Geodatabase: Best Practices

Robert A. LeClair
Outline
Geodatabase: Best Practices

- Geodatabase Creation
- Data Ownership
- Data Model
- Data Configuration
- Geodatabase Behaviors
- Data Validation
- Extending
- Performance
Geodatabase Lifecycle

- Maintain
- Extend
- Manage Privileges
- Consume
- Create
- Configure
- Load Data
- Share Data
Geodatabase – Extending the Database

• Database
  - Tables and Native geometry types
  - Transaction/Security model

• Geodatabase
  - Information model, short/long transaction model, metadata tables, and an API/SDK

• Benefits
  - Improved data quality
  - Editing efficiency
  - Web model (WebGIS/services)
  - API (ArcMap) / SDK (ArcGIS Pro)
Geodatabase – Transaction Models

• **Transaction models:**
  - Single User - 1 editor at a time
  - Multi-user - versioning supports long transactions

• **Archiving**
  - Built on top of versioning
  - Enables historical reference of what has happened to data over time
  - Enables a time slider in ArcMap and ArcGIS Pro
Outline

Geodatabase: Best Practices

- Geodatabase Creation
  - Data Ownership
  - Data Model
  - Data Configuration
  - Geodatabase Behaviors
  - Data Validation
  - Extending
  - Performance
### Geodatabase Types

**What type to use?**
- Single User
  - File Geodatabase
- Multi-user
  - Enterprise Geodatabase
  - Database Server Desktop/Workgroup

**Considerations**
- Size
- Functionality
- Concurrent viewers/editors
- Storage platform
- Plans for publishing/services

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Personal geodatabase</th>
<th>File geodatabase</th>
<th>Desktop geodatabase</th>
<th>Workgroup geodatabase</th>
<th>Enterprise geodatabase</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functionality</strong></td>
<td>Original Desktop format</td>
<td>Improved Desktop format</td>
<td>Distributed data or project-level use</td>
<td>Departmental projects or small organizations</td>
<td>Large capacity and user base</td>
</tr>
<tr>
<td><strong>Storage mechanism</strong></td>
<td>Microsoft Access database (.mdb)</td>
<td>File folder; displays .gdb extension in ArcCatalog</td>
<td>Microsoft SQL Server Express</td>
<td>Microsoft SQL Server Express</td>
<td>SQL Server, Oracle, PostgreSQL, DB2, Informix</td>
</tr>
<tr>
<td><strong>Storage limit</strong></td>
<td>2 GB per geodatabase; effective limit ~500 MB</td>
<td>1 terabyte (TB) per object, configurable to 256 TB</td>
<td>10 GB per database server</td>
<td>10 GB per database server</td>
<td>Limited by relational database and hardware</td>
</tr>
<tr>
<td><strong>User limit</strong></td>
<td>One editor per database</td>
<td>One editor per object</td>
<td>Three concurrent users, one can edit</td>
<td>Ten concurrent users, all can edit</td>
<td>Unlimited</td>
</tr>
<tr>
<td><strong>Platform</strong></td>
<td>Windows</td>
<td>Any</td>
<td>Windows</td>
<td>Windows</td>
<td>Any</td>
</tr>
<tr>
<td><strong>Licensing:</strong></td>
<td>All ArcGIS for Desktop licenses can view, create, edit, and manage.</td>
<td>• All ArcGIS for Desktop licenses can view. • Standard or Advanced license required to create, edit, and manage.</td>
<td>Not required</td>
<td>Workgroup Edition</td>
<td>Enterprise Edition</td>
</tr>
<tr>
<td><strong>ArcGIS for Server</strong></td>
<td>Not required</td>
<td>Not required</td>
<td>Not required</td>
<td>Workgroup Edition</td>
<td>Enterprise Edition</td>
</tr>
</tbody>
</table>
Geodatabase Creation

• Single user geodatabases
  - Create File GDB

• Multiuser geodatabases
  - Create Enterprise Geodatabase
    - All in one (creates database, geodatabase admin account, etc.)
    - Requires access to DBA account
  - Enable Enterprise Geodatabase
    - No access to DBA account
    - Two steps – setup database and then enable the geodatabase
Best Practice - Geodatabase Creation

• For local geodatabases
  - Use the “Create File Geodatabase” GP tool – supported across the platform

• For Enterprise Geodatabases
  - Use the “Create Enterprise Geodatabase” GP tool
  - Fallback to the “Enable Enterprise Geodatabase” GP tool
Enterprise Geodatabase Administration

• **Geodatabase Administration toolbox**
  - Geodatabase creation/upgrade

• **Geodatabase configuration tasks**
  - Logfile configuration
  - DBTUNE keyword customization

