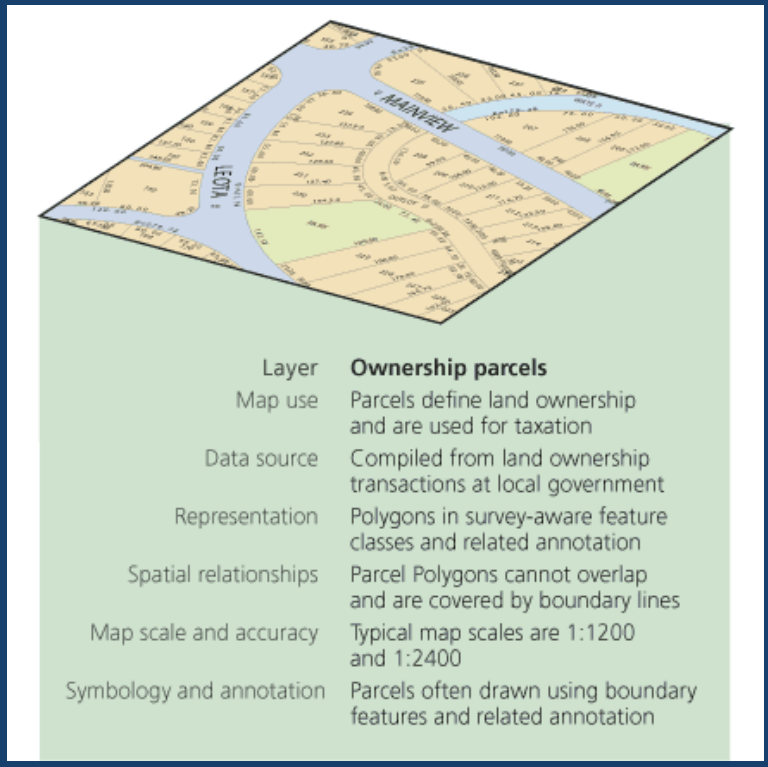
GEODATABASES

Sandeep Talasila, GISP

INTRODUCTION

- A collection of geographic datasets of various types held in a multiuser relational database management system (DBMS).
- A Geodatabase design is about identifying the thematic layers and specifying…
  - Representations (geometry)
  - Attributes (properties)
  - Relationships (dependencies)
  - Integrity Rules (behavior)



Layer	**Ownership parcels**
Map use	Parcels define land ownership and are used for taxation
Data source	Compiled from land ownership transactions at local government
Representation	Polygons in survey-aware feature classes and related annotation
Spatial relationships	Parcel Polygons cannot overlap and are covered by boundary lines
Map scale and accuracy	Typical map scales are 1:1200 and 1:2400
Symbology and annotation	Parcels often drawn using boundary features and related annotation

Source: ESRI

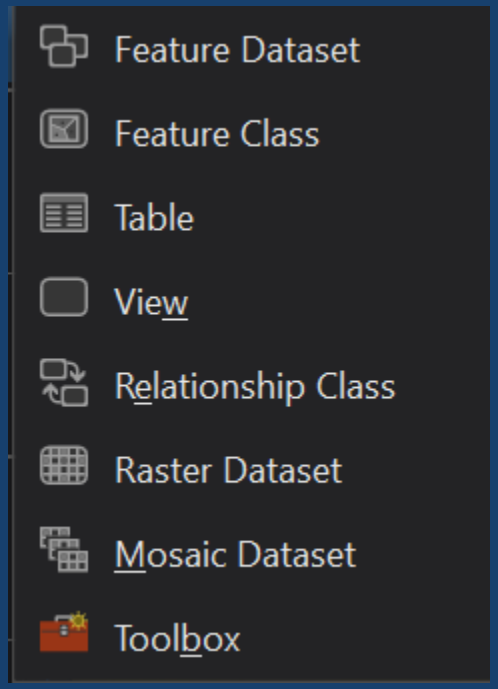
MORE…

- Native data storage and data management framework for ArcGIS.
- Object-Relational Database Management Systems (ORDBM) to store graphic and attribute data.
- Supports Multiuser editing through versioning.
- Implement subtypes and domains.
- Build relationships

TYPES OF GEODATABASES

- **File geodatabases** – Stored as folders in a file system. Each dataset is held as a file that can scale up to 1 TB in size. File extension .gdb
- **Personal geodatabases** – Stored within a Microsoft Access data file, which is limited in size to 2 GB, file extension .mdb
- **Enterprise geodatabases** – Stored in a relational database using Oracle, Microsoft SQL Server, IBM DB2, IBM Informix, or PostgreSQL. Uses ArcSDE technology. These multiuser geodatabases can be unlimited in size and numbers of users.

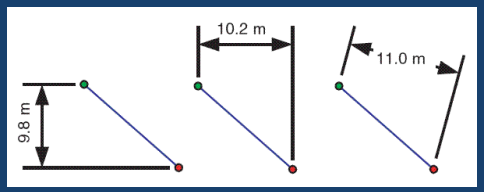
INSIDE A GEODATABASE…



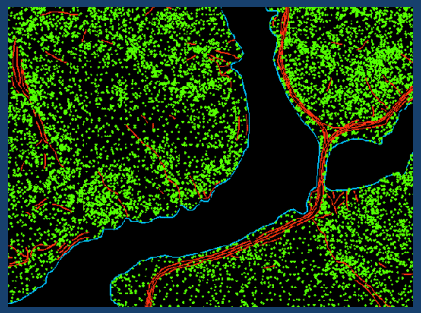
FEATURE CLASSES

- Points
- Lines
- Polygons
- Annotations
- Dimensions
- Multipoints
- Multipatches

