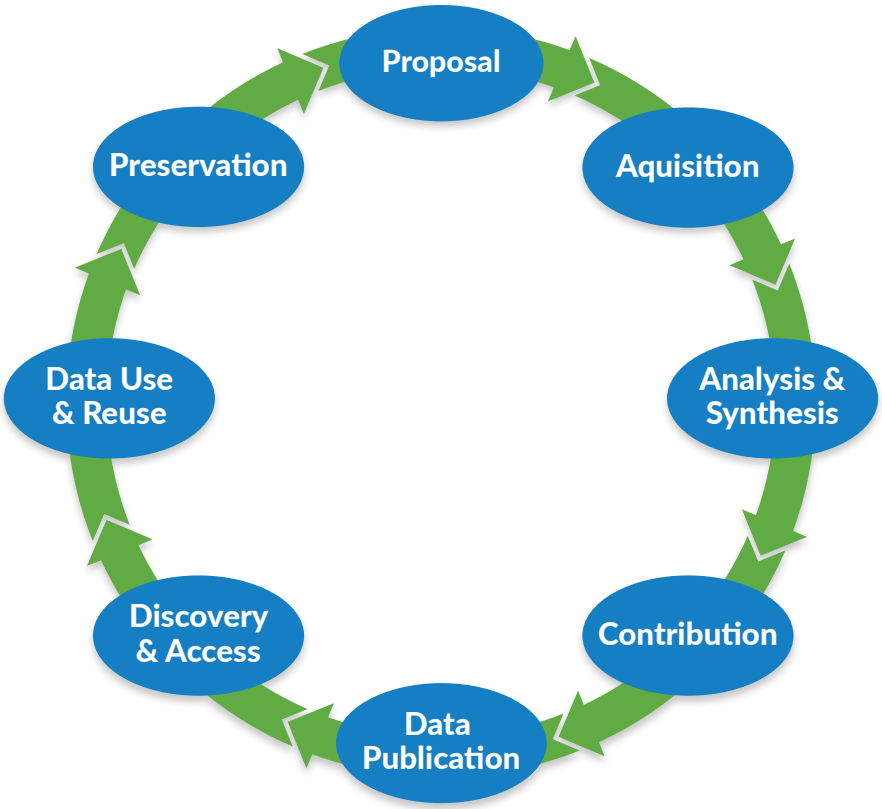
— BCO-DMO Quick Guide —

BCO-DMO

The Data Life Cycle



Curating and providing open access to research data is a collaborative process. This process may be thought of as a life cycle with data passing through various phases. Each phase has its own associated actors, roles, and critical activities. Good data management practices are necessary for all phases, from proposal to preservation.

BCO-DMO, a repository funded by the National Science Foundation (NSF), supports the oceanographic research community's data needs throughout the entire data life cycle. This guide describes the services available from BCO-DMO from proposal to preservation and highlights phases where researchers engage significantly with the office.

Data management services are provided free of charge for projects funded via:

- NSF-OCE Biological Oceanography Section
- NSF-OCE Chemical Oceanography Section
- Division of Polar Programs' Antarctic Organisms & Ecosystems Program

Not funded from the programs above? We can assist in determining the appropriate repository for your project data. See a list of other data management centers[1]. Under certain circumstances, we may negotiate services for data not covered by these NSF sections.

BCO-DMO provides the following services:

- Proposal: Help with your NSF Data Management Plan (see page 2).
- Acquisition: Advice on collecting good metadata and data.
- Contribution: Submission to the database, ensuring compliance with NSF OCE Sample and Data Policy[2] (NSF 17-037). One-on-one assistance with your data submission (see page 4).
- Data Publication: Datasets are published online at BCO-DMO[3]; citations in just one click; DOIs available (see page 8).
- Discovery & Access: BCO-DMO search tools (see page 9).
- Data Use & Reuse: Data are freely accessible*; many types of data are available for new and collaborative research/ modelling/synthesis projects.
- Preservation: BCO-DMO works with the appropriate national data center for long-term archiving (see page 13).

* If you need a limited additional period of time while manuscripts are prepared for publication before your data are publicly available at BCO-DMO, have a conversation with your Program Manager.

[1] https://www.bco-dmo.org/how-get-started#other_data_centers
[2] NSF 17-037; https://nsf.gov/pubs/2017/nsf17037/nsf17037.jsp?org=NSF
[3] https://www.bco-dmo.org

BCO-DMO Data Management Plan Template

Proposals submitted to NSF must include a supplementary document of no more than two pages labeled "Data Management Plan" (DMP). This supplementary document should describe how the proposal will conform to NSF's policy on the dissemination and sharing of research results.

Investigators working under awards granted by the NSF Division of Ocean Sciences (OCE) have additional conditions to which they must adhere, as described in the Division of Ocean Sciences Sample and Data Policy[1].

How BCO-DMO can help

BCO-DMO has developed a Data Management Plan template to assist investigators in submission of plans that meet the NSF OCE Sample and Data Policy requirements. The template can be found and completed on the DMPTool website.

DMPTool

Build your Data Management Plan

About DMPTool

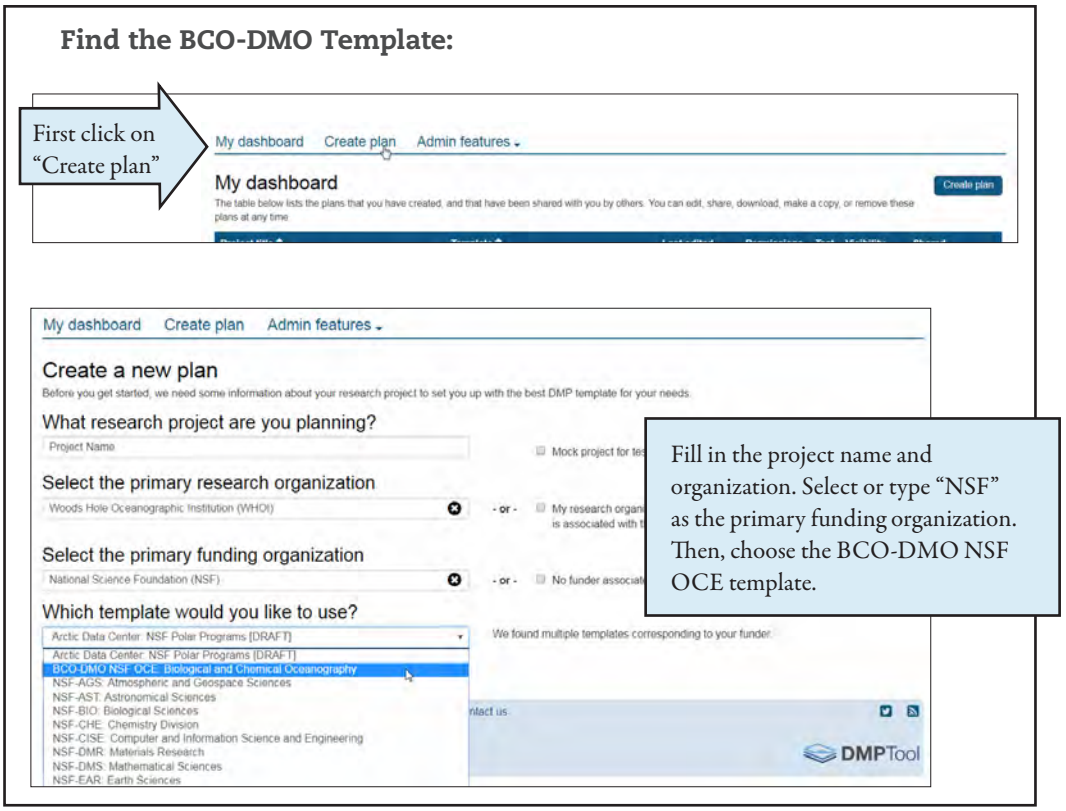
DMPTool[2] is a free, open-source, online application that helps researchers create data management plans. The DMPTool provides detailed guidance and links to informational resources and walks researchers through the process of generating comprehensive plans tailored to specific DMP requirements, in this case, the NSF OCE requirements.

