FINAL

OREGON GIS UTILITY PROJECT—PHASE 1
REQUIREMENTS ASSESSMENT AND BUSINESS CASE

GEOGRAPHIC DATA STEWARDSHIP BEST PRACTICES
Deliverable 3E

Submitted to:

Cy Smith, Statewide GIS Coordinator
State of Oregon
Dept. of Administrative Services—IRMD
Geospatial Enterprise Office
1225 Ferry Street, 1st Floor
Salem, OR 97301-4280
(503) 378-6066

Submitted by:

Peter Croswell, Executive Consultant
PlanGraphics, Inc.
112 E. Main Street
Frankfort, KY 40601
(502) 223-1501

Draft: August 19, 2005
Final Revision: December 30, 2005

TABLE OF CONTENTS

List of Tables

SECTION 1
INTRODUCTION

1.1 WHAT IS DATA STEWARDSHIP?

According to thefreedictionary.com, stewardship [is the] “responsibility for taking good care of resources entrusted to one.” Applying this meaning to the GIS Utility, “geographic data stewardship” encompasses all aspects of effective data maintenance and provides access to that data by users. This report provides a context for an effective organizational and technical environment to maintain high-quality geographic data and to make this data available to users. Basic principles that serve as a foundation for data stewardship are explained in Section 2. A set of best practices that draws on documented experiences of large GIS programs is presented in Section 3.

1.2 THE NATURE OF GEOGRAPHIC DATA

In a broad sense, geographic data is any type of data that has an explicit reference to a location. This includes traditional vector GIS data, as well as other non-map data types that have logical locational ties (e.g., map feature link, site address, street segment, parcel, district, etc.). Table 1-1 describes the major types of geographic data.

Table 1-1: Types of Geographic Data

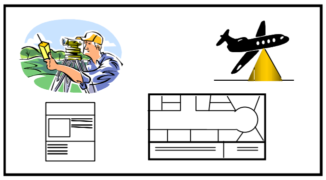
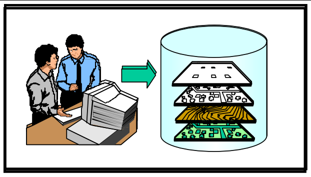
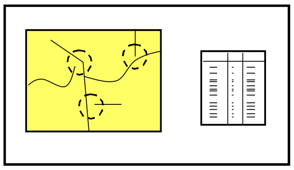
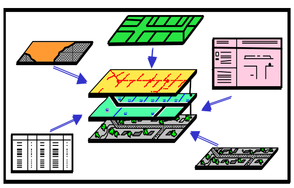
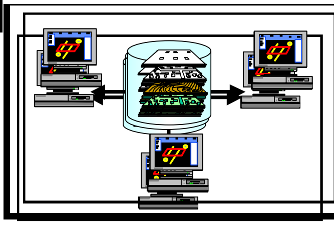
Data Types	Explanation
Vector GIS	Vector representations of map features and locations as point, line, area, or complex objects (e.g., networks).
GIS Attributes	Tabular attribute data (text or number values) stored in a recognized database format logically associated to map features and managed through GIS software.
GIS Annotation	Permanently stored text in GIS software package associated directly with vector map features or areas in the GIS database.
Raster Geographic Image	A geographically registered “layer” that is in raster form. This could be a digital orthoimage (compiled from aerial photography or satellite data), a geo-registered image without full ortho control (e.g., scan of a rubber-sheeted aerial photograph), or a map stored in grid format with values assigned to cells.
Digital Elevation Model (DEM)	Elevation data stored in one of several proprietary formats (e.g., ESRI TIN) that can be used by GIS software for display and analysis.
CAD File	A vector graphic file usually containing engineering design information (usually for small project areas) that is stored in an accepted CAD format (usually AutoCAD or MicroStation). CAD files are differentiated from vector GIS files in that they store features using CAD-based parametric definitions and do not explicitly reference a true spatial coordinate system or map projection.
Document	Any record (e.g., form, report, legal document, site photograph, meeting minutes, etc.) usually scanned and stored in a standard raster format and associated with database index data to support queries and retrievals.
External Tabular Database	Any tabular database normally maintained using commercial database management software that is maintained separate from GIS. It contains one or more key fields that allow the data to be logically related or ”joined” with features or locations in a GIS database.

While the GIS Utility will have a major focus on vector and raster map data sets, sound practices for data stewardship are important for all of these data types.

1.3 STATEWIDE COMMUNITY OF DATA MAINTAINERS AND PROVIDERS

The GIS Utility initiative, referred to as "navigatOR", is a statewide effort that will include users and data providers from many public sector and non-public organizations. Potentially, all organizations and groups participating in navigatOR could be in a position to provide or update geographic data and therefore may have a lead or supporting role as a data steward. Table 1-2 summarizes the likely roles that participating organizations will play in geographic data provision or update. In this table, five main roles are identified:

Table 1-2: Data Stewardship Roles

	**Produce and Provide Source Material or Data**	Collection or compiling of data or records in digital or hard copy form which are the primary sources in the creation or update of a GIS database. This covers a range of sources that depend on the particular data element and organization responsible. Sources could include orthoimagery, GPS-based field data, inspection of inventory forms, etc.
	**GIS Database Update**	Modifying an existing GIS database, using map and database editing or data import tools, and acceptable source material, to reflect actual real world changes. Ideally, updating procedures and timing follows specific rules that vary for different data elements.
	**GIS Data Quality Assurance**	Automated and manual checks (and accompanying corrections if needed) on a new or updated database to ensure that it meets documented specifications and quality criteria.
	**GIS Data Integration**	Includes horizontal and vertical integration operations that involve certain types of processing and packaging of data from multiple datasets. Horizontal integration involves collecting pieces of data from multiple stewards, merging them into a seamless dataset and providing easy access by users, usually through a Web-based service. The integrator may also carry out some value-added service, such as data format translation, data restructuring, and combining data sets to provide more effective offerings for the user community. Vertical integration assures topological integrity among thematic datasets. This requires familiarity with multiple data standards, thematic content, integrators and working groups. For example, vertical integration is necessary for roads, bridges and culverts with surface waters over orthoimagery
	**Providing GIS Data and Related Services**	The process and service for providing GIS data, normally through a Web-based service. Services include metadata query and search, direct interactive viewing, on-line GIS applications, and data download.

