



**United States
Department of
Agriculture**

Service Center
Implementation
(SCI)

Interim Policy for Data Management

Prepared by
Data Management Technical Working Group

Abstract: Initial guidance for Business Process Reengineering (BPR) projects to coordinate data development and create data sharing opportunities amongst the partner Service Center agencies. Establishes roles and responsibilities for data management and defines the foundation for coordinating data development and sharing efforts.

Keywords: data management, policy, data sharing.

Published and distributed in June 1998

Introduction

(This introduction is not part of SCI Std #####-##-###, Interim Data Management Strategy.)

This policy establishes initial guidance for coordinating data development effort to facilitate the definition and development of data resources that improve communication about and shareability of data across the partner Service Center agencies.

The individuals who contributed to the development of this guide are:

Jim Schmidt, Leader

Tom Richards

Richard Verhoff

Dave Butler

Chuck Calcaterra

Lee Stieger

Dorn Egley

Jerry Kleiman

Mindy Gehrt

Norma Westbrook

Figure 1—Working group list

Table of contents

1. Overview 1
 1.1. Scope..... 1
 1.2. Purpose 2
 2. Data management policy..... 3
 2.1. Data quality..... 3
 2.2. Data sharing 3
 2.3. Database administration..... 4
 2.4. Information access 4
 3. Data management roles and responsibilities 5
 3.1. Management Review Board 5
 3.2. Information Resources Management (IRM) Group 5
 3.3. Executive Data Sponsor..... 5
 3.4. Data Steward..... 6
 3.5. Data Team..... 6
 4. Procedure for Maintaining the Interim Policy for Data Management 9
 Appendix A – Authorities 10
 Appendix B – References 11
 Appendix C – Definition of Terms 12

INTERIM POLICY FOR DATA MANAGEMENT

1. Overview

Data is an important asset which has value, and which must be managed effectively with the same diligence as other important agency assets. It is the policy of the Agencies to implement data management in ways that promote and enhance the mission of the Agencies through the effective acquisition, integration, use, and dissemination of data and metadata.

The intent of this policy is to establish initial guidance to the Business Process Reengineering projects for coordinating data development efforts and creating data that can be shared among the partner Service Center Agencies. The policy establishes important roles and responsibilities for data management and establishes the basic infrastructure for coordinating data development efforts.

The Service Center Data Team (Data Team) developed the material contained herein. This team is composed of representatives from the Service Center Partner Agencies: the Natural Resources Conservation Service (NRCS), the Farm Service Agency (FSA), and the three mission areas within Rural Development. Sponsorship and direction for the Data Team comes from the Executive Director of the National Food and Agriculture Council (FAC) of the U.S. Department of Agriculture (USDA) and the partner Service Center agencies Chief Information Officers (CIOs).

1.1. Scope

This interim policy pertains to all data associated with new or re-engineered systems using National Databases. It applies to all data created, collected, obtained from external sources, processed, disseminated, stored, or utilized from "National Databases", including Data Warehouses, by the Agencies both directly or indirectly (e.g. through partnerships or contracts). This policy also applies to contractors, consultants, partners, and universities providing development and maintenance services to the Service Center Partner Agencies.

This interim policy does not pertain to data created, collected, stored or utilized from "Local Databases" or legacy systems not being reengineered.

This policy pertaining to Data Management/Administration has been reviewed and does not conflict with any existing departmental policy.

1.2. Purpose

The purpose of this Interim Policy for Data Management for the Service Center Partner Agencies is to establish processes under which data (Tabular and Spatial) is developed and managed so as to protect the value of the data asset, ensure data quality, improve data organization, and enhance the ability to share and reuse data. This **interim** policy will be transferred to the Information Technology Services Branch (ITSB) data management organization when it becomes operational.

2. Data management policy

This policy addresses the management of data in a manner to ensure the quality of data, promote maximum utility and sharing of information resources, and to ensure that databases are managed effectively and efficiently.