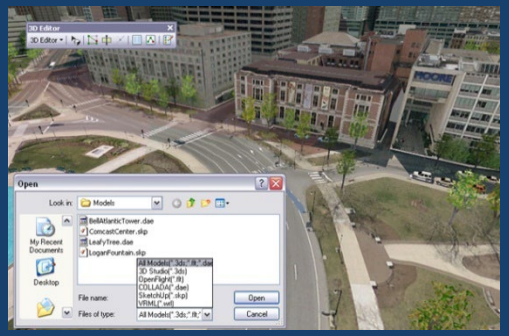
Annotations



Dimensions



Multipoints

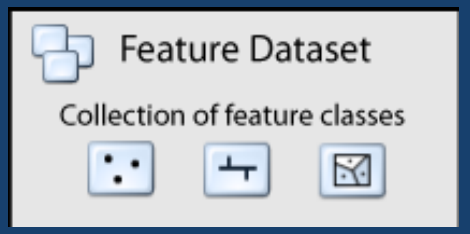


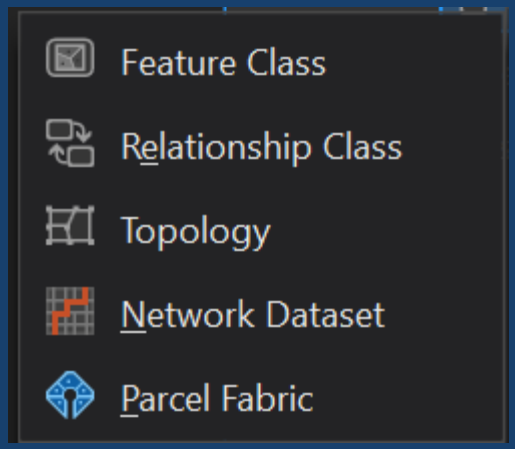
Multipatches

Image source: ESRI

FEATURE DATASETS

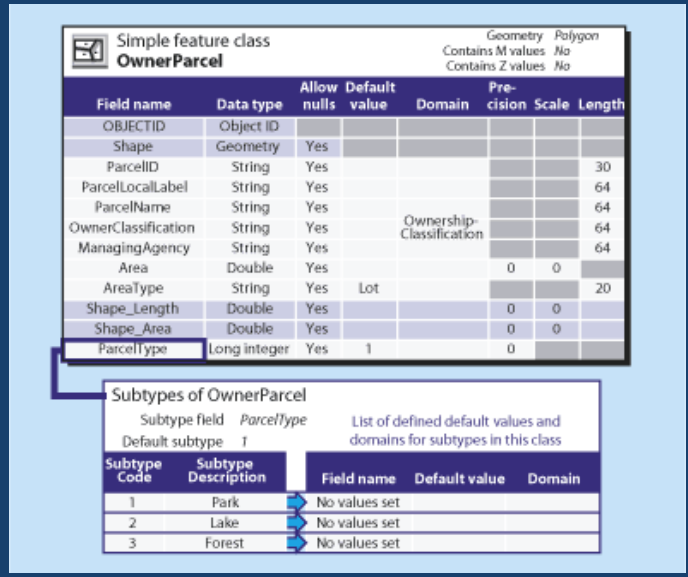
- A feature dataset is a collection of related feature classes that share a common coordinate system.
  - Used to spatially or thematically organize related feature classes
  - Can also be used to control access based on database privileges
  - Need to define spatial reference prior to organizing feature classes
- Controller/extension datasets
  - Topology
  - Terrain (3D Analyst)
  - Geometric Networks (Network Analyst)
  - Parcel Fabric





TABLES

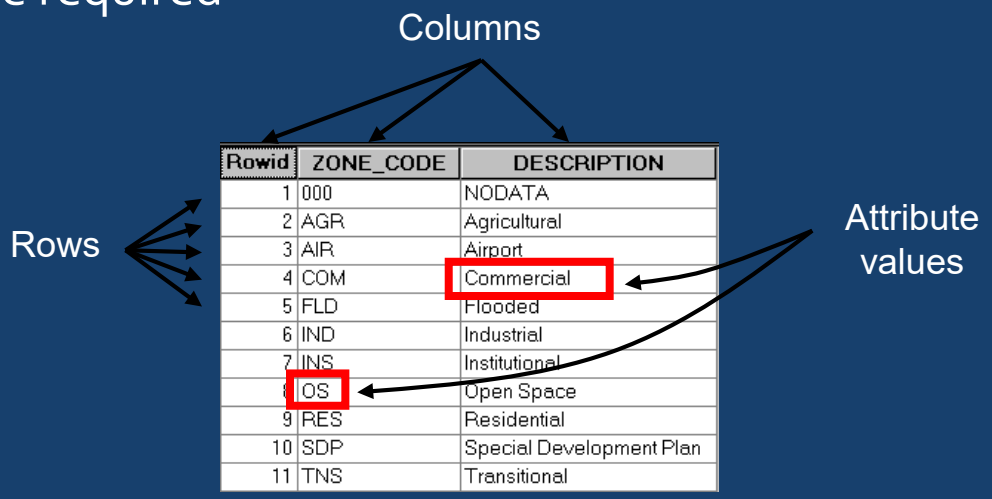
- Manage attributes
- Spatial and Non-spatial formats
- Each table contains rows with same columns
- Data Types:
  - Numbers: Long Integer, Short Integer, Float, Double
  - Text
  - Date
  - BLOB
  - Geometry
  - Identifiers: Object ID, Global IDs



https://desktop.arcgis.com/en/arcmap/latest/manage-data/geodatabases/arcgis-field-data-types.htm

TABLE ANATOMY

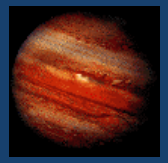
- Basic table properties
  - Records/rows and fields/columns
  - Column types can store numbers, text, dates
  - Unique column names are required



DATA FIELD TYPES

- Different field types store different kinds of values
- Choose the right field type for the right value
- Field types vary according to table format

Name: Jupiter
Moons: 16
Diameter: 142,984 km
Date of Comet Shoemaker-Levy impact: 7/16/1994
Rotation period: 9.8 hr



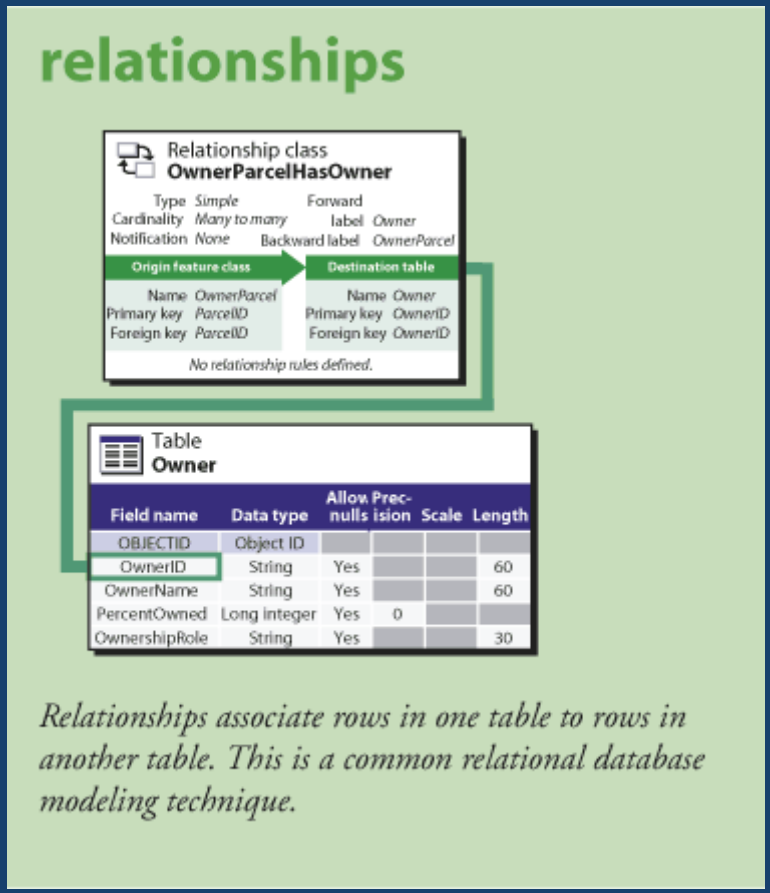
Text	Date	Short	Long	BLOB	Float
Jupiter	7/16/1994	16	142984		9.8

