Migrating to the Geodatabase

Transcript

Welcome to the ESRI Instructional Series Podcasts. This broadcast is titled *Migrating to the Geodatabase*. My name is Derek Law and I work at ESRI's main office in Redlands, California. I specialize in geodata management and ArcSDE technology.

In this podcast, I will talk about how to migrate existing GIS data to the geodatabase. This podcast is intended for GIS managers and GIS analysts who are just starting to work with geodatabases and would like to learn more about how to convert their GIS data into the geodatabase.

To quickly review: The geodatabase is the native data structure in ArcGIS; it is a container of spatial and attribute data, and supports all the different types of data that can be used by ArcGIS. In this podcast, I will discuss the migration strategies of five very common GIS data formats into the geodatabase. This includes: tables, shapefiles, coverages, CAD data, and imagery (i.e., raster data).

ArcCatalog is the ArcGIS Desktop application that enables you to easily manage and access GIS data. Within ArcCatalog, there are two general workflows to migrate GIS data into the geodatabase:

1. Import data into the geodatabase: In the Catalog tree, select the geodatabase you want to place your GIS data into, right-click and select Import from the context menu, then choose the appropriate data import tool.
2. Export data into the geodatabase: In the Catalog tree, browse to and select the GIS dataset that you want to put into the geodatabase, right-click and select Export from the context menu, then choose the appropriate data export tool.

Both approaches utilize the same geoprocessing (GP) tools that are readily available directly within ArcToolbox, specifically in the *Conversion Tools* toolbox, in the *To Geodatabase* toolset. Recall that ArcToolbox is available in both the ArcCatalog and ArcMap Desktop applications. Since the *To Geodatabase* toolset migration tools are GP tools, you can leverage them within the ArcGIS geoprocessing framework.

This means that you can execute the migration tools in several different ways: as a regular tool dialog by activating the tool in ArcToolbox, or you could execute the tool as a model in ModelBuilder, or you could run the tool within a Python script, or lastly, via the ArcGIS Desktop command line interface. This option is for those die-hard ArcInfo Workstation users who loved typing in commands and writing AML scripts.

For more detailed information on the different ways to execute a GP tool in ArcGIS, look in the Geoprocessing help book in the ArcGIS Desktop help.

Next, let's examine the migration to geodatabase strategies for each of the five common GIS data formats:

For tables

You use the *Table to Table* and *Table to Geodatabase* GP tools to import or export dBase, INFO, VPF, and OLE DB tables into the geodatabase. The *Table to Table* tool works on a single table. You need to specify the input rows (i.e., the table you want to migrate), the output location (which would be the geodatabase), and the name of the output table. There are also three optional input parameters:

- Expression enables you to define a SQL expression to select a subset of the rows in the input table to migrate.
- Field Map provides options to add, rename, or delete fields when the input table is converted to the output table.
- Configuration Keyword is used to specify storage parameters in ArcSDE geodatabases and file geodatabases.

The *Table to Geodatabase* tool works on multiple tables. You specify one or more input tables to convert and the output geodatabase location. The converted tables will have the same names as the original tables when they are migrated into the geodatabase.

For shapefiles

A shapefile is the native data format for ArcView GIS 3.x technology. It is a type of feature class, and therefore directly maps to the geodatabase as a single feature class when you migrate it to the geodatabase. You use the *Feature Class to Feature Class* or *Feature Class to Geodatabase* GP tools to import/export shapefiles into the geodatabase.

The *Feature Class to Feature Class* tool works on a single shapefile. You need to specify the input shapefile, the output location (which could be the geodatabase or a feature dataset within the geodatabase), and the name of the output feature class. There are also three additional input parameters: Expression, Field Map, and Configuration Keyword, which were discussed previously.

For the Field Map parameter, we recommend that you drop the FID, shape_length, and shape_area attribute fields, because these will be replaced by new fields when the shapefile becomes a feature class in the geodatabase.

The *Feature Class to Geodatabase* tool works on multiple shapefiles. You specify one or more shapefiles to convert and the output geodatabase location. By default, the shapefiles will have the same names as the original shapefiles when they are migrated into the geodatabase.

For coverages

A coverage is the native data format for ArcInfo Workstation. Its structure can contain multiple geometry types and it is conceptually similar to a feature dataset in the geodatabase that contains feature classes. Each geometry type in a coverage would map to a separate feature class in the geodatabase, and typically they would be collected within a feature dataset to maintain their spatial relationships.