For each organization, Table 1-3 shows one of three designations for each role. These designations are a subjective assessment of the organizations' contribution to overall database development and ongoing update for navigatOR.

Table 1-3: Potential Geographic Data Roles for navigatOR Participants

		Potential Geographic Data Roles[1]				
Organization Type or Group	**Explanation**	**Provide Source Material or Data**	**GIS Database Update**	**Data Quality Assurance**	**GIS Data Integration**	**Providing GIS Data and Services**
Federal Agencies	Federal agencies operating in Oregon, specifically those with a major mission for GIS data capture and update (e.g., USFS, BLM, NRCS, etc.).	**P**	**P**	**P**	**S**	**P**
State Agencies	Any Oregon State agency (Department, special office, commission) that collects or maintains maps or geographically referenced information—with a focus on those agencies with major geographically related missions (e.g., DOT, DEQ, DFW, ODF, DOGAMI, DLCD, DOR, OSP).	**P**	**P**	**P**	**S**	**P**
Tribal Governments	Any recognized native American tribal government with jurisdiction over a specific land area.	**P**	**S**	**S**		**S**
Regional Agencies	Any regional organization or special district operating independently or semi-independently from state or local government with some type of responsibility or jurisdiction over an area of land (often more than one county). Examples include COGs, school districts, etc.	**S**	**S**	**S**	**P**[2]	**P**[2]
Local Governments	City and County governments, including all departments, offices, commissions, etc., which operate as part of the local government. Cities and counties in the state with active GIS programs play multiple roles aimed at supporting local needs.	**P**	**P**	**P**	**P**[2]	**S**
navigatOR Core Operations (GEO)	The core operational function of navigatOR, responsible for operational oversight of the utility and stewardship of certain data themes.		**S**	**S**	**P**	**P**
Utility Companies	Public and private utility companies and organizations that have responsibility for utility service (water, sewer, gas, electric, telecommunications, etc.) over a specific territory.	**P**	**P**	**P**		**S**
Private Engineering or GIS Product and Service Companies	Private companies with a primary business focus on providing services for the development of GIS and GIS databases. This includes many engineering firms, aerial survey and mapping firms, data producers and re-sellers, and other companies that specialize in the collection and provision of geographic data.	**P**	**S**	**S**	**P**[2]	**P**
Private Sector Geographic Data Users	Includes a wide range of retail, commercial, and industrial product and service companies that use geographic data for planning and operations.	**S**		**S**		
Academic and Research	Public, private, and not-for-profit organizations whose primary mission is research and education. This includes public and private colleges and universities, special programs and institutes associated with universities, as well as other research organizations	**S**			**S**	**S**
Not-for-Profit Associations or Societies	Includes professional or trade associations, citizen or political advocacy groups, or organizations providing information or supporting a societal cause.	**S**	**S**	**S**		

[1]P = primary role, S = secondary role, and blank = no significant role.
[2]Many organizations in this category are not active in this role, but selected ones do play a major part.

SECTION 2
BASIC PRINCIPLES OF DATA STEWARDSHIP AND CONTEXT FOR navigatOR

2.1 BASIC PRINCIPLES

Best practices for geographic data stewardship will be based on the following overriding principles:

1. Geographic data is a valuable asset that supports the business needs of users. Its value should be maintained over time through an effective, efficient update program.

2. Organizations with missions encompassing geographic data collection should have a lead role in updating and providing that geographic data for the navigatOR.

3. The processes supporting geographic data stewardship will be based on a clear, inclusive, and well-documented data architecture.

4. Policies, procedures, and technical processes for data update should be well documented and widely communicated.

5. Metadata providing information about the content, format, quality, authority, and availability of geographic data is vital, and this metadata should be updated along with the data itself.

6. Maintaining a high-level of geographic data quality is critical.[1] Data should be maintained at a specified quality level that is well documented. This quality level should be met unless there is a good reason to deviate from it, and any deviations should be documented.

7. Geographic data should be shared widely among the entire user community in Oregon with proper consideration to legal and policy concerns that may restrict access and distribution.

---

[1]Data quality encompasses a number of interrelated aspects of the GIS databases content and format. First, quality relates to conformance with the overall technical specifications of the structure of the databases and the map feature and attributed data content. Quality levels relate directly to such parameters as positional accuracy, attribute accuracy, completeness, currency, and graphic integrity. Quality levels should be defined in metadata.

2.2 GIS UTILITY GEOGRAPHIC DATABASE

The navigatOR *Conceptual Design* includes a description of the content and format of the GIS Utility database. That document explains the primary focus on GIS data organized under 14 Framework data themes defined by the Oregon Framework Implementation Team (FIT). These themes are described below.

