

GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata

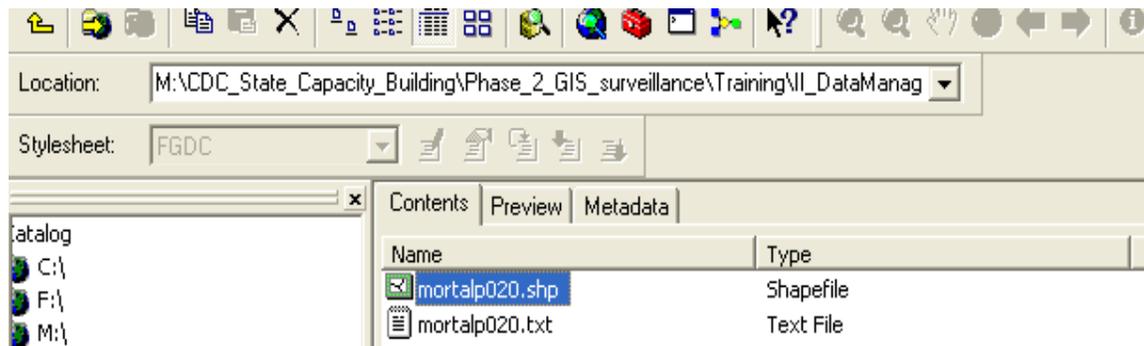
*** Files needed for exercise: *mortalp020.shp*

Goals: Our goal for this exercise is to examine metadata in both a text file and in ArcCatalog. We will also learn how to use the ArcCatalog metadata editor and to search for data using its metadata.

Skills: After completing this exercise, you should be familiar with examining metadata contained in a text file and how to associate this information with your data, as well as viewing metadata in ArcCatalog. You should also know how to use the ArcCatalog search function to find data using metadata keywords.

Importing metadata from a text file:

1. Open ArcMap. Click on the Add Data button and browse to *mortalp020.shp*.
2. Open the attribute table of the layer by right clicking the layer and choosing Open Attribute Table. We can see that while there are many attribute fields in this table, the labels are not informative. Close ArcMap.
3. Open ArcCatalog and browse to the location of *your mortalp020.shp* file. If there are metadata associated with the file as an .xml file, the software will read it and provide you with useful information on the dataset. You can check this by highlighting your shapefile and clicking on the metadata tab. If you do not see it, turn on your metadata toolbar (View > Toolbars > Metadata).



In the stylesheet dropdown select FGDC ESRI, this is an easily readable standard format that should provide some useful information about your data.



GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata

Contents | Preview | Metadata

mortalp020
Shapefile

Description	Spatial	Attributes
<p>Keywords</p> <p>Theme: REQUIRED: Common-use word or phrase used to describe the subject of the data set.</p> <p>Description</p> <p>Abstract REQUIRED: A brief narrative summary of the data set.</p> <p>Purpose REQUIRED: A summary of the intentions with which the data set was developed.</p> <p>Status of the data</p> <p>Time period for which the data is relevant</p> <p>Publication Information</p> <p>Data storage and access information</p> <p>Details about this document</p>		

Not much help here, in fact there appear to be no metadata associated with the shapefile at all!

- No description
- No spatial info
- No attribute info

4. These data are from a federal source, so you should expect some compliant metadata. Open Windows Explorer and browse to the text file: *mortalp020.txt* that came with the shapefile. This is a piece of comprehensive metadata that was created to accompany this data. This text file contains metadata for the shapefile and was packaged with the shapefile when downloaded from its source.
5. Open the *mortalp020.txt* with notepad to take a look.

```

C:\mortalp020 - Notepad
File Edit Format View Help
mortalp020.txt
Identification Information:
Citation:
Citation Information:
  Originator: U.S. Geological Survey
  Publication Date: 1999
  Title: United States Mortality Database, 1968-1992
  Publication Place: Reston, VA
  Publisher: U.S. Geological Survey
  Online Linkage: http://nationalit.us.gov/atlasfta.html
Description:
Abstract:
  This map layer contains mortality information for United States health service areas (HSA groups or counties). Included are mortality rates by sex and race (white and black) for the 11 leading causes of death, 8 major causes, and all causes combined. For 100,000 age-adjusted rates are presented along with indicators of whether each rate is considered reliable and whether it is significantly different from the corresponding U.S. rate. This is a revised version of the 1999 data set. This data set was also distributed with the name U.S. Mortality Database.
Purpose:
  These data are intended for geographic display and analysis at the national level, and for large regional areas. The data should be displayed and analyzed at scales appropriate for 1:1,000,000-scale data. The U.S. Geological Survey assumes no responsibility for the use of these data.
Supplemental Information:
  This data comes from Picelle L.W., Hargrove M., Jones G.R., White A.A., Atlas of United States Mortality, Department of Health and Human Services (DHHS) Publication, no. (DHD) 97-1401, Hyattsville, MD, National Center for Health Statistics, 1996. Additional information may be obtained from the home page of the U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, at: http://www.cdc.gov/nchs/
  The user is cautioned not to over-interpret apparent visual correlations between spatial patterns of mortality and environmental and/or socio-demographic factors. In many cases, patterns appear similar because some other unmeasured factor causes both the risk factor and the mortality patterns, not because the risk factor actually causes higher mortality. Users are urged to confirm suspected correlations by a more complete, multivariate analysis which includes known risk factors for the disease.
Time Period of Content:
Time Period Information:
Range of Dates/Times:
  