Before you perform the data conversion (import/export) of a coverage to the geodatabase, we recommend that you do two things:

1. Prepare the coverage for migration. This includes applying BUILD/CLEAN operations to clean the coverage topology; executing JOINITEM on tables associated with the coverage; running a PROJECTDEFINE to ensure the coverage has a documented spatial reference, and a DISSOLVE operation to drop any edges.
2. Plan ahead and map the data transfer workflow. Decide how you want to model each coverage geometry type in the geodatabase; for example, if a coverage contains arc and polygon features, do you want to model them as two stand-alone feature classes in the geodatabase, or do you want to store them both within the same feature dataset?

To convert a coverage's geometry type (for example, a coverage's polygon features), you use the *Feature Class to Feature Class* GP tool to import/export it into the geodatabase. This is the same GP tool that I mentioned previously for migrating a shapefile to the geodatabase.

For the Field Map parameter, we recommend that you drop the following attribute fields: FNode#, TNode#, LPoly#, RPoly#, Length, cover #, and coverID, because these will not be used by the new feature class in the geodatabase.

There is an excellent document in the ArcGIS Desktop help called *Getting started migrating from ArcInfo Workstation*, which guides you through the migration process of a coverage to the geodatabase.

For Computer Aided Design (CAD) data

ArcGIS supports CAD drawing files with the following extensions: .DGN, .DWG, and .DXF.
The structure of CAD drawing files are different in nature to GIS data, therefore there is not always a one-to-one correspondence between CAD objects in a drawing file and GIS features. With pre 9.3 ArcGIS releases, the process of migrating CAD drawing files into the geodatabase is a multi-step process.

First, you use the *Import from CAD* GP tool, which generates results in a staging geodatabase. It contains some tables and a feature dataset with feature classes that represent the geometry in the CAD drawing. In the tool, you need to specify: the input CAD drawing, an output staging geodatabase (which is a new geodatabase generated by the tool), and optionally, a spatial reference for the output feature dataset. You can also simplify the complex CAD objects by exploding them into simpler features. The output result in the staging geodatabase is an optimized representation of the CAD drawing in a geodatabase. Within its simplified structure, each piece of geometry translated from the CAD drawing will have multiple attributes that can be associated with it.

Next, you can perform some post-processing on the data inside the staging geodatabase, such as removing redundant records with the *Pivot Table* GP tool, to construct GIS feature classes from the original CAD data that meet your application needs. At the 9.3 release, CAD data conversion to the geodatabase is a single-step process with the *Feature Class to Feature Class* GP tool.

Note that there is a separate GP tool called *Import CAD Annotation* for migrating annotation objects in CAD drawing files into the geodatabase.

A comprehensive discussion on migrating CAD data to the geodatabase is beyond the scope of this podcast. For more information look in the *How to Import from CAD works* and *Using the Feature Class to Feature Class tool* help topics in the ArcGIS Desktop help.

The last data type is imagery

This is also commonly referred to as raster data. You use the Raster to Geodatabase GP tool, which can convert one or multiple raster files. You specify the input raster files to convert and the output geodatabase location. By default, the output raster datasets will have the same names as the original raster files when they are migrated into the geodatabase. The tool also has a Configuration Keyword option for specifying storage parameters in ArcSDE geodatabases and file geodatabases.

When loading rasters into an ArcSDE geodatabase, there are additional parameters to consider, such as pyramids, compression, and tile size. This workflow describes migrating imagery into the geodatabase as individual raster datasets. Alternatively, you could have loaded the imagery into a raster catalog in the geodatabase, where each record in the table stores a raster dataset, or you could have loaded a series of adjacent images into a single raster dataset by mosaicking them together (i.e., combining them into a single raster dataset file).

For more information on migrating imagery into the geodatabase, look in the *Importing raster datasets* help topic in the ArcGIS Desktop help.

To summarize, this podcast discussed the migration strategies of five very common GIS data formats to the geodatabase. The five data formats were: tables, shapefiles, coverages, CAD data, and imagery. As you learned, GIS data migration into the geodatabase is a relatively easy procedure that can be done within ArcCatalog, by either importing or exporting them into the geodatabase, using a collection of data conversion GP tools available in ArcToolbox. These can also be leveraged within the ArcGIS geoprocessing framework.

For more detailed information on the various GIS data migration strategies to the geodatabase that were discussed in this podcast, read the help topic titled *An overview of importing datasets* in the ArcGIS Desktop online help located at support.esri.com.

ESRI offers several instructor-led training classes on the topic of GIS data migration to the geodatabase. ESRI also has several self study classes available online.

Thank you for listening, and stay tuned for future podcasts.

