

## LTERCCE Data Management Practices

For an LTER CCE dataset to be submitted as 'core data' and reported in the LTER CCE data repository requires coordination in terms of data file structure and context captured through metadata. Metadata includes associated CCE dictionaries and directories.

### 1. Data Files

Data file features include structure, indexing information and associated metadata.

-Structure: each data file should be a matrix of comma or tab-delimited data values containing rows (records) and columns (measurements) of regular data with annotation appearing only in header rows that may contain the designation of column names

- Index: some type study/cruise event number may be associated with each record or file if the dataset is to be related to other datasets in a collection.

-Metadata: a data file or collection of similar files requires an associated metadata file containing dataset metadata including the dataset name, associated projects/study designation, abstract, methods, analysis, and contact information.

The following topics need to be considered:

- Project/Study Names: How is the dataset to be related to other CCE data?
- Dataset Type: What is the granularity or scope in length and frequency, ie single file per cruise, multiple casts per cruise
- Dataset Name: A cruise unique identifier for the dataset, ie chl, nutrients, etc
- Event#, Bottle#: What relation is the dataset to cruise event# activity (ie CTD, NetMoc)
- Time: How is time represented in your data (format, zone)
- Location: What type of location information is recorded with the data
- Calculations: any calculations used, can be internal or post-processed
- Standardized 'no data' (recommend -999 but need to identify it, ie NaN, -99, blank)

Example

Project Name	CCE LTER												
Study Name	0605KN												
Dataset Name	chl												
Assoc Activity	CTD												
			intended	intended	yyyy	mm	dd	hh (GMT)	mm(GMT)	m	mg/m3	mg/m3	
index	event	bottle#	station#	line#	year	month	day	hour	min	depth	chl	phae	
	1	7											
	2	15											
	3	23											

\*The index is at investigator's discretion.

## **2. Attribute, Unit, and Code Dictionaries**

These will be available online for reference and input.

-Attribute: to be selected or entered online. It is defined according to a CCE site attribute dictionary template. The definition includes attribute definition, calculation, information for derived values as well as storage and precision information.

-Unit: to be selected or entered online. It is defined according to the CCE site unit dictionary template

-Code: to be created by the investigator and submitted with the dataset. It is defined within a reference (ie CalCOFI station locations) or within a code dictionary file with a unique dictionary name, a list of files to which the code is pertinent, and with at minimum two columns (the code and its definition)

## **3. Personnel Directory**

Individuals associated with any part of the data process should be provided in a list with a name, email, contact information, and project affiliation. From this a CCE LTER personnel directory entry will be made and further information can be input via a web form.

### Reference

R.B.Cook, R.J.Olson, P.Kanciruk, and L.A.Hook. 2001. Best Practices for Preparing Ecological Data Sets to Share and Archive. Bulletin of the Ecological Society of America 82(2): 138-141.