2.1. Data quality

Data quality refers to the degree of correctness of the data. Data will be managed to improve, maximize, and protect its quality throughout its lifecycle. The following statements describe policy pertaining to data quality:

- The Agencies will establish data management authority and responsibility for each National Database by assigning an Executive Data Sponsor and a Data Steward for each business application where data is permanently stored. The Agencies will ensure that these roles and responsibilities are identified and described in project documentation.
- The Agencies will provide training on data management roles and responsibilities with the goal of implementing effective data management throughout each phase of the systems lifecycle.
- The Agencies will ensure that new or reengineered information systems using National Databases have effective data security plans, and clearly defined operational procedures.
- The Agencies will ensure that adequate metadata is collected and maintained for each National database.
- The Agencies will ensure that policies, standards, processes, and procedures regarding data and metadata quality are implemented for each new or reengineered information system.
- The Agencies will ensure that these policies are transferred to the ITSB data management organization when it becomes operational.

2.2. Data sharing

Data sharing refers to the practice of managing and accessing that information for which there is a common or shared use across the enterprise through a set of common databases and/or data files that are accessible to authorized users throughout the enterprise, and are shared across functional business areas and agencies.

All data assets will be organized, integrated, and managed to promote maximum utility and sharing of data and information resources as appropriate.

- Data assets will conform to the Service Center adopted standards for documentation and implementation.
- A centralized metadata repository will be established and maintained for use by the Agencies, and a common database management process will be established and used for documenting, creating and maintaining National Databases.
- All new or reengineered information systems and application projects will ensure that existing "Agencies" enterprise data assets are reused as appropriate before new data elements are created.
- All new or reengineered information systems and application projects will ensure that all data assets created will be fully integrated into the overall "Agencies" enterprise data architecture.

2.3. Database administration

Database administration involves functions related to the day-to-day monitoring and maintenance of databases. The Agencies are charged with ensuring that databases are managed effectively and efficiently. The following statements further describe policy pertaining to this function:

- The Agencies will provide for the efficient and coordinated operation of National Databases.
- The Agencies will ensure all National Databases have an effective security plan and disaster recovery plan approved by the appropriate system security manager.

2.4. Information access

Information access refers to the availability of and access to data, and descriptive information that enables the audience to transform the data into meaningful information.

- The Agencies will streamline information access by documenting available data assets, and by making access to descriptive information about data assets as effortless and economical as possible within the constraints of the Privacy and Freedom of Information Acts.

3. Data management roles and responsibilities

This section outlines the roles and responsibilities of the various parties involved in the management of business and metadata.

3.1. Management Review Board

The Management Review Board consists of the Chief Information Officers of the Partnering Agencies and Deputy Administrators for Management (DAMS) and Deputies for Programs. The Management Review Board will:

1. Approve Service Center Data Management Policies.
2. Provide coordination and means for implementing policies within agency organizations.

3.2. Information Resources Management (IRM) Group

The IRM Working Group consists of the Chief Information Officers of the "Agencies". The IRM Working Group will:

1. Recommend Service Center Data Management Policies.
2. Approve standards and Procedures.
3. Administer and ensure the enforcement of the Service Center Data Management Policies and standards within their respective agencies.
4. Approve waivers to Policy and Standards.

3.3. Executive Data Sponsor

The Executive Data Sponsor(s) is a business-area manager who is accountable for the collection, management, and use of data assets. The person has overall responsibility for the definition of the data and all issues that deal with data content. In some cases this may be a shared responsibility between several business-area managers from different agencies.

The Executive Data Sponsor(s) for a national database will:

1. Determine if data which the Agency plans to collect has already been collected by the Agency, or whether cooperative efforts are possible to obtain the data from other existing sources (See Authorities: Executive Order 12906).
2. Coordinate funding for data collection, storage and maintenance; and for software application development, support and maintenance.

