



United States Department of Agriculture

Service Center Data Management Plan

Implementing Sound Practices for Managing Service Center Data Resources

April 13, 1998

Foreword

This document describes a vision and plan for implementing Data Management processes within the service centers and the agency business areas supporting them. The Service Center Data Team (SCDT) developed the material contained herein. This team is composed of representatives from the Natural Resources Conservation Service (NRCS), the Farm Service Agency (FSA), and the Rural Development (RD) information technology communities.

Sponsorship and direction for the SCDT comes from the Executive Director of the National Food and Agriculture Council (FAC) of the U.S. Department of Agriculture (USDA) and the service center partner agency Chief Information Officers (CIOs).

Executive Summary

The Service Center initiative is bringing about major changes in the delivery and stewardship of data. With a goal to improve delivery of information and services to USDA customers, this initiative is greatly expanding the need to share data, changing lines of authority, pushing advances in systems and telecommunications technology, and has the potential to significantly effect how agency missions are carried out. Data management processes are changing in response to these developments. Effective management of data resources is a key component of Service Centers implementation, and for protecting the Service Center partner agencies' massive investment in electronic data.

Mission Drives Data Management Objectives

The goals and objectives of "Data Management" are a direct reflection of the missions of the agencies. The specific business goals and objectives that implement agency missions drive the requirements for data. Data and information requirements, in turn, help drive the technical architecture for hardware, software, and telecommunications.

Data Management is the business of managing a resource. The service center agencies (FSA, NRCS, and RD) have billions of dollars invested in electronic data and in the infrastructure of software, hardware and staff necessary to collect and maintain this data.

Data Management Vision

A vision for data management has emerged from efforts to reengineer the delivery of products and services to USDA customers. The vision simply states that Data Management is:

**A PARTNERSHIP between BUSINESS and
INFORMATION TECHNOLOGY
to deliver
the RIGHT INFORMATION
to the RIGHT PEOPLE
at the RIGHT PLACE,
at the RIGHT TIME
and at the RIGHT COST.**

The focus is on the business and the customer. This focus sets the mark against which to evaluate which data management processes and practices are really important for accomplishing the mission.

Data Management Infrastructure

Data Management is a partnership between business areas and information technology staffs. The stewardship, or management, of data is a business area function that manages a resource needed to accomplish the mission.

Data Management staffs in the Information Technology organizations have as their business function the process of managing data.

There must be a partnership between the business areas, who have responsibility for the data, and the IT staffs who have responsibility for the process of managing data. This partnership is key to maintaining the focus on the data management vision.

Goals for Data Management

Four goals have been identified as a means of achieving the Data Management Vision. They are stated as follows:

To institute good Data Management practices within the business disciplines and the information technology community to:

1. Enhance Customer Service by Making Data Available When and Where It Is Needed,
2. Manage and Protect Agency Data Assets,
3. Improve The Effectiveness Of Program Delivery, and
4. Improve the Return on Investment in Data Collection and Maintenance.

A number of objectives are identified in the plan for each of these goals. The objectives will be addressed by work groups within the IT community with help and review by business areas. A number of these objectives offer technical support to business areas in achieving their data resource management requirements.

Implementation

Where possible, specific actions and timeframes are specified in the plan. In other cases, accomplishing the objective relies on timetables set by business areas and other IT functions. Much of the work that directly related to implementing Service Centers would be coordinated through the Business Integration Center. Early agreement on a basic system development lifecycle among the agencies will help provide a common framework and common terms for the data management process.

Table of Contents

LIST OF EXHIBITS	1
LIST OF ACRONYMS	2
1. THE BUSINESS OF MANAGING DATA.....	4
1.1 MANAGING DATA RESOURCES TO MEET AGENCY MISSIONS.....	4
1.2 DATA ENDURES WHILE THE BUSINESS CHANGES.....	5
1.3 A PARTNERSHIP BETWEEN BUSINESS LEADERS AND INFORMATION TECHNOLOGY	5
1.4 DATA IS A VALUABLE AGENCY ASSET	6
1.5 AGENCIES ARE ORGANIZED TO MEET THEIR DIFFERING MISSIONS.....	6
1.6 SERVICE CENTER DATA TEAM.....	7
2. REENGINEERING THE BUSINESS.....	9
2.1 FACTORS PRODUCING CHANGE	9
2.2 THE FOCAL POINT FOR REENGINEERING.....	11
2.3 BUSINESS PROCESS REENGINEERING (BPR)	12
3. THE VISION, GOALS AND OBJECTIVES FOR SERVICE CENTER DATA MANAGEMENT.....	14
3.1 THE SERVICE CENTER DATA MANAGEMENT VISION.....	14
3.2 GOALS AND OBJECTIVES FOR DATA MANAGEMENT	15
3.2.1 <i>Goal 1 - Enhance Customer Service By Making Data Available When and Where it is Needed</i> 16	
3.2.1.1 Develop a Data Architecture to Describe the Physical Placement of Data across the Network	16
3.2.1.2 Make Data Available Directly To The Customer and Support ‘One-Stop Shopping’	18
3.2.1.2 Develop Data Warehousing Strategies	19
3.2.2 <i>Goal 2 - Manage and Protect Agency Data Assets,</i>	<i>21</i>
3.2.2.1 Implement Data Administration/Management Policies	22
3.2.2.2 Establish Data Administration Operations and Standards.....	23
3.2.2.3 Implement a Centralized Metadata Repository	24
3.2.2.4 Implement Database Administration Processes and Standards	26
3.2.2.5 Implement a Data Stewardship Program.....	27
3.2.2.6 Implement Standard Data Management Tool Configurations.....	28
3.2.2.7 Develop and Implement a Data Security Program.....	29
3.2.3 <i>Goal 3 - Improve the Effectiveness of Program Delivery.....</i>	<i>30</i>
3.2.3.1 Establish Common Databases to Support the Service Centers.....	30
3.2.3.2 Develop and Maintain an Enterprise Data Model.....	33
3.2.3.3 Support the Acquisition, Integration, and Management of Spatial Data.....	34
3.2.3.4 Provide Ongoing Support for the BPR Planning and Implementation Process	36
3.2.4 <i>Goal 4 - Improve the Return on Investment in Data Collection</i>	<i>37</i>
3.2.4.1 Promote Maximum Sharability and Re-Use of Data Resources	37
4 IMPLEMENTATION APPROACH.....	39
4.1 INFORMATION TECHNOLOGY ARCHITECTURE PROVIDES AN ENGINEERING PERSPECTIVE	39
4.2 IMPLEMENTING A DATA MANAGEMENT INFRASTRUCTURE	39
4.3 FRAMEWORK FOR SERVICE CENTER DATA MANAGEMENT	40
4.3.1 <i>Data Program Implementation and Management.....</i>	<i>41</i>
4.3.2 <i>Enterprise Data Administration</i>	<i>42</i>
4.3.3 <i>Data Stewardship.....</i>	<i>43</i>
4.3.4 <i>Application Development</i>	<i>44</i>
4.3.5 <i>Database Administration.....</i>	<i>44</i>

4.3.6 *Data Asset Use*.....45

5. APPENDIX46

APPENDIX 5.1 AUTHORITIES AND REFERENCES.....47

APPENDIX 5.2 DATA MANAGEMENT ROLES AND RESPONSIBILITIES OF THE BUSINESS DISCIPLINES .133

APPENDIX 5.3 FEDERAL DIRECTIVES ESTABLISH REQUIREMENTS FOR DATA MANAGEMENT175

APPENDIX 5.4 GLOSSARY177

APPENDIX 5.5 ZACHMAN FRAMEWORK181

LIST OF EXHIBITS

EXHIBIT 4.1: FRAMEWORK FOR ENTERPRISE-WIDE DATA MANAGEMENT

EXHIBIT 5-1: ENTERPRISE INFORMATION SYSTEM COMPONENTS

EXHIBIT 5-2: THE ZACHMAN FRAMEWORK

LIST OF ACRONYMS

BIC	Business Integration Center
BPR	Business Process Reengineering
CASE	Computer Aided Software Engineering
CCC	Commodity Credit Corporation
CCE	Common Computing Environment
CD	Compact Disc
CDRS	Central Data Repository System
CIO	Chief Information Officers
CIP	Customer Information Profile
COTS	Commercial Off-the-Shelf
DR	Departmental Regulation
DBMS	Database Management System
EIS	Executive Information System
FAC	Food and Agricultural Council
FGDC	Federal Geographic Data Committee
FSA	Farm Service Agency
FY	Fiscal Year
GILS	Government Information Locator System
GIS	Geographic Information Systems
GPRA	Government Performance and Results Act
I/O	Input/Output
IT	Information Technology
ITA	Information Technology Architecture
ITSB	Information Technology Services Bureau
LAN/WAN	Local Area Network/Wide Area Network
L/W/V	LAN/WAN/Voice
MRB	Management Review Board
NRCS	Natural Resources Conservation Service

NSDI	National Spatial Data Infrastructure
OIP	Office Information Profile
OMB	Office of Management and Budget
PC	Personal Computer
RBS	Rural Business-Cooperative Service
RD	Rural Development
RUH	Rural Housing Service
RUS	Rural Utility Service
SCDT	Service Center Data Team
SCIT	Service Center Implementation Team
SSB	Support Service Bureau
TWG	Technical Working Groups
USDA	United States Department of Agriculture

1. The Business of Managing Data

The value of any information system comes from the services it helps to deliver and the information that is created from data. The Service Center partner agencies (NRCS, FSA, RD) have billions of dollars invested in electronic data. They have similar investments in computer software to enter and maintain this data, plus a large number of people dedicated to data-related tasks. It is this data that allows the agencies to service a large clientele, to keep a historical perspective of agency activities, and to support managerial and technical decision-making. Data is an asset, and its creation and maintenance are day-to-day activities as common to the corporate world as to government operations.

