an unknown or a source of uncertainty. Further:

    1. Like an Assumption, a Risk has two components: uncertainty (probability) and consequence (impact).

    2. The uncertainty (probability) is expressed mathematically on a range from 0 – 1.

    3. The consequence might be undesirable (aka a 'downside risk' to be avoided) or it might be desirable (aka an "upside risk' to be cultivated). The consequence is typically expressed in terms of $ impact. In the case of a downside risk the impact is increased cost, typically the $ cost to recover from the undesirable event. In the event the downside impact includes schedule delays, which by extending the length of the project / program incur additional cost, which are included in the calculation of consequences.

4. There are many sources of risk:
    - political, organizational, or business issues
    - emerging technologies and techniques
    - budget and funding
    - start-up activities, staffing, and organizational issues
    - program management practices.

- an Issue is defined as a planning parameter, with a confirmed value, that acts as constraint or limitation the ability of the project / program to deliver a successful outcome. Typical Issues include the level of available: budgets, staffing, infrastructure, facilities, and information technology.

1 – As an Acquisition Project Manager or Program Manager, when I am organizing a new MDAP by capturing material Issues for planning purposes, I want to be able to specify:

- Issue Title
- Issue Description
- Issue Impact (Cost, Schedule, project / program performance, combat capability delivered)
- Issue Owner
- Issue Action/Decision Owner
- Issue Action/Decision Description
- Issue Action/Decision Start Date
- Issue Action/Decision Target Close Date

2 – As an Acquisition Project Manager or Program Manager, when I am organizing a new MDAP by capturing material Assumptions for planning purposes, I want to be able to specify:

- Assumption Source
- Assumption Category, including:
  - » Ally Decision or Action
  - » Client (MilDep or COCOM) Commitment & Environment
  - » Competitor / Enemy Decision or Action
  - » Delivery & Operations
  - » External Risks
    - ◊ Economics
    - ◊ Defense Supply Chain
    - ◊ Dual-use Supply Chain
  - » Financial
  - » Team Experience & Capability
  - » Legal & Contractual
  - » Project Complexity
  - » Project Governance
  - » Project Resourcing /HR
  - » Project Schedule
  - » Security
  - » Scope/ Engagement Definition
  - » Sub-contractors
  - » Technology
    - ◊ Hardware
    - ◊ IT - Hardware
    - ◊ IT – Software
    - ◊ IT - Data
  - » Transition & Transformation
- Assumption Owner
- Assumption Cost Category (BLIN)
- Assumption Status
- Assumption Title
- Assumption Description
- Impact Description
- Probability Assumption is Incorrect (0 – 1)
- Assumption Verification Action Plan
- Assumption Impact (Cost, Schedule, Program Performance, Combat Capability Delivered)
- Action Strategy
- Action Owner

3 – As an Acquisition Project Manager or Program Manager, when I am organizing a new MDAP by capturing material Risks for planning purposes, I want to be able to specify the following characteristics of each Risk:

- Threat (downside) or Opportunity (upside)
- Risk Source
- Risk Category, including:
  - » Ally Decision or Action
  - » Client (MilDep) Commitment & Environment
  - » Competitor / Enemy Decision or Action
  - » Delivery & Operations
  - » External Risks
    - ◊ Economics
    - ◊ Defense Supply Chain
    - ◊ Dual-use Supply Chain
  - » Financial
  - » Team Experience & Capability
  - » Legal & Contractual
  - » Project Complexity
  - » Project Governance
  - » Project Resourcing /HR
  - » Project Schedule
  - » Security
  - » Scope/ Engagement Definition
  - » Sub-contractors
  - » Technology
    - ◊ Hardware
    - ◊ IT - Hardware
    - ◊ IT – Software
    - ◊ IT - Data
  - » Transition & Transformation
- Risk Owner
- Risk Cost Category (BLIN)
- Risk Status
- Risk Title
- Risk Description (There is a risk that X will happen because of Y)
- Impact Description
- Risk Response Description
- Risk Response Strategy
- Risk Trigger Description
- Risk Trigger Date
- Risk Probability (0 – 1)
- Cost Impact ($s)
- Schedule Impact (Time and $s)
- Expected Monetary Value (Risk Probability * (Cost Impact and $ value of Schedule Impact))
- Estimated Cost of Risk Response
- Contingency Included in Budget

4 – As an Acquisition Project Manager or Program Manager, when I am organizing a new MDAP by capturing material Risks for planning purposes, I want to be able to prioritize Risks.

