**Spatial Data Documentation Guidelines**

The easiest way to document spatial data is to use ArcCatalog and the metadata editor. The editor will populate many fields for you and you can build templates to deal with data that is the same for all your spatial data (such as contacts). You can save your entries at anytime, however, the save button will exit you out of the metadata editor. You then would need to select the **edit metadata** button to return to the editor. Your entries are not saved until you click on the **save** button.

The first thing you will want to do is build a template for your repeating information. Create a file and call it meta_template.txt. In ArcCatalog, click on the file, and click on the metadata tab in the right side frame. You will see default values. Go to the toolbar to the right of Stylesheet and select the edit metadata tool.

| Location: | J:\hja27\boundary |
| Stylesheet: | FGDC Classic |

⊞ 📁 radiation
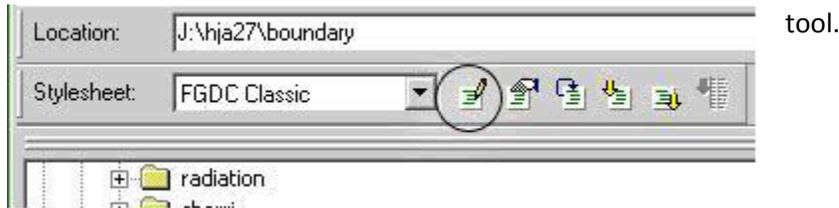
Complete any repeating information. Examples include: (inputs in bold) You can copy and past into the metadata editor.

1. General tab:

    A. Access Constraints (should be **Available on-line** unless restricted)

    B. Use constraints: **See data access policy at www.fsl.orst.edu/lter (especially the data use policy)**

2. Contact tab: fill out for person (as the primary contact)

    A. Primary Contact (name/organization/position)

    B. Phone and address information

    C. Citation: series/publication information:

    D. Publication place: **Corvallis, OR**

    E. Publisher: **Forest Science Data Bank**

3. Keywords:

    A. Place: Consider putting the following if your data is in the Andrews:
        Oregon, Willamette Basin, Blue River Watershed, HJ Andrews Experimental Forest (4 separate entries, using the + button). If you need to add additional place keywords, you can add them for each spatial data layer you document.

    B. Theme: Consider putting **Andrews Forest LTER Thesaurus** for the Theme Thesaurus (if this is where you will be getting your theme keywords).

4. Metadata Reference:
   fill in the contact information like you did for the contact tab. This might be the same information as your contact information, but it might be different (you might want one person as contact for the data, another about the metadata).

5. Distribution Tab:

   A. General

      a. Resource Description: **Downloadable Data**

      b. Distribution Liability: **While substantial efforts are made to ensure the accuracy of data and documentation, complete accuracy of data sets cannot be guaranteed. All data are made available "as is". The Andrews LTER shall not be liable for damages resulting from any use or misinterpretation of data sets.** (go ahead and copy and paste this block into the template) This is from the Andrews data use policy.

      c. Custom Order Process: Call contact person for instructions and costs.

      d. Distributor: Fill in the contact information like you did for the contact tab. This might be the same information as your contact information, but it might be different.

      e. Standard Order Process: Digital Form should be selected

      f. Fees: **none**

      g. Turnaround: **as time permits**

      h. Ordering Instructions: **Obtain information off of WWW site, call contact person for special requests.**

   B. Save and exit the editor. This text file should be filed somewhere where you can find it whenever you need to document spatial data. When you saved the data, you created an .xml file with the same name as your original text file.

For each geospatial data set (Coverage, shapefile, feature class, image, grid):

To start your documentation of spatial data, with ArcCatalog,

1. Select the data coverage, shapefile, featureclass, or image.

2. Select the Metadata tab in the right frame/window.

3. Click on the **import metadata** button on the toolbar.

4.
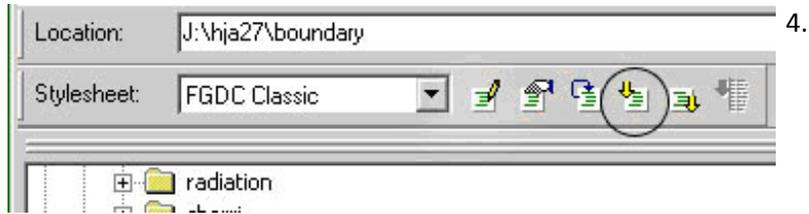
5. Select your template file (browse to where it is stored on the computer and select the .xml file to import). Your template information will be now visible in the right hand frame.