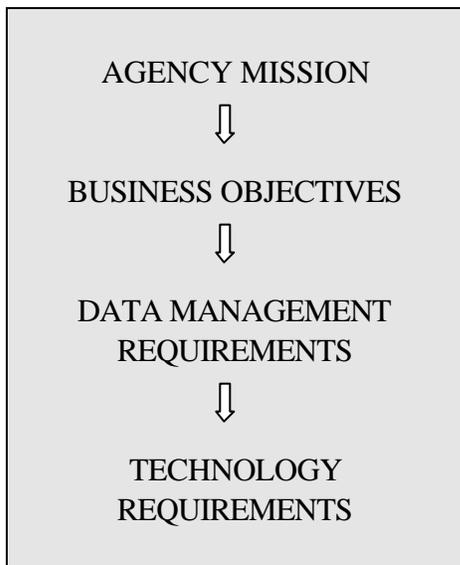
1.1 Managing Data Resources to Meet Agency Missions

Managing data is a business function. Business-area managers identify the data they need to accomplish their mission. They are likewise responsible for the collection, maintenance, delivery and utilization of this data in a cost-effective, business-like manner. They are the stewards of these data assets for as long as the data has value to the business.

Data Management
is a
managerial function
which
focuses data assets
on meeting
program delivery goals.

“Data Management” is the discipline of managing data. It encompasses a number of technical functions, standards, procedures and tools that a manager employs in managing business data.

*Data Management is a Reflection
Of the Business Mission*



The Information Technology (IT) community has long had an interest and responsibility to ensure that good data management practices are used throughout the organization. Business managers have often looked to IT to provide the Data Management functions they need.

The goals and objectives of “Data Management” are a direct reflection of the missions of the agencies. The specific business goals and objectives that implement the agency missions also drive the requirements for data and the information derived from that data.

Data and information requirements, in turn, help

drive the technical architecture for hardware, software, and telecommunications which also help managers accomplish their missions.

The Service Center initiative is bringing about major changes in the delivery and stewardship of data. With a goal to improve the delivery of information and services to USDA customers, this initiative is greatly expanding the sharing of data, changing lines of authority, pushing further advances in systems and telecommunications technology, and has the potential to significantly effect how agency missions are carried out. Data Management processes are changing in response to these developments. This strategic plan helps map the implementation path for these changes.

1.2 Data Endures while the Business Changes

One constant, in the midst of change, is the basic store of data built up over many years of operations. The storage media may change, but the hard facts endure and must be protected and used.

The need for sound data management exists throughout the lifecycle of the data, starting with the conception of a new information service, extending through the development of procedures and software, reaching full force with the collection of the first data, and continuing undiminished until the data is finally retired.

1.3 A Partnership Between Business Leaders and Information Technology

The difference between the function of managing a resource (i.e. data) and the function of providing the technical framework for data management is a key distinction.

The stewardship, or management, of data is a business area function. It is managing the resources needed to accomplish the mission.

IT Data Management, as an organization within the Information Technology community, has a business function and a support function. It is a business in that it has its own goals and objectives that cut across organizational and mission-area boundaries. It is, at the same time, a support function since it often neither creates, owns, nor uses the data it manages.

Data Management is
a **Partnership**
between
Business Leaders and
Information Technology.

Its success depends on strong partnerships with the business community.

Good Management of Data
Requires Changes in the
Agency Culture

The IT Data Management community is charged with providing the framework for good data management throughout the agencies, including the standards,

processes, tools, and training needed by employees to accomplish their business objectives. The IT Data Management organization crosses agency and organizational lines to bring continuity, cost effectiveness, and the best technical environment for using and maintaining agency data assets. The future of the data management function is in a partnership between business leaders and IT to use resources as effectively as possible while responding to a fast changing environment.

There must be strong communication conduits between Information Technology and the business disciplines, with a focus on meeting both long-term business strategies and short-term project objectives. Changes in the agency culture will need to take place so that stewardship of data and a commitment to the protection and maintenance of data assets are accepted business functions.

1.4 Data is a Valuable Agency Asset

Data, and information derived from it, are strategic assets that strengthen the ability of the agencies to provide high quality service to customers.

Rural Development's automated accounting system manages 900,000 loan accounts with an outstanding loan balance of over \$80 billion. Obligations for new loans in Fiscal Year (FY) 1997 totaled over \$5 billion.

In FY 1997, Farm Service Agency systems' processed approximately \$12.7 billion worth of Commodity Credit Corporation (CCC) checks, collected over \$7 billion and managed a direct Farm Loan portfolio of approximately \$10 billion.

The Natural Resource Conservation Service investment in natural resource data is estimated to be over \$8 billion. Over \$120 million is spent each year in collecting and managing natural resource data. Most of this investment is in the National Soil Survey and National Resources Inventory. However, investments in plant and climatic information are also significant. NRCS is accelerating the digitizing of soil survey information, with a completion target in 2002. The digitizing cost for the period 1995-1999 is projected to be \$30 million.

Orthophotography is recognized as an important asset for Service Center missions. The total USDA investment in Orthophotography since the start of the program in 1995 projected through Fiscal Year 1999 is \$59 million.

Agency investments in data are high, making the value of these assets equally high. It is important to keep in mind that data is perishable, and can be quickly devalued or destroyed by lack of attention and proper stewardship or by loss of confidence in the data's accuracy.

1.5 Agencies are organized to meet their differing Missions

The missions, and therefore the data management requirements, of the Agriculture Agencies are vastly different. The Farm Service Agency (FSA) and the Rural Development Agency (RD) have a largely centralized system development environment. With a massive amount of client-specific data, privacy and security are a paramount concern. The handling of billions of dollars in loan and commodity payments makes tight accounting controls mandatory. While its systems development and database management are centralized, FSA requires that much of its data be updated from local levels, including the dynamic updating of common land units and related tabular data. Immediate access to client and program information at the local level is also needed.

NRCS organizes around its mission to distribute information to the public, and to provide services in a mobile environment (data anywhere, anytime). NRCS has traditionally had a highly distributed application development environment. Applications and data are developed at the Agency's Information Technology Center, in state and regional offices, by universities, and by private consultants.

The Data management solution must encompass all of these development environments. There is not a single Data Management solution to cover all agencies and all situations. The IT community is currently restructuring itself through the "Administrative Convergence" initiative into a more cohesive organization. The objective is to provide continuity in the delivery of information services over multiple agencies. But, the resulting centralized IT Data Management Division will still have to address the multiple business cultures and demands of its agencies.

1.6 Service Center Data Team

A Service Center Data Team (SCDT) has been established as a part of the Service Center Implementation Team (SCIT) to develop and implement a common data management infrastructure for the partner agencies. The Team is tasked to work in concert with agency business areas, FSA, RD, and NRCS software development centers, and the Business Integration Center to achieve Business Process Reengineering goals.

The Service Center Data Team is made up of data management professionals from Service Center partner agencies – FSA, RD, and NRCS, supported by contract staff. It is intended to be a partnership of the three agencies in reengineering the way data is managed across the information system development centers.

The Service Center Agency CIOs and the FAC Executive Officer act as a governing body for the Data Team to ensure that data management initiatives support the agencies businesses and missions.

The IT staffs of the three Service Center partner agencies are being converged into a single organization, the Support Service Bureau (SSB). A major function of the Information

Technology Service Branch of SSB will be data management. The Data Team will turn over its responsibilities to this organization once it is established under Administrative Convergence. The policies, practices and principles being established by the SCDT will provide a solid framework for the SSB.

2. Reengineering the Business

Constant change is a factor in the businesses of the USDA Agencies and in the delivery of Information Technology. Many of the current activities taking place in the Data Management discipline are aimed at proactively addressing changes in basic mission-area responsibilities, and in preserving valuable data resources as technology and business requirements are in transition.

Data Management is aimed at effectively creating, using, and administering a very valuable USDA asset. Electronic data has a large cost for its creation, storage, and distribution but it creates great value when used to support agency missions.

Therefore, it is important to plan for its protection and promote its usage during periods of organizational and business change.

2.1 Factors Producing Change

Agency business areas and the Information Technology community are going through periods of intense change. Raised expectations of USDA customers, agency staff, and federal oversight groups have fueled much of this change. The proliferation of home computers is making both customers and USDA employees more aware of the potential for computer technology in the workplace. Business Process Reengineering efforts currently underway are a reflection of many of these broadening expectations.

Many government programs are large, and it often takes forces built-up over time to cause the abrupt changes being seen today. Factors currently serving as instigators for change include:

- Business Changes – Changes include program cuts, changing focuses for conservation, reorganization of service delivery points and mechanisms, Congressional scrutiny and mandates for change, declining budgets and personnel ceilings, etc. The 1996 Farm Bill established new directions, and the next Farm Bill will likely do the same. Advances in natural resource science and agriculture technology are being translated into computerized decision-support tools requiring new types, and vast amounts, of data. The dramatic reorganization of the agencies has changed lines of authority, and regrouped functions.
- Changing Clientele – Changes include the growth of corporate farming, growth of private agricultural consultants, increased emphasis by State Governments on agriculture and environmental concerns, a global perspective on agricultural products, and a more highly-educated customer base, etc.
- Customer Expectations – Customers are demanding services that are faster, more accurate, and easier to access. They want business processes that are less complicated and more

responsive. Customer service is being measured against private sector competitive standards. There is growing emphasis on taking the business to the customer and more customers have computer technology in their businesses and homes.