5 – As an Acquisition Project Manager or Program Manager, when I am planning my Risk response, I want to be able to determine and record a Response Strategy for each Risk.

- For downside Risks, the available strategies are: Transfer, Mitigate, Accept, Avoid.
- for upside Risks, the available strategies are: Ignore, Enhance, Pursue.

6 – As an Acquisition Project Manager or Program Manager, when I am planning my Risk response, I want to be able to determine and record a specific Response for each Risk. The response may include a mitigation/enhancement plan; a contingency plan; an estimate of the $ cost of the Risk response.

7 – As an Acquisition Project Manager or Program Manager, when I am planning my Risk response, I want to be able to determine and record a project / program risk reserve for known unknowns and unknown unknowns.

Data Story 2 – As an Acquisition Project Manager or Program Manager, when I am managing an MDAP I want to be able to monitor and control Risks. Typical actions include:

- record actual data in a Risk log
- review and update risk status
- monitor early warning indicators, with automatic feeds of commercial and open source data
- execute Risk responses and evaluate results for problems and windfalls
- communicate with stakeholders and project / program team members

Data Story 3 – As an Acquisition Project Manager or Program Manager, when I am closing out the project / program I want to:

- document lessons learned
- register proposed improvements to the Risk management process, tools and templates
- register proposed improvements to other project / program processes, templates and tools to create future leverage / reduce future exposure

Data Story 4 – As an Acquisition Portfolio Manager, when I conduct an IAPR I want a data product that enables me to answer the following questions:

- What are the key unknowns (Risks) that might impact the success of a given MDAP as measured by:
  - » combat capability delivered
  - » schedule
  - » cost
  - » program performance
- Is the pattern of Risks correct and complete for this MDAP?
- Are the Risks being closed at an appropriate pace given the schedule?
- Are there any obvious risks missing, especially upside risks? Are we taking adequate steps to cultivate upside risks?
- Looking across my portfolio of MDAPs, what are the common sources of:
  - » material technology risk?
    - ◊ do I have adequate mitigations / contingencies in place?
  - » material schedule risk?
    - ◊ do I have adequate mitigations / contingencies in place?
  - » material cost / budget risk?
    - ◊ do I have adequate mitigations / contingencies in place?
- Looking across my portfolio of MDAPs, what are the common Assumptions? Are these Assumptions still valid and accurate? Do I need to revise, or instruct PEOs to revise, any of the material Assumptions?
- Looking across my portfolio of MDAPs, what are the common Issues? Are these Issues still valid and accurate? Do I need to revise, or instruct PEOs to revise, any of the material Issues?

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. Executing the tasks under the Enabling Objective will move the acquisition community further towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.

DATA PRODUCT 6: CLIMATE CRISIS

Empowering DoD to enhance infrastructure and tackle the climate crisis

SUMMARY

This data product will support the analytical workload regarding the integration of climate considerations into DoD's policies, strategies, partner engagements, war-gaming, modeling and simulation in order to bolster mission resilience and reduce DoD's carbon footprint. The data product will support visualizing DoD land, built infrastructure, and installation operations, providing analytical insights to develop strategy and answer mission-critical questions. The data product will empower the DoD with an authoritative view of its energy profile, assets and inventory, and climate risk to make insight-driven decisions, tackle the climate crisis, and build resilient infrastructure.

OPPORTUNITY SPACE

Secretary of Defense Hon. Lloyd J. Austin has stated, "We face a growing climate crisis that is impacting our nation's missions, plans and capabilities and the DoD must take immediate action." In line with President Biden's direction, the DoD will elevate the climate as a national security priority, integrating climate considerations into DoD's policies, strategies and partner engagements. The DoD will incorporate climate-risk assessments into war-gaming, modeling and simulation, and bolster mission resilience and deploy solutions that optimize capability, and reduce our carbon footprint.

