**Best Practices for documenting geospatial data Version 5.0 3/8/2012 --Draft-**

**Updated by Leanne Lestak, 12/4/12**

**Theresa Valentine 12/18/2012 (raster file information)**

**Theresa Valentine 1/30/2013 (continued with raster information)**

**File Naming Conventions:**

1. No blanks between words
   1. Yes: boundary\_hja.shp
   2. No: boundary hja.shp
2. Keep your file names short (too long and they truncate). You can put long title in metadata
3. Keep your attribute titles short, as they tend to truncate as well, especially when you make several overlays

**Step 1: Create metadata using an editor**

1. Using ArcGIS 10.0 or higher editor:
   1. Create a template of repeating information to import into the editor.
   2. Be sure that under **Customize** tab, ArcCatolog options, metadata that you select the FGDC CSDGM Metadata style (not the default).
      1. Open the ***Options*** dialog box for your ArcGIS Desktop application.
         1. In ArcMap, click **Customize**>**ArcMap Options**.
         2. In ArcCatalog, click **Customize**>**ArcCatalog Options**.
         3. In ArcGlobe, click **Customize**>**ArcGlobe Options**.
         4. In ArcScene, click **Customize**>**ArcScene Options**.
      2. The ***Options*** dialog box appears.
      3. Click the **Metadata** tab.
      4. Click the drop-down arrow and click the style of metadata you want to create. (FGDC)
      5. If you're using the **Description** tab when you choose a new metadata style, you won't immediately see the results of that change. Click another tab in ***ArcCatalog*** or the ***Item Description*** window, such as the **Preview** tab, then click the **Description** tab again for the new metadata style to take effect.
      6. You can synchronize metadata updates when you like by using the  [Synchronize Metadata](http://help.arcgis.com/en/arcgisdesktop/10.0/help/0012/001200000011000000.htm) geoprocessing tool
   3. If you have ESRI metadata already completed in a prior version of ArcGIS, then you need to upgrade this to ArcGIS metadata. Note: a metadata document may have been created for your data in a previous version of ArcGIS. It’s possible that this default document was never edited, but it contains some default items that may show up in your final products. Two ways to avoid this problem are:
      1. Create a template of your repeating data and import that first. Doing this will solve the potential problem.
      2. Or look at the FGDC read-only data (view it) at the bottom of the description, and if you see any REQUIRED text, open the xml file in an editor, **This step should only be used if you have added metadata to the esri editor, and don’t want to lose what you have entered by importing your contact information.** Find the Binary block, and delete the following:
         1. <Enclosure><Descript>**original metadata**</Descript> <Data SourceMetadata="**yes**" OriginalFileName="**source\_metadata.xml**" and another big block of gibberish.    You need to get rid of everything in the <enclosure> all the way to the final </Enclosure>…then save your xml file, exit out of ArcCatalog, then go back in to ArcCatalog and look at the metadata.  It should have some of your data in the “FGDC metadata” but not any “REQUIRED” text. Continue with editing (**d. iv below**), follow the step of exporting the arcgis metadata to FGDC, and the correct fields will be generated.
      3. In Toolbox, under Conversion tools>Metadata: Select the Upgrade Metadata Wizard, select the source metadata (your original data layer), upgrade type (FGDC\_TO\_ARCGIS) This should update your original ESRI FGDC metadata into the new ArcGIS metadata format.
      4. Begin editing your metadata now that it is converted (you may need to make some changes).   
         **Note: Upgrading or importing metadata will over-write any metadata already entered. It’s important to do this step first!**
         1. Click on the **Description** tab in ArcCatalog
         2. Select the **Edit** button;
         3. and move through the tabs.
         4. **EML Best Practices:**
            1. Only enter one name (first name, middle initial, last name) per contact role. No prefix/suffix. You can add another name with another role.
            2. Do not list more than one on-line distribution.
            3. Fill in link to the data under:

Resource>Distribution>Digital Transfer Options>Online Resources (copy a url to the data here). For the GeoNIS, this should be a zipped file with the data that meets the specified requirements.

