

## Getting Started in the GIS Data Library

[Introduction](#)

[Frequently Used Folders](#)

[Metadata](#)

[Layers and Predefined Symbology](#)

[Templates and Print Ready PDF Maps](#)

[Documentation and Tutorials](#)

### Introduction

A great deal of GIS and GIS related information is already stored on the GIS server. This document is designed to give you an overview of the contents of the folder M:\GISDATA.

The GIS Data Library can be navigated using Windows Explorer. The folder contains GIS formatted data, as well as documentation, tutorials, service packs and patches for the ArcView software and predefined templates to make it easier to make a basic map. In order to preview the geographic files, you must use ArcCatalog. (When in ArcCatalog, only the geographic files will be visible.)

**M:\GISDATA\GISdocumentation** - contains many sets of instructions on how to perform various tasks in GIS. Please take a moment to review the titles in this folder. You may wish to print some out to have handy while you are working on a new task.

### Geographic Data – Most Frequently Used Folders

[M:\GISDATA\Census2000](#)

[M:\GISDATA\District\\_bounds](#)

[M:\GISDATA\Facility](#)

[M:\GISDATA\Natural](#)

[M:\GISDATA\Trans](#)

[Aerial Photos](#)

The parent folder or directory for GIS data will always appear as the network drive M:\GISDATA.

Census2000 - census tracts, blocks, block groups and some associated data

The basic building units of census boundaries are tract, block group and block.

Counties are broken into tracts that contain roughly the same number of people. These tracts are then subdivided further into first block groups and then blocks. In an urban area, a census block is often the same as a city block. Below the census tract, in block groups and blocks, fewer attempts are made to keep similar population number in each.

Each county is assigned a county Federal Information Processing Standards code (FIPS). FIPS codes are regularly used to assign numbers instead of names to the counties (eliminates issues with spelling etc). Within each county, the census tracts are numbered, starting with one. Since each county starts at one, in order for the ID of the tract to be unique across the state, it must be combined with the state FIPS code. (If you're interested, the states all have FIPS codes too, and you'd need to add that code at the beginning to make the tract ID unique across state boundaries. FYI – Missouri's FIPS code is 29).

The numbering scheme continues through into block groups and blocks.

Demographic information is not included with these boundaries but must be downloaded as tables, generally in dBase (DBF) format. Demographic data is generally referred to as either SF1 or SF3 data. SF1 data is based on the short census form; SF3 is based on the long form. Because not everyone fills out the long form, the data in the SF3 table has been extrapolated for the total population. SF3 data is not available at the block level; certain SF3 data is not available at the block group level.

Demographic data must be joined to the boundaries to be used in GIS. Both the geographic file and the demographic file contain an ID number (called Tractlink, BGLink or blocklink) that allow these two files to be joined. There is also data aggregated to the county level.

Not all demographic data is available at all units. Data is limited at the block level because of privacy issues or estimation limits.

A few files already have this join done. Look for files in the census folders with demo at the end of the name (short for demographic, not demonstration).

There are also files for the 1980 and the 1990 census. The census created electronic boundaries for the first time in 1980. (We do not have a full set of boundaries for this census.) The 1990 census contains all the various levels of census boundaries, as well as demographic information corresponding to these boundaries. Tract, block group and block boundaries may change between censuses. Comparisons between decades may not be possible, but the Bureau of the Census does provide equivalency tables containing 1990 IDs and the relevant 2000 ID.

District bounds – political boundaries, district boundaries, city limits, ZIP codes etc. The majority of the layers in this folder are political boundaries, such as state boundaries, county boundaries, and city limits.

A note about ZIP codes: The layer that is called zipcode is based on the ZIP code tabulation areas (ZCTA) as defined by the U.S. Bureau of the Census. ZIP codes are not actually closed shapes at all, but are segments of roads assigned to a mail carrier for delivering the mail. Because ZIP codes are a popular unit of geography, the ZCTA boundaries were created. These boundaries are only

updated during a census year. The U.S. Postal Service (USPS) makes changes to ZIP codes regularly. These changes will not be reflected in the zipcode layer until the next census. There are private companies that offer more regular boundary updates, but their accuracy has not yet been determined.

A really good website that describes the issues with trying to do GIS using ZIP codes can be found at:

<http://www.oseda.missouri.edu/jgb/ZIP.resources.html>

A description on the creation of the ZCTA boundaries can be found at:

<http://www.census.gov/geo/ZCTA/zcta.html>

Facility – hospitals, daycares, nursing homes, trauma centers, schools, etc  
Facilities are represented as a single point for each facility. When a facility has multiple buildings, such as a hospital campus, the point is placed centrally. Locations are generated by geocoding the facility address. Several locations were further researched with aerial photographs and USGS quadrangle maps. Check the metadata for each layer for the methodology used.