- Geodetic Control: Geodetic control provides the means for determining locations of features referenced to common, nationally used horizontal and vertical coordinate systems. Geodetic data provide the basic reference framework for all geospatial data, as well as a method for using different layers and sets of spatial data together. Geodetic data are essential in developing a common coordinate reference for all other geographic features. Horizontal or vertical location is used as a basis for obtaining locations of other points. Latitude, longitude, and ellipsoid height are determined relative to the Geodetic Reference System of 1980 ellipsoid, a mathematical model of the earth. Orthometric height is determined relative to the most current geoid model for the United States, GEOID93, developed by the National Geodetic Survey.

- Digital Orthoimagery: An orthoimage is a georeferenced image prepared from a perspective photograph or other remotely sensed data in which displacements of images due to sensor orientation and terrain relief have been removed. Many geographic features can be interpreted and compiled from the orthoimage. Orthoimages can serve as a backdrop in addition to linking the results of an application to the landscape.

- Elevation: Elevation refers to a spatially referenced vertical position above or below a datum surface. Digital, georeferenced elevation data exists in several forms, including digital elevation models (DEMs), triangulated irregular networks, vector contour files, and spot elevations. The other forms of elevation data can be derived from DEMs, so the DEM serves as the minimum element for elevation data within the Oregon Framework.

- Administrative Boundaries: Administrative and governmental boundaries are the district, service, governmental, election, and census polygons that serve to organize administrative and governmental functions. Administrative and governmental boundaries define geographic areas within which resources are targeted and services are managed.

- Hydrography: Hydrography includes geographic features that represent streams, lakes, reservoirs, rivers, and other surface drainage features. The elemental feature of the Pacific Northwest Hydrography Framework is the stream, defined from its headwaters to the mouth. Its unique ID is referred to as the LLID (Longitude/Latitude Identifier). Each stream segment is also referenced by its unique NHD identifier.

- Transportation: The road, rail, air and water transportation network, along with associated features, facilities, and attributes constitute the transportation theme. This theme includes a linear referencing system important for locating incidents within the network. In addition, a feature ID is being developed and applied to every segment of the network and to all associated features and facilities.

- Cadastral (Property Ownership): Cadastral information is the spatial and attribute data describing real property and the interests associated with real property. Cadastral data serves as the foundation upon which the majority of local thematic spatial data is compiled. In Oregon, the Public Land Survey System (PLSS) serves as the cadastral reference grid to which real property features and attributes are linked.

- Public Safety/Preparedness: This theme includes elements of specific importance for emergency planning and management and for response and management of specific emergency events, including natural hazards, local public safety (fire, law enforcement, emergency medical), terrorist threats, or other concerns. Addresses are an important location key for meeting these needs. Addresses are typically attribute information linked to individual tax lots as discrete site addresses, and it is advisable to develop point locations for site addresses. Many other elements are an important part of emergency management and public safety.

- Land Cover/Land Use: This theme includes elements that describe the natural and built cover on the earth's surface and the human use of land. Areas are categorized according to a specific land use or land cover classification scheme.

- Bioscience: Bioscience features relate to biological data sets of statewide concern. Many of the bioscience features relate to the biological information needs of the Oregon Plan for Salmon and Watersheds established by Executive Order. The goal of the Oregon Plan is to enhance, restore, and protect Oregon's native salmonid population, watersheds, fish and wildlife habitat, and water quality, while sustaining a healthy economy.

- Climate: The climate theme includes information on the geographic distribution and variability of temperature, precipitation, wind, and related data elements.

- Utilities: The utilities theme encompasses all infrastructure for the provision of water, sewer, telecommunications, electric, gas, and waste management services to citizens and businesses. This includes major generation, storage, treatment, and transmission facilities, as well as local networks for service provision to individual customers. Depending on the type of service and specific areas, utility services may be provided by government agencies, special service districts, private companies, or utility cooperatives.

- Geoscience: This theme provides information about the physical characteristics of the rocks, soils, and landforms in Oregon that are important to large sectors of the economy (agriculture and natural resources), construction (road and building), hazards planning, and many facets of public and private activity.

- Hazards: This theme includes the most likely natural hazards threatening life and property in Oregon. Some elements record occurrences, while others reflect the results of modeling to anticipate potential hazards.

An additional data category supplementing the 14 themes is called "Reference." This theme includes elements and data that provide general locational reference and background information to support map display, geographic queries, and GIS applications that involve general map reference.

As previously mentioned, the GIS Utility initiative will place emphasis on the development and maintenance of the Framework Data, but will do this within the context of broader geographic data types (see Table 1-1) that may be used in geographic data applications. It is also recommended that the navigatOR database accommodate certain selected GIS data sets (see the *Conceptual Design for the GIS Utility*, Section 4) which are not included in the Framework Themes or do not fully comply with data standards but which still may be of interest and value to GIS users. This manner of "opening up" the GIS Utility to non-Framework or non-standard data by no means implies that all data generated by groups in Oregon will be accepted or accessible through navigatOR. Only selected data sets that meet certain criteria for content and quality will be accepted.

SECTION 3
GEOGRAPHIC DATA STEWARDSHIP BEST PRACTICES

A "best practice" may be defined as "a technique or methodology that, through experience and research, has proven to reliably lead to a desired result." This implies that a best practice supports a practical, cost-effective, and results-oriented approach to accomplish a goal. For the purpose of geographic data stewardship for the Oregon GIS Utility, this goal is to "build and maintain, in the long-term, a comprehensive, high-quality geographic database and to make these data available to a large user community in Oregon."

Best practices are organized into the four categories explained below:

1. Database Organization and Design: Includes practices that call for sound, effective approaches for enterprise data modeling, physical database design, and establishing specifications and standards for data format and quality.

