An Overview of the Geodatabase and Five Reasons Why You Should Use It

Transcript

Hello, and welcome to the ESRI instructional podcast series. 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 the geodatabase and the five reasons why you should be using it. There will be several upcoming podcasts related to the geodatabase.

The first, which you are listening to now, will present an overview of the geodatabase and discuss five reasons why you should be using this superior data format offered by ESRI. This podcast is intended for ArcGIS Desktop users and GIS managers, specifically for those who may still be working with shapefiles and coverages. For users who are already working with geodatabases, you might learn some new tricks in leveraging the power and functionality of the geodatabase.

First, let's define what a geodatabase is. Simply put, it is a container for spatial and attribute data and has been the default data format for ArcGIS since its initial release at version 8.0. The geodatabase supports all the different types of data that can be used by ArcGIS, such as attribute data, geographic features, satellite and aerial imagery (in other words, raster data), surface modeling or 3D data, network systems, and survey measurements. ArcGIS has a complete suite of conversion tools within ArcToolbox and ArcCatalog to easily migrate existing GIS data into the geodatabase. By storing GIS data within a geodatabase, you can then take advantage of its data management capabilities, which we will discuss in more detail later in this podcast. Vector GIS data is stored within the geodatabase as thematic layers called a feature class. A feature class is a collection of geographic features with the same geometry type, such as point, line, or polygon; the same attributes; and the same spatial reference—in other words, the same coordinate system information. Raster GIS data is stored as raster datasets. Typically, each raster image is stored as its own thematic layer. Multiple rasters can be collected together into a raster catalog, or they can be aggregated into a mosaic. With respect to older ESRI spatial data formats, a coverage is a collection of feature classes. A shapefile is a single feature class, and the geodatabase can contain many feature classes within its structure. Behind the scenes, the geodatabase is implemented as either a collection of files in a file system or as a collection of tables stored within a relational database management system—in other words, a DBMS.

Next, let's discuss the five reasons why you should be using a geodatabase.

1. A geodatabase is a uniform repository for geographic data, and it is scalable. The geodatabase supports many different types of GIS data, and all of it can be placed within its structure.

Therefore, GIS data is stored in one centralized location. This makes it easier to manage and access. You no longer have to keep track of various data in different formats stored at different locations within a network. You can also use the geodatabase's replication functionality to easily share and duplicate the contents of a geodatabase with other users at other locations. The geodatabase is also a scalable storage solution. You can start with a small, personal geodatabase for an individual project then easily upgrade it to a larger workgroup or enterprise geodatabase as your GIS data volume increases or your project scope expands.

2. Data entry and editing are more efficient. By storing GIS data in a geodatabase, you are able to apply rules and constraints on the GIS data to reduce the chances of error being introduced into the datasets. The geodatabase offers business logic functionality such as domains, valid values for fields and tables. There are two different types—range and coded value. Subtypes enable you to categorize features in a feature class based on an attribute field of type integer. Default value—when new features are created, you can specify a default value for an attribute field. And also, split and merge rules. So, when features are edited, you can set default rules on what happens to their attribute field values. As a GIS manager, you can specify these rules and constraints that your geodatabase users will follow. This will help to enforce database integrity and hopefully reduce the chances of errors occurring in your GIS data.

3. The geodatabase can model advanced spatial relationships. The geodatabase not only defines how data is stored, accessed, and managed, but it can also implement and model spatial relationships of features in a feature class or between feature classes. When feature classes are grouped together within a feature dataset, you can easily model advanced geospatial relationships such as topologies, geometric networks, network datasets, terrains, and survey datasets. One good example is that the geodatabase supports a full topology model. It enables you to implement 25 different spatial relationship rules on your GIS data in any combination you can think of. You control when topology validation occurs, and you can still work with the data even if the topology has errors. This enables you to leverage your GIS data to its full potential and maintain a consistent, accurate database.

4. The geodatabase has multiuser editing capabilities. Geodatabases with ArcSDE technology support a multiuser editing environment, without data duplication or data extraction. In other words, you can have two or more users accessing the same data at the same time,

simultaneously making edits. This is possible with versioning, the ability to track and record all changes to the contents of a geodatabase. Versioning is used to support complex business workflows and is conceptually the idea of being able to support long transactions within a DBMS. Multiple editors can make changes to data in the geodatabase. Then, later in the business workflow, the different edits can be reconciled and aggregated back together. GIS managers would control user access and read-write permissions to the contents of a geodatabase, such as individual tables or feature classes.

5. The geodatabase enables your GIS data to be integrated with information technology systems. Geodatabases with ArcSDE technology are implemented in a DBMS, so GIS data could be used with other non-GIS data stored in the same DBMS. You could create database views and, in some instances, directly query GIS data with a native SQL or sequel language of the DBMS. If your business model is not exclusive to GIS, you could utilize your GIS data for other uses. For example, a feature class of parcels might be used by another third-party program that also accesses the DBMS for property tax calculations.

To review, the geodatabase is a container of spatial and attribute data and enables you to store many different types of GIS data within its structure. It is implemented as either a collection of files in a file system or tables in a DBMS. Five reasons why you should be using a geodatabase are: (1) a geodatabase is a uniform repository for geographic data and it is scalable; (2) data entry and editing are more efficient; (3) the geodatabase can model advanced spatial relationships; (4) the geodatabase has multiuser editing capabilities; and (5) the geodatabase enables your GIS data to be integrated with information technology systems. Much of this functionality is not available with shapefiles or coverages.

I have discussed a lot of functionalities available in geodatabases at a very high level. For more information, you can go to ESRI's main Web site, www.esri.com/geodatabase. Another resource is the ESRI support site, support.esri.com. On the left side of the main landing page, there are shortcut links to the ArcGIS online help documentation. In addition, ESRI also offers an instructor-led class on the geodatabase called *Building Geodatabases.*

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