

federal
enterprise architecture
program



DRM

The Data Reference Model

Volume I, Version 1.0

September 2004

Document Organization

Document Organization	2
Executive Summary	3
Overview of the DRM	9
DRM Foundation	12
Use of the DRM	17
DRM Roadmap	23
Glossary	26

1

Executive Summary

Two of the major goals of the Federal Enterprise Architecture (FEA) are information sharing and the improved effectiveness of federal IT investments. Achieving these goals requires the ability to identify and use common data across the federal government. The Office of Management and Budget (OMB) has developed the Data Reference Model (DRM) to assist in the identification and use of common data. The DRM's primary purpose is to promote the common identification, use, and appropriate sharing of data/information across the federal government.

DATA REFERENCE MODEL

DRM

Volume I

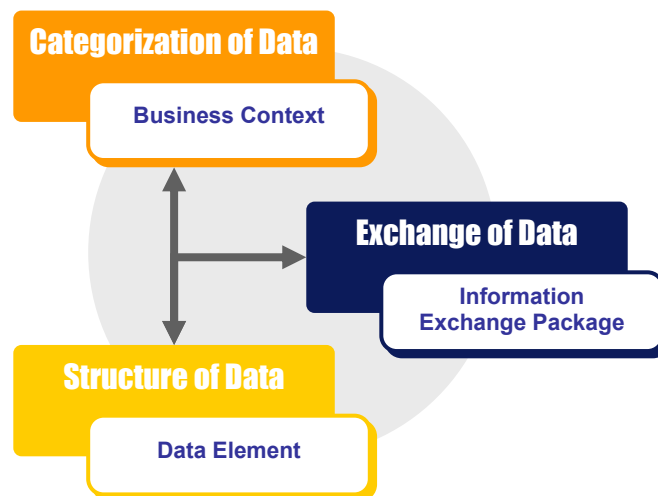
Version 1.0

To achieve this purpose, the DRM describes three basic standardization areas:

- **Categorization** of data
- **Exchange** of data
- **Structure** of data

Information sharing can be enabled through the common categorization and structure of data. By understanding the business context of data, DRM users will be able to communicate more accurately about the content and purpose of the data they require. This improved communication on the content and purpose of data will improve the ability to share information throughout the federal government. Exhibit A illustrates the basic areas of the DRM.

Exhibit A: DRM Structure



Categorization of Data: The DRM establishes an approach to the categorization of data through the use of a concept called **Business Context**. The business context represents the general business purpose of the data. The business context uses the FEA Business Reference Model (BRM) as its categorization taxonomy.

Exchange of Data: The exchange of data is enabled by the DRM's standard message structure, called the **Information Exchange Package**. The information exchange package represents an actual set of data that is requested or produced from one unit of work to another. The information exchange package makes use of the DRM's ability to both categorize and structure data.

Structure of Data: To provide a logical approach to the structure of data, the DRM uses a concept called the **Data Element**. The data element represents information about a particular thing, and can be presented in many formats. The data element is aligned with

the business context, so that users of an agency's data understand the data's purpose and context. The data element is adapted from the ISO/IEC 11179 standard.

Potential Outcomes of the DRM

The DRM's primary purpose is to promote the common identification, use, and appropriate sharing of data/information across the federal government.

Effective use of the DRM's approach may facilitate several outcomes, including:

- Improvement of federal agencies' ability to share information
- Improvement of the effectiveness of federal IT investments

Information sharing is improved through the use of an integrated DRM. In order to successfully share information, users must fully understand the context or business purpose of the information that is needed or produced. Using the DRM's business context facilitates an agency's ability to categorize its data in a common way. Once the categorization of the information is understood, agencies can use the DRM's structure to consistently describe the actual data element. In this way, the DRM's information exchange package uses the common approach to the categorization and structure of data to facilitate the sharing of information.