The Energy, Installations, and Environment (EIE) Analytics portfolio supports visualizing DoD land, built infrastructure, and installation operations, providing analytical insights to develop strategy and answer mission-critical questions. The DoD has more than 550,000 installation facilities worldwide, which include a massive installation and operational energy footprint. As a result, the DoD is one of the world's largest producers of greenhouse gases. The portfolio's suite of products and applications empowers the DoD with an authoritative view of its energy profile, assets and inventory, and climate risk to make insight-driven decisions, tackle the climate crisis, and build resilient infrastructure.

STRATEGY FOR IMPROVEMENT

Develop a tranche of new EIE data products. The initial focus is on assessing the Department's existing EIE data and analytics maturity. Later the focus will expand to improving EIE data and analytics capabilities and delivering new insights to drive improved Department performance in Readiness and Environmental Protection Integration. See Figure 11 below.

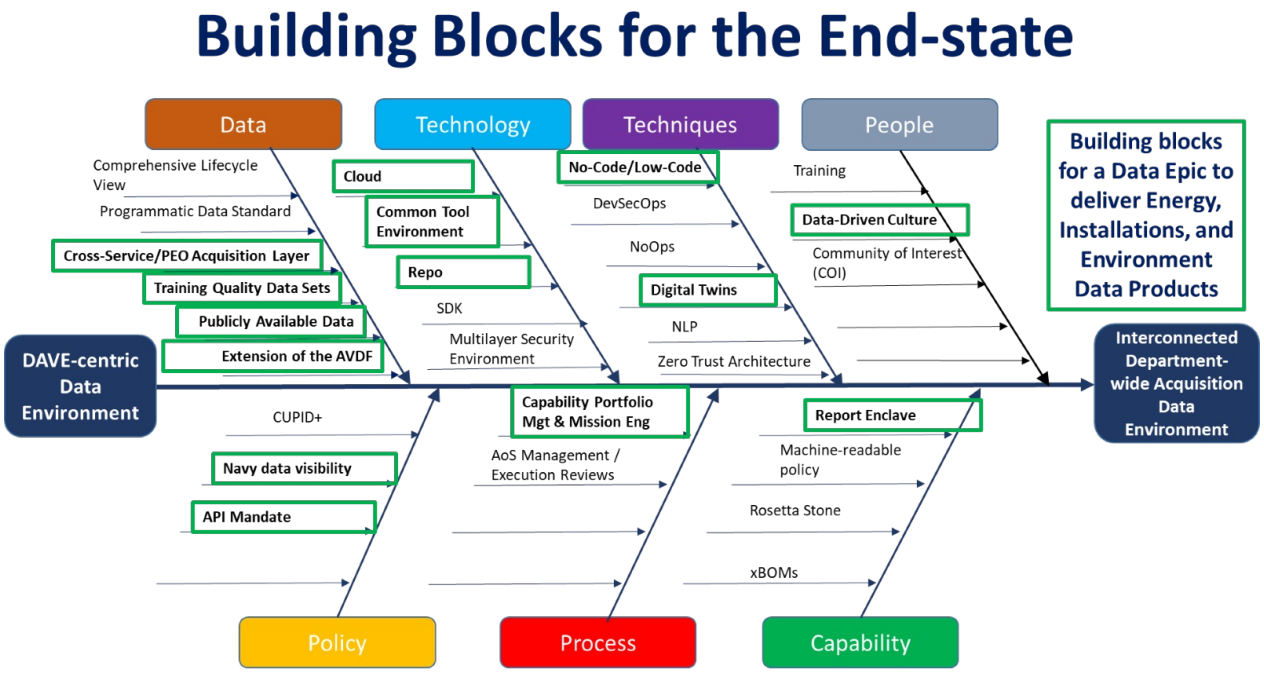

Figure 11. Building blocks for a EIE use case

DAY-IN-THE-LIFE:

