Introduction to the Geodatabase

Eric Rodenberg
Workshop Agenda

• Welcome – 1:30
• Geodatabase Overview 1:40 – 2:30
• Hands on Tutorial 2:30 – 4:30
• Wrap Up and Questions 4:30 – 5:00

• No Scheduled Breaks – Please take breaks as needed after 2:30
Geodatabase Overview

• The Geodatabase
  – What is it?
  – Why use it?
  – What types are there?

• Inside the Geodatabase

• Advanced Behavior

• Editing Geodatabases

• Geodatabase Potpourri
What is the Geodatabase?

- **Core ArcGIS data model**
  - A comprehensive model for representing and managing GIS data

- **A physical store of geographic data**
  - Scalable storage model supported on different platforms

- **A transactional model for managing GIS workflows**

- **Set of COM components for accessing data**
Geodatabase Data Management Approach

• The geodatabase is built on an extended relational database.
  – Base relational model
  – Base short transaction model
  – Relational integrity
  – Reliability, Flexibility, Scalability
  – Supports continuous, large datasets

• Built on the simple feature model
  – Open access (OGC, C, COM, SQL)
Geodatabase Data Management Approach …

• Simple features + logic
  – All geographic data stored as tables in a DBMS
  – Extend functionality and data integrity
  – Functionality is consistent across DBMS’

• Application logic (software)
  – Works on standard DBMS tables
  – Implements GIS integrity and behavior
  – Business rules, topology, networks
    • Data Integrity
Geodatabase Data Management Approach …

• Editing and data compilation
  – Rich set of editing tools
  – Maintain spatial and attribute integrity
  – Undo and redo edits
  – Multiple users editing the same data

• Versioning work flows
  – Long transactions
  – Distributed data management
  – Archiving

• Robust, customizable framework
  – Build and manage your own specific GIS solution
Geodatabase Data Management

- Schema is defined in ArcCatalog
  - Define feature classes, datasets, relationships, etc
- Import and convert data from other formats
  - Shapefile
  - Coverage
  - CAD
  - Raster
- Copy and Paste
- Geodatabase XML Export / Import
  - For transferring Schema or Features and Schema
- Use an ESRI Data Model
  - Industry specific data models available
  - Copy geodatabase template
3 Types of Geodatabases

• **Personal Geodatabase**
  - Single user editing
  - Stored in MS Access
  - Size limit of 2 GB

• **File Geodatabase**
  - 1 TB per table
  - Reduced storage requirements

• **ArcSDE Geodatabase**
  - Stored in an enterprise DBMS
  - Supports multiuser editing via versioning
  - Requires ArcEditor or ArcInfo to edit
<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage format</td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td>Storage capacity</td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td>Supported O/S platform</td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td>Number of users</td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td>Distributed GDB functionality</td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
3 Types of Geodatabases...

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage format</td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td>Storage capacity</td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td>Supported O/S platform</td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td>Number of users</td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td>Distributed GDB functionality</td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
## 3 Types of Geodatabases...

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
Session Path

• The Geodatabase

• Inside the Geodatabase
  – Object class, Feature class, Raster dataset
  – Feature datasets
  – Validation rules
  – Domains, Subtypes, Relationship classes
  – Annotation, Dimensions
  – Exploring a Geodatabase DEMO

• Advanced Behavior

• Editing Geodatabases

• Geodatabase Potpourri
Inside the Geodatabase

• A geodatabase contains datasets

• Datasets represent collections of information with a real-world interpretation

• Types of geographic datasets:
  – Tables
  – Object classes, feature classes, relationship classes
  – Feature datasets
  – Networks, Topologies, Raster and cadastral datasets

• Datasets have associated information to help manage integrity, behavior, and interpretation
  – Domains, Relational integrity, Topology, Metadata
Inside the Geodatabase...

- Feature dataset
  - Feature class
    - Polygon
    - Line
    - Point
    - Annotation
    - Dimension
    - Route
  - Relationship class
  - Topology
  - Geometric network
  - Network dataset
  - Terrain

- Table

- Raster dataset

- Raster catalog

- Schematic dataset

- Survey dataset
  - Project folder
  - Project

- Toolbox
  - Tool
  - Model
  - Script

- Behavior
  - Attribute domains
  - Attribute defaults
  - Split/merge policy
  - Relationship rules
  - Connectivity rules
  - Topology rules
Objects and Object Classes

- Objects are entities with properties and behavior
- An object is an instance of an object class
- All objects in an object class have the same properties and behavior
- An object can be related to other objects via relationships

A row stores an Object
A table stores an ObjectClass
Features and Feature Classes

• Builds on the Relational Model
• A feature is a spatial object
• A feature is an instance of a feature class
• Extended the relational model with
  – Geometry attribute types

A feature class is a table of rows, where each row has a geographic column
Field data types

