

**XML in the Wild Blue Yonder:
A Survey of USAF XML Implementations & Implications
Authors: Cheryl Connors, Amit Maitra, Mary Ann Malloy
Keywords: E-Government, XML, Knowledge Management**

Abstract

A recent survey of XML implementations found that many United States Air Force (USAF) communities are incorporating XML as a foundational step in their migration to a net-centric vision. Although the survey was limited to publicly available resources – and thus only a partial view of total USAF efforts – thoughtful analysis of the survey results nonetheless reveals both strengths and weaknesses in the approaches inspected. In this paper we summarize the survey results and what it implies for how the USAF is progressing towards net-centricity. We note potential positive impacts XML technologies could have on USAF business practices, and some potential shortfalls. We discuss some strategies that may help favorably impact the nature and progress of USAF XML adoption. This discussion supports our position that XML technologies can be an essential part of an overall interoperability solution and the migration to net-centricity, but like other technologies that have gone before them they are still no “silver bullet.” Technologies, however expedient, must be embedded in an overarching framework that supports sustainment, divestiture, standardization, harmonization, fallback, redirection, synchronization and re-implementation as needed and when appropriate to meet the visionary goals set by an enterprise. We briefly look at the Federal Enterprise Architecture (FEA) Data and Information Reference Model (DRM) as one enterprise-scale solution with an XML implementation that may be a blueprint for USAF success. We note there are outstanding issues for community consideration, and we hope this paper stimulates further discussion of them.

Introduction

The United States Air Force (USAF) is implementing a “net-centric vision.” Migrating to this vision will evolve the physical and logical aspects of present point-to-point, pre-engineered information exchange solutions into a web-enabled information infrastructure that leverages the latest internet technologies. Resultant efforts include the articulation of a data migration strategy that promises quantum leap increases in information exchange and knowledge management capabilities. Why is such revolutionary change viewed as desirable or necessary?

It is no secret that the US Department of Defense (DoD) and Federal government at every hierarchical level have been subjecting themselves to “woodshedding” on the question of the use they make of precious data resources. It is essential to smooth and effective operations to ensure that decision-makers have decision-quality information where and when they need it. The inability to share or exchange data efficiently costs time and money and is contrary to the prevailing citizen-centered focus of contemporary business practices. Data inefficiency results in ever-increasing costs to manage and integrate more

and more data. It can contribute to corruption of the data, negatively impact interoperability and increase the burden of finding and accessing the right data.

We sought evidence for how likely these shortfalls are to be mitigated by current USAF efforts. We conducted a survey of XML implementations in March 2005 and found that many USAF communities are incorporating XML as a foundational step in their data migration processes. A thoughtful analysis of the survey results reveals both strengths and weaknesses in USAF approaches and, by implication, the processes of other enterprises that are turning increasingly to the promise of XML-based solutions to these problems.

We posited several desirable impacts XML technologies could have on USAF information exchange and business practices, such as:

- Adoption of XML could increase the ability to exchange information among the scores of heterogeneous systems currently active in the USAF.
- Implementation of authoritative ontological and taxonomic definitions could support automatic transformations between diverse lexical representations.
- Integration of common operational processes and data resources, enabled by XML technologies, could result in more efficient and effective operations.

At the same time, we admitted these potential gains co-exist with encumbering shortfalls and challenges, such as:

- Lack of a common technical roadmap with definitive metrics that is applicable for each USAF organization.
- Limited knowledge management process in place to disseminate news and lessons learned about existing, often disparate efforts.
- Lack of proven approaches for how to make new and dynamic sources of information available in a trusted and interoperable environment, while at the same time leveraging presumably cheaper commercial off-the-shelf (COTS) solutions and standards-based technologies.

To understand the relationships between the positive impacts we anticipated and the challenges we could foresee, we reviewed the current adoption of XML in the USAF as reflected in a recent survey. Because implementing XML-based interoperability solutions is a foundational step in the Air Force's migration to achieving the net-centric vision, this survey can provide some perspective on how these efforts are progressing.

XML Implementation Survey in the USAF

The DoD recognizes the potential that exists in Internet related technologies for improving information exchanges among the heterogeneous systems that are used across the spectrum of operations. A DoD memorandum published on the Net-Centric Data Strategy [1] describes the need for the DoD for “ensuring data are visible, available and usable” and the “‘Tagging’ of all data”. The memorandum also mentions the employment of XML metadata registry efforts – acknowledging the importance of this technology – that are already in place. This memorandum followed an earlier recommendation by the Air Force Scientific Advisory Board to implement the Joint Battlespace Infosphere (JBI) [2] which describes the importance of the JBI as an infrastructure to integrate, aggregate, and distribute information, and to use XML as a key enabling technology.

Since then, many USAF agencies and systems have developed an XML representation for some or all of their data. Our interest in determining the level of XML implementation by the USAF led to the development of the survey.

Survey Resources

We derived all of the information for the survey from publicly available resources, such as on-line articles, web sites and government doctrine that has been published without distribution restrictions. It should be noted that there are potential biasing factors and limitations in the scope of the survey data. For example, many of the resources describing USAF XML implementations have been published by product vendors. These articles generally focus on announcing the successful product or service sale to the USAF rather than the extent or success of any resulting XML implementations.

The DoD Metadata Clearing House and Gallery [3] is a repository for XML information objects developed by all US Armed Services. A review of the Gallery provides clear indications that there are many ongoing DoD efforts to migrate to XML, but many have not been made public. In addition, it is not possible to trace the Gallery’s individual subsections to the Services that have helped populate them; so it was not possible to include the Gallery as a resource for this survey. In short, the survey results are limited and they cannot be used to determine exact percentages or degrees of completion per transformation efforts underway in the USAF.

The survey results can be interpreted as indicators that some effort has been made to implement XML by numerous USAF agencies and programs. Similarly, the data provides a glimpse of the level of USAF interest in acquiring and implementing XML solutions, and what this implies for the move to net-centricity.

Summary of Survey Results

We have summarized the XML implementation survey results in Table 1. The categories across the top – “War-Fighting”, “Business Support”, “Intelligence”, and “Prototype Development” – provide further elaboration regarding the type of XML implementations being conducted by the USAF programs listed in the “Programs” column.