```

GIS II: Data Management: Creation, edition and maintenance of geographic data

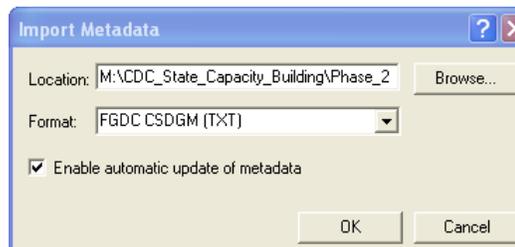
Module 3: Getting started with metadata

That is more like it; you can see some crucial information about the dataset. The text file is formatted to a metadata standard and so can be imported and associated with your shapefile as an .xml file that is easily read by ArcCatalog.

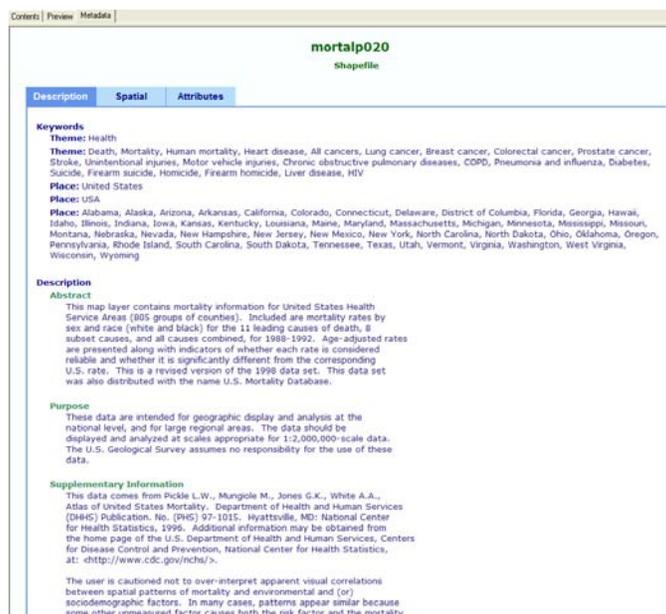
- Click on the import metadata icon to the right of your style sheet.



Import your metadata by browsing to the location of your *mortalp020.txt* file and selecting the TXT format option (this works because the data in your text file are formatted to a standard) and check the Enable automatic update of metadata box. If you want more information about anything in this dialogue use your Contextual question mark. When you are ready, click OK.



- Now take a look at your metadata, click through the three tabs and you should see that the essential information is now present.



This information is now associated with your shapefile as an .xml file. Take a look in window to confirm this.



GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata

8. You may notice that most of your attribute fields are still not very descriptive, however if you open the *mortalp020.txt* again using notepad you can see that the attribute fields are defined. If you want to add information to the metadata of your shapefile, you will need to do so manually using the metadata editor.

```
Attribute:
Attribute_Label: RateXXYY
Attribute_Definition:
  The age-adjusted death rate due to cause YY (see table below) per
  100,000 population, for group XX (see table below).
  >XX indicates the gender and ethnic group to which the rate applies,
  >where:
  > wm = white males
  > wf = white females
  > bm = Black males
  > bf = Black females
  >YY indicates the cause of death to which the rate applies, where:
  > 01 = Heart disease
  > 04 = All cancers
  > 05 = Lung cancer
  > 06 = Breast cancer
  > 07 = Colorectal cancer
  > 08 = Prostate cancer
  > 09 = Stroke
  > 10 = Unintentional injuries
  > 11 = Motor vehicle injuries
  > 12 = Chronic obstructive pulmonary diseases (COPD)
  > 13 = Pneumonia and influenza
  > 14 = Diabetes
  > 15 = Suicide
  > 16 = Firearm suicide
  > 17 = Homicide
  > 18 = Firearm homicide
  > 19 = Liver disease
  > 20 = HIV
  > 25 = All causes
Attribute_Definition_Source:
```