3. Promulgate and implement the policies and procedures necessary for ongoing management of the physical data content, the standards for the acquisition and certification of data, the policies for the collection and usage of metadata, and the procedures for the protection of the physical data assets.
4. Designate a Data Steward, and other critical data management roles and responsibilities as appropriate.
5. Authorize the release of data and application software to internal and external customers.

3.4. Data Steward

The Data Steward is a business area expert who is assigned responsibility for the content of the database.

The Data Steward for national databases will:

1. Act as the designated authority and point of contact for all business-area decisions concerning the database. Responsibilities include obtaining the needs/requirements from the users, and coordinating with the Data Team on metadata and other data management issues.
2. Establish and maintain business rules and consistent definitions for data elements, identify data domains and relationships, establish data quality and certification standards associated with the contents of the database, and recommend availability, security and access authority for the data
3. Ensure that metadata is collected, approved, and certified for release according to the adopted industry, Federal and USDA metadata and data management standards.
4. Ensure the validity, accuracy and completeness of the physical data and supporting metadata; certify that data meets quality standards; and certify that data is ready for release for internal and/or public use.
5. Provide training within the Data Steward's business area on data management roles and responsibilities.

3.5. Data Team

The **Data Team** is an interim interagency team established by the Service Center Agencies and BPR Office. The team is responsible for implementing data management principles, policies, standards, and for establishing the overall data architecture. This Data Team will turn over its responsibilities to the data management organization within the Information Technology Services Branch (ITSB), when it becomes operational.

The Data Team will:

1. Establish a core Data Architecture, to include:
 - Maintaining the Enterprise Data Model for all new/reengineered applications
 - Coordinating the collection of metadata for spatial and tabular data
 - Maintaining the business rules supporting the Enterprise Data Model
 - Modeling the physical layout and location of National data used by the Agencies.
2. Provide strategic planning for the acquisition and use of data assets to meet program goals.
3. Perform Data Administration functions for applications, to include:
 - Resolution of conflicting data names, establishing common lookup tables, setting common domains for sharable data elements, and establishing unique keys and identifiers
 - Coordinating data administration/management training
 - Coordinating the release of data to the public
4. Provide management of the ongoing Data Management process, to include:
 - Reviewing, at a minimum annually, and recommending modifications to the data management policies, as appropriate
 - Developing standards, procedures, and shared utilities and tools for data management
 - Maintaining a shared Data Management Handbook that documents common data management standards and procedures
 - Coordinating implementation of a metadata repository, CASE and modeling tools, and other supporting data management software
5. Maintain a shared, central metadata repository for use by the Agencies to store and provide access to metadata, to include:
 - Establishing standard and sharable data elements to promote data reuse
 - Making metadata accessible to system developers and other users

- Maintaining a central source of data name abbreviations, and common acronyms
6. Provide a consolidated voice to the Department and to other government committees on data management issues.
 7. Coordinate and support Database functions, to include:
 - Coordination and review of logical and physical models
 - Providing performance measures and metrics
 - Implementing and coordinating security rules
 - Monitoring databases for performance
 - Coordinating and supporting reconstruction of databases to accommodate new information or to facilitate changes in the physical deployment of the database.
 8. Coordinate and support establishment of Data Warehouses to include:
 - Coordinating data that crosses business areas
 - Coordinating the development and maintenance of the warehouse data model

