U. S. Railroad Retirement Board

Data/Object Domain Architecture

Data/Object Domain Architecture
Table of Contents

DATA/OBJECT DOMAIN DEFINITION **4**

DOMAIN TECHNOLOGY CATEGORIES **4**

DATA/OBJECT DOMAIN PRINCIPLES SUMMARY **4**

DOMAIN RELEVANT TRENDS **5**

BACKGROUND OF DATA/OBJECT TECHNOLOGIES AT THE RRB **5**

DETAILED DOMAIN PRINCIPLES **6**

DOMAIN PRINCIPLE 1 6
Centralized Data Administration *6*
Enterprise data will be managed through a dedicated data administration function. This function will manage a centralized metadata repository. *6*
DOMAIN PRINCIPLE 2 7
Data Ownership *7*
All data will have an identified owner. The data owner is accountable for the data, identifies its metadata and is responsible for coordinating change management. *7*
DOMAIN PRINCIPLE 3 7
Data Availability and Recoverability *7*
Appropriate data management techniques will be used to maintain data availability and recoverability. *7*
DOMAIN PRINCIPLE 4 8
Data Design *8*
Data design must be integrated in the appropriate phases of the systems development life cycle. *8*
DOMAIN PRINCIPLE 5 8
Data Operation *8*
Data that is accessible from existing sources, both internal and external, should be reused wherever practical. *8*
DOMAIN PRINCIPLE 6 9
Data Security, Privacy and Confidentiality *9*
Use of enterprise data must adhere to the agency's security, privacy and confidentiality policies. *9*
DOMAIN PRINCIPLE 7 9
Reporting Requirements *9*
Production, analytical and ad hoc reporting on data must be addressed in requirements definition and data design. *9*
DOMAIN PRINCIPLE 8 9
Employ Data Constraints *9*
Design teams will employ technological features of the DBMS/RDBMS to provide more robust validation and editing of the data when appropriate. *9*
DOMAIN PRINCIPLE 9 10
Data Management Tools *10*

Purchased data management tools must be industry-proven (have a significant market share) and adhere to industry standards. *10*

DOMAIN TECHNOLOGY CATEGORIES, STANDARDS AND PRODUCTS **11**

QUERY TOOLS 12
DATA BASE MANAGEMENT SYSTEMS 13
DATA WAREHOUSE 14
IMAGING/OPTICAL DISK 15
MICROGRAPHICS/PAPER 16
DATA MODELING 17
DATA REPLICATION 18
DATA BACKUP AND RECOVERY 18
DATA ADMINISTRATION 19
DATA QUALITY ASSURANCE/QUALITY CONTROL 20

PREFERRED DOMAIN DESIGN OR CONFIGURATION PATTERNS **21**

PATTERN 1 21
Capture and use metadata. *21*

DOMAIN PARTICIPANTS **22**

APPENDIX 1: DOMAIN GLOSSARY **23**

APPENDIX 2: CONCEPTUAL TO DOMAIN PRINCIPLE MATRIX **24**

Data/Object Domain Definition

The scope of the Data/Object Domain is to define the mechanics for managing, securing, designing, defining and maintaining the integrity of the data and objects without regard to platform. It addresses the need for high-quality, consistent data and objects in support of the business functions of the RRB. It requires that the data and objects be accurate and easily accessible. It provides standards and guidelines for accessing data and objects for decision support and analytical processing. The life cycle of data and objects begins with data entry, continues through transactional, operational and decision support, and ends with obsolescence and/or archival of data.

For the purposes of efficiency, all references in this document to data include data and objects.