Table 1. Summary of USAF XML Implementation Survey Results

<i>No.</i>	<i>Program</i>	<i>War-Fighting</i>	<i>Business Support</i>	<i>Intelligence</i>	<i>Prototype Development</i>	<i>Comment</i>
1	XML-MTF (USMTF)	X		X	X	The DoD Messaging standard “MIL-STD-6040” has a formally approved XML representation
2	XML-MTF (NATO)	X		X		NATO Messaging standard “ADatP-3” has a formally approved XML representation
3	DoD Metadata Gallery	X	X	X	X	
4	NATO XML Registry (not yet implemented)	X	X	X	X	Impacts interoperability between USAF and NATO
5	XML Binary				X	(Proposed W3C initiative). Actively supported by the USAF as a research effort.
6	Ground Moving Target Indicator (GMTI)	X		X		NATO Messaging standard “Stanag 4607” has an XML representation that is in the approval process
7	Coalition War Fighter Interoperability Demonstration				X	Participants included Australia, Canada, Denmark, France, Germany, Hungary, Italy, Republic of Korea, the Netherlands, New Zealand, Norway, Poland, Spain, Turkey, United Kingdom, and the United States of America.
8	Cursor On Target (CoT)				X	XML format used for exchanging targeting information
9	Joint Expeditionary Force Experiment (JEFX)				X	Focuses on experiments that enhance the fusion of data and improved dissemination across the battlespace. Article: http://esc.hanscom.af.mil/Esc-PA/NEWS/2004/April%202004/ESC%2004-14.HTM
10	Theater Battle Management Command System (TBMCS)	X		X		Employs XML to exchange information via web services
11	TBone	X		X		XML implementation is currently under development
12	DCGS	X				(Future: currently under development)
13	AF GFM		x			(Future: currently under development)

14	AF Weather JMBL and JMIS	X				
15	AFKN		X			
16	TMSS		X			
17	ECATS		X			
18	BOXML					Under development, operational area unknown
19	ICML			X		Under development

Questions Raised by the Survey

As noted earlier, the USAF is infusing XML technologies in its interoperability solutions to migrate to the net-centric vision. While XML technologies are potent enablers in achieving improved information exchange, they alone cannot solve such a complex and multi-faceted problem. While focusing only on those parts of the interoperability problem that programs have sought to resolve using XML technologies, a review of the survey results leads to some logical implications, raising questions and issues that need to be resolved. We have summarized the main points of this discussion in Table 2.

Table 2. Summary of Issues Areas Raised by the Survey

<i>Issue</i>	<i>Explanation</i>	<i>Implications</i>
<i>Scope</i>	The net-centric precept states that “ALL data are to be made available across the enterprise”; not just the pre-engineered exchanges that represent the status quo.	Believing that middleware will magically automate exchanges without a lot of hard mediation, standardization and/or harmonization is naïve, even when the exchanges are limited and pre-engineered. As the scope broadens, and exchanges become ad hoc, the complexity becomes huge. Questions about how to handle satisfiability and loss-y exchanges need more thought.
<i>Funding</i>	Traditionally funds are tied to functionality and follow point-to-point lines. This distribution does not put the money where it is needed to enable the new model.	The lack of adequate funding and/or failure to break down old interface/funding lines will impede successful migration of systems and information to the net-centric model. Due to draw down, many programs lack the resources to sustain, let alone move to a new paradigm.
<i>Guidance & Measures</i>	A plethora of documents related to net-centricity have been issued. How do implementers determine which of the many are relevant to their context? How do they assess compliance and/or measure successes?	There is no overarching roadmap to frame and guide implementation efforts to successful outcomes. Standardized means for measuring progresses and/or compliance are lacking. Objective metrics or criteria must be readily accessible to assess progress towards and compliance with net-centricity.
<i>Orchestration</i>	The number of repositories and registries for	With no orchestrating plan in place, ensuring

	collecting XML artifacts is proliferating; but there is no cross-cutting plan in place for coordinating them.	interoperability among many resources won't improve or may even degrade interoperability. Should this ultimately turn out to be the case, then what has the migration gained us?
Culture	How do we evolve our exchange models – or more importantly, our thinking – from “need to know” to “need to share?” Is this really a binary choice, or are there hybrid possibilities (e.g., “need to share SOME”)?	How to establish proper access authority when arbitrary, unanticipated data aggregations are possible is a complex and as yet unresolved problem. Program managers lack the money and resources to support arbitrary exchanges when they can't see an operational reason to do so. We are still grounded (and funded!) in a “need to know” mindset.

Now, we discuss each issue area in turn and provide additional details.

Scope: What do you mean, I have to share ALL my data?

A net-centric principle is that ALL data are made available across the enterprise. Making properly annotated data publicly available does add value when appropriate middleware tools are there to assist those in the enterprise to leverage it. These tools must automate the processes for finding the correct data and automatically mediate differences in the data's physical representation; for example, automating the transformation of latitude measured in degrees to another meaningful coordinate system.

Data experts will scoff at the ease this simplistic middleware solution suggests. They know from experience that the complexities of automating the exchange of data between heterogeneous formats are huge even when those involved in the exchange are known *a priori*. There are rarely simple answers in the point-to-point model; the net-centric interchange model introduces other problems we have yet to definitively resolve; for example:

- What happens when System A requires a specific data item to complete an information exchange with System B, but cannot discover it?
- What happens when there are different levels of granularity in requirements for reporting?
- What happens when System A's data field allows a range of “red,” “green” and “blue” but System B only allows “red” and “green”?

Questions of satisfiability and information loss need greater attention than they have received to date, to develop mitigating alternative approaches that still support sharing.

Funding: Where are the net-centric dollars and who will “pay the piper?”

Traditional funding avenues are aligned with specific functionality. That is, interfaces between systems are generally worked as point-to-point solutions and are funded according to those lines. The net-centric guidance to which the USAF (and all DoD

entities for that matter) must comply generally changes the interface model to one of publish-and-subscribe. But there has not yet been a change in the way programs and capabilities are funded in order to make this vision real and to break down the old interface/funding walls.

