The Data Reference Model

Management Strategy

Draft

October 17, 2005

FEAPMO

FEDERAL ENTERPRISE ARCHITECTURE PROGRAM MANAGEMENT OFFICE

The Data Reference Model Management Strategy

Preliminary versions of this document were vetted with a limited number of senior executives who indicated a vital interest in the future of government data. The final version is due to OMB on November 17, 2005 and OMB plans to publish the DRM Management Strategy on December 17, 2005. This draft Management Strategy document is offered to the Federal CIO Council as a starting point for their discussions. The outcome of their deliberations will be reflected in the Architecture and Infrastructure Committee submittal to OMB.

Table of Contents

Executive Summary

The Federal Enterprise Architecture (FEA) Data Reference Model (DRM) Management Strategy is written for Assistant Secretaries, Directors, Chief Information Officers, Chief Financial Officers, Chief Architects, Project Managers, Programmatic Business Leaders, and Information Technology (IT) Managers and Specialists.

It is common knowledge that information technology advancements have made it easier to search, exchange, and manipulate information. However, these advancements, which show no signs of abating, can neither replace nor abrogate the senior leadership's responsibility to understand its organization's business needs for data, and design the most efficient solutions for delivering them to all necessary and appropriate users. Within the federal government—and often vertically with the state and local/tribal levels—there are mission critical needs to understand information as a whole. This is simply because delivering services to citizens requires seamless "data" cooperation between and amongst departments and agencies, and often beyond with appropriate and legal stakeholders and partners—termed Communities of Interest (COI). Government has heard the lessons from dramatic and harrowing events loud and clear, and cannot afford inefficiency attributable to an inability to use, reuse, share or understand our data and the information power they give us. The DRM now provides a concrete means for improving the capacity for mission-related information access, sharing and reuse across agencies while also increasing the downstream value of our government's strategic information assets.

Among past criticisms of government information management are contentions that we locked up our data in unconnected stovepipe legacy systems; that we did not define or manage data appropriately so that they are authoritative and consistent; and that we allowed replications of the same data over and over, requiring vast data storage resources that added no real value beyond local convenience, which now, is rarely necessary. The DRM Management Strategy premise, described herein, is that we must create a coordinated FEA, or more simply, a guiding architectural blueprint for all government information investments. Further, this blueprint must align our data, information and the systems enabling their use with our strategic goals and business lines; without which, our future could look very much like the past. We all know that is simply unacceptable.

We cannot just upgrade our stovepipe legacy systems with advanced information technology. The result would be, in large part, a future of search engine-driven stovepiped information networks, with data still hidden from appropriate discovery, reuse and sharing. Quick accessibility of pertinent, highly meaningful data from different networks will remain a persistent problem.

As Internet-based approaches create the potential for mission results at higher performance levels, constructive "line of sight" communications by agency officials are necessary to transcend complexity and multiple perspectives in order to achieve the service results expected by citizens. The DRM, together with the other reference models serves as the "common language" for vital communications and strategic planning. Becoming agile today means establishing common ground and trust among all linked process stakeholders sharing and reusing data. Without the "grounding" of reference models, including the DRM, to counter and anchor the daunting complexities of scale and specialization, strategic dialogue among the right parties is hampered. The result is often un-agile, uncoordinated and ad hoc approaches that are sub-optimal at best.

The DRM development itself demonstrated the next level of maturity for Enterprise Architecture showing how all major agencies collaboratively created the new DRM which:

1) Enables formation of credible agreements around the meanings of fundamental data concepts

2) Embraces a governance approach that accelerates multi-agency harmonization and information exchange agreements

3) Implements an iterative approach to validate and test the value of joint agreements for improved mission performance

Until now, there has been no effective common data architecture framework for departments and agencies to follow in support of cross-agency and intergovernmental shared business goals. The FEA DRM is designed to provide that common framework for the effective sharing of government information across organizational lines while respecting the security, privacy and appropriate use of that information. In other words, the DRM is a framework to enable information sharing and reuse across the federal government through the standard description and discovery of common data and the promotion of robust data management practices. The DRM offers guidance for agency agility in drawing out the value of information as a strategic mission asset. The DRM also provides a starting point and organizing mechanism through authoritative COIs for mission-related information exchanges across agencies. Through the application of Enterprise Architecture, this data sharing is always done in response to identified strategic goals and associated business needs.

The DRM is the flexible structure or semantic sinew that will enable the existing common framework of the FEA Reference Models to be used by agencies to accelerate more rapidly from the “declarative – what is” to the “dynamic what can be.” It is a catalyst toward multiplying the value of existing “silo” and “solo” data holdings through better description and understanding of what the data mean. This makes more data “high-performance” ready to meet critical needs for real-time relevance—the right information to the right people at the right time, from multiple sources. Implementing the DRM should logically lead to a decrease in the number of small incompatible information systems and a future where data reuse, shared systems and core capabilities, used by many, deliver better information to those who need it, at far less cost. This will surely be appreciated by our public, stakeholders and especially our state and local/tribal partners.

The DRM Management Strategy answers five key questions (and corollaries) about the DRM:

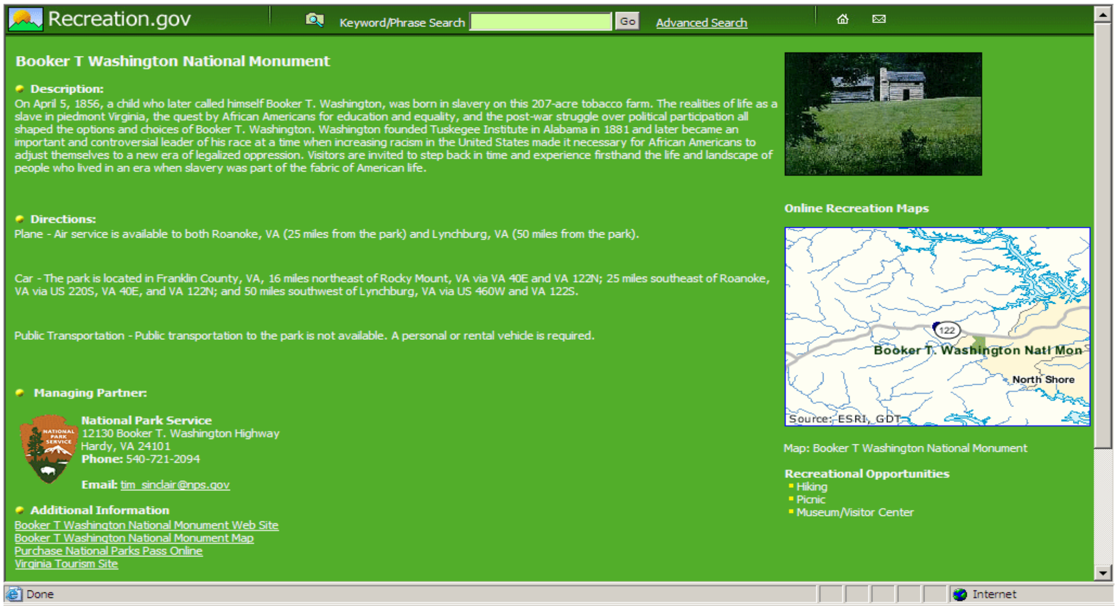
- What is it? (What it is not?)
- Why is it important? (What is its value?)
- Who are the key players? (Why are they key?)
- How will it work? (What is the concept and vision?)
- What is required of agencies? (Must data architectures be redone?)