Domain Technology Categories

- Database Management Systems
- Data Warehouse
- Imaging/Optical Disc
- Micrographics/Paper
- Data Modeling
- Data Replication
- Data Backup and Recovery
- Data Administration
- Data QA/QC
- Query Tools

Data/Object Domain Principles Summary

1. Enterprise data will be managed through a dedicated data administration function. This function will manage a centralized metadata repository.
2. All data will have an identified owner. The data owner is accountable for the data, identifies its metadata and is responsible for coordinating change management.
3. Appropriate data management techniques will be used to maintain data security, availability and recoverability.
4. Data design must be integrated in the appropriate phases of the systems development life cycle.
5. Data that is accessible from existing sources, both internal and external, should be reused wherever practical.
6. Use of enterprise data must adhere to the agency's security, privacy and confidentiality policies.
7. Production, analytical and ad hoc reporting on data must be addressed in requirements definition and data design.
8. Design teams will employ technological features of the DBMS/RDBMS to provide more robust validation and editing of the data when appropriate.
9. Purchased data management tools must be industry-proven (have a significant market share) and adhere to industry standards.

Domain Relevant Trends

- There is an increasing need to access RRB applications and data via the Internet.
- There is a greater need to access data in a distributed computing environment (use and process data crossing PC and mainframe platforms).
- Changes in law and regulations involving the Railroad Retirement Act may impact enterprise business and data needs. For example, the pending Railroad Retirement and Survivors' Improvement Act legislation would increase reporting requirements and require revisions to existing code and processes as well as defining and storing additional data.
- eXtensible Markup Language (XML) – The Federal CIO Council has established an interagency XML Working Group to facilitate the efficient and effective use of XML through cooperative efforts among government agencies, including partnerships with commercial and industrial organizations. The eXtensible Markup Language (XML) embodies the potential to alleviate many of the interoperability problems associated with the sharing of documents and data. However, realizing the potential requires cooperation not only within but also across organizations.
- SQL is being used with increased frequency, especially in the distributed processing environment.
- The future viability of IDMS is in question, and alternative systems such as IBM's DB2 will be investigated.
- The RRB beneficiary population is declining.
- There has been an increase in the legislative and regulatory changes involving RRB business partners and their resulting impact on enterprise business needs (e.g., CPI corrections, SSA, IRS).
- The enterprise is using and relying on the RRB Intranet with increased frequency.
- There is a potential for future outsourcing of certain business functions.
- RRB workforce changes are expected due to an aging workforce and retirements over the next several years.
- The RRB is looking to consolidate existing databases.
- There is an increase in the amount of redundant data stored at RRB.
- The number of RRB employees who will be telecommuting (due to legislative mandate) is expected to rise.
- The demand for customized data and data presentation is increasing.
- Systems and processing should be designed to allow "One and Done" handling. Customers, either internal or external, should be able to complete their request in a single interaction/contact with the RRB.

Background of Data/Object Technologies at the RRB

Over the years data at the RRB has been available on paper, punch cards, magnetic tape, microfiche, micrographics, disk and recently imaging. The RRB acquired the first IDMS DBMS (DATABASE MANAGEMENT SYSTEM) in the seventies from Cullinane Corporation. It provided navigational access to IDMS mainframe databases. The database management services included automatic journaling, logging, reporting and error recovery. The data was stored on mainframe disk drives and made online transactions processing available.

CA-IDMS navigational access supports the processing of hierarchical and network structured databases. It allows the program to keep record currencies and establish and maintain relationships between records. It also provides a higher level of control within program processing. The centerpiece of the CA-IDMS environment is the Integrated Data Dictionary. The dictionary is the

central repository of information about the IDMS environment. The dictionary maintains definitions of data structures, application components, systems and users. Computer Associates acquired IDMS from the Cullinane Corporation in 1989.

With the 12.0 release in 1995, CA provided the SQL option, which provided relational access to IDMS DBMS with the capability of creating, and managing a RDBMS (Relational Database Management System). This feature allowed current applications to continue to run unchanged, and SQL could be used to access the same data in new and existing applications. Through the use of CA-IDMS SERVER software, PC SQL-based tools can access the mainframe legacy data. The SQL option enables open access to CA-IDMS mainframe data. Because CA-IDMS SQL and CA-IDMS Server provide support for Microsoft Open Database Connectivity (ODBC) interface, Graphical User Interface (GUI) application development and end-user tools can be used with CA-IDMS mainframe data.