The lack of adequate funding will impede successful migration of systems and information to the net-centric model. The *Net-centric Enterprise Solutions for Interoperability (NESI)* effort, a cross-service effort between USAF (ESC) and Navy (PEO C4I & Space) sums up the issue this way:

Moving to a net-centric environment is a high priority of DoD leadership ... However, there are few or no additional dollars available for net-centricity. [4]

Guidance & Measures: Can someone hand me a roadmap and a yardstick?

A plethora of documents related to net-centricity – among them, mandates and user guidance – have been issued. Figure 1 shows one graphic that has been created to attempt to summarize all of them. [5] But implementation guidance is lacking, or when issued lacks the level of detail needed. Due to the sheer volume of information they represent, implementers have difficulty forming a comprehensive view of how these documents interrelate or how to build solutions in compliance with them. Neither have they been provided an implementation roadmap from leaders higher in the organization with an eye across the entire enterprise.

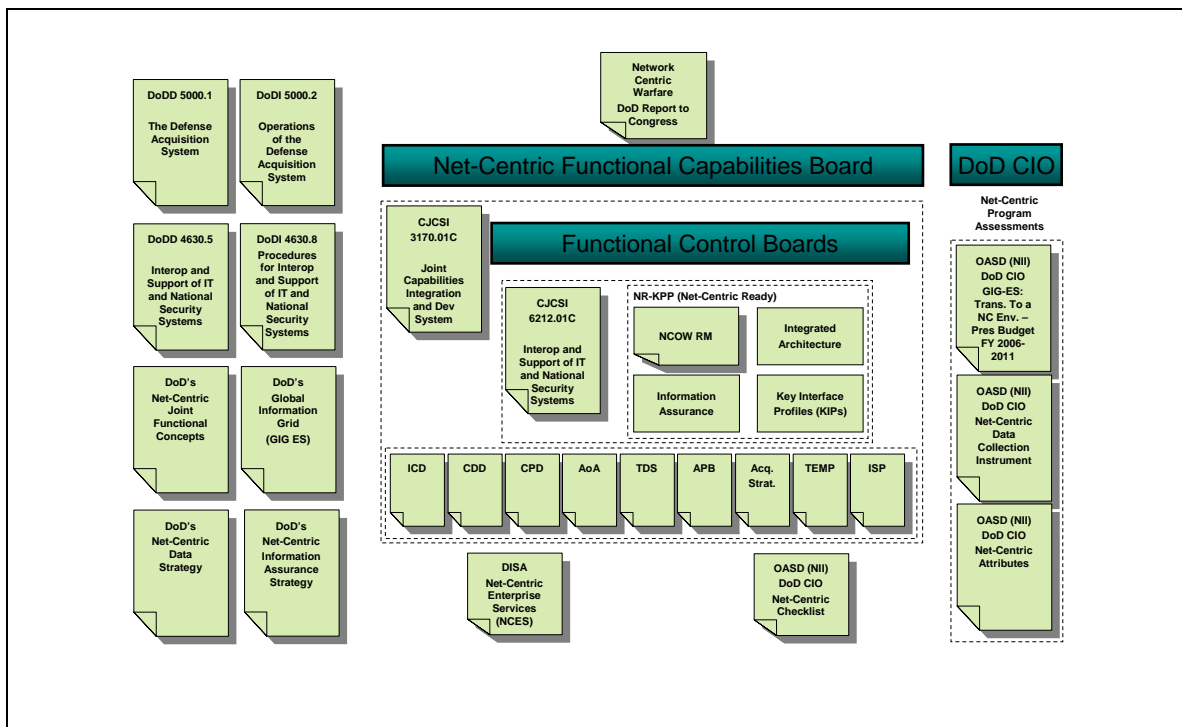


Figure 1. Navigating the Net-Centric Documents

The USAF has made it a priority to migrate all exchanged data into an XML representation and to have the resultant products registered in the DoD Metadata Gallery. Unfortunately, it is sometimes difficult for newcomers to understand how to access the appropriate XML representations and supporting data for their domain. A core purpose of the Gallery is to support reuse. Reuse of published XML representations are expected to aid in achieving interoperability and reduce implementation costs. Currently, the Gallery lacks information services that would support developers attempting to find and reuse the Gallery's XML representations. For example, in the context of the Metadata Gallery itself, implementers find it difficult to determine which namespaces they should review when seeking to reuse XML definitions in support of a particular functional area. A classic example is the word "tank" – how do you determine in which namespace to look to find an XML representation for "tank" when talking Army tanks vice fuel tanks, when the word "Army" or "fuel" may not appear as part of the XML tag name "tank"? . It should be noted that the Gallery is being revised and hopefully in the future will provide a better set of tools for these purposes. What is really needed is additional funding and resources to evolve the Gallery's current information services to meet the desired functionality.

Further, despite diligent efforts to locate them, we have been unable to find "laundry lists" of the criteria that should be used to determine if a particular implementation has achieved or failed its net-centric goals. The Net-Centric Checklist [6] is often pointed to; but it was actually designed to help programs understand the net-centric attributes they need to implement to move into the net-centric environment rather than rigorously stating requirements in a testable form. Summing the issues:

- There is no roadmap in place to guide implementers along a practicable path
- There are no appropriate supporting tools and processes to ease the development burden
- There are no checklists for assessing which guidance is relevant
- There is no standardized means for assessing the degree to which resulting implementations comply with the net-centric checklist (i.e., no binary-like metrics that can be used to determine pass or fail.)

The Net-centric Checklist is not reduceable to a set of binary-like metrics that can be used to determine pass or fail. Objective metrics or criteria must be readily accessible and interpretable as measures for implementers. This allows them to demonstrate and document the maturity, reliability, and functionality of their products and the degree to which they are in compliance with both expectations and prescriptive requirements.

Orchestration: How can we ensure the proliferation of Registries and Repositories will not impede future interoperability?