- Changing Technologies - Changes include precision farming and other advanced technologies, delivery of services and information through the Internet, faster personal computers to handle geospatial data and complex resource models.
- Greater Trust is being placed on Computerized Data - Fewer field service people, and higher turnover in personnel are forcing more of the corporate memory to be housed in computer files. There is a greater need to trust data, and a greater need for data to be self-explanatory for easier interpretation. It is frustrating for customers when data, such as acreage amounts, change without explanation.
- Increased Mobility of Data – A proliferation of smaller computers brings a corresponding high demand to move data between computers, and to move the computer to where the work is actually being done. The workforce is going mobile, and the data must follow.
- Data Security is a growing Issue – Aside from the problems caused by computer hackers, the growing reliance on computerized data has a direct correlation to the health and profitability of agricultural businesses. Competitive business information getting into the wrong hands can have adverse impacts, as can getting the wrong information.
- Mainstreaming of Hardware and Software – There is growing reliance on commercially-available hardware and software to not only reduce systems costs, but to be more compatible with the computers used by USDA customers.
- A Growing Respect for the Value of Governmental Data Stores – USDA managers are recognizing the large investments that have been made in data collection, and are taking better care of these valuable assets. USDA customers are also becoming aware of the vast amounts of useful information to be found in USDA databases.
- Individual and Business concerns for Data Privacy – As larger amounts of data are stored in electronic form, the public is increasingly concerned that certain portions of data specific to them be kept confidential. Systems must respect and protect confidential data.

2.2 The Focal Point for Reengineering

The focal point for reengineering is the “Service Center” initiative. Service Center’s are sites throughout the country which provide services directly to USDA customers. A keystone of this initiative is the establishment of multi-agency offices at the county level where a customer can go to obtain all types of USDA services. The Secretary of Agriculture has listed four primary objectives for establishing the Service Centers.

The Secretary’s Service Center Objectives and how they will be achieved are detailed in the following:

1. Provide One-Stop Service by making Data available through a Single Source

To the external customer, this means getting needed information from one office, or from one on-line computer source, rather than having to collect bits and pieces from many locations. For internal customers, such as software developers, this means having a central, authoritative reference about the data and Metadata on which information systems can be constructed. A Central Data Repository System (CDRS) that captures, standardizes and stores metadata (data about data) used in the various Service Center databases is a means to gain these benefits.

2. Provide Quality Customer Service to the point of exceeding Customer Expectations

A first step is to recognize the current and future customers for USDA data. Incremental steps can be made to provide accurate, consistent data that directly supports internal USDA users and external customers. The data must be made easily accessible in the required formats. Methods for efficiently transmitting data to the customer must be explored. Establishing a single integrated source for natural resource data is one example. Close cooperation with state and local governments will help increase the flow of information to customers, and provide outreach opportunities to under-served customers.

3. Reduce Operating Costs through Resource Sharing

Cost savings can be achieved in several ways: 1) by reducing the cost of gathering and maintaining data, 2) by sharing and using data repeatedly, and 3) by using the data to more easily accomplish a task.

Sharing data between agencies and between computer systems to identify incidents of benefit fraud and abuse can reduce the time and, therefore, cost needed to investigate these incidents. Eliminating the manual reconciliation of inconsistent data held in multiple databases will also reduce cost. Some of the first data that must be re-engineered for sharing and access are common names and addresses, common land units, common office information, and common benefit eligibility information.

While shared data will improve internal USDA operations; it can also produce real cost savings for USDA's customers. If customers can get information on properties owned in multiple states and counties by visiting only their local service center, then they are getting better service from USDA. They are saving the cost currently incurred to travel to multiple counties and agency offices to transact business.

4. Establish Partnerships with Stakeholders, Agencies and Customers

There is a need to create ongoing, long-term partnerships among agency data management staffs, software developers, business areas, other Federal agencies, state and local governments, partner organizations, and end-users to set goals for collecting, managing, and sharing data to ensure that the right information will be available in the future, and that it can be accessed in a timely, business-like manner.

2.3 Business Process Reengineering (BPR)

Reengineering the basic delivery of mission services is key means for reaching the Secretary's goals for Service Centers. To meet these expectations, the National Food and Agriculture Council (FAC) was charged with implementation responsibility for the Service Center initiative.

The FAC began a formal Business Process Reengineering (BPR) study in December 1996 to establish more streamlined and more customer-focused business practices. Initiatives coming out of this BPR study, and the corresponding Data Management functions, include:

BPR Initiatives

- Customer Interface
 - Effective customer outreach targeting under-served and rural communities.
 - Timely and effective program outreach by employees, enabling service delivery and information provided directly to customers at their home or work site.
- Program Delivery
 - Comprehensive customer and program information integrated across agency and program lines.
 - Improved access to information provided to service center employees, customers, USDA partners, and the public.
- Geospatial Systems
 - Common definition of land unit boundaries.
 - GIS tools for acquiring, managing and using spatial data.

- Administrative Management
 - Streamlined administrative processes for travel, hiring, fleet management, and directives.

The BPR study identified 17 projects to be undertaken to improve the delivery of services. Of these 17 projects, 14 have a requirement for managing data and information resources. The challenge in managing the BPR data is to build a unified approach that addresses technical issues and supports business requirements. The foundation for this approach is built upon the following:

- Establish data policies and standards,
- Produce and maintain the enterprise data model,
- Standardize data tools,
- Establish data warehouse and data mart strategies, and
- Establish the data architecture.

3. The Vision, Goals and Objectives for Service Center Data Management

3.1 The Service Center Data Management Vision

Data assets must be managed and used in order to gain the greatest return on the investment made. It is essential to know where one wants to be in the future. A clear vision for data management has emerged from ongoing efforts to reengineer the delivery of products and services to USDA customers. A number of factors will continue to shape this vision, including the Secretary's goals for Service Centers, further refinement of mission-area BPR requirements, and outside federal directives.

THE DATA MANAGEMENT VISION

A PARTNERSHIP
between
BUSINESS and
INFORMATION TECHNOLOGY
to deliver
the RIGHT INFORMATION
to the RIGHT PEOPLE
at the RIGHT PLACE,
at the RIGHT TIME
and at the RIGHT COST.

A vision is a far-off look at an ideal situation.

The Data Management "VISION" must be translated into specific goals that will make the vision a reality. Goals are steps that can be prioritized, the means to attain them planned, resources acquired, and progress measured.

These goals apply equally to the current Business Process Reengineering efforts and to the legacy functions of the agencies that must continue to be supported.

The following statement contains four goals for Data Management:

GOALS FOR DATA MANAGEMENT

To institute good Data Management practices within the business disciplines and the information technology community so as to:

- 1. Enhance Customer Service by Making Data Available When and Where It Is Needed,**
- 2. Manage and Protect Agency Data Assets,**
- 3. Improve The Effectiveness Of Program Delivery, and**
- 4. Improve the Return on Investment in Data Collection.**

Data Management functions must address a very complex business environment:

- the BPR projects,
- the ongoing program work of the individual agencies,
- the handling of legacy systems,
- rapid changes in technology and customer needs,
- cost reduction initiatives,
- organizational changes,
- computer mobility,
- large data volumes and complex data issues associated with geospatial data management,
- the Web and other new data access technologies, and
- constant assaults on data integrity.

A number of strategy initiatives to begin handling this complex environment are described below. These strategies will be implemented through both short and long-term projects and activities.

3.2 Goals and Objectives for Data Management

The four major goals for Data Management are translated into a number of short and long-term objectives, as follows:

3.2.1 Goal 1 - Enhance Customer Service By Making Data Available When and Where it is Needed

3.2.1.1 Develop a Data Architecture to Describe the Physical Placement of Data across the Network

- **Business Objective**

To maximize the usage of network and computer resources across organizational levels and locations of the Service Center agencies, and to forecast the impacts of proposed new systems on these resources.

- **Background**

This objective will 1) define the physical architecture of the data, i.e. the location of data across the network. 2) describe what data and databases are located on a central server, decentralized servers, and local clients. 3) make use of metrics to assess movement of data through the USDA network to maximize the functionality and responsiveness of software systems. 4) account for where, and how often, data is updated or refreshed and 5) promote a high level of data sharing and data mobility to reduce program delivery costs.

The placement of data will be dictated as much by business needs as by technical issues. Data used for national-level reporting, for example, may need to be stored and managed centrally. Data integrity and data sensitivity may likewise be determining factors.

Two underlying efforts upon which the Data Architecture is structured are the installation of a national telecommunications network (LAN/WAN/Voice) and the acquisition of computer systems for Service Centers and supporting offices. Remote access to data and the movement of data between computers will provide local offices with many new capabilities. The Data Management function is supporting forecasting of capacity needs for the data networks, and is developing an overall Data Architecture for the placement of databases and other data stores at locations that help ensure rapid information retrieval.

The Common Computing Environment (CCE) project is tasked with setting requirements for the hardware and software tools to be installed in the Service Centers. The Data Management function includes specifying the requirements for data base management systems, and determining data storage media, capacity, and tools needed to use and maintain special data types such as photographs, document images, and geospatial data.

- **Actions**

A Technical Working Group (TWG) has been established to plan the enterprise data configuration and to develop a data architecture that will:

- Distribute data across servers in one or more locations to achieve optimal system response time.
- Establish an overall mapping of data interrelationships and integration points.
- Maintain an overall plan for the ‘seamless’ movement and sharing of data.
- Formulate a framework for the integrated storage and movement of spatial and tabular data.