Introduction

In the government, information is generated all around us. Regardless of where we are and who we are, we must all make decisions using and based on information and data. The never-ending data challenges are how to find the right pieces of information, how to grasp appropriate understanding, and how to effectively manage and maintain that data so we can successfully meet the needs of our government at all levels. We work in a knowledge-intensive world; government employees and contractors, citizens and business partners, including the non-profit stakeholders that we support and work with, and state and local/tribal governments, are all depending on the federal government for consistent, quality, and timely data and information.

Data and information management are multi-dimensional and represent a tough business problem that is at the heart of a knowledge-intensive government. While some government data need to be exposed and discovered, other data have to be managed and protected. Getting the right information to the right people at the right time is at the heart of second generation of E-Government. We need a data and information strategic approach and a roadmap for the challenges we currently face and the unknowns ahead of us. We also know we can always do better.

What makes data management challenging is that the data are either hidden in the processes that produce desired results or are not commonly understood. For example, when a citizen makes a reservation at a National Park, that citizen sees a simple web display illustrated in Figure 1.a. What the citizen does not see are the data behind the display as illustrated in Figure 1.b.



Booker T Washington National Monument

- **Description:**

On April 5, 1856, a child who later called himself Booker T. Washington, was born in slavery on this 207-acre tobacco farm. The realities of life as a slave in piedmont Virginia, the quest by African Americans for education and equality, and the post-war struggle over political participation all shaped the options and choices of Booker T. Washington. Washington founded Tuskegee Institute in Alabama in 1881 and later became an important and controversial leader of his race at a time when increasing racism in the United States made it necessary for African Americans to adjust themselves to a new era of legalized oppression. Visitors are invited to step back in time and experience firsthand the life and landscape of people who lived in an era when slavery was part of the fabric of American life.

- **Directions:**

Plane - Air service is available to both Roanoke, VA (25 miles from the park) and Lynchburg, VA (50 miles from the park).

Car - The park is located in Franklin County, VA, 16 miles northeast of Rocky Mount, VA via VA 40E and VA 122N; 25 miles southeast of Roanoke, VA via US 220S, VA 40E, and VA 122N; and 50 miles southwest of Lynchburg, VA via US 460W and VA 122S.

Public Transportation - Public transportation to the park is not available. A personal or rental vehicle is required.

- **Managing Partner:**

National Park Service
12130 Booker T. Washington Highway
Hardy, VA 24101
Phone: 540-721-2094

Email: tim_sinclair@nps.gov

- **Additional Information**

Booker T Washington National Monument Web Site
Booker T Washington National Monument Map
Purchase National Parks Pass Online
Virginia Tourism Site