* + - 1. Save and exit (saving will exit you out of edit mode)
      2. When finished, **move to step 3.**
  1. If you don’t have FGDC metadata started for your data, do the following:
     1. Suggest you create a template to store your repeating information. This would be your contact information, data use polices, etc (see example 1). Basically, you create a text document with nothing in it, and create metadata for that document.
     2. **Example 1:** **Contact template:** These are samples of fields filled out for the HJ Andrews Geospatial databases. Create a small text file for the purpose of creating your metadata template.
        1. **Tags:** Blue River Watershed, Willamette Basin, Oregon, HJ Andrews Experimental Forest
        2. **Use Limitation:** 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.
        3. **Place Keywords:** put only one keyword per line
           1. Blue River Watershed
           2. Willamette Basin
           3. Oregon
           4. HJ Andrews Experimental Forest
        4. **Resource Citation Contacts:**
        5. **Organization:** Forest Science Data Base
        6. **Role:** Publisher
        7. **Metadata Contacts:** (put in info here about person who is responsible for metadata)
        8. **Security Constraints: classification:** Unclassified (all of our data is unclassified)
        9. **Resource:** 
           1. **Contacts:** enter the person who is the point of contact for the datasets (in this case, it’s the same person for all the data, so it is entered in the template)
        10. **Distribution Information:** Add contact here if one person responsible for distribution and add ordering process if same for all your data.
     3. Import the template into the data set.
        1. (use import tool on Description tab)
     4. Begin editing your metadata. **Note: Upgrading or importing metadata will over-write any metadata already entered. Remember to do this first**! Look at **section c** above, as there may have been an xml file created in a previous version of ArcGIS. Remember that you can copy/paste from other documents. Several tabs will be completed for you. The attribute information will be completed under **Fields.** The major attribute file will be under Details (with the name of the layer). Click on attributes that are not computer generated, and add your definitions. Refer to the LTER EML Best Practices for attribute information. ArcCatalog metadata editor will allow you to add more than one domain for one attribute. EML does not like this, so you might need to edit your eml file. Be careful about using unrepresentable domains, as it appears that you can’t delete them once they are added.
     5. Click on the Description tab in ArcCatalog
        1. Select the Edit button
        2. Edit environment will appear, and move through the tabs.
        3. **EML Best Practices:**
           1. Only enter one name (first name, middle initial, last name) per contact role. No prefix/suffix. You can add another name with another role.
           2. Do not list more than one on-line distribution.
           3. Fill in link to the data under:

Resource>Distribution>Digital Transfer Options>Online Resources (copy a url to the data here). For the GeoNIS, this should be a zipped file with the data that meets the specified requirements. Note that you can put in the full url or just the filename, (see section below on customizing).

* + - * 1. Type the entity name (computer name of the file) in the Overview section of the ArcGIS metadata, under Citation, scroll to Other Details and put the name in the text block. This is important to get the entity name into your EML document. (see below on customizing)
      1. Save and exit (saving will exit you out of edit mode)
      2. When finished, move to step 3.
      3. **Hint:** If you have multiple data sets with similar metadata (for example, the month is the only change on a series of precipitation grids), you can complete one metadata record, import this metadata record into the dataset for the next month, and then edit the title/etc. This saves a lot of time when dealing with similar data sets.
      4. **Link to esri metadata help:** <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/What_is_metadata/003t00000001000000/>
      5. **For Raster Data:**
         1. Note that you will need to make sure that your cell size resolution is indicated under Spatial Representation Information, under grid spatial representation, dimensions for the x and the y axis. You might want to check the same dimensions under the georectified representation also (same spot under dimensions). Note there may be more information you need to include with Rasters, that aren’t in FGDC.  See reference here: <http://knb.ecoinformatics.org/software/eml/eml-2.1.1/eml-spatialRaster.html>
         2. Adding the entity information: When editing in ArcGIS Version 10.1, under **Resources, Fields**, you need to create **a new entity and attribute detail**. Under **label,** put the computer name of the raster (this is what becomes the entity title). Next, call it a “raster”  for **Object** and gave it a definition “raster data set”, and put in a **definition source** of “dataset”.  The last two items aren’t really used, but you need to fill them in.
         3. Add the definition of your raster value as a new attribute, and fill that in.   This information is also captured under the content/image description, but it’s not getting transferred to the FGDC standard from ESRI from the image location. This is because image metadata is not part of the core FGDC standard.  Be sure to add the new domain. The computer program doesn’t automatically fill the information in here, (like it does under the image description), so you have to fill it in yourself.
         4. If you have a VAT or an RAT  (value or raster attribute table),  you would add them in here as additional entity and attribute details.
    1. **Go to step 2.**

1. Using ARCGIS 9.3 editor:
   1. Edit data with the metadata editor
   2. Export your file to an FGDC xml format (usually the default in this editor)
   3. **Go** **to step 2**
2. Using ArcGIS 10.1 editor: This is similar to the 10.0 editor, but it offers additional functionality in that there are edit checks. You will get red checks when elements aren’t completely filled in. This is useful, but sometimes you can’t figure out where to put the changes in, so you might have to leave them x’s. This does make for more complete, and fewer errors when you transform to EML. Remember to export out to a FGDC xml document and then go to step 2.
3. Using non-esri editor:
   1. Save your metadata information/file to an FGDC xml document.
   2. **Go to** **step 2.**

