Division of Ocean Sciences Sample and Data Policy



National Science Foundation

May 2011

Table of Contents

I. Purpose	3
II. NSF Philosophy	3
III. OCE General Data Policy	3
IV. OCE Proposal Requirements	4
V. OCE Reporting Requirements	5
VI. More Specific Data Submission Guidance	5
VII. More Specific Sample Submission Guidance	6
Appendix I. National Data Centers	8
C. National Geophysical Data Center (NGDC) D. National Snow & Ice Data Center (NSIDC) E. Carbon Dioxide Information Analysis Center (CDIAC)	9 9
Appendix II: Program Specific Requirements A. Biological and Chemical Oceanography Data Management Office B. Marine Geology and Geophysics Data Management Office C. Ocean Drilling Program D. U.S. CLIVAR – Climate Variability and Predictability	10 10 11
Appendix III: Other Database Activities	14
Appendix IV. Sample Repositories	15

I. Purpose

This document provides updates and revisions to the Division of Ocean Science's (OCE) Data and Sample Policy to reflect changes in (1) NSF Data Management Plan requirements and (2) the location and contact information for designated Data and Sample Centers that have occurred since the last issuance of this policy in 2004.

II. NSF Philosophy

The overall NSF philosophy of data management and dissemination is embodied in the NSF Award and Administration Guide (AAG) Chapter VI.D.4:

Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. Privileged or confidential information should be released only in a form that protects the privacy of individuals and subjects involved. General adjustments and, where essential, exceptions to this sharing expectation may be specified by the funding NSF Program or Division/Office for a particular field or discipline to safeguard the rights of individuals and subjects, the validity of results, or the integrity of collections or to accommodate the legitimate interest of investigators. A grantee or investigator also may request a particular adjustment or exception from the cognizant NSF Program Officer. Investigators and grantees are encouraged to share software and inventions created under the grant or otherwise make them or their products widely available and usable.

III. OCE General Data Policy

In accord with NSF policy, Principal Investigators (PIs) working under OCE awards have additional conditions to which they need to adhere. These are specified below.

- A. PIs are required to submit, at no more than incremental cost and within a reasonable time frame (but *no later than two (2) years after the data are collected)*, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF/OCE grants to the appropriate Data Center (See appendices below or consult with the cognizant NSF Program Officer).
- B. Where no data or sample repository exists for collected data or samples, metadata must be prepared and made publicly available over the Internet and the PI is required to employ alternative strategies for complying with the general philosophy of sharing research products and data as described above. Where appropriate, samples should be curated in a manner that preserves the quality and integrity of the samples. When in

doubt about what to do, the PI should consult with the NSF Program Officer handling the award.

- C. All standard underway data collected at sea aboard NSF-supported oceanographic research vessels will be submitted to the appropriate long-term archive through the R2R (Rolling Deck to Repository) program. For these data sets, this relieves the PI and the ship operator of post-cruise data management responsibility for the underway datasets (See Section VI-B for more specific information). For other at-sea, special ops or research-specific datasets collected on a cruise (such as water samples, ROV sampling, dredge hauls, etc.) the PI is responsible for collecting and making public the metadata associated with sample collection and the data or results from the research.
- D. PIs and their institutions, and ship-operating institutions are responsible for meeting all legal requirements for submission of data and research results that are imposed by foreign governments as a condition of that government's granting research clearances. Each PI and institution must determine their legal obligations in this respect, with the assistance of the Department of State and NSF, as necessary.

IV. OCE Proposal Requirements

For each proposal submitted to NSF, the NSF Grant Proposal Guide requires that proposals include a supplementary document of no more than two (2) pages that is labeled "Data Management Plan".

