



White Paper

ENABLING THE VISION FOR THE DOD DATA STRATEGY



Enabling the Vision for the DoD Data Strategy

The DoD recognizes that data is a strategic asset that must be operationalized in order to provide a lethal and effective Joint Force. Improving data management will enable operators and military decision-makers to harness data to capitalize on strategic and tactical opportunities that are currently unavailable. The DoD Data Strategy Sections 2.2.4 (Data Collection) and section 4.3 (Making Data Understandable) are 20 percent of the objectives that will produce more than 80 percent of the intended objective results and mission achievement.

In February, 2020 during an AFCEA Signal Magazine interview, Mr. Tom Sasala, Chief Data Officer for the U.S. Navy, acknowledged that cleaning and raising the quality of data is the bulk of the directorate's effort right now. "My primary drive over the next 12 months is really what I'm calling investing in the data management infrastructure, which is the ability to find the data sets, bring them into an environment, clean and curate them and make them available for analysis. Once we have that kind of highly integrated, highly curated data, it really does provide exceptional value to the warfighter and the analysts and data scientists." Tamr agrees with Mr. Sasala's approach.

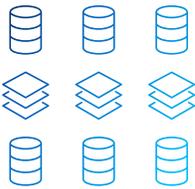
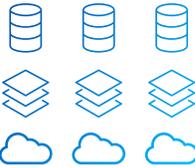
The DoD's Chief Information Officer, Dana Deasy, stated, "Data is the ammunition in the Digital Modernization Strategy and is increasingly central to warfighter advantage on and off the battlefield. The National Defense Strategy directed us to be more lethal, efficient, and interoperable with partners. This strategy is our first step to making that ammo persistently available to the men and women of the DoD regardless of echelon or geographic location." Think "federated MDM" and domain data mastering capabilities.

Data sets, bring them into an environment, clean and curate them and make them available for analysis

DoD data is not yet widely available nor accessible by mission commanders, warfighters, and decision-makers in a real-time, usable, secure, and in a linked manner. “The public sector’s reflexive habit of classifying data as the default further complicates this challenge. This combination effectively locks away the data, preventing anyone from using it,” said Michael Conlin, then DoD Chief Data Officer in February 2020.

Like DevOps which has moved from waterfall to agile methods saving time and money and improving quality, similarly DataOps (Data Operations) should involve more agile data management adhering to specific architectural principles:

Technology - Architectural Principles

Source	Process, Technology, Organization	Consumers
<p>Internal Tabular Data</p> 	<ul style="list-style-type: none"> • Scale Out/Distributed <ul style="list-style-type: none"> • Cloud First • Collaborative (Humans at the Core) <ul style="list-style-type: none"> • Highly Automated - automate whenever possible • Bi-Directional (Feedback) • Open/Best of Breed (not one platform/vendor) <ul style="list-style-type: none"> • Service Oriented (clear endpoints for data) • Loosely Coupled (Restful Interfaces Table(s) In/Out) 	 <p>Citizens</p>
<p>External Tabular Data</p> 	<ul style="list-style-type: none"> • Continuous (assume data will change) <ul style="list-style-type: none"> • Both aggregated AND federated storage • Both batch AND Streaming • Lineage/Provenance is essential 	 <p>Analysts</p>  <p>Data Scientist</p>  <p>Developers</p>

An essential part of realizing the vision of the DoD strategy is through streamlined DataOps. All pertinent data needs to be managed as a strategic asset and established by building out the data pipelines required in order to serve that data to all the components and users that need it, in a timely manner. The quality of this data will determine the quality of the decisions made on the front lines for the operational community and the warfighter. Tamr is able to streamline this paramount step by improving the quality of the data by collating and correlating all of the component data sources to drive more accurate analytics. Starting with high quality data is an essential beginning towards a journey of digital transformation and being able to tackle more expedited decisions downstream as components become more data-driven.

The quality of data determines the quality of decisions made on the front lines

For DoD missions, Tamr delivers decision makers information that is up to date, accurate and unified across a myriad of sources for more granularity. Tamr connects and integrates data that is siloed across services and across the DoD enterprise. This data unification platform surfaces more comprehensive insights faster. Tamr is used for automation of entity resolution, record deduplication and classification of lists of information into a desired enterprise-wide ontology. This can make your AI algorithms, as well as any analytics your organization may need, much more effective.

Tamr collates and correlates all data sources to drive more accurate analytics

Three Examples of DoD Data Quality Management



The Joint Artificial Intelligence Center (JAIC)

The JAIC is seeking to improve readiness and cost avoidance of preventive aviation maintenance and repair. Tamr Government Solutions is working with the JAIC to produce a common ontology for predictive preventative maintenance. As of October 2020 the prototype project is maturing to the satisfaction of the JAIC and is expected to move to production in 2021 with expanded scope.



US Air Force SEEK EAGLE Office, Eglin AFB

Tamr was selected to enable the digital transformation of the Air Force SEEK EAGLE Office. Tamr leveraged its machine learning, human-guided platform to automate a process of understanding and providing searchable context to 40 years of test flight data. Inbound requests for review of new aircraft configuration are assisted with automated discovery of relevant antecedent flight tests. Furthermore, Tamr automated the process for analogy-based assessment of specific engineering disciplines pursuant to aircraft configuration recommendations. These capabilities are driving efficiency and quality improvements to core functions within the SEEK EAGLE's mission.



US Navy

Tamr is automating the fusing and correlating of collections from multiple, advanced sensors observing common areas of interest. This includes spatial-temporal feature-rich data sets. Reliance on manual processes in dealing with the mountains of data created by these sensors has put the Navy at a disadvantage against our adversaries. The data volume and complexity make it impossible for current human-centered, manual processes to be effective, meaning threats will be missed, putting lives at risk in a combat environment. Tamr's machine-learning based, human-guided solution is able to recognize and correlate relationships from those disparate data sets which reference the same entity, detect new anomalous data outside of the normal range of existing feeds, and provide new user/analyst visualizations for improving warfighter understanding of the battlespace environment.

Tamr's Data Mastering Platform is built on an open architecture and open APIs.

From a data rights assertion standpoint, all of the data is owned by the DoD. Tamr may be deployed on DoD infrastructure and is also the only data mastering solution with native capability on all three major cloud service providers: AWS, Azure and GCP.



Conclusion

Over time, data sprawl creates a large gap in institutional knowledge, hindering the ability to find, access, and understand the useful data needed for critical business and mission questions.

Given the substantial challenges of the DoD data landscape, our recommendation is to start an investigation of focus areas of Senior Leader Decision Support and Business Analytics starting with Category Management that would yield improved combat readiness and cost avoidance. Senior leaders' dashboards are historically populated with dirty, outdated and inaccurate data. Populating these dashboards with clean, mastered data should be the highest priority for the DoD.



To learn more, visit tamr.com/public-sector

This White Paper is company confidential and proprietary.