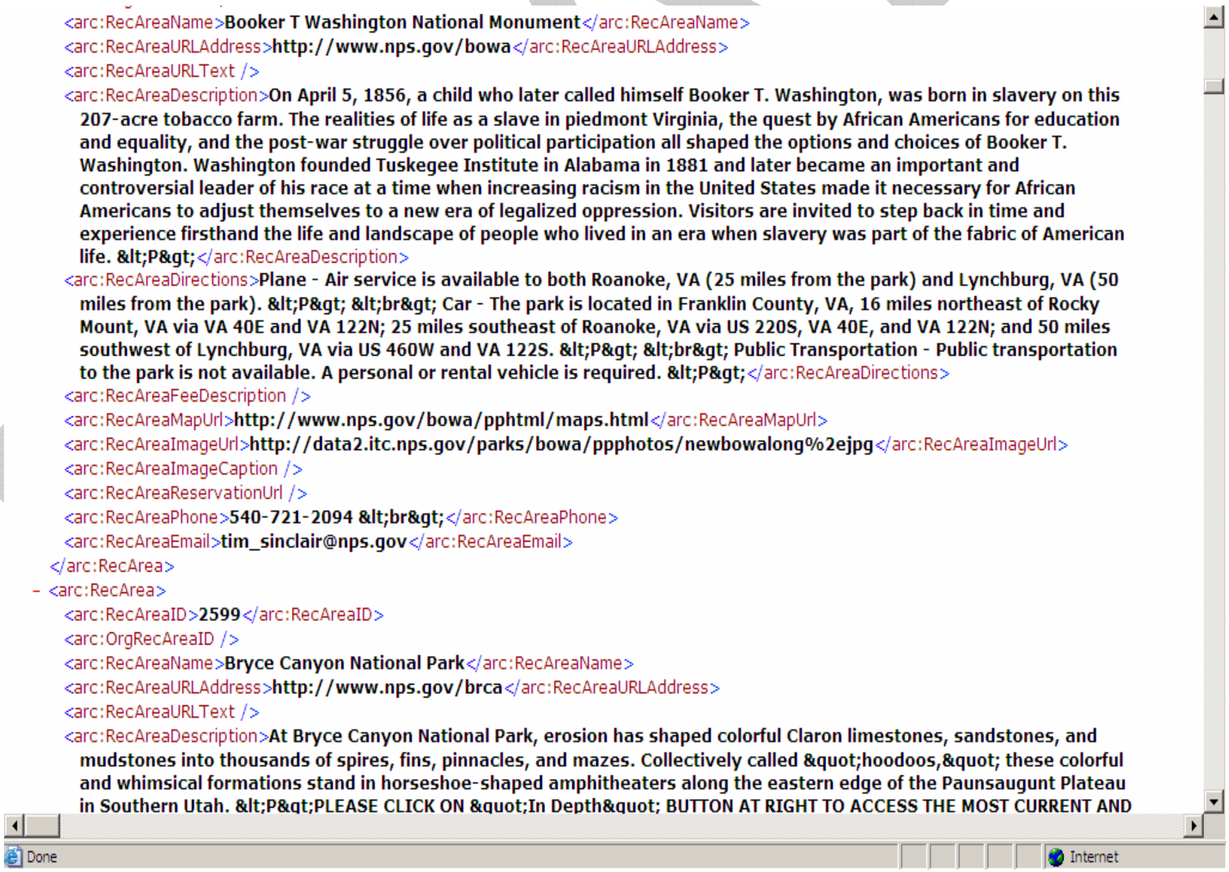
Online Recreation Maps

Map: Booker T Washington National Monument

Recreational Opportunities

- Hiking
- Picnic
- Museum/Visitor Center

Figure 1.a: Citizen View of Recreation Reservation



```
    <arc:RecAreaName>Booker T Washington National Monument</arc:RecAreaName>
    <arc:RecAreaURLAddress>http://www.nps.gov/bowa</arc:RecAreaURLAddress>
    <arc:RecAreaURLText />
    <arc:RecAreaDescription>On April 5, 1856, a child who later called himself Booker T. Washington, was born in slavery on this
      207-acre tobacco farm. The realities of life as a slave in piedmont Virginia, the quest by African Americans for education
      and equality, and the post-war struggle over political participation all shaped the options and choices of Booker T.
      Washington. Washington founded Tuskegee Institute in Alabama in 1881 and later became an important and
      controversial leader of his race at a time when increasing racism in the United States made it necessary for African
      Americans to adjust themselves to a new era of legalized oppression. Visitors are invited to step back in time and
      experience firsthand the life and landscape of people who lived in an era when slavery was part of the fabric of American
      life. &lt;P&gt;</arc:RecAreaDescription>
    <arc:RecAreaDirections>Plane - Air service is available to both Roanoke, VA (25 miles from the park) and Lynchburg, VA (50
      miles from the park). &lt;P&gt; &lt;br&gt; Car - The park is located in Franklin County, VA, 16 miles northeast of Rocky
      Mount, VA via VA 40E and VA 122N; 25 miles southeast of Roanoke, VA via US 220S, VA 40E, and VA 122N; and 50 miles
      southwest of Lynchburg, VA via US 460W and VA 122S. &lt;P&gt; &lt;br&gt; Public Transportation - Public transportation
      to the park is not available. A personal or rental vehicle is required. &lt;P&gt;</arc:RecAreaDirections>
    <arc:RecAreaFeeDescription />
    <arc:RecAreaMapUrl>http://www.nps.gov/bowa/pphtml/maps.html</arc:RecAreaMapUrl>
    <arc:RecAreaImageUrl>http://data2.itc.nps.gov/parks/bowa/ppphotos/newbowalong%2ejpg</arc:RecAreaImageUrl>
    <arc:RecAreaImageCaption />
    <arc:RecAreaReservationUrl />
    <arc:RecAreaPhone>540-721-2094 &lt;br&gt;</arc:RecAreaPhone>
    <arc:RecAreaEmail>tim_sinclair@nps.gov</arc:RecAreaEmail>
  </arc:RecArea>
- <arc:RecArea>
    <arc:RecAreaID>2599</arc:RecAreaID>
    <arc:OrgRecAreaID />
    <arc:RecAreaName>Bryce Canyon National Park</arc:RecAreaName>
    <arc:RecAreaURLAddress>http://www.nps.gov/brca</arc:RecAreaURLAddress>
    <arc:RecAreaURLText />
    <arc:RecAreaDescription>At Bryce Canyon National Park, erosion has shaped colorful Claron limestones, sandstones, and
      mudstones into thousands of spires, fins, pinnacles, and mazes. Collectively called &quot;hoodoos,&quot; these colorful
      and whimsical formations stand in horseshoe-shaped amphitheaters along the eastern edge of the Paunsaugunt Plateau
      in Southern Utah. &lt;P&gt;PLEASE CLICK ON &quot;In Depth&quot; BUTTON AT RIGHT TO ACCESS THE MOST CURRENT AND
```

Figure 1.b: Technical View of Recreation Reservation

There are many repeated instances when data, not unlike the data displayed in Figure 1.b, are vitally needed and could be used by one or more other agencies. Agencies that have a need for this type or source of data, but lack the awareness that it already exists, can end up spending additional time and scarce resources to recreate the same or similar data to support their missions. The lack of data visibility and accessibility results in costly redundancy of data within and among government agencies. Until now, we have not had the catalyst for a solution that the DRM provides.

Additionally, the government must work in real time, critically, constantly and dynamically gathering and processing data whether they are about:

- Protecting us from terrorists
- Tracking and responding to storms
- Providing rapid disaster recovery
- Collecting and refunding tax payments
- Distributing citizen and public benefits
- Sharing critically important research findings

Our ability within the federal community to quickly discover, access, and understand data and information must mature beyond what has been accomplished in our first wave of E-Government solutions. We must start using best practices to further enable appropriate data and information sharing and reuse. The problem of information sharing is known to all senior government officials but has become mission critical to many, especially in programs requiring rapid data acquisition and analysis to save lives or prevent unthinkable events. This effort required to prepare the federal government for this revolution in data awareness and readiness cannot be addressed exclusively with a top-down approach or merely with a set of data communities and grassroots efforts; rather it demands a more strategic and tactical approach recommended and described in this version 1.5 of the new DRM. New technologies and the standards for their use and implementation are increasingly available, but it is the strategic visioning—contained herein and in the revised DRM and incubated in the leadership to implement it—which is needed to enact real results like stopping crimes, protecting lives and reducing pain and suffering.

The DRM provides a framework to begin to address the following key questions everyone is asking:

- How do agencies discover what data are available for sharing and re-use?
- How do agencies with data available for sharing and re-use, make them visible and accessible?
- How do agencies ensure procedures for security and appropriate use of the data shared are considered and followed?
- How do agencies reduce unnecessary redundancies in the collection and storage of data?
- How do agencies drive down IT system costs through effectively managing data?

And perhaps the most critical question is: how does the federal government create rapid information sharing in responding to a time sensitive event or crisis?

DRM – What is it?

The DRM represents current best thinking and working agreements by CIOs and their staff around key data concepts that need to be commonly understood for strategic dialogue around structured, repeatable business processes. In other words, the DRM is a framework to enable information sharing and reuse across the federal government through the standard description and discovery of common data and the promotion of robust data management practices. The DRM offers guidance for agency agility in drawing out the value of information as a strategic mission asset. The DRM also provides an authoritative starting point and organizing mechanism through authoritative Communities of Interest (COIs) for mission-related information exchanges across agencies. As illustrated in Table 1, the DRM addresses three important standardization areas to support deliberation and joint action around structured, repeatable processes.