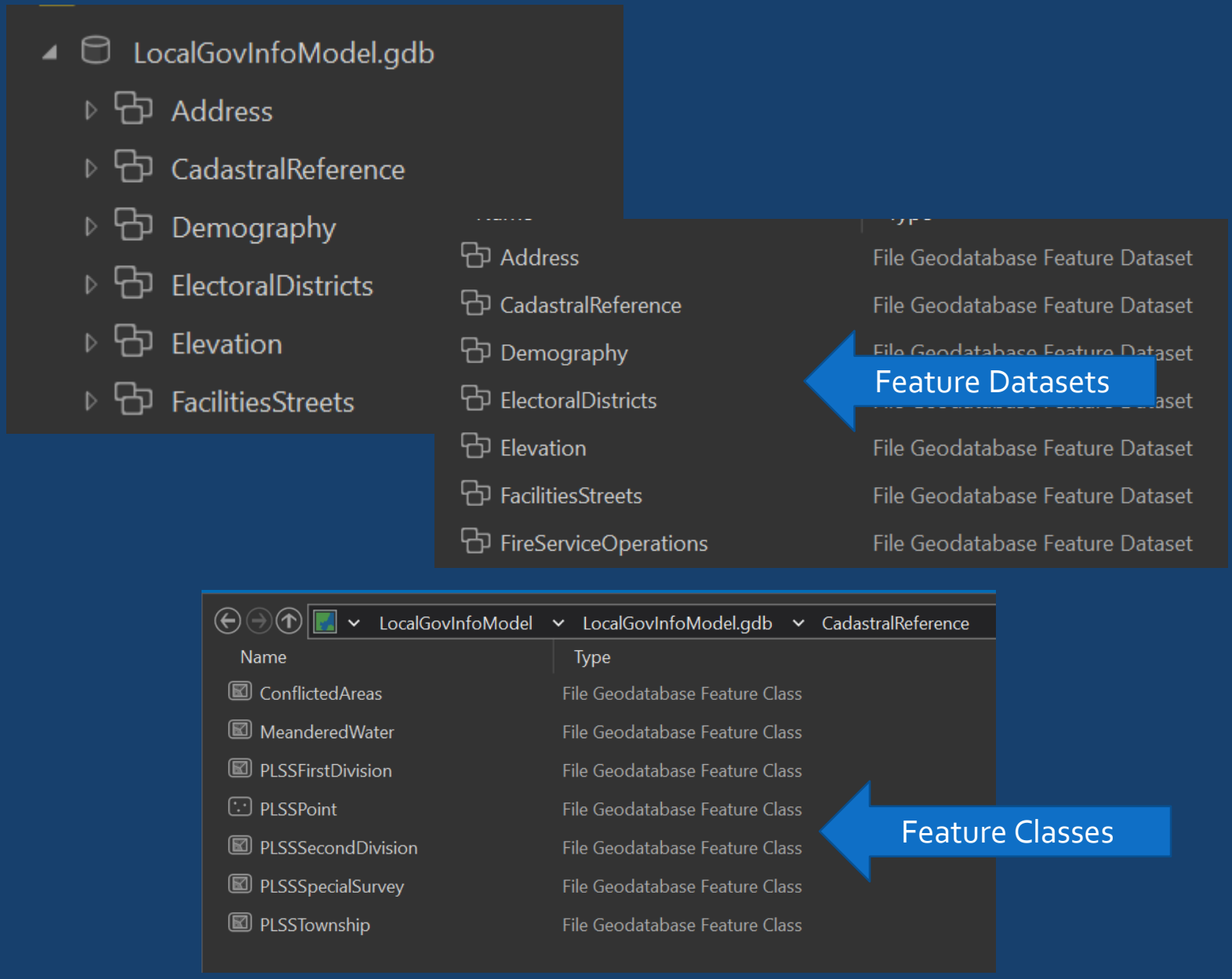
RASTERS IN A GEODATABASE

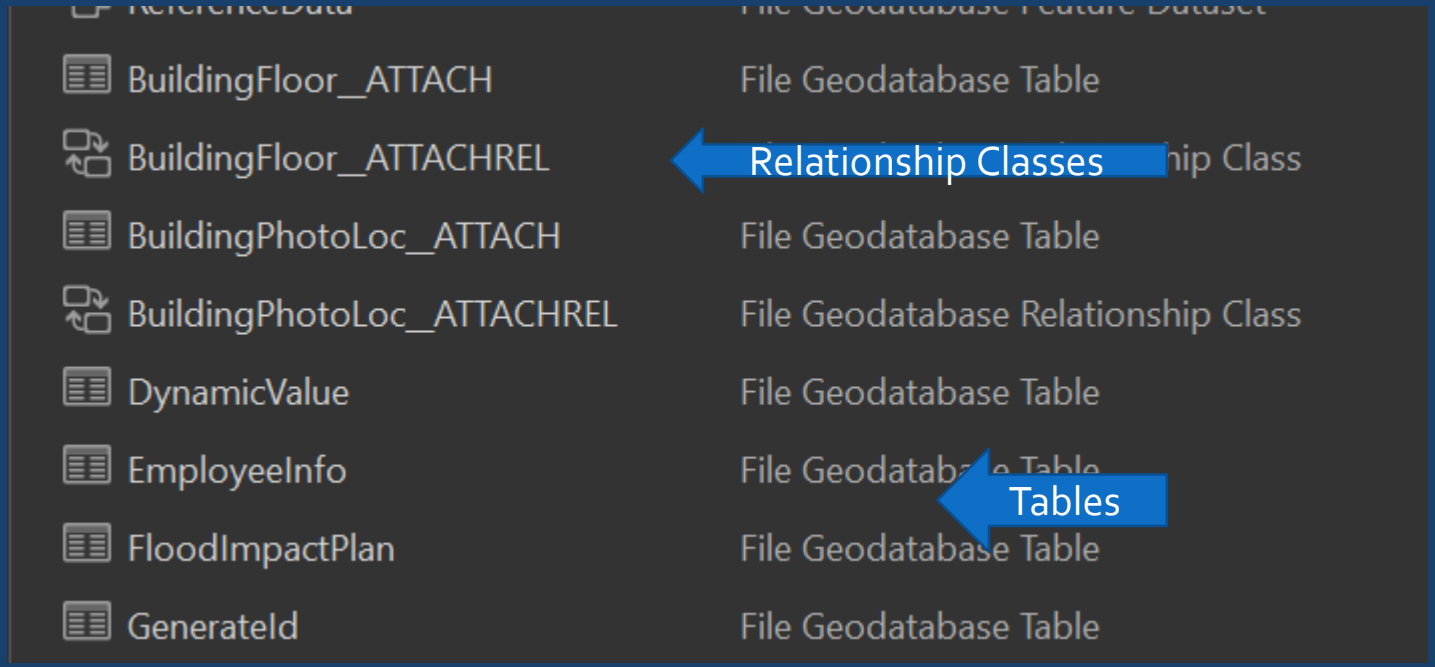
- Data Management Strategies
  - Raster Provisioning
  - Rasters in the geodatabase
- Structures
  - Raster dataset: Any raster data model stored in a geodatabase can be referred as a raster dataset.
  - Mosaic dataset: A collection of raster datasets stored as a catalog and viewed or accessed as a single mosaicked image or individual images.