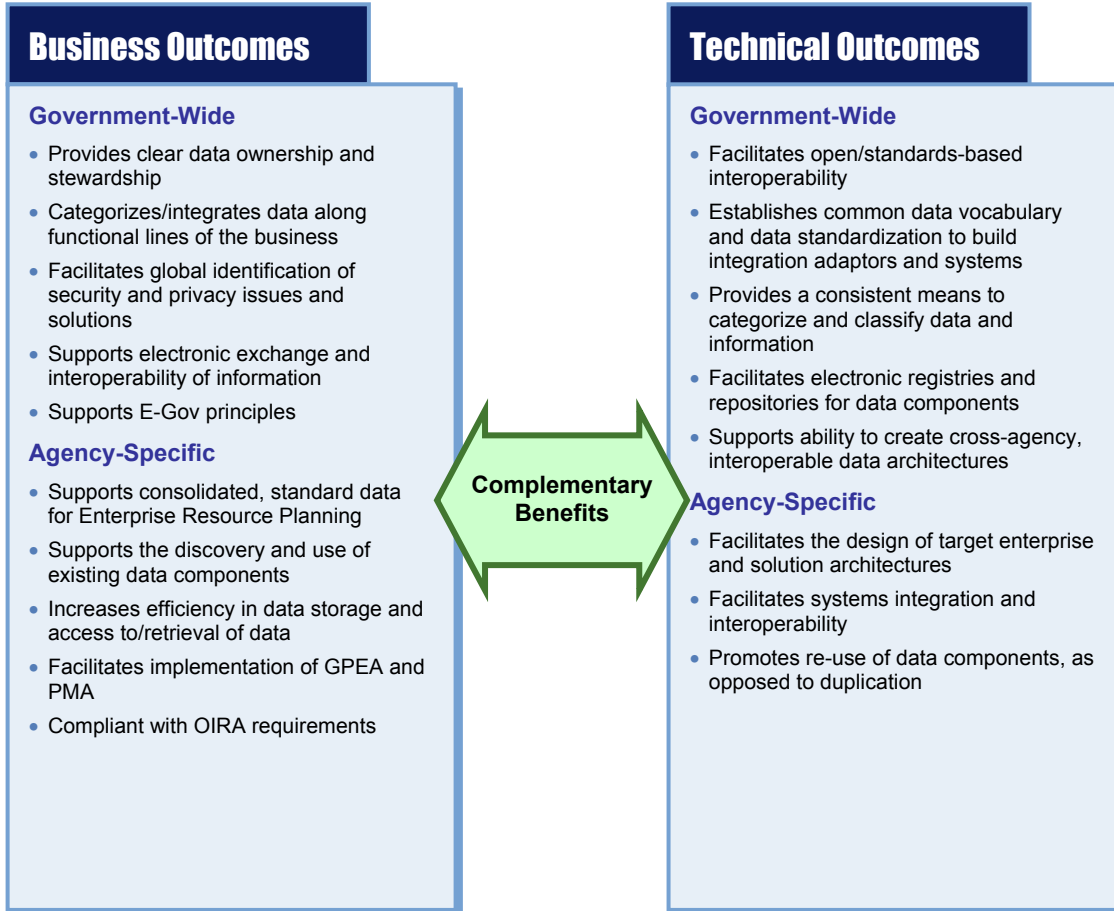
Adopting the DRM's approach to the categorization, structure, and exchange of data promotes the effectiveness of agency IT investments. Agencies that define and categorize their data using a common approach can identify IT applications that meet user requirements prior to proposing new IT investments. Agencies that exchange their data in a common structure increase the likelihood that other agencies can re-use IT investments as shared services within their own architectures.

In addition, use of the DRM supports legislative requirements such as the Information Quality Act (IQA)¹. The IQA, for example, establishes guidelines that are focused on the management and quality of data generated by federal agencies. Effective data management requires that agencies fully understand the use and purpose of the data they are managing. Furthermore, agencies can only manage the quality of their data if they

¹ Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001

can validate that its content and structure are accurate. Exhibit B illustrates several additional technical and business outcomes.

Exhibit B: Outcomes of the DRM



Use of this Document

Volume I of the DRM (this document) establishes a high-level overview of what the DRM is in the context of its ability to support a common approach to the categorization, exchange, and structure of data. The following table illustrates the content of each section within Volume I of the DRM and the type of user to which it applies.

Volume I Guide

Users	Categorization of Data	Exchange of Data	Structure of Data
General	Executive Summary	Executive Summary	Executive Summary
CIOs, Senior Managers	Section 2 provides a high-level overview.	In addition to Section 2, the content in Section 4 (Use of the DRM) may also be considered.	Section 2 provides a high-level overview.
Resource and Budget Managers	Section 2 provides a high-level overview.	In addition to Section 2, the content in Section 4 (Use of the DRM) may also be considered.	Section 2 provides a high-level overview.
Program Managers	Sections 2 and 4 provide perspective on how data is categorized and how that categorization is used in the exchange of information.	Section 4 provides perspective on how the DRM uses categorization and definition to facilitate the exchange of information.	Section 2 provides a high-level overview.
Architects	All sections.	In addition to Section 3, Section 4 provides perspective on the use and integration of DRM concepts.	Section 3 provides the detailed structure used to define and categorize data.

DRM Roadmap

This is the first of four volumes of the DRM. Future volumes of the DRM will focus on providing users with more details regarding the categorization, exchange, and standardization of data. OMB will also release a complementary Data Management Strategy. This strategy document results from collaboration between the CIO Council and the FEA PMO. It will address the governance and management of data.

The following table is a guide to the use of future volumes of the DRM. The table illustrates which volumes that users of the DRM will find most applicable. Detailed information regarding each volume is contained within Section 5 of this document.

Roadmap Guide

Volume I: DRM Overview and Purpose			
This volume identifies the high-level overview and purpose of the DRM. It applies to all users.			
Volume II: Exchange of Data	Volume III: Technical Data Standards	Volume IV: Technical Data Structure	Data Management Strategy
<ul style="list-style-type: none"> • Program Managers • Architects 	<ul style="list-style-type: none"> • Program Managers • Architects 	<ul style="list-style-type: none"> • Architects 	<ul style="list-style-type: none"> • CIOs, Senior Mgrs • Program Managers • Architects

In addition to collaborating with the CIO Council, OMB will continue to develop the DRM through a series of ongoing pilots with various federal agencies and programs. Pilots represent existing government programs that will contribute to the development and refinement of future volumes of the DRM.