Table 1: Three Standardization Areas of the DRM

The three standardization areas are:	Meaning it:
Data Sharing	Supports the reuse and exchange of data where sharing consists of ad-hoc requests (such as query of a data asset), and exchange consists of fixed, re-occurring transactions between parties. Data Sharing is enabled by capabilities of the Data Context and Description areas.
Data Description	Provides a means to richly describe data, thereby supporting its discovery and sharing.
Data Context	Facilitates discovery of data through an approach to the categorization of data according to taxonomies, including linkages to the other FEA reference models.

The DRM is a federal-wide framework to meaningfully capture information on what categories of data we have, how we exchange or access data and how we understand what the data means. It simply requires that an agency understand, from its own perspective, what categories of data support its mission, what methods are used to exchange or access the data and how the data are or can be understood. Agencies already managing their data should be able to easily map the way it is done to the DRM, using it as a common framework. In the context of agencies, use of the DRM constitutes guidance that should be useful and useable, but that will not unnecessarily burden them with a lot of data architecture rework. On the other hand, agencies that have yet to establish their data architectures and data management processes should consider alignment with the DRM in choosing how they will proceed.

The only cases where the DRM is mandated to be followed or used, beyond the use as a high level framework, are when repeatable data architecture blueprints are needed for new Lines of Business (LoB) charted by OMB and the CIO Council, and in government cross-cutting missions and programs (including with state and local/tribal governments). The DRM then provides a common framework to enable data interoperability, harmonization and standardization across the federal government and with its stakeholders in response to identified business needs. It will allow all organizations within those LoBs at any level of the government to come to the table and quickly create information sharing mechanisms and focus appropriate attention on the data rather than acquisition and implementation of long-range, costly systems related to IT investments. This is the real value of the DRM; to put forth a repeatable framework which ensures the federal government can rapidly create information sharing and/or reuse data in response to LoBs or governmental cross-cutting needs. Thus, the DRM has the potential to contribute to a significantly improved strategic dialogue and broad-based action within and across agencies.

The DRM is not a prescriptive standard that compels departments or agencies to either categorize their data, or exchange, describe or manage data and systems in a specific way. It was designed to support the legislative and policy requirements the federal government already has in place. The DRM does not mandate information sharing. That is only done in agreements between departments and agencies in response to a business need and within their correct and proper legislated authorities and legal practices.

To gain a deeper understanding of the DRM, it must be looked at in the context of the Federal Enterprise Architecture. Enterprise Architecture provides the most strategic of all our tools to manage the complexity of the information challenge. It allows leadership to ask the tough questions: Do I have all the information to achieve my goals or to make the strategic decisions I face every day? As the common vocabulary of the reference models gains currency throughout the federal government, the transformative potential of individual agency transition strategies on behalf of the whole of government is amplified. The greatest potential of the DRM today to contribute to this transformation will be its ability to balance, unify and accord respect to all the multiple perspectives present when vital missions are rapidly stood up among familiar and unfamiliar partners.

The DRM is now at the heart of the FEA even as we know it needs further cross-referencing and alignment within the other reference models as they get updated. The complete and current set of reference models as the date of this publication are the:

- Performance Reference Model (PRM) Version 1.0
- Business Reference Model (BRM) Version 2.0
- Data Reference Model (DRM) Version 1.5b
- Service Component Reference Model (SRM) Version 1.0
- Technical Reference Model (TRM) Version 1.1

In other words, these interrelated Reference Models are touch points or a frame of reference allowing departments and agencies to design their own Enterprise Architecture while providing a management mechanism to identify, analyze and, where appropriate, facilitate cross-agency analysis and business-driven opportunities for collaboration and cost efficiencies. The DRM forms a real foundation for this collaboration and establishment of cross-cutting efficiencies.

To further illustrate how the DRM works in an Agency's Enterprise Architecture consider the following scenario outlined in Figure 2. A senior government official can strategically look and "see" that to achieve a specific performance goal (PRM), a strategy he or she may pursue is to publish a rule (BRM), there is data needed to develop that rule (DRM) which is dependent on information technology (SRM/TRM) to deliver the information necessary. So asking questions of the Enterprise Architecture, will point out, from a performance goal and business perspective, when information is missing or if exist where it is. It allows leadership to configure their IT systems for fast, efficient low cost delivery of information.

