**INTRODUCTION TO ARCGIS SOFTWARE**

I.      History of Software Development
        a.  Developer – ESRI - Environmental Systems Research Institute, Inc., in 1969 as a
            privately held consulting firm that specialized in landuse analysis projects.
                i.   Software Developers originated at Harvard graphics lab
        b.  ArcInfo – early version of software
                i.    Main-frame, unix-based software
                ii.   "arc" refers to line segments of map elements
                iii.  "info" refers to information in database system
                iv.   Command driven software package (analogous to "DOS")
        c.  ArcView (1990's)– first release of a windows-based, GUI (graphical-user interface)
            based GIS system for desktop computers; still available but phasing out.
        d.  ArcGIS (2000's) – next generation GIS software merging codes and routines from
            best of ArcInfo and placed in a GUI-Windows environment of ArcView, compatible
            with desktop computing

II.     Basics of Spatial Data Model
        a.  Georeferencing - To establish the relationship between digital map elements and real-
            word geographic coordinate location systems.
                i.   Real-world coordinates = an x,y coordinate system used to represent
                     geographic locations.
        b.  Vector Model – map elements composed of points, lines, and polygons; with each
            point located at a georeferenced X,Y coordinate.
                i.    Feature classes – points, lines, polygons
                        1.  Attributes – information attached to map elements
                ii.   Vector models
                        1.  Topological models – individual map elements are spatially related to
                            one another
                        2.  Spaghetti models – individual map elements are independent and piled
                            on top of one another like a plate of spaghetti
        c.  Raster Model – map features are comprised of a matrix array of cells ("pixels" or
            "grid cells") that are arranged in rectangular blocks
                i.    Columns and Rows of cells
                ii.   Cells are georeferenced to X-Y locations
                iii.  Real-world X-Y dimensions of grid cells = "resolution"

III.    Overview of ArcGIS Sofware - Software collective referred to as "ArcGIS Desktop"
        a.  Software Modules
                i.    ArcMap – software used to display, analyze, and create GIS data
                        1.  ArcView level install – basic level application
                        2.  ArcEditor – 1 step up, includes data editing capabilities
                        3.  ArcInfo – highest level install, with advanced editing and analysis
                            features
                ii.   ArcCatalog- tool for viewing and managing spatial data files (analogous to
                      Windows Explorer)
                iii.  ArcToolbox – set of tools and functions used to convert data formats, manage
                      map projections, perform analysis, modify data.

b. Extensions
    i. The Extensions dialog allows you to load and unload software capabilities, allowing you to enhance your working environment with additional objects, scripts and customization.
    ii. You can use extensions provided by ESRI and you can also create your own.
c. Data Files in ArcGIS
    i. Shapefiles – developed for ArcView

ArcView shapefiles are a simple, non-topological format for storing the geometric location and attribute information of geographic features. A shapefile is one of the spatial data formats that you can work with in ArcView. The shapefile format defines the geometry and attributes of geographically-referenced features in as many as five files with specific file extensions that should be stored in the same project workspace. They are:

.shp - the file that stores the feature geometry.

.shx - the file that stores the index of the feature geometry.

.dbf - the dBASE file that stores the attribute information of features. When a shapefile is added as a theme to a view, this file is displayed as a feature table.

.sbn and .sbx - the files that store the spatial index of the features.

WARNING: shapefiles come in connected bundles of 3+; when you are managing your files, do not delete or rename any of the related files, it will corrupt the data source, and ArcGIS will not be able to load the data.