2

DATA REFERENCE MODEL

DRM

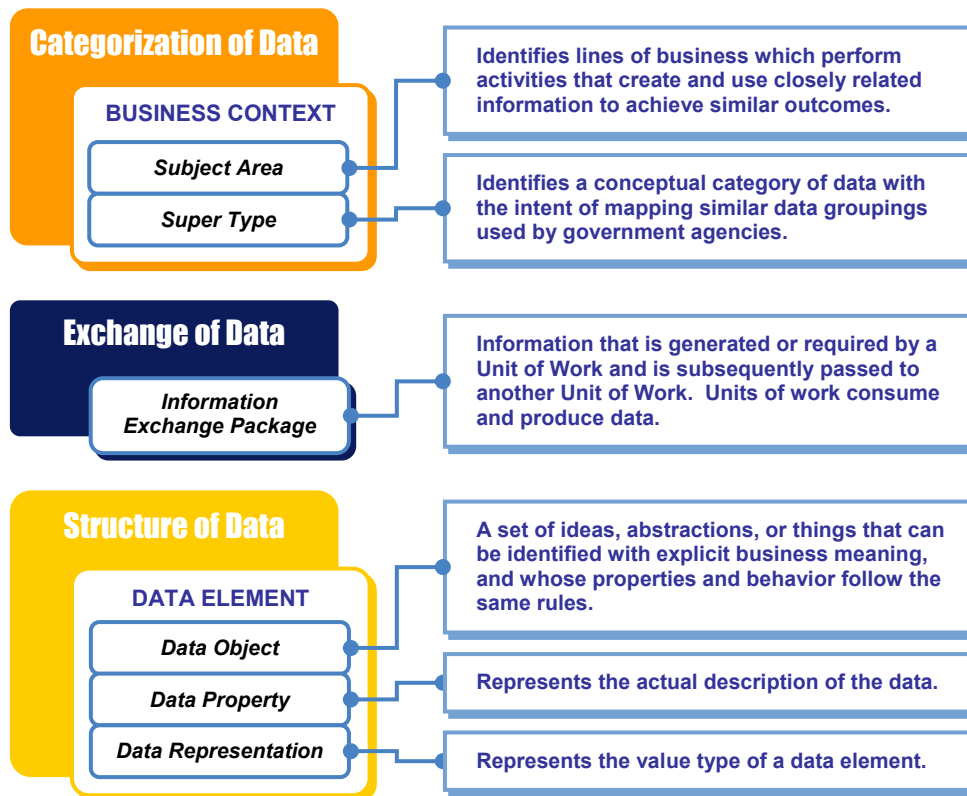
Volume I

Version 1.0

Overview of the DRM

The DRM's primary purpose is to promote the common identification, use, and appropriate sharing of data/information across the federal government. To meet this purpose requires an approach to the common categorization, exchange, and structure of data. Each of the areas of the DRM is described in the sections following Exhibit C. Exhibit C illustrates the integrated approach of the DRM.

Exhibit C: DRM Approach



The DRM presents a common approach to the categorization of data. To categorize data in a common way, the DRM establishes a **Business Context**. The business context represents the business use of a given set of data and makes use of the **Subject Area** and **Super Type** to further describe the business context of a given set of data. Subject areas represent a high-level set of business functions and are obtained from the FEA's Business Reference Model (BRM). Super types represent an additional level of definition with the business context and are generally related to specific business activities and/or processes that support the subject area.

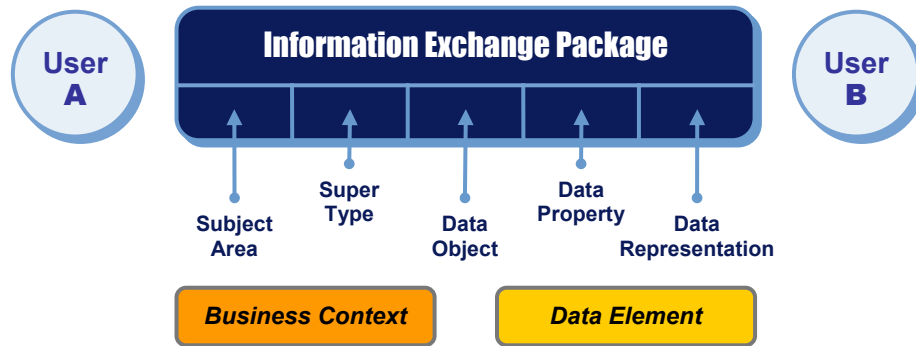
Exchange of Data

The exchange of data can be enabled through the DRM's information exchange package concept. The **Information Exchange Package** represents the actual message or combination of data that is exchanged between users of the data. The information exchange package brings the business context and data element (described in the structure of data section) together to define how a common transaction (the exchange of information and data) might appear. Exhibit D illustrates the information exchange

package concept. Future volumes of the DRM will continue to expand on the definition and scope of the information exchange package.

Note: The information exchange package can apply to data that is transmitted or to data that is shared or retrieved.

Exhibit D: Information Exchange Package



Structure of Data

The DRM uses the **Data Element** to describe the structure of data within a given business context. The structure of data is comprised of three elements adapted from the ISO/IEC 11179 standard. The **Data Object** is the set of ideas, abstractions, or things that can be identified with explicit business meaning and whose properties and behavior follow the same business rules. In the context of population health management, for example, the data object could be a vaccination. To further define the data object, the DRM's approach uses a **Data Property** and a **Data Representation**. The data property describes the data element. In the population health management example, the data property could be name, weight, potency, etc. (of a vaccine). The data representation is the value type of the data object. For example, representations could be plain text, integers, whole numbers, etc.

