Test and Evaluation and Digital Data Strategy

- Leveraging a Tester's Experience to Build a Successful Enterprise Data Strategy
- Assessing the Resilience of Critical Infrastructure Systems and Supply Chains Against Cascading Disruptions
- Optimizing Antenna Placement Using Modeling and Simulation
For the test professional, the truth has always been in the data. Particularly when it comes to the Evaluation steps of the job description, accurately characterizing systems under test without the appropriate data can be impossible. Testers have always cared about access to data, collection methodologies, data storage, and security. Indeed, testers have been in the business of data management since before Big Data was cool.

But far too often, these concerns were most loudly voiced at the level of the test engineer or lowly data analyst. The lessons learned about the velocity of data have not yet been completely or effectively leveraged by leaders in acquisitions at the enterprise level. The insights available from these lessons and their applicability to Test and Evaluation (T&E) business practices are just as important to enterprise decision making and risk assessment as the performance characterization of the systems evaluated by test professionals during T&E campaigns.

“Every organization, be it a publicly traded company, a private for-profit company, a non-profit organization, a governmental agency, or a quasi-governmental entity all have strategic goals.”¹ An enterprise data strategy should support these strategic goals, as well as the improvement of and progress for data processes in that organization—regardless of how mature the underlying data framework. Defense Acquisitions is no different.

Often the biggest improvements in data process lie at the intersection between operations (or execution) and information science. In the development life cycle of a novel, complex system, this intersection is where testers like to operate. T&E is full of challenges and lessons in intersection. This article will identify some of these challenges experienced by test professionals and posit their contributions to an integrated, enterprise-level acquisitions data strategy.

The first step in understanding where these lessons can fit into a data strategy is to briefly define the nuances of a Data Strategy at the enterprise level.

What the Heck is a Data Strategy?

There is no singular, authoritative definition of a Data Strategy. There are several informative books and
numerous high-quality articles written about different types of Data Strategies and their composition. However, there are some generally agreed-upon elements that are the basic composition of a good data strategy. This section discusses these key elements in order to provide context for capitalizing on the lessons learned in T&E. This section also highlights why data are a strategic asset and why a strategy is needed for generating the highest return on investment.

**Data Vision**

Just like an organization develops its unique mission and vision statements, based on its values, goals, and principles, a good data strategy is birthed from a data vision. A “data vision statement indicates how data will help an organization realize its mission and strategic goals.” In the absence of a vision of how data, data processes, and data assets contribute to organizational objectives, a data strategy will provide little strategic guidance for the practitioner or manager.

This vision should project into the future what an organization’s picture for success in their data processes looks like in five years. A clearly stated data vision that reflects an organization’s values and its focus on data as an asset will guide the next handful of elements that comprise the data strategy.

**Data Acquisition**

“One of the most basic constructs for using and sharing data within a company is establishing a means to identify and represent the content.” This process includes data identification, prioritization, capture, and curation to include the management of data quality. It should include assessing data quality levels, utilizing the most valid source of data, and data standardization in terms of content and format. For testers, this should seem reasonable and part of the day-to-day job responsibilities for anyone who conducts test events.

**Data Governance**

Data Governance is both ubiquitous and often poorly understood. Governance is not only the set of principles and processes by which an organization performs audits and control, checks and balances, and assigns quality labels to data, but it also consists of information technology (IT) policies and mechanisms for effective data utilization. "Importantly, data governance is about people and business processes more than it is about data, and while IT professionals should participate in data governance, it should not be relegated to or led by an institution’s IT unit.” These processes should flexibly address data dictionaries and acceptable metadata that drive the rules and procedures for today’s data processes and beyond.

Governance should pay attention to how data operators and analysts interact with IT personnel, and it should facilitate fruitful conversations about security policy among stakeholders from operations, security, IT, and management. These conversations are often political or contentious, as more lenient policies typically generate a perception of more work for IT professionals.

Another crucial element of governance is establishing an office for a champion for data processes, as well as encouraging (formally or informally) data stewardship at the tactical level. The process of establishing this office for a data champion includes outlining the appropriate roles and responsibilities of everyone in an organization who even thinks about data.

While this article is not about the importance of an independent Chief Data Officer outside of the IT functional domain, it is worth mentioning that establishing a Chief Data champion and data management functional community in an organization is only the first step toward the successful implementation of a value-generating data strategy.