Phase 1 (FY 1998)

- Develop initial requirements for data distribution plans for initial BPR projects.
- Develop initial data architecture for BPR pilot implementation.
- Establish requirements for county and mainframe operations not included in the established BPR projects.

Phase 2 (FY 1999)

- Refine data configuration plan from results of BPR pilots.
- Feed requirements to CCE to guide implementation requirements.

Phase 3 (FY 2000 – 2003)

- Implement complete data configurations for major Agency and Service Center databases.
- Provide training on data architecture.

3.2.1.2 Make Data Available Directly To The Customer and Support ‘One-Stop Shopping’

- **Business Objective**

Support BPR and other agency efforts to 1) establish the means for customers to download data directly to personal or business computers to support various agricultural modeling and planning tools; 2) establish the framework for making data available to a customer about their loans and other financial dealings with USDA; 3) establish processes for publicizing the availability of agency data, including use of Web technology; and 4) support Government initiatives, such as the Government Information Locator System (GILS), to advertise existing data stores available for customer use.

- **Background**

This objective embodies the concept of making current and accurate information about USDA programs accessible to employees, customers and the public through communications technology. It includes the use of Web technology and other means to advertise the availability of data that the public may be interested in using.

The project includes the packaging of operational data on financial transactions and natural resources for convenience downloading by customers. The project also includes the organizing of data resources throughout the enterprise to support customer service at any Service Center a client may contact.

- **Actions**

Phase 1 (FY 1998)

- Data Management tools selected for pilot sites.
- EIS and other Query tools selected for pilot sites.

Phase 2 (FY 1999)

- Program area databases fielded at pilots.
- EIS fielded at pilots.
- Data Management, EIS, and other tools selected for CCE

Phase 3 (FY 2000 – 2003)

- Program area databases fielded nationally.
- EIS fielded nationally.

3.2.1.3 Develop Data Warehousing Strategies

- **Business Objective**

This task will perform user requirement studies to determine the business needs at all levels of the organization that can be filled by a data warehouse. Objectives for data warehouses will be mapped to the enterprise and agency missions and goals they support. Technical requirements for warehouse systems to meet identified needs will be developed. Data migration and data collection strategies for populating the selected data warehouse(s) will be developed with the business areas.

- **Background**

Operational data warehouses and data marts provide the basis for decision support systems and for both internal and public access to information. They can support both executive decision-making and normal agency operations, and be utilized at both national and local levels. Warehouses can contain raw operational data as well as refined, summarized data.

A key difference between data warehouses and regular operational databases lies in the migration, cleansing and structuring of the data to be housed in a warehouse or mart. Data is organized to allow searching, correlation of related data, summarization, and downloading of specific portions of a dataset within established privacy and security guidelines. A data migration strategy, or more likely strategies, will be put in place to move identified operating data to data warehouse(s) or mart(s).

Data warehouses can be constructed to produce national-level, consolidated reports for Congress and other high-level executives. These reports (frequently produced on an ad hoc basis) reinforce the need for centralized access to data pertaining to all parts of the organization.

- **Actions**

A Data Management Technical Working Group has been identified to perform studies of business needs that can be filled by a data warehouse. Technical requirements will be developed for the identified warehouse systems needs. Data migration and data collection strategies for populating the selected data warehouse(s) will be developed.

The Technical Working Group (TWG) will provide assistance to business areas to:

- Establish and manage data warehouses to support management information requests.
- Establish data clearinghouses for packaging and distributing data to internal and external customers.

- Migrate data from disparate systems into a cohesive data structure
- Advertise the availability of data through data registries, search engines, and published descriptions.
- Perform normal data and database administration functions for warehouses and clearinghouses.

The Data Management Technical Working Group will support business areas in the following actions:

Phase 1 (FY 1998)

- Begin development of the business case for data warehouses and a high-level architecture for their implementation.
- Select BPR project(s) as pilot warehouse projects.

Phase 2 (FY 1999)

- Migration, cleansing and structuring of the data for the initial data warehouses and marts.
- Implement Initial data warehouse and mart pilots.
- Provide training on data warehouse and mart implementation, use and, management.

Phase 3 (FY 2000 – 2003)

- Mature data warehouses and marts in operation.
- World Wide Web access to do business with, and receive information from, the service center at home.
- Provide training on data warehouse and mart implementation, use and, management.

3.2.2 Goal 2 - Manage and Protect Agency Data Assets

‘Data stores are a valuable, but perishable, commodity.’

Data Management, the managing of data resources, is a key component for improving work processes and delivering accurate and timely land, customer and program information.

Data is an asset with considerable value that can be easily devalued or destroyed by lack of attention and proper stewardship. The data management function specifically addresses the usage and protection of this asset. Managing data, like the management of any government asset, deals with:

- Describing the asset.
- Accounting for it.
- Establishing stewardship and responsibility for its management.
- Protecting and assuring its integrity ,and
- Providing for its most effective use.

The need for sound data management exists throughout the data’s lifecycle, starting with the conception of a new information service, extending through the development of procedures and software, reaching full force with the collection of the first data, and continuing until the data is finally retired.

3.2.2.1 Implement Data Administration/Management Policies

- **Business Objective**

Establish policies that implement common data management processes and standards for the Service Center partner agencies.

- **Background**

Policies are necessary for implementation of effective data management. These policies will be established prior to BPR application development and communicated to project managers for their use in developing BPR applications. Project deliverables will be evaluated for compliance with data policies through procedures implemented in the Business Integration Center.

- **Actions**

A Technical Working Group has been established to develop and implement a common set of data administration policies. Actions include:

Phase 1 (FY 1998)

- Implement common data management policies.
- Initial BPR projects reviewed and conformed to data policies.
- Provide information to program managers, system developers and data stewards on data management policies.

Phase 2 (FY 1999)

- Common data management policies evolve to meet changing requirements and technology improvements.
- Provide information to program managers, system developers and data stewards on data management policies.

Phase 3 (FY 2000 – 2003)

- Maintain data management policies.
- Provide information to program managers, system developers and data stewards on data management policies.

3.2.2.2 Establish Data Administration Operations and Standards

- **Business Objective**

Establish an operational data administration process and standards for Service Center application development.

Establish standards for data management deliverables produced in all phases of the system development life cycle.

- **Background**

Standards for documenting metadata are an initial priority. Proposed standards will build upon those previously developed by the Service Center Data Team. A Data Management Handbook will be made available to business disciplines and software developers.

- **Actions**

A Technical Working Group has been established to develop and implement a data administration process and standards. The process will provide on-going planning, management, and quality assurance activities. Standards will be used in agency software development centers and the Business Integration Center for systems documentation. Actions include:

Phase 1 (FY 1998)

- Develop and implement data administration processes for BPR and other agency application development.
- Establish a process for capturing and utilizing business rules as a vital link between the business specialists and software developers.
- Provide data administration services to BPR projects.
- Develop and implement data administration standards and procedures.
- Provide initial training to software developers and data stewards on data administration standards and procedures.

Phase 2 (FY 1999)

- Develop, implement, and manage a data administration change control function.
- Provide data administration services to BPR and agency projects.
- Provide training to software developers and data stewards on data administration standards and procedures.

Phase 3 (FY 2000 – 2003)

- Provide data administration services to BPR and agency projects.

- Provide training to software developers and data stewards on data administration standards and procedures.
- Manage the change control process.

3.2.2.3 Implement a Centralized Metadata Repository

- **Business Objective**

Establish and maintain a central metadata repository that fully describes partner agency data assets. The repository is used by system developers, data stewards, and the public to locate data sources, and to access specific metadata about data objects (data definition, data models, ownership, privacy, availability, etc.). The repository is a tool for coordinating the reuse and sharing of existing data, and establishes a basis for data warehousing and management information initiatives.

- **Background**

This task will include an evaluation, selection, and implementation of metadata administration tools (metadata, repository, CASE, etc.) needed to implement repository-based data administration. Processes will be established to populate the repository and for managing enterprise and project metadata.

- **Actions**

A Technical Working Group has been established to develop requirements, evaluate, select and implement a metadata administration tool configuration. Actions include:

Phase 1 (FY 1998)

- Develop a concept of operations that describes repository administration and metadata management processes.
- Select and acquire, as needed, central metadata repository tools and integrated CASE tools for development centers.

Phase 2 (FY 1999)

- Implement central metadata repository software and integrate software development centers' CASE tools.
- Provide training in CASE tool interfaces with the repository.
- Provide assistance and support to system developers and data stewards in use the of metadata management tools.
- Maintain metadata tool configuration.

Phase 3 (FY 2000 – 2003)

- Tools evaluated to determine changing availability and technology, particularly as needed to meet changing business requirements.
- Provide training in CASE tool interfaces with repository.

3.2.2.4 Implement Database Administration Processes and Standards

- **Business Objective**

This task includes identifying issues and concerns regarding administration of mainframe and decentralized mid-range databases. Functions such as database server installation and configuration, software upgrades, regular database operations, disaster recovery, and data security are topics to address.

- **Background**

Areas of concern include: installation and upgrading of DBMS operating software; management of data distribution including replication services; performance monitoring and resolution; data backup/recovery and archiving; implementation of user access security; implementation of database structure changes across the organization, mainframe and server middleware support, and routine maintenance. It also includes the implementation of data-related business rules and security rules, and overall protection of data assets.

- **Actions**

A Technical Working Group has been established to develop database administration processes and standards which will be used across the partner agencies and development sites. Actions include:

Phase 1 (FY 1998)

- Develop initial standards to implement in BPR pilot operations.
- Provide training for database administration.

Phase 2 (FY 1999)

- Develop and Implement database administration standards and procedures.
- Provide training for database administration.