    ii. Coverages – vector data developed for Arc/Info, using a topological model.
        1. coverages come in bundles of multiple files and multiple folders; the main coverage folder is assigned a *.adf (arc data file) extension. All coverage files must be grouped and located in the same directory.
        2. INFO files are the database attributes that accompany the digital map element files in a coverage.
        3. A folder containing one or more coverages is termed a "workspace" in ArcInfo lexicon
    iii. Geodatabases
        1. A new data model developed for ArcGIS. All map elements, feature classes, and attribute data are stored in a relational database file format.
            a. Personal Geodatabase – for use by individuals, compatible with MS Access software environment
            b. Enterprise Geodatabase – for use by multiple, network users; uses server-based relational database management systems (RDBMS) such as ORACLE or SQL-Server
        2. ArcSDE = Arc spatial database engine – software that links enterprise geodatabases to ArcMap and Arc Catalog

      iv.  Layer Files – DOES not contain actual spatial data, but is a "pointer" file that references the data source, downloads data into ArcMap, and formats the map layers for appearance and style.
          1.  Several spatial data files can be combined into a single "layer file"
      v.  Raster Data - Raster data records spatial information in a regular grid or matrix organized as a set of rows and columns.  Each cell within this grid contains a number representing a particular geographic feature, such as soil type, elevation, land use, slope, etc. Raster data is commonly, but not exclusively, used to store information about geographic features that vary continuously over a surface, such as elevation, reflectance, groundwater depths, etc. ARC/INFO grids are raster data. Image data is a form of raster data in which each cell or pixel stores a value recorded by an optical or electronic device.

        Raster data is highly dependent on the resolution of the regular grid in which it is recorded. The size of the cells in the grid is fixed, so as you zoom in on raster data displayed on a view, you will eventually see the shape of the cells.
      vi.  Tables – spreadsheet tables (e.g. in Excel) that are not linked to digital map elements, may be stored in a *.dbf or *.txt format.
      vii.  Grids – raster data format developed by ESRI for use in ArcInfo and ArcMap; requires use of the "Spatial Analyst" extension
      viii.  TINS and DEMs – raster-based elevation models of the Earth's surface.  The attribute associated with georeferenced grid cells is land elevation.
IV.    Introduction to Metadata
    a.  'data about the data': who created, where did it come from, what coordinate system in being used in georeferencing, what to the attribute fields mean, etc.
    b.  Metadata standards managed by Federal Geographic Data Committee (FGDC)
    c.  Metadata is critical to make the GIS meaningful and to understand what the data is about.

Geographic Names Information System [downloaded file]. (2008) Reston, VA: Unites States Geological Survey. Available FTP: http://geonames.usgs.gov/domestic/download_data.htm [May 21, 2008].

1:24,000 Digital Line Graphs (DLG) [downloaded file]. (2000) Sioux Falls, South Dakota: USGS EROS Data Center. Available FTP: http://edc2.usgs.gov/geodata [January 7, 2007].

**Data provided upon request by agency or individual**

Data name [type of medium]. Producer location city/state/province: Producer name, year.

Badlands National Park GIS Data [CD-ROM]. Interior, South Dakota: National Park Service—Badlands National Park, by request, 2006.

## About ArcGIS

### ArcGIS overview

ArcGIS is developed and sold by Environmental Systems Research Institute, Inc. (ESRI). It has a long history and has been through many versions and changes. Originally developed for large mainframe computers, in the last 10 years it has metamorphosed from a system based on typed commands to a graphical user interface (GUI), which makes it much easier to use. Because of the size and complexity of the program (actually a suite of programs), and because users have come to depend on certain aspects of the software, much of the code is carried forward and included in the new versions. Knowing this background helps a student of ArcGIS understand the nature of the ArcGIS system and helps explain some of its odd characteristics.