If you are a researcher from one of the DMPTool partner institutions, you can log in using your institutional credentials. If your institution is not a partner, you can create your own account using any email address. In each section of the DMP template, you will see instructions containing the question or a description of information that should be provided to meet the specific requirement. Each question/requirement also has an example answer and links to additional guidance. The plan may be saved at any point, and can also be shared with collaborators. Once complete, your DMP can be exported in several different formats for inclusion in your NSF proposal.

Additional Resources

More information on NSF's data management requirements is available on BCO-DMO's website[3]. Detailed instructions on how to use BCO-DMO's DMPTool template are also available in our "Getting Started with DMPTool" guide[4].

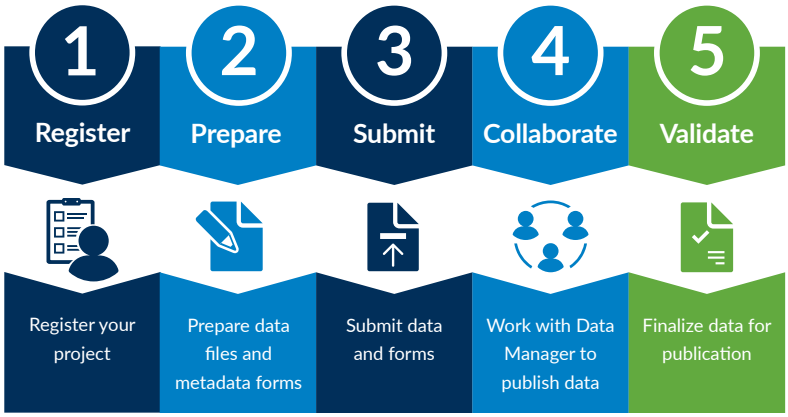
Image from BCO-DMO's "Getting Started with DMPTool" guide".



[1] https://www.nsf.gov/pubs/2017/nsf17037/nsf17037.jsp
[2] https://dmptool.org/
[3] https://www.bco-dmo.org/nsf-two-page-data-management-plan
[4] https://www.bco-dmo.org/sites/default/ les/page_ les/BCODMO_Getting_Started_with_DMPTool.pdf

Contributing Data to BCO-DMO

By depositing your project information and data into BCO-DMO, your data become shareable, citable resources available for community reuse. We are continually striving to make this process more efficient and streamlined, and welcome your feedback



1 Register a Project

A. Search BCO-DMO to see if your award is already in our system. If not,

B. Complete a Project Metadata Form[1] to provide information about projects that are not already registered at BCO-DMO.
- A project oversees a collection of one or more datasets.
- There is usually one project per NSF award with the exception of Collaborative Research awards where one project is funded by multiple award numbers. Some time-series projects may contain multiple awards as well.
- If you do not yet know the NSF award number, please provide as much information as possible including the project title and investigator contact information.

C. Submit your NSF Data Management Plan[2] with the Project Metadata Form.

2 Prepare Data and Metadata

A. Prepare the data files, including error checking and formatting. Understand what is meant by a 'dataset' relative to BCO-DMO (see page 4).

B. Complete a Dataset Metadata Form[3] to provide information about each unique dataset. See 'Preparing and contributing metadata' (page 4).

C. If data were collected from a research vessel, mooring, glider, or other unique deployment, complete a Deployment Metadata Form[4].
- Complete this form only if it is applicable to the dataset(s) you are submitting.
- Deployments help describe the geographic and temporal scale of datasets and provide context for mapping the associated data.
- The Rolling Deck to Repository, R2R[5], provides cruise data for vessels in the UNOLS fleet. For data on these cruises, you need only reference the R2R cruise identifier and BCO-DMO does the rest.

3 Submit

Send applicable metadata forms and data files to info@bco-dmo.org.

- You will receive a reply from one of our Data Managers confirming receipt of your forms and data files, if applicable.
- If a data set is too large to send as an email attachment, please contact us for instructions on the best way to contribute your data.
- NOTE: We strongly encourage you to submit data at least one month in advance of any pressing deadlines (e.g. NSF reports, manuscript publication) to provide adequate data processing time.

4 Collaborate

A Data Manager will begin the process of making the data available online. We strive to develop robust metadata that will ensure the data are easily discoverable and reusable. Your Data Manager will contact you with follow-up questions or requests for more information to ensure that the metadata is complete. This may be an iterative process, so your patience and cooperation are greatly appreciated.

5 Validate

Once your datasets are online, you'll be asked to review the data and metadata for completeness and accuracy. This validation stage is the final step in the process, and necessary for assignment of DOIs and long-term archive.

Once datasets are reviewed and validated by the contributor, BCO-DMO ensures that the data are archived properly at the appropriate National Data Center (e.g. National Centers for Environmental Information, NCEI[6]).

[1] http://www.bco-dmo.org/ les/bcodmo/PROJECT.rtf
[2] http://www.bco-dmo.org/nsf-two-page-data-management-plan#DMP_Template
[3] http://www.bco-dmo.org/ les/bcodmo/DATASET.rtf
[4] http://www.bco-dmo.org/ les/bcodmo/DEPLOYMENT.rtf
[5] http://www.rvdata.us/
[6] https://www.ncei.noaa.gov/

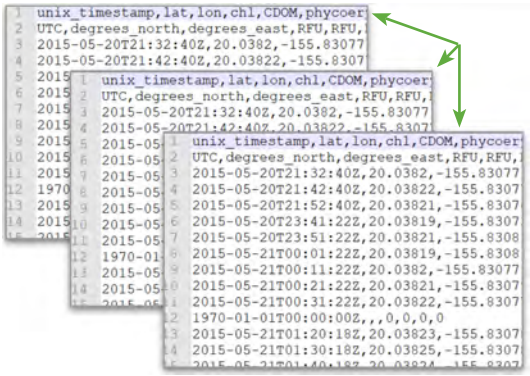
Preparing Data and Metadata

You should submit data in the format most appropriate for your community. If this format is proprietary or non-tabular, Data Managers will create a tabular version of your data to import into the BCO-DMO data system. If the most appropriate format is not one of the various output formats provided by the BCO-DMO data system (e.g. .csv., .tsv, .nc, .mat), we will work with you to arrive at the best data representation possible.