For proposals submitted to OCE core programs (i.e., unsolicited proposals), this supplementary document should describe how the proposal will conform to the OCE policy on the dissemination and sharing of samples and research results. If the proposal is associated with a specific solicitation it should include a discussion of any additional solicitation-specific data management or reporting criteria. Examples of information that may be appropriate to include in a Data Management Plan may include:

- the types of data, samples, physical collections, software, derived models, curriculum materials, and other materials to be produced in the course of the project;
- the 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, where needed);
- mechanisms for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
- policies and provisions for re-use, re-distribution, and the production of derivatives; and

 plans for archiving data, samples, and other research products, and for preservation of access to them.

V. OCE Reporting Requirements

Annual Reports, required for all projects, should address progress on sample, data, and research product sharing. OCE requires that Final Reports document compliance or explain why it did not occur. In cases where the Final Report is due before the required date of sample or data submission, the PI must report submission of metadata and plans for final submission. The PI should notify the cognizant Program Officer by e-mail after final data and/or sample submission has occurred, even if this is after the expiration date of the award.

VI. More Specific Data Submission Guidance

A. For most ocean data there are designated National Data Centers where data must be deposited. These represent long term archives of data from NSF funded research. These Centers and a brief description of the data they support are described in Appendix I. Submission of data to alternate databases does not eliminate the requirement for final deposition of data in these National Data Centers. Some NSF-funded data management groups deposit data submitted to them into the National Data Centers. If this is the case, the PI does not need to submit an additional copy to the Center. Please check with the data managers to verify for which datasets this service is provided.

It should be noted some OCE core programs have specified data management groups to which most data funded by their program should be submitted (Chemical Oceanography, Biological Oceanography, and Marine Geology and Geophysics). These data management groups and associated specific requirements for data submission are shown in Appendix II.

B. All underway data collected at sea aboard NSF-supported oceanographic research vessels will be submitted to the appropriate long-term archive through the R2R (Rolling Deck to Repository) program. The R2R program maintains a central shore-side system through which underway data from oceanographic expeditions are routinely cataloged and securely transmitted to the national long-term archives including the National Geophysical Data Center (NGDC) and National Oceanographic Data Center (NODC). Data submission to R2R is accomplished by vessel operators rather than individual science parties, ensuring routine preservation and dissemination of underway data. Effective with the release of this policy, the underway data from NSF-funded cruises will be placed in the public domain for unrestricted open access after 60 days of the cruise end date, unless a request for restricted access is submitted through R2R. If such a request is submitted, a proprietary hold period of *up to two (2) years* will be maintained by R2R with approval of the cognizant NSF Program Officer. Requests for a longer

proprietary period will be maintained by R2R only with the approval of the cognizant NSF Program Officer. Data will be transmitted to the appropriate National Data Center for dissemination only after the conclusion of the stated proprietary hold. R2R does not provide data management services for datasets acquired with special operations or research-specific science party instrumentation such as instrumentation associated with the National Marine Seismic Facility, National Deep Submergence Facility, and the OBS Instrument Pool. Responsibility for these datasets lies with these facility operators and/or the science party. These data should be submitted to the appropriate National Data Center (Appendix I).

See http://www.rvdata.us/about/underway for the specific underway data maintained by the vessel operator in the R2R program.

- C. Focused community-driven science programs supported by OCE core programs may have alternate or more stringent data submission procedures to meet the needs of these programs. Pls funded by these programs should follow the data submission procedures specific to those research efforts. Requirements for these special programs are listed in Appendix II.
- D. For some special programs and focused community initiatives, alternative database activities exist. Some of these databases are listed in Appendix III. These databases may be funded by NSF, by other Federal Agencies, or by non-governmental organizations. Principal Investigators are encouraged to submit their data to these databases when appropriate. Since such databases may not provide long-term archival capabilities, such submission will satisfy the PI's obligations only if the database submits the data to one of the National Data Centers.
- E. Community standards for handling genomic data are still evolving. PIs who employ genomic techniques should articulate a strategy for providing timely community access to the data collected and for establishing links between genomic and environmental data. Sequence data should be submitted to a publicly accessible data repository (e.g., National Center for Biotechnology Information). The human genome community has recently articulated a philosophy of pre-publication access to sequence data, which is available at: http://www.genome.gov/page.cfm?pageID=10506537.