RELATIONSHIP CLASSES

- Create an association between two tables.
  - Ex: a building can be associated with a parcel, parcel can be associated with an owner name.
- Cardinality
  - One-to-one
  - One-to-many
  - Many-to-many




LocalGovInfoModel.gdb
Address
CadastralReference
Demography
ElectoralDistricts
Elevation
FacilitiesStreets
Address
File Geodatabase Feature Dataset
CadastralReference
File Geodatabase Feature Dataset
Demography
ElectoralDistricts
Feature Datasets
Elevation
File Geodatabase Feature Dataset
FacilitiesStreets
File Geodatabase Feature Dataset
FireServiceOperations
File Geodatabase Feature Dataset
LocalGovInfoModel
LocalGovInfoModel.gdb
CadastralReference
Name
Type
ConflictedAreas
File Geodatabase Feature Class
MeanderedWater
File Geodatabase Feature Class
PLSSFirstDivision
File Geodatabase Feature Class
PLSSPoint
File Geodatabase Feature Class
PLSSSecondDivision
File Geodatabase Feature Class
Feature Classes
PLSSSpecialSurvey
File Geodatabase Feature Class
PLSSTownship
File Geodatabase Feature Class


BuildingFloor__ATTACH
File Geodatabase Table
BuildingFloor__ATTACHREL
Relationship Classes
BuildingPhotoLoc__ATTACH
File Geodatabase Table
BuildingPhotoLoc__ATTACHREL
File Geodatabase Relationship Class
DynamicValue
File Geodatabase Table
EmployeeInfo
Tables
FloodImpactPlan
File Geodatabase Table
GenerateId
File Geodatabase Table

GEODATABASE SCHEMA

- A schema defines the physical structure of the geodatabase along with the rules, relationships, and properties of each dataset in the geodatabase.
- Schema remains fixed when geodatabase is in use with help of schema locks
- Schema locks: shared or exclusive
- Geodatabase Schema can be shared as an XML document.

AN ARCGIS FEDERAL LANDS DATA MODEL

for the U.S. Bureau of Land Management

NILS conceptual model

The NILS implementation is based on a conceptual model comprising three tiers. The foundation is the survey fabric—networks of measurements, computations, coordinates, and survey points. The survey fabric is used to construct and maintain the legal description fabrics, containing Public Land Survey System (PLSS) boundaries, and other surveys such as metes and bounds. The legal description fabrics forms the building blocks for the parcel fabric.

Parcel fabric

Feature datasets with ownership parcels, rights, segregations, use authorizations, and land status.

Legal description fabric

Feature datasets with corners and boundaries, Public Land Survey System, special surveys, and unsurveyed features.

Survey fabric

Survey datasets with projects containing measurements, computations, coordinates, and survey points.

Topology rules

Survey links

The geodatabase structure

survey to GIS

Survey dataset

A survey dataset is a logical grouping of measurements and information. It is managed as an integrated data layer in the geodatabase in a set of DBMS tables.

A survey project represents a unit of work and manages a group of measurements, points, coordinates and computations.

Survey projects

Project ID	Name	Coordinate System
1	Blue	NAD_1927_UTM_Zone_12N
2	Red	NAD_1927_StatePlane_Arizona_Central

Measurements

Project ID	FromTo	Type
1	B1-B2	Reference Direction
1	B1-P1	Angle
1	B1-P1	Distance
1	P1-B1	Reference Direction
1	P1-P2	Angle
1	P1-P2	Reference Direction
1	P1-P2	Distance
2	R1-R2	Reference Direction
2	R1-P2	Angle
2	R1-P2	Distance
2	R2-R1	Reference Direction
2	R2-P3	Angle
2	R2-P3	Distance

Survey measurements measure slope, horizontal and vertical distances between points, and angles between lines of sight.

Survey points

Project ID	Point ID	GIS Coordinate	ID
1	B1	(x, y, z)	101
1	B2	(x, y, z)	101
1	P1	(x, y, z)	103
1	P2	(x, y, z)	104
1	P3	(x, y, z)	105
2	R1	(x, y, z)	106
2	R2	(x, y, z)	107

Survey points are named locations observed through multiple surveys. A GIS coordinate is derived from multiple coordinates associated with a survey point and reflects the shape of features in survey-aware feature classes.

Computations

Project ID	Name	Type
1	Computation1	Traverse
2	Computation1	AngleDistance
2	Computation2	AngleDistance

Computations define points that can be used as input for other computations. Survey Analyst stores algorithm definitions computations in a computation network.