Recently the RRB has begun to use image technology. The RRB is in the process of creating applications that will access the mainframe legacy data within non-relational databases and make the information available over the Internet to our clients. Future plans will provide RRB clients access to additional information using the Internet.

Detailed Domain Principles

Domain Principle 1

Centralized Data Administration

Enterprise data will be managed through a dedicated data administration function. This function will manage a centralized metadata repository.

Rationale:

- Helps in data standardization.
- Helps associate data with business processes.
- Reduces unintended redundancy of data.
- Helps identify global data.
- Recognizes data as a valuable resource.
- Helps RRB maintain data integrity.
- Enables reuse.
- Develops and enforces standards to ensure consistency (definition, naming, use, size) across applications and platforms.
- Enables more efficient use of data storage.
- Provides technical support/self-service.

Implications:

- Extends the time line for development.
- May need additional data administration staff.
- Governance process must have clear responsibilities and consequences.
- To be effective, it has to apply to all development efforts where agency data is involved.
- Requires education and training.
- Will require big effort to capture all data dictionary inventory.
- May speed development where data assets exist.
- Will require consensus of stakeholders for initial definition and changes.
- Ensures that metadata is developed and maintained.
- Enforceable standards need to be developed if not already in place.

Domain Principle 2

Data Ownership
All data will have an identified owner. The data owner is accountable for the data, identifies its metadata and is responsible for coordinating change management.

Rationale:

Promotes one set of validation rules.
Promotes data standardization.
Associates data with relevant business processes.
Recognizes data as a valuable resource.
Helps maintain accuracy, consistency and appropriate use of the data.
Promotes reuse.

Implications:

Data owners and data users must reach consensus.
Data owners must be identified and documented in metadata.
Current data must be inventoried to create metadata.
Owners/users have the responsibility to communicate the need for changes.
Users have the responsibility to communicate problems.
Requires education and training.
Requires procedures for responsible use of data.
Owners will be responsible for obsolescence and archiving.
Archiving must be addressed during the design phase.
Owner will be responsible for ensuring that archived data remains accessible and usable.
Ongoing quality assurance and periodic data review must be provided.

Domain Principle 3

Data Availability and Recoverability
Appropriate data management techniques will be used to maintain data availability and recoverability.

Rationale:

Minimizes disruption of critical business functions.
Maintains reliability of data.
Complies with standard enterprise security procedures.

Implications:

Appropriate backup and recovery plans must be established and maintained.
Routine backup schedule must be established.
Business resumption plans must be included and maintained in addition to a backup site.
Roles and responsibilities need to be defined.
Coordination is needed with other affected domains.
Systems should be developed with backup and recovery in mind.
Must adhere to appropriate security, confidentiality and privacy principles.
Must be applied to all platforms.

Domain Principle 4

Data Design
Data design must be integrated in the appropriate phases of the systems development life cycle.
Rationale:

Ensures that proper database design principles are followed.
Promotes better implementation planning.
Broadens design team understanding of business needs and issues.
Reduces unintended redundancy of data.
Helps identify global data.
Recognizes data as a valuable resource.
Helps RRB maintain data integrity.
Enables reuse.
Helps ensure consistency (definition, naming, use, size) across applications and platforms.

Implications:

A database administrator must be included in the design team.
May extend time line due to constraints in DBA availability.
Applies to efforts that extend beyond single/limited use.
Databases will be promoted into production only after DBA review and sign off.
May need additional DBA staff.
Database designers must regularly exchange information about the databases under their jurisdiction.

Domain Principle 5

Data Operation
Data that is accessible from existing sources, both internal and external, should be reused wherever practical.
Rationale:

Lessens storage space required.
Improves accuracy, timeliness and currency of data.
Promotes retrieval of only required data.
Supports real-time, one and done.
Reduces redundancy across platforms and business partners.
Lessens RRB security burden.
Reduces overall maintenance costs.

Implications:

Assumes a reliable connection.
Assumes data availability.
Will need agreements from data owners (especially for external sources).
Certain business needs may require storage of redundant data.
Requires education, training and communication of reusability techniques.
Requires a metadata repository and public view access to it.
Will be greater impact if shared data or metadata is changed.
Allows use of batch processes when necessary.
May result in more inefficient processing such as additional I/O or network traffic.

