

Step: 9

Conduct Data Standardization

Version 1.0, February 2005

1 Step Description/Objectives:

Step 9, *Conduct Data Standardization*, is intended to reduce the life cycle cost of data through data integration, standardization and the identification and use of authoritative data sources. Documenting data in a structured way will promote the understanding and meaning of data. A data standard provides the framework for how data will be formatted for implementation within systems and in the data exchanges between systems. This task relies on the identification of data by data stewards or communities of practice and on the functional community that will eventually implement and use standardized data.

2 Step Purpose:

Step 9 is performed for the following purposes:

- To identify priority data that are important to the Department especially data that are shared across systems and organizations.
- To enhance interoperability among DOI information systems by promoting data structures and data exchanges that are based on a common understanding of data.
- To identify and reuse quality data.
- To expose data standards to a broader DOI community and seek collaborative input from that community on how shared data is best defined and structured.
- To provide database schemas and data exchange standards that are based on standardized data.

Impact of Not Performing this Step

Step 9 is the primary output of the Data Resource Management and Standardization Program and the data product of the Enterprise Architecture. This step is designed to incrementally improve data sharing and interoperability within and without DOI. If Step 9 is not performed:

- Data exchange will continue to be problematic and costly
- Data will not be aligned with the enterprise architecture
- Data integrity, quality, and consistency will not improve

3 Tasks to Perform Step:

Step 9 starts with the following inputs:

Inputs:

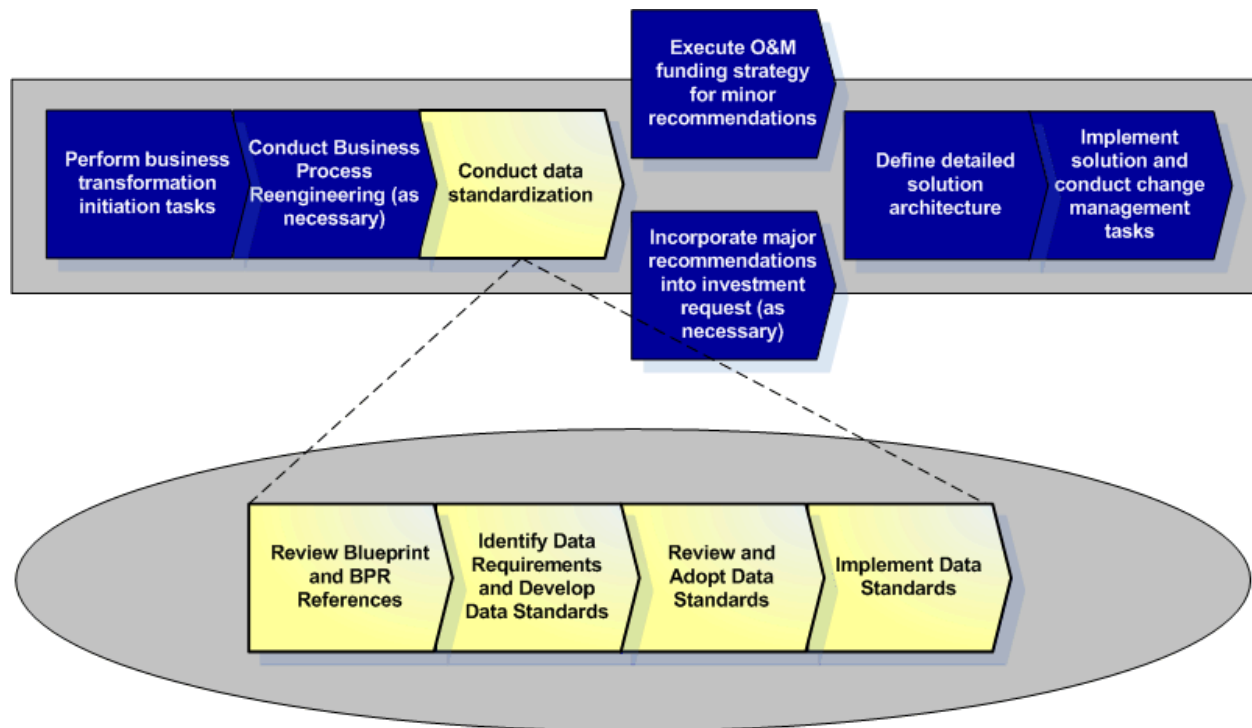
- Modernization Blueprint Data and Data Exchange Requirements
- BPR References
- Current Version of the Interior Data Reference Model