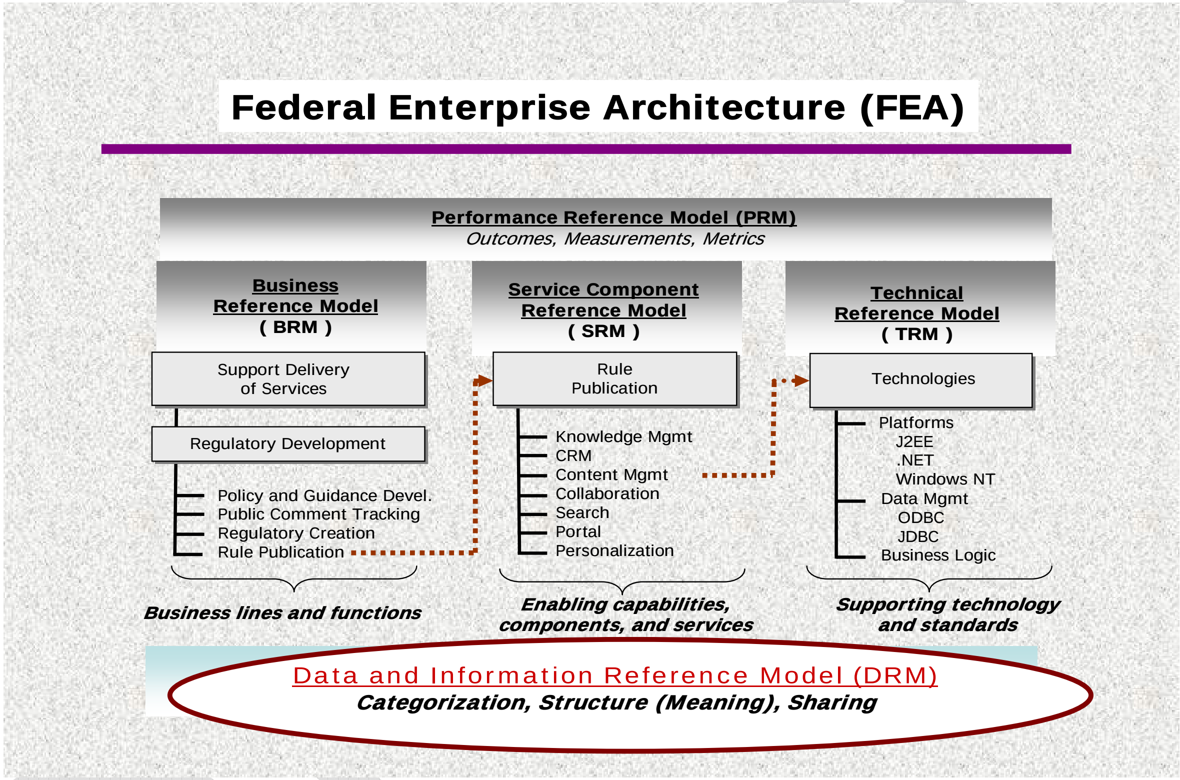

Federal Enterprise Architecture (FEA)
Performance Reference Model (PRM)
Outcomes, Measurements, Metrics
Business Reference Model (BRM)
Service Component Reference Model (SRM)
Technical Reference Model (TRM)
Support Delivery of Services
Regulatory Development
Policy and Guidance Devel.
Public Comment Tracking
Regulatory Creation
Rule Publication
Rule Publication
Knowledge Mgmt
CRM
Content Mgmt
Collaboration
Search
Portal
Personalization
Technologies
Platforms
J2EE
.NET
Windows NT
Data Mgmt
ODBC
JDBC
Business Logic
Business lines and functions
Enabling capabilities, components, and services
Supporting technology and standards
Data and Information Reference Model (DRM)
Categorization, Structure (Meaning), Sharing

Figure 2: FEA Architecture Reference Model Relationships

A simplistic view of the Data Reference Model is depicted in Figure 3. This model is based on three key questions a business owner really cares about: 1.) How do I find data? 2.) How do I exchange data? 3) How do I understand what the data means?

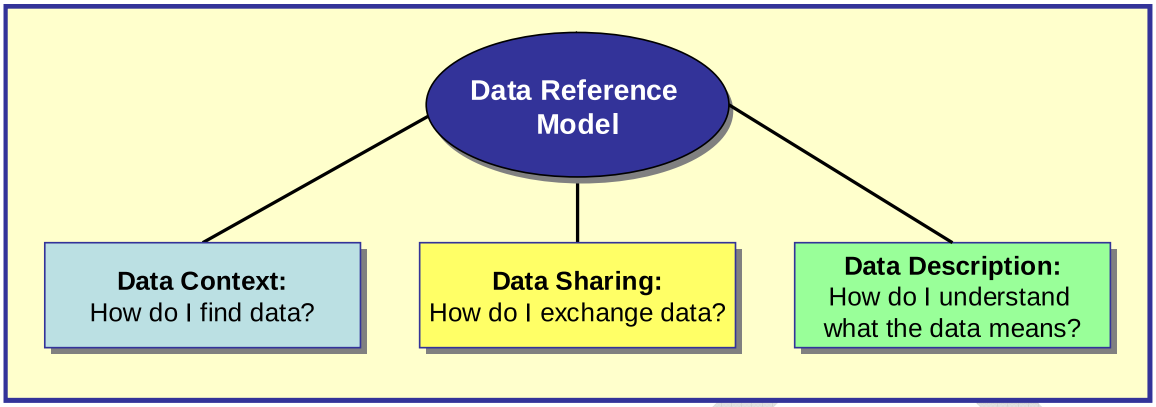

Data Reference Model
Data Context: How do I find data?
Data Sharing: How do I exchange data?
Data Description: How do I understand what the data means?

Figure 3: DRM Framework

Data Context: How do I find data?

Data Context facilitates the categorization of data in a business context to help an agency understand what data (no matter what format, text, maps, data in analytical systems) are needed to support a particular line of business. Data categorization will also help in the identification of duplicative data resources, discovery of data for reuse and enhancement of search capabilities.

Data Sharing: How do I exchange data?

Data Sharing is the process that provisions the data from an information source to an information consumer in order to meet a business requirement. Streamlining our information sharing mechanisms will be greatly appreciated by, and create cost savings for, state and local partners who deal with multiple federal departments and agencies.

Data Description: How do I understand what the data means?

Data Description facilitates the understanding of data once they are obtained. It is not enough to be able to find and access data; for decision-making data must be of trusted quality.

These three components of the DRM are powerful in their simplicity. It was designed to address a very real problem; how to make the most efficient use of our investment in data. If one agency already collects the data and they can be appropriately reused, why place an additional reporting burden on the regulated community or state and local governments? It provides the guidance to strategically support a department or agency in understanding the data it has so it can be deployed cost effectively and with appropriate use. The DRM is also the framework to enable data interoperability, harmonization and standardization across the federal government in response to an identified business need.

In doing so, the DRM serves as a support to government leaders in fulfilling their responsibilities under existing policy and law.

- Maximize the practical utility of and public benefit from information collected by or for the federal government and reduce collection burdens on the public (GPRA & OMB Circular A-130)
- Seek to satisfy new information needs through interagency or

intergovernmental sharing of information, or through commercial sources where appropriate, before creating or collecting new information (OMB Circular A-130)

- Implement processes to organize and categorize government information (E-Gov Act)
- Identify how information and data are created, maintained, accessed and used (OMB Circular A-130)
- Define agency data and describe relationships between mission and program performance and information resources to improve the efficiency of mission performance (OMB Circular A-130)
- Define data and describe relationships among data elements used in the agency's information systems and related information systems in other agencies, state and local governments and the private sector (OMB Circular A-130)
- Adopt a basic standard of data quality (including objectivity, utility and integrity) as a performance goal and incorporate information quality criteria into agency dissemination practices (Data Quality Act)
- Establish and maintain inventories of all agency information dissemination products and develop other aids to locating government information dissemination products such as catalogs and directories (OMB Circular A-130 & OMB M-05-04)

DRM – Why is it important?

Most senior executives in public or private service, when asked whether they have ready access to data within their respective enterprises, will answer "yes." Although it may take some insistence on their parts, they can obtain data. If they are asked whether they have all the data they need to manage their respective enterprises, they will most likely say "no." The data they receive in many instances are either: packaged to reflect the equities of those providing them in the best light; not accurate because they are not well maintained; not current; not provided in a timely manner; or worst of all, have multiple, competing "right" answers to any given question, again depending on the equities of those presenting these "right" answers.