Coordinates

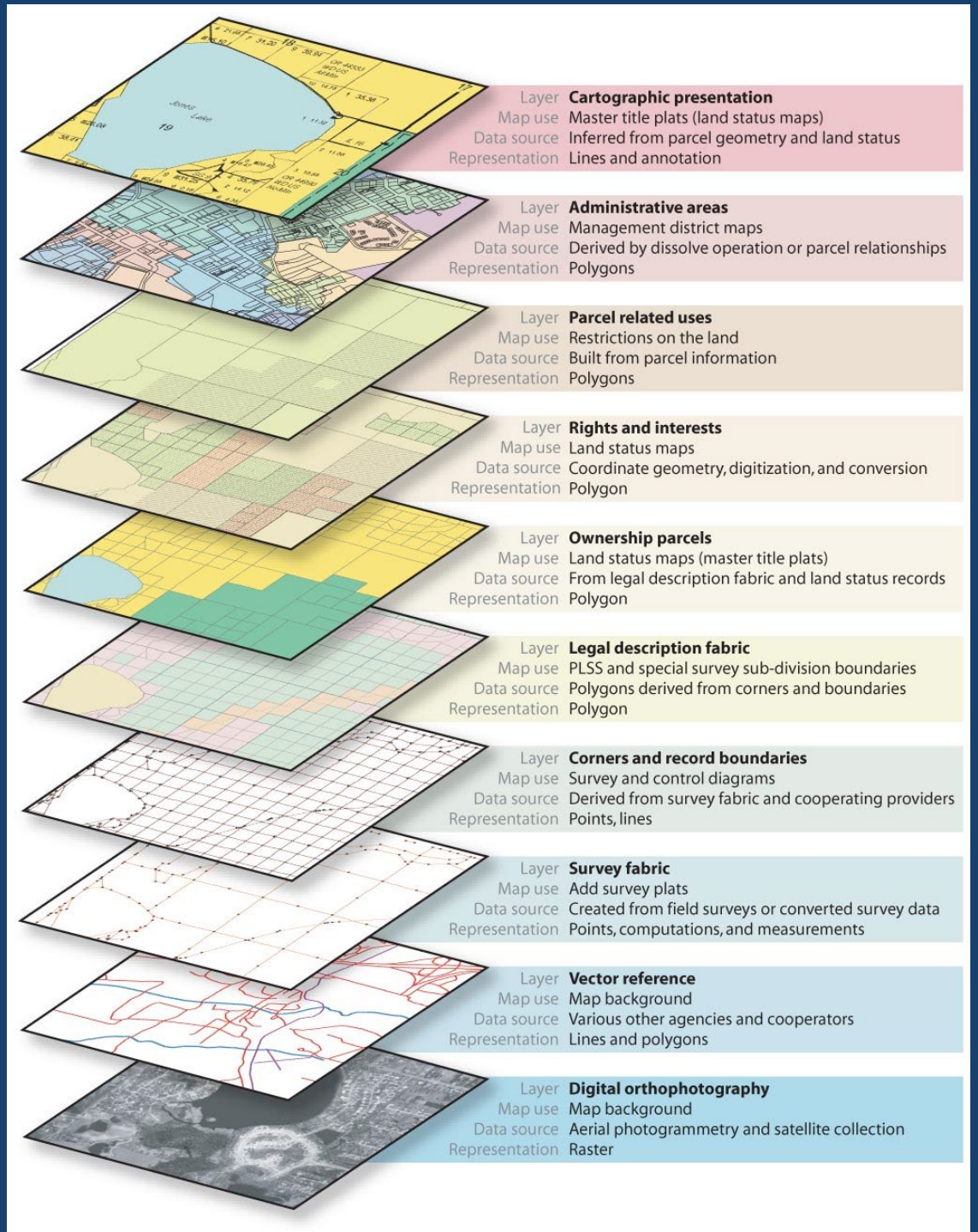
Project ID	Survey Point ID	Type	Coordinate
1	101 (B1)	Imported	(x, y, z)
1	101 (B2)	Imported	(x, y, z)
1	103 (P1)	Computed	(x, y, z)
1	104 (P2)	Computed	(x, y, z)
1	105 (P3)	Computed	(x, y, z)
2	106 (R1)	Imported	(x, y, z)
2	107 (R2)	Imported	(x, y, z)
2	104 (P2)	Computed	(x, y, z)
2	105 (P3)	Computed	(x, y, z)

Coordinates for the natural survey points are calculated through a series of survey measurements and computations.

https://navigator.blm.gov/

Thematic Layers in NILS

https://www.esri.com/news/arcnews/winter0304articles/nils-data-model.html



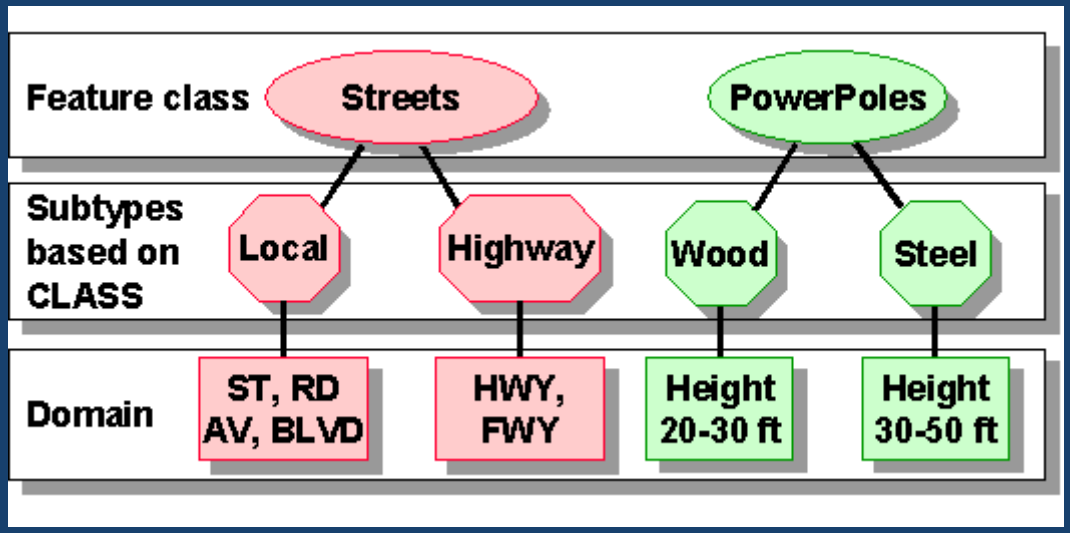
DOMAINS

- Rules that indicate valid values for a field in a table in a geodatabase.
- Preserves data integrity by restricting irrelevant data values
- Domains are used only if there was a definable set or range of specific values possible for that field.
- Types
  - Coded Domain – A list of values
  - Range Domain – A range of values (min, max is defined)
- Supports Split and Merge policies

SUBTYPES

- Subtypes are classifications within a feature class or table in a geodatabase.
- Logically group features based on a unique characteristic or behavior of the data.
- Using subtypes to store groups of related features can improve query performance.
- Rules:
  - Only one field in a table or feature class can have subtypes applied to it.
  - To use subtypes, the field on which you base the subtype must be a long or short integer field.
  - You can apply different topology and relationship rules to different subtypes.
  - You can apply different attributes or coded domains to other fields in the table based on subtypes.