Step 9 Guidance: Conduct Data Standardization

- Other Information Requirements
- IDEF0 Activity Models

Detail:

Within this step, there are four major tasks:



Task 1: Review Blueprint and BPR References. The Modernization Blueprints provide a basis for developing and prioritizing data requirements. Within the Goals and Objectives and the Findings sections of the blueprint is information that relates to improvement opportunities in data and data exchange. Specific data requirements are developed and/or projects initiated for the discovery and documentation of data requirements that satisfy the findings of the blueprint. The resulting data requirements are formatted in the form of one or more logical data model diagrams using the DOI standard IDEF1X modeling method, and an accompanying metadata dictionary in the Popkin SA tool, or in some other compatible format.

Activities:

1. Create logical data model diagrams based on the data requirements in the Goals and Objectives and in the Findings sections of the Modernization Blueprints.

Task Outputs:

- Logical data model diagrams with metadata dictionary

Step 9 Guidance: Conduct Data Standardization

Task 2: Identify Data Requirements and Develop Data Standards. This task applies information engineering structure and provides an initial review of data that will be proposed as a DOI standard. There are two primary objectives involved with this task:

1. The development of data standard proposal packages based on the identification, collection, validation, design, and documentation of data requirements. During this task, the data requirements gathered from sources such as the Modernization Blueprint and reverse engineered systems are formatted and structured based on a prescribed template according to the draft Department Procedures for Data Standardization. A logical data model diagram and detailed information exchanges that are in harmony with the Interior Data Reference Model (IDRM) is a mandatory product of this task.
2. The initial functional coordination of new data standards, modification to existing data standards, archiving of existing data standards, and the preparation and submittal of a data standard submission package. The developers of a proposed DOI data standard are responsible for conducting a preliminary coordination among a community of practice, a business process, a functional community, with IS managers, and with the one or more DOI Data Stewards who have purview over the proposed data standards. This is an iterative process that may require frequent contact with stakeholders.

Such coordination typically centers on the logical data model(s) containing the proposed standards. Suggested changes and harmonization of similar concepts are typical. Many model and definitional changes may occur and changes to relationships within the logical data model are possible.

This step serves these purposes:

- It serves as a validation mechanism for the proposed data standards by having varying levels of expertise with similar information needs review the proposed data standard
- It provides a way to review the proposed data standard in advance of a formal review. This may reduce the rework and possible delays that could result had such coordination not taken place.
- The coordination ensures that the proposed data standards do not already exist and that they comply with the guidance set forth in this manual.
- The coordination also ensures that the logical data model associated with the proposed data standard is functionally integrated with the DOI DRM.

This process results in a functionally coordinated proposed data standard.

Activities:

1. Collect and identify data requirements
2. Validate data requirements
3. Identify existing data standards for reuse
4. Develop logical data models
5. Document data element metadata
6. Conduct Initial functional coordination
7. Prepare proposal package

Task Outputs:

- Data Standard Proposal package

Task 3: Review and Adopt Data Standards. This task addresses the specific subtasks involved with reviewing proposed data standards for completeness, semantic correctness, and functional fit within the DOI community. It culminates with a proposed data standard that is either accepted as a DOI standard, tabled, or rejected. The data standard proposal package is formally distributed throughout the DOI data community for comment. The comments are collected, reviewed, and a response is prepared. Agreed-upon changes are reflected in the proposal package model and metadata and a final resolution document is distributed. The DOI DRM is updated with the adopted data standards and a new release of the IDRM is prepared and issued.

Activities:

1. Perform a technical review of the proposal package to ensure compliance with technical and completeness requirements.
2. Conduct the formal review process.
3. Perform comment resolution.

Task Outputs:

- Adopted or rejected DOI Data Standards
- Possible changes to a proposal package for resubmission
- Comment Resolution Document

Task 4: Implement Data Standards. Once a data standard has been established, it can only provide tangible benefit if it is used throughout DOI systems and data exchanges. There are two primary objectives of this task:

Activities:

1. Integrate the data standard into the IDRM and disseminate knowledge about that data standard to the data community. This provides an updated version of the IDRM, with new or modified associations with other Enterprise Architecture products, such as mappings between entities and the BRM. This also ensures that the DOI data community is notified of a newly adopted or modified data standard.
2. Provide the data standard in the form of database schemas, xml standards, information exchange packages, and listings of authoritative sources.