Domain Principle 6

Data Security, Privacy and Confidentiality
Use of enterprise data must adhere to the agency's security, privacy and confidentiality policies.

Rationale:

Prevents unauthorized release of data.
Ensures authenticity of data.
Ensures consistency of security/privacy regardless of platform.
Limits liability exposure.

Implications:

Requires awareness training in regard to policies and current security approaches.
Requires adherence to RRB security policy through reporting/monitoring with appropriate support tools.
Guidelines for the use of enterprise and cross-enterprise data must be maintained, updated and communicated. This includes vehicles such as email, Internet, file transfer communication, etc.

Domain Principle 7

Reporting Requirements
Production, analytical and ad hoc reporting on data must be addressed in requirements definition and data design.

Rationale:

Lowers cost in the long term.
Improves reporting capabilities.
Improves customer satisfaction.
Provides robust data accessibility.

Implications:

Extends lifecycle. Additional tasks will be added to requirements definition and data design.
May impact efficiency of non-reporting functions.
May require more user and developer involvement in requirements definition.
System owner needs to determine priority of requirements.
Allows design team to determine the most appropriate data design.
May require additional resources for data storage.

Domain Principle 8

Employ Data Constraints
Design teams will employ technological features of the DBMS/RDBMS to provide more robust validation and editing of the data when appropriate.

Rationale:

Maintains consistency of the metadata, data content and business rules.
Provides opportunity to edit data as it is stored.
Promotes 'one and done'

Implications:

May increase response time when adding creating or modifying data.
Conflicts between data accuracy and business process needs will have to be resolved.
Requires DBMS/RDBMS training and education.
Requires coordination across platforms and organizations.

Domain Principle 9

Data Management Tools
Purchased data management tools must be industry-proven (have a significant market share) and adhere to industry standards.

Rationale:

Provides consistency.
Reduces overhead.
Don't have to learn many different tools.
Enhances knowledge base.
Eases enterprise analysis because the data will be more comparable.

Implications:

Requires education and training.
May result in higher upfront cost, but lower long term cost.
Provides greater number of 'off-the-shelf' options.
Must establish who is responsible for selecting the standard tools.
Tools must be agency approved.

This section represents a view of the standards and products that are used to implement solutions.

To better illustrate each technology category, the following attributes of that category are documented when pertinent information exists:

Standards – This table represents standards that have been adopted for central support across the enterprise. Support for a standard may be direct, or indirect by virtue of a supported product that relies heavily on a particular standard. Standards refer to those sanctioned by national or international standards bodies, or industry groups that dictate how products are developed, deployed or interoperate with each other.

Other Standards in Use – This table lists standards that have been adopted by specific areas of the company either directly or by virtue of using a particular product.

Products – This table list those products that:

1. Are centrally supported or used widely across the enterprise
2. Are being considered for central support
3. Have been centrally supported in the past

Other Installed Products – This table lists those products that:

- Are being used by a limited (usually one) number of areas
- Are not centrally supported

Product Lifecycle

The following key refers to the products and standard timelines.

Code	Term	Meaning
ID	Identified	The company is aware of the item. It has not been formally evaluated. It is not approved for current use.
RD	Research and Development	Business units may use RD components only as selected participants in a company sanctioned research effort. These items may not be used in any other context and may not be used in production.
P	Pilot	Approved for production, but widespread use not yet encouraged.
I	Invest	Appropriate use encouraged.
M	Maintain	New implementations are not encouraged. Existing systems may continue to rely upon these components and extend existing implementations.
D	Disinvest	In process of being phased out.
O	Obsolete	Vendor or industry support is gone and use in not recommended.
R	Rejected	Investigated and found to not meet the organization's needs.

THE FOLLOWING TABLES ARE POPULATEDWITH PRODUCTS AND STANDARDS, INDICATING LIFECYCLE CATEGORIZATION OVER TIME.