Across the Federal government many XML registries and repositories have been or are in the process of being built. The DoD Metadata Gallery mentioned earlier is only one of many XML registries and/or repositories. The developers of the "xml.gov" web site put together a list of XML registries and repositories across the US Federal government [7]. We have summarized this list in Table 3.

Table 3. XML Registries and Repositories Related to the US Government

<i>Organization</i>	<i>Purpose</i>	<i>Registry URI</i>
DoD XML Registry (now called the "DoD Metadata and Clearing House" and the "DoD XML Gallery")	Constitutes guidance in the generation and use of XML among DoD communities of interest and is the authoritative source for registered XML data and metadata components.	http://diides.ncr.disa.mil/xmlreg/user/index.cfm
Navy Logistics/Tech-Data XML/SGML Repository	Promotes the sharing of DoN DTDs, Schemas, and Styles (e.g., FOSIs, style sheets, XSL) with the aim of reducing DoN investment by avoiding redundant development. The DTDs in this Repository are official Navy DTDs and should be used when appropriate. The DTDs are also provided in an HTML format to enable navigation through a DTD using hyperlinking. It includes the Navy Baseline Tag Set, a data dictionary of commonly used element types. The Repository's purpose is to promote the use of these element types in tagging similar constructs (e.g., different paragraph levels, procedural steps, etc.) in Navy document instances.	http://navycals.dt.navy.mil/xml-sgm-rep/index.html
DLA/DLIS Integrated Metadata Repository System (IMRS)	A metadata management tool used to support a corporate data management function and is intended to provide metadata management services. Thus, the IMRS will support the engineering and configuration management of data environments incorporating e-business transactions, complex databases, federated data environments, and data warehouses/data marts. The metadata contained in the IMRS used to support application development, data integration, and the system administration functions needed to achieve data element semantic consistency across a corporate data environment, and to implement integrated or shared data environments.	http://www.dlis.dla.mil/imrs/
EPA's Environmental Data Registry (EDR)	A comprehensive, authoritative reference for information about the definition, source, and uses of environmental data. It supports the creation and implementation of data standards that are designed to promote the efficient sharing of environmental information among EPA, states, tribes, and other information trading partners. It also catalogs data elements in application systems. The EDR does not contain environmental data; it provides descriptive information to make the data more meaningful.	http://www.epa.gov/edr/
Justice Standards Registry	A repository of standards and specifications that help practitioners increase the nation's safety. This dynamic Web site was developed as part of the US Department of Justice interoperability effort to facilitate information sharing. The Web site captures existing standards and alerts users of new or emerging standards. It also provides comment areas for users to offer support for the listed standards. The latest version of the Global Justice XML Data Dictionary, Version 3, has been released for comment. Please visit http://it.ojp.gov/jxdm for additional information on this project.	http://www.it.ojp.gov/jsr/public/index.jsp
IRS Tax XML Developments	Lists various XML schemas for different tax forms.	http://www.taxadmin.org/fta/edi/xmldev.html#IRSXML
eGov/Federal Enterprise Architecture (FEA) Schemas & Documents	Contains two XML related information sets: 1) Revised Exhibit 300 Schema, Version 2.95 (for FY06) This document describes and defines the type of content including the entities, attributes, elements and notation used for the submission of Exhibit 300s for the FY06 budget cycle. 2) FEA Federal Reference Models (BRM,SRM,TRM) Version 1.2: XML Document This document contains the content of the Federal Enterprise Architecture relating to the Business Reference Model(BRM) v2.0, Service Component Reference Model(SRM) v1.0, and the Technical Reference Model(TRM) v1.1. IT includes the various layers of the Federal Reference Models and their detailed descriptions.	http://www.whitehouse.gov/omb/e-gov/e-5-documents.html

Education Community Registry	A central access point for XML Core Components, XML Schemas, and supporting documentation for the Education Standards Community. Created and reviewed through a collaborative effort between the Department of Education's Office of Federal Student Aid (FSA), the Postsecondary Electronics Standards Council (PESC), and the education standards community.	http://fsaxmlregistry.ed.gov/XMLRegistry/pages/welcome.jsp
LANL XML Repository	A repository of records in XML format. These records contain all essential bibliographic, keyword (semantic), and citation (structure) information about published documents that we require for ARP research and development. The XML marked up format allows much higher ease of access and retrieval of documents and was designed with the future needs of other Library Without Walls (LWL) efforts in mind.	http://informatics.indiana.edu/rocha/lww/XMLRep.html
NIST STEP Repository - Demo	NIST is developing a web-based repository to serve as the core of a modular environment for developers of STEP, a family of product data exchange standards. Modules are represented in the repository as XML, enabling them to be treated both as documentation and as software components. A prototype implementation of the module repository uses XML and other emerging technologies to dynamically render a module's content in a web browser in response to user input. No server-side processing is required.	Paper: http://www.mel.nist.gov/msidlibrary/doc/xmlrepo.pdf Other info: http://www.itl.nist.gov/div897/ctg/regrep/
UN XML Repository	"The EEMA EDI/EC Work Group has proposed to CEFAC the establishment of a Global Repository for the translation of XML tags in UN/EDIFACT and human language on the Internet. The EEMA EDI Working Group is prepared to assist in the set up and operation of such a repository, which could be crucial in the advancement of the use of EDI over the Internet."	http://www.edi-tie.nl/edifact/xml-edi.htm

A central issue with repositories in general is the need to harmonize the artifacts each contains among the different users all posting their metadata there. This problem escalates when multiple repositories are involved in the operational scenario. The implementation of multiple repositories as potential enterprise resources with no orchestrating plan in place seems to indicate that ensuring interoperability among XML components across the Federal and DoD branches of government will continue to be an issue. Should this ultimately turn out to be the case, then what has the migration to XML gained us?

We have been unable to find any definitive explanation or systematic plan to support interoperability between these different resource development efforts, although some material indicates that the "xml.gov" group is going to be responsible for the "harmonization of data elements across agencies." [8] Although the results, degree of harmonization, processes and participants are not clearly described, this claim is an acknowledgment that harmonization is important. But can anyone really build and effectively leverage a federated metadata repository without a plan in place to address these critical path aspects of its care and feeding?

