

Florida Coastal Everglades LTER Program Data Management Policy

September, 2014

Objectives

As a member of a Long Term Ecological Research Network, all researchers participating in the Florida Coastal Everglades Long Term Ecological Research Program (FCE LTER) recognize their responsibility to contribute all NSF FCE LTER funded data to the FCE LTER database and to publish data via the FCE website in a timely manner. The FCE LTER Data Management Group has established a set of protocols and services for data collection, quality assurance, organization, data archive, data access, and data distribution to facilitate our scientific work and provide research data to those participating in Everglades restoration work. Given the potentially important role that the FCE LTER program may play in Everglades restoration, another important objective of the FCE Information Management Group is to increase public and private awareness of our Everglades research activities.

FCE Information Management Advisory Committee

A five member Information Management Advisory Committee (IMAC), consisting of a voting member of the FCE Internal Executive Committee (IEC), a project collaborator, a student, a technician, and an education and outreach representative, assists the FCE Data Management Group with informatics, primarily the FCE website. Members of this standing committee are volunteers and hold a 1-year term beginning in January of a new calendar year. The committee will perform an annual internal review of the FCE website using a questionnaire based on the LTER Network Web Site Design and Content document, a reference for the LTER Network IM Review Criteria. The main focus of the IMAC is to continually assess the scientific content of the FCE website and to ensure that the posted information is accurate and comprehensive. Other duties include testing the website for functionality and making recommendations for a more user-friendly interface, creating and contributing to web content, and suggesting innovative changes to the website such as adding specific web-based tools that will facilitate site science. As the committee is responsible for web scientific content, it will take a majority vote from its members to make changes to that content on the FCE website. Any changes suggested by IMAC involving the FCE Oracle database and FCE website designs and structures will be vetted by the FCE Information Manager and Project Manager.

Research Projects

Project management at the FCE LTER is very complex because of the large number of research personnel and immense study area. The Florida Coastal Everglades LTER sites are located in freshwater marsh, estuarine mangroves, and sea grass estuary ecosystems in Everglades National Park, an area that covers approximately of 4300 km² or 1,062,553 acres. With 23 'official' FCE LTER and 260 'related' sampling sites, the task of keeping track of the different types of research, sampling parameters, sampling frequencies, and intra-site sampling point locations is difficult. The FCE Information Management System (IMS) is committed to collection and organization of FCE LTER project information and the IMS team has developed a web-based interactive mapping application called the 'FCE LTER Interactive Everglades Map' to their website at http://fcelter.fiu.edu/data/GIS/interactive_map/ to facilitate information and project management. There are three criteria are used to define a FCE LTER Project: 1) The funded grant included a FCE letter of support at the time of proposal submission; 2) The work supported by funded grant occurs at one or more FCE LTER sites and these sites were explicitly described in the proposal as FCE LTER sites, or 3) The funded grant cited the FCE LTER program in the proposal. Researchers are asked to submit FCE LTER project information, via a project information form (http://fcelter.fiu.edu/research/information_management/documents/) to the FCE Information Manager upon notification of project funding, but within 6 months of notification of project funding. Additionally, researchers will also be asked to update their project information annually.

Data Collection

To ensure that FCE LTER data are properly archived and protected, researchers are strongly encouraged to routinely submit data from ongoing research to the FCE LTER Information Manager. To this end, FCE researchers must submit all data (digital format) collected during a calendar year to the FCE LTER Information Manager by no later than June 1 of the following year. The Information Manager will schedule quarterly data collection dates (Jan 1, April 1, July 1, & October 1) throughout the year whereby electronic collection reminders, with the appropriate data submittal information, will be sent out to all participating researchers. The purpose of these reminders is to encourage researchers to continually submit their data for archiving and protection, rather than waiting until June 1 of each year to submit the previous year's data. All data must be submitted with a standard FCE LTER digital EML 2.1 metadata form describing the format and content of the data together with a list of persons that have permission to make data available to others during its protected period (see Data Distribution Section). A full discussion of the FCE data submission and data distribution policies can be found at http://fcelter.fiu.edu/research/information_management/documents/.