Query Tools

Definition: Technologies used for queries and reporting of data.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
SQL	I	I	I	I												

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
OLQ	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
FileAid	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Easytrieve Plus	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
SAS Mainframe	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
SQL Server Query Analyzer	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Microsoft Query	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Visual Interdev Query Window	M	M	M	M	M	M	D	D	D	D	D	D	D	D	D	D
VB.Net	RD	RD	RD	RD	RD	RD	P	P	P	I	I	I	I	I	I	I
Eastman Imaging Routebuilder Query	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
dBase 4 Query	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
Microsoft Access	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Cobol	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Microsoft Visual Basic	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
Webconnector Query	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
HSL	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
IBM Utilities	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
IDMS Utilities	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
DMLO	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
OPAL	P	P	P	P	P	P	P	P	P	P	P	P				
Decision Analyzer			RD	RD	RD	RD	RD	RD	RD	RD	RD	RD	RD	RD	RD	RD

Other Installed Products:

Product Name	Usage Scope
SAS PC	Number of available licenses.

Data Base Management Systems

Definition: Technologies that manage data for multiple users ensuring data integrity and security.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Relational Database																
Network Database																
ANSI																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
CA-IDMS/DB 14.x	I	I	I	I	I	I	I	I	O	O	O	O	O	O	O	O
CA-IDMS/DB 15.x									I	I	I	I	I	I	I	I
Microsoft SQL Server 7.x	M	M	M	D	O											
Microsoft SQL Server 2000	ID	ID	ID	I	I	I	I	I	I	I	I	I	I	I	I	I
Microsoft Access 97	M	M	M	D	O											
Microsoft Access 2000	ID	ID	ID	I	I	I	I	I	I	I	I	I	I	I	I	I
Microsoft Access XP	ID	ID	ID	ID												
DBASE4	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
IBM DB2 for OS/390	ID	ID	ID	P	P	P	P	P								

Other Installed Products:

Product Name	Usage Scope
dBase 4	Office of Programs
Microsoft Exchange Server 5.50	
Microsoft Exchange Server 2000	

Data Warehouse

Definition: Technologies that support database decision making.

Standards:

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4

Other Installed Products:

Product Name	Usage Scope

Imaging/Optical Disk

Definition: Technology to store, search, retrieve, manipulate, display and print digital representations, i.e. 'images', of objects by computers.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4

Other Standards in Use:

Standard	**Usage Scope**
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Eastman Archive Manager V1.x	M	M	D	D	O											
Eastman Archive Manager Enhanced	ID	I	I	I	I	I	I	I								
Eastman Imaging Server V3.x	M	M	D	D	O											
Eastman Imaging Server V4.x	I	I	I	I	I	I	I	I								
Eastman Workflow Imaging Client V3.x	M	M	D	D	O											
Eastman Work Desk	RD	RD	P	I	I	I	I	I								
Eastman Web Desk	ID	ID														
Eastman Web Connector V1.x	I	I	I	I	I	I	I	I								

Other Installed Products:

Product Name	**Usage Scope**
None	

Micrographics/Paper

Definition: Storage media that are historical and are inaccessible electronically.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Installed Products:

Product Name	Usage Scope
Micrographics	Contracted out.

Data Modeling

Definition: Products that assist in creating models to analyze information needs and transform them into database solutions.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
System Architect (Popkin Software)	RD	RD	RD	P	P	P	P									
Microsoft SQL Server 7.x	M	M	M	D	O											
Microsoft SQL Server 2000	ID	ID	ID	I	I	I	I	I	I	I	I	I	I	I	I	I

Other Installed Products:

Product Name	Usage Scope
None	

Data Replication

Definition: Technologies that efficiently copy DBMS data and manage the distribution of the copies.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Microsoft SQL Server 7.x (replication included in the RDMS core functionality)	M	M	M	D	O											
Microsoft SQL Server 2000 (replication included in the RDMS core functionality)	ID	ID	ID	I	I	I	I	I								

Other Installed Products:

Product Name	Usage Scope
None	

Data Backup and Recovery

Definition: Technology that is used for business resumption needs.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Data Backup and Recovery continued

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Fast Dump Restore (FDR)	I	I	I	I	I	I	I	I								
IBM Utilities	I	I	I	I	I	I	I	I								
CA-IDMS Utilities	I	I	I	I	I	I	I	I								
Tivoli Storage Manager	I	I	I	I	I	I	I	I								

Other Installed Products:

Product Name	Usage Scope
None	

Data Administration

Definition: Technology for the definition and support of an enterprise-wide data architecture.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
CA-IDMS Data Dictionary	I	I	I	I	I	I	I	I								
CA-IDMS Catalog	I	I	I	I	I	I	I	I								
MS SQL Server Catalog	I	I	I	I	I	I	I	I								
CA-Endevor R3.x	I	I	I	I	I	I	I	I								
BMC Control-M R5.x	I	I	I	I	I	I	I	I								
MS SQL Server Enterprise Manager	I	I	I	I	I	I	I	I								

Other Installed Products:

Product Name	Usage Scope

Data Quality Assurance/Quality Control

Definition: Technologies to enhance data reliability and consistency.

Standards:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
None																

Other Standards in Use:

Standard	Usage Scope
None	

Products:

	2001				2002				2003				2004			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Microsoft SQL Server	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
CA-Enforcer	ID	ID	ID	ID	ID	ID	ID	ID								
Eastman Imaging Utilities	I	I	I	I												
CA-IDMS Utilities	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I

Other Installed Products:

Product Name	Usage Scope
None	

Pattern 1

Capture and use metadata.

Purpose

To promote reusability of data.

Applicability

Use this pattern in data design and application design.

Assumptions

- It is assumed that this pattern will be included in the systems development life cycle.
- An enterprise metadata repository will need to be created.
- A data administration function will need to be clearly defined.
- Tools for storage and query of the metadata must be evaluated and purchased.
- Training for use of the products will need to be done.
- Training on metadata guidelines, standards and procedures will need to be done.
- Quality assurance for data will be a data administration function.
- The data administration staff may delegate functions. For example, in the initial development of an application, a project team member may handle some data administration tasks.
- In product selection, we want a product with multiple access paths and a search capability.

Structure Overview

None

Detailed Pattern Description

- A metadata entity will minimally include the following entries. Each type of entity will require specific entries. For example, a data element will require a picture entry. A file would not.
  - Name
  - Owner
  - Use cross reference
  - Definition
  - Type (e.g., data item, object, image, file, etc.)
- Develop guidelines for:
  - Data administration
  - User reference

Benefits

1. Promotes consistent documentation for the data.
2. Supports and improves data reusability.
3. Improves change management because you know who owns the data and who uses the data.
4. Supports the concept of data being a corporate asset.
5. Improves systems development life cycle by accessing a repository of information about enterprise data.

Consequences

1. Initially, the systems development life cycle may be increased as the data repository is being populated.
2. Additional personnel will be required to accomplish the data administration function.
3. The acquisition, training and maintenance of software tools will require funding.
4. Errors in the central repository will have greater impact.

Variations

None.

Related Patterns

None.

Known Uses

None.

Domain Participants

Domain Team Leader: Bob Schempp

Line of Business Representatives: Catherine Leyser, Nancy Marks, Marla Huddleston

Domain Participants: Marie Leeson, Steve Gehrke, Eileen Leiderman, Sue Feltman, Julia Pelt, Robert Piech, Frank Restivo, Tim Conheady