Admittedly, no interoperability can or should be expected between certain component repositories; but progressing towards a paradigm that lacks a single source or methodology (other than brute force) for accessing all of the enterprise's metadata artifacts seems to be at odds with the principles of net-centricity. For example, if information on terrorism is uncovered on the battlefield in Iraq, how can we support its timely delivery to and interpretation by an FBI field agent when there isn't a shared

understanding of how the information is represented? And when there isn't even a shared Repository between these communities?

Without an attempt to harmonize the XML representations, even when limiting the exchange efforts to closely associated domains, all that will happen is that the DoD will create yet more stovepipes. Adding to our worries on this point, the UN and NATO are also creating their own XML registries and repositories! How will we ensure continued interoperability with our international partners in the context of the net-centric vision, when it's not even clear we will be able to do it in our own backyard?

From an even broader perspective, each of the individual services and agencies appears to have one or more major programs or initiatives they are following to achieve net-centric goals and objectives, over and above the repository issues we have just pointed out. While it is clear there is an intent to leverage and forge partnerships to coordinate these ongoing, parallel efforts, it is unclear the extent to which this is occurring.

Culture: Can the enterprise culture efficiently evolve from "need to know" to "need to share"?

The move to net-centricity also requires a revolutionary change in the culture climate from "need to know" to "need to share." In other words, in this context, to the net-centric principle "make ALL data available across the enterprise" we really have to add the caveat "...to ALL *authorized* users." How to establish proper access authority in a model where arbitrary, unanticipated data inquiries and aggregations are possible is a complex and as yet unresolved problem.

In addition, there is no incentive for program managers to ensure their data are interoperable with anyone else's, unless there are expedient reasons for them to do so. At present, programs ensure there is a common mechanism or format in place for partners with whom they already exchange information. But with budgets strained or drawn down, there is no incentive to attempt to reach beyond their current immediate partners. The degree of interoperability expected between disparate implementations that nonetheless may reference a common information domain has not been clarified.

To pursue the logical consequences of the sharing model when no operational reason is evident cannot be justified and takes resources away from other tasking this is likely to seem more important and immediately urgent. As a reasonable migratory step, initially it may make more sense to require programs to use XML to represent information that is already exchanged with other systems. The resulting artifacts can form a foundation for building web services in the future for more general consumption of those data across the enterprise. In the interim, communities need to work out the problems of determining access authority and appropriate sharing guidelines in the fluid net-centric environment.

Potential Solutions and Successes

The issue areas enumerated and discussed above point to the disjointed nature of USAF XML adoption and encumbrances to achieving the net-centric vision. At the core of this migration is how to accomplish the integration of data from disparate sources. Besides grappling with the questions of scope, funding, guidance and measures, orchestration of efforts and cultural change, stewards need to consider such mundane factors required for configuration management including data age and quality, the application(s) and partners that will use the merged products, and the availability of resources (particularly staff with the appropriate skills) to handle a data integration project of the nature and scope required.

Integration Options

An integrated solution minimally must satisfy the prerequisites of the Net-centric data strategy, as spelled out in the checklist: make data visible; make data accessible; make data understandable; make data trustable; make data interoperable; provide data management; and be responsive to user needs. Satisfying all these tenets in a net-centric operational environment will require registration and harmonization and/or standardization of data and business rules.

Those tasks in turn will require agreements between participating parties, and that is where the enduring challenge lies. A consistent, enterprise-wide method is needed for organizing data. XML can be a critical enabling technology, but in and of itself it does not solve such basic needs as getting players to agree on the specifics of the metadata and artifacts, where to register them and how to synchronize, deconflict and harmonize them. This is where a supportive framework, embedded in an overarching strategic plan, is needed.

Today, a few integration options are being explored, including:

- Extract, Transform, and Load (ETL) tools that extract data from various sources, transform them into a common format and then load them into a single database. Federal agencies such as the US Geological Survey, the Environment Protection Agency, the FBI, USAF, and the Defense Information Systems Agency, and a number of state and local governments use ETL.
- Enterprise Information Integration (EII) approaches, which use visualization to give the appearance of a consolidated data store, while leaving datasets in their native locations. EII is relatively new, with only \$250 million in US market share in 2004. Until now only small companies have been marketing EII tools. More recently, however, bigger players like IBM and BEA Systems have entered the market, thereby giving EII more presence and depth. The advantages of EII are significant: it requires little movement of data and little actual data transformation. Most important of all, EII offers costs savings.

- The Federal Enterprise Architecture (FEA) Data and Information Reference Model (DRM) is yet another alternative that promotes the common identification, use, and appropriate sharing of data/information across the Federal government.

This list is by no means exhaustive. Like the proliferation of repositories, the proliferation of potential solutions does not in and of itself recommend one over another or mean we aren't making any forward progress at all; the context or overarching framework within which one or several solutions are applied ultimately will determine that. For illustrative purposes we focus here on one solution option, the DRM approach, as a potential path to success that is being pursued by the USAF.

Federal Enterprise Architecture (FEA) Data and Information Reference Model (DRM)

As mentioned above, we can synopsize the key net-centric enablement challenge as a daunting data organization effort. This is an over-simplification of course; but the current lack of organization is an essential problem which, left unresolved, will become an insurmountable impediment to net-centricity. The FEA DRM, summarized in Figure 2, provides a structure through which the USAF (and others) can put their data houses in order.

In general, DRM was created to promote a government-wide understanding of its mission, vision, goals, objectives, and measures. The next few paragraphs present a high-level summary of the key elements of the DRM as of June 2005. For a more comprehensive discussion, see [9].

