Foothills Model Forest

Natural Resource Information Management Workshop

Christian Weik
Julie Dugas

September 6th, 2002
“I am not a teacher: only a fellow-traveler of whom you asked the way. I pointed ahead-ahead of myself as well as you.”

George Bernard Shaw

“Database Design for Mere Mortals”
– Michael Hernandez
Outline (Presentation)

- Workshop objectives
- Glossary
- Why data management?
- Existing practices
- Evolving role of GIS Personnel
- RDBMS vs. Spreadsheet
Outline (Presentation)...

- Database design process
- Mock design process
Workshop Objectives

- Emphasize the importance of good data management
- Promote awareness to integrate better practices
- Teach simple skills for design and reporting
- Emphasize software independence
- Spread the word
- Tell us how this could be better
- Note that some principles we mention may not yet be the accepted culture in all organizations
Glossary

- RDBMS
- Spreadsheet
- Row/record
- Column/field
- Database design
- Query/view
- Table
- Constraint
- Form
- Key
- Relationship
- Domain/validation
- Script/code
- Function
- Null
- Data integrity
- Metadata
- Parent table
- Child table
Why Data Management?

- “...15 percent to greater than 20 percent of a company's operating revenue is spent doing things to get around or fix data quality issues.”
- “...Data quality issues cost US businesses $600 billion per year.”
- FMF fish and watershed crew spent 2 man months in the last two years fixing data problems.
- 80% of GIS effort spent on data management.
Existing Practices

- Very little focus on data quality and management in natural resource management
- Little design effort
  - Little foresight for growth
  - Little consideration for long-term usability (if the owner leaves the organization)
- No integrity
- No documentation
- Poor spatial descriptions of data
Existing Practices...

- Local Level Indicators (LLI) example.
  - Inability to report at a landscape scale.
  - Different information, captured with poor documentation and very different formats.
- Enormous amount of time spent to compile relatively simple data.
Evolving Role of GIS Personnel

- Traditional role of GIS
  - Very isolated, specialized
  - Reactive to needs of users
  - Poor data planning and design resulting in poor information
  - Enormous amounts of time repairing poor data for reporting purposes
Evolving Role of GIS Personnel...

- Integration of non-spatial and spatial data management
- Proactive communication (inclusion) with project teams to define needs at project outset
- Facilitation of end-user application of GIS - less dependence on GIS team for simple GIS tasks (mapping, simple reporting)
Evolving Role of GIS Personnel...

- Changes driving new role of GIS
  - GIS group forced to deal with poor data
  - Simpler tools allowing end users to perform mapping, querying and simple analysis
  - Convergence of GIS and mainstream information technology (IT)
  - Storage of spatial information in business databases (eg. Oracle, SQL server etc.)
Evolving Role of GIS Personnel...

• Results
  – Better data; More accurate, timely and accessible
  – Less time spent on maintenance
  – Empowered users, less dependence on specialized staff
  – Minimized data duplication, single source
  – Data are more usable by other parties
Evolving Role of GIS Personnel...

- GIS (or IT) group can play a more active role in data management planning
  - Enable coordination with other groups (and data) in the organization
  - Enable users to take ownership of the data
RDBMS Vs. Spreadsheet

- Spreadsheets are excellent tools for reporting and summarizing data
- Spreadsheets are NOT designed to store relational databases
- Better data quality, access speed and versatility using a RDBMS
- Database integrity
  - The validity, consistency and accuracy of the data in a database
  - We will use database tools to ensure database integrity
End of sermon...
Database Design Process

- GIS can be involved in the formal database design process
- Mission statement and mission objectives
- Conduct the user needs analysis
  - Analyse the current database
  - Interviews with users
- Establish table structures
Database Design Process...

- Establish column list
- Establish key columns
- Establish table relationships
- Establish data integrity and business rules
- Define view and reports
- Perform preliminary testing
- Develop forms
Database Design Process...

- Develop reports
- Final testing
- Load historical information if required
- User training
- Apply an iterative approach to development, review, testing
  - Most problems occur in the gap between user needs and expectations and the database design
Effort

Time

Needs Analysis
Design
Development
Loading
Training
Maintenance

Reactive, exclusive
Proactive, integrated
Data entry

Integrity checks

RDBMS
Access
Oracle etc

ODBC/OLEDB

Excel

ArcView

Word

SPSS

TOAD/Oracle

Access
Mock Design Process

- Vegetation survey database
- Criteria as follows:
Mock Design Process...

- Vegetation plots
- Plots may be revisited
- Circular plot
- Capture tree, shrub, slope, aspect, photos
- Data will be key punched from field sheets
- Analysis will be performed using Excel and SAS

- Record data entry personnel
- Plots are in Willmore Wilderness area
Mock Design Process

- Define the mission statement
- Define mission objectives
- Conduct the user needs analysis
  - No existing database in place
  - Interviews with users
  - Discuss the business process with the ‘users’
- Establish table structures
- Establish column list
- Establish domain/validation lists
- Establish key columns
- Establish table relationships
Mock Design Process...

- Establish data integrity and business rules (briefly)
- Define views and reports (not today)
- Perform preliminary testing (not today)
- Develop forms (not today)
- Develop reports (not today)
Hands-on Creation of Database

- Create domain/validation tables
- Create tables
- Create columns
  - Create validation rules
- Establish a key field
  - Why use an auto-number primary key
  - Auto-numbers vs. Codes in domain tables
  - How to ensure no duplicates – indexes, unique constraints

Document columns, tables
Hands-on Creation of Database...

- Create relationships
  - Enforcing integrity
  - Cascading updates and deleted
  - Relationships in queries
Hands-on Query and Analysis

- Set database to confirm changes (options, edit, confirm)
- The query design grid
- Relationship properties
- Types of queries
  - Select
  - Update
  - Delete
  - Append
Hands-on Query and Analysis...

- Data summarizing
  - Max
  - Min
  - Average
  - Count

- Data manipulation functions (worksheet functions)
  - Calculated columns
  - Left, right, mid$
  - IIF statement
  - Distinct
Query Data Using Other Tools

- Direct export
- Open database connectivity (ODBC)
- Excel, ArcView 3.X and 8.X
Geodatabase

- Storage of spatial data in Microsoft Access (personal Geodatabase)
Geodatabase...

- Spatial (GIS) Data
- Non-Spatial Data
- Integrated Database (Geodatabase)
- Access
- Oracle
- SQL Server
Closing

- Questions?
- Feedback
- Complete evaluation form
- Follow-up survey