It is important to do this import before you start editing the metadata, as the import will write over any existing metadata.

6. Click on the **edit metadata** button on the toolbar

7.
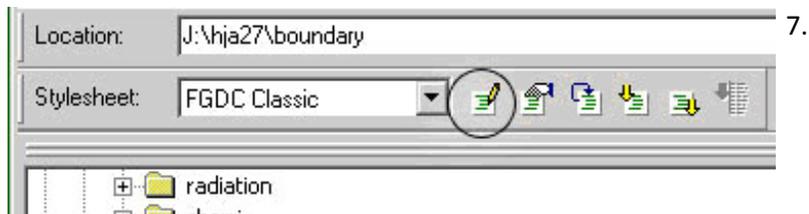
8. The metadata editor will pop up and you can begin editing your metadata.

Editing 'Adminstrative boundary, Andrews Experimental Forest'

Identification   Data Quality   Data Organization   Spatial Reference   Entity Attribute   Distribution

General | Contact | Citation | Time Period | Status | Spatial Domain | Keywords | Browse Graphic | Security |

Description

Abstract: The Administrative Boundary of the HJ Andrews Experimental Forest

Purpose: The boundary was developed to show the administrative and watershed boundary of the HJ Andrews Forest.

Language: en

Supplemental Information: 158_01

Access Constraints: Available on-line

Use Constraints: See data access policy at www.fsl.orst.edu/lter/data (especialy the data use policy)

Data Set Credit: Data Set created by Tere Turner, Blue River Ranger District, WIllamette National Forest

Native Data Set Environment: Microsoft Windows XP Version 5.1 (Build 2600) Service Pack 2; ESRI ArcCatalog 9.2.1.1332,Microso Version 5.1 (Build 2600) Service Pack 2; ESRI ArcCatalog 9.2.1.1332

Native Data Set Format: ArcInfo Coverage

Save   Can

9.

10. Identification Tab:

    A. General:
      Complete the abstract, purpose and data set credit (if there is another person you would like to identify). Follow the instructions for abstract/purpose found at the top of this document. The rest of the data fields should be completed for you.

    B. Contact Information:
      This tab should be completed for you from your metadata template. You can edit this if you need to make changes.

    C. Citation Information: General tab

        A. Provide a descriptive title for the data set. The default is the computer name of you data. You will want a more descriptive title.

        B. Originator: You can list one or more people

        C. Publication Date (you can put the date the data was/is going to be published. If this is going to be it is first published date, then you can put in this year).

D. Online linkage: copy and past this for the spatial data link on the Andrews website: **http://www.fsl.orst.edu/lter/data/spatialcatalog.cfm?topnav=160**

E. Other Citation Details: you can enter the computer name for the dataset here as a cross-reference. You can also put any other FSDB study codes that relate to this data (if applicable).

D. Citation Information: Series/Publication tab

A. If this is part of a series, indicate name and number

B. For publication place, you can put in Corvallis, OR with Forest Science Data Base as the publisher (you can also list another publication source if applicable). If this is part of a larger piece of work, fill out the information in the larger work citation tab.

E. Time Period:

A. Currentness Reference: choose **publication date** or **ground condition**. Ground condition would indicate date indicates when the data was collected. Publication date indicates the time when the data was recorded or published. You can type in another value, such as **photo date**.

B. Indicate if this is a single date/time, multiple dates/times or range of dates.

C. Add calendar date (you do not have to enter time of day unless it is relevant).

F. Status:

A. Progress: indicate if complete, in work, or planned

B. . Update Frequency: chose from pull down list. Most will be **as needed**

G. Spatial Domain:

A. General: Enter min/max altitude for the study area. Enter altitude units.

B. Bounding coordinates will be calculated by the computer.

C. G-Ring information is optional, but it allows you to group areas that are not connected (for example, the Hawaiian Islands as a state are a group of islands). You can also exclude areas with G-Ring information (for example, your area is in the Willamette Basin, but not in the HJ Andrews. This **donut hole** would be excluded from your description by using G-Rings). See spatial data manager if you want to use G-Rings to identify the spatial domain. For most studies, bounding coordinates will work.

