

# National Science Foundation (nsf.gov): BCO-DMO NSF OCE: Biological and Chemical Oceanography

## Data Policy Compliance

Identify any published data policies with which the project will comply, including the NSF OCE Data and Sample Policy as well as other policies that may be relevant if the project is part of a large coordinated research program (e.g. GEOTRACES).

*Example Answer:*

The project investigators will comply with the data management and dissemination policies described in the NSF Award and Administration Guide (AAG, Chapter VI.D.4) and the NSF Division of Ocean Sciences Sample and Data Policy.

*Guidance:*

- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)
- [NSF Guidance on Data Management Plans](#)
- [NSF GEO Directorate Data Policies](#)

## Pre-Cruise Planning

If the proposed project involves a research cruise, describe the cruise plans. (Skip this section if it is not relevant to your proposal.) Consider the following questions:

1. How will pre-cruise planning be coordinated? (e.g. email, teleconference, workshop)
2. What types of sampling instruments will be deployed on the cruise?
3. How will the cruise event log be recorded? (e.g. the Rolling Deck to Repository (R2R) event logger application, an Excel spreadsheet, or paper logs)
4. Will you prepare a cruise report?

*Example Answer:*

Pre-cruise planning will be done via teleconferencing and a planning workshop. Detailed plans for station locations, instrument deployment, water sampling strategy, and water sample allocation will be written up as a science implementation plan for the cruise. The actual sampling events will be recorded on paper logs (scanned into PDF documents) and/or in a digital event log using the R2R event logger application (if available).

*Guidance:*

If cruise plans are not known at this time, it is appropriate to omit this section or to state that cruise plans will be made at a later date. Funded projects that involve deployments (including research cruises as well as deployments of moorings, floats, and gliders) will be expected to provide deployment metadata to BCO-DMO along with dataset metadata.

- [BCO-DMO "How-To" Pages](#)
- [University-National Oceanographic Laboratory System \(UNOLS\)](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)
- [R2R Scientific Sampling Event Log](#)

## Description of Data Types

Provide a description of the types of data to be produced during the project. Identify the types of data, samples, physical collections, software, derived models, curriculum materials, and other materials to be produced in the course of the project. Include a description of the location of collection, collection methods and instruments, expected dates or duration of collection. If you will be using existing datasets, state this and include how you will obtain them.

*Example Answer:*

The project will produce several observational and experimental datasets, described in the list below. In addition to the datasets described below, educational resources produced by the project, including data and images, will be made available for public use on the COSEE.net website. Observational data will be collected on a North Atlantic research cruise planned to take place during the summer months (July-August).

### **Observational Datasets:**

1. **CTD and Niskin bottle data:** CTD data collected using a SeaBird SBE CTD package; processing to be done using SeaBird's SeaSave software; data will include standard environmental measurements (such as pressure, temperature, salinity, fluorescence). File types: Raw (.con, .hdr, .hex, .bl) and processed and .cnv, .asc, .btl) ASCII files. Repository: BCO-DMO
2. **Event log:** Cruise scientific sampling event log; will include event numbers, start/end dates, times & locations of instrument deployments. Will be recorded using the R2R event logger (if available) and on paper log sheets. File types: Excel file converted to .csv; scanned PDFs. Repository: BCO-DMO and Rolling Deck to Repository (R2R).
3. **Cruise underway data:** Routine underway data collected along the ship's track (including meteorological data, sea surface temperature, salinity, fluorescence, ADCP). Will be collected by the shipboard instrumentation. File types: .csv ASCII files. Repository: BCO-DMO and R2R.
4. **Zooplankton sampling logs and images:** Zooplankton will be sampled via Reeve net trawls and MOCNESS (Multiple Opening/Closing Net and Environmental Sensing System) tows during the cruise. Species identified, tow numbers, locations, depths, dates, and times will be recorded by hand on log sheets. Information from log will be transferred into an Excel spreadsheet. Photographs of each tow/rawl will be taken on the ship using a digital camera. File types: PDF files of scanned log sheets; Excel files of sampling logs; images (.jpg files). Repository: BCO-DMO.

### Experimental Datasets:

1. **Pteropod respiration:** Physiological experiments carried out on pteropods captured at sea raised under controlled pCO<sub>2</sub> conditions; dataset will include data on the experimental treatments and the observed respiration rates. Animals will be captured using a Reeve Net or MOCNESS. Experiments will be conducted in the ship's lab. File types: Excel file(s). Repository: BCO-DMO.
2. **Genetic sequencing:** mRNA and DNA sequences from animals collected at sea. Sequencing will be performed at the PI's lab in Woods Hole, MA following the research cruise. File types: Short-read archive (.sra) and .fasta files. Repository: NCBI; accession numbers to be provided to BCO-DMO.

### Guidance:

It may be useful to group data into four categories:

1. Observational (e.g. in-situ, collected in the field). Examples may include: shipboard underway data; mesozooplankton samples collected by a net system; copepod specimens collected, identified, and preserved; hydrographic casts; alongtrack data; remote sensing (e.g. ocean color); acoustic data.
2. Experimental (e.g. generated in a lab or under controlled conditions). Examples: controlled carbonate chemistry experiments; DNA and RNA sequences.
3. Simulations (e.g. machine-generated). Example: models and their output.
4. Derived (e.g. synthesized from existing datasets). Examples: compiled database, products, reports.