- The geodatabase supports **eight** field data types

<table>
<thead>
<tr>
<th>Data type</th>
<th>Bytes</th>
<th>Range / format / notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Integer</td>
<td>2</td>
<td>-32,768 to +32,767</td>
</tr>
<tr>
<td>Long Integer</td>
<td>4</td>
<td>-2,147,483,648 to +2,147,483,647</td>
</tr>
<tr>
<td>Float</td>
<td>4</td>
<td>About –3.4e38 to +1.2e38 (~7 significant digits)</td>
</tr>
<tr>
<td>Double</td>
<td>8</td>
<td>About –2.2e308 to +1.8e308 (~14 significant digits)</td>
</tr>
<tr>
<td>Text</td>
<td>varies</td>
<td>Up to ~64,000 characters</td>
</tr>
<tr>
<td>Date</td>
<td>8</td>
<td>mm/dd/yyyy hh:mm:ss am/pm</td>
</tr>
<tr>
<td>BLOB</td>
<td>varies</td>
<td>Store large binary content or other multimedia</td>
</tr>
<tr>
<td>Raster</td>
<td>varies</td>
<td>Store images</td>
</tr>
</tbody>
</table>

- Supported field data types are generic
  - Data types specific to an RDBMS are not supported
Geodatabase Supports Advanced Geometry

- Points, lines, polygons
  - Single and multipart features

- Text and surfaces

- Flexible coordinates
  - XY, Z, M

Feature with many parts

One record in feature class table
Geodatabase Raster Data

• Support for many formats
  – tiff, bmp, GRID

• Raster dataset
  – Separate rasters
  – Mosaicking

• Raster catalog
  – A collection of raster datasets
  – Accessed as one entity
  – Each member can be accessed as a raster dataset
  – Each member can have its own storage properties
  – Managed/Unmanaged
Feature Datasets

- A container object for other datasets
  - Same spatial reference

- Analogous to a coverage
  - Less restrictive

- Contain geometric networks and topologies
  - Optionally relationship classes

```
Subdivision
  ParcelCorner
  Parcel
  ParcelAnno
  LotLines
  Parcel_Top
  LotDimensions
  BoundryLines
```
Validation Rules

- Store attribute, connectivity, and relationship rules on objects as part of the geodatabase

- Predefined, parameter driven
  - Attribute range rule
  - Attribute set rule
  - Connectivity rule

- On demand

- Perform custom validation by writing code
Domains

• Describe the legal values of a field type
  – Used to ensure attribute integrity

• Defined at the geodatabase level

• Types of domains:
  – Range
    • A tree can have a height between 0 and 300 feet
    • A road can have between one and eight lanes
  – Coded Value
    • A tree can be of type oak, redwood, or palm
    • A road can be made of dirt, asphalt, or concrete
Subtypes

• Partition the objects in a class into like groups

• Defined at the class level

• Defined by the value of a subtype field
  – Have the same attribute\behavior schema
  – Can have different default values and domains for each field
  – Can define topology rules between subtypes
Relationship Classes

• An association between two object classes
  – A class may participate in multiple relationship classes

• Simple relationships

• Composite relationships
  – Related objects can message each other
  – Can trigger behavior (cascade delete, move to follow, custom, etc.)

• Associate rules with relationship classes
  – Each Parcel can have between 1 to 3 Buildings

<table>
<thead>
<tr>
<th>Parcel</th>
<th>Buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>Zone</td>
</tr>
<tr>
<td>28</td>
<td>Commercial</td>
</tr>
<tr>
<td>794</td>
<td>Residential</td>
</tr>
<tr>
<td>858</td>
<td>Residential</td>
</tr>
</tbody>
</table>
Annotation

- Annotation feature classes may be
  - Feature linked or Non-feature linked

- Composite relationship manages link

- Can store text as well as other graphics
  - Lines, arrows, boxes, etc…
Dimension Features

- Type of annotation that displays specific distances on a map
- Graphic features stored in a dimension feature class
- "Smart" feature
  - Special drawing
  - Special editing
Object Behavior

• You can:
  – Instantiate classes with predefined behavior. *(Dimensions and Annotation)*
  – Control the default value and acceptable values for any attribute in a class. *(Domains and Validation)*
  – Partition the objects in a class into like groups. *(Subtypes)*
  – Control the general and network relationships in which an object can participate. *(Relationship Classes)*

• Out of the Box in ArcGIS!
  – Configurable, no programming required
Exploring a Geodatabase Demo

- Explore a Geodatabase
  - Tables
  - Feature Classes
  - Subtypes
  - Domains
  - Relationship Classes
Session Path

• The Geodatabase

• Inside the Geodatabase

• Advanced Behavior
  – Geometric Networks
  – Network Datasets
  – Geodatabase Topology
  – Advanced behavior DEMO

• Editing Geodatabases

• Geodatabase Potpourri
Geometric Networks

• Used to model network systems

• Connectivity relationships between feature classes
  – Can associate connectivity rules with the network
  – Connectivity is based on geometric coincidence, *always live*

• Each feature class has a role in the network
  – A network may have multiple feature classes in the same role
Geometric Networks

• A geometric network is associated with a logical network
  – Each network feature is associated with one or more elements in the logical network

• Trace solvers on the logical network provide
  – Connectivity tracing, cycle detection, flow directions
  – Upstream/downstream tracing, Isolation tracing
Network Datasets