Data before and after submission

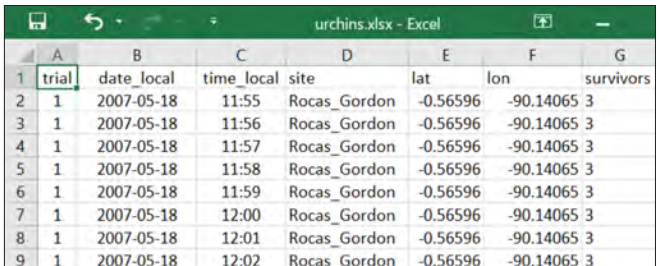
Before

These are plain text files of the same type (e.g. one file per cast) so they can be combined into one dataset



Excel file

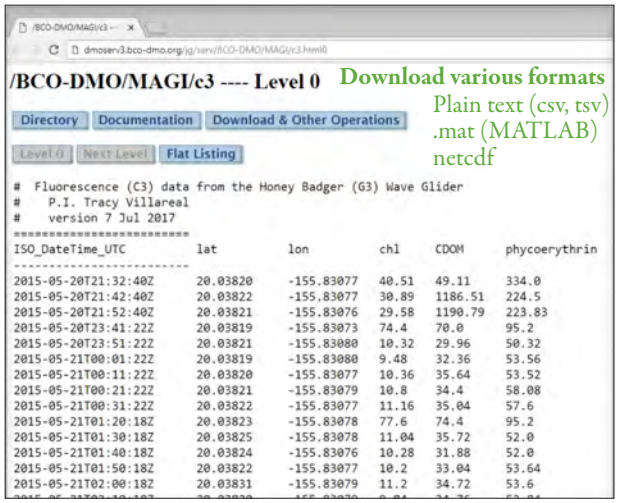
Different data types cannot be combined; these are split into separate datasets.

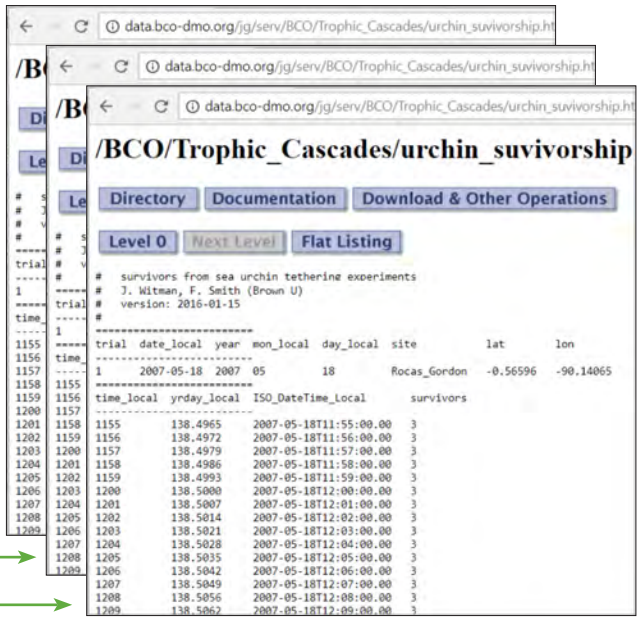


trial	date_local	time_local	site	lat	lon	survivors
1	2007-05-18	11:55	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	11:56	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	11:57	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	11:58	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	11:59	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:00	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:01	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:02	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:03	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:04	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:05	Rocas_Gordon	-0.56596	-90.14065	3
1	2007-05-18	12:06	Rocas_Gordon	-0.56596	-90.14065	3

After

Imported into BCO-DMO's data system





Data Preparation Tips

General tips:

- Round your data to the appropriate number of decimal places.
- Make sure all flags and codes are documented in your metadata.
- Submit measured or observed values, not just statistical and calculated values.

Excel files:

- Remove formatting that will not be preserved when exported as a plain text file (e.g. color, merged cells, plots, etc.)
- Only include one tabular dataset (i.e. table) per Excel sheet.
- Ensure cells contain intended values. Check formula results, references to other sheets, hyperlinks, etc.

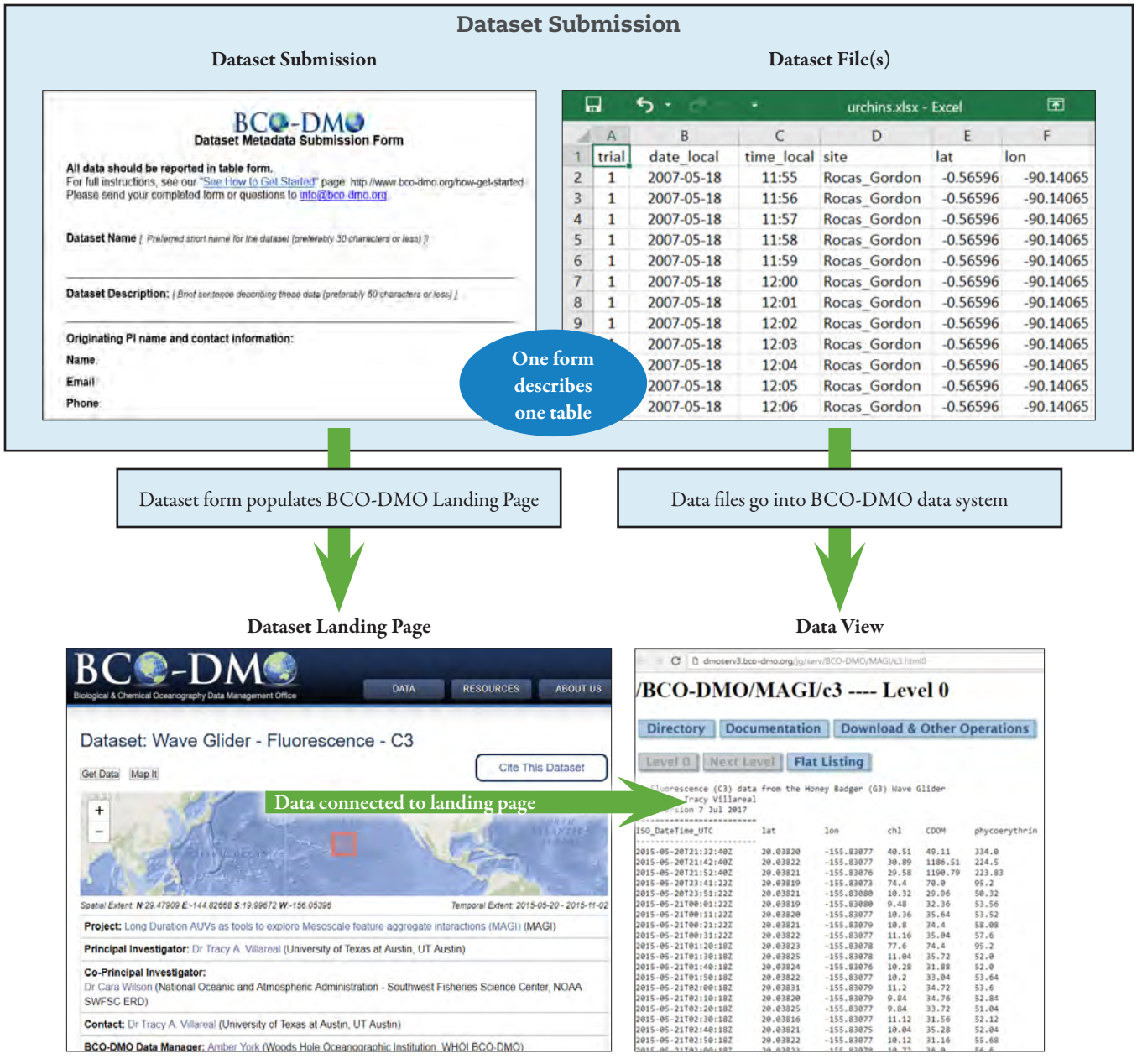
Error checking:

- QA/QC your data before submitting.
- Check species name for correct spelling and use taxonomically accepted names.

Dates & locations:

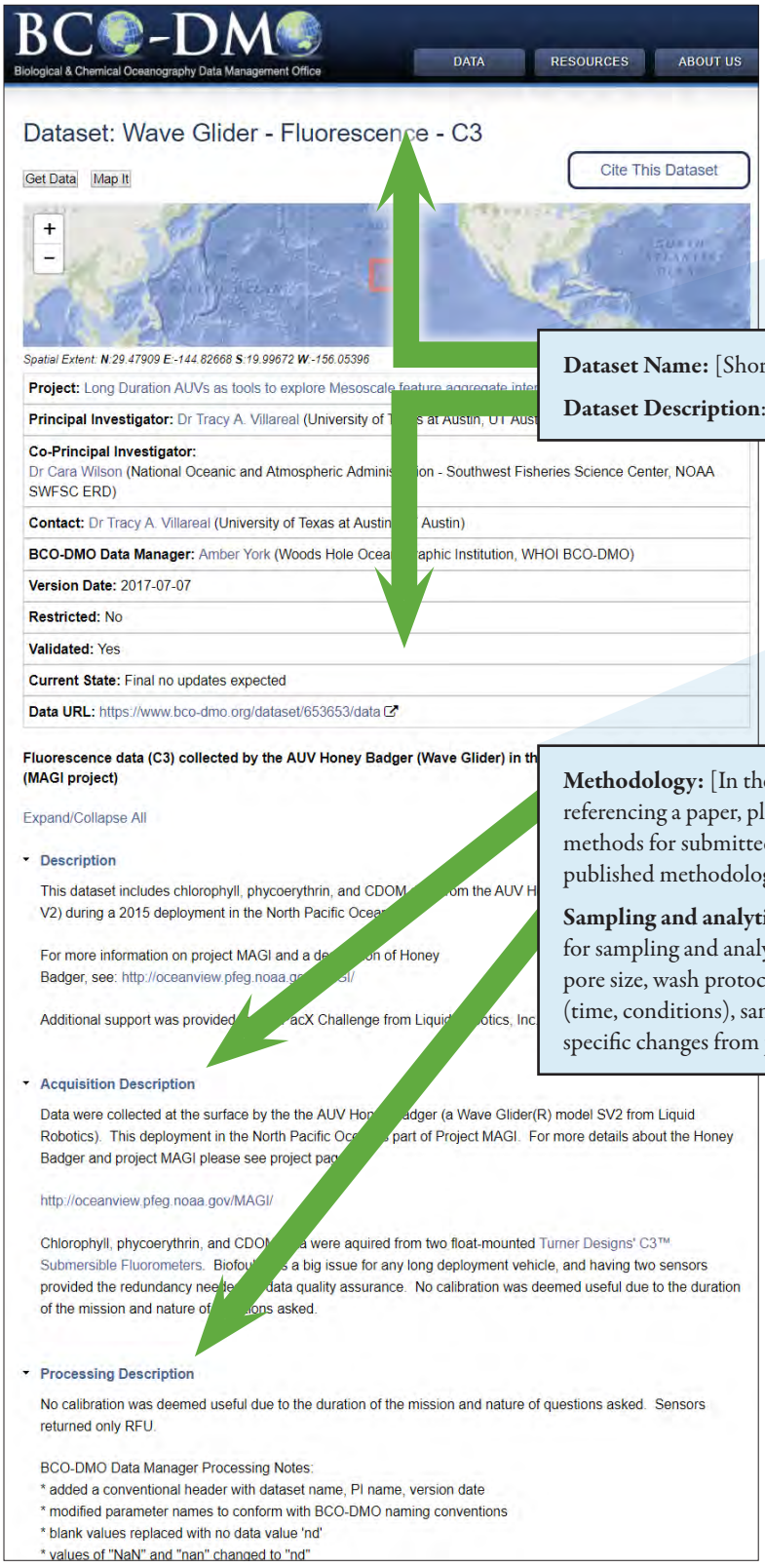
- Document your time and date format including time zone (e.g. UTC, UTC+02, local EST).
- Check for inconsistent date/time formatting.
- In-situ data: include date/time and lat/lon.
- Experimental data: include date/time of experiments if applicable.

Connecting Data and Metadata

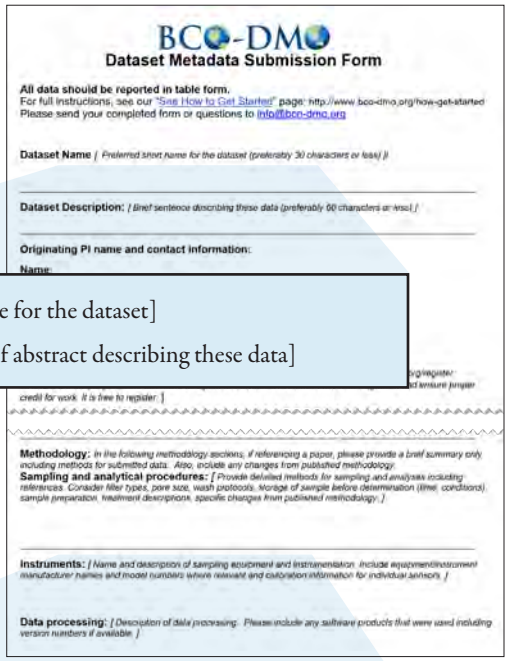


Metadata

BCO-DMO uses a form to capture important information about your dataset, such as where and how it was collected, analysis methods, and funding sources. This information is known as "metadata". The metadata you provide about your data through the form should be thorough, complete, and publication ready. The contents of your metadata form are directly used to populate the public Dataset Landing Page.