2. Database Processes: Defines workflows, procedures, and tools for effective and efficient database conversion, maintenance, quality assurance, and maintenance tracking.

3. Organizational Structure and Responsibility: Establishment of organizational responsibilities for database maintenance and the organizational elements to support ongoing maintenance.

4. Database Administration and Accessibility: Practices and policies for ongoing database administration and security and for making GIS data accessible to users.

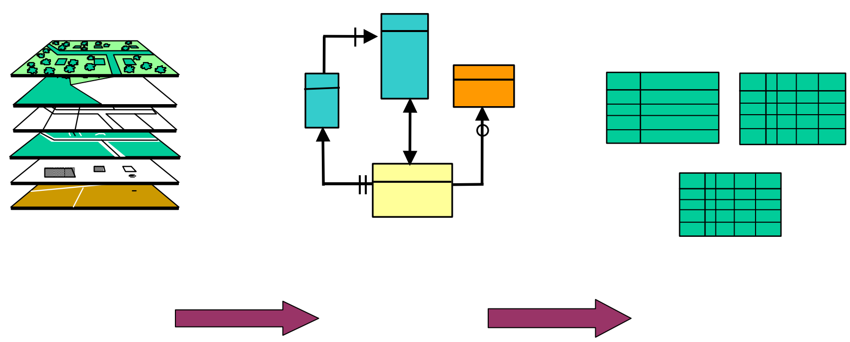
Table 3-1 at the end of this section summarizes best practices under each of the four categories and includes a reference (when appropriate) for a source (see Appendix A) that further explains best practices background and approach. Appendix B includes more information about selected Best Practices in cases where elaboration is needed.

Most of these best practices relate to data maintenance—ongoing update and administration of the navigatOR database after the data has been captured. This is the principal focus of the concept of data "stewardship." However, recognizing that a major part of navigatOR development will include major database development efforts for data themes not yet populated, several best practices address this data conversion activity.

3.1 DATA MODELING AND DESIGN

Data modeling and database design are formal steps in defining the content and structure of the geographic database. Just like the planning and design stages in building construction, it is important to carry out an effective modeling and design for databases—the larger and more complex the database, the greater the need and benefits of a formal modeling and design process. The data model is a representation of the real world and defines, in a clear way, the basic content and relationships among different database components. Data models may be developed at different levels of detail that suit the needs of particular projects. A conceptual model is a simple description of geographic data components or themes. Logical models (or "entity-relationship models") are somewhat more rigorous and include a formal definition and description of data entities (e.g., map feature set, attribute data set), identification of at least primary key identifiers for the entities, and definition of logical relationships between the entities. The modeling-design process follows a progression of steps from a general conceptual model, to a formal logical model, to a physical design as shown in Figure 3-1.

Figure 3-1: Major Database Design and Modeling Stages



3.2 DATA UPDATE SCENARIOS

Data updates will occur under a variety of different circumstances. A main premise for navigatOR maintenance is that data update will be the primary task of organizations that are responsible for documenting changes to certain types of data or that receive documents (e.g., engineering plans or project documents) that are sources for database updates. Some of the common circumstances driving database update activities include:

- Major Data Collection or Inventory Project: A major well-planned and -resourced project that involves map compilation or data collection of a certain type for a specific area. The project results in specific source materials (e.g., maps, compiled databases, engineering drawings). Examples include—a) tax lot map conversion for a county, b) natural resource inventory for a specific project area, and c) major road construction projects.

- Transaction-based Changes: Data updates or enhancements that occur more or less continuously over a period of time and which are tracked and documented. Examples include—a) updates to taxlot maps as boundaries change (parcel splits or merge) or new subdivisions are created, or b) road or utility maintenance that modify infrastructure features.

- Map Feature and Attribute Edits and Quality Improvements: Ongoing edits to map features or attributes of the GIS databases that correct or improve the accuracy or general quality of the database. The edits result from use of ancillary information sources that reveal quality problems with the database.

- Data Migration: Any restructuring or reformatting of the existing GIS database (mainly an automated process) to adapt the database to physical database design changes or to translate data format in the case of migration to a new software version or package.

3.3 QUALITY CONTROL VS. QUALITY ASSURANCE

Quality for GIS databases encompasses the following:

- Adherence to file and directory naming conventions
- Graphic construction and integrity
- Map feature completeness
- Map feature placement and positional accuracy
- Attribute accuracy
- Symbology consistency
- Annotation placement and format.

In addition to these specific criteria relating to the quality of the specific data elements being converted, overall quality concerns exist such as file and feature class naming, proper population of feature classes, and all project workspace tables.

"Quality control" (QC) includes the processes and tools used during the data conversion projects to produce data that meets quality requirements. "Quality assurance" (QA) refers to an independent process—often employing a sampling technique—that checks and ultimately accepts the converted data. QC and QA are, therefore, complementary constituents of the overall data conversion project. Both involve a number of automated and manual processes designed to ensure data quality. QC and QA may use many of the same or similar processes and automated tools. While QC and QA checking are operationally independent, the design and development of tools and procedures will overlap and, therefore, should be coordinated.