3

DATA REFERENCE MODEL

DRM

Volume I

Version 1.0

DRM Foundation

The DRM uses a flexible and standards-based approach to describe the categorization, exchange, and structure of data.

The categorization of data is achieved through the use of the BRM as the organizational construct for identifying the data's business context.

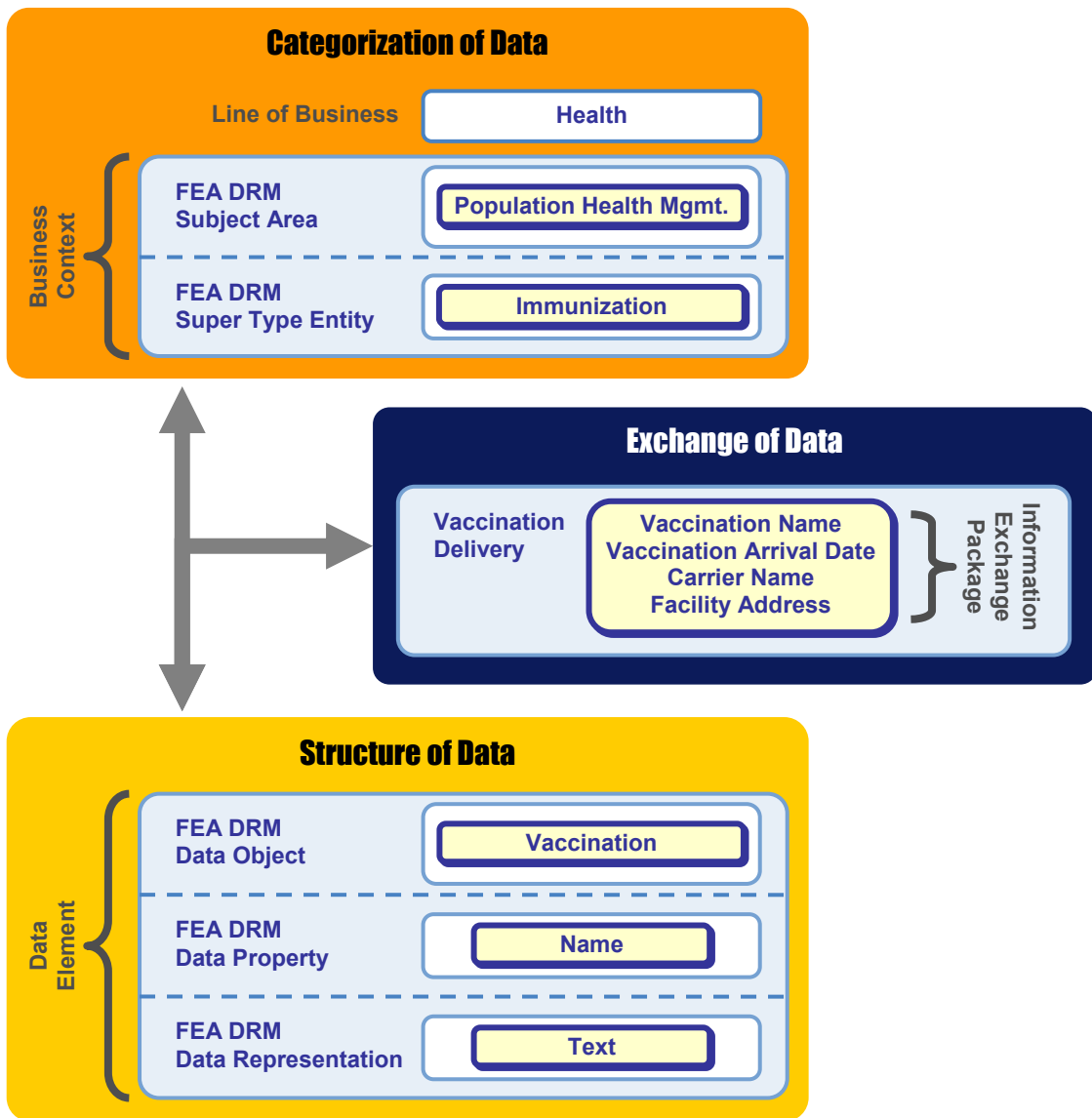
The exchange of data is facilitated through the information exchange package and describes a packaged set of data categorized into a message that can be re-used by other users. The specific standards associated with this concept will be defined in future volumes of the DRM.

In the DRM, a common approach to the structure of data is realized through an adaptation of the ISO/IEC 11179 standard as a guide. This standard provides the structure by which data can be defined in terms of its business context. The common structure implements a basic set of constraints and requirements while providing agencies the flexibility to use the DRM in a way that is consistent with their own business needs.

The DRM addresses business needs through its common approach to the categorization, exchange, and structure

of data. Exhibit E illustrates the DRM's approach in the context of an example involving organizations that provide health services. Through this example, one can see how the common categorization, exchange, and structure of data will allow agencies engaged in health services to share data regarding a variety of topics in a common way.

Exhibit E: DRM Example



Categorization of Data

Leveraging the BRM, categorization is achieved through the use of the BRM's sub-function. Use of the BRM's sub-function establishes the business context of a given set of data.