H. Keywords:
Please use the Andrews LTER Preferred Keyword List available on the web for your theme keywords. For place, if you put the general Andrews place keywords in your template, then add any specific place keywords, including specific sites (watershed1).

Check the place keyword list for specific place keywords already defined. Include at least 3 of each.

I. Browse Graphic:
The browse graphic is an image of the data set. The image allows prospective users to more beyond textual descriptions and see what the data set looks like. The image might show a simple display of the data set, the results of an application that used the data set, different aspects of the quality of the data set, or other information. This is an optional field.

  A. Give the file a name, description, and type. You can have more than one browse graphic. You would want to include this file in a zip file when you submit the data to be added to the FSDB.

J. Security:
Select security classification from pull down list. Most cases should be unclassified, unless restricted per data use policy.

K. Cross Reference:
This is where you fill in a citation to any publications you know of that have used this data or any other data sets that are related. Fill out the form.

11. Data Quality:

  A. General:

i.   Logical Consistency Report: This describes the relationships encoded in the data structure of the digital data. The report shall detail the tests performed and the results of the tests. For legacy data, this field tends to be used as a general statement about the data quality, such as fair, poor, good). If you did run any tests to check for label errors, topology problems, this is where you would indicate the tests you ran. You can also specify the date of the tests in the main text block.

ii.   Completeness Report: Information about omissions, selection criteria, generalizations, definitions used, and other rules used to derive the data set. Again, you might want to use a general description here, for example: data consists of vegetation inside the boundary of the HJ Andrews and was limited to managed stands).

iii.   Cloud cover: this would be for remotely sensed imagery where the amount of cloud cover is relevant.

  B. Attribute Accuracy: An assessment of the accuracy of the identification of entities and assignment of attribute values in the data set.

.   Accuracy Report: Text explanation of the accuracy of the attribute values in the data sets and a description of any tests used, for example: Attribute accuracy is excellent. A report was generated to identify all the possible attributes and their values, all errors were corrected.

i.   Value: an estimate for the accuracy of the attribute values in the data set. Options include unknown, good, bad, very bad, very good, +-95%.

ii.   Attribute Accuracy Explanation: the identification of the test that yielded the attribute accuracy value.

        C.   Positional Accuracy: An assessment of the accuracy of the positions of spatial objects. (* while good to identify the accuracy of the positions, if you do nott know what the accuracy is, put unknown or write in the report your best guess) You can have more than one assessment, especially if your data is made up of a mixture of data sources or data collected at different dates, with different GPS accuracies.

.   Horizontal Accuracy: an estimate of accuracy of the horizontal positions of the spatial objects

i.   Accuracy Report: Text explanation of the accuracy of the horizontal positions in the data sets and a description of any tests used, for example, +- 40 meters (or some distance based on GPS readings and report from GPS recorder).

ii.   Value: an estimate for the accuracy of the horizontal positions in the data set. Options include unknown, good, bad, very bad, very good, +-95%.

iii.   Explanation: the identification of the test that yielded the horizontal accuracy value.

iv.   Vertical Accuracy: an estimate of accuracy of the vertical positions of the spatial objects

v.   Accuracy Report: Text explanation of the accuracy of the vertical positions in the data sets and a description of any tests used, for example, +- 40 meters (or some distance based on GPS readings and report from GPS recorder or from the accuracy of a contour map).

vi.   Value: an estimate for the accuracy of the vertical positions in the data set. Options include unknown, good, bad, very bad, very good, +-95%.

vii.   Explanation: the identification of the test that yielded the vertical accuracy value.

        D.   Source Information:

.   General: (you can have more than one source)

        i.   The denominator of the representative fraction on a map (for example, on a 1:24,000 scale map, the Source Scale Denominator is 24000).

        ii.   Type of Source Media: choose from drop down list or type in another free text value. (where your data came from)

        iii.   Source Citation Abbreviation: short-form alias for the source citation (do not worry about this one)

        iv.   Source Contribution: brief statement identifying the information contributed by the source to the data set, for example, data was copied from USGS topographic map.

i.   Source Citation: same set up as other citations. Here you are giving credit to your source. Fill this out if it you need to give credit to the data source, otherwise leave it blank.

ii.   Source Time Period of Content: Time period(s) for which the source data set corresponds to the ground or was published.