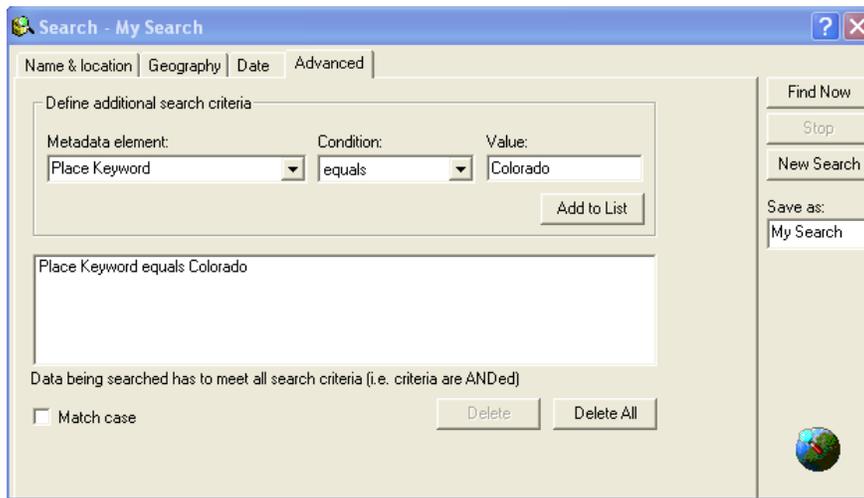
9. We will add information about selected attributes by editing the metadata a little later.

Searching by metadata:

1. Open ArcCatalog.
2. We want to locate a file that contains the latest sanitation districts for the state of Colorado. You know that it is located in a geodatabase that a coworker gave to you. Expand the *COdata.mdb* geodatabase and note that it contains multiple files with obscure names. Any one of them could be the sanitation districts we are looking for. We prefer not to go through them one by one.
3. Right click on the geodatabase and choose Search.
4. Click on the Advanced tab.
5. First we are going to look for data that is from Colorado and is recent. In the Metadata Element pull down list choose Place Keyword. In the Condition pull down list choose equals. In the Value list type Colorado.
6. Click Add to List. This adds a search string that will look for any metadata that has a Place Keyword Colorado.

GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata



7. Now repeat the process, but choose Temporal Keyword, equals, and 2008.
8. When these two search criteria are complete, click Find Now. ArcCatalog should find two of the files in the geodatabase. Both these files contain Colorado data from 2008. Because these files hold appropriate metadata, we can conclude that the other files in the geodatabase are either not actually data from Colorado or contain less recent data.
9. Now follow steps 3 - 8 as outlined above, but include one more search criteria: Theme Keyword, includes the word, sanitation.
10. You should only get one result, and we can conclude that this is the Sanitation District data that we are looking for. We can confirm this by looking at the attributes of this file, in the Name field. In this exercise we only had a handful of files to search through. However, imagine if you needed to find one file in a folder of hundreds. In such a case, metadata search is a very useful tool.
11. Note that you can actually search the entire Catalog with this tool. There is no need to limit yourself to one folder or database. This is useful when you know what file you need, but not where it is located.
12. Delete your search by right clicking My Search and choosing delete.



GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata

If you have time... update attribute fields using the ArcCatalog metadata editor:

1. Let's go back to *mortalp020.shp*, recall that we successfully imported some useful metadata but were not able to automatically update information on the mortality fields.
2. Open the *mortalp020.txt* using notepad and scroll down to attribute information.

```

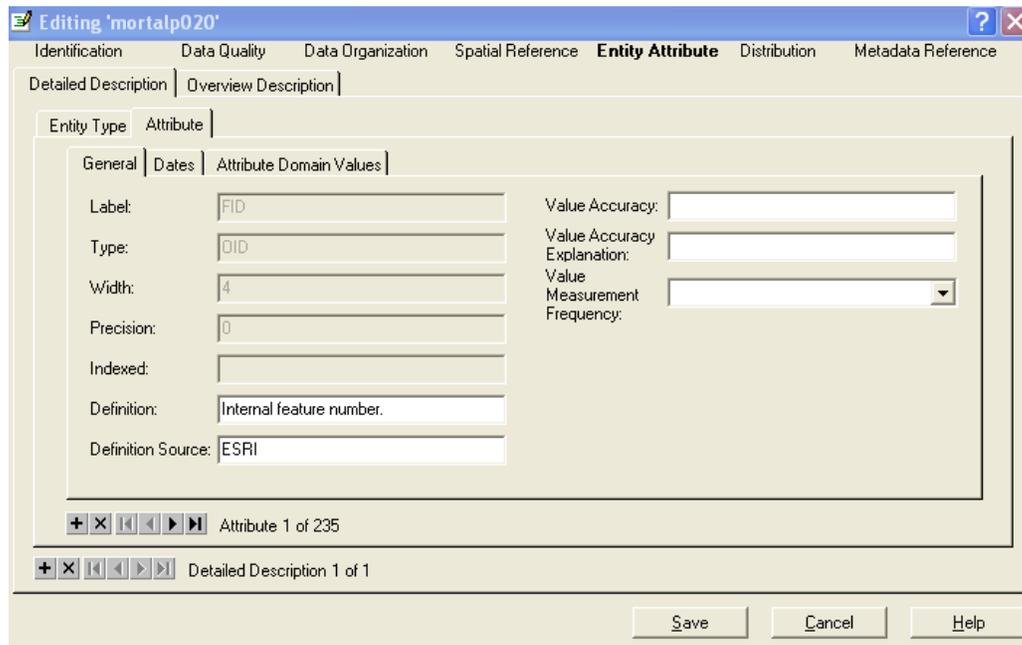
Attribute:
Attribute_Label: RateXXYY
Attribute_Definition:
  The age-adjusted death rate due to cause YY (see table below) per
  100,000 population, for group XX (see table below).
  >XX indicates the gender and ethnic group to which the rate applies,
  >where:
  > wm = white males
  > wf = white females
  > bm = Black males
  > bf = Black females
  >YY indicates the cause of death to which the rate applies, where:
  > 01 = Heart disease
  > 04 = All cancers
  > 05 = Lung cancer
  > 06 = Breast cancer
  > 07 = Colorectal cancer
  > 08 = Prostate cancer
  > 09 = Stroke
  > 10 = Unintentional injuries
  > 11 = Motor vehicle injuries
  > 12 = Chronic obstructive pulmonary diseases (COPD)
  > 13 = Pneumonia and influenza
  > 14 = Diabetes
  > 15 = Suicide
  > 16 = Firearm suicide
  > 17 = Homicide
  > 18 = Firearm homicide
  > 19 = Liver disease
  > 20 = HIV
  > 25 = All causes
Attribute_Definition_Source:
  
```

Keep this information open and handy.

3. We are going to add information to the metadata using the metadata editor. Open ArcCatalog and navigate to the *mortalp020.shp* file.
4. Open the Metadata Toolbar. Click Edit Metadata . Click the "Entity Attribute" tab then click the "Attribute" tab.

GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata



Here you can scroll through the attributes for the shapefile and make edits. You can also see the total number of attributes.

5. Scroll to 8 of 235 - RATWM01.
6. Using the attribute information included in the *mortalp020.txt* file we can decode this field:
 - RATE= The age-adjusted death rate per 100,000 population
 - WM= White Males (ethnic group)
 - 01= Heart disease (cause of death).
7. We are going to update two basic attribute fields in this shapefile. In the Definition box, write the definition that we have decoded. In the Definition source box, write the cited definition source for the field: *U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics*. These are two useful elements of attribute data that you have added.
8. Save your changes and take a look at the attribute field you have edited by selecting RATEWM01 in the attribute tab. You should see the definition and definition source for the attribute field you edited. Try the process with some other attributes.



GIS II: Data Management: Creation, edition and maintenance of geographic data

Module 3: Getting started with metadata

Description	Spatial	Attributes
Details for mortalp020		
<i>Type of object:</i> Feature Class		
<i>Number of records:</i> 3623		
Description		
Attributes		
FID		
Shape		
Area		
Perimeter		
Mortalp020		
Hsa		
Hsa_name		
RATEWM01		
<i>Alias:</i> RATEWM01		
<i>Data type:</i> Number		
<i>Width:</i> 20		
<i>Number of decimals:</i> 9		
<i>Definition:</i>		
death rate/10000 white males heart disease		
<i>Definition Source:</i>		
U.S. Department of Health and Human Services, Centers for Disease		