The older core of the ArcGIS system was called Arc/Info and included a basic set of programs—Arc, ArcEdit, and ArcPlot (Fig. 1.10)—which utilized the coverage data model and was built in a database program called **INFO** that appears primitive today. All of the programs were command based, meaning that the user typed commands into a window to make the program work.
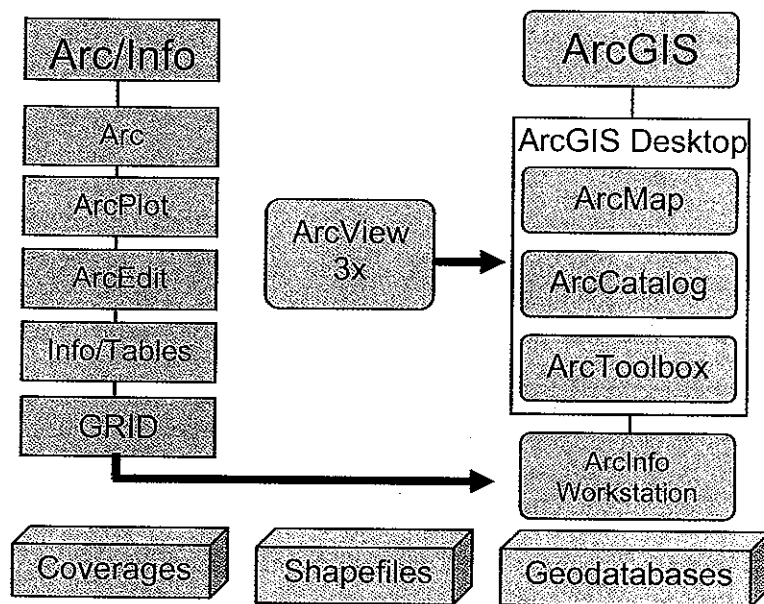


Fig. 1.10. Relationship between ESRI products and data formats

The difficulty of learning Arc/Info prompted ESRI to create ArcView, which was easier to use but not as powerful as Arc/Info. ArcView was designed primarily to view and analyze spatial data rather than create it. ArcView also used a simpler data model, called the shapefile, although it could read coverages and convert them to shapefiles. Beginners in GIS often learned ArcView first and then began learning Arc/Info as their needs and abilities advanced.

ArcGIS, released in 2001, is a synthesis of the powerful Arc/Info system with the easy-to-use interface of ArcView, updated to use the latest advances in desktop computing and database technology. It contains two programs, collectively referred to as ArcGIS Desktop.

> ArcMap provides the means to display, analyze, and edit spatial data and data tables. Similar in appearance to its ArcView predecessor, it nevertheless contains powerful new functionality.

> ArcCatalog is a tool for viewing and managing spatial data files. It resembles Microsoft Windows Explorer, but it is specially designed to work with GIS data. It should **always** be used to delete, copy, rename, or move spatial data files.

In addition, ArcGIS Desktop contains ArcToolbox, a collection of tools and functions for operations in ArcCatalog and ArcMap, such as converting between data formats, managing map projections, and performing analysis. Users may create and add their own tools or scripts for special or often-used tasks. The ESRI Web site at www.esri.com has a large library of scripts and tools that can be downloaded to extend the ArcGIS functionality.

Finally, the original Arc/Info command-line software can still be accessed in the additional module called Workstation ArcInfo, which is still used by organizations that may be tied to the older coverage model for various reasons, such as having a large number of specialized programs written in the older AML programming language.

The ArcGIS system also provides different levels of functionality that all use the same basic interface. Users can save money by buying only the functions they need.

> ArcView provides all of the basic mapping, editing, and analysis functions for shapefiles and geodatabases and is the level of functionality most users will require on a regular basis. It includes ArcMap, ArcCatalog, and a subset of ArcToolbox functions.

> ArcEditor includes all the functions of ArcView but adds editing capabilities needed to work with the advanced aspects of the geodatabase, such as topology and network editing. Additional functions reside in ArcToolbox at this level.

> ArcInfo provides access to the full functionality of the ArcGIS Desktop tools and the full version of ArcToolbox. In addition, it includes the original core Arc/Info software, now called Workstation ArcInfo.

This book focuses almost exclusively on the functions available with an ArcView license, although it mentions some of the additional capabilities as appropriate. Users can read the software documentation to learn more about the advanced topics.
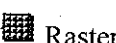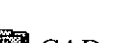
The ESRI system of GIS programs, then, is a fairly complex set of tools with a long history, designed to work with a number of different data formats, also with a long history. We turn now to a discussion of how ArcGIS stores data.