> i.   Currentness Reference: published, ground condition, or perhaps photo date.
>
> ii.   Single date/time, multiple date/times, or range (select one)
>
> iii.   record calendar date, do not worry about time unless important.

iii.   Process Step: This is where you document your methods to create the spatial data. You can have multiple process steps. Some times the software will add a step when you edit the metadata or perform other functions on the data. Some of the fields may be filled in automatically.

> i.   Process Description: Describe the step/steps you took to create/generate the spatial data, for example, The coverage was digitized from USFS Primary Base Series topographic maps and edited in ArcEdit. The dataset was built (completed topology). You can copy/paste from existing documentation.
>
> ii.   Process Software and Version: Identify the software you used and version, for example: Arc/Info version 9.1
>
> iii.   Process Date: The date the process was completed. This can be unknown.
>
> iv.   Process Time: Optional, do not spend any time trying to find this unless you have it handy.
>
> v.   Process Contact: Contact name for the person who did the process. Similar to other contact information.
>
> vi.   Source Used Citation Abbreviation and Source Produced Citation Abbreviation: This section is important in documenting the lineage of data, especially as data is updated and/or generated from different sources. Used Citation Abbreviation of a data set used in the processing step (would have filled out the source citation section in the Source section) and the produced citation abbreviation is from an intermediated data set that is 1) significant in the opinion of the data producer, 2) is generated in the processing step, and 3) is used later in processing steps. You might not document or keep this intermediate data set.

12. Data Organization:
    This tab should be complete by the computer program. If it is not, you will want to use the

update metadata button from the toolbar. If this section is not updated after using the toolbar, there might be a problem with your data. Contact the spatial data manager.

13. Spatial Reference:

    A. General: This information should be complete by the computer program. If it is not, you will want to use the update metadata button from the toolbar. If this section does not update after using the update toolbar, your spatial reference most likely is not defined. Contact the spatial data manager for help in defining your spatial reference. If the spatial reference does not look correct, contact the spatial data manager for assistance.

    B. Horizontal Coordinate System:
       Again, these fields should be completed for you. Note the Planar Distance Units: Values would be meters, feet, some fraction of meters.

    C. Vertical Coordinate System: **Applies only if you are dealing with elevation data.** You will need to fill out this section, depending on if you are dealing with above water level or below water level:

    . Altitude System Definition:

          i. Datum Name: select from pull down list.

          ii. Distance Units: Units in which altitudes are recorded: select from pull down list.

          iii. Encoding Method: the means used to encode the altitudes: select from pull down list. (for DEMs, this would be Attribute values)

          iv. Resolution: the minimum distance possible between two adjacent altitude values, expressed in Altitude Distance Units of measure.

    i. Depth System Definition: (bithemitry)

          i. Datum Name: select from pull down list.

          ii. Distance Units: Units in which depths are recorded: select from pull down list.

          iii. Encoding Method: the means used to encode the depths: select from pull down list. (for DEMs, this would be Attribute values)

          iv. Resolution: the minimum distance possible between two adjacent depth values, expressed in Depth Distance Units of measure.

14. Entity Attribute:
    Users of a data set need to know the meaning of entity, attribute, and attribute value information associated with the spatial information. For example, a data set might include the entity "road". A "road" might have the attribute "road type," which can be assigned the

attribute values of "heavy duty," "medium duty," "light duty," or "trail." The producer of the data set may have different definitions for "road," "road type," "heavy duty," "medium duty," "light duty," or "trail" than a user. The Entity and Attribute Information section provides the way for a producer to describe the meaning of this nonspatial entity, attribute, and attribute value information so a user can understand the information content of a data set and use the data appropriately. If you have metadata already completed for these attributes, you can skip this section and reference them through #12 Overview Description.

A. Detailed Description: description of the entities, attributes, attribute values, and related characteristics encoded in the data set.

. Entity Type: the definition and description of a set into which similar entity instances are classified.