• **Database/Geodatabase admin tasks**
  - User/role creation
  - Rebuild indexes
  - Analyze Datasets (Update Statistics)
  - Change privileges
  - Compress (versioned geodatabase)
Outline
Geodatabase: Best Practices

- Geodatabase Creation
- **Data Ownership**
- Data Model
- Data Configuration
- Geodatabase Behaviors
- Data Validation
- Extending
- Performance
Data Ownership Considerations

- POLP – Principle of least privilege
  - Database Admin / Geodatabase Admin / Data Owner / End User

**Geodatabase admin responsibilities:**
- Ownership and maintenance on repository
- Manages the DEFAULT version
- Runs the COMPRESS GP tool
- Upgrades
- Changes to DBTUNE/log file configurations

**Data owner Responsibilities:**
- Manage privileges
- Manages indexes
- Manages fields
- Register data as versioned
Best Practices - Data Ownership

• Grant the fewest permissions to the fewest number of users

• Have a unique Geodatabase Admin user

• Have a non-Geodatabase Admin create data (i.e. Data Owner or End User)
Demo
Geodatabase Creation
Outline
Geodatabase: Best Practices

- Geodatabase Creation
- Data Ownership
- **Data Model**
- Data Configuration
- Geodatabase Behaviors
- Data Validation
- Extending
- Performance
Data Model

What is a Data Model?
- Collection of things in a Geodatabase
- Feature classes, attributes, subtypes, domains, etc.

Picking a Data Model
- What data are you representing
- ArcGIS Data Models
  - http://solutions.arcgis.com/
- Other published data models

Custom Data Model
- Is your data unique?
- Do you just not like an existing data model?
Best Practices – Data Model

Use an ArcGIS Data Model - http://solutions.arcgis.com/

Use a Data Model provided by a business partner

Build a custom Data Model (if):
- You have an understanding of your data
- You have an understanding of the geodatabase
- You have an understanding of how Data Model decisions will impact the entire Platform
- You follow the guidelines provided in the following section
Outline
Geodatabase: Best Practices

- Geodatabase Creation
- Data Ownership
- Data Model
- **Data Configuration**
  - Feature Datasets
  - Feature Classes
  - Fields
- Geodatabase Behaviors
- Data Validation
- Extending
- Performance
Data Configuration – Feature Datasets

What is a Feature Dataset?
- Logical container of feature classes
- Enforces a common Spatial References to all classes within
- Enforces a common database privilege level
- Required for Geodatabase controller datasets (i.e. topology, terrains, networks, etc.)
What is a Feature Class?

- A table which stores spatial data
- Stores attributes in fields about features
  - Many data types supported
- Single Geometry representation (point, line, polygon, annotation, dimensions, multipoints and multipatches)
- Used throughout the ArcGIS Platform
Best Practices – Data Configuration

Use a Feature Dataset when you want to:
- Use a Controller Dataset
- Group similar themed classes

Minimize the number of Feature Classes
- Group together similar features
- Find a balance between grouping and null or empty attributes

Field data types and field order:
Understand your data when choosing field types:
- Consider field order during creation
- Keep text fields short, expand later
- Will data be entered via web/mobile clients
Demo
Geodatabase Configuration
Outline
Geodatabase: Best Practices

- Geodatabase Creation
- Data Ownership
- Data Model
- Data Configuration
- **Geodatabase Behaviors**
  - Subtypes
  - Default values
  - Domains
  - Relationship Classes
- Data Validation
- Extending
- Performance
Subtypes

• Modeling multiple types of features in a single feature class

• Features that shares similar attributes

• Short or Long Integer field - short is almost always appropriate
Default Values

- Most fields in a feature class or table support a default value
  - This is a value entered during editing, if the user does not specify a value
  - Raster fields do not support default values

- Default values can be configured at the field level or subtype level

- Once subtypes are created – default values should be configured at subtype level
  - Subtype “Commercial” can have a default value of 2 for a field
  - Subtype “Other” can have a different default value for the same field
Domains

- Define a set of permissible values

Range Domains
- A range domain specifies a valid range of values for a numeric attribute (minimum and maximum valid values).
- May be applied to short/long integers, float/double and date attribute types

Coded Value Domains
- Apply to any attribute type.
- Uses a Code & Description
- In ArcGIS editing tools, drop downs are provided with valid descriptions
- Field calculation – code can be used

Validation will compare actual values against their domain to determine if the entries are correct
Best Practices – Geodatabase Behavior

• Subtypes
  - Use subtypes where possible, instead of adding additional feature classes
  - If features share a large majority of their attributes with other features
  - Group the features into a single class differentiated by subtype