Table 3-1: Summary of Best Practices for Data Stewardship

Best Practice	Explanation	References and Sources for More Detail*
A. Database Organization and Design		
A.1: Base Design on Enterprise Architecture	Geographic database modeling should be based on a sound business-oriented architecture. In essence, this means that data included in navigatOR should reflect business needs of users in all organizations and that this business need is a basis for prioritizing data. The GIS Utility Business Case provides an overall foundation for defining enterprise architecture and this is sufficient, at a minimum, as an acceptable architecture from which database modeling design may begin. See references for more formal approaches to developing enterprise architecture frameworks. Each of these describes an approach and template format for developing detailed enterprise framework architecture—not necessarily completely applicable to the GIS Utility initiative—but they may provide some insights for a more formal enterprise architecture description.	Formal Enterprise Architecture formats: a) Zachman Framework (Zachman Institute, www.zifa.com) and b) the Federal Enterprise Architecture (www.whitehouse.gov/omb/egov/a-1-fea.html)
A.2: Develop and Maintain Enterprise Data Model	A formal data modeling and design process should be followed for the navigatOR database beginning with a conceptual model, moving to logical data modeling and then to physical design for selected data elements. For navigatOR, the recommended best practices are: • A.2a: Create a General Comprehensive Logical Data Model for all geographic data to be included in navigatOR. • A.2b: Prepare a Conceptual Database Model for Framework Geographic Data • A.2c: Prepare a Detailed Logical Model for high-priority Framework Data (for selected elements for which database development or deployment is imminent). • A.2d: Use the logical data models in A.2c as the basis for the physical design. • A.2e: Use applicable and approved existing data models and template designs (see reference for FGDC and ESRI) as the basis for design work. • A.2f: Use automated tools for data modeling and design when database complexity warrants use of such tools.	References: A, C, L, M, O Other Sources: Template data models available from: a) FGDC: www.fgdc.gov/standards b) ESRI Template data models: http://support.esri.com/index.cfm?fa=downloads.dataModels.gateway c) *GIS Utility Conceptual Database Design*

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
A. Database Organization and Design (continued)		
A.3: Define/Maintain Data Format, Content, Quality Standards	Building on current policies and procedures established and overseen by the Oregon Geographic Information Council (OGIC) and the Geospatial Enterprise Office, formalize procedures for defining data standards for Framework themes. Develop a more complete set of data standards and keep them updated over time. In addition, publish and communicate these standards in a manner that makes them accessible and easy to understand by users and developers, including easy Web-based access. Data standards cover the following key areas: • Compliance with database design (map feature and attribute content and format) • Spatial reference system and coordinates • Spatial data import and export • Data coding and classification • Graphic integrity • Completeness • Attribute accuracy • Map position accuracy • Data resolution • Map presentation and product standards (symbology, annotation, map layout, etc.). In addition to standards for fully compliant Framework Themes, identify a less demanding set of standards for non-Framework data—data that does not fully adhere, in content or format, with fully compliant Framework data but which is available to navigatOR participants and may be accessed through navigatOR.	References: C, D, E, F, G, H, I, N, O, P, Q Other Sources: a) *GIS Utility Conceptual Database Design* (Deliverable 4C) b) OGIC Approved standards: www.oregon.gov/DAS/IRMD/GEO/standards/standards.shtml
A.4: Identify Volatility, Trigger, and Source for All Data Update	Each of the 14 Framework Themes listed above includes several data elements. It is important to define detailed information about each of these elements in order to set up an effective database update program. Each of these elements has their own "volatility" (frequency of change) and sources that will drive the update process. These are defined in preliminary form in the *Conceptual Design*. This must be completed in more detail as a step in operational planning for a data update program. A table that includes the name of the element, a "volatility category," and ideal update frequency should be completed and updated, as source information is available.	References: E, F, G, J, Other Sources: a) *GIS Utility Conceptual Database Design* (Deliverable 4C)

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
A. Database Organization and Design (continued)		
A.5: Define Data Format Rules	Data update and quality control procedures should make use of encoded rules that govern the format and "behavior" of map features. ESRI software allows for the definition of rules that are associated with specific map features and which can be used in applications in data validation during data capture and editing. Specific rules will depend on the type of map feature and user requirements. The main categories of feature-based rules include: • GIS feature placement rules • Feature topology rules • GIS feature connectivity rules • Map annotation rules.	References: A, G, H, N Other Sources: a) Appendix B
A.6: Develop and Implement Efficient Physical Design	Use a firm data model as the basis to prepare a physical database design for implementation as an ArcSDE Geodatabase. Include all Framework data elements for which data population is planned.	References: A, H, L, N, O Other Sources: a) Appendix B
A.7: Design Standards-based Metadata	For all Framework Data Themes, follow the FGDC Content Standard for Geospatial Metadata as approved by OGIC. See the site, www.oregon.gov/DAS/IRMD/GEO/standards/standards.shtml, for the accepted OGIC metadata standard and the document, http://egov.oregon.gov/DAS/IRMD/GEO/standards/docs/Metadata_Opportunity.pdf for metadata information. The Oregon standard calls for at least the use of the FGDC mandatory metadata items. The implementation of this standard should adhere to the following basic guidelines: A.7a: All populated Framework Data elements should have "data set" (i.e., ArcSDE Feature Class) level metadata, including at least the mandatory elements defined by the FGDC standard. A.7b: Where appropriate and feasible, include detailed feature and element level metadata to track quality and lineage information at this level of detail. A.7c: Use, to the greatest extent possible, packaged software for metadata update and management—including ESRI's ArcCatalog. A.7d: Define minimal requirements for metadata for non-Framework, including non-GIS data (e.g., geographically referenced documents). NOTE: It is an acceptable practice to define specific metadata profiles, based on the standard, but including sets of metadata elements that are specific to certain themes or data elements. This practice is used by the FGDC and is recommended for Oregon—to address specific user needs for metadata associated with particular data themes.	References: A, C, D, H, N, Q, R Other Sources: a) *GIS Utility Conceptual Database Design* (Deliverable 4C)