**Data Storage and Access**

The SAS Institute succinctly defines data storage as follows: “Persist data in a structure and location that supports easy, shared access and processing.” This might seem like an oversimplification; however, many organizations manage their data storage in an inconsistent manner that precludes easy sharing and makes collaborative processing and analytics a challenge. A good data strategy will settle upon a flexible, consistent storage solution to enable cross-functional analysis and data handling as well as sharing the right data at the right time with both internal and external stakeholders.

Access refers to the implementation of the security policy as well as the connectivity requirements for authorized users to get their hands on the appropriate data so they can exploit data for business purposes. Access includes written plans for both current and future infrastructure modernization in a rapidly evolving information domain.

**Data Use and Analytics Tools**

If your organization has articulated a vision and future in which data as an asset drive the strategy and the process (a rather large if), and you have satisfied general administrative requirements for data acquisition, governance, storage, and access, your strategy can next address how data will be used and the tools your
organization will support or develop to best expedite data to decisions in a value-generating way.

The idea of a single, authoritative version of truth underpins an organization’s data use, and this single version of truth is a critical asset for any organization that enables systems and digital modeling. This single source of truth “is a logical, often virtual and cloud-based repository that contains one authoritative copy of all crucial data.” A single source of truth can be one of a company’s most valuable assets. This source of truth ensures that assessments use accurate, valid data. And those data and core analytics activities are based on defensible processes for data standardization, information extraction, the use of analysis software, and the development of advanced analytics tools to facilitate business processes.

A data strategy should establish best practices and in general elaborate on a baseline for the types of tools that can be used to store, manipulate, and analyze data. However, this guidance should not be so prescriptive as to become restrictive over time. A data strategy is a long-term strategy document, and the rapid pace of software development can quickly make overly specific guidance on software tools obsolete.\(^\text{12}\)

**Data Literacy, Information Extraction, and Reporting**

“Data is useless without the skills to analyze it.”\(^\text{13}\) Even the best data strategy will break down when confronted with a *data illiterate* user base. A well-constructed data strategy will include a plan to ensure users understand the organization’s data, how it’s stored, how it can be accessed, the tools available, and how information or analytics can be extracted from it. Relevant data competencies should be addressed in this part of the document. These competencies could include mathematical reasoning, big picture analysis, and data experimentation, but they also might include code sharing/re-use, compilation skills, and tool containerization.\(^\text{14,15}\) These competencies should include references to the types of skills, models, and use cases that a user might find helpful while accessing the organization’s data.

This section of a data strategy should also include best practices in how to extract information from the data, whether it is through a proprietary software suite or using a one-size-fits-most commercial-ready tool. Once the data are “in,” the user now needs guidance on how to get the information “out” while accurately reporting results. This should include guidance on how to appropriately cite data sources while ensuring consistencies in documentation.\(^\text{16}\)

**You Want to Put Together a Data Strategy...Now What?**

Once an organization has recognized both the value of data as a strategic asset and the importance of a data vision and strategy, it is time to scope and shape the organization’s goals with respect to data processes to best achieve its overall organizational objectives, mission, and vision. No one is going to argue that these processes can help refine the state of the art, particularly as business practices move toward a more connected digital environment.

As program offices of major programs and the acquisition arms of each of the military services attempt to address the key elements of a data strategy that were mentioned above, there is a long list of lessons and challenges already faced in the T&E community. Ignoring these lessons in concocting and implementing a data vision and strategy is like leaving money on the table, metaphorically speaking. Let’s recount some of the most applicable lessons that can be used to inform weapon system-level or program executive officer-level acquisitions data strategies.

**Information Overload**

Whether you agree with it or not, Moore’s Law has had an effect on the computing power of the hardware inside of major defense acquisition programs. This author posits that development has shifted from a hardware-centric, platform-focused mindset toward a digital environment where software is ubiquitous. This, if done well, can unlock a myriad of new capabilities: software-defined apertures, rapidly re-programmable operating systems, and a general abstraction of capabilities development away from the underlying physical phenomena that software is controlling. This list of pros is abruptly shortened because the pivot toward more capable computation and malleable software capabilities is not the point. When the rubber actually meets the road, this new era for development has two major drawbacks: increased complexity and an explosion in the data generated during routine operations.