Quality Assurance

A goal of the FCE LTER Data Management Group is to create and maintain an archive of FCE LTER data files that are error free and fully documented. Good communication and interaction between the Information Manager and scientists are both critical to preventing data loss and to maintaining data integrity. Any FCE LTER researcher who submits data to the FCE LTER database will be responsible for the quality assurance of that data, including the metadata content. Once those data have been submitted to the FCE Information Manager, the LTER program and Information Manager will be responsible for supervision and stewardship of datasets.

Data Organization

File Naming and Site Name Protocol

FCE LTER data files and their accompanying metadata files will carry the same name with the exception of 'eml' added to the end of the metadata file name. Data and metadata files will be submitted in either Microsoft Excel (.xls) or tab delimited text files (.txt) formats.

Datasets and metadata submitted by FCE LTER researchers will be named by using 1) the type of dataset, 2) researcher's major research working group acronym, 3) researcher's last name, and 4) a dataset number (based on the number of datasets previously submitted by researcher to the Information Manager).

For example, the first short-term experiment dataset and dataset metadata submitted by John Doe, a member of the Primary Production Workgroup, would be assigned a dataset name of ST_PP_Doe_001 and ST_PP_Doe_001_eml, respectively. The first long-term experiment dataset and dataset metadata submitted to the LTER Information Manager by John Doe2, also a member of the Primary Production Workgroup, would be assigned a dataset name of LT_PP_Doe2_001 and LT_PP_Doe2_001_eml, respectively.

For those researchers who will be updating data files already submitted to the Information Manager (e.g. monitoring data), the updated data file should retain the old data file name and contain both the previously submitted data plus the newly appended data. The dataset will retain the original file name. The accompanying metadata files will be updated to reflect the appended data and will retain its original file name.

Type of Dataset	Code
Short-Term*	ST
Long-Term**	LT
Physical Data***	PHY
Climate Data ****	Climate
GIS Data	GIS

*Data collected for a period <= to 3 years.

**Data collected for a period > than 3 years.

***Data collected in the abiotic environment that influences the growth and development of organisms of biological communities such as water levels, air temperatures, rainfall amounts, and evapotranspiration.

**** Data collected from NOAA weather stations.

***** Data generated in a geographic information system environment.

FCE Research Phase	FCE Research Workgroup	Acronym
I	Nutrients & DOM	ND
I, II, III	Primary Production	PP
I	Soils & Sediment	SS
I	Consumer Dynamics	CD
I, II, III	Trophic Dynamics & Community Structure	TDCS
I	Education & Outreach	EO
I	Ecological & Social Modeling	ES
II	Modeling & Synthesis	MS
II, III	Biogeochemical Cycling	BC
II, III	Organic Matter Dynamics	OMD
II	Climate & Disturbance	CDI
II	Hydrology	HY
II	Human Dimensions	HD
III	Hydrology & Water Policies	HYW
III	Carbon Dynamics	CCD
III	Modeling & Scenarios	MSN
III	Climate & Disturbance Legacies	L

FCE Metadata Protocol

The FCE LTER Program adheres to the LTER network-wide Ecological Metadata Language (EML 2.1.0) standard developed at the National Center for Ecological Analysis and Synthesis (NCEAS). This new standard includes distribution of both geospatial and non-geospatial metadata in parsable XML formats based on EML. EML exists as a set of XML Schema documents that allow for the structural expression of metadata necessary to document a typical data set in the ecological sciences. A full description of EML 2.1.0 can be found at the following URL: <http://knb.ecoinformatics.org/software/eml/>.

Currently, the FCE LTER Information Management Group uses a Metadata template that follows the EML content standard and is based on the LTER EML Best Practices Version 2

(August, 2011). A downloadable version of the FCE Metadata template, available in Microsoft Excel (.xls), can be found under the *Information Management Policies and Documents* section on the FCE public website (http://fcelter.fiu.edu/research/information_management/documents/).

Data Archive and Protection

The FCE LTER Data Management Group is dedicated to developing and maintaining a high-quality program for long-term data archival storage. All FCE LTER data submitted to the Information Manager will be stored in the FCE data archives. FCE project information, minimal research data metadata will be stored in the FCE LTER Oracle Database.