Some facilities are listed from multiple years. The most current data is the file WITHOUT any year on it at all.

When layers are updated annually, the previous year's file will be renamed to the year it was based on, and the latest file will be created with no year in the name. For example, when the hospital layer was updated in March of 2004, the existing layer, simply called hospital.shp, was renamed to hospital03.shp. The newest layer was again called hospital.shp.

This naming convention allows us to create projects that reference the most current data. Otherwise, a project updated on an annual basis would need to have all the old facilities deleted and all the newer layers added.

Natural – basic geology, fault lines, rivers, streams, lakes, land use  
Information in this folder shows details about the natural features. There are several different layers dealing with water features, starting with the major features, such as the Missouri and Mississippi Rivers and ending with very small streams.

Trans – transportation: airports, railroads, highways, local streets  
The majority of the layers in this folder deal with the road network maintained by the Dept of Transportation (MoDOT). The road layer moroads is updated several times throughout the year. The various levels of highways (interstate, US highway MO highway) are selected from moroads into their respective layers. The updates in the moroads file are generally limited to roads maintained by MoDOT.

Quads and Aerial Photos – USGSdrg and AerialPhoto  
M:\GISDATA\USGSdrg

M:\GISDATA\AerialPhoto  
M:\GISDATA\AerialPhoto03

A full set of USGS 7.5 minute quads or quadrangles is on the server. The 7.5 minutes refers to 7.5 minutes of a degree of latitude and longitude covered by each map. The folder is USGSdrg. The naming convention for the subfolders (i.e. USGS11, USGS12) and files (i.e. C38094A1.TIF) are extremely confusing. To make it easier to identify which quad you need, a quad index shape file, quad24k.shp is provided. The MXD file, quadlocator.mxd, contains the basic map layers necessary to identify which quad you might need. In this MXD file, the electronic quad file is linked to each polygon in the quad24k layer. Use the hyperlink (shown as a lightning bolt icon), click on a quad feature, and the corresponding USGS quad will open in a separate window.

This set of quads is 'collar on'. This means that each quad contains all the borders, legends and other marginal data of the original document. This is handy when using a single quad, but these types of documents make a very poor backdrop to a map, because the margins from adjacent quads overlap. A set of quads 'with no collar' is available for download. If this is a dataset that would be more useful, please contact GIS staff and we will see about acquiring the quads in this format.

Aerial photos are available for each county. The folder is AerialPhoto. Because these files are so large, they have been divided into subfolders alphabetically. These photos are from the mid-1990s statewide flyover. You may also hear the term 'DOQQ'. This stands for Digital Ortho Quarter Quadrangle. These images were originally  $\frac{1}{4}$  of a USGS 7.5 minute quadrangle. Once they have been mosaicked into a county-wide image, this term DOQQ is not technically accurate for these images.

In 2003, a new flyover was done. These photos are leaf-on, infrared (false color) imagery. There is a sample in the AerialPhoto03 folder c-d for Cole County. These color images are even larger than their black and white counterparts and will only be added to the library as needed.

## **Metadata**

Metadata is data about data. Metadata exists for all layers in the most frequently used folders. There are a few folders containing data that does not have any metadata, but these folders are in the process of being deleted from the system and will soon not be available. The parent folders (i.e. Facility, Natural, Trans) also have basic metadata describing the content of the folder.

New data will not be placed in the library without metadata. If you find data without metadata, please contact someone on the GIS staff, because this is an oversight.

Metadata can be added to any layer using the metadata layer editor. There are required fields to meet the current federal metadata standards which are easily

identified while in the editor by the red piece of text that says REQUIRED, followed by a brief description of what data should go in that field.

## **Layers and Predefined Symbology**

### Layers

M:\GISDATA\Base Map Layers

The ESRI layer file (file extension LYR) can be easily confused with the generic term we use all the time for the layers of information we load into an ArcMap project. This folder can be accessed using either ArcCatalog or ArcMap.

The LYR file is extremely handy. It looks a lot like any of the other geographic files we have (e.g. geodatabase, shape file, coverage). The LYR file stores not only the location and type of geographic file, but also allows you to store symbology, label definitions, scale ranges and even allows you to give a more user –friendly name. Any of these geographic files can be made into a LYR file.