This new world generates so many potential interactions among cyber-physical components as well as copious amounts of data that these revolutionary changes have moved beyond the limits of a human’s ability to “intellectually manage” the underlying complexity.\(^\text{17}\) Aside from challenges in simply digesting the underlying software–hardware process coupling, these new systems generate more data during standard operations than legacy systems produced when instrumented with exquisite data acquisition systems.
Combine this explosion in the quantity of data with a tester’s desire for special instrumentation recording every bit on every bus, and you have a recipe for disaster. Test points that previously generated a few hundred megabytes of data today yield several gigabytes of data. Test campaigns for legacy systems that generated a few gigabytes of data now, in 21st century systems, produce terabytes of data.

For the test community, this has presented a number of challenges. Instrumentation systems need to be more robustly designed to better capture useful data during execution. Data acquisitions, transfer, and storage requirements have exploded in the last decade, driving increased costs for on-premise storage capabilities as well as stretching the Test Enterprise’s IT subject matter expertise to its limits. Finally, the quantity and disparate types of data have challenged T&E’s analytics professionals.

“The rapid growth of data creates new opportunities for smart analytics,” but only if the test community’s IT professionals, data scientists, test engineers, and leadership can work together.18 As the military internet of things eventually proliferates, increased reliance on machine-to-machine processes and data derived from IoT-enabled sensors will open up entirely new paradigms for data processes and the information available within brontobytes of leverageable data.19 The challenges in this data explosion are obvious. With constant growth in the quantity and types of data collected during test campaigns, identifying logical solutions to this problem has not yet proven fruitful.

Testers can always triage data sources and information streams to alleviate some of the challenge of this information overload. However, the broader test community has not yet settled on standards that can help govern the types of data and data products that should be preserved. In the absence of guidance, the most conservative approach is to avoid accidentally deleting potentially valuable information and instead coping with this issue by committing to store vast amounts of data.

**Analysis Capabilities and Human Capital**

Practitioners are often analytically baselined to the tools with which they grew up. That is to say that analysts who first learned to implement a particular analysis technique using a proprietary Mathworks Matlab or an open-source tool like Python or R naturally gravitate back toward those analysis tools. Subscribing to commercial software creates one burden for an analytics organization, while using open-source tools generate another.

Proprietary licenses are costly, often requiring specific toolboxes and add-on capabilities. These powerful tools are not cheap, and because of IT and security policy, rarely are these licenses centrally managed and available to all data professionals in a data enterprise.

Organizations that employ open-source tools routinely struggle in managing the software configuration as well as just-in-time updates to required analytics packages, if they can make it past the system owner’s approval to actually install the tools they need on a machine that has access to the data in question.

Few organizations have the luxury of juggling both types of capabilities. IT and security policies have at times created data stovepipes that have both prevented insightful analytics and stifled innovation.20 The use of air-gapped private networks makes managing analytics tools a challenge. Data science professionals are confronted with the unenviable task of turning a bolt with a hammer; oftentimes, they are not resourced with the best software for the job at hand. However, in today’s hyper results-focused environment, testers have nobly pressed ahead with the tools they have available.

Aside from the costs to purchase commercial licenses or to manage routine updates for open-source analytics packages, there is a bureaucratic cost for these tools that often goes overlooked. Leadership should remain apprised at the consternation associated with any IT updates, including the associated contracting process required by government policy to actually solicit proposals for software licenses, the sometimes-unproductive interaction between IT tacticians and a system’s authorizing official, and the occasional inflexibility of local IT professionals to support on-demand configuration changes required in the high-pressure T&E environment.

It would be remiss in any discussion of data strategy or analytics to not mention the clear-cut single most important asset in any organization: its people. Human capital is critical in the pursuit of data excellence. No strategy can be well implemented without the right personnel.

“Unfortunately, data people are relatively rare.”21 Data people understand data processes, the life cycle of useful data, and the velocity of data. They are integral in documenting and implementing a data strategy. These rare pieces of human capital truly recognize data as a strategic, value-generating asset, and they understand the opportunity costs associated with the mismanagement of data.