Hayden Mills is an Acquisition professional responsible for supporting the Deputy Assistant Secretary of Defense (DASD) for EIE. A Deputy's Management Action Group (DMAG) briefing is scheduled for the end of the month. Hayden is getting organized to prepare the DASD for the DMAG. Hayden logs into the A&S data repo in one browser screen, and Advana Executive Analytics on another browser screen. The A&S Data Products for EIE are the place Hayden does deep analytics, performs scenario planning, and tweaks performance targets. The impact of those activities are automatically fed into the Advana Executive Analytics. Hayden starts with the A&S repo, Hayden "thinks out loud" about what the Data Products show:

> "Hmm. We're making good progress with root cause analysis on the main drivers of the Department's Greenhouse Gas Emissions (GHG). And I can see the impact of rising temperatures on HVAC power consumption. Which will take our GHG emissions in the wrong direction. That's not what we want.
>
> What if we increase the number of High Climate Exposure Resilience Projects in Quartiles 3&4 by 20%? OK, that partially compensates for the higher GHG emissions. Let's try 27%. Ok that neutralizes the expected increase in GHG emissions. But to make the improvements the Secretary wants, we're going to need to increase the number of High Climate Exposure Resilience Projects in Quartiles 3&4 by 35%. Do we have the funding for that? Ooops, we do not. Looks like we need to get ready for a budget discussion."

DATA EPIC SCOPE:

As an Acquisition professional, when I am performing oversight and management of the Department's Readiness and Environmental Protection Integration (REPI) performance, I want access to A&S Data Products that enable me to answer the following questions:

1. What # of policy and technical guidance updates has been published? What is our mathematical confidence in the precision and recall of this number?

2. What % of mission assurance priority installations are complete? What is our mathematical confidence in the precision and recall of this percentage?

3. What % of major installations are loaded into DoD's climate assessment tool? What is our mathematical confidence in the precision and recall of this percentage?

4. What % of major installations have integrated installation resilience plans? What is our mathematical confidence in the precision and recall of this percentage?

5. What are the main drivers of the Department's GHG? What is our mathematical confidence in the validity & accuracy of this analysis?

6. What are the existing or desired performance metrics to manage the Department's carbon footprint? What is our mathematical confidence in the validity & accuracy of this analysis?

7. What is the Department's operational energy profile? What is our mathematical confidence in the validity & accuracy of this analysis?

8. What is the profile of the Department's installations with respect to energy and resilience? What is our mathematical confidence in the validity & accuracy of this analysis?

9. What are the Department's investments to bring down its carbon footprint? How effective are they? What are the opportunities to scale? What is our mathematical confidence in the validity & accuracy of this analysis?

10. Are more resilience projects continuing to be funded by the REPI Program? What is our mathematical confidence in the validity & accuracy of this answer?

11. Is more funding being allocated to REPI resilience projects? What is our mathematical confidence in the validity & accuracy of this answer?

12. Are installations with greatest exposure to climate change impacts being prioritized? What is our mathematical confidence in the validity & accuracy of this analysis?

13. Is the Department meeting the FY targets for number, and rate of increase, of Installation Resilience Projects? What is our mathematical confidence in the validity & accuracy of the answers to these questions?

14. Is the Department meeting the FY targets for allocation, and rate of increase, of Installation Project Funding? What is our mathematical confidence in the validity & accuracy of this analysis?

15. Is the Department meeting the FY targets for the number of High Climate Exposure Resilience Projects in Quartiles 3&4 receiving funding? What is our mathematical confidence in the validity & accuracy of this analysis?

16. How can the Department leverage its buying power to reduce emissions? What is our mathematical confidence in the validity & accuracy of this analysis?

CONCLUSION

Completion of this implementation plan enables A&S to better align efforts being executed across the enterprise aimed at accomplishing the mission objectives outlined in the DoD Data Strategy. The implementation plan template defines tasks and outcomes that support strategic data goals. Executing the tasks under the Enabling Objective will move the acquisition community further towards a data driven culture where there is a renewed focus on performance, accountability, and collaboration to improve acquisition outcomes.