The entire network of FCE servers and workstations undergo continual updates and patches to their operating systems. There is a secure socket layer (SSL) on all servers and all FCE computers have dynamic firewalls. The FCE Information Management System implements 3 levels of data protection:

- 1) Level 0 – Nightly incremental backups of the Oracle Database server, the FCE Web server and IMS personnel workstations
- 2) Level 1 – Two sets of weekly full backups of all FCE LTER physical and virtual servers and workstations are made to external hard drives. One set of hard drive backups is stored in an off-site location in a fully padded, waterproof, rolling case and the second set is stored in the FCE LTER Office firebox safe. There are a total of six external hard drives used in the weekly backup rotation, giving the FCE data management group 3 weeks of system backups at any given time.
- 3) Nightly Syncs and/or manual syncs when necessary made between the 3 production virtual servers housed at FIU and identical virtual servers residing at the Northwest Regional Data Center (NWRDC) located roughly 400 miles north of Miami on the campus of Florida State University in Tallahassee, Florida. This 'off-site' disaster recovery plan will allow the FCE website to be continually available throughout disaster events such as hardware failures and hurricanes.

In addition to all FCE LTER Program digital data, any pertinent hardcopy data, such as completed field notebooks, may also be stored in the FCE LTER office fireproof cabinet. Researchers are encouraged to take advantage of the LTER office for these kinds of non-digital backup protection.

Data Access

Part A- Data

Data and information derived from publicly funded research at the Florida Coastal Everglades (FCE) LTER sites under the National Science Foundation cooperative agreements #DEB-1237517, #DBI-0620409, and #DEB-9910514 are made available online with as few restrictions as possible on a nondiscriminatory basis. FCE researchers make every effort to release data in a timely fashion.

The FCE Information Management protocol is that researchers must provide an "unlock" date for all datasets (meaning a date after which these data can be unlocked and made available to the public). This "unlock date" must be no more than 2 years from the submission date if data were collected using FCE LTER NSF funding. The FCE LTER researcher associated with a given dataset controls access to their data within this 2 year protected period. No locked datasets will be released to anyone (including FCE-affiliated researchers) without the expressed permission of the researcher, transmitted in writing to the FCE Information Manager (email is acceptable), whose data have been requested. If a longer protection period is necessary for a particular dataset, the FCE researcher may petition the FCE LTER Internal Executive Committee (FCE LTER IEC) for a protection period extension. Without this special permission, no dataset may remain locked for more than the time periods described below after submission to the FCE database.

The FCE LTER Data Management Group will make FCE LTER data and metadata available for on line access via the web according to the data availability status as defined by their data type described below:

Type I – data are to be released to the general public according to the terms of the general data use agreement (see *Data Use Agreement Section below*) within 2 years from collection and no later than the publication of the main findings from the dataset.

Type II - data are to be released to restricted audiences according to terms specified by the owners of the data. Type II data are considered to be exceptional and should be rare in occurrence. The justification for exceptions must be well documented and approved by the lead PI and Site Data Manager. Some examples of Type II data restrictions may include: locations of rare or endangered species, data that are covered under prior licensing or copyright (e.g., SPOT satellite data), or covered by the Human Subjects Act, Student Dissertation data and those data related to the FCE LTER Program but not funded by the National Science Foundation (NSF) cooperative agreements #DEB-1237517, #DBI-0620409, and #DEB-9910514. Researchers that make use of Type II Data may be subject to additional restrictions to protect any applicable commercial or confidentiality interests.

While the spirit of this document is to promote maximum availability for ecological data in either Type I or II status, there are criteria by which priority for data release may be determined. Primary observations collected for core research activities directly supported by LTER research must receive the highest priority for data release. Data collected by other sources to which LTER supported research has added value is also a high priority. Other types of data including non-LTER data that was acquired for LTER research, student thesis data, schoolyard LTER data, or legacy data that already suffer from inadequate documentation or format obsolescence may be ranked a lower priority by the FCE LTER. Legacy data will be released as resources become available.

Conditions Justifying Type 2 Statuses

1. Legal Issues

- a. The data includes the location of sensitive resources that might be endangered by revealing the location. This would include locations of artifacts, threatened or endangered species.
- b. Data will not be released that may compromise one's personal privacy. This might include types of survey and census data involving human subjects.
- c. Data that are covered by copyright laws (e.g. TM and/or SPOT satellite data).