1. Label: The tables/entities will be listed by the computer program. You will move through the tables associated with the GIS layer. Document the definition of the table and the source, if known. Spatial data table examples include:

a. layer_name.pat (polygon or point attribute table) for coverages

b. layer_name.aat (arc attribute table for coverages)

c. layer_name.rat.routename (route attribute table)

d. layer_name (dbf table for shapefiles and tables for geodatabases)

e. grid_name.vat (value attribute table for grid)

B. Attribute: the name of the attribute (go to this tab for every table associated with the layer). This is where you define your fields/attributes. The editor will populate the fields, and you just need to define them. If the values are completed in the form, you only need to change if they are incorrect.

. General:

. Label: this is the name of the field in the table. Use the + button at the bottom left of the form to move to the next value. You will want to fill out the Definition, Definition Source (if you have one), Value Accuracy (if you can determine), the value accuracy explanation (if you have a value accuracy),, and the value measurement frequency (from pull down list or add another)

i. Dates: provide a beginning and ending date(s) for the values if known/applicable (one for each field). This could be useful if you are collecting data values that vary in time frame.

ii. Attribute Domain Values: the valid values that can be assigned for an attribute. A domain is the set of possible data values of an attribute. From the example used above, the domain for the attribute "road type" consists of "heavy duty," "medium duty," "light duty," and "trail." (choose one of the following) and select the circle:

a. Enumerated Domain: An enumerated domain is one comprised of a list of values. The "road type" attribute has an enumerated domain which contains the values "heavy duty," medium duty," "light duty," and "trail." In this case, the list of possible values, the definitions of the values, and the sources of the definitions should be provided. Complete the following for each value in the table: This is where a summary of the values in the field will come in handy.

    a. value

    b. value definition

    c. value definition source

b. Range Domain: A range domain is one comprised of a sequence, series, or scale of (usually numeric) values between limits. For example, an attribute of age might have a range domain of integers from 0 to 100. In this case, the minimum and maximum values should be provided.

    a. Minimum

    b. Maximum

    c. Mean

    d. Standard Deviation

    e. Attribute Units of Measure

    f. Attribute Measurement Resolution

c. Codeset Domain: A codeset domain is one in which the data values of defined by a set of codes. Examples include the Federal Information Processing Standards that contain numeric codes for nations, States, and counties. In this case, the title of the publication containing the code set and the source of the codeset should be provided.

i.   Name

ii.   Source

d. Unrepresented Domain: An unrepresented domain is one for which the set of data values cannot be represented. Reasons include attributes whose values do not exist in a known, predefined set (for example, the values for an attribute of names), or attributes whose values cannot be depicted using the forms of representation (available character set, etc) used for the metadata. In these cases, the information content of the

set of values should be provided. Type in a description of the code.

15. Overview Description: **If you have completed #11 Entity Attribute, you can skip this section.**

    A. Overview Description.

    B. Dataset Overview: summary of, and citation to detailed description of, the information content of the data set.

.   Entity and Attribute Overview: detailed summary of the information contained in a data set.

i.   Entity and Attribute Detail Citation: reference to the complete description of the entity types, attributes, and attribute values for the data set

16. Distribution: (this should have been filled out from your template). Make any updates/changes as necessary

17. Metadata Reference: This references the metadata standard that was used to complete the metadata. It will be completed by the editor. You will want to check and make sure your metadata contact is correct (this would be imported from your template)

18. **Save and exit. Congratulations!**

Your metadata is now stored with your spatial data. You can update it at any time by viewing the metadata, and clicking on the edit metadata tab.

19. If you have a series of data sets, where the information is the same, except for changing title, year or perhaps date, (data by month/year), you can complete metadata for one shapefile/coverage and import it into the others. You then need to edit the metadata through the editor, to change any needed fields. Be sure to update any process steps that might be be specific to the data.

20. As you update your spatial data, and create new data through geoprocessing in ArcGIS, these processes will be updated in the metadata under Geoprocessing History (view metadata with the FGDC Stylesheet to see).

21. Provide a copy of your data to the spatial data manager. The manager will check for completeness and discuss the organization of the data within the FSDB data structure.

last updated: 05/14/2008