Note: *The categorization of subject areas at the Line of Business (LoB) level does not imply that data is applicable only within a line of business. Future volumes of the DRM will address the categorization and exchange of data across LoBs.*

Line of Business: Many organizations within the federal government provide services that contribute to the health of citizens and residents of the United States of America. These organizations are part of the Health LoB within the Services for Citizens business area (from the BRM). The LoB represents organizations that have a common business or program interest. The organizations in this case include federal programs and activities that ensure and provide for the health and well-being of the public.

Note: *Line of Business categories are obtained from the BRM.*

Subject Areas: Organizations within the Health LoB perform many business activities. One of these activities is population health management; this is equal to the DRM's subject area. The subject area is the first element used to represent the business context of a particular set of data. Population health management involves activities associated with the management and monitoring of health, health planning, and health management. Population health management is a sub-function and is located under the Health LoB within the BRM. BRM sub-functions are supported by lower-level activities that represent more detailed views of the business function. For example, immunization is an activity supporting the BRM sub-function of population health management.

Note: *Sub-functions are obtained from the BRM.*

Super Types: Super types represent lower level business activities that represent data that is used in support of the subject area. Super types provide an additional level of detail regarding the subject area. In the example illustrated in Exhibit E, the super type is an immunization which represents an activity in support of the population health management subject area.

Note: *Super types are obtained from agency EA. Future volumes of the DRM will address the process of identifying super types from agency EAs.*

Exchange of Data

Data that is categorized around a particular business context can be exchanged in support of a business function or process. The DRM uses the information exchange package as a structure to enable the exchange of data:

Information Exchange Package: The information exchange package represents a set of data that is transmitted for a specific business purpose. It makes use of the ISO/IEC 11179 concept of Information Interchange.

The information exchange package is used to fulfill business requests that make use of agency business processes. In this scenario, the information exchange package provides data resulting from a business process that is engaged in supporting the population health management BRM sub-function. The actual content of the information exchange package is dependent upon the particular business process accessed. In this case, it could communicate information about immunization records and/or disease characteristics.

Structure of Data

Structured data has the standards and definitions necessary to describe the data that is associated with a business context. The **Data Element** concept is used to structure data within the DRM. To clarify the business context of a particular set of data, the subject areas and super types of the data set are supported by additional levels of detail described within the data element. A collective set of three layers, the data element enables a more accurate description of the business purpose of the data. It is consistent with the ISO/IEC 11179 standard and includes a **Data Object**, a **Data Property**, and the **Data Representation**. In practical terms, the data element provides a set of information that is used in a given business context.

Data Element: A data element is a representation of a data object, a data property, and a data representation. The data element defines a particular concept or item that is of interest within the super type.

Data Object: In describing the super type of immunization, it is necessary to more specifically define the particular concept or item that is of interest within the immunization super type. This item is called a data object, and, in this scenario, represents a vaccine. The vaccine represents a particular item of interest within the super type of immunization.

Data Property: The DRM uses a data property to distinguish or describe the actual vaccine. The data property represents the elements used to describe an object and can include characteristics such as type, weight, potency, etc. In this scenario, the data property is the name of the vaccine.

Data Representation: The DRM uses a data representation or value domain to represent the type of value that can be associated with the data element. Representation values can include integers, whole numbers, dollars, etc. In the case of vaccine, the value is plain text.

Security and Privacy

The successful categorization, exchange, and structure of data are dependent on the implementation of security regarding the data being exchanged. Security requirements must be considered at each level of the DRM and, in particular, regarding the exchange-of-data transaction. The DRM is designed to allow for the integration of existing federal information security and privacy policies within each of its elements. It provides for this integration through its common approach and use of standards. Exhibit F generally describes the relationships between the DRM and several sets of security/privacy policies and legislation.

Exhibit F: DRM - Policy and Legislation Relationships

Policy/Legislation	Linkage to DRM	Description
Federal Information Security Management Act (FISMA) (see Title III – Information Security)	Entire Model	FISMA applies to all levels of the DRM. FISMA requirements are applicable to the data of a classified or unclassified information exchange package.
National Institute of Standards and Technology (NIST) FIPS (See NIST FIPS 199)	Entire Model	FIPS standards can be applied to specific data elements and the information exchange packages used to exchange them.
E-Government Act of 2002 (see Title III, Section 208 – Privacy Provisions)	Entire Model	The DRM's approach to the categorization, structure, and exchange of data are consistent with the requirements of the E-Government Act privacy provisions.
OMB Circular A-11 (see Section 31-8)	Entire Model	Effective use of the DRM within an agency's EA ensures that security is applied correctly at the data element and information exchange package level.
NIST 800-60 (see Volume I)	Entire Model	The use of agency-common information and agency-specific information is supported through the business context, information exchange package, and data element layers of the DRM.