Dataset Metadata Form



Dataset Name: [Short name for the dataset]

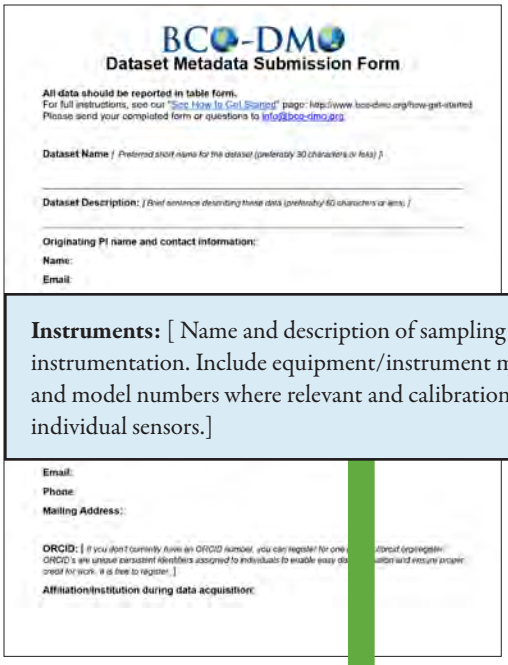
Dataset Description: [Brief abstract describing these data]

Methodology: [In the following methodology sections, if referencing a paper, please provide a brief summary only including methods for submitted data. Also, include any changes from published methodology.]

Sampling and analytical procedures: [Provide detailed methods for sampling and analyses including references. Consider filter types, pore size, wash protocols, storage of sample before determination (time, conditions), sample preparation, treatment descriptions, specific changes from published methodology.]

Metadata

Dataset Metadata Form


BCO-DMO
Dataset Metadata Submission Form

All data should be reported in table form.
For full instructions, see our "See How to Get Started" page: http://www.bco-dmo.org/how-get-started
Please send your completed form or questions to info@bco-dmo.org

Dataset Name [Preferred short name for the dataset (preferably 30 characters or less)]

Dataset Description: [Brief sentence describing these data (preferably 60 characters or less)]

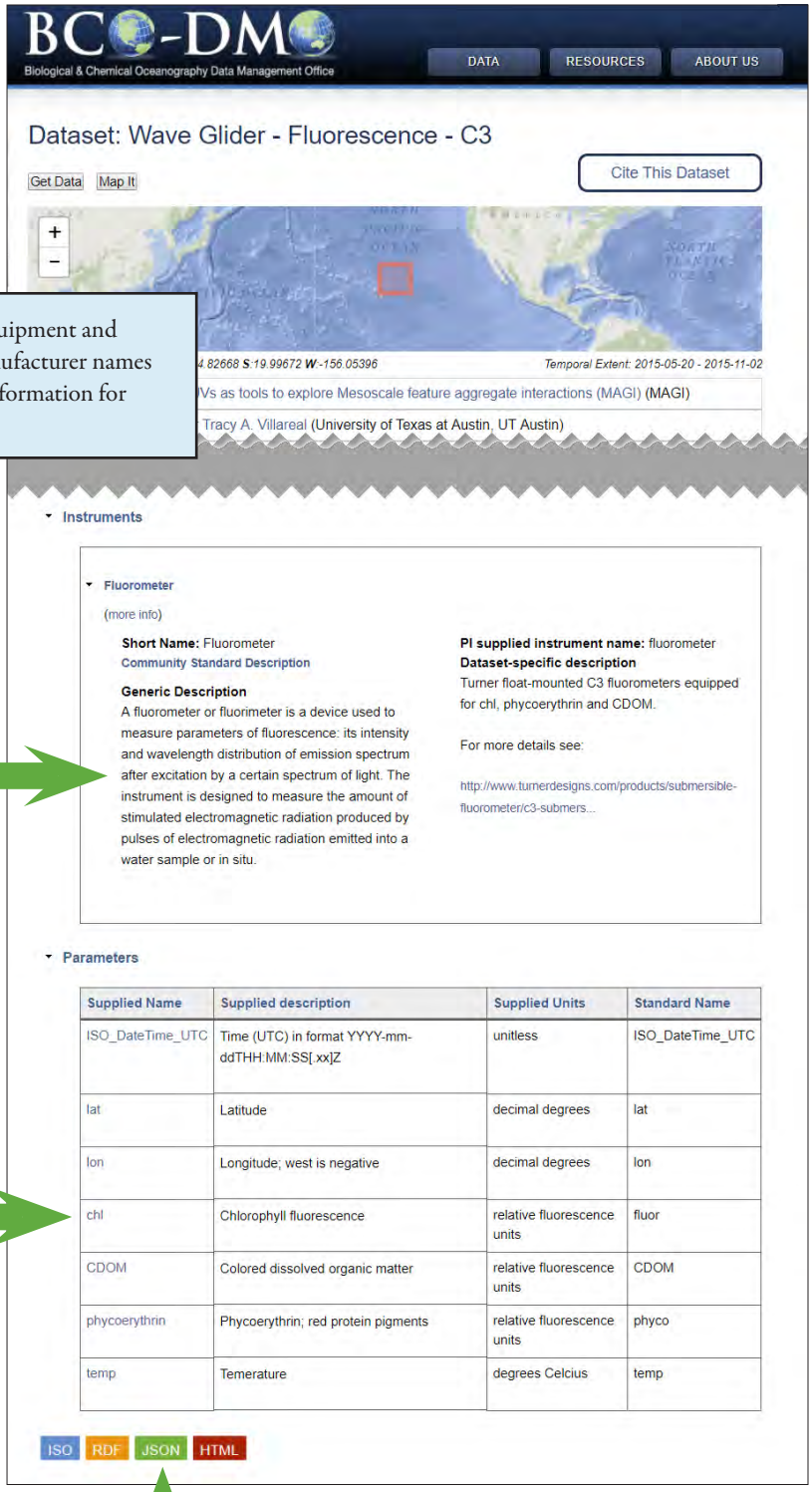
Originating PI name and contact information:
Name:
Email:

Email:
Phone:
Mailing Address:

ORCID: [If you don't currently have an ORCID number, you can register for one ... ORCID's are unique persistent identifiers assigned to individuals to enable easy ... creation and ensure proper credit for work. It is free to register.]

Affiliation/Institution during data acquisition:

Instruments: [Name and description of sampling equipment and instrumentation. Include equipment/instrument manufacturer names and model numbers where relevant and calibration information for individual sensors.]


BCO-DMO
Biological & Chemical Oceanography Data Management Office
DATA RESOURCES ABOUT US

Dataset: Wave Glider - Fluorescence - C3

Get Data Map It Cite This Dataset

4.82668 S:19.99672 W:-156.05396 Temporal Extent: 2015-05-20 - 2015-11-02

Vs as tools to explore Mesoscale feature aggregate interactions (MAGI) (MAGI)

Tracy A. Villareal (University of Texas at Austin, UT Austin)

Instruments

Fluorometer

(more info)

Short Name: Fluorometer
Community Standard Description

Generic Description
A fluorometer or fluorimeter is a device used to measure parameters of fluorescence: its intensity and wavelength distribution of emission spectrum after excitation by a certain spectrum of light. The instrument is designed to measure the amount of stimulated electromagnetic radiation produced by pulses of electromagnetic radiation emitted into a water sample or in situ.

