Geodatabase Best Practices

Dave Crawford | Erik Hoel
Geodatabase best practices - outline

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance
Geodatabase - extending the database

• Database
  - Tables and native geometry types
  - Transaction / security model

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

• Benefits
  - Improved data quality
  - Editing efficiency
  - Web model
  - API / SDK
Geodatabase - transaction models

• Transaction models
  - Single User - one 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  - Enables historical reference of what has happened to data over time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 Pro
Geodatabase best practices

- **Geodatabase creation**
  - Data ownership
  - Data model
  - Data configuration
  - Geodatabase behaviors
  - Data integrity and validation
  - Administration
  - 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
Geodatabase creation

• Single user geodatabases
  - Create File Geodatabase

• Multi-user 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
Geodatabase creation - best practices

• For File Geodatabases
  - Use the geoprocessing tool - supported across the platform

• For Enterprise Geodatabases
  - Use the Create Enterprise Geodatabase geoprocessing tool
    - Oracle - prerequisite of using an existing database
  - Fallback to the Enable Enterprise Geodatabase geoprocessing tool
Geodatabase best practices

- Geodatabase creation
- **Data ownership**
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance
Data ownership considerations

• POLP - principle of least privilege
  - Database admin / geodatabase admin / data owner / end users

Geodatabase admin responsibilities:
• Ownership and maintenance on system tables
• Compress and DEFAULT version
• Upgrades
• Changes to DBTUNE/log file configurations

Data owner responsibilities:
• Manage privileges
• Schema changes
• Register data as versioned
Data ownership - best practices

- Grant the fewest permissions to the fewest number of users
- Have a unique Geodatabase Admin user
- Have a non-Geodatabase Admin create data
DEMO – Geodatabase creation and configuration
Geodatabase best practices

- Geodatabase creation
- Data ownership
- **Data model**
  - Data configuration
  - Geodatabase behaviors
  - Data integrity and 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?
Data model - best practices

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
Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- **Data configuration**
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance
Data Configuration - Prototype everything!

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

- Use a Feature Dataset when you want to:
  - use a Controller Dataset
  - enforce common spatial reference / privileges
  - group similar themed classes

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

- Geodatabase creation
- Data ownership
- Data model
- Data configuration

**Geodatabase behaviors**
- Data integrity and validation
- Administration
- 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 permit values to be valid in a field when they are within the specified range

• Coded value domains
  - Description (what you see in ArcMap or Pro) and code (the value that is stored)
  - 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
Geodatabase behavior - best practices

• 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
Relationship classes - best practices

- 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)
  - Copy tool or manual copy/paste

• Updating existing objects:
  - Simple data loader vs. object loader
    - 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 (requires a destination feature class or table to already exist)
Data loading - best practice

- Load all data before registering as versioned – this limits records in deltas

- If you need to load data into a versioned system, consider unregistering as versioned prior to loading the data
DEMO – Geodatabase behaviors and data loading
Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors

**Data integrity and validation**

- Administration
- Extending
- Performance
Attribute integrity while editing

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

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

- The Geodatabase provides functionality to help maintain attribute quality while editing
  - Attribute domains
  - Subtypes
  - Default values for fields
  - Relationship classes
Attribute integrity 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 integrity 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
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/Pro extension for a more interactive validation process
  Attribute Assistant - editor add-in/extension from Solutions team
DEMO - Editing and data validation
Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation

**Administration**

- Extending
- Performance
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)
Enterprise geodatabase administration

- Rebuild indexes / analyze datasets
  - Data owner can analyze the tables they own
  - GDB Admin can analyze all tables, including system tables
  - Defaults to rebuild delta tables only
Enterprise geodatabase administration

• In order to maximally compress the geodatabase:
  - Remove locks (disconnect users, stop services consuming gdb datasets)
  - Reconcile, post, and delete versions
  - Synchronize any replicas

• Reconcile Versions tool
  - Batch reconciles when ‘Post Versions After Reconcile’ option is checked
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 Def
Geodatabase best practices - outline

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance
Please take our Survey
Your feedback allows us to help maintain high standards and to help presenters

Find your event in the Esri Events App

Find the session you want to review

Scroll down to the bottom of the session

Answer survey questions and submit