Phase 3 (FY 2000 – 2003)

- Continue to provide training for database administration.

3.2.2.5 Implement a Data Stewardship Program

- **Business Objective**

The primary objective of the data stewardship program is to establish accountability for data assets, and to provide the guidelines and training needed to institute a good data stewardship process across the organization.

Implement strategic data management planning focused on how data is acquired, stored, and used to ensure that it is employed in the most efficient manner to support the mission.

- **Background**

The Data Steward role is critical for instilling good data asset management practices within the business disciplines and IT community. This is an important role in the partnership between the business area, which owns the data, and the information technology community which supports the management of data.

Although the Data Steward role is primarily a business-area function, the Data Stewardship program is also a partnership between business and IT to establish responsibility for the identification, specification, and standardization of data assets within a specific business area; for achieving data integrity; and for managing data assets to improve reusability, accessibility, and maintain quality .

- **Actions**

Phase 1 (FY 1998)

- Develop a Data Stewardship plan and get authority from management to implement the plan.
- Identify data stewards and sponsors for mission critical databases.
- Establish roles, responsibilities, and performance measurements.
- Provide training to data stewards on roles and responsibilities.

Phase 2 (FY 1999)

- Provide training to data stewards on roles and responsibilities.
- Establish feedback mechanisms to strengthen the business/IT partnership.
- Provide support to data stewards in their role and responsibilities.

Phase 3 (FY 2000 – 2003)

- Provide maintenance and support for data stewardship program.

3.2.2.6 Implement Standard Data Management Tool Configurations

- **Business Objective**

Support the procurement and implementation of enterprise-wide data management tools, including Database Management Systems (DBMS), server-based Geographic Information System (GIS), desk-top GIS, and Data Warehousing tools, and other tools as identified.

- **Background**

The selection, implementation, support, and management of a common suite of data management tools is necessary to provide orderly and effective management for Service Center data. A plethora of data management tools are currently being supported by Service Center partner agencies. Efficiencies can be gained in implementing a common suite of tools and concentrating expertise on management and use of these tools. The procurement of data management tool configurations will be guided by the Common Computing Environment (CCE) project.

- **Actions**

A Technical Working Group has been established to develop requirements for data management tools and to support the implementation of these tools.

Phase 1 (FY 1998)

- Develop requirements for data management tools to support BPR pilots.
- Support the evaluation, selection and implementation of data management tool suites for development centers and BPR pilots.
- Provide training in use of tools.

Phase 2 (FY 1999)

- Develop requirements and select data warehouse tools for initial pilot efforts.
- Finalize requirements for data management tools as a result of BPR pilot testing.
- Tools market evaluated to determine changing availability as needed to meet changing requirements.
- Provide training in use of tools.

Phase 3 (FY 2000 – 2003)

- Implement enterprise data management tool suites within CCE.
- Provide training in use of tools.

3.2.2.7 Develop and Implement a Data Security Program

- **Business Objective**

Implement procedures to support security and Privacy Act provisions.

- **Background**

The issue of data privacy and confidentiality is of growing concern in the public's mind, particularly with the explosion of Web-based access, and the establishment of data warehouses. The Privacy Act of 1974 and the computer security Act of 1987 set guidelines, but greater efforts are likely needed with the subsequent vast changes in information technology.

- **Actions**

Phase 1 (FY 1998)

- Develop data security and data privacy guidelines for pilot sites.

Phase 2 (FY 1999)

- Develop, in coordination with agency security officers, a common set of data security procedures and guidelines for Service Centers.

3.2.3 Goal 3 - Improve the Effectiveness of Program Delivery

3.2.3.1 Establish Common Databases to Support the Service Centers

Service Center data is classified as common, shared, or unique. Common data is jointly owned, used and managed by the Partner agencies. Shared data is data that is the primary responsibility of one agency. It is managed by that agency but shared by other partners. Unique data is used only for the mission of an agency. The emphasis is on promoting the cross-agency and cross-program usage of common and shared data.

The BPR process has identified four categories of common data that will be addressed immediately, 1) customer data, 2) land unit data, 3) administrative data and 4) spatial data (discussed as a separate topic).

3.2.3.1.1 Customer Data

- **Business Objective**

To share a common database of USDA customer information between the three Service Center agencies.

- **Background**

There are two perspectives for customer data – local and national. Local customer data (i.e. client name, address, telephone number) is used for managing customer files and mailing lists in field offices and other service delivery points. A pilot project under the name “Customer Information Profile (CIP)” was conducted jointly between FSA, NRCS, and RD in FY96/97. The pilot is currently inactive.

There is currently a high demand for national customer data to support FSA and RD loan and payment systems.

- **Actions**

Establishing a customer database is a priority BPR project. Active participation on the team implementing this project is proposed, to include developing the data model, data business rules, and associated metadata.

Phase 1 (FY 1998)

- Logical data model and related metadata established.
- Database tools selected.

Phase 2 (FY 1999)

- Customer database fielded at pilots.

Phase 3 (FY 2000 – 2003)

- Customer database fielded nationally.

3.2.3.1.2 Land Unit Data

- **Business Objective**

Provide a common definition and concept of a “land unit” among the partner agencies to form a framework for linking tabular data to the spatial dimension, i.e. tying existing data to a specific location.

- **Background**

The Service Center partner agencies have long needed a system to share and coordinate information between various USDA programs. It has been very difficult to define a term and boundary delineation for a common land unit that covers all land uses and earth covers, and yet maintains some recognition and understanding by internal and external customers. Work has been done in the past on this issue, and basic definitions have been developed. These definitions have not yet been integrated into agency operations.

- **Actions**

Support business areas in establishing standards for land units, and the means to uniquely identify a land unit for data systems. Develop the means to integrate spatial land unit data with the tabular data that describes it. Develop specifications for Database Management Systems that incorporate the ability to interface a database directly with GIS tools, and provide for the management of spatial data, descriptive metadata, and tabular information for land units within a database environment.

Phase 1 (FY 1998-1999)

- Support establishment of Common Land Unit standards
- Establish DBMS requirements related to maintaining Land Unit data.
- Integrate Land Unit concepts into requirements for GIS tools.

3.2.3.1.3 Administrative Databases

- **Business Objective**

To share common databases related to offices, human resources, property inventory, and telecommunication management.

- **Background**

Several BPR-related projects are currently underway in this area. A major pilot project for a reengineered Human Resources system is currently nearing implementation. A system for identifying agency offices and their locations is currently being developed in the form of the “Office Information Profile (OIP)” system. Work is also in the planning stages to maintain Service Center inventories of computer and communication equipment acquired through the LAN/WAN/Voice and Common Computing Environment projects. Telecommunication forecasting and load planning is a related effort.

It is important that Departmental Administration, the Office of the Chief Financial Officer, and the Office of the Chief Information Officer be included in determining the definition of administrative data.

- **Actions**

Provide direct support to the Human Resources project in analyzing and maintaining the data model, and supporting integration with other agency systems. Provide project management and project development resources to complete the office database. Provide analysis and system integration support for inventory and telecommunication forecasting systems. Work with Departmental Administration, the Office of the Chief Financial Officer, and the Office of the Chief Information Officer on the collection of metadata and the defining of administrative data.

3.2.3.2 Develop and Maintain an Enterprise Data Model

- **Business Objective**

Establish, populate and maintain a management view of the data resources that exist throughout the agencies. Work in partnership with business areas to maximize data sharing, implement the best access to data, and provide the most cost-effective means for storing and transmitting data.

- **Background**

To achieve Service Center goals, data held throughout the partner agencies must be managed as an enterprise resource. This enterprise view of data is represented through an enterprise data model. The enterprise data model represents the basic data classes and shows their relationships throughout all levels of the partner agencies. It includes centralized and decentralized systems. It presents a global view across locations and levels of agency operations for a total, integrated enterprise view. It becomes the focal point upon which to impose the discipline of data management and the basis from which application development can be coordinated. It is a master plan for the sharing and reuse of data. The enterprise data model will continue to evolve in response to business needs.

- **Actions**

Phase 1 (FY 1998)

- Establish a data administration process with the enterprise data model as its cornerstone.

Phase 2 (FY 1999)

- Increase efforts to share and reuse data.
- Enterprise data model evolves to meet changing requirements.

Phase 3 (FY 2000 – 2003)

- Enterprise data model continues to evolve to meet changing requirements.

3.2.3.3 Support the Acquisition, Integration, and Management of Spatial Data

- **Business Objective**

Support BPR efforts to define a nationally consistent set of core cartographic and business geospatial data themes that provide a foundation upon which to base business applications. Develop standards for these themes. Support business-areas in the establishment of standards for identifying basic land units. Establish the means to relate new and existing tabular data to the spatial feature it describes and maintain those relationships across databases, applications, and data tools.

- **Background**

A good deal of information used in Service Center offices is geographic in nature in that it relates to a place on the ground. Current geographic, or geospatial, information processing in offices is map-based and very labor intensive. A cornerstone of reengineering Service Center business processes is implementing digital geospatial information processing using Geographic Information System (GIS) technologies.

Service Centers will have a recurring need for several themes or layers of basic geospatial data, including both natural resource and business data themes. These themes and accompanying data can be managed as an enterprise resource that is shared and administered collectively by Service Center Partners.

Data themes that have been identified as important to Service Center offices are digital orthoimagery (replaces the current aerial photography), elevation, transportation, hydrography (surface water features), governmental units, and cadasteral (Public Land Survey and large, publicly administered parcels). BPR projects will likely identify additional themes, and opportunities to relate themes to standard tabular data from such applications as production adjustment, and compliance, etc.