The temptation of many senior executives is to view this disparity as an "IT problem," because IT systems store the data. The "fix", in an IT perspective, becomes better tools, better systems development, better databases, more sophisticated approaches to data and data management, and so on. The reality is that betting on technology is betting on a promise. The solutions to the issues of data management are simple, yet hard to accept. Data must be viewed as an enterprise resource, and they must be managed as such. Senior executives must accept responsibility for understanding their organization's business needs and the data required to support those needs, within the context of the nation's interests and objectives.

We must change our perspective. Currently, all data, much like politics, are local. Data are currently created and maintained to support the specific business processes that individual organizational elements are responsible for executing. The criticism of the management of government information in the past is that we locked up our data in unconnected stovepipe legacy systems. We must now view our data as a national asset.

CIOs are in a unique position to ensure that information required for a strategic goal or business objective is quickly and easily located, accessed and understood by those who need it. They must manage both the strategic, tactical and operational data and information needs of their agencies and they must be a partner with the total government infrastructure.

The DRM assists specifically by serving as strategic guidance for federal business leaders and CIOs constructing their respective Strategic Data and Information Plan used:

- To map from business and mission goals to data and information usage
- To integrate data and information architecture, and data and information management into complete business-driven enterprise architectures
- Provide a critical link between the business needs of an agency and its services and technologies.

The real test for the DRM is whether it improves the federal government's collective ability to deliver the right data to the right person at the right time as illustrated in Figure 4. The approach is driven by the needs of each agency and department to find the "right information" to meet their stakeholder and user needs.

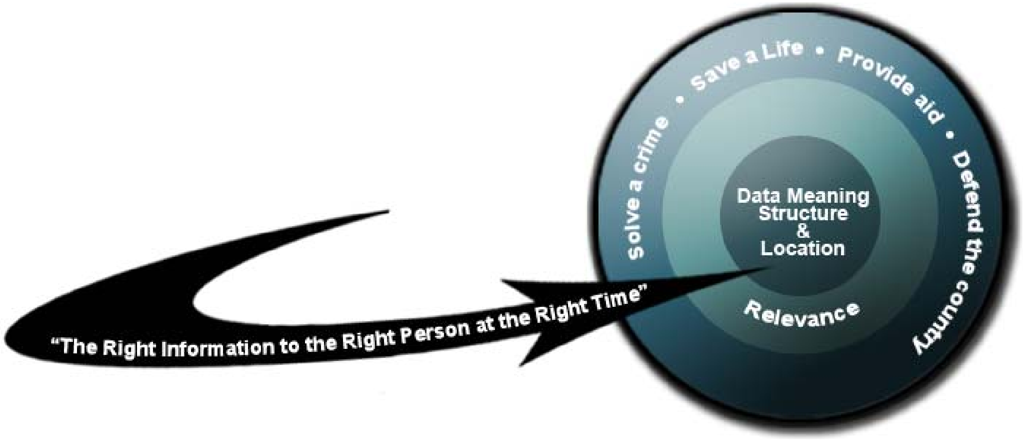

Save a Life • Provide aid • Defend the country
Solve a crime •
Data Meaning Structure & Location
Relevance
"The Right Information to the Right Person at the Right Time"

Figure 4: Strategic Goal of the DRM

As depicted in Figure 5, if departments and agencies choose to implement the DRM at a more detailed level than the high level framework, they will discover a way that their Data Architectures can support privacy and other attributes associated with data. While information sharing is an important goal in improving our ability to achieve the President's goal of a citizen-centered government and reducing duplication, we must not lose sight of the increasing importance of security and privacy issues. The DRM, allows for a way to tag data with security and privacy attributes, creates a mechanism to help us ensure appropriate use of data in accordance with existing legislation, directives and policies.

DRM – Who Are the Key Players?

The DRM framework will support the emerging prescribed repeatable processes that enable agencies to discover, share and seamlessly exchange data and information relevant to meet a business objective. At this stage, there is no federal mechanism for federal data governance, and this strategy does not create one. Instead it recognizes that that capacity must be developed over time and can only be accomplished with the involvement of many.

A key to success is that this capability must be developed through working the real problems of government. We must also recognize that the way business gets done today for a particular performance goals or business lines relies on what are being referred to as COIs

Communities of Interest Defined

COIs are collaborative groups of users who require a shared vocabulary to exchange information in pursuit of common goals, interests, missions or business processes.

COI members may include any number of members representing tribal, local, state, federal, public, private or other non-governmental organizations. COI members may also include cross-functional members representing data consumers, data producers, program managers, application developers and data sharing governance groups. The federal health care business is one example of a COI collaborating around a particular business need. It clearly spans the gamut of participants as subject matter experts.

COIs work to resolve common issues affecting their communities. COIs develop products to support an increase in information sharing, volume, speed and reach to known and unanticipated authorized users. They also provide organization and maintenance disciplines for the data. COIs can be formed on an ad hoc basis or they can be formally established. Data panels (working groups) at the COI level facilitate data sharing through common COI vocabulary development activities across and between COIs through:

- Establishment of COI data processes
- Creation of subject area categories or vocabulary
- Preparation of integrated data access plans or information exchange schemas
- Development of information exchange agreements
- Brokerage of conflict resolution among data stewards
- Identification of Authoritative Data Sources (ADS)

The COI concept is in practice today in the Department of Defense (DoD). COI products developed enable a consumer-oriented, ubiquitous global network of data assets and information services. COI product development includes the following:

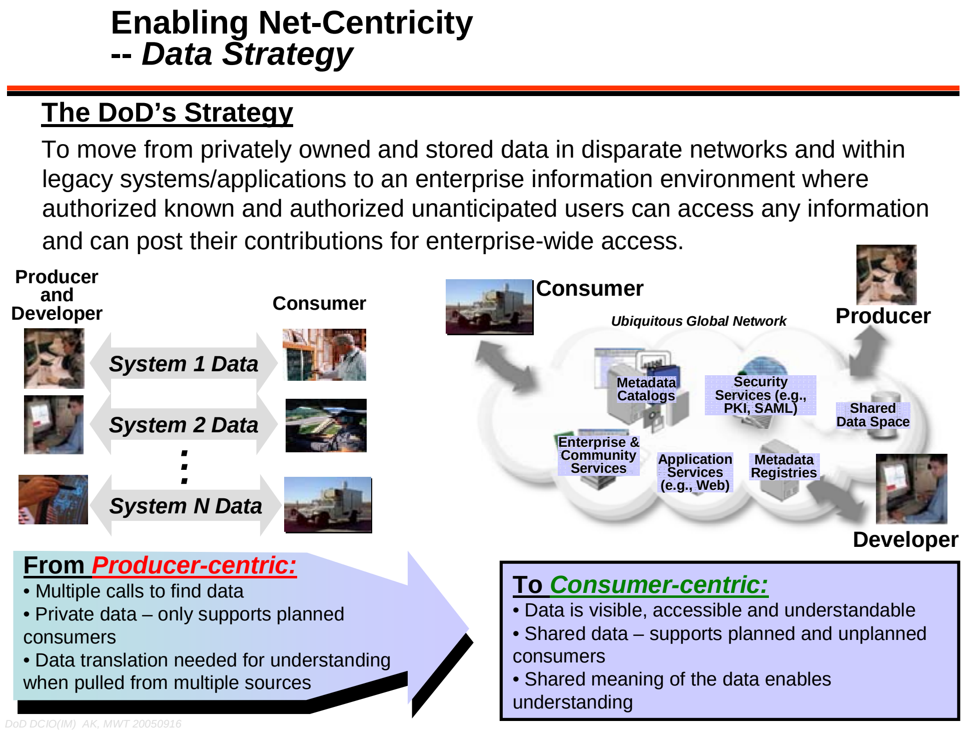
- Identify the COI data sharing challenge to be addressed
- Identify COI membership (cross functional, spanning organizations)
- Identify COI governance structure and process
- Identify data assets such as files, databases and information services
- Define shared vocabularies and taxonomies
- Define discovery metadata with extensions as needed to the DoD

Discovery Metadata Specification (DDMS)

- Register semantic and structural metadata to the DoD Metadata Registry
- Pilot COI product capabilities and integrate into programs of record
- Make the data visible, accessible and understandable across the DoD enterprise by tagging data assets and posting to searchable catalogs
- Expose data assets to include discovery metadata and information services to known and unanticipated authorized users (consumers, producers or developers)

As depicted in Figure 5, the COI data products are key enablers for the Net-Centric Data Strategy. The Department of Defense COI concept is in practice today and used to:

- Understand and remove the barriers in obtaining and using data
- Make the data visible, accessible and understandable
- Meet the goals of their net-centric data strategy



Enabling Net-Centricity
-- *Data Strategy*

The DoD's Strategy

To move from privately owned and stored data in disparate networks and within legacy systems/applications to an enterprise information environment where authorized known and authorized unanticipated users can access any information and can post their contributions for enterprise-wide access.

From *Producer-centric:*
- Multiple calls to find data
- Private data – only supports planned consumers
- Data translation needed for understanding when pulled from multiple sources

To *Consumer-centric:*
- Data is visible, accessible and understandable
- Shared data – supports planned and unplanned consumers
- Shared meaning of the data enables understanding

DoD DCIO(IM) AK, MWT 20050916

Figure 5: DOD COI Framework Supporting the Consumer-Centric Data Services Environment

The Management Strategy advocates adopting the COI as the approach to use when chartering new lines of business by OMB and the CIO Council. The fundamental questions of how to govern the core overlap between communities of interest, for data or exchange methods, needs more careful investigation and will be further explored by OMB and the CIO Council in 2006.

DRM – How will it work?

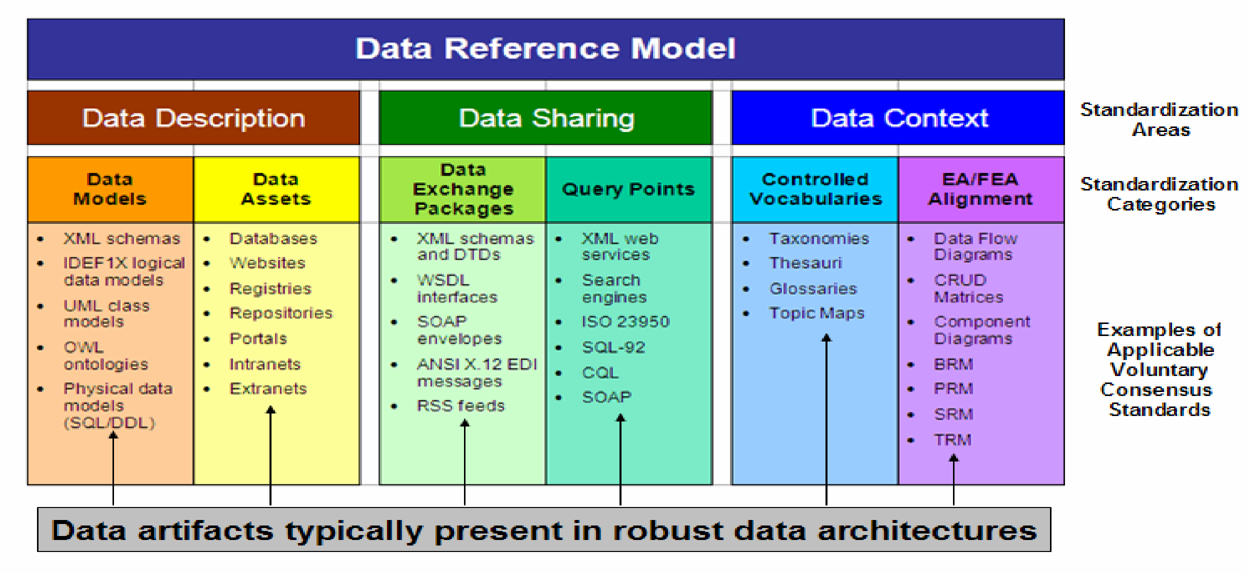
Key Strategies

1. Assist the departments and agencies in developing robust data architectures and mature data management practices as an integral part of their enterprise architecture program.

The management strategy advocates that department and agencies align with the two top standardization levels of the DRM as depicted in Figure 6. At the first level, the agencies need to bring focus to their data management activities within the DRM framework in three standardization areas: 1) Data Description, 2) Data Sharing and 3) Data Context, and at the next lower level identifying and documenting agency's use of various standards in each of these categories. For departments and agencies with more established data architectures, many of which already address advanced data capacities, the DRM does not in any way impede their efforts but provides a framework, in which their current efforts can be understood throughout the federal government.

The DRM will be of most immediate support to agencies that are just beginning their corporate data architectures. If adopted, agencies may find that they can leapfrog to a data architecture which supports necessary policy and business capabilities in security and privacy, geospatial and records management. Perhaps most importantly, a more detailed DRM approach to data architecture enables enterprise architecture to promote business objectives through rapid information sharing.

As for the 'data management' part of the equation, once senior leaders can "see" through data architecture what data are most important to them, they can ensure that they are well managed.