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
A. Database Organization and Design (continued)		
A.8: Define Mapping Rules to Support Conversion and Update	As part of the detailed planning for data conversion and data update, prepare and maintain a manual defining "mapping rules." Mapping rules are directions to data developers or updaters on how to handle actual data capture or update instances that may not be clear. Often, this addresses cases where source material is not entirely clear, there is missing information, or when map feature position or annotation on sources deviates from conversion specifications.	References: F, G, J
A.9: Design GIS Databases to Support External Integration	The data modeling and database design processes should be carried out within the context of external systems—and the need for future integration between GIS and "non-GIS" databases. The modeling and design work should identify key systems that are likely candidates for integration and should include data elements that support integration (i.e., secondary database keys in the GIS database that allow for direct database joins or data import/export from and to the GIS).	References: A, N, O
A.10: Establish and Employ an Approach for Temporal Data Management	Keep a record of all database update transactions (posting of updated data to the navigatOR database). At a minimum, this should be a text-based log that records the type of change (addition, deletion, change of a map feature or attribute) with the information on the time and the party responsible for data update. For each Framework Data Theme (and each element), determine the need for capturing database changes (actual graphic and attribute edits), and implement selective and automated means to do this using ArcSDE tools.	References: O
A.11: Have Process for Submittal and Execution of Database Design Changes	Allow database design changes to occur if there is a business case to support such changes. This involves modifying the data model and physical database design for any of the following circumstances: a) addition of new data (new map features or attributes), b) deletion of data previously included in the design, or c) modification to the format of map features and elements already in the design. Have a clear procedure allowing users or developers to submit a request for a change to be followed by review, approval or denial, and execution of the change. Also have a process in place to make modifications to existing data that might be necessary to accommodate the change.	References: N

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
B. Database Processes		
B1: Prepare a Long-term Operational Plan for Data Conversion	To guide long-term development of Framework Data, develop an operational plan for data conversion. The plan should include tasks, a schedule, resource requirements, and roles/responsibilities. It will be used as the basis for budgeting and overall initiation and management of multiple data conversion projects.	References: E, M, N, P, R
B2: Establish Project Management Procedures for Data Conversion	For all data conversion projects requiring the allocation of significant staff and monetary resources, establish clear technical specifications and project management procedures for work of in-house and contracted personnel. Follow project management guidelines of the Project Management Institute (PMI). This includes assigning a qualified project manager and team, making and following a clear plan, tracking and reporting on progress, identifying and managing risk, and putting in place effective quality assurance review procedures.	References: C, H, D, E, F, I, P, R
B3: Define and Document Workflows for Update Processes	For each Framework Theme (including each element), prepare an update workflow process that defines the steps for data update. This workflow should include the following: a) identification of the "trigger" for an update transaction (the event or circumstance that occurs to initiate an update), b) description and diagram of the major steps in the update process, c) identification of the sources used for the update, and d) identification of the parties responsible for the update steps (see C4).	References: C, H, D, E, F, I, P, R
B4: Acquire/Build Effective Tools for Database Update and QC	Design, acquire, and develop applications for database update and quality control that occur during the update process. The specific applications will vary widely depending on the type of data, the nature of the update, the source of the update, and the party responsible for the update. Data update applications should use, to the greatest extent possible, off-the-shelf tools and software, with customizing carried out when necessary to make them efficient for the user. Many (perhaps most) updates will be performed in batch mode and then posted to a permanent database for access after quality assurance steps, but some updates may be allowed interactively using direct, Web-based data edits.	References: D, E, F, H, I, J, K, P
B5: Set up and Follow Procedures for DB Update during Conversion Projects	Establish a process to manage database changes that occur during a major database development effort. The specific type of process will depend on the type of data and its volatility. At a minimum, this should include detailed logging of changes that occur in the area of conversion so that these changes can be included in the final delivered database.	References: J, K, P

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
B. Database Processes (continued)		
B6: Set up an Effective Process for Submittal of Updates	Design and document a process for batch data submittals (updated data) for data stewards that have been assigned database update responsibility. This includes an easy to use Web-based tool for submittal of data updates and associated metadata.	References: H, N, R
B7: Set up a Process and Tools for Final QA and Posting	The administrators of the Core navigatOR Data Center and individual data stewards should design and document processes (for each data theme and element,) for quality assurance, acceptance, and final posting of data for access by navigatOR users. The QA process should include automated checks on data content, format, and quality, and, in some cases, visual spot checks. QA processes should be applied to data received from a designated steward or a contractor. After QA is performed and the data is accepted, it is posted for general access.	References: H, N, P, R
B8: Acquire Updated Data from Commercial Source	Explore and act on all opportunities to purchase updated data or data update services from a commercial source. Weigh important trade-offs on cost, timing, and data quality, and select qualified commercial providers when it is beneficial to do so (as opposed to generating or updating data using in-house staff).	
C. Organizational Responsibilities and Resources		
C1: Organize and Sustain an Active Framework Implementation Team	Augment the role and responsibility of the Framework Implementation Team (FIT). This includes confirming a clear lead person and organization and assignment of responsibilities for defining data content and format for each data element for all 14 Framework Themes. Ensure a well-defined and documented role for the FIT and its organizational relationship with the navigatOR governing board and management unit.	References: B, C, N
C2: Assign a Formal Data Standards Body	Create a formal standards body with an aggressive agenda to complete data content, format, and quality standards for high-priority data themes. The standards body should operate under the oversight of the GIS Utility governing board and should use a well-documented process for standards proposal, review, and approval.	References: C, H, N, Q
C3: Assign Project Teams and Management for Data Conversion	For each data conversion effort initiated for Framework database development, a project manager and project team should be assembled to manage the conversion work and to carry out designated elements of the project, including a) source material preparation, b) data capture, and c) quality control and quality assurance.	References: C, M, N