VII. More Specific Sample Submission Guidance

The infrastructure to support sample deposition and archiving varies among OCE programs.

A. Marine Geological Samples

PIs are required to archive and curate sediment, core, and dredge samples and to make them available to other investigators as soon as possible but *no later than two (2) years after the samples are collected*. NSF anticipates that most sediment, core, and dredge samples will be archived at NSF-approved repositories, examples of which are listed in Appendix IV. Principal Investigators may choose to archive their materials at their home or alternative institution, provided that the following conditions are met:

- Samples should be curated in an institutional facility that has a written and publicly discoverable sample distribution policy.
- Samples should be available to any US investigator upon their request.
- Metadata on samples, including where they are archived must be submitted to the appropriate National Data Center and to IEDA, the Marine Geology and Geophysics Data Management Group (see Appendix II), within 60 days post cruise.
- If samples are transferred to a new location for permanent archiving, the metadata at the National Data Center or IEDA, the Marine Geology and Geophysics Data Management Group (see Appendix II) must be updated when the transfer takes place.

The cognizant Marine Geology and Geophysics Program Officer must approve any exceptions to the above sample policy in writing.

B. Biological Samples

Academic, private, and community facilities have traditionally been sites where biological materials are curated. Not all material can (or should) be accommodated in these facilities. Pls should archive voucher and type specimens as dictated by community standards and practices, as required by journals for publication, and as appropriate to support research results. Sharing of valuable sample material is highly encouraged and can be facilitated by providing metadata, indicating that samples are available early in the development of a research program.

For further information on sample repositories, see Appendix IV.

Appendix I. National Data Centers

The final repositories for OCE funded data are the following designated National Data Centers. Data are to be submitted according to formats and via the media designated by the National Center.

A. National Oceanographic Data Center (NODC)

Ocean physical data: Temperature, salinity, light transmission or attenuation, currents, waves, pressure, sea level, and sound speed.

Ocean chemistry data: Nutrients such as phosphates, nitrates, nitrites and silicates; chemical tracers such as helium, tritium, freon and argon; pollutants such as petroleum hydrocarbons, organochloride and organophosphorus pesticides, polychlorinated biphenyls (PCBs) and heavy metals. Data may represent chemicals in water samples or biota.

Ocean biology data: Primary productivity; concentrations of pigments in phytoplankton, such as chlorophyll-a; species lists; biomass of phytoplankton, zooplankton, benthos and nekton; and bioluminescence.

National Oceanographic Data Center (NODC)
Data Officer, E/OC
1315 East-West Highway

ssmC3, 4 floor
Silver Spring, Maryland 20910
Phone: (301) 713-3267 x151

Fax: (301) 713-3301

http://www.nodc.noaa.gov/

B. National Climatic Data Center (NCDC)

Surface meteorological data: meteorological data in appropriate World Meteorological Organization formats as part of the Voluntary Observing Ship (VOS) program: air temperature, sea-surface temperature, dew point temperature, pressure, wind speed and direction, wind and swell waves, weather, short- and long-term radiation, visibility, cloud cover and type, and ice accretion.

National Climatic Data Center (NCDC) Federal Building 151 Patton Ave. Asheville, NC 28801-5001

Phone: (828) 271-4800 Fax: (828) 271-4876

http://lwf.ncdc.noaa.gov/oa/ncdc.html

C. National Geophysical Data Center (NGDC)

Geophysical, geological and geochemical data: bathymetry, magnetics, gravity, seismic and other quantitative geophysical data; geological data including station locations, collection/storage locations, preliminary descriptions of seafloor samples recovered, and all descriptions and analytical data, including geochemistry, derived from sediment and rock samples, including data from the Ocean Drilling Program