Metadata (data describing a spatial dataset) must be extensive and complete to ensure the full use of this resource for both internal agency operations and for distribution of this data to outside customers.

- **Actions**

A BPR project has been identified to develop a standard for geospatial data themes; to define a basic nationally-consistent set of core cartographic and business geospatial data that provides a foundation upon which to base business applications; to support the

development of standards for these themes; and to coordinate the implementation of government and industry standards for the accompanying metadata. Data Management will support the geospatial BPR team activities and coordinate the development and implementation of common geospatial data standards.

Specific actions and schedules are provided in the geospatial BPR plan.

3.2.3.4 Provide Ongoing Support for the BPR Planning and Implementation Process

- **Business Objective**

Bring an enterprise perspective on data issues to high-level BPR planning and management functions. Establish the cornerstones for interagency data sharing. Provide strategic data planning for proposed BPR initiatives. Ensure that planned BPR projects will integrate into a consistent Enterprise Data Architecture.

- **Background**

Reengineering the basic delivery of mission services is a key means for reaching the Secretary's goals for Service Centers. The high-level planning being done for the reengineering initiatives requires the inclusion of a data management perspective. The Service Center Data Team was an integral part of the early requirements gathering and planning activities for the current BPR efforts.

- **Actions**

Through membership in the Service Center Implementation Team and as a component of the Business Integration Center, provide overall enterprise planning support to the reengineering initiative.

Provide guidance and support to projects, and manage a data change control function for the BPR projects. Establish a data administration process as part of the Business Integration Center. Validate new BPR projects against the enterprise data model and the data architecture.

3.2.4 Goal 4 - Improve the Return on Investment in Data Collection

3.2.4.1 Promote Maximum Shareability and Re-Use of Data Resources

- **Business Objective**

Ensure the maximum utilization of enterprise data resources, and that cost is not incurred to collect or reenter data that already exists.

- **Background**

Every new information project should ensure that required data does not exist in USDA databases or from external sources before requiring that it be collected and entered manually. Data managers will promote data sharing and reuse by making it easier to locate existing data stores, by advertising the existence of potentially sharable data, and by analyzing the data requirements of proposed new and reengineered systems.

- **Actions**

Work with individual projects to identify potential sources of existing electronic data within the enterprise or from external sources where cost effective. Work to increase partnerships with state and local governments, and other organizations, to collect and maintain data. Work closely with these groups to establish, or incorporate existing metadata standards to facilitate data sharing. Participate in governmental and industry data standards groups.

4 IMPLEMENTATION APPROACH

4.1 Information Technology Architecture provides an engineering perspective

The Service Center Data Management Plan provides the strategy for implementing the Service Center data management component under the USDA Information Technology Architecture (ITA). The ITA was developed by the U.S. Department of Agriculture to satisfy legislative requirements, and present USDA with the opportunity to improve the management of strategic information, plan for IT investment, implement interoperable systems, facilitate change management, improve productivity, and establish true data sharing. The ITA provides a very high level framework pertaining to:

- Business Process
- Information Flow
- Applications
- Data
- Technology

A correctly documented and implemented ITA links strategic requirements to systems that support them. This is accomplished by analyzing the business processes and relating those processes throughout all successive layers of an ITA including data, applications, etc. This approach allows the architecture to be built on core business practices that are important to the Service Centers. Thus, the Service Centers will create an architecture that supports the business, not the business supporting the technology architecture. The plan adds specificity and detail for implementing ITA.

The Zachman Framework has been chosen as the basis for the Service Center ITA. A more detailed discussion of the Zachman Framework is available in Appendix 5.5.

4.2 Implementing A Data Management Infrastructure

The goals of the Service Center initiative for improved delivery of information and services to USDA customers will require a significant change in the way data management be viewed within the Service Centers partner agencies. Software applications, and the data collected by them, can no longer stand alone. Applications co-exist in a community of other applications that share a common computing environment share common data, and are delivered to a common community of customers. Data and information must be viewed and organized from the perspective that the Service Center agencies function as a single enterprise. The concept of enterprise data management is emerging as an important component for large distributed organizations.

4.3 Framework for Service Center Data Management

The SCDT has adapted a Model for Enterprise Data Management as a framework for implementing enterprise-wide data management principles and practices [Exhibit 4.1]. The framework structures the enterprise's many data management activities into six high level components. These components fall into two major categories: those that provide direction to data management activities, and those that deal with implementation. The framework describes the functions, roles and responsibilities that constitute key components of enterprise data management, and describes the boundaries and dependencies between the functions.

The framework approach is used to implement sound data management principles for the development of Business Process Reengineering initiatives and to lay the foundation for the data management function of the Information Technology Service Bureau.

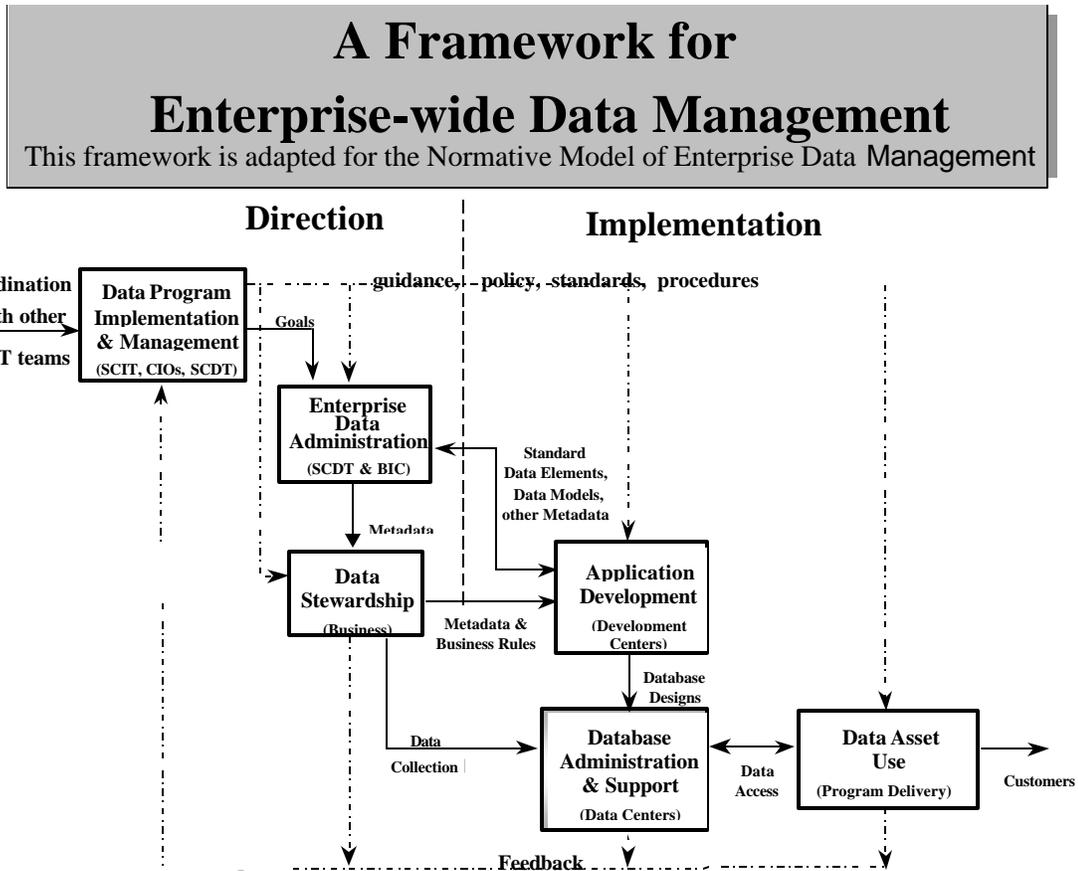


Exhibit 4.11: A Framework for Enterprise-wide Data Management

The six components included in the framework are Data Program Implementation and Management, Enterprise Data Administration, Data Stewardship, Application Development, Database Administration and Support, and Data Asset Use.

The framework components are described as follows:

4.3.1 Data Program Implementation and Management

This component of enterprise-wide data management is responsible for the definition, coordination and implementation of a Service Center wide data management program. The function establishes the vision and goals from an enterprise perspective. It establishes the organization, processes, policies, plans, standards, metrics and resources, that is, the

¹ This framework is adapted for the Normative Model of Enterprise Data Management which originated from work at MITRE in 1994 and 1995. This original work is presented in the Auerbach 1998 Handbook of Data Management. Burton Parker, one of the original contributors to the MITRE work, has evolved the original concepts from three to its current six components. This work is used by permission.

programmatic infrastructure, for enterprise data management activities. These activities will be implemented and managed as a coherent set of activities and closely coordinated with other Service Center implementation teams.

The Data Program component will be implemented through an appointed Service Center Data Team (SCDT) that develops data management policies, standards, and operating procedures. Nine Technical Working Groups (TWG) have been commissioned to develop the data management infrastructure. Oversight, guidance, and direction to TWG is provided by the SCDT Executive Team which is made up of a Data Management program leader from each Service Center Partner. All policies, standards, and operating procedures are reviewed and approved by the partner agency CIOs and the FAC Executive Officer who coordinate agency reviews and who function as a steering body for the SCDT. Data Management policies are presented to the Management Review Board (MRB) which is responsible for their final acceptance and implementation.

4.3.2 Enterprise Data Administration

This component of enterprise -wide data management is responsible for ensuring data is shareable across business areas, across the enterprise, as well as with the public and other interested parties outside of the Service Center. Specifically, enterprise-wide data administration is responsible for identification, coordination, and organization of the data to be shared. This focus on the enterprise view of data assets leads to maximized data sharing and minimized data redundancy.