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
C. Organizational Responsibilities and Resources		
C4: Establish Formal Agreements with GIS Utility Participants	Establish a formal registration process for all users seeking access to navigatOR data and on-line services. This would be a no-cost, simple registration process (best implemented on-line) that would allow navigatOR management to have basic information about the user community. For data stewards, prepare and ratify a standard data sharing agreement that defines terms for provision and use of data.	References: B, C, D, N, R Other Sources: a) *GIS Utility Conceptual Database Design* (Deliverable 4C)
C4: Assign Formal Organizational Responsibilities for Data Update	For each data element for all 14 Framework Themes, assign update responsibilities to one or more organizations in a manner that is clear and unambiguous. Identify a specific timing for update and the associated workflow and sources (see B3).	References: H, N, R
C5: Assign Resources for GIS Utility Database Administration	Assign well-qualified staff to the navigatOR management unit for database administration tasks (user accounts and data access, security, data back-up, etc.) Appropriate staff should be assigned to the navigatOR management unit at all organizations that have a major role in updating and maintaining data that will be accessed through navigatOR.	References: A, H
D. Database Administration and Accessibility		
D1: Establish and Track Data Security and Access	Set up access rights for navigatOR users and developers that match their need for access. Maintain access accounts over time, and monitor and log access on a regular basis.	References: A, H, R Other Sources: a) Appendix B
D2: Establish Database Back-up and Recovery Procedures	Industry-standard practices for protection of GIS data should be instituted at all locations where navigatOR data is being maintained and updated. This includes a) on-site daily back up, b) off-site storage of data back-ups on stable media, and c) effective power loss and disaster recovery. For the Core navigatOR Data Center, standards and procedures established for the new consolidated State data center should be applied to navigatOR. Similar procedures for other sites at which data is maintained should be put in place.	References: A, D, C, H Other Sources: a) Appendix B.4
D3: Track Database Changes and Notify Users	Update transactions should be documented and reported on a regular basis (e.g., weekly) and distributed in a clear way (Web posting) to navigatOR users.	References: H, G, R

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

Table 3-1: Summary of Best Practices for Data Stewardship (continued)

Best Practice	Explanation	References and Sources for More Detail*
D. Database Administration and Accessibility (continued)		
D4: Establish, Monitor, and Respond to Legal Impacts on Data Access	Legal and policy constraints on access to and distribution of navigatOR data should be clearly defined. Potential access/distribution limitations may relate to regulatory controls, privacy restrictions, security, and limitations imposed by specific agreements or contracts with data suppliers (public or private sector). Such restrictions should be captured as part of the data set metadata and should be modified as changes occur.	References: B, C
D5: Set up a Process and Mechanism for Data Export	A potential service of navigatOR should be to provide geographic data to users in a translated format (a different spatial file format than its native format as stored in navigatOR). This would include industry standard formats involving translation using off-the-shelf software. Appropriate software tools and procedures should be put in place to provide this translation/export service in an efficient manner. Ideally, this would be implemented as a Web-based service, where, at a minimum, users could request data sets on-line and have them processed by staff. The possibility exists to automate most of this process with a Web application that allows data to be extracted, translated, and downloaded—with little staff intervention.	
D6: Provide Technical Support to Users	Provide “help desk” services to users seeking to access data and services from navigatOR. This includes qualified technical staff available to help new users gain access to navigatOR and ongoing support when users have questions. This technical support should employ helpdesk software that allows users to submit questions on-line and often obtain an on-line response and resolution to the problem. The helpdesk software and administration procedures should be used to log all requests and to report on the history of user help requests.	References: B, C

*See the list of references in Appendix A and the elaboration of selected topics in Appendix B.

APPENDIX A
REFERENCE SOURCES FOR GIS DATA STEWARDSHIP BEST PRACTICES

APPENDIX A
REFERENCE SOURCES FOR GIS DATA STEWARDSHIP BEST PRACTICES

A	*System Design Strategies*, ESRI Technical Reference Document, June 2005, www.esri.com/systemsint/kbase/strategies.html
B	URISA National Geographic Information Cooperation, Coordination, Collaboration Task Force (3CTF), 2002-2004 Report & Recommendations, April 2005
C	*The National Map Best Practices*, draft report prepared by a USGS-sponsored cross-sector working group, March 2005
D	*Geographic Information Policies and Guidelines of the Ohio Department of Natural Resources*, GIMS Policy and Planning Committee, and the Ohio DNR Office of Information Technology, December 2004, www.dnr.state.oh.us/gims/includes/GIMS_Standards04.pdf
E	*GIS Best Practices White Paper*, PlanGraphics, Inc., project report for the Sichuan Urban Environment Project Office, March 31, 2003
F	"Geodatabase Quality Control—More Important than Ever," ESRI International User Conference, 2003
G	"Data Integrity and Validation for Enterprise-wide ESRI Implementations," ESRI International User Conference, 2002
H	*Best Practices for GIS within King County* (WA), King County GIS Technical Committee, October 2002
I	"Using Digital Quality Assurance Routines for Checking Converted Data," ESRI International User Conference, 2002
J	"Transaction Management and Versioning for Enterprise-wide ESRI Implementations," ESRI International User Conference, 2002
K	"GIS Quality Assurance-Data Acceptance Specifications and Quality Control," URISA Annual Conference Proceedings, Orlando, 2000
L	*Data Modeling,* Quick Study publication published by URISA, 2000, www.urisa.org/store/data_modeling_qs.htm
M	*Geospatial Data Infrastructure: Concepts, Cases, and Good Practice, Oxford University Press*, ISBN: 0-19-823381-7, 2000
N	Florida's Geographic Data Stewardship Program, Florida Geographic Information Board, December 1999
O	*Modeling our World—the ESRI Guide to Geodatabase Design*, ESRI Press, 1999
P	"GIS Data Conversion Quality Assurance and Database Acceptance—An Overview of Concepts and Procedures, PlanGraphics whitepaper, 1999
Q	*Spatial Information Technology Standards and System Integration*, published by URISA, December 1997
R	"Spatial Data Maintenance at the City of Calgary," URISA Annual Conference Proceedings, San Antonio, 1995