OEA Representative: Robert LaBerry

Appendix 1: Domain Glossary

Data	Distinct pieces of information, usually formatted in a special way. Data can exist in a variety of forms, as numbers or text on pieces of paper, as bits and bytes stored in electronic memory.
Data Modeling	Used to develop an accurate model, or graphical representation, of the client's information needs and business processes. The data model can act as a framework for the development of the new or enhanced application. Data modeling can occur in logical design and physical design.
Data Owner	The enterprise organization (bureau, division or section) responsible for development, maintenance and stewardship of the data. The data owner's responsibilities include: • coordination with users to ensure that data meets expected needs; • participation in developing and validating metadata components related to business needs; • coordinating and reporting metadata changes to data administration; • monitoring the data for accuracy and validity; and • change management.
Enterprise	In the computer industry, this term is used to describe any large business organization that utilizes computers. In our case, Enterprise refers specifically to the RRB.
Enterprise Data	Data that supports the business conducted at the RRB and is made available to all approved RRB and business partners. Enterprise data may be housed either at the RRB itself or by one of the RRB's business partners. Data created and used only by individual users or business units is not considered to be enterprise data. This type of data may be used for analysis, for creating specialized, limited use, internal reports, or as temporary work files. An example of this type of data is an Access database used to answer a research question and housed on a PC's hard drive.
Industry Standard	Industry standards usually refer to those sanctioned by national or international standards bodies or industry groups that dictate how products are developed, deployed or interoperate with each other. However, a data processing item like a data format, programming language, communication protocol, or operating system can often be considered a standard, not because it has been approved by a standards organization, but because it is widely used and recognized by the industry as being standard. Standards are extremely important because they are indicative of technologies that are strongly supported by many customers and a broad information technology industry base.
Metadata	Metadata is data about data. It is a way of documenting information about data. Metadata includes things like the name, owner, length, valid values and description of a data element. It could describe how and when and by whom a particular set of data was collected, and how the data is formatted. In database management systems, metadata files are such things as index files and data dictionaries, and store administrative information.
Object	Generally, any item that can be individually selected and manipulated. This can include text, shapes, pictures, sounds, etc.
Platform	The underlying hardware or software for a system. The platform defines a standard around which a system can be developed. Once the platform has been defined, software developers can produce appropriate software and managers can purchase appropriate hardware and applications. The term is often used as a synonym of operating system. The term cross-platform refers to applications, formats, or devices that work on different platforms.

Appendix 2: Conceptual to Domain Principle Matrix

	Relationship Between RRB's Domain Principles And Conceptual Architecture Principles																								
Domain Principle	***Conceptual Architecture Principles***																								
	CA 1	CA 2	CA 3	CA 4	CA 5	CA 6	CA 7	CA 8	CA 9	CA 10	CA 11	CA 12	CA 13	CA 14	CA 15	CA 16	CA 17	CA 18	CA 19	CA 20	CA 21	CA 22	CA 23	CA 24	CA 25
D-1		X			X		X							X											
D-2					X		X	X						X											
D-3							X									X	X								
D-4					X		X							X									X		
D-5							X							X											
D-6							X									X									
D-7				X	X				X														X		X
D-8								X																	
D-9											X									X					

Conceptual Architecture Guiding Principles:
1. Use guidelines consistent with the Federal Enterprise Architecture. 2. Support a single Enterprise Wide Technical Architecture (EWTA). 3. IT projects are to be consistent with the Enterprise Architecture. 4. IT projects are to be consistent with the Enterprise Architecture. 5. Reduce integration complexity. 6. Technical architecture must be extensible and scalable. 7. Manage information and data as enterprise-wide assets. 8. Validate information as close to its source as possible. 9. Enhance the ability to capitalize on and exploit business information. 10. Support multiple data types. 11. Make an informed buy versus lease versus build decision before proceeding with any new development project. 12. Require shorter development cycle times. 13. Keep current with emerging technologies and their applicability to enterprise architecture. 14. Maximize infrastructure asset reuse. 15. Sustain reliable connectivity. 16. IT systems will be implemented in adherence with the agency's security, confidentiality and privacy policies. 17. The agency will use a consistent set of security interfaces and procedures. 18. Reduce total cost of operation (TCO). 19. Extend E-Mail to Become a Corporate Information Exchange Vehicle. 20. Adopt Open Systems Standards. 21. Reduce duplicate information systems. 22. Reduce duplicate information systems. 23. Maximize and exploit Internet and Intranet technologies and approaches. 24. Integrate Enterprise Architecture into the investment management process. 25. Customer perception is a measure of the quality of the automation processes.