4. Procedure for Maintaining the Interim Policy for Data Management

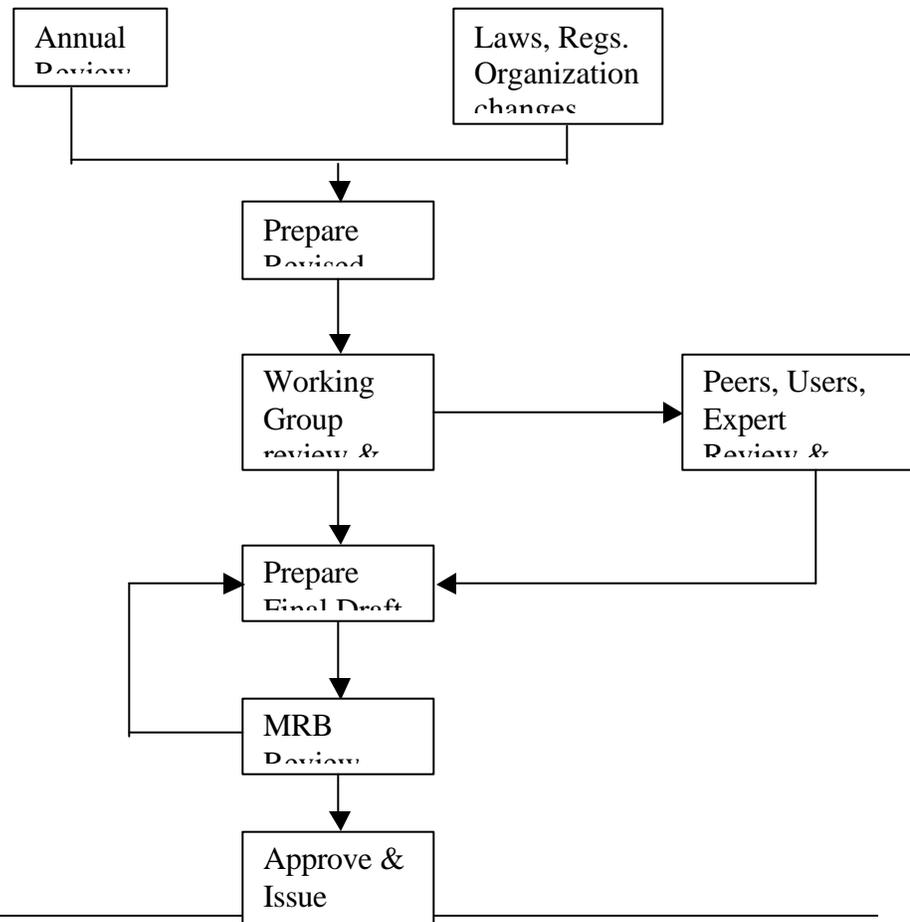
The updating of the Interim Policy for Data Management will be prompted by two events:

- First, a regular annual review to ensure the policies are accurate, easy to read, and understandable.
- Secondly, on an as needed basis due to changes in laws, regulations or the organizational structure of the department and or partnering agencies.

The Data Management Policy Working Group Leader will send a reminder to all Agency representatives on the Working Group to poll their agencies for changes to the policy annually.

Anyone that becomes aware of new or changes in laws, regulations or Organizational changes, that may necessitate modification to the current Data Management Policy; should notify the Data Management Policy Working Group Leader.

The leader will draft a revised policy and actively seek input from the working group, users, peers and experts. A final draft will be prepared for submission to the Management Review Board for approval.



Appendix A – Authorities

1. Computer Security Act of 1987 (Public Law 100-235).
2. Departmental Regulation (DR) 3400-4, *Departmental Data Administration Program*.
3. Executive Order 12906, *Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure*, April 11, 1994.
4. Office of Management and Budget (OMB) Circular A-130, *Management of Federal Information Resources*.
5. OMB Circular A-16, *Coordination of Surveying, Mapping and Related Spatial Data Activities*, revised October, 1990.
6. The Paperwork Reduction Act of 1980 (Public Law 96-511) as amended by the Paperwork Reduction Reauthorization Act of 1986.

Appendix B – References

1. Charter of the Service Center Data Team
2. Interagency Project Manual (IP 3430-2) “InfoShare, FISVIS, MAP Data Element Standardization Manual”, InfoShare Data Team.
3. FGDC, Content Standards for Digital Geospatial Metadata, June 8, 1994.
4. FIPS 173, Spatial Data Transfer Standard.