Enterprise -wide data administration entails definition of the enterprise data architecture, management of an integrated enterprise data model, enforcement of standards and conventions, management and administration of a central metadata repository, identification and management of standard data elements, and delivery of training to those who create, manage and use the enterprise's data assets. These enterprise data administration activities are primarily carried out through the SCDT and SCDT's role in the Business Integration Center.

The Business Integration Center (BIC) is the focal point for all Business Process Reengineering (BPR) activities. This facility, located in the Washington, D.C. area, is dedicated to the integration of reengineered business applications. The focus is upon integration of business, data, and technical platforms (L/W/V and CCE), and testing of cross-agency as well as agency-specific projects to improve the delivery and cost-effectiveness of USDA programs.

4.3.2.1 Data Team Role in the Business Integration Center

The Service Center Data Team (SCDT) plays an important role in coordinating data management activities of all new systems and Business Process Reengineering (BPR) projects that pass through the Business Integration Center.

- **Training** -The SCDT will provide training needed to effectively perform project activities related to data management.
- **Data Sharing & Integration** – Standards for the metadata collected and stored by project teams are required to maximize data sharing and reuse. The SCDT will assist project teams in documenting project data and data models according to standards adopted by the partner agencies. Review of project deliverables by the data team will promote system integration, reuse of existing data, adherence to naming standards, and that sufficient metadata is collected. Reviews will also identify possible system performance problems early in the development lifecycle. SCDT reviews will assist the projects in ensuring that data moves seamlessly and is accompanied by its metadata.
- **Central Metadata Repository** - The SCDT will maintain a central metadata repository to store, maintain, and distribute service center metadata related to the new and BPR projects. Assistance will be provided to the project teams to ensure that new and changed metadata are coordinated and properly documented in the repository and available for sharing with other projects.
- **Enterprise Data Model** -An Enterprise Data Model for the Service Centers will be maintained by the SCDT. Each project's data model will be reviewed and integrated into the Enterprise Data Model. Metadata will be reviewed to ensure data ownership is maintained, absence of key clashes, and maximum sharing and reuse of data and related processes.
- **Configuration Management** -The SCDT will have a role in Configuration Management. As projects move into production, previous versions of models and related metadata must be preserved and impacts of proposed changes must be assessed.

4.3.3 Data Stewardship

The data stewardship component of enterprise data management is a business function responsible for identification, specification, and standardization of data assets within a specific business area. Data stewardship is a process for achieving data integrity. The objective of data stewardship is managing data assets to improve reusability, accessibility, and quality of data assets.

Data stewardship is implemented at a business or subject area level. Each business area database will have a designated executive data sponsor and data steward. The data steward in particular is a high profile position because the steward acts as a conduit between the end user community and IT. There is a great deal of collaboration and communication that must take place to ensure that data assets of the enterprise are used to achieve the maximum return on investment.

A detailed listing of the roles and responsibilities of the Data Steward is given in Appendix 5.2.

4.3.4 Application Development

Application development responsibilities include requirement analysis, business area data modeling, and the design, organization, storage, and distribution of data assets needed for a business area's activities. The project's responsibility includes integrating with enterprise data assets and maximizing use of enterprise data assets that are available for sharing. The majority of application development activity takes place at the partner agencies' software development centers located in Kansas City, Missouri (FSA); St. Louis, Missouri (RD); and Fort Collins Colorado (NRCS).

System developers are responsible for interacting with data stewards and converting the business information (i.e., metadata, business rules, etc.) into models. Computer Aided Software Engineering (CASE) tools may be used to capture data models and metadata. Information is transferred between the project specific CASE tools and the central metadata repository and enterprise data model. System developers will interact with the SCDT and the Business Integration Center throughout the development process.

SCDT is currently developing standards for metadata and interoperability with the central repository.

4.3.5 Database Administration

This component of enterprise-wide data management is responsible for supporting the physical or real instances of data within the enterprise. Activities include initialization, operation, tuning, maintenance, backup and recovery, and archiving of the data assets. Database administration is responsible for providing end-user access to the data assets so business activities can be accomplished. With the insight gained from managing day to day operations, database administrators can provide valuable feedback on issues relating to performance, capacity, etc.

Database administration activities normally are functions performed within the data centers. The processes for database administration have yet to be fully established. The SCDT will work to establish common standards and processes for the Service Center database administration function.

4.3.6 Data Asset Use

This component of enterprise-wide data management is the actual use of the data assets to carry out the business of the Service Centers, that is, the actual program delivery. Since data assets exist to meet the needs of the customer, this is where the value of the data assets is realized. Feedback from this component should determine whether the right information has gotten to the right people at the right place at the right time and for the right cost.

5. APPENDIX

Appendix 5.1 - Authorities and References

Appendix 5.2 - Data Management Roles and Responsibilities of the Business Disciplines

Appendix 5.3 - Federal Directives Establish Requirements for Data Management

Appendix 5.4 - Glossary

Appendix 5.5 - Zachman Framework

Appendix 5.1 Authorities and References

Authorities

- a. Computer Security Act of 1987. (Public Law 100-235)
URL: <http://csrc.nist.gov:80/secplcy/csa87.txt>
- b. Departmental Regulation (DR) 3400-4 Departmental Data Administration Program.
- c. Executive Order 12906, Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure, April 11, 1994.
URL: <http://library.whitehouse.gov/Search/Query-ExecutiveOrders.html>
- d. Office of Management and Budget (OMB) Circular A-130, Management of Federal Information Resources.
URL: <http://www1.whitehouse.gov/WH/EOP/OMB/html/circular.html>
- e. OMB Circular A-16, Coordination of Surveying, Mapping and Related Spatial Data Activities, revised October 1990
URL: <http://www1.whitehouse.gov/WH/EOP/OMB/html/circular.html>
- f. The Paperwork Reduction Act of 1980 (Public Law 96-511) as amended by the Paperwork Reduction Reauthorization Act of 1986.
URL: <http://www.law.cornell.edu/uscode/rr/3501.shtml>

References

- a. Charter of the Service Center Data Team.
- b. Interagency Project Manual (IP 3430-2) "InfoShare, FISVIS, MAP Data Element Standardization Manual", InfoShare Data Team.
- c. FGDC, Content Standards for Digital Geospatial Metadata, June 8, 1994.
URL: <http://www.fgdc.gov/Metadata/metav1-0.html>
- d. Spatial Data Transfer Standard, FIPS 173.

Appendix 5.2 Data Management Roles and Responsibilities of the Business Disciplines

Each store of data will have the following roles and responsibilities assigned for it:

The **Executive Data Sponsor** for each application or business area will:

- 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 (as required by Executive Order 12906)
- Coordinate funding for data collection, storage, and maintenance; and software application development, support, and maintenance
- Develop and implement the policies and procedures necessary for ongoing management of the physical data content, including the retention and archiving of the data
- Designate a Data Steward, and other critical data management roles and responsibilities as appropriate
- Authorize the release of data and application software to internal and external customers

The **Data Steward** for each application or business area will:

- 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 data management issues.
- Establish and maintain business rules and the definition of data elements; identify valid data values; establish data quality and certification standards; and establish availability, security, and access authority for the data.
- Ensure that metadata is documented, approved, and certified for release according to the current agency standards, as appropriate.
- Review and update data to ensure its validity, accuracy, and completeness; certify that data meets quality standards; and certify that data is ready for release for internal and/or public use.
- Provide training on data management roles and responsibilities.

Appendix 5.3

Federal Directives Establish Requirements for Data Management

In addition to USDA business requirements for Data Management, the Federal government has enacted a variety of data-related policies, guidelines, and executive orders in recent years. One major emphasis is on better delivery of data to the public. These Federal policies have a direct impact on agency responsibilities in data, metadata, and database management. Some of the applicable policies include:

- **Raines Rules** - These set criteria for investment in major information systems. Systems must support core/priority mission functions; support simplified work processes that are redesigned to reduce cost; make use of commercial off-the-shelf technology; reduce risk by using fully-tested pilots; and be implemented in phases.
- **Government Performance and Results Act** - This act emphasizes the gathering of data to measure program effectiveness, and thereby effects the design and implementation of technology and systems for the Service Centers.
- **Office of Management & Budget (OMB) Circular A-130** - This establishes that agencies shall: "Disseminate information products on equitable and timely terms... Agencies have the responsibility to provide information to the public consistent with their missions: Their responsibility requires that agencies distribute information at the agency's initiative, rather than merely responding when the public requests information."
- **USDA Regulations on Public Access to USDA Electronic Information** - These regulations establish that "Facilitating public access to government information shall be considered an integral part of an agency's mission." It further specifies that "agencies shall employ an active approach in information dissemination services where feasible."
- **National Spatial Data Infrastructure (NSDI)** - This establishes, by Executive Order, a coordinated program to develop the NSDI. It calls for heads of agencies to submit a schedule and coordinated funding plan to the OMB and authorizes the OMB to periodically conduct a budget crosscut to ensure a coordinated and coherent Federal effort in spatial database development.
- **Office of Management & Budget (OMB) Circular A-16** - This establishes the Federal Geographic Data Committee (FGDC) with responsibilities to promote the coordinated development, use, sharing, and dissemination of spatial data. A-16 also establishes responsibilities for Federal agencies to cooperate, supply necessary information, and conduct its surveying, mapping, and related spatial data gathering and product distribution activities in a manner that provides effective Government-wide coordination.