2. Publication Issues

- a. The protection of a vulnerable PI is required. These data are those collected by graduate students, Post-Docs and others who have limited number of data to their name and whose professional development might be compromised by releasing the data too early.

3. Non-LTER Funded Data

- a. This includes data related to the FCE LTER program but not funded by the NSF LTER grants #DEB-1237517, #DBI-0620409, and #DEB-9910514. These data are funded by other agencies and/or institutions but the data are tied to the LTER objectives or were collected at FCE sites. In these cases, the researchers are strongly encouraged to either a) submit links to their databases, allowing direct connection to their web accessible data via the FCE LTER website and through the FCE Database, or b) to also submit their data to the FCE Information Manager at the time the data are released to the agency. In either case, the researcher should seek

permission of the agency in question to make this link. The recommended lock period is identical to that mandated by the funding agency.

4. Data Quality Assurance and Control Issues

- a. Data with quality assurance and control issues such as data with low quality or data that include measurements using new techniques that require further study before their value and limitations are understood.

Part B- Metadata

1. Metadata documenting archived/online data sets of all types listed above will be made available when, or before, the dataset itself is released according to the terms above.
2. All metadata will be publicly available regardless of any restrictions on access to the data.
3. All metadata will follow LTER recommended standards and will minimally contain adequate information on proper citation, access, contact information, and discovery. Complete information including methods, structure, semantics, and quality control/assurance is expected for most datasets and is strongly encouraged.

Data Use Agreement

Type I data sets available via the FCE LTER Program web site (<http://fcelter.fiu.edu/>) are freely available and can be downloaded for academic, research, or professional purposes subject to the following user terms:

1. **Acceptable use:** User must notify designated FCE researcher when any future work based on or derived from FCE data is published and they must agree to send 2 reprints of any publications resulting the use of the data and documentation to the following address:

Florida Coastal Everglades LTER Program
C/O Mike Rugge, Program Manager
Florida International University
11200 SW 8th Street
MMC, SERC, OE 148
Miami, FL 33199

Publications may also be E-mailed to fcelter@fiu.edu.

2. **Redistribution:** User agrees not to redistribute original FCE LTER data and Documentation.
3. **Citation:** User will acknowledge the support of the FCE LTER Program and appropriate NSF Grant numbers in any publications using these data with the following citation:

'(Dataset Originator's name(s), dataset publication date: Title of dataset. Dataset Digital Object Identifier (DOI). Data sets were provided by the Florida Coastal Everglades Long-Term Ecological Research (LTER) Program. This material is based upon work supported by the National Science Foundation under Grant # (please choose one of the following options in terms of Grant numbers and date of materials used, to insert in above citation):

Option 1) Work from December 2012 to Present (NSF Grant No. DEB-1237517): "This material is based upon work supported by the National Science Foundation under Grant No. DEB-1237517.

Option 2) Work from 2007-2012 (NSF Grant No. DBI-0620409):

"This material is based upon work supported by the National Science Foundation under Grant No. DBI-0620409."

Option 3) Work through 2006 (NSF Grant No. DEB-9910514):

"This material is based upon work supported by the National Science Foundation under Grant No. DEB-9910514."

OR if work spans multiple grants:

Option 4) "This material is based upon work supported by the National Science Foundation under Grant No. DBI-0620409 and Grant No. DEB-9910514."

Option 5) "This material is based upon work supported by the National Science Foundation under Grant No. DEB-1237517, DBI-0620409, and Grant No. DEB-9910514."

Citation Example:

'Childers, Daniel; Troxler, Tiffany (2008-03-14): Water Quality Data (Grab Samples) from the Taylor Slough, just outside Everglades National Park (FCE), for August 1998 to November 2006. Long Term Ecological Research Network. <http://dx.doi.org/10.6073/pasta/ab34172874c2517430e2d98940a3b98e>. Data sets were provided by the Florida Coastal Everglades Long-Term Ecological Research (LTER) Program. This material is based upon work supported by the National Science Foundation under Grant No. DEB-1237517, DBI-0620409, and Grant No. DEB-9910514.'

By using or copying these data and documentation, the Data User agrees to abide by the terms of this agreement.