PI supplied instrument name: fluorometer
Dataset-specific description
Turner float-mounted C3 fluorometers equipped for chl, phycoerythrin and CDOM.

For more details see:

http://www.turnerdesigns.com/products/submersible-fluorometer/c3-submers...

Parameters

Supplied Name	Supplied description	Supplied Units	Standard Name
ISO_DateTime_UTC	Time (UTC) in format YYYY-mm-ddTHH:MM:SS[.xx]Z	unitless	ISO_DateTime_UTC
lat	Latitude	decimal degrees	lat
lon	Longitude; west is negative	decimal degrees	lon
chl	Chlorophyll fluorescence	relative fluorescence units	fluor
CDOM	Colored dissolved organic matter	relative fluorescence units	CDOM
phycoerythrin	Phycoerythrin; red protein pigments	relative fluorescence units	phyco
temp	Temerature	degrees Celcius	temp

ISO RDF JSON HTML

Parameter names, descriptions, units:
[Parameter names are the column headings]

You can provide this information in the metadata form or as a separate text file or Excel spreadsheet.

Other critical metadata to be included in your form

- Project and Funding information
- Geographic Location
- Cruise/Deployment information
- Access Restrictions - dependent on funding policies
- Related files and references
- Data Processing
- Problem report

Metadata can be downloaded in various formats by clicking the badges at the bottom of the Dataset Landing Page.

Data Publication

1. BCO-DMO publishes data and metadata, fostering data discoverability, access, reuse, and attribution.

2. DOIs are generated for every dataset and will be assigned for all submissions. All datasets must be final and validated before a DOI is assigned.

3. We provide a recommended citation, so that users can properly cite each dataset.

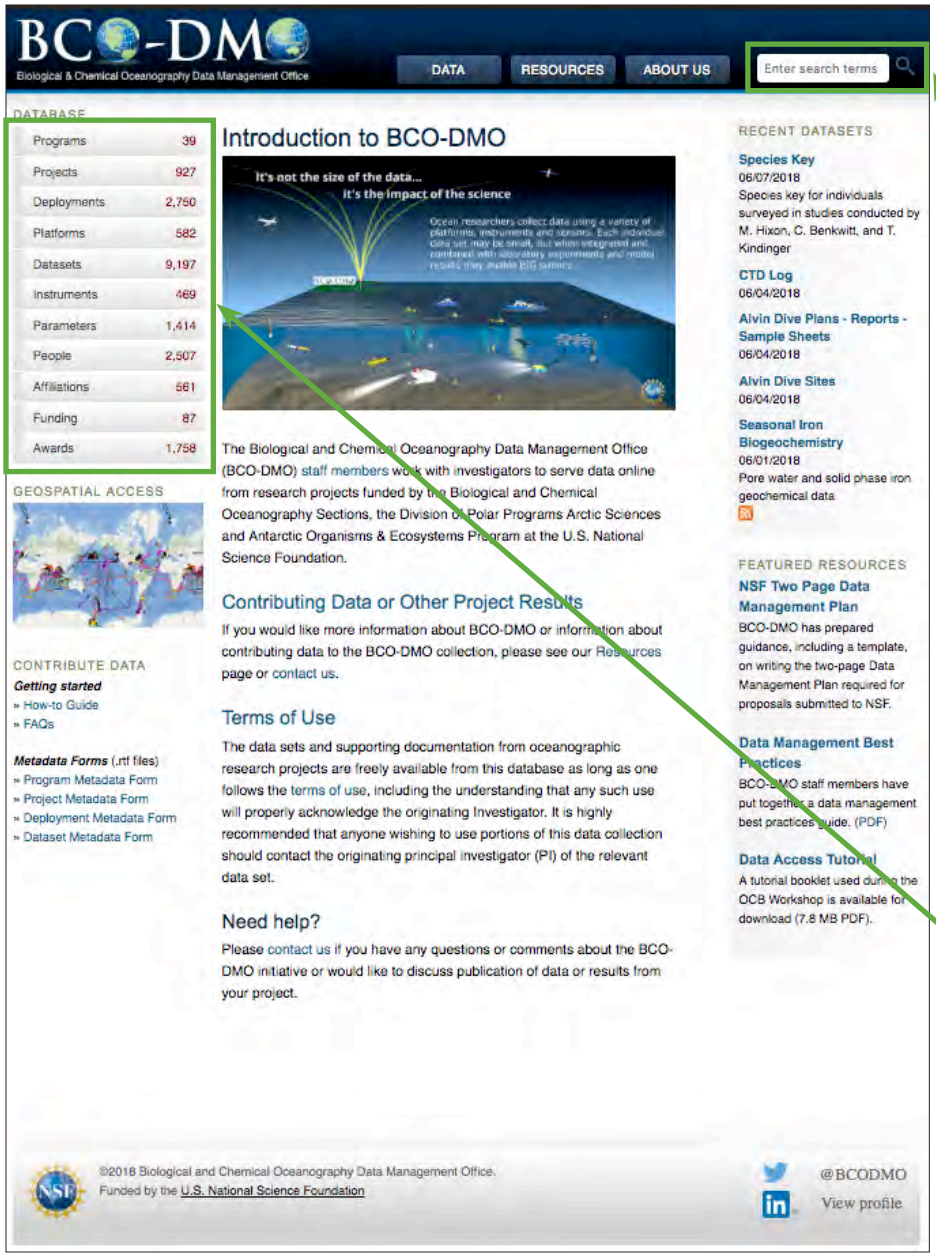
4. All datasets available at BCO-DMO are licensed under a Creative Commons Attribution 4.0 International license, ensuring that each data contributor will receive proper credit.

5. All DOIs are minted for archive by the WHOI Open Access Server (WHOAS), and resolve to WHOAS landing pages.

Discovery and Access

Once data are processed and published online, the BCO-DMO website enables data discovery via text and geospatial search interfaces, making it easy for users to find datasets of choice. Through text-based searches, the database can be searched by cruise, project, person, or any keyword provided in metadata upon submission. Access to data is made possible from the Dataset Landing pages, and data may be subsetted, plotted, and reformatted prior to download. The BCO-DMO database encompasses the full range of oceanographic measurement types from limnological, physical, chemical, biological and/or ecological, and biogeochemical sub-domains.

http://bco-dmo.org/



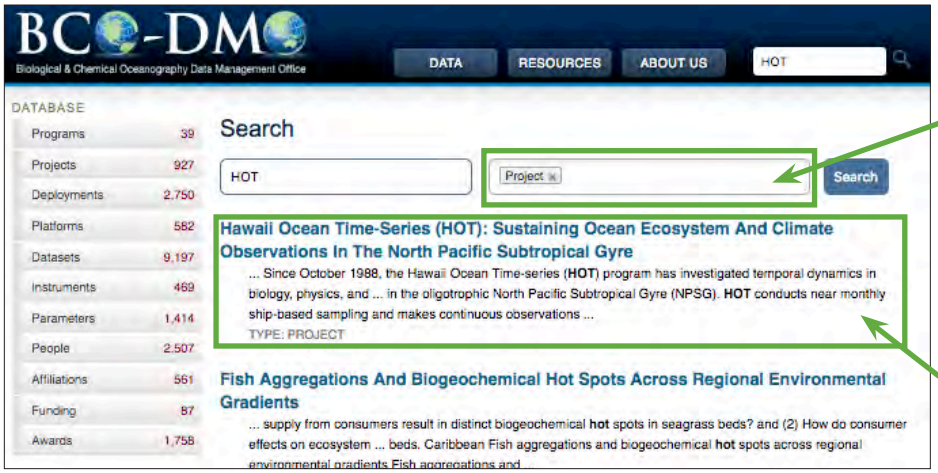
Search for any keyword. This can be a type of data (pH), a project name or acronym (HOT), a person's name (John Smith), or a funding number (OCE-0926766).