Keep in mind that code or software developed through your project must also be made publicly available in compliance with NSF Policies. Code may be contributed to BCO-DMO or published with a DOI through Zenodo.

If the expected dates or duration of collection are not known at this time (e.g. due to funding schedules or ship availability), it is appropriate to give approximations, the ideal dates/duration, or to state that these details will be determined at a later date.

- [BCO-DMO "How-To" Pages](#)
- [BCO-DMO Guidance on Models, Software, and Code](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)

### Data and Metadata Formats and Standards

Identify the formats and standards to be used for data and metadata formatting and content. Where existing standards are absent or deemed inadequate, these formats and contents should be documented along with any proposed solutions or remedies. Consider the following questions:

1. Which file formats will be used to store your data?
2. What type of contextual details (metadata) will you document and how?

3. Are there specific data or metadata standards that you will be adhering to?
4. Will you be using or creating a data dictionary, code list, or glossary?
5. What types of quality control will be used? How will data quality be assessed and flagged?

*Example Answer:*

Field observation data will be stored in flat ASCII files, which can be read easily by different software packages. Field data will include date, time, latitude, longitude, cast number, and depth, as appropriate. Quality flags will be assigned according to the ODS IODE Quality Flag scheme (IOC Manuals and Guides, 54, volume 3; [http://www.iode.org/mg54\\_3](http://www.iode.org/mg54_3)). Metadata will be prepared in accordance with BCO-DMO conventions (i.e. using the BCO-DMO metadata forms) and will include detailed descriptions of collection and analysis procedures.

*Guidance:*

Usually, data served by BCO-DMO are submitted as comma- or tab-separated ASCII files (.csv, .txt) or as spreadsheet files (.xls, .xlsx). However, BCO-DMO is flexible and willing to work with whatever reasonably organized format the investigator uses. Whenever possible, quality flags should be included as columns within each dataset. The use of standard vocabularies/naming conventions is highly recommended.

- [BCO-DMO "How-To" Pages](#)
- [BCO-DMO Guidance on Organizing Data Tables](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)

## Data Storage and Access During the Project

Describe how project data will be stored, accessed, and shared among project participants during the course of the project. Consider the following:

1. How will data be shared among project participants during the data collection and analysis phases? (e.g. web page, shared network drive)
2. How/where will data be stored and backed-up?
3. If data volumes will be significant, what is the estimated total file size?

*Example Answer:*

The investigators will store project data (including spreadsheets, ASCII files, images, and PDFs of scanned logs) on laboratory computers that are backed up by the University's central IT organization. The Principal Investigator (PI) has also established an account with the San Diego Super Computer's enterprise class Cloud Service for data storage and sharing among project investigators. Personal computers in all laboratories are backed up daily using Apple Time Machine to an onsite external hard drive, and weekly to an offsite hard drive.

*Guidance:*

- [BCO-DMO "How-To" Pages](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)

## Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution

Describe mechanisms for data access and sharing, and describe any related policies and provisions for re-use, re-distribution, and the production of derivatives. Include provisions for appropriate protections of privacy, confidentiality, security, intellectual property, or other rights or requirements. Consider the following:

1. When will data be made publicly available and how? Identify the data repositories you plan to use to make data available.
2. Are the data sensitive in nature (e.g. endangered species concerns, potential patentability)? If so, is public access inappropriate and how will access be provided? (e.g. formal consent agreements, restricted access)
3. Will any permission restrictions (such as an embargo period) need to be placed on the data? If so, what are the reasons and what is the duration of the embargo?
4. Who holds intellectual property rights to the data and how might this affect data access?
5. Who is likely to be interested in re-using the data? What are the foreseeable re-uses of the data?

*Example Answer:*

Immediately after completion of the research cruise, underway data and metadata will be submitted

to the Rolling Deck to Repository (R2R) project. DNA sequences will be deposited in the National Center for Biotechnology Information (NCBI) database GenBank upon submission of manuscripts. GenBank accession numbers and related metadata will be provided to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) in an Excel spreadsheet or .CSV file and metadata will be provided using the BCO-DMO online Submission Tool. Data sets produced by the science party will be made available through the BCO-DMO data system within two-years from the date of collection. The project investigators will work with BCO-DMO data managers to make project data available online in compliance with the NSF OCE Sample and Data Policy. Data, samples, and other information collected under this project can be made publicly available without restriction once submitted to the public repositories.

Data produced by this project may be of interest to chemical and biological oceanographers, and climate scientists interested in the role of biogeochemistry in the global climate system. We will adhere to and promote the standards, policies, and provisions for data and metadata submission, access, re-use, distribution, and ownership as prescribed by the BCO-DMO Terms of Use (<http://www.bco-dmo.org/terms-use>).