All LYR files continue to reference their underlying geographic file. If a geographic file is deleted or renamed, an LYR file that references it will still exist, but will now have a broken link. LYR can be copied and moved around and will continue to reference their geography. If you wish to copy a LYR to local drive, laptop or CD, the associated geographic file must be copied. Check the properties of the LYR file after you move it to make sure it is properly referencing the new data source location.

The folder Base Map Layers contains a set of the most commonly used map features. This folder was made to simplify adding data to a map project.

LYR files also exist in all if the frequently used folders. If you see a LYR file in the same folder as the layers you want to use, try selecting the LYR file over the 'plain' geographic file. These files are meant to save the user several steps.

LYR files can be created for groups of data. In the Base Map Layers folder are LYR files such as Medical Facility.lyr and Transportation – Highways.lyr. These are group layers. When you load these files, multiple features will be added. For instance, the Transportation – Highway.lyr will load Interstate, US Highway, Missouri Highways and lettered county roads.

The May 2003 GIS user group presentation more fully describes and illustrates all the things a LYR can do. This is located under M:\GISDATA\GISusergroup\presentations\May 03 - DHSS GIS User Group\_Layers.ppt.

### Symbology

M:\GISDATA\dhss.style

“The set of conventions, rules, or encoding systems that define how geographic features are represented with symbols on a map. A characteristic of a map feature may influence the size, color, and shape of the symbol used.” (ESRI – GIS Dictionary online)

A variety of symbol sets is provided as part of the ArcView software installation. Each set is categorized by topic or industry. Symbols are created for point line and polygons, as well as labels. Style sets are accessed using ArcMap.

You can create your own symbol set, copying the symbols you used the most into a single file for easier access. This is covered in the Advanced Symbology course.

The file extension for symbol sets is \*.STYLE. The DHSS style set is located under M:\GISDATA\dhss.style. The symbols in this style set correspond to the symbology defined in the LYR files.

### **Templates and Print Ready PDF maps**

ArcMap templates can be accessed using ArcCatalog or ArcMap. PDF print ready maps should be accessed using Windows Explorer.

#### Templates

M:\GISDATA\ArcMap Templates

Templates look very similar to ArcMap MXD files. They are a quick method of creating a map, frequently containing not only data, but also a predefined map layout (north arrow, title, legend etc). Templates are located under M:\GISDATA\ArcMap Templates.

#### Print Ready PDF maps

M:\GISDATA\PDF\_printableMaps

Maps that are frequently requested can be saved in PDF format for quicker access. PDF documents are located in M:\GISDATA\PDF\_printableMaps.

### **Upgrades – Service Packs and Patches**

As upgrades are available, they will be downloaded into this folder. Major upgrades will be pushed out through Application Explorer.

#### Service Packs (SP) and Patches

M:\GISDATA\ArcGIS Patches and Service Packs

Service Packs and patches are released between software versions to correct specific problems in the software. When software is installed, all available SP or patches are also installed. Most service packs and patches relate to functionality that is not relevant to DHSS.

SP and patches are available under M:\GISDATA\ArcGIS Patches and Service Packs. There are subfolders for 8.2, 8.3, 9.0, 9.1, 9.2. Only the latest version of SP or patch

need be installed, all the changes in any previous version are included. Service packs are generally pushed out to users from ITSD client support staff.

To find out what service pack you have installed, look under ArcGIS > Desktop Administrator.

## **Documentation and Tutorials**

### M:\GISDATA\GISdocumentation

The GISdocumentation folder contains a variety of 'How-to' documents for common procedures in ArcMap, ArcCatalog etc. This includes a full set of the documentation for ArcMap, ArcCatalog and ArcToolbox (included in the subfolder ESRI Digital Books). Also included is a extensive glossary of GIS terms. This folder should be accessed using Windows Explorer.

### M:\GISDATA\GISusergroup

Minutes and Presentations from the GIS user group are available in this folder under the GISusergroup folder.

### Tutorials

M:\GISDATA\ArcTutor

M:\GISDATA\GISdocumentation\ESRI Digital Books\Tutorials

This folder contains all the data for the tutorials provided by ESRI for their products. The tutorial documents are in PDF format, which must be accessed from Windows Explorer. Depending on which tutorial you are working, you may need ArcCatalog, ArcMap or ArcToolbox to complete the tutorial. Tutorials are included for several of the ESRI extensions, you must have access to the relevant extension to work through that tutorial.

*Updated September 2007*