**LTER EML Microsoft Excel Metadata Template Documentation**

**(xlsx2EML-04\_Metadata\_Template\_FCE.xlsx)**

The metadata fields listed in the LTER EML Microsoft Excel metadata template are in compliance with the Ecological Metadata Language (EML) 2.2.0 Standards and Formats (<http://knb.ecoinformatics.org/software/eml/> ) and the [EML Best Practices, Version 2](http://intranet2.lternet.edu/sites/intranet2.lternet.edu/files/documents/Scientific%20Reports/Informatics/emlbestpractices-2.0-FINAL-20110802.pdf) document released in August 2011.

The LTER EML Microsoft Excel metadata template was created to facilitate the capture of EML metadata. The ultimate goal for Information Managers is to produce a valid EML document which could be submitted to Metacat, a flexible XML database. There are 3 steps necessary in producing a valid EML document when using the EML Microsoft Excel metadata template:

***Step 1:*** Capture EML (Tier 5, Except for the Constraint Element) metadata using the

EML Excel metadata template.

***Step 2:*** Load completed LTER EML Microsoft Excel metadata template file into the

Perl EML converter program.

***Step 3:*** By default, the Perl program will automatically validate the new EML file.

It is suggested that the user re-validate the converted EML (xml) file by loading

or pasting file into the KNB EML 2.2.0 validation service at <http://knb.ecoinformatics.org/emlparser/index.html> or into an XML editor like XMLSpy or Oxygen to validate document. *Note: if* *an XML editor is used for validation, the user MAY have to have a copy of the EML schema stored locally and the eml.xsd path at the top of the generated EML document will have to be changed to the local path, as well as, the stmml.xsd path at the bottom of the document in the additional metadata section.*

**Basic Instructions**

* Please fill in all blank, underlined rows as completely as possible unless directed otherwise by notes in the field. In some cases, field may not be applicable for datasets, such as 'Dataset Creator Organization Name'. Each metadata field has an accompanying comment box directing users as to what type of information should be entered.
* Separate multiple entries using vertical line delimiters (|). (Applies to the Street Address and Research Project Temporal Coverage fields).
* Please do NOT insert or delete rows from the template (long entries will wrap within cells, except as noted below) as the Perl converter program will not work. Note that extremely long entries may not be completely displayed even when the cell is enlarged, but the full contents will be retained and will export properly.
* EML tags used for metadata field are listed in Column A of same metadata field row.
* Optional Metadata fields are magenta.
* Blue Metadata fields will most likely be completed by the Information Manager.
* If a row background is light blue, please enter multiple first names, addresses, keywords, etc. in new columns. Please note: Creator, Contact, Geographic description, Data Entity Taxon, and Research Project personnel sections must include all information for one person, taxon, or location in the SAME column.
* Move the mouse pointer over a field name to display instructions and comments.

Click +/- icon in column 1 to expand or hide EML category (i.e. Dataset Creator).

* Once the EML metadata template is complete, save the file and then load the completed metadata file into the Perl EML converter program. The metadata will be converted into a valid EML xml document. Please read the Perl documentation for instructions on how to launch and use the program.

***Description of Template Worksheets***

**I. General Metadata Worksheet**

**A. LTER Dataset Information**

General dataset metadata including Dataset Creator, Dataset Coverage,

Dataset Distribution, Dataset Associated Party, Dataset Contact, Dataset

Publisher, Dataset Metadata Provider, Dataset Access, Dataset Methods,

Dataset Quality Control and Dataset maintenance.

**B. LTER Dataset Table Information**

General Entity and Attribute metadata such as Entity Name, Entity Description

and Number of Data Records. The detailed Attribute information is entered

separately in the ‘DataTable’ worksheet.

**C. LTER Research Project Information**

A link is followed to a general research project information worksheet labeled

‘ResearchProjects’.

**D. Additional Metadata - Dataset and Project**

Additional metadata such as Dataset Research Type, Dataset Submission Date,

and Dataset Purpose can be entered in this section. These metadata fields are

NOT part of the required EML nor were they listed in the LTER EML Best Practices document. Users may modify this section as needed but MUST be

sure to make the appropriate changes to BOTH column A and column B.

**II. MethodsCitation Worksheet**

Enter methods citation information beginning with Dataset Methods Citation Number 5.

The title, author’s name, and publication date are common to ALL types of citations and

are found in the first section of the citation block. On Row 11, the user is prompted to

‘SELECT A CITATION TYPE’ and once the choice is made, the appreciate citation fields

will appear in the following rows. Remember that the citation information is entered as a

block so if there are multiple citations, the next citation entry would occur under Dataset

Methods Citation Number 28. Do NOT change the citation numbers (citation number = row number) for any reason as the Perl program will offset the entered information. If

additional citation blocks are needed, copy a complete section and paste at the bottom

of the worksheet. **The citation IDs must be entered in ROW 117 of the General Metadata Worksheet.**

**III. MethodsProtocol Worksheet**

Enter methods protocol information beginning with Dataset Methods Protocol Number 5.

Remember that the protocol information is entered as a block so if there are multiple protocols, the next protocol entry would occur under Dataset Methods Protocol Number 28. Do NOT change the protocol numbers (protocol number = row number) for any reason as the Perl program will offset the entered information. If additional protocol blocks are needed, copy a complete section and paste at the bottom of the worksheet. **The protocol IDs must be entered in ROW 118 of the General Metadata Worksheet.**

**IV. ResearchProjects Worksheet**

Enter research project information beginning with Research Project Number 5.