4

DATA REFERENCE MODEL

DRM

Volume I

Version 1.0

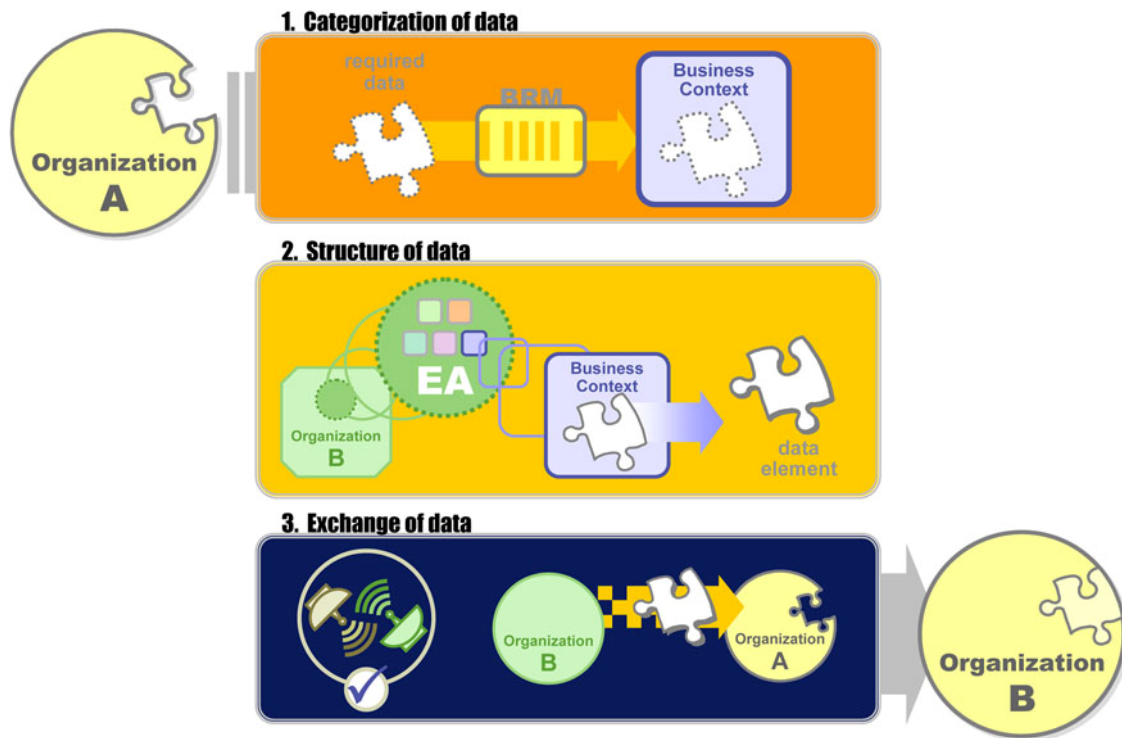
Use of the DRM

A conceptual process to identify data commonalities across agencies (using the DRM) helps illustrate the context in which agencies use the DRM. The process uses an agency’s EA in combination with the DRM’s common approach to the categorization, exchange, and structure of data to share information. This section offers a potential collaboration process and the results of a pilot conducted by the Department of the Interior (DOI) to demonstrate how the DRM might be used.

Collaboration Opportunities

Collaboration opportunities can be identified through the DRM's approach to the categorization, exchange, and structure of data. The process illustrated in Exhibit G lays out (at a conceptual level) the steps an agency might go through in its use of the DRM.

Exhibit G: DRM Collaboration Process



In this scenario:

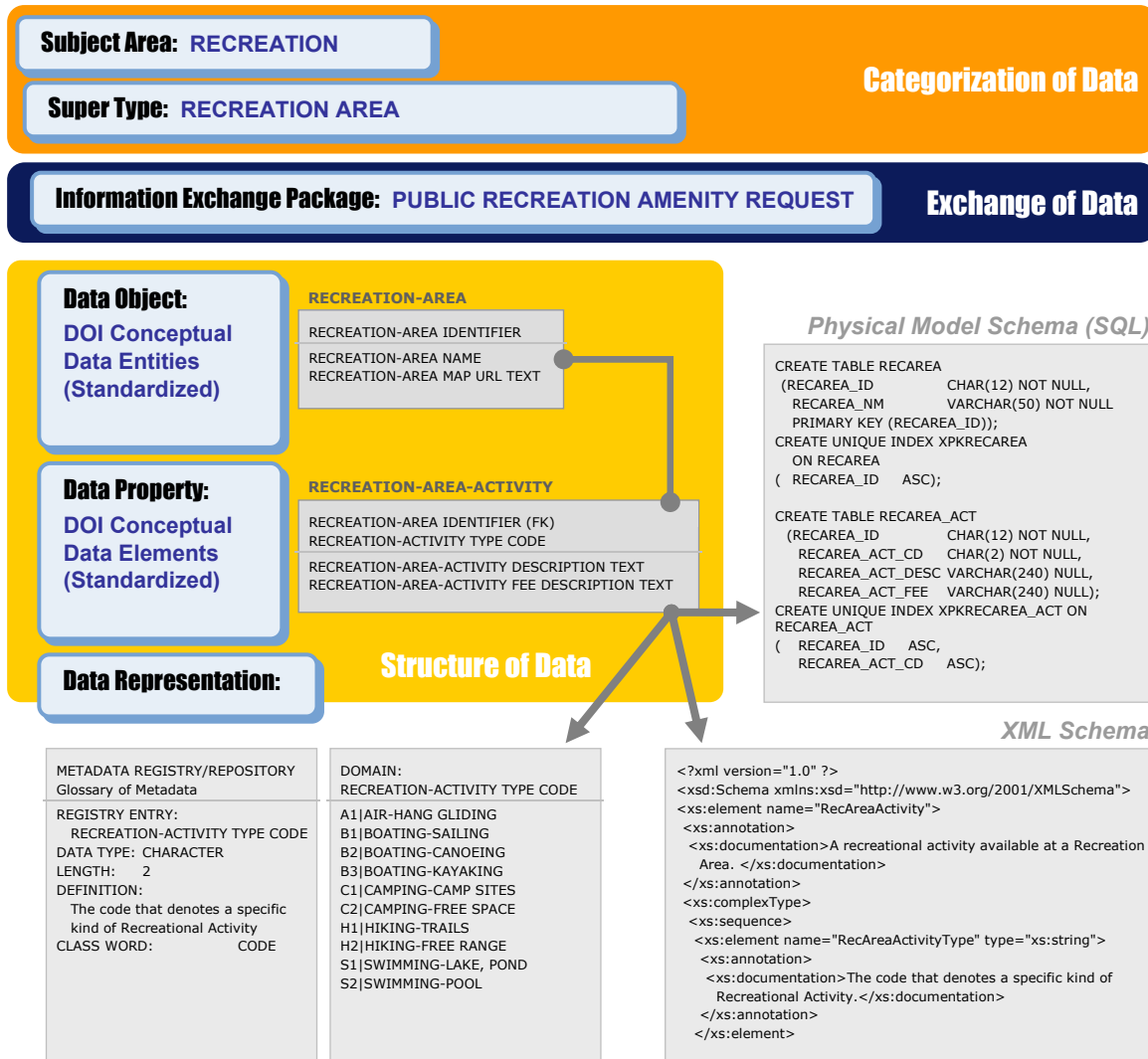
1. Organization B determines that it has the need for a particular set of data that might be available from Organization A.
2. Organization A uses the DRM to categorize its data (using the BRM) into a business context.
3. Organization B identifies Organization A's available data through its business context.
4. Organization A uses the DRM to publish the detailed structure of the actual data element (in support of the business context).
5. Organizations A and B determine if, in fact, the data produced by Organization A will meet the needs of Organization B.
6. Once Organizations A and B determine that the data can be re-used, the information exchange package is used to transmit the data.