Appendix C – Definition of Terms

The following definitions apply to terms used in this document:

Business Rule: A statement that defines or constrains some aspect of the business as it is implemented in the data model. Data-related business rules are statements, phrased in absolute terms, about data (i.e. A telephone number must have 10 digits), and about relationships between data (i.e. If a phone number is entered, the phone type must also be entered.)

Data: A discrete fact or value. Data is the raw material, which through its use and interpretation can provide valuable information. Data is the content of databases or data files.

Data Administration: The technical function of acquiring, defining, certifying, organizing, protecting, and delivering data and the metadata that describes it.

Data Administrator: The person who defines, organizes, manages, controls, protects and standardizes data models, data elements and other metadata.

Data Architecture: An orderly arrangement of data resources to achieve: 1) a common understanding of available data resources; 2) a planned approach to data acquisition, storage and retrieval to achieve a high degree of responsiveness to user demands; and 3) a high degree of data sharing and data mobility to reduce program delivery costs.

Data Dictionary: A database about data and database structures. A catalog of all data elements, containing their names, structures, and information about their usage. Normally, data dictionaries are designed to store a limited set of available metadata, concentrating on the information relating to the data elements in the databases, files and programs of implemented systems.

Data Management: The managerial function of taking responsibility for data and the processes that support it. It focuses the strategic planning and operational data functions (technical planning, data administration, database administration, data warehouse administration) on meeting program delivery goals.

Data Model: A pictorial view of data, groupings of data, relationships between data groupings, and the organization of data groupings by dependencies. A “logical” data model is a view that does not depend on the characteristics of the computerized system or of the physical storage. A “physical” data model typically refines the logical model by adding the constraints incumbent to the database system or physical storage method.

Data Repository: A database of information describing the characteristics (metadata) of data. Typically, the repository also stores a broad range of descriptive information, including business rules, data models, and process models that help to elaborate on the usage of data in various systems. Repositories can also store metadata for the purpose of identifying and retrieving sets of actual data. Metadata that describes a map is an example.

Data Warehouse: An informational database, or collection of databases, used to store shareable data to support business and managerial decision processes. The warehouse is usually created through data extracts from operational databases. The warehouse adheres to a single SCIT enterprise data model to ensure consistency of decision-support data across the SCIT enterprise. The warehouse typically allows users to tap into an organization's vast store of operational data to track and respond to business trends, and to facilitate forecasting and planning efforts.

Database: A collection of related data organized to serve one or more applications. In the broader sense, it describes any organized collection of data regardless of the physical storage method.

Database Administration: Encompasses the day-to-day technical functions that support ongoing business operations. It includes the collecting, defining, certifying, organizing, protecting, and delivery of both data and metadata (data about data).

Domain: A listing of all the possible valid values that can be stored in a data element.

SCIT Enterprise Data Model: An overall pictorial view of the many applications and databases making up the participating agencies' combined data assets. The intent is to manage the overall data assets to achieve optimal integration, sharing, access, and utilization of technology resources and infrastructure.

Geospatial Data: Information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth. This information may be derived from sources such as remote sensing, mapping and surveying technologies. It includes both attribute (text) as well as spatial (map) data.

Information: A commodity derived from data through analysis, or by the orderly presentation of data for human interpretation.

Local Database: A database, which is used by an agency organization (e.g. Region, State, Center, Institute, etc.) but does not meet the definition of a

National Database. Examples include: local spreadsheets, report extracts, and PC-based databases

Metadata: Data about data. Metadata describes how, when, and by whom a particular set of data was collected, and how the data is formatted. Metadata includes attributes such as data name, length, domain of valid values, and definition. Metadata can also identify and describe a set of data or a complex data type such as a map, photograph, spatial data set, etc.

National Database: A permanent database which:

1. has international, national, USDA, or agency-wide application, or
2. is included in a standard software suite, or
3. contains data that is used/shared directly in making national program decisions, or
4. is used/shared in multiple offices, states, or other internal/external organizations.