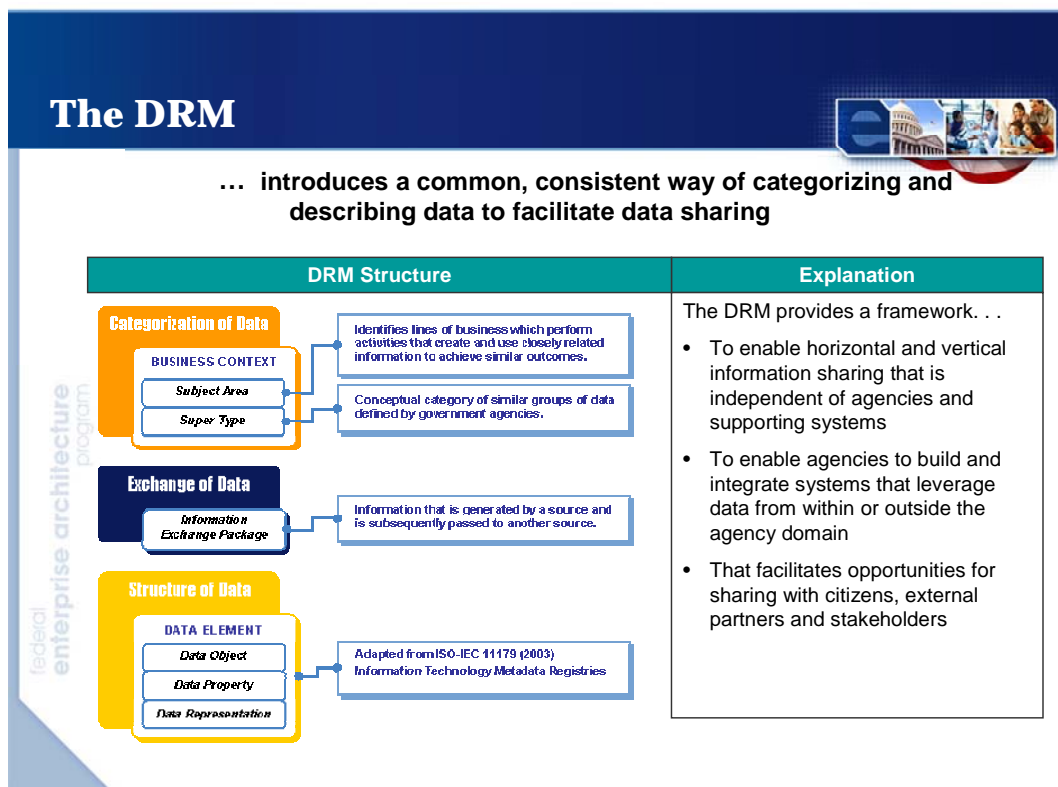


Figure 2. DRM Summary

The DRM Volumes

The DRM is being developed in a bottom-up collaborative manner within communities of interest (COI's) instead of the usual top-down collaborative approach. When completed, the DRM will comprise three volumes: a Management Strategy, an Overview, and an Implementation Guide.

The Management Strategy provides an introduction to the key elements of governance based on guidelines, principles, and issues from various data management perspectives. It also identifies the DRM business drivers which are issues voiced by managers or identified by working groups trying to build cross-domain information sharing solutions.

The Overview volume will help participating entities find common ground to begin a dialog on collaboration and sharing. It outlines the underlying drivers of DRM and outlines its benefits. It discusses how DRM fits in with the other reference models of the FEA, as illustrated in Figure 3. The FEA interrelated reference models are designed to facilitate analysis of investments and business functions across entities, while identifying opportunities for collaboration and cost efficiencies. While the other FEA reference models address performance, business, service and technical aspects, the DRM targets the business context, data classification / categorization and sharing.

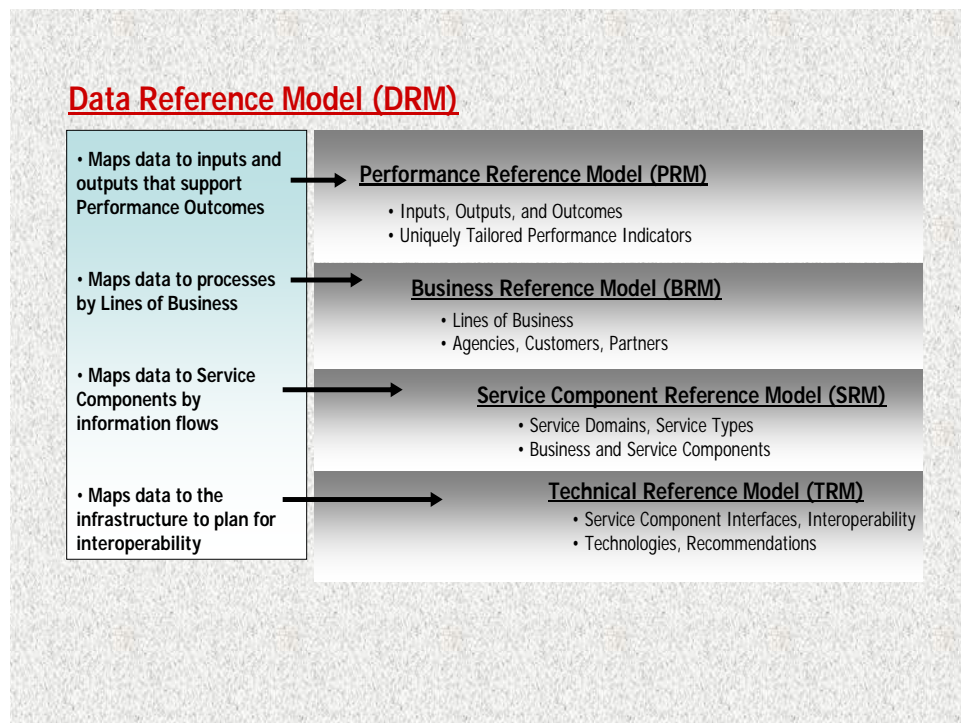


Figure 3: DRM “fits” with the other FEA Reference Models

Finally, the Implementation Guide is intended to facilitate sharing, accessing and exchanging information and data in a networked environment. The Guide promotes common identification, use and appropriate sharing of data/information across the enterprise through its standardization of data in the following three areas:

- ***data description***: This refers to the use of a standard approach to describing structured, semi-structured or unstructured data.
- ***data sharing***: This refers to a standard approach to describing the characteristics and requirements of data exchanges, including sources, and the definition of a standard message structure known as an Information Exchange Package.
- ***data context***: This refers to a standard approach to representing taxonomies for categorizing data. This allows the business context of the data to be well-understood.