## Data files in ArcGIS

ArcGIS can read a variety of different file formats. Many of these come from older versions of the software. Some can come from other programs such as image processing packages and

computer-aided design (CAD) systems. Table 1.1 lists many of the data sets than can be used in ArcGIS with the icons showing how they appear in ArcCatalog.

Table 1.1. Types of files and data sources used by ArcGIS

| File type | Description |
|---|---|
| ▨▨ ▨ Shapefiles | Shapefiles are vector feature classes developed for the early version of ArcView and have been carried over into ArcGIS. |
| ▨ ▨▨ Coverages | A coverage is the vector data format developed for Arc/Info and is the oldest of the data formats. |
| ▨ ▨▨▨▨ Geodatabases | Geodatabases represent an entirely new model for storing spatial information with additional capabilities. |
| ▨ Database connections | Database connections permit users to log in to and utilize data from an RDBMS geodatabase. |
| ◈ ◈ ◈ Layer files | A layer file references a feature class and stores information about its properties, such as how it should be displayed. |
| ▨ Rasters and grids | Rasters represent thematic maps or images by arrays of numbers stored in binary format (base 2). Grids are a special raster format used with the Spatial Analyst extension. |
| ▥ ▥ Tables | Tables can exist as separate data objects that are unassociated with a spatial data set. |
| ▨ Internet servers | Many organizations now make data available over the Internet. Users can connect to these data sources and download information for their work. |
| ☑ TINs | TINs are Triangulated Irregular Networks that store 3D surface information, such as elevation, using a set of nodes and triangles. |
| ▨ CAD drawings | Data sets created by CAD programs can be read by ArcGIS, although they cannot be edited or analyzed unless they are converted to shapefiles or geodatabases. |

### Shapefiles

Shapefiles are georelational spaghetti data models developed for the early version of ArcView. A shapefile contains one feature class composed of points or lines or polygons but never a mixture. The attributes are stored in a **dBase** file. Shapefiles can, however, store **multipart features (or multifeatures),** which are single features made of multiple objects. For example, the state of Hawaii requires multiple polygons to represent the different islands, but it can be stored as a multifeature so that it has only one record in the attribute table.

Although a shapefile appears as one icon in ArcCatalog, it is actually composed of multiple data files that can be seen individually in Windows Explorer (Fig. 1.11). The rivers shapefile has eight different files associated with it. The .shp file stores

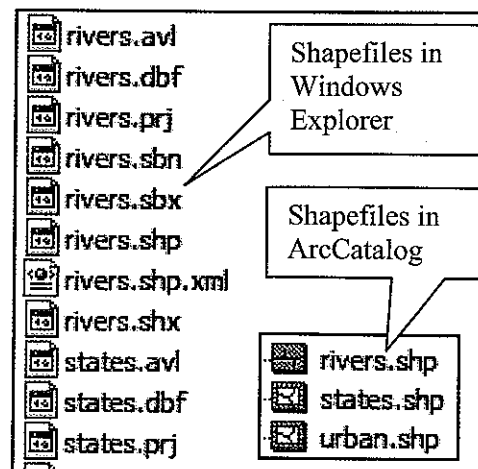| | |
|---|---|
| ▨ rivers.avl | Shapefiles in Windows Explorer |
| ▨ rivers.dbf | |
| ▨ rivers.prj | |
| ▨ rivers.sbn | |
| ▨ rivers.sbx | Shapefiles in ArcCatalog |
| ▨ rivers.shp | |
| ▨ rivers.shp.xml | |
| ▨ rivers.shx | |
| ▨ states.avl | ▨ rivers.shp |
| ▨ states.dbf | ▨ states.shp |
| ▨ states.prj | ▨ urban.shp |

Fig. 1.11. Shapefiles are groups of files but appear as single entries in ArcCatalog.