

Florida Coastal Everglades LTER Program Data Management Policy **March, 2006**

Updated November 2007

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 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. 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, seagrass 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 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).

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 the word 'meta' 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 theme acronym abbreviation (specific to dataset; with the exception of those data types which fall in the 'Physical Data' category), 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_meta, 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_meta, respectively. If John Doe2 submitted his second long-term experiment dataset and dataset metadata related to a different research workgroup, Soils and Sediments, the dataset would be LT_SS_Doe2_001 and the metadata would be named LT_SS_Doe2_001_meta. In the case of FCE LTER datasets which fall under the category of 'Physical Data', those dataset names would only include the PHY abbreviation, the researcher's name, and the dataset number. For example, the first water level dataset and dataset metadata submitted by John Doe would be labeled PHY_Doe_001 and PHY_Doe_001_meta, 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 with a version extension added. For example, if the old data file is named LT_PP_Doe_001, the first update of the file will be named LT_PP_Doe_001.v1 and the second update of the file will be named LT_PP_Doe_001.v2. The accompanying metadata files will be named LT_PP_Doe_001.v1_meta and LT_PP_Doe_001.v2_meta.

| <i>Type of Dataset</i> | <i>Abbreviations</i> |
|-------------------------|----------------------|
| Short-Term* | ST |
| Long-Term** | LT |
| Physical Data*** | PHY |

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

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

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

| <i>FCE Research Phase</i> | <i>FCE Research Workgroup</i> | <i>Acronym</i> |
|---------------------------|---|----------------|
| FCE I & FCE II | Primary Production | PP |
| FCE I & FCE II | Trophic Dynamics & Community Structure (Formerly Consumer Dynamics (CD)) | TDCS |
| FCE II | Biogeochemical Cycling | BC |
| FCE II | Organic Matter Dynamics | OMD |
| FCE II | Hydrology | HY |
| FCE II | Modelling & Synthesis | MS |
| FCE II | Human Dimensions | HD |
| FCE II | Climate & Disturbance | CD |
| FCE I | Nutrient & DOM | ND |
| FCE I | Soil & Sediment | SS |
| FCE I | Ecological & Social Modeling | ES |
| FCE I | Education & Outreach | EO |

FCE Metadata Protocol

The FCE LTER Program adheres to the LTER network-wide Ecological Metadata Language (EML 2.0.1) 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 2.0.1 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 Tier 5 (EXCEPT for the "Constraint" Element (September, 2004)). 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 added to the FCE LTER Oracle Database and stored in ASCII text format.

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 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.

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 Distribution

The LTER Network recognizes that researchers on the FCE LTER have first priority for use of their data in publications. 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. 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 "unlocked" FCE LTER data and metadata available for on line access via the web according to the data availability status as defined by the following data 'Types':

- Type 1:** Published data and metadata are available upon request without restrictions other than those set out in the 'Data Users Agreement'. These data cannot be "locked".
- Type 2:** Collective data of the FCE LTER site, such as climatological and hydrological measurements, are available for specific scientific purposes within 1 year after collection. Maximum lock period = 1 year.
- Type 3:** Original FCE LTER experimental data collected by individual FCE researchers are available for scientific purposes 2 years after collection. Maximum lock period = 2 years.
- Type 4:** Data related to the FCE LTER Program but not funded by the National Science Foundation (NSF) LTER Grants #DEB-9910514 and DBI-0620409 are available for scientific purposes no more than 4 years after collection. Recommended maximum lock period = 2 years. In some cases, these data are funded by other agencies 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 they are released to the agency. In either case, the researcher should seek permission of the agency in question to make this link. In these cases, the recommended lock period is identical to that mandated by the funding agency.
- Type 5:** In rare cases, data may require permanent restriction, and thus always be "locked" except with the permission of individual researcher. Data may receive this Type 5 status due to following conditions:
- A) Data that include sensitive resources that might be endangered by the release of information. This may include but not be limited to the locations of endangered species and locations of permanent plots in sensitive areas.
 - B) Data that may compromise a person's personal privacy such as census data involving human subjects.
 - C) Data that are covered by copyright laws such as TM and/or SPOT Imagery data.
 - D) 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.

Data User Agreement

Unlocked data sets available via the FCE LTER Program Web site are freely available and can be downloaded for academic, research, or professional purposes subject to the following user terms:

- A) User must notify designated FCE researcher when any future work based on or derived from FCE data is published.
- B) User agrees not to redistribute original FCE LTER data and Documentation.
- C) 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:

'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 2007-2012 (NSF Grant No. DBI-0620409):

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

OR if work spans both grants:

Option 2) "This material is based upon work supported by the National Science Foundation under Grant No. DBI-0620409 and Grant No. DEB-9910514."

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."

- D) User agrees 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
Southeast Environmental Research Center
OE 148, University Park
Florida International University
Miami, Florida 33199

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