**Step 2. Export to FGDC or raw ARCGIS metadata xml file**

1. There are two methods for creating valid EML: when creating FGDC or other metadata files, do not name them the same as your source data. This will cause problems: for example, if you have an shapefile called roads.shp and you create an FGDC xml file called roads.xml, then this will not show up to import or transform. It is suggested you place these files in a different folder on your computer (where you could call it roads.xml, but don’t put any other spatial data there. ESRI gets confused). Create a folder called Metadata\_docs
   1. Export the ArcGIS metadata to FGDC format and run that through the parser. There are some metadata tags that will be dropped during this process. This is particularly a problem with raster data. Use the export button on the description tag, or you can use the toolbox to export (data conversion/metadata)
      1. This will create a new FGDC xml file. You will indicate an output file location to put the file. **Go to step 3.**
   2. Use the ArcGIS metadata file (raw xml) and run it through the transformation. If you use this method, you would **go directly to step 3** and don’t create another file.
   3. If you have FGDC data outside of esri products, then **go to step 3**.

**Step 3. Customize esri102eml21.xsl stylesheet and prepare for transformation:**

Modify the esri102eml21.xsl stylesheet for your site. Available here: (<http://im.lternet.edu/node/243/release>) look for the latest version:

* 1. If you don’t have your own site data policies, the LTER Network Defaults will be used to populate that section of your EML document.
  2. Search for “AND” in the stylesheet, and adjust for your site.
  3. There are a couple of additional places you might want to customize.
     1. <spatialRaster> the link in the current esri2eml stylesheet is set up to look for the entity name under the following area in the ArcGIS metadata: under overview citation, you need to place the entity name (file name) under “Other Details”. If you want to pull the entity name from another location, you need to change the stylesheet: The links are highlighted from the portion of the stylesheet, under the spatialRaster tag. **Note that putting the entity name in alternate title in the citation will not transfer over to the FGDC metatdata (using the current esri translation).**
        1. <xsl:if test="metadata/idinfo/citation/citeinfo/othercit!=''"> this vaue can ge <alternateIdentifier>  
            <xsl:value-of select="metadata/idinfo/citation/citeinfo/othercit"/> </alternateIdentifier>  
            </xsl:if>  
            <entityName> <xsl:choose>

<xsl:when test="metadata/eainfo/detailed/@Name!=''">

<xsl:value-of select="metadata/eainfo/detailed/@Name"/>  
</xsl:when>  
<xsl:whentest="metadata/idinfo/citation/citeinfo/othercit!=''">  
<xsl:value-of select="metadata/idinfo/citation/citeinfo/othercit"/> </xsl:when>  
<xsl:otherwise>  
 <xsl:value-of select="'Name of the resource not known'"/>  
 </xsl:otherwise>  
 </xsl:choose>  
 </entityName>

* + 1. Also note where you need to change the link to your data (the link provided is where the Andrews stores the data)…  if you put the path in your arcgis metadata, then this line in the stylesheet should look like this:

<online>

<url> <xsl:value-of select="metadata/idinfo/citation/citeinfo/onlink"/></url>

</online>

be sure to include the full path to the data: example <http://fsl.orst.edu/mydatastores/here_is_my_file.zip> Or ftp://ftp...........

* + 1. This part of the EML shows what the format of your distribution file is and the on-line distribution link.

<xsl:choose>  
       <xsl:when test="/metadata/spdoinfo/direct ='Vector'">  
            <dataFormat>  
                <externallyDefinedFormat>  
                    <formatName>zip file with ARC/INFO shape file   <!--the format we use on the AND-->  
            </formatName>  
            </externallyDefinedFormat>  
                 </dataFormat>  
         </xsl:when>

<xsl:when test="/metadata/spdoinfo/direct ='Raster'">  
      <dataFormat>  
            <externallyDefinedFormat>  
                  <formatName>zip file with ascii file    <!---again, the format we are going to be using-🡪  
                 </formatName>  
          </externallyDefinedFormat>  
     </dataFormat>  
   </xsl:when>  
 </xsl:choose>  
        <!-- check for content -->  
       <distribution>  
                <online><url>[ftp://ftp.fsl.orst.edu/pub/lter\_web/<xsl:value-of](ftp://ftp.fsl.orst.edu/pub/lter_web/%3cxsl:value-of) select="metadata/idinfo/citation/citeinfo/onlink"/> </url>  <!--path to our data.. use this type of format if you don’t want the full link in your arcgis metadata-->  
       </online>  
    </distribution>

* + 1. PackageID: be sure to put in your packageID according to LTER guidelines. Example: knb-lter-and.145.1 (for package 145 version 1 from the Andrews LTER). You can add your prefex to the stylesheet if you want to just add the number at the end. You would have to modify the stylesheet to make this happen.