DOMAINS & SUBTYPES



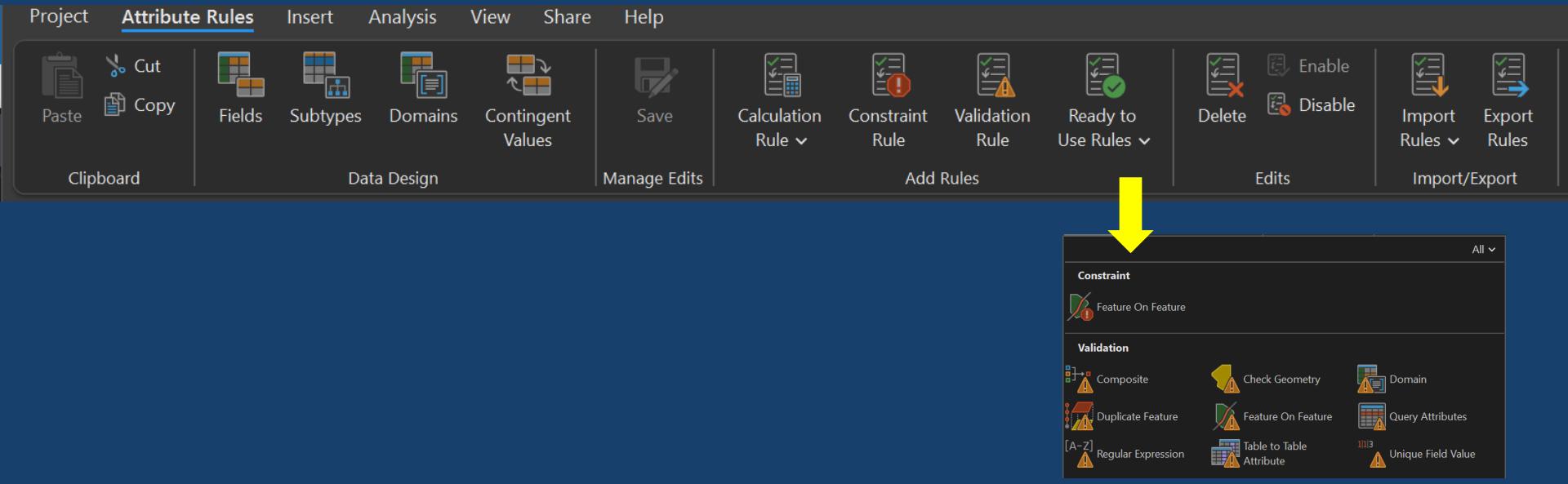
Source: ESRI

ATTRIBUTE RULES

- User-defined rules that can be used to:
  - Automatically populate attributes,
  - Restrict invalid edits during editing,
  - Perform quality assurance checks

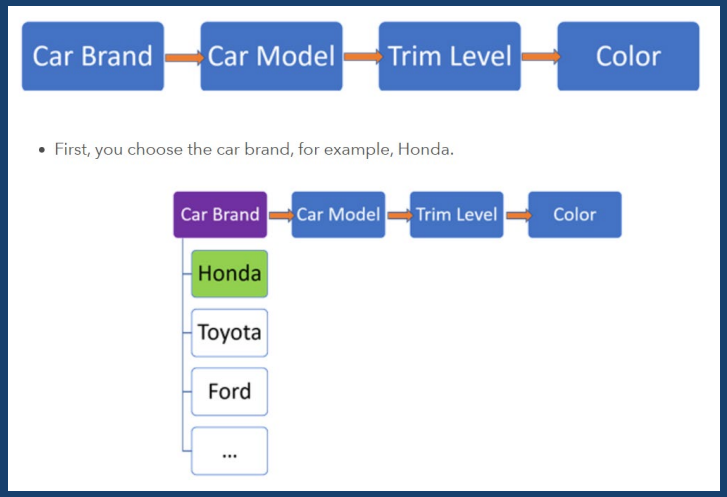
Rule Types

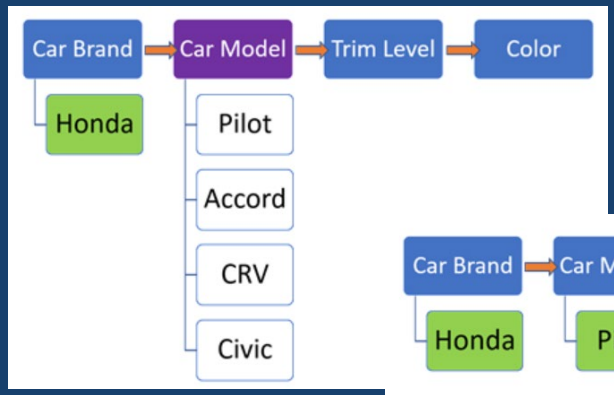
Calculation
Constraint
Validation

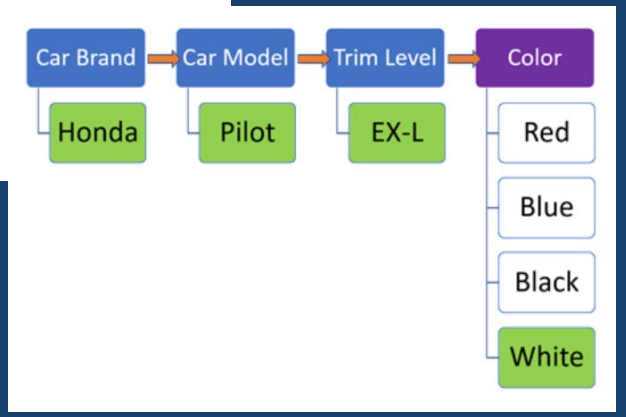


CONTINGENT VALUES

- Allows creating values in one field dependent on values in another field
- Domains must be created first, and contingent values can be used to further restrict the data values