Any keyword that might be associated with the data you are interested in.

If you know the type of keyword you are searching for, the left-hand tabs allow search of the specific database fields.

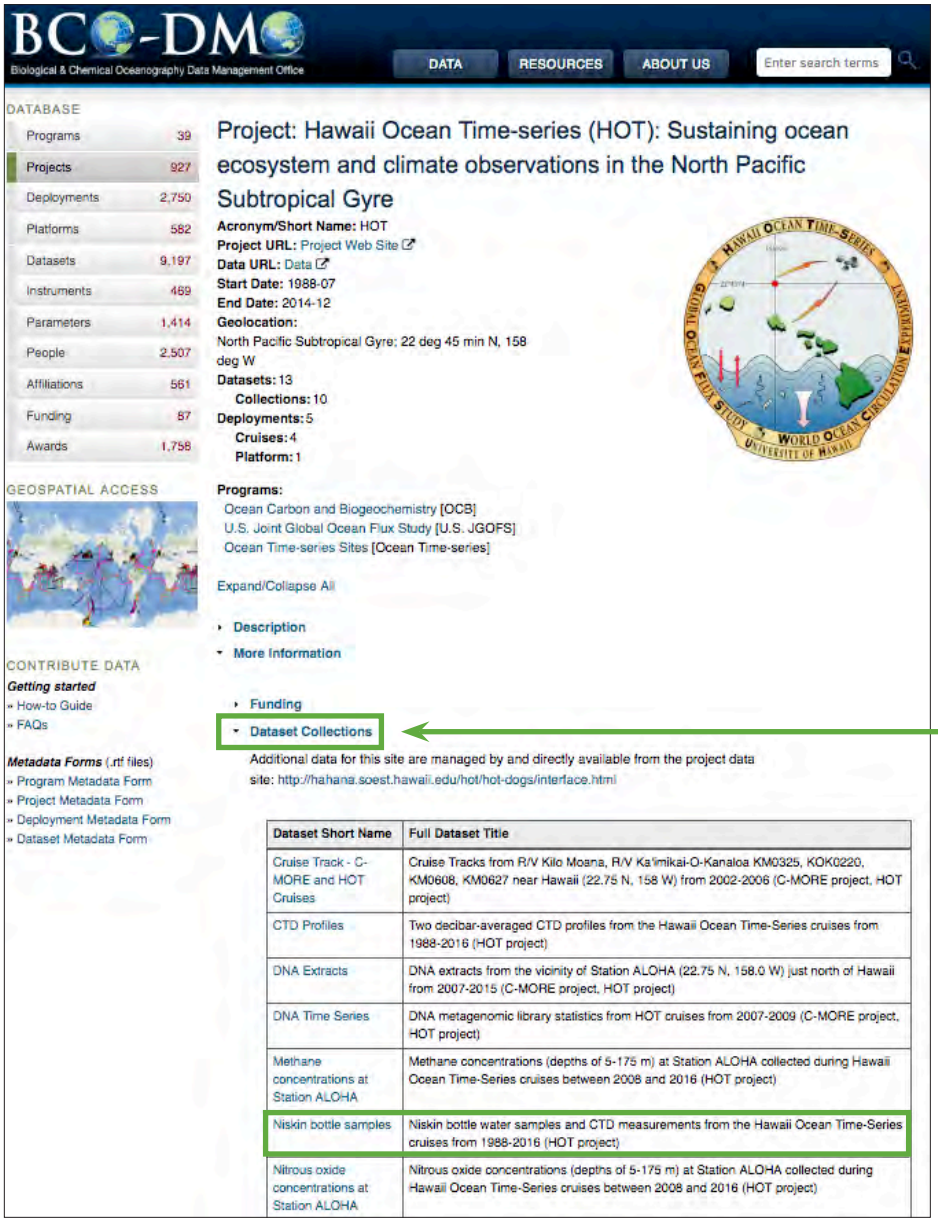
The following example searches for the project HOT (the acronym for Hawaiian Ocean Time-series) and downloads the niskin bottle data.

Discovery and Access



If your search does not return the result of interest, try to filter the search. In this case, we can filter by the type "Project" since we know it's a project in our system.

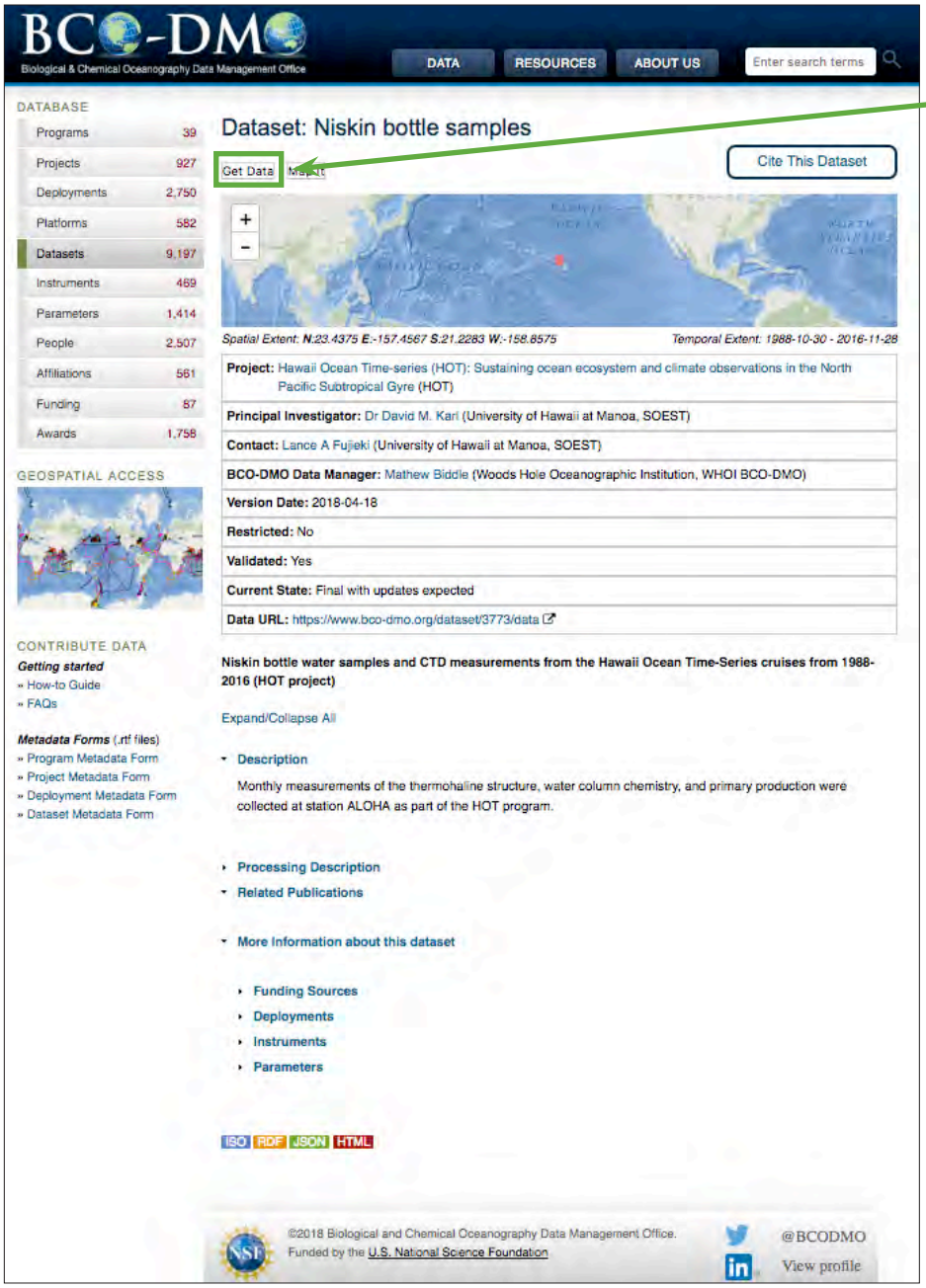
Clicking on the title will take you to the metadata page for the type of record you select (defined in grey here, as TYPE: PROJECT). There you can see various metadata elements describing the record. This includes individual datasets associated with that record.