1. If you are using ArcGIS Version 10.0: Copy the following files to your computer:
   1. esri102eml21.xml (<http://im.lternet.edu/node/243/release>)
      1. Place in <WHERE YOU HAVE STORED ARC SOFTWARE ON YOU R COMPUTER>\ArcGIS\Desktop10.0\Metadata\Translator
         1. EXAMPLE: C:\Program Files\ArcGIS\Desktop10.0\Metadata\Translator
   2. esri102eml21.xsl
      1. Place in subdirectory \ArcGIS\Desktop10.0\Metadata\Translator\Transforms).
   3. Export your metadata using the esri102eml21.xml stylesheet:
      1. Using the export button on the description tab or the export wizard from ArcToolbox <conversion tools><metadata> , select the Export Metadata wizard. Browse to <WHERE YOU HAVE STORED ARC SOFTWARE ON YOU R COMPUTER>\ArcGIS\Desktop10.0\Metadata\Translator and select the esri102eml21.xml as the translator. (**Hint:** if you copy this link to your clipboard or a document, you can paste it into the dialog box and save time.)
      2. Indicate the directory and filename you want for your new EML file.
      3. Run the wizard.
   4. Note: if you want to convert many files, you can use the “Export Metadata Multiple” wizard. You select the source metadata files, the translator, and then provide an output folder destination. The output EML files are named by appending \_export.xml to the item's name and are stored in the specified folder.
2. You can run the stylesheet through a transformation on any xml editing program or from the DOS prompt (instructions: <http://im.lternet.edu/project/Esri2Eml/docs/use>). Get the most recent xalan and xerces code here: <http://mirrors.gigenet.com/apache/>. I grabbed this combined set here: <http://mirrors.gigenet.com/apache/xalan/xalan-c/binaries/xalan_comb-1.11-x86-windows-VC90.zip>. Put xerces dll in same directory as xalan.
3. You will create a new xml document after the transform. It is suggested that you save these files to a directory where they can be edited or harvested by Metacat.
4. **RASTER Data Problems**. ArcGIS 10.0 and 10.1 are not properly transposing the cellsize and number of bands for rasters. You will have to manually edit this in the EML file. From an xml editor, search for <cellSizeXDirection>. If the value is unknown, and you know the value, replace the value with it’s correct value. For example, for a cell size of 1 meter by 1 meter, this is how the tag would look: <cellSizeXDirection>1</cellSizeXDirection>. The same process should be completed for the YDirection, and if you know the number of bands in the image, replace the unknown with the correct value. (Check in the previous section above about properly documenting raster attributes.)

**Step 5. Run through the EcoInformatics Parser and correct errors.**

1. The Parser is located here: <http://knb.ecoinformatics.org/emlparser/>
2. You can select a file to parse, or copy and past your eml into the text block.
3. Contact Theresa Valentine ([theresa.valentine@oregonstate.edu](mailto:theresa.valentine@oregonstate.edu) or Inigo San Gil ([isangil@lternet.edu](mailto:isangil@lternet.edu)) for help with errors and/or to report problems with the stylesheet.

**Step 6. Prepare your data for the geospatial Data Package:**

**Geospatial Data Packaging;**

The idea of data packaging is to prepare spatial data sets that can be harvested for ingestion into PASTA and to the GeoNIS geospatial database. Best practices for the contents of a geospatial data package are as follows:

1. File to be harvested is a ,zip or .targz
2. File location is located under distribution information in the (spatialvector or spatialraster tags) in EML document for the database. The link should point directly to the data package, without download sign-ins required. This can be accomplished using the Data Access Server (DAS) or direct link to the data from your site.
3. Contains native GIS/spatial data (in native format, refer to attachment list from www.geomapp.net)
4. Vector: preferred format should be shapefile or KMZ
5. Raster: (refer to attachment list from www.geomapp.net)
6. Metadata File: (xml that travels with ArcGIS data, or FGDC xml)
7. Projection file: (usually included with ArcGIS products): should make sure it’s there.
8. Symbology layers would be another entity
9. Color ramps would be an associated file or entity.
10. ReadMe File with the following information: unique layer name (with no spaces) and not to exceed 10 characters,
11. This data package needs to be placed in the location specified in your EML document. The EML document should be validated by the PASTA portal-s and when valid, uploaded into PASTA. Eventually there will be spatial data quality checks performed prior to acceptance by PASTA. The GeoNIS working group are developing tools to perform these tests.