

Adding ERDDAP Access to the BCO-DMO Data Management System



Pingyu Qiao, Robert Groman, Cyndy Chandler, Adam Shepherd, Dicky Allison, Danie Kinkade, Shannon Rauch, Peter Wiebe, David Glover



Summer Student Research Forum – August 5, 2016

Abstract: The Biological and Chemical Oceanography Data Management Office (BCO-DMO) serves data online from ocean researchers who collect data from a variety of platforms, instruments, and sensors. BCO-DMO provides both geospatial and text-based access to the data and allows one to retrieve data in several formats including tab and space separated tables, Matlab format, netCDF, and ODV file formats. To expand the number of data download options, we are implementing an ERDDAP file server and an interface to our system so that users can take advantage of several other download options.

ERDDAP requires detailed metadata to fully describe the corresponding data, including descriptions of the data columns, data units, investigator names, etc. While BCO-DMO has this information, it is necessary to extract this information from the BCO-DMO system and provide it to the ERDDAP system as an XML-formatted file

Project Progress: The first step was to understand what metadata tags ERDDAP uses and to map these to the appropriate metadata attributes in the BCO-DMO system.

```

--<addAttributes>
  <att name="cdm_data_type">Other</att>
  <att name="Conventions">COARDS, CF-1.6, Unidata Dataset Discovery v1.0</att>
  <att name="creator_name">Benjamin Walther</att>
  <att name="creator_email">benjamin.walther@tamucc.edu</att>
  <att name="infoUri">
    http://data.bco-dmo.org/jg/info/BCO/OtolithHypoxia/Sampling_Sites%7Bdir=data.bco-dmo.org/dir/BCO/Otolithhypoxia/
  </att>
  <att name="institution">
    Biological and Chemical Oceanography Data Management Office(BCO-DMO)
  </att>
  <att name="keywords">
    bco, bco-dmo, count, cruise, dataset, date, dmo, iso, noaa, seamp, station, statistics, time, width
  </att>
  <att name="license">[standard]</att>
  <att name="Metadata_Conventions">COARDS, CF-1.6, Unidata Dataset Discovery v1.0</att>
  <att name="sourceUri">(local files)</att>
  <att name="standard_name_vocabulary">CF-12</att>
  <att name="summary">
    2014 NOAA Seamp Fall Groundfish Cruise Stations where approximately 20 Atlantic croaker per station were collected for this project.
  </att>
  <att name="title">Dataset:2014 NOAA Seamp Cruise Station</att>
</addAttributes>
--<dataVariable>
  <sourceName>Station</sourceName>
  <destinationName>Station</destinationName>
  <dataType>String</dataType>
</dataVariable>

```

Tag name	description	available	comment
type	Identifies the dataset type (eg:EDDTableFromAsciiFiles)		a required attribute
DatasetID	The id of dataset eg:565679	1	a required attribute
iso19115File	Tell ERDDAP to use a per-made ISO 19115 file not to generate the file Usage:<iso19115File>fullFileName</iso19115File> fullFileName refer to a local or the URL.	1	An optional tag
reloadEvertNMinutes	Specifies how often the dataset should be reload	2	
updateEveryNMinutes	Work with datasets that change very frequently. Specifies how often the dataset should be update.	2	An optional tag
fileDir	The directory with the data files.eg:/data302/OtolithHypoxia	1	
recursive	Indicates if subdirectories of fileDir have data files (Boolean)	2	
charset	The character set of the source files. eg:UTF-8	1	UTF-8
columnNameRow	The number of the row with the column names in the files.	1	
firstDataRow	The number of the first row with data in the files	1	
preExtractRegex	No description	3	An optional tag
postExtractRegex	No description	3	An optional tag
extractRegex	No description	3	An optional tag

The next step will be to write the program to extract the information from the BCO-DMO metadata database.

Further work will be required to design and implement how best to make the data accessible to ERDDAP.

Acknowledgments: We would like to thank Bob Simon, primary creator and supporter of the ERDDAP system, for help in understanding the ERDDAP system and its XML encodings.

DATABASE	
Programs	36
Projects	701
Deployments	2370
Datasets	8131
Instruments	415
Parameters	1363
People	2114
Affiliations	489
Funding	83
Awards	1386