• Default Values
  - Always setup default values to limit the <Null> entries in a table
  - Editing using web/mobile
  - Limits the amount of entry required

• Domains
  - Use Domains to simplify editing
  - Use Domains to make sure valid values are being entered
Relationship Classes

• A persisted relationship between objects in different feature classes or tables
• Visible using the attribute dialog in ArcMap and ArcGIS Pro
• Support rules for data integrity

• No relationship rules (constraints) = no validation
  - Business logic by defining constraints on the classes’ objects or features that is accessed during validation
  - Setup after relationship class creation
Best Practices – Relationship Classes

• Use relationship classes for those relationships that will enforce some data integrity

• Create rules to further ensure data integrity
  - These rules are not reactively checked

• Avoid creating relationship classes between two feature classes residing in separate feature datasets
Data Loading - Tools

• Loading new objects
  - Import/Export XML Workspace Document, Feature Class to Feature Class, Feature Class to Geodatabase GP Tools
  - Copy/Paste or Drag/Drop

• Updating existing objects:
  - Simple data loader vs. Object loader (only in ArcMap)
    - Object loader requires an edit session but triggers geodatabase behavior
    - Object loader trigger validation against geodatabase rules during load
    - Object load may be undone as it is inside an edit session
  - Append GP Tool (requires a target feature class or table to already exist)
Data Loading - Tools

- Load all data before registering as versioned – this limits records in delta tables.

- If you need to load data into a versioned system, consider unregistering as versioned prior to loading the data to prevent state doubling.
Demo
Geodatabase Behaviors and Data Loading
Attribute Editing while Editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
  - Attribute domains
  - Subtypes
  - Default values for attribute fields
  - Relationship classes
Attribute Editing while Editing

- Attribute domains
- Subtypes
- Default values for fields
- Relationship classes

The Geodatabase provides functionality to help maintain attribute quality while editing.
Attribute Editing while Editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
  - Attribute domains
  - Subtypes
  - Default values for fields
  - Relationship classes
Attribute Editing while Editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
  - Attribute domains
  - Subtypes
  - Default values for fields
  - Relationship classes

- Attribute Assistant
  - Editor add-in/extension from Solutions team
  - Reduced mouse clicks during editing
  - Accurate attribute info
Attribute Editing while Editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
  - Attribute domains
  - Subtypes
  - Default values for fields
  - Relationship classes
Data Validation

• Verifies an object against any rules that are defined for objects class or subtype

• Even with geodatabase behavior such as attribute domains there is still a need to perform ‘Validate Features’ to review data that violates rules

• Examples of editing operations that require validation:
  - Editing of Range Domains
  - Editing Attribute domains (coded value/range) via field calculator
  - Bulk loading of data into existing feature classes

Validate Features – currently only in ArcMap
ArcGIS Pro – use Data Reviewer Extension
Data Validation

- **Multiple Invalid Features:**
  - Warning/message received will not include the specific reason
    - Tip: make a single sub selection and re-validate
  - Validation process stops at the first rule where a feature is found invalid
    - Always re-validate
  - **Data Reviewer - ArcMap/ArcGIS Pro Extension** for a more interactive validation process
    - **Attribute Assistant** - editor add-in/extension from Solutions team
Demo
Editing and Data Validation
Maintain – Enterprise Administration

- In order to maximally compress the geodatabase:
  - Remove locks (disconnect users, stop the services)
  - Reconcile, Post, and delete Versions
  - Synchronize any replicas
Maintain – Enterprise Administration

• In order to maximally compress the geodatabase:
  - Remove locks (disconnect users, stop the services)
  - Reconcile, Post, and delete Versions
  - Synchronize any replicas

• Rebuild indexes / Analyze datasets
  - Data owner can analyze the tables they own
  - GDB Admin can analyze all tables, including system tables
Demo
Database Maintenance
Performance

- Fat/Wide feature classes vs. narrow feature classes
- Subtypes vs. new feature classes
- Lumping feature classes together
- Relationship class messaging – turn off for simple relationship classes
- Layers in a map
- Prototyping – What you are actually going to do with the data!
- Query Definition
Please Take Our Survey on the Esri Events App!

Download the Esri Events app and find your event

Select the session you attended

Scroll down to find the survey

Complete Answers and Select “Submit”
Print Your Certificate of Attendance
Print stations located in the 140 Concourse

Monday
12:30 PM – 6:30 PM  
GIS Solutions Expo,  
Hall B

5:15 PM – 6:30 PM  
Expo Social,  
Hall B

Tuesday
10:45 AM – 5:15 PM  
GIS Solutions Expo,  
Hall B

6:30 PM – 9:30 PM  
Networking Reception,  
Smithsonian National Air and Space Museum
Thank-you and Questions