“Data science itself is an eclectic mix of a myriad of skills and engineering disciplines; from data engineering to big data analytics to advanced applied mathematics, no individual data scientist possesses all of the relevant skills to tackle each and every analytics
Parts of the defense T&E world still depend on the technologies of the late-1900s. Imagine attending a wedding or a birthday party for a close friend or family member next weekend. Someone would take photos, whether professional or amateur, and these photos can easily be shared through the internet. OneDrive, Google Drive, DropBox, pick your sharing resource of choice, and you can have a read-only link distributed via email to all of your family and friends in minutes. For the DoD tester, things aren’t that simple. To juxtapose with current standard practices, the T&E professional would do one of two things: print the photos of interest (after going through a burdensome approval process to even connect a printer to their enclave!) or put them on an external hard drive (after going through a separate burdensome approval process to whitelist this particular drive for this particular task). After this transfer, most testers are forced to physically mail this media across the country. In the most well-funded military machine in the industrialized world, access is not instantaneous, nor is it convenient.

Even these laborious processes are not foolproof. Anyone who is seasoned in this IT shuffle will explain to you that they are highly personality dependent. A system owner is empowered, in the name of security obviously, to deny these types of requests. In some cases, government testers and support personnel are denied access to government data that is part of their project or program. This is ironically counter to Schofield’s first law of computing, “Never put data into a program unless you can see exactly how to get it out.”

Data science professionals in industry have likely never experienced these security- and process-related challenges. How could they expect to continuously generate value and return on data with these types of constraints? The velocity of data and how it flows throughout its life cycle is just as important as its content, and apparently that has not yet been realized in T&E analysis. DoD testers are also steeped in the catastrophic effects of vendor lock and proprietary, incompatible solutions. As the DoD attempts to digitally transform itself into a 21st-century fighting force, solving these access problems is a major hurdle.

Applying these Lessons to an Enterprise Data Strategy

Over the years, T&E professionals have learned a slew of lessons that are applicable to an enterprise-level data strategy for acquisitions systems or for a portfolio of systems. At the program executive officer level, there is a virtual treasure trove of useful data that can be better utilized if an organization effectively implements a

---

Data Accessibility and Collaboration

Testers are often entrusted with highly sensitive information; whether it is proprietary contractor data, important weapon system performance data, or state secrets, Defense T&E professionals take security very seriously. This forces IT professionals to implement strict security policies, controls, and access management measures.

Some of these measures include air-gapped enclaves that are connected only to a power supply. No internet, no networks, just a couple of computers tied together with an ethernet cable. Additionally, once testers have accomplished the “data acquisition” step of the data life cycle, there is often little hope of ever transferring data off of that computer or enclave, all in the name of security. These data black holes exist throughout the Defense test enterprise, and handle one type of data for one single purpose, but sadly they do not facilitate cross-platform collaboration in a world that is pivoting to an all-domain operating environment. While from an IT and security standpoint these solutions are secure, from a data use and reuse perspective such solutions are stove-piped and inefficient. These types of solutions do not appropriately balance security risk, innovation, and mission needs, severely hampering the insights that can be garnered from the terabytes of data that exist on Department of Defense (DoD) hardware all over the world.
useful data strategy. This data includes design documentation, specification data, requirements documents, results from analyses of alternatives, threat data and associated red team models, results from test and evaluation events, deployment and sustainment performance data, object-oriented system and components models, model-based digital artifacts, and other authoritative data sources that attempt to support a digital approach to acquisitions portfolio management. The lessons from earlier in this paper have helped develop recommendations to best support this Program Executive Officer-level data strategy.

Deriving T&E data requirements has traditionally been a linear, decompositional process. System performance specifications drive requirements for test objectives; these objectives determine which test techniques the test team will employ and the test data that will be captured during an overall test program. When systems were hardware defined and underpinned by a level of complexity that could be intellectually managed by a small engineering team, this was an acceptable methodology and cornerstone of the T&E data process.

However, as software has become ubiquitous and users’ demands on new systems have manifested in the form of feature explosion, the new normal for a test campaign includes an overall T&E strategy that cannot possibly be exhaustive. As the functionality of different layers and components of a complex system has exploded, the result is a “combinatorial explosion” in the number of test events, test techniques, and the overall amount of test data that need to be collected to accurately characterize a system under test.