Although the process illustrated in Exhibit G is simplistic in nature (and it will likely require a level of detail on the part of the organizations seeking to share data), it does provide a conceptual view to the steps an organization might take in order to use the DRM.

Department of the Interior (DOI) Pilot

The pilot conducted by DOI in Exhibit H provides a perspective on how the DRM might be used to improve the ability to share information and efficiently use IT investments. DOI's pilot uses the DRM to share information regarding its recreational amenities in a common approach that can be easily interpreted and employed by many users. Users who wish to take advantage of DOI recreational information need only understand the DRM to understand and use DOI data properly. Exhibit H illustrates the DOI pilot and the use of the DRM's common approach to the categorization, exchange, and structure of data. This diagram also describes the use of actual data models (schemas). These schemas represent a potential view of the implementation of an information exchange package. Future volumes of the DRM will define in more detail the relationship of schemas to the information exchange package.

Exhibit H: Use of the DRM



The DOI pilot demonstrates the potential outcomes provided by the DRM. The pilot’s use of the DRM is described in the following sections.

CATEGORIZATION OF DATA

DOI used the DRM to categorize data through the identification of activities performed within the recreational resource management and tourism sub-functions of the BRM. With the BRM categorization identified, DOI further identified a super type of “recreation area.” This super type is a more detailed categorization of the type of data consumed/produced through this business function.

EXCHANGE OF DATA

Defining data in common terms related to the request of a public amenity enables it to be shared with other users. By using the DRM's approach to the exchange of data, DOI identified a set of information (information exchange package) that directly supports the request for a recreational amenity. DOI used the DRM's categorization approach to relate the information exchange package to a particular business context, and then made it available as a re-usable set of data.

STRUCTURE OF DATA

After using the categorization approach to identify the subject area and super type, DOI used the DRM's common approach to identify the data elements. The data element includes the data object (names of the recreational areas), data property (the types of recreation activities aligned with specific recreational areas), and data representation (specific value of the data element).

Potential Outcomes of the DOI Pilot

INFORMATION SHARING

Information sharing related to recreation areas is facilitated because users understand the categorization, exchange, and structure of the data needed to satisfy their business needs. In the pilot, DOI creates information about its recreation areas and the many activities available within them. An agency that wants to make a recreation amenity request would, for example, look in the BRM for a sub-function that describes the activity it is seeking (recreational management and tourism). Once the agency knows the sub-function, it can use the Federal Enterprise Architecture Management System (FEAMS) to identify investments currently supported by DOI that provide recreational-amenity management capabilities. With the investments identified, the agency can work with DOI to determine whether the functions and data supported by the investment meet its needs. Once this is confirmed, the information exchange package would be used to actually transmit the request from the agency's systems to the DOI systems that manage recreational amenities. Once the DOI data is made available to multiple users, it increases the department's ability share information.

IMPROVED EFFECTIVENESS OF IT INVESTMENTS

The DOI pilot illustrates how an agency might improve the effectiveness of its IT investments by making the data produced by its investments available to others. The DRM's common approach to data categorization, exchange, and structure provides a mechanism whereby an agency does not need to create a new investment when the data it

requires is available from another source. Agencies engaged in the DOI pilot now have a common way to describe the business purpose of the data required by their uses and agencies participating in the DOI pilot can re-use the various existing IT investments to meet its business needs.