DRM Implementation Approach

To facilitate implementation of the DRM, the Office of Management and Budgets (OMB) will provide a template for building an information architecture. This model – reusable across the enterprise – will require answers to three key questions:

- How can I find the right data?
- How can I share the data?
- How do I understand what the data means?

The answers to these three questions depend on modeling what is meant by *right*. Recognizing the highly decentralized nature of data management in the enterprise, the DRM team advises participants to implement a standard template for building their information and data architecture. But using a standard template, in and of itself, is not going to create information sharing. This must be coupled with common data categorization and data structure to make data more visible and usable across the enterprise, hence facilitating information and data sharing wherever common mission needs exist.

To give credence to this categorization and structuring theme, Figure 4 shows a revised DRM structure that incorporates those themes.

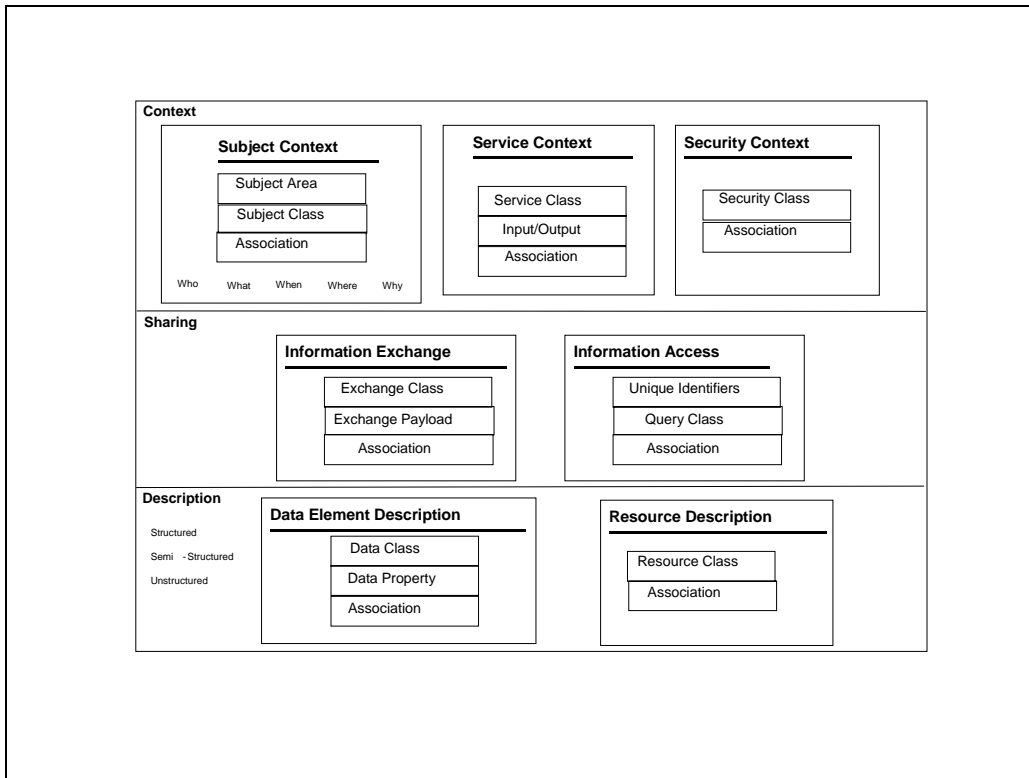


Figure 4. Revised DRM Structure

The DRM XML Profile and Schema

Based on this revised structure, an XML Profile of the DRM as created, shown here in Figure 5. To facilitate the inventory, cataloging and discovery of information and data holdings, an FEA DRM XML Schema was built based on this profile. Best-practice modeling, linking and modularity principles were applied to its design.

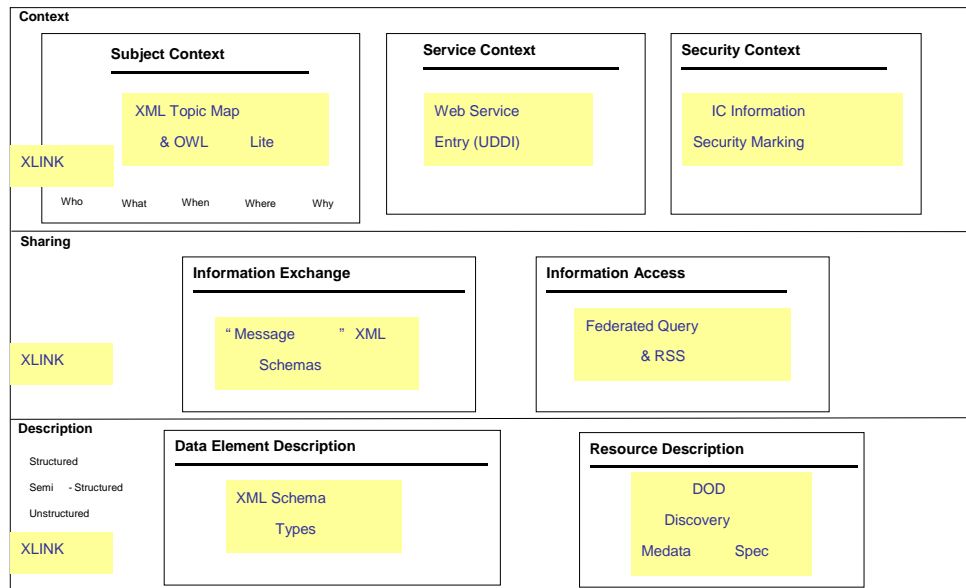


Figure 5. XML Profile of DRM

The DRM XML Schema is a useful template for collecting metadata because it represents all three major DRM standardization areas: categorization, exchange and structure. The Data Description section provides a standard means for describing data and data sources clearly, concisely and unambiguously. The Data Sharing section provides a standard means for describing and collecting information on data exchanges and data sharing capabilities. The Data Context section provides a standard means for representing taxonomies used to categorize data.