One way to manage this increased complexity and data overload is thoughtful applications of systems theory. In systems theory, complex systems are modeled as a hierarchy of levels of organization, each more complex than the one below, where a level is characterized by having emergent or irreducible properties. This approach is contrary to traditional scientific reductionism: “systems theory states that complex systems must be considered holistically.” This holistic approach can guide both the test team and program management functional leadership in identifying targeted complex interactions among systems within a portfolio, critical sub-system level data, and performance data to support a model-based approach throughout a system’s life cycle. This prioritization will drive key contributions to a data strategy by helping determine the data life cycle, acquisitions, and use of the highest priority data. Learning from the T&E community’s experience in this arena will be critical in informing high-level strategy in the knowledge age.

Testers are also familiar with advanced architectures and their implications. They already know that “architecture plays a major role during systems analysis, design and development.” As systems become increasingly distributed in an all-domain, digital operating environment, architecture decisions become even more influential in the data produced by weapon systems. These insights surrounding systems architectures can inform both an architecture-driven digital approach and the data strategy for a particular series of architectures.

Logically these system design architectures may utilize similar subsystem functionality and component interactions across a portfolio “making them suitable for representation by a reusable specification architecture.” Testers are already familiar with this construct, and while this lesson can inform data acquisition, governance, and finding the right tools for analysis within a data strategy. This architecture-based modelling approach extends beyond what is simply in a single system’s interface control document or at the level of abstraction of a system’s software architecture description document. This approach includes both high-level systems’ architecture insights and the backbone of the data process architecture that the data strategy seeks to address: the data acquisitions, storage, analysis, reporting, and extraction components of the strategy.

Testers are accustomed to a fast-paced, operations-driven work environment. Always willing to get the job done, analysts and data scientists in T&E utilize the tools available to them, even if they aren’t the best tool for the job. At times this can be cumbersome and result in suboptimal data processes. In an organization that truly values data and its inherent value, this should not be the case.

In generating a data strategy, the data champions at the leadership levels should continuously seek to resource their data science teams and practitioners with the best tools for the job. In a data strategy this must be addressed in how data is acquired, how information is extracted and reported, and specifically in the tools available and use of an organization’s data assets. This lesson is critical in the knowledge age, and if acquisitions portfolios sincerely desired to shift toward a digital, model-based approach, it cannot be ignored.

If you have your finger on the pulse of the desired shift toward a connected, all-domain fighting force, you might be familiar with the idea of passing the right data from any sensor, to any shooter, at the right time. While this is not an article about some of the barriers to that level of connectivity capability, it is difficult to imagine this future state when testers are still mailing hard drives and hardcopy, printed reports to stakeholders.
because of connectivity, IT, and security policy challenges.

Data people often turn the popular phrase: bring the analysis to the data. Unfortunately, in today’s system, that is at times impossible because of high-level policy or organizational inertia. Testers have recognized that if the data is available to a competent data science team with the right tools for the task, data science as a service is not a farfetched concept in T&E. Bringing the right resources to the analytics edge is a key enabler of this data process.

Including distinct verbiage to promote this enabler within a data strategy will better serve the defense acquisitions community. Whether it is increased resourcing for connectivity into a cloud-based or government-owned data lake or networked access to the appropriate collaboration and analysis tools, data strategy will be more successful if its authors include this specific lesson from the test community.

Lessons Learned

Testers are uniquely qualified to influence organizational and enterprise-level policy when it comes to data processes. In one way or another, they are familiar with the key elements of a data strategy, and the test enterprise has both struggled and succeeded in implementing key elements of such a strategy.

Lessons and challenges from how testers have dealt with feature explosion, complexity, and the exponential increase in the quantity and types of data can be better managed using systems theory and an architecture-based approach. The principles of these ideas can help identify and prioritize the data, its use, and the types of tools required within a data strategy.

Portfolio-level leadership should also pay attention to the appropriate analytics capabilities while enabling a competent data science team to execute precision analytics at the edge. This lesson direct from the T&E’s book will influence how data is acquired, how information is extracted and reported, and specifically in the tools available and use of an organization’s data assets; these critical insights will drive a data strategy. Lastly, testers have struggled with connectivity, access, and collaboration for years. Stove-piped enclaves and air-gapped enclaves preclude insightful analytics, and a data strategy should account for the imperative to work through these challenges.