*Guidance:*

Explain how and when data will be made available. Also describe any re-use and re-distribution policies and how the data sharing plans are related to those policies. Identify who will be allowed to use your data and whether or not they will be allowed to disseminate your data. If access to, use of, or dissemination of data will be restricted, explain how you will codify and communicate these restrictions.

**Data Sharing via BCO-DMO**

If the proposal is being submitted to the NSF Division of Ocean Sciences' (OCE) Biological or Chemical Oceanography Sections or Division of Polar Programs (PLR) Antarctic Sciences (ANT) Organisms & Ecosystems Program, then the two page plan can state that BCO-DMO staff will work with you to manage the data, and that data or model results generated during the proposed research project will be contributed to the BCO-DMO system. BCO-DMO provides data management services at no additional cost to projects funded by these NSF sections/programs.

Project investigators funded by these sections/programs can work with BCO-DMO to make project data available online, in compliance with the NSF OCE Sample and Data Policy. PIs of funded projects will be expected to submit project metadata to BCO-DMO beginning with the award/proposal number and DMP for proposals that are recommended for funding.

BCO-DMO can deal with a wide variety of data, including but not limited to biological, chemical, and physical oceanography measurements. BCO-DMO data managers routinely serve in-situ data including standard hydrographic, biogeochemical, biological, ecological, and microbial measurements and chemical tracers; experimental and model results; images and movies, etc. If you are uncertain if BCO-DMO is an appropriate repository for your data, please contact [info@bco-dmo.org](mailto:info@bco-dmo.org).

Refer to the OCE Sample and Data Policy for information on other suggested databases and repositories for physical samples.

**Genomic Data and Other Specialized Repositories**

Provisions should be made for the sharing and storage of genetic and molecular data in a publicly accessible, permanent database such as the various NCBI databases (e.g. GenBank), RAST, MG-RAST, etc. Information about the types of data accepted by GenBank is available on their website at [http://www.ncbi.nlm.nih.gov/genbank/submit\\_types](http://www.ncbi.nlm.nih.gov/genbank/submit_types). If known, it'd be helpful to disclose the details and availability of bioinformatics pipelines that will be used in next-generation methods. Accommodations for sharing of metagenome, metatranscriptome, and proteomics data should also be described in the DMP.

BCO-DMO can enable discovery of data that have been contributed to specialized repositories (e.g. NCBI, LTER data catalog, CDIAC). For example, genetic sequence data are best served by an NCBI repository such as GenBank. Metadata and accession numbers/URLs/unique identifiers can then be provided to BCO-DMO so that access to these alternate repositories can be provided through the BCO-DMO website.

**Underway Shipboard Data**

All routine underway data collected by vessel-resident instrumentation aboard UNOLS-supported oceanographic research vessels will be submitted to the appropriate long-term archive through the Rolling Deck to Repository (R2R) program. The PI is responsible for disseminating the data and metadata produced by the science party's research.

- [BCO-DMO "How-To" Pages](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)
- [BCO-DMO's Terms of Use](#)

- [R2R Cruise Catalog](#)
- [Creative Commons License Types](#)
- [National Center for Biotechnology Information \(NCBI\)](#)
- [GenBank](#)
- [System for Earth Sample Registration \(SESAR\)](#)

## Plans for Archiving

Describe the plans for long-term archiving of data, samples, and other research products, and for preservation of access to them. Consider the following:

1. What is your long-term strategy for maintaining, curating, and archiving the data?
2. What archive(s) have you identified as a place to deposit data and other research products?

*Example Answer:*

R2R will ensure that the original underway measurements are archived permanently at NCEI and/or NGDC as appropriate. BCO-DMO will also ensure that project data are submitted to the appropriate national data archive. The PI will work with R2R and BCO-DMO to ensure data are archived appropriately and that proper and complete documentation are archived along with the data.

*Guidance:*

After data contributed to BCO-DMO are online and fully documented, BCO-DMO ensures that the data are archived for long-term storage. The Rolling Deck to Repository program (R2R) is responsible for archiving routine underway data at the appropriate national archive, including the National Geophysical Data Center (NGDC) and the National Centers for Environmental Information (NCEI). Refer to the OCE Sample and Data Policy for information on other suggested databases and repositories for physical samples.

- [BCO-DMO "How-To" Pages](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)

## Roles and Responsibilities

Describe the roles and responsibilities of all parties with respect to the management of the data. Consider the following:

1. If there are multiple investigators involved, what are the data management responsibilities of each person
2. Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?

*Example Answer:*

Each PI will be responsible for sharing his/her subset of data among the project participants in a timely fashion. J. Doe will be responsible for collecting and analyzing the zooplankton sampling data. P. Smith will oversee the molecular biology work and will submit the resulting sequences to the National Center for Biotechnology Information's (NCBI) GenBank database. The Lead PI, R. Jones, will coordinate the overall data management and sharing process and will submit the project data, including GenBank accession numbers, and metadata to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) who will be responsible for forwarding these data and metadata to the appropriate national archive.

*Guidance:*

- [BCO-DMO "How-To" Pages](#)
- [NSF Division of Ocean Sciences Sample and Data Policy](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)