

Tips for submitting FCE LTER data and metadata for publication

Updated 2024-02-09

Consult the FCE Information Manager

Your first step when you plan to submit data is to contact the FCE Information Manager, Gabe Kamener (gkamener@fiu.edu), with details about the data you intend to archive. He will need to know the number of data files you will submit and how they relate to each other. He can then advise you on how to organize the data files for archiving. Let him know at this time if you will be submitting code, photographs or other non-tabular data to include with the archived data package (a data package = metadata + data and other resources archived together).

Data should be organized into logical units based on data structure and theme for archiving. If you are submitting data for multiple years, for instance, all data should be compiled into a single file. If you are submitting several data tables that were all collected as part of a single experiment, they should probably be archived together in a single dataset. The IM will advise you on the best organization for your dataset.

Format Your Data for Publication

Before submitting your data to the FCE Information Manager, please ensure that the data are formatted according to the following rules:

- a. Data may be submitted to the IM for initial review as Microsoft Excel files or comma-separated files (.csv) but will need to be formatted as csv files for publication.
- b. Column header names should not include spaces or special characters such as '/' (forward slash). Underscores are permitted.
- c. Format dates as YYYY-MM-DD. Spreadsheets are notorious for "coercing" date and time formats into proprietary numerical formats when data is entered. It is your responsibility to ensure that the spreadsheet you submit shows the correct dates.
- d. Do not include commas in cell contents. For instance, do not enter in a Comments field "missing tag, could be 602." Instead, enter "missing tag; could be 602." The data will be archived as a comma-separated file, and commas within the data will cause some records to have too many fields.
- e. Do not submit data with empty cells. Instead, enter a code for the missing observations, such as -9999 for numerical variables or NA for text variables. Define the missing value code in the metadata.

Create Metadata

- a. Metadata describe the who, what, where, when and why of a dataset, and provide the details that a secondary user will need to understand and reuse your dataset. The information you submit will be encoded as Ecological Metadata Language (EML), the metadata standard used by the LTER Network. EML makes the metadata machine readable, enabling discovery and automated processing.

- b. If you are submitting a new dataset, please use the [ezEML editor](https://fcelter.fiu.edu/data/_assets/ezeml-instructions_fce.pdf) from the Environmental Data Initiative (EDI). This is a form-based online metadata editor. You can login to this tool using your ORCID, Github or Google credentials. When you have finished editing your metadata, please share it with the FCE IM (gkamener@fiu.edu) through the ezEML's "Collaborate" for final review before your package is published in the FCE LTER data catalog. Linked below are some instructions to help you get started using ezEML.
- https://fcelter.fiu.edu/data/_assets/ezeml-instructions_fce.pdf

Selected tips on describing metadata in ezEML (sourced from FCE LTER and EDI guidance)

Title

The title should be descriptive, including what, where, and when. **Do not** use a manuscript title; the dataset should have its own unique title.

Example: "Monthly Water Quality Data from Shark River Slough, Florida Everglades, Florida, USA (FCE LTER): 2010-2019"

Data Table

Keeping a data dictionary with your data file to describe your columns will make it easier to enter Data Table metadata into ezEML.

As ezEML uploads the data file (which must be a csv), it infers whether each variable is of type "Numerical", "Text", "DateTime" or "Categorical". In this window you can edit the "Type" of variable if ezEML incorrectly inferred it. Then you must "Edit Properties" to finish describing each variable. For each variable, you will need to enter a Definition, choose a Standard Unit from the dropdown list, or create a Custom Unit. Specify custom units using camelCase, not symbols. Specify missing value codes. Filling out other fields is optional, although precision is desirable.

Abstract

Your abstract should describe the dataset. It will not be the same as an abstract provided for a paper derived from the dataset.

The abstract will be used for full-text searches, and it should be rich with descriptive text.

In particular, descriptions should include information that does not fit into structured metadata, and focus on the "what", "when", and "where" information, general taxonomic information, as well as whether the data package is ongoing or completed.

Some general methods description is appropriate, and broad classes of measured parameters should also be included. For a large number of parameters, use categories instead of listing all

parameters (e.g. use the term “nutrients” instead of nitrate, phosphate, calcium, etc.), in combination with the parameters that seem most relevant for search.

Example: “Asian swamp eels (*Monopterus albus/javanensis* complex) were first reported from Florida in 1997 and the Everglades in 2007; swamp eels have been established in Taylor Slough of Everglades National Park since 2014. This dataset incorporates plot-level mean densities (# of individuals per square meter) of common aquatic animals collected during 1996–2022 from 24 sites across four regions of the Everglades: Taylor Slough, Shark River Slough, Water Conservation Area 3, and the C-111 Panhandle. Prey species included are the six most common small fishes prior to swamp eel invasion of Taylor Slough (1996–2009) and the three common decapod species (two crayfish species and grass shrimp). An annual index of mean wet season electrofishing catch-per-unit-effort of swamp eels, Mayan cichlids (*Mayaheros uruphthalmus*), and the three other large 'top predator' fishes (*Amia calva*, *Lepisosteus platyrhincus*, *Micropterus salmoides*) is included for plots where electrofishing was performed from 1997-2021. Hydrologic measures used in analyses are included.”

Methods

The description of the methods should be detailed enough to answer any questions someone might have about how the data were collected. Include a list of cited references (if applicable) at the bottom of your methods description.

Keywords

Select keywords from [LTER Controlled Vocabulary](#) when possible. You can easily do so in the dropdown menu on the Keywords page in ezEML.

Taxonomic coverage

Make sure to check spelling of names is accurate.

Notes about other sections when using ezEML

- Creators: These are the individuals whose names you want to appear in the citation to the dataset.
- Contacts: You must enter at least one.
- Associated parties are field crew members, lab technicians, etc.
- Intellectual Rights: Please select CC-BY
- Geographic Coverage: Coordinates of sampling sites or regions should be in decimal degrees
- Project: Should include information about any projects (including award information) that supported data collection.