The Dataset Collections section of the Project metadata page provides links to datasets associated with this specific project.

The Dataset Short Name link will take you to the Dataset Landing page for that dataset.

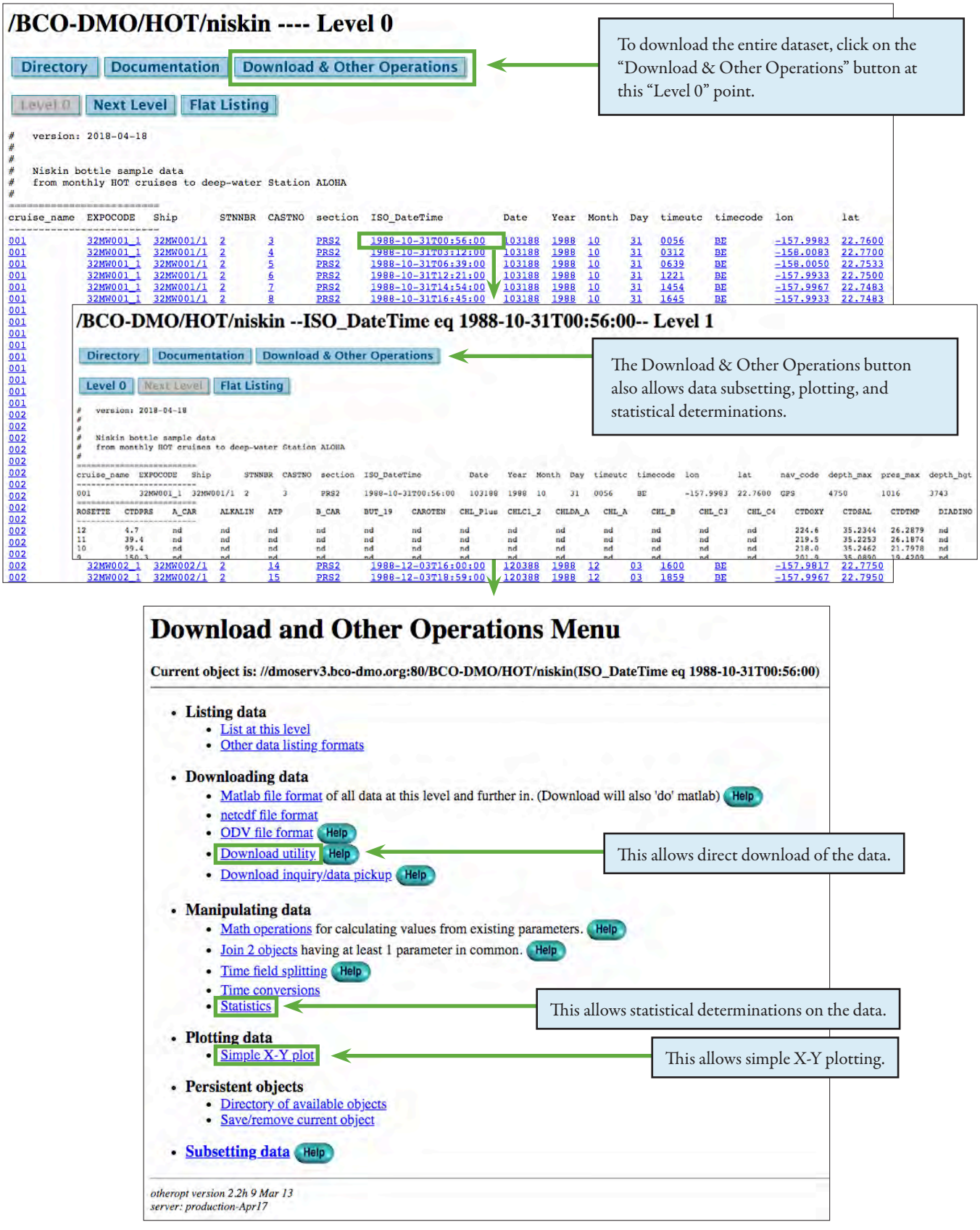
Discovery and Access



The "Get Data" button allows you to look at all of the data values submitted to BCO-DMO.

Discovery and Access

This page displays data values in a hierarchical view (beginning at "Level 0"). Blue text indicates clickable values that expand to uncover more data.



Preservation

Preservation marks a maturity level that allows data to begin the data life cycle again in new research endeavors. BCO-DMO serves as a domain specific, intermediate data repository, and as such does not function as a long-term archive for data preservation. BCO-DMO provides data management support throughout a project award's period of performance which, prepares project output for reuse and reanalysis by the community. Once a project's data and metadata are published online at BCO-DMO, they are then submitted to an appropriate national data center for long-term preservation (e.g., the National Centers for Environmental Information).

FAQ's

Many Frequently Asked Questions are also addressed on our website at https://www.bco-dmo.org/faq-page. Still have questions? Feel free to contact the office at info@bco-dmo.org and a team member will respond.

Acknowledgements

The Biological and Chemical Oceanography Data Management Office (BCO-DMO) is funded by the U.S. National Science Foundation award 1435578. The office gratefully acknowledges the Woods Hole Oceanographic Institution's Graphics Department for their support and creativity in website development and outreach materials.

Find us on twitter at @BCODMO