The Schema will enable participants to share XML instance documents regarding data categorization, exchange and structure across the boundaries of their entities. It will also facilitate data modeling efforts by supporting such capabilities as deriving physical data models from logical data models. Instances of the Schema may also potentially be used for configuration management and operational purposes, such as automatically configuring and/or categorizing a data source and supporting exchanges based on a service-oriented architecture (SOA).

Thus the DRM XML Schema provides an open and well documented standard to enable organization and categorization of information, in ways that are searchable and interoperable across domain boundaries. And it does not preclude the use of other metadata processes or standards as appropriate.

USAF DRM Implementation

The DRM XML Schema presents a design pattern for the Air Force; nonetheless, AF participants still need to build to that pattern. To implement DRM within the USAF enterprise environment, the Headquarters Architecture and Standards Group in coordination with the Operations Support and Warfighting sub-enterprise domains has undertaken the initial steps in defining the categorization of data, based on business context; exchange of data based on information sharing; structure of data based on data descriptions. The end goal is to create an integrated USAF DRM to support data sharing.

Through the Architecture Integration Working Group, much has already been accomplished towards the discovery of common data across USAF Lines of Business (LOB) and communities of Interest (COIs). There is early evidence of data sharing opportunities and redundancy elimination by mapping systems to data subject areas and information classes, and a high percentage of reuse has been observed. Current efforts are focused on categorizing data using external standards as needed.

Thus the DRM may prove equal to the challenge of “hooking” the USAF into an overarching framework (i.e., FEA) that will provide the management oversight, direction and implementation guidance that seem to be missing when we looked across current AF XML implementations directed towards net-centricity. It imposes a consistent, enterprise-wide method for organizing data, although participants still need to agree on the specifics of the metadata and XML artifacts they will generate and where to register them. In other words, the DRM provides a blueprint for addressing the organizational aspects of the data migration, if the AF can continue to incentivize communities to build to the pattern.

Conclusions and Future Directions

We began this paper by observing that the USAF is migrating to net-centric operations and is building its interoperability solutions around XML technologies. This paper responded to some of the issues and questions raised in part because of the pivotal role XML technologies play in this migration. The USAF is not alone in facing data sharing issues in its business practices, some attributable to legacy choices, some attributable to migration choices; nor is it alone in facing questions about how best to evolve current practices toward enterprise-wide solutions. For example, in Iraq today, the Army is dealing with 184 different battlefield systems operating on 11 different networks. Those systems only partially share information, particularly at the tactical level, which is key to operational success. [10]

During a speech he wrote for the Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance Conference on 2 March 2005, Col. Stuart Whitehead, director of the Army’s Training and Doctrine Command’s (TRADOC) Program Integration Office for Battle Command, said “We had created the moral equivalent of a Battle Command ‘Tower of Babel.’” The assumption that battle command systems can be developed in a linear fashion and successfully merged after the fact has not been realized in practice. This increases our suspicion of current XML-driven “develop in

parallel now, merge later” approaches towards achieve net-centricity and compels us to urge rethinking this plan.

All information-age interoperability pathfinders face challenges; not the least among them the cut-over, migration and roll out of new technologies and business practices. The community has struck down previously unexplored roads in the past; they collectively understand many of the tradeoffs and have learned potential risks they face in striving for information that is accessible, visible, understandable, interoperable and integratable. Reliance on XML technologies introduces some new, perhaps less understood challenges, some of which we have discussed here. It is our position that technologies, and piecemeal solutions, however expedient, must be embedded in an overarching framework that supports the end-to-end data sharing problem in the context of the visionary goals set by an enterprise.

We note some positive actions in addressing these challenges from such a broad perspective. From the USAF, as pointed out in this paper, the DRM effort is the largest solution yet attempted to address enterprise-scale interoperability. It exists in a supporting, well-thought-out framework as part of a DoD-wide solution and aims at the heart of the problem – how to organize and integrate data from disparate sources. To begin remodeling its ‘Tower of Babel,’ the Army’s TRADOC has created a Battle Command migration plan that will look horizontally and vertically across the Army to link efforts in acquisition, testing, funding, leader development, training and fielding, addressing the end-to-end problems from the foundation up, instead of precariously stacking the solution building blocks and hoping they don’t topple over. These steps in the right direction increase our confidence that none of the challenges we have spelled out in this study is insurmountable; but some course correction may be needed and is certainly possible. We call upon the communities-at-large to engage – as we continue to do – in thoughtful self-analysis and to formulate actionable responses and solutions for the way-ahead.

Acronyms

CIO	Chief Information Officer
CoT	Cursor on Target
COTS	commercial off-the-shelf
CWID	Coalition Warfighter Interoperability Demonstration
DoD	Department of Defense
JB	Joint Battlespace Infosphere
GMTI	Ground Moving Target Indicator
NATO	North Atlantic Treaty Organization
USAF	United States Air Force
USMTF	United States Message Text Format
XML	eXtensible Markup Language

References

- [1] DoD Memorandum: “Net-Centric Data Strategy”, May 9, 2003, Mr. John P. Stenbit, DoD CIO
- [2] “Report on Building the Joint Battlespace Infosphere, United States Air Force Scientific Advisory Board, Volumes 1 and 2”, SAB-TR-99-02, December 2000
- [3] See http://diides.ncr.disa.mil/xmlreg/user/ns_inv_count.cfm?namespace_list=All.
- [4] *NESI Net-centric Guidance*, <http://nesipublic.spawar.navy.mil/docs/part3>
- [5] Modeen, Ray, *Net-Centric Fundamentals*, presented at *Net-centric Bootcamp*, MITRE Corporation, February 2005
- [6] Office of the Assistant Secretary of Defense, Networks and Information Integration, Chief Information Officer (ASD NII/CIO), *Net-Centric Checklist version 2.1.4*, 30 July 2004
- [7] See <http://xml.gov/registries.asp>
- [8] See <https://xml.core.gov/>
- [9] See <http://cosine.cim3.net/cgi-bin/wiki.pl?DataReferenceModel>
- [10] DiMascio, J, “Army’s Battle Command Systems Not Working Together As Planned,” *Inside the Army*, 2 March 2005.