- Network designed for the transportation industry
- Does not replace the Geometric Network
- Multimodal
- Edges, Junctions & Turns
- Attributes
  - On-the-fly calculation of costs
  - Improves analysis
    - Cost, restriction, descriptor
Network Dataset Functionality

• Multimodal
  – Points span multiple connectivity groups
  – used to create connectivity between lines in different groups

• Turns
  – Turns do not alter connectivity, but traversability
    (e.g. U-Turn restriction)
Geodatabase Topology

• A topology manages a set of simple feature classes that share geometry

• Topology is used to
  – Integrate feature geometry
  – Validate features
  – Control editing tools
  – Define relationships between features
  – Ensure the quality of your data
Topological Integrity

• Topology defines integrity rules for associated feature classes
  – Participating feature classes / subtypes
  – Cluster tolerance, ranks and rules
    • Cluster Tolerance for XY and Z

• Rules are evaluated during Validation
  – Define rules when creating the Topology

• Violations of these rules are expressed as error features managed in the database as a part of the topology
  – Error and Exceptions
  – Examine and Fix errors in ArcMap
Topology Error Examples

- Rules enforced to maintain topological integrity
  - 25+ topology rules in ArcGIS
Editing with a Topology

• Editing creates a **dirty area**
  – Area has been edited and may contain errors
  – Can be symbolized

• Errors are found during **validation**
  – Errors have properties
    • What rule was violated
    • Which feature(s) created the error

• Your options:
  – Ignore the error
  – Mark as exception
  – Fix the error
Geodatabase Behavior Demo

• Explore a Geodatabase
  – Topology
  – Geometric Network
• The Geodatabase

• Inside the Geodatabase

• Advanced Behavior

• Editing Geodatabases
  – Transaction model
  – Geodatabase editing solutions
  – Versioning

• Geodatabase Potpourri
Editing Geodatabases

• ArcGIS datasets stored in the geodatabase are editable
  – Merge adjacent parcels in a topology
  – Add water mains to a network
  – Update land owners in a relationship class
  – Etc…

• There is a rich transaction model for editing in ArcGIS
  – Edits are performed in an edit session
    • Open session – edit – save edits / don’t save edits
  – A series of edit operations constitutes a transaction
    • Unit of work performed against the database
    • The transaction is either committed or rolled back
Editing Geodatabases...

• **Personal Geodatabases**
  - Single user, cubicle editing on small datasets
  - Multiple readers
  - Editing locks at geodatabase level
    • Two editors cannot edit within the same geodatabase at same time

• **File Geodatabase**
  - Single user, Workgroup editing on small to very large datasets
  - Multiple readers
  - Editing locks at the feature level
    • Two editors cannot edit the same object/feature class at same time
Editing Geodatabases…

• ArcSDE Geodatabases
  – Extend the transaction model with Versions
  – Enterprise level editing
  – Multiuser editing without locking
    • Unique isolated view of the geodatabase

• Benefits of versioned editing
  – Long Transactions
  – Undo / Redo
  – Archiving
  – Replication / mobile GIS
Session Path

- The Geodatabase
- Inside the Geodatabase
- Editing Geodatabases
- Advanced Behavior
- Geodatabase Potpourri
  - Terrains
  - Cartographic representations
  - Cadastral
Terrains

• Massive point datasets in a multi-resolution, on-the-fly generated TIN
  – Dataset for modeling 3D surfaces
  – Modeled within a feature dataset
  – User defined terrain (pyramid) levels
    • Different resolutions & vertical tolerances

• Requires 3D Analyst
  – Extension to define & edit
  – No license needed to view
Representations

• Property of a feature class
  – Stores info about feature symbology

• One feature class - multiple representations

• Rules and overrides

• Representation Management Toolset
Cadastral Editor

- Solution for parcel data management
  - Survey Analyst extension

- Uses COGO attributes and survey control to improve spatial accuracy

- Cadastral editing
  - Cadastral editor toolbar
  - Cadastral fabrics
    - Group layer with sublayers
    - Jobs

![Cadastral fabric data model](image-url)
Summary

• The Geodatabase
  – Data model, Storage, Transaction model, COM components

• Inside the Geodatabase
  – Datasets, Validation rules, data behavior and integrity

• Advanced Behavior
  – Geometric Networks, Network Datasets, and Topology

• Editing Geodatabases
  – Transaction model, Editing solutions, Versions

• Geodatabase Potpourri
  – Terrains, Representations, Cadastral
Hands on Tutorial

- Work at your own pace
- Take at least two breaks
- Each exercise should take no more than 15 minutes
- Please raise your hand if you have a question
- Please remain quiet as your neighbors may still be working
Hands On Tutorial

- Organizing your data in ArcCatalog
- Importing data into your Geodatabase
- Creating subtypes and attribute-domains
- Creating relationships between objects
- Building a geometric network
- Creating annotation
- Create Cartographic Representations
- Creating a topology
- Replicate and Synchronize the Geodatabase
QUESTIONS?