5

DATA REFERENCE MODEL

DRM

Volume I

Version 1.0

DRM Roadmap

Future volumes of the DRM will continue to address the standardization areas necessary to promote a common approach to the categorization, exchange, and structure of data. This is advanced through a focus on three areas:²

- **Information Management** - Use of data and information in support of business operations
- **Information Architecture** - Definition of the data linked to specific business operations (functions)
- **Information Exchange** - Standard, repeatable processes and technologies in support of exchanging data

Exhibit I illustrates the DRM roadmap in context of these focus areas.

² The focus areas are adapted from, *The Three Pillars, an Adaptation of Information Management and Data Quality*, by Bryan Aucoin, Panel 1, proceedings of the Eighth International Conference on Information Quality, (ICIQ-03)

Exhibit I: DRM Roadmap

Content		DRM Volumes			
		I	II	III	IV
Categorization of Data	Information Management				
	Business Context	■			
	Use of DRM in EA		■		
	Information Indexing		■		
	Security of Data			■	
Data Standardization Requirements				■	
Exchange of Data	Information Exchange				
	Core DRM Elements	■			
	Information Exchange Package	■			
	Federated Data Classifications		■		
	Data Patterns			■	
	Data Exchange Requirements			■	
Metadata Requirements				■	
Structure of Data	Information Architecture				
	Data Element Definition	■			
	Information Categories		■		
	Data Groups			■	
	Data Structures				■

Each future volume of the DRM will address different topics within the various focus areas. All topics addressed in future volumes of the DRM will continue to advance the common approach to the categorization, exchange, and structure of data. Each topic is described in Exhibit J.

Exhibit J: DRM Future Volume Topics

DRM Volume	Topic	Description
I	Business Context	Business purpose of data.
	Core DRM Elements	Approach to the categorization, exchange, and structure of data.
	Information Exchange Package	Common approach to the exchange of information between units of work. Units of work represent consumers and producers of data.
	Data Element Definition	Categorization and identification of the actual data used in support of a given business context.

DRM Volume	Topic	Description
II	Use of DRM in EA	Use of the DRM in investment and collaboration decisions.
	Information Indexing	Approach to the indexing of information. Cataloging of information ensures that information is available to its many consumers.
	Federated Data Classifications	Approach to identifying commonalities and opportunities for re-use (at the data level).
	Information Categories	Detailed definition and standards associated with the categorization and structure of data.
III	Security of Data	Information protection, assurance, and privacy.
	Data Patterns	Define data groups (tables, records, messages, text) and attributes that reflect business process needs.
	Data Exchange Requirements	Define data transformation patterns and key attributes that facilitate the sharing of information.
IV	Metadata Requirements	Define the data required to provide or support a specific community of interest or LoB.
	Data Structures	Detailed data design and format requirements.
	Data Standardization	Detailed data composition and requirements.

6

DATA REFERENCE MODEL

DRM

Volume I

Version 1.0

Glossary

Baseline Architecture

The set of products that portray the existing enterprise, the current business practices, and technical infrastructure (commonly referred to as the “As-Is” architecture)

BRM

Business Reference Model of the Federal Enterprise Architecture

Business Context

The purpose or use of the data (generally related to the BRM)

Business Functions

High-level set of business activities that are performed by the organization

Communities of Practice

Lines of business within the government that are dedicated to the support of business functions

Data Element

Physical description of the data used within an information exchange package

Data Management Strategy

Forthcoming document that will describe the role of data and data governance within an agency EA

Data Model

Representation of the information required to support the operation of any set of business processes and/or the systems used to automate them

Data Object

Basic definition of the data element

Data Property

Description of the data element in context of the data object

Data Representation

Describes how data is described within the property and object layers

DRM

Data Reference Model of the Federal Enterprise Architecture

Domain

High-level approach to the categorization of a set of data elements for purposes of organization and standardization

FEA

The Federal Enterprise Architecture, a set of reference models intended to support the use of agency EAs

FEA Business Integration Patterns

Conceptual approach to using the data from an agency EA in a common way to identify re-usable information technology investments

FISMA

Federal Information Security Management Act

Information Exchange Package

Set of data elements used to support the sharing of data within a particular business context

Information Flow

The process by which information exchange packages are shared

Metadata

Represents information about the data and could include value constraints, naming rules, etc.

PRM

Performance Reference Model of the Federal Enterprise Architecture

Relational Table

Set of elements within a database that organizes data in a meaningful way

Schema

The structure of a data set, database, information exchange package, etc.

SRM

Service Component Reference Model of the Federal Enterprise Architecture

Sub-Function

Detailed business activities performed in support of a particular business function

Subject Areas

Broad classification of data and super types related to a business context

Super Types

Generic groupings of data related to a specific subject area

TRM

Technical Reference Model of the Federal Enterprise Architecture

Target Architecture

The set of products that portrays the future or end-state enterprise, generally captured in the organization's strategic thinking and plans; commonly referred to as the "To-Be" architecture

Transition Plan

A document that defines the strategy for changing the enterprise from the current baseline to the target architecture; it schedules multiple, concurrent, interdependent activities, and incremental builds that will evolve the enterprise