Data Reference Model

Data Description		Data Sharing		Data Context		Standardization Areas
Data Models	Data Assets	Data Exchange Packages	Query Points	Controlled Vocabularies	EA/FEA Alignment	Standardization Categories
XML schemas IDEF1X logical data models UML class models OWL ontologies Physical data models (SQL/DDL)	Databases Websites Registries Repositories Portals Intranets Extranets	XML schemas and DTDs WSDL interfaces SOAP envelopes ANSI X.12 EDI messages RSS feeds	XML web services Search engines ISO 23950 SQL-92 CQL SOAP	Taxonomies Thesauri Glossaries Topic Maps	Data Flow Diagrams CRUD Matrices Component Diagrams BRM PRM SRM TRM	Examples of Applicable Voluntary Consensus Standards

Data artifacts typically present in robust data architectures

Figure 6: Data Reference Model Standardization Levels

2. Increase departments and agencies ability to reuse data.

As a first step in improving the federal government's ability to discover and reuse the data it already has, departments and agencies will post their major "categories of data" in CORE.gov with a point of contact. CORE.gov is the federal government's portal which currently contains "reusable" services. Because Enterprise Architecture ensures that data is linked to the business it serves, federal, state and local departments and agencies will be able to search for data partners who are doing similar missions or basic functions for the first time. The pay off is finding reusable data to reduce new information collections, closing Government Performance Results data gaps and the exploring of collaborations leading to more standardized, understandable data.

3. Create the ability for rapid information sharing and data standardization through the DRM.

As new lines of business, E-Government initiatives or mission-driven interagency collaborations are jointly chartered by OMB and the CIO Council, the federal departments and agencies involved will be required to adopt a DRM framework as a common vocabulary between their data architects. Each effort will not have to reinvent the wheel in understanding the way multiple departments and agencies handle data architecture. This is the real value of the DRM—to put forth a repeatable approach which ensures the federal government can rapidly create information sharing or reuse data in response to a business need.

These newly chartered efforts will have their COIs officially sanctioned and their data standardization efforts will create the authoritative data for that line of business, E-Government initiative or mission area.

4. Explore how to govern the small set of government information which cuts across the federal government.

The federal government is highly decentralized in the way it manages data. Yet it appears that some core data could be standardized, because they are so common across multiple players. The Architecture and Infrastructure Committee (AIC) and OMB propose to explore this issue in 2006. The value is some "core" standards could increase data interoperability between systems.

5. Conduct a one-year test on use of the DRM to create rapid information sharing and more efficient data capacity within agencies.

The AIC, working with interested departments and agencies, will explore how the DRM can be used to create more agile, responsive information capabilities including more rapid information sharing.

Measuring DRM for Success

How can government agencies assess the capabilities of their DRM in context of the Enterprise Architecture and in support of their strategic objectives? One tool the OMB provides is the Enterprise Architecture Maturity Assessment Framework. This framework is designed to help agencies better understand the current state of their Enterprise Architecture and integrate their decision-making processes with Enterprise Architecture. The assessment contains criteria for five maturity levels for each of the architecture layers (performance, business, data, service component and technology) and references the DRM in the Data Architecture section.

The four maturity levels for data architecture leads to level five as quoted in part below:

"Activities: When applicable and required by law and policy, the agency has:

- documented procedures to ensure information is properly managed (i.e., created, collected, categorized, inventoried, preserved, disseminated, searched for, retrieved, and shared) in a manner consistent with applicable information policies and procedures;
- implemented such policies; and
- prepared and published inventories and otherwise made them available for use by all interested and authorized parties including other agencies and as appropriate, the general public, industry, academia, and other specific user groups.

Where applicable, the agency is using data standards to fulfill mission needs and meet the requirements of law and policy and has published the nature and use of such standards centrally for access by all interested parties, including the general public."

Requiring agencies to designate data sensitivity and quality is an option that could help to reduce risks in this area. Data quality involves getting the right data that are accurate and complete to the right person at the right time. Future measurement of the DRM within agencies and/or within COIs can be based to the extent which data can be:

1) Discovered (e.g., content made consistently findable or present)
2) Identified (e.g., content that is semantically consistent and reasonable)
3) Standardized (e.g., content that has syntactic and structural integrity)
4) Re-used (e.g., content that can be leveraged within and across domains to minimize redundancy)
5) Trusted (e.g., content that is 'reliable')
6) Good Quality (e.g., content that embodies and shares 'conformance, integrity, timeliness' among many business processes)
7) Protected (e.g., content that can be shared free of inappropriate disclosure or compromise)

Regardless of the options used for DRM measures, every measure must be defined in the context of a business or performance goal. If performance of the data can be directly associated with a business impact, this may become the most important measure among agencies, LoBs and COIs.

Conclusion – What is required of Agencies?

Three Simple Things:

1. Implementation of the DRM will be an evolutionary and iterative process. During the first year, 2006, the Management Strategy advocates that departments and agencies align with the two top standardization levels of the DRM. This harmonizing framework is simple but powerful. Through asking questions of their own Enterprise Architecture staff, federal leadership with a business need can look across its organization or the federal government to see what data is available, how to access it and understand it.

2. As a first step in improving the federal government's ability to discovery and reuse the data it already has, department and agencies will post their major "categories of data" in CORE.gov with a point of contact. CORE.gov is the federal government's portal which currently contains "reusable" services. Because of the way Enterprise Architecture works, federal, state and local departments and agencies will be able to search for data partners who are doing similar missions or basic functions for the first time. The pay off is finding reusable data to reduce new information collections, plug Government Performance Results data gaps, and the exploring of collaborations leading to more standardized, understandable data.

3. As new lines of business or E-Government initiatives are chartered by OMB and the CIO Council, the federal departments and agencies will be required to adopt a DRM framework as a common vocabulary between their data architects. Each effort will not have to reinvent the wheel in understanding the way multiple departments and agencies handle data architecture. This is the real value of the DRM—to put forth a repeatable approach which ensures the federal government can rapidly create information sharing or reuse data in response to a business need.

By adopting the Communities of Interest approach, these newly chartered efforts will drive the path to the adoption of standardized data that supports specific business lines of government.

The Action Oriented Road Map:

1. March 2006: Agency Enterprise Architecture submissions to OMB. Agencies who are ready will use the Core.gov registration process to register their data context architecture products

2. April 2006: OMB feedback to agencies concerning DRM alignment (as part of the OMB EA Assessment Framework)

3. June 2006: OMB/AIC uses Feb 06 submissions and CIO Council recommendations to identify/charter new lines of business with specific COIs

4. June 2007: Business-driven COIs publish their recommended data standards

5. June 2008: Deadline for agencies to implement applicable developed data standards