- **Privacy Act of 1974** - This act provides privacy protection by allowing an individual to decide what records kept by the Government are important to him/her, and permits him/her to insist that those records be used only for their intended purpose.
- **Paperwork Reduction Act of 1980** - This act establishes policies and procedures for controlling paperwork burden imposed by Federal agencies on the public. It minimizes the public burden and Federal cost of data gathering. It also mandates OMB approval prior to conducting surveys, interviews, or questionnaires involving more than 10 people.
- **Paperwork Reduction Act of 1995** - This act requires agencies to plan for the development of any new collections of information, and the extension of ongoing collection. Agencies must seek public comment on proposed collections, certify to OMB that efforts have been made to reduce the burden of the collection, and have in place a process of independent review of information collection requests prior to submission to OMB.
- **Computer Security Act of 1987** - This act is aimed at improving the security of sensitive information in Federal computer systems through Federal worker awareness training , establishment of responsibilities for developing computer systems security standards and guidelines, and the requirement that agencies establish computer system security plans.

Appendix 5.4 Glossary

This section clarifies terminology 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.)

CASE Tools

Software tools used to automate software development activities.

Data

A discrete fact or value. Data is the raw material of a system supplied by data producers and used by information producers.

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 Service Center data resources to achieve (1) a common understanding of data resources available; (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 Clearinghouse

A facility for advertising and distributing datasets. Metadata describing available datasets is made available to a chosen audience (i.e., the general public, partner agencies, etc.). Packaged datasets (i.e., diskettes, CDs, etc.) or online retrieval of selected subsets of data are distributed within security guidelines. There may be a charge for obtaining datasets.

Data Dictionary

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

Data Integrity

The state that exists when data is handled as intended and is not exposed to accidental or malicious modification, destruction, or disclosure.

Data Management

The managerial function that focuses the technical planning and operating data functions (i.e., strategic planning, data administration, and database administration) on meeting program delivery goals.

Data Mart

A type of data warehouse that contains smaller subsets of data and focuses on a particular business discipline or organizational component.

Data Model

A pictorial view of data, groupings of data, relationships between data groupings, or 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 constraint 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.

Data Steward

A business area expert who is assigned responsibility for the data content of the database. The data steward establishes business rules, defines data elements, identifies valid data values, establishes certification standards, and establishes the completeness and availability of the data.

Data Validation

Applying a set of rules, comparisons, or decisions to a data element to determine if it falls within the pre-established boundaries of values for that element.

Data Warehouse

An informational database, or collection of databases, used to store shareable data. The warehouse is usually created through data extracts from operational databases. The warehouse adheres to a single enterprise data model to ensure consistency of decision-support data across the 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 interrelated data organized to serve one or more applications.

Database Administration

The operation of data storage systems such as a DBMSs or file systems, to include the implementation of data-related business rules and security rules, maximization of performance, and protection of data assets.

Database Administrator

The person who creates, manages, controls, and protects a database.

Domain

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

Enterprise Data Architecture

See "Data Architecture". Emphasizes that the data architecture extends to the entire business enterprise of the Service Center agencies.

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 attributes (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 that 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.

National Database

A permanent database that (1) has international, national, USDA, or agency-wide application, (2) is included in a standard software suite, (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.

Reverse Engineering

The development of a data model and associated specifications for an existing system by manual, or through specialized software, analyzing existing databases and software. This is usually done for applications that do not have a data model or specifications in existence or for which the model and specifications have become obsolete. The products of reverse engineering are used for maintenance, movement of the system to another platform, or provision of a baseline for re-design or re-engineering of the system.

Strategic Data Management Planning

Planning how data is acquired, stored, and used in the most efficient manner to support the agency mission.

Appendix 5.5 Zachman Framework

The implementation of Service Center Data Management is based upon an accepted industry information engineering approach. The Data Management Team recognizes the importance of working in concert with the business and technical communities.

The Zachman Framework for Enterprise Architecture uses a two-dimensional matrix to represent the components of an enterprise information system, as depicted in Exhibit 5-1. The following table shows one dimension (the columns) of the matrix and represents the components, or domains, of a Zachman information system and the types of information involved in each domain.

Domains	Type of Information
Data	What information and data are important to the business?
Function	How does the business function?
Network	Where is the business conducted?
People	Who accomplishes the business?
Time	When are the significant business events?
Motivation	Why is this the business?

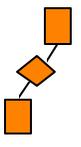
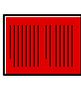
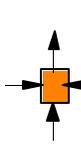
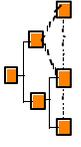
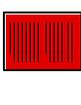
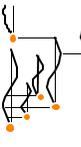
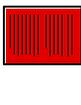
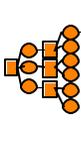
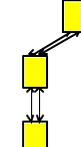
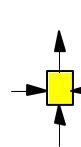
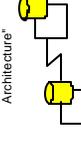
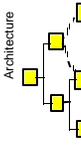
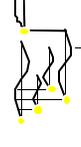
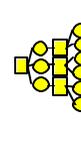
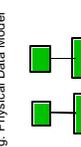
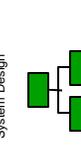
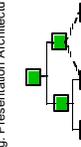
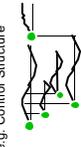
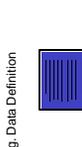
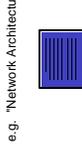
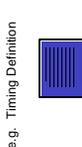
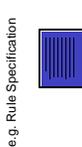
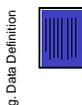
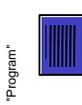
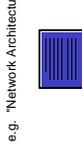
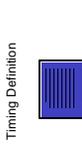
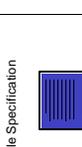
Exhibit 5-1: Enterprise Information System Components

The rows of the Zachman matrix (shown in Exhibit 5-2) represent stages in the evolution of an information system. The first three rows, Scope, Enterprise Model, and System Model, constitute a description of the functional architecture. The fourth row, Technology Model, describes the technology that is available to implement the functional architecture. The remaining rows, Detailed Representations and Functioning Enterprise, describe the system design. The suggested business interpretations for these rows in the Zachman Framework are:

- Concept of Operations or Scope and Objectives
- Business Rules or Model of the Business System
- Model of the Information System
- Technology Availability or Model of the Technology
- Detailed Description or System Design
- Functional System Level

As the characterization of an information system progresses through the stages or views of the Zachman Framework, the level of detail becomes increasingly crisp and finely drawn. While each domain is initially treated as an independent description of the component, as the matrix begins to fill out, the contents of the domains are balanced against one another. The result is a balanced representation of the components of an information system. It is important to note that the domains and layers of the Zachman Framework matrix are all interconnected and have a distinct relationship with each other. For example, the motivation of a particular organization will directly drive the data that is important to its business.

ENTERPRISE ARCHITECTURE - A FRAMEWORK TM

SCOPE (CONTEXTUAL)	DATA	FUNCTION	NETWORK	PEOPLE	TIME	MOTIVATION	SCOPE (CONTEXTUAL)
<i>Planner</i>	List of Things Important to the Business  ENTITY = Class of Business Thing e.g. Semantic Model 	List of Processes the Business Performs  Function = Class of Business Process e.g. Business Process Model 	List of Locations in which the Business Operates  Node = Major Business Location e.g. Logistics Network 	List of Organizations Important to the Business  People = Major Organizations e.g. Work Flow Model 	List of Events Significant to the Business  Time = Major Business Event e.g. Master Schedule 	List of Business Goals/Strat  Ends/Mean=Major Bus. Goal/ Critical Success Factor e.g. Business Plan 	<i>Planner</i>
<i>Owner</i>	Ent = Business Entity Rel = Business Relationship e.g. Logical Data Model 	Proc = Business Process I/O = Business Resources e.g. "Application Architecture" 	Node = Business Location Link = Business Linkage e.g. "Distributed System Architecture" 	People = Organization Unit Work = Work Product e.g. Human Interface Architecture 	Time = Business Event Cycle = Business Cycle e.g. Processing Structure 	End = Business Objective Means = Business Strategy e.g. Business Rule Model 	<i>Owner</i>
<i>Designer</i>	Ent = Data Entity Rel = Data Relationship e.g. Physical Data Model 	Proc = Application Function I/O = User Views e.g. "System Design" 	Node = I/S Function (Processor, Storage, etc) Link = Line Characteristics e.g. "System Architecture" 	People = Role Work = Deliverable e.g. Presentation Architecture 	Time = System Event Cycle = Control Cycle e.g. Control Structure 	End = Structural Assertion Means = Action Assertion e.g. Rule Design 	<i>Designer</i>
<i>Builder</i>	Ent = Segment/Table/etc. Rel = Pointer/Key/etc. e.g. Data Definition 	Proc = Computer Function I/O = Screen/Device Formats e.g. "Program" 	Node = Hardware/System Software Link = Line Specifications e.g. "Network Architecture" 	People = User Work = Screen Format e.g. Security Architecture 	Time = Execute Cycle = Component Cycle e.g. Timing Definition 	End = Condition Means = Action e.g. Rule Specification 	<i>Builder</i>
<i>Sub-Contractor</i>	Ent = Field Rel = Address e.g. DATA 	Proc = Language Stmt I/O = Control Block e.g. FUNCTION 	Node = Addresses Link = Protocols e.g. NETWORK 	People = Identity Work = Job e.g. ORGANIZATION 	Time = Interrupt Cycle Cycle = Interrupt Cycle e.g. SCHEDULE 	End = Sub-condition Means = Step e.g. STRATEGY 	<i>Sub-Contractor</i>
FUNCTIONING ENTERPRISE	e.g. DATA	e.g. FUNCTION	e.g. NETWORK	e.g. ORGANIZATION	e.g. SCHEDULE	e.g. STRATEGY	FUNCTIONING ENTERPRISE

Copyright - John A. Zachman, Zachman International