APPENDIX B
ELABORATION ON DATA STEWARDSHIP BEST PRACTICES

APPENDIX B
ELABORATION ON DATA STEWARDSHIP BEST PRACTICES

B.1 FEATURE-BASED RULES

Feature-based rules are encoded and associated with specific map features. These rules will govern certain aspects of the position and spatial relationships of the features and can be used as a basis for building data capture and data update processes and validating results. Common categories of feature rules include:

- Feature Placement Rules: Rules that govern the horizontal placement and stored coordinates for certain features in the geodatabase relative to other features. One example is a rule that assigns a coordinate location for a manhole, relative to a base map or orthoimage, when its actual coordinate gathered from GPS is more precise and accurate but would show an obvious position error relative to the existing base (e.g., manhole represented as being located outside a street curb when its actual location is within the street).

- Feature Connectivity Rules: Networks have established connectivity requirements that define what types of point and line features can be physically attached. Network in the geodatabase should reflect the actual conditions of the network and should not show "illegal" types of connections (e.g., 8" pipe connected to 6" pipe without the proper type of reducer fitting"). These rules should be formally defined in the geodatabase as part of a geometric network and used in data capture and update to ensure data quality.

- Topology Rules: Rules can be established and stored in the geodatabase that govern or restrict certain types of spatial relationships among features. These rules may relate to adjacency and overlap of polygon features, overlap and intersection of line features, and various types of line-in-polygon, point-in-polygon, and point-on-line spatial relationships. Such topology rules are used in data update and quality control, as well as spatial analysis applications.

- Annotation Rules: As a general rule, annotation and labeling of map features will be attribute-driven (label values dynamically accessed from attributes in the GIS database and displayed interactively). For certain map products, stored map annotation classes exist in which the position and appearance of the annotation labels are adjusted and customized. Annotation rules are a mix of cartographic design standards, geodatabase configuration issues, and the functional requirements for their use.

B.2 MAJOR GEODATABASE COMPONENTS

The geodatabase is built on an underlying relational database management system (RDBMS) such as Oracle or SQL Server. However, in an effort to make a uniform user interface across multiple database products, ESRI has chosen to hide the familiar database components. Instead, these have been replaced with geodatabase equivalents, which may have similar names but which work quite differently. The most important of these are briefly described below, grouped by function.

Data is stored within a geodatabase using three fundamental components:

1. An **object class** is the ArcSDE equivalent of the RDBMS table containing attributes about a GIS feature. For instance, an object class may be represented as an SQL Server table that stores physical attributes or maintenance data about a pipe segment or manhole.

2. A **feature class** is defined, based on the object class, to hold records that include a graphic attribute. Feature classes are just like object classes except that they additionally include the capability to store and manage a shape column that holds the graphic representation of that individual. For instance, a VALVES entity may be defined, within which each water valve in the city may be represented by a record that includes both descriptive attributes and a SHAPE column that holds a point graphic representing the location of the valve.

3. A **feature data set** is an element that defines a group of feature classes. These belong to the data set due to (a) having the same projection and coordinate system, and (b) participating in some shared rules-based behavior. They ensure consistency among the participating classes and are foundational for topologies and geometric networks.

Classes that hold the actual feature data in the geodatabase can be related to one another in two ways. Two classes can be related to one another using a relationship class. Some more-complex topological relationships can also be used to relate feature classes together, using the geometric network or the topology constructs.

Relationship classes are the geodatabase replacement for RDBMS relationship constraints. A relationship class says the two participating object and/or feature classes share a real-world relationship of such importance that the computer must ensure it is enforced when changes are made to the geodatabase contents.

The geodatabase version works differently than the relationship in an RDBMS, but it has most of the equivalent functionality. A relationship class can be simple or complex. A ***simple*** relationship class represents a situation in which each of the two records can exist independently of one another. In a ***composite*** relationship class, in contrast, the "parent" record can exist independently of the "child" records, but the child cannot.

Geometric networks are a means of consistently implementing behaviors important to managing connectivity among linear feature classes such as utility infrastructure or street networks. It is the geodatabase implementation of the coverage route. A geometric network is a feature data set containing a feature class of connected arcs and a feature class of the nodes that connect them, for instance, road segments and intersection nodes within a road network. Definition of a geometric network adds additional feature classes and imposes a set of predefined behaviors on the participating feature classes. It also allows imposition of **connectivity rules**, which are a rules-based means of managing connectivity logic.

Topology is a set of rules about how features within collections of feature classes share geometry. Topology imposes predefined behaviors on the feature classes. These allow specification of behavior rules for individual features, e.g., census tracts have planar topology. They also allow specifications of rules governing relationships between members of two classes, e.g., census tracts must be wholly contained within a single county. For more information about topology, see the 2003 ESRI whitepaper: *ArcGIS: Working with Geodatabase Topology (www.esri.com/library/whitepapers/pdfs/geodatabase-topology.pdf)*