Task Outputs:

- An extended/updated IDRM
- XML Tags and Schemas
- Information Exchange Packages
- Physical database schemas

4 Step Participants: (Use Bullets)

- DOI Data Architect: **Person**
- Principal Data Stewards: **Person**
- Bureau Data Architects: **Person**
- Business Data Stewards: **Person**
- Subject Matter Experts: **Person**
- Database Administrators: **Person**
- Data Reference Model Steering Group DRMSG: **Org**
- Functional Representatives: **Person**

5 Dependencies:

Data stewards must be identified and performing their role. A repository tool is required to manage and disseminate data. This step is based on the recommendations of a Modernization Blueprint that has been approved by the IRB.

6 Step Deliverables:

Logical Data Model: This work product is an artifact that documents data requirements in an IDEF1X or equivalent diagram with associated metadata and definitions. The DOI preferred format is a fully attributed IDEF1X logical data model. Data requirements are based on goals and objectives and findings of the Modernization Blueprint.

- **TEMPLATE :** Hyperlink to Template (**Deliverable Type: IDEF1X Model for Popkin SA or ERwin**)

Data Standard Proposal Package: This work product is a package of three primary products:

A fully attributed logical data model:

- **TEMPLATE :** Hyperlink to Template (**Deliverable Type: Adobe pdf Diagram or Popkin SA diagram**)

An excel spreadsheet containing the Target Information Exchange Matrix:

- **TEMPLATE:** Hyperlink to Template (**Deliverable Type: MS Excel**)

An excel spreadsheet containing Entity metadata:

- **TEMPLATE:** Hyperlink to Template (**Deliverable Type: MS Excel**)

An excel spreadsheet containing Data Element metadata:

- **TEMPLATE:** Hyperlink to Template (**Deliverable Type: MS Excel**)

An cover page containing submitter and point of contact information and principle steward information:

- **TEMPLATE:** Hyperlink to Template (**Deliverable Type: MS Word**)

Step 9 Guidance: Conduct Data Standardization

Comment Resolution Document. This document provides the tool for the documentation of the formal Data Standard Review process.

- **TEMPLATE :** Hyperlink to Template (**Deliverable Type: MS Word**)

Implemented Data Standards: This work product is the data standards that are implemented in the databases, file structures, data exchanges, and as authoritative sources of data. This work product also provides to the DOI community the notification and dissemination of knowledge about adopted data standards and associated authoritative sources where applicable. Tangible work products:

1. XML Tags: This product consists of data entity and data element names that have been formatted for inclusion into XML files and schemas. The XML tags are typically formatted as part of the standardization process and are subject to formal review as part of the data standard proposal package.
 - **TEMPLATE :** Hyperlink to Template (**Deliverable Type: MS Excel**)
2. XML Schemas: This product is a collection of XML Tags that are placed into a single XML Schema file for use in defining standard data exchanges.
 - **TEMPLATE :** Hyperlink to Template (**Deliverable Type: MS Word**)
3. Information Exchange Package: An Information Exchange Package represents a set of data that is transmitted for a specific business purpose. It is the actual data exchange instance (such as an XML instance) that delivers the payload or information. An Information Exchange Package typically contains a collection of artifacts that describe the structure and content of a data exchange including, for example, an XML Schema, purpose and scope statements for the exchange, timing, and security dependencies.
 - **TEMPLATE :** Hyperlink to Template (**Deliverable Type: n/a**)
4. Physical Database Schemas: The constructs for data storage and/or the encoding for data exchange that are implemented to process data standards.
 - **TEMPLATE :** Hyperlink to Template (**Deliverable Type: n/a**)

7 Communications Considerations:

Primary communication regarding data standardization occurs as part of the Data Resource Management Program itself. Communication within this step is primarily focused on two points:

- The communication that must occur before the final development of a Data Proposal Package takes place within the Data Steward community and often crosses functional and organizational domains. This is a crucial step to coordinate and resolve data issues to the greatest extent possible prior to submitting a proposal package for formal review.
- Communication that markets newly adopted data standards to the DOI data community. This is necessary to ensure that data standards are infused into systems and data exchanges.

8 References:

- Draft Department Procedures for Data Standardization
- Policy: Draft Department Manual for Data Resource Management