Remember that the project information is entered as a block so if there are multiple projects, the next project entry would occur under Research Project Number 28. Do NOT change the project numbers (project number = row number) for any reason as the Perl program will offset the entered information. If additional project blocks are needed, copy a complete section and paste at the bottom of the worksheet. **The project IDs must be entered in ROW 160 of the General Metadata Worksheet.**

**V. DataTable Worksheet**

Detailed data attribute information and data values are entered in the DataTable worksheet.

* Do NOT fill out or delete the custom unit section (red text, yellow cells). These cells will be filled out automatically (from the 'Units' worksheet) when you select a value from the 'Units' cell dropdown list.
* Descriptions of each header field are listed in the cell comments (hover the mouse pointer over the field name to view)
* Make sure the column formatting is appropriate for your data - values will be export 'as is' so undisplayed digits will be lost.
* Fill out the header section (cyan cells) for each data column as completely as possible. Fields in bold are required.

**VI. Annotations Worksheet**

* Annotations for attributes can be included in this worksheet. The attribute name, attribute label, and attribute definition rows (rows 2-4) must be the same as rows 11-13 in the Data Table Worksheet. The program currently allows one annotation per attribute, although more than one annotation per attribute is possible in EML 2.2.0.
* Multiple dataset annotations can also be included in this worksheet. Enter each new dataset annotation into the subsequent column (ie. the first dataset annotation is entered in rows 17-20 of column B, the second dataset annotation is entered in rows17-20 of column C).

**VII. Awards Worksheet**

* Enter awards for each research project in this worksheet. Please skip two rows between research projects, since the program is expecting each award to have six rows followed by two blank rows.

**VIII. References Worksheet**

The Long-Term Ecological Research Network (LTER) metadata is based on a metadata specification call Ecological Metadata Language (EML) Version 2.1.0 developed by the ecology discipline and for the ecology discipline, based on prior work done by the Ecological Society of America and associated efforts. EML is implemented as an XML schema that can be used to document ecological data. The following URL will provide in depth EML information including the EML 2.1.0 Specification, EML project information, and EML development information: http://knb.ecoinformatics.org/software/eml/.

\*Original Excel Metadata Template created by Wade Sheldon, Information Manager, Georgia Coastal Ecosystems LTER and modified by Linda Powell (Information Manager) and Mike Rugge (Program Manager) Florida Coastal Everglades LTER (September, 2004). Contributors to this Excel metadata template include Linda Powell (FCE), Mike Rugge (FCE), Wade Sheldon (GCE), and Kristin Vanderbilt (FCE).

The LTER Core Research Areas are Primary Production, Disturbance, Organic Matter, Inorganic Nutrients and Populations.

**VII. IM Use Only Worksheet**

This worksheet contains 3 types of lists that are used in the DataTable and MethodsCitation worksheets. **DO NOT DELETE any items within the list nor the entire IM Use Only worksheet.**

**VIII. Units IM Use Only Worksheet**

This worksheet contains a listing of EML and Customs units used to populate Row 24 (Units) of the DataTable worksheet. You may add a custom unit to this worksheet and the addition will appear as a choice in the drop-down list on Row 24. Note: Make sure all the pertinent unit information is complete when adding a custom unit to the list. **DO NOT DELETE any items within the list nor delete the entire Units IM Use Only worksheet.**

**Acknowledgements and References**

The Long-Term Ecological Research Network (LTER) metadata is based on a metadata specification called Ecological Metadata Language (EML) Version 2.1.0 developed by the ecology discipline and for the ecology discipline, based on prior work done by the Ecological Society of America and associated efforts. EML is implemented as an XML schema that can be used to document ecological data. The following URL will provide in depth EML information including the EML 2.1.0 Specification, EML project information, and EML development information: http://knb.ecoinformatics.org/software/eml/

The original Excel Metadata Template created by Wade Sheldon, Information Manager, Georgia Coastal Ecosystems LTER and modified by Linda Powell (Information Manager) and Mike Rugge (Program Manager) Florida Coastal Everglades LTER (October, 2004).

***Contributors to this Excel metadata template include:***

* Linda Powell, Kristin Vanderbilt, and Mike Rugge from Florida Coastal Everglades LTER Program (<http://fcelter.fiu.edu>) at the Florida International University.
* Wade Sheldon from Georgia Coastal Ecosystems LTER Program (<http://gce-lter.marsci.uga.edu/lter/>) at the University of Georgia.
* Kristin Vanderbilt from the Sevilleta Long-Term Ecological Research LTER Program ([http://**sevilleta**.unm.edu](http://sevilleta.unm.edu)) at the University of New Mexico.
* Youngmi Kim and Travis Brooks, programmers for the Canopy Database Project and graduates of The Evergreen State College Software Engineering Program (<http://canopy.evergreen.edu/>).
* Judy Bayard Cushing, Ph.D., a member of the Faculty (Computer Science), The Evergreen State College, Olympia, Washington and a principal investigator of the Canopy Database Project (<http://academic.evergreen.edu/j/judyc/home.htm>, <http://canopy.evergreen.edu/>).

Working in cooperation with the Evergreen State College contributors are Professor Barbara Bond, Department of Forest Science, Oregon State University and her students, Georgianne Moore, Texas A&M University and Kate George, USDA.

***References***

Cushing, J.B., N.M. Nadkarni, M. Finch, A.C.S. Fiala, E. Murphy-Hill, L. Delcambre, and D. Maier. In press. Component-based end-user database design for ecologists. Journal of Intelligent Systems.

Cushing, J.B., N.M. Nadkarni, B. Bond, and R. Dial. 2003. How trees and forests inform biodiversity and ecosystem informatics. Computing in Science & Engineering 5:32-43.