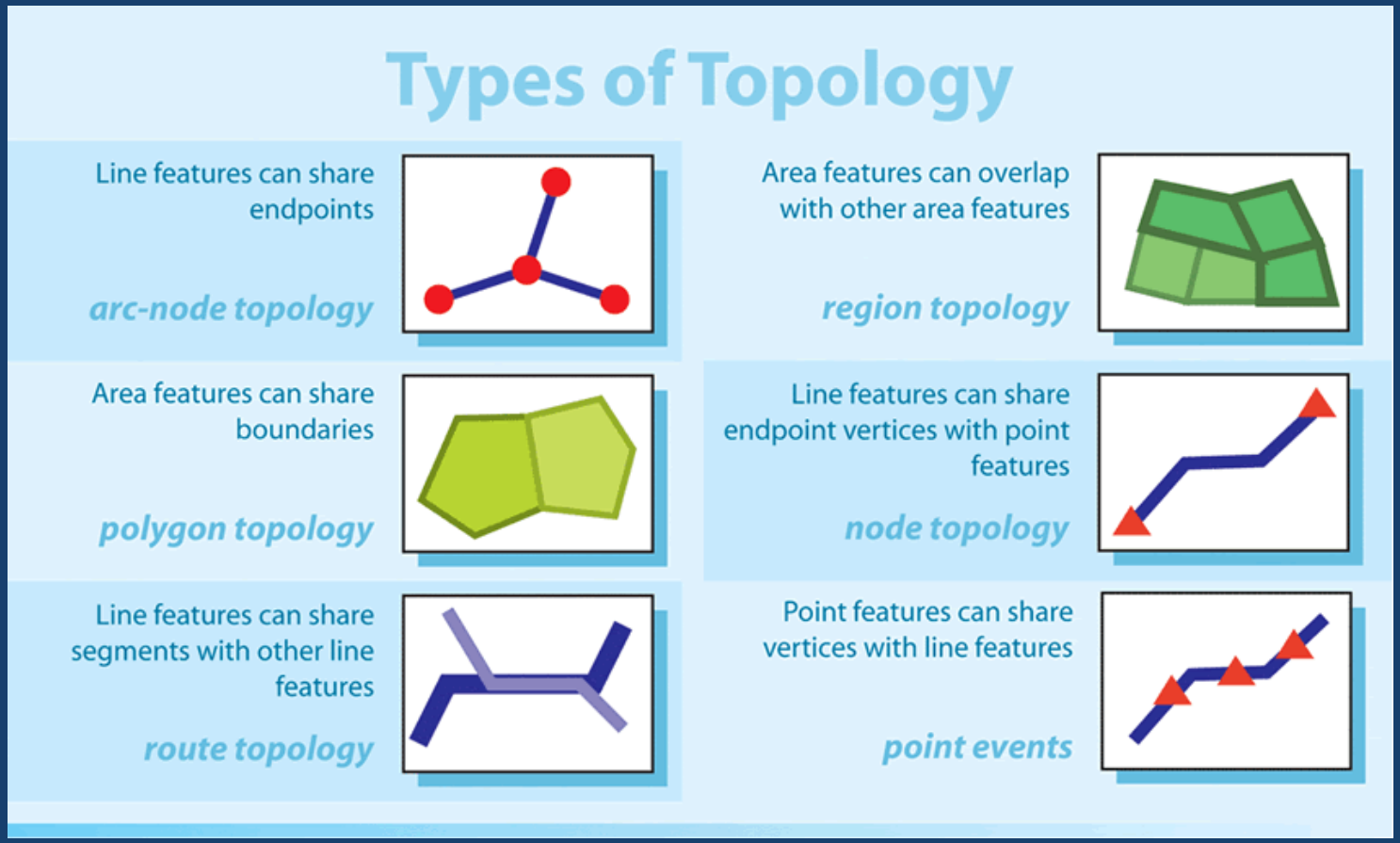


pro.arcgis.com

TOPOLOGY

- Rules to preserve data integrity by managing spatial relationships.
- Can only be created inside a feature dataset.
- A geodatabase can have multiple topologies, but one feature class can participate in only on topology.
- Topology rules can be defined between subtypes of feature classes.
- If errors are acceptable, can be marked as exceptions.
- Stores error features as: Point errors, Line errors, and Area errors

TYPES OF TOPOLOGY



ESRI

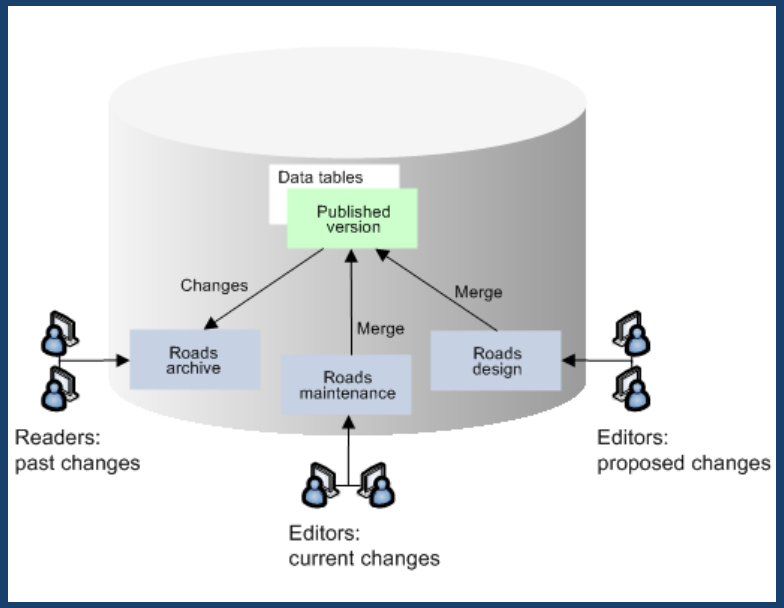
BENEFITS OF TOPOLOGY

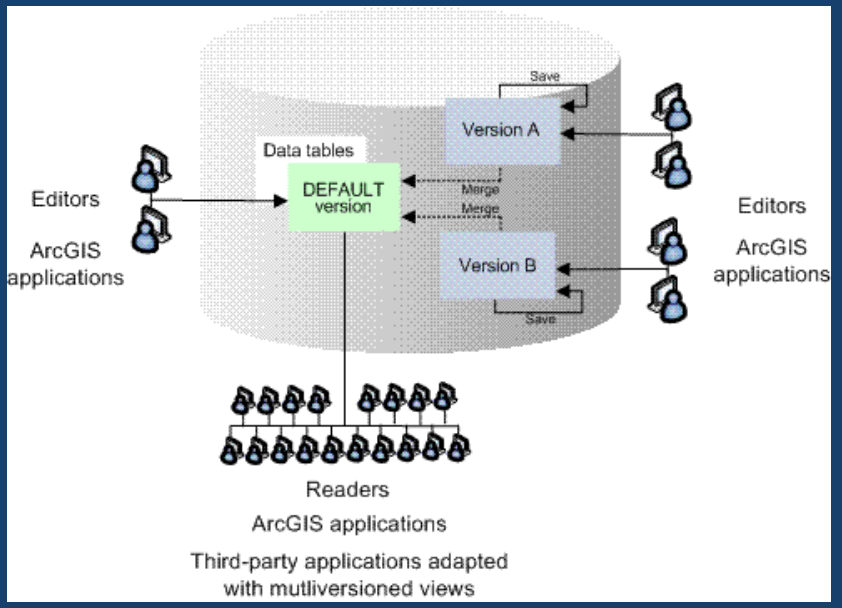
- Better data management
- More flexibility
- Improved data integrity
- More opportunities for data modeling
- ArcSDE multiuser environment
- Large map layers

EDITING DATA

- Versioned and Nonversioned data
- A version is a snapshot of the geodatabase, managed as system tables.
- Multiple users can connect to the database and edit
- Edits are recorded in delta tables
- Access permissions can be set on users
- Reconcile and Post operations allow to view conflicts and update edits to the default database.