ALEXANDER “HEFOE” HILLMAN is a Ph.D. candidate in aeronautical and astronautical engineering at the Massachusetts Institute of Technology. An active-duty Air Force officer, Major Hillman is an Air Education and Training Developing Airmen We Need PhD Fellow and a Military Fellow, MIT Lincoln Laboratory. A United States Air Force Test Pilot School graduate, HEFOE has served in various technical and leadership positions in both developmental and operational test inside of the Air Force’s test enterprise. Alex holds master’s degrees in operations research, systems engineering, flight test engineering, and military operational art and science. Prior to his doctoral studies, Major Hillman was most recently assigned as the Chief Data Officer of the Air Force’s only Cyberspace Test Group and as Deputy Director of the Air Force Test Center’s Emerald Flag.

Endnotes

4 Hosch, “Key Elements of a Data Strategy.”
5 Adelman, Data Strategy.
7 Statistical Analysis System Institute, The Five Essential Components of a Data Strategy.
8 Hosch, “Key Elements of a Data Strategy.”
9 Statistical Analysis System Institute, The Five Essential Components of a Data Strategy.
10 Hosch, “Key Elements of a Data Strategy.”
14 Ibid.
16 Ibid.


8th Cybersecurity Workshop

NEW DOMAINS IN CYBERSECURITY T&E

WHO:
Cybersecurity Testers (DT, OT, Red Teams), Researchers, Test Event Directors, Test Planners, DoD and Industry Leaders

WHEN:
October 18–20, 2022

WHERE:
Embassy Suites by Hilton Destin Miramar Beach, Destin, FL

At this workshop, we will offer hands-on tool demonstrations and interactive sessions, and expand your knowledge of leading technologies and best practices.

You will see the government, industry, and academia showcase capabilities, research, tools, and services across a wide range of cybersecurity applications.

KEYNOTE SPEAKERS:
Maj. Gen Evan C. Dertien, USAF, Commander, Air Force Test Center, Edwards AFB

James Wells, (SES), Director, Office of T&E, Department of Homeland Security

George Rumford, (SES), Director (acting) and Principal Deputy, Test Resource Management (TRMC)

Sarah Standard, Cybersecurity/Inter-operability Technical Director, OUSD R&E, Developmental Test, Evaluation, and Assessments

TRACKS WILL INCLUDE:
- Cybersecurity & Test Technology
- Securing the SW Supply Chain and Infrastructure
- Test Design, Regulations, and Best Practices
- Best Practices for Cyber Evaluation Tools
- Assessing and Mitigating Risk

Sponsor sign-up, contact info@itea.org
Registration open www.itea.org
Forging the 21st Century T&E Tools in an Era of Great Power Competition

September 27-30 | 2022
The Westin Virginia Beach Town Center | Virginia Beach, VA

Why this T&E Symposium is a “Must Attend” Event

The current Global Power Competition (GPC) climate can pose a change to the acquisition and fielding process of most systems. It is important to understand the 21st Century T&E Tools to improve testing to meet the GPC demand. Join us as we ask the tough questions and gain the answer to: How will the military budget, DoD, DHS, and Commercial Industry adjust in a time of risk and change?

Join us as we present the latest insights on:
- Rapid and Mid-Tier Acquisition
- Incorporating T&E into Acquisition Contracts
- Big Data and Cyber Playbooks
- AI and Machine Learning: How to Test the Trustworthiness & Human System Integration
- Cyber War
- Digital Twin in T&E
- Accelerating T&E for Hypersonic
- Critical and Emerging Autonomous Systems Technologies
- From Low Orbit to Mars, Challenge of Deep Space Exploration

Who Should Attend?
- Government Program Managers, Testers & Engineers
- Principal & Chief Scientists
- Sr. Marketing Managers, Defense Systems
- Sr. Analysts, Defense Weapon & Testing
- Sales Directors, Space & Defense Systems
- University Professors & Students
- University R&D Directors, Faculty
- DoD Branch Chiefs
- Commanding Officers & CCCOMs
- Presidents / Vice Presidents Defense & T&E Programs

Limited availability for sponsors and exhibits. Sign up today.

For more information, go to itea.org, select Symposium under Events or contact:
Erwin Sabile, CTEP
ITEA 2022 Symposium Chairperson
ITEA Hampton Roads President
sabile_erwin@bah.com

www.itea.org