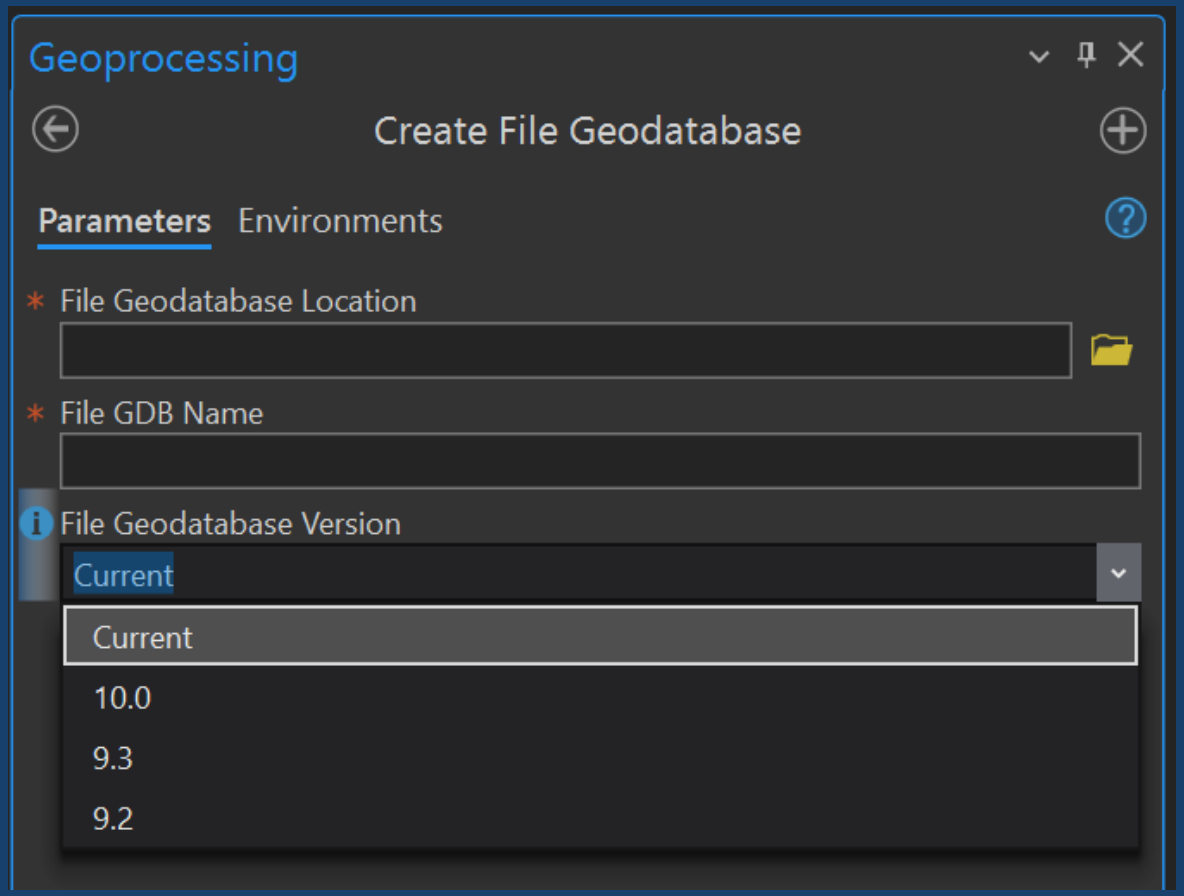
VERSIONING DATA





ESRI

OLDER RELEASE GEODATABASE

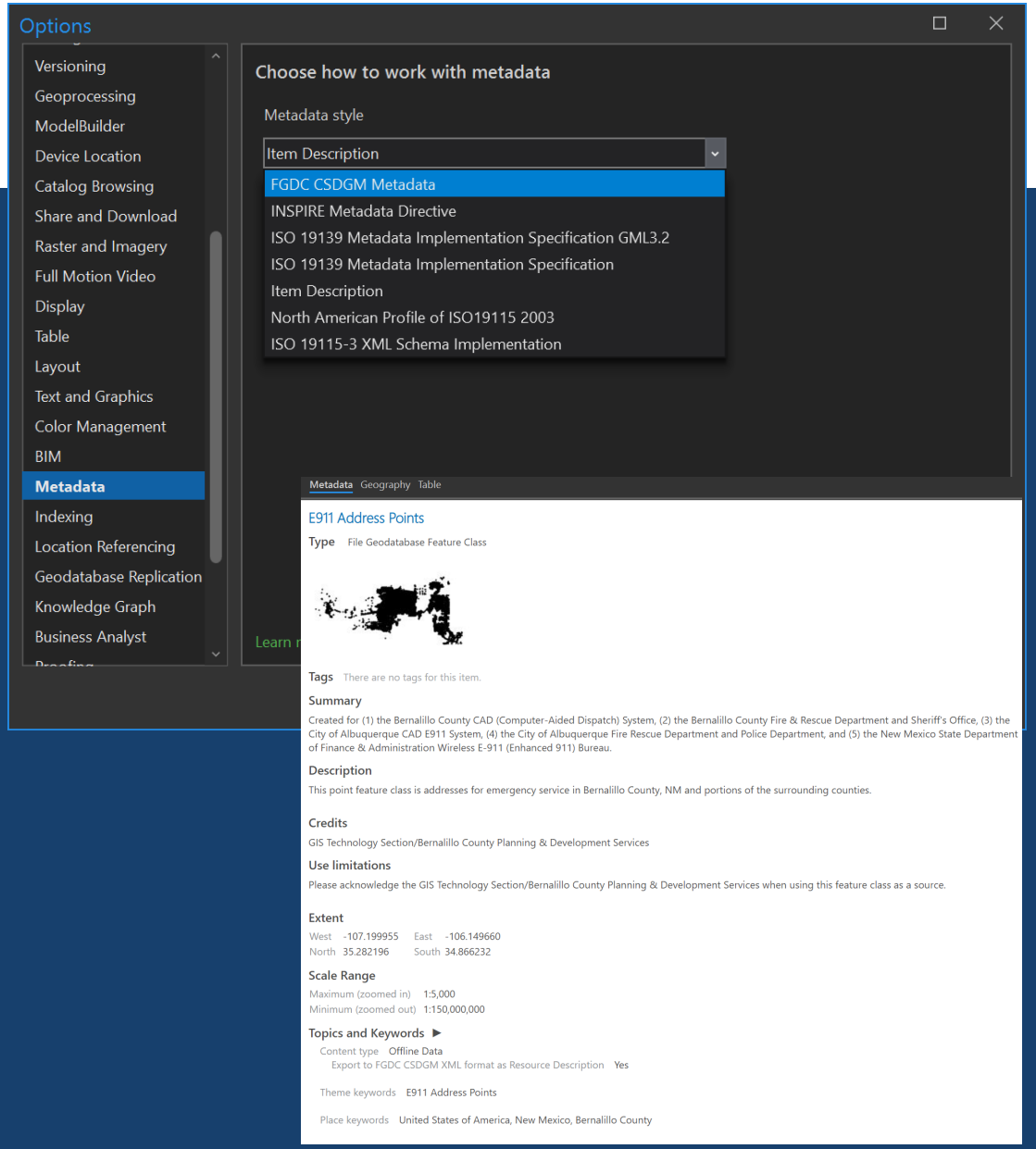


FILE NAMING CONVENTIONS

- The name must begin with a letter, cannot contain spaces or special characters (such as *, &, !, %, ., +, or -)
- Maximum length can be of 31 characters when combined with your server name.
- Multiple-band grid cannot have more than 9 characters in its filename, and a single-band raster dataset cannot have more than 13 characters.
- Feature class or table name – 160 characters
- Field names:
  - 64 characters (file and personal geodatabases)
  - 30-31 characters (Other enterprise formats)
- Field names in dBASE tables – 10 characters

METADATA

- At minimum fill in basic information such as project description and date.
- Descriptions for field names.
- Data access constraints.
- FGDC-CSDGM Metadata Standard



READINGS

- https://pro.arcgis.com/en/pro-app/latest/help/data/geodatabases/overview/what-is-a-geodatabase-.htm
- https://desktop.arcgis.com/en/arcmap/latest/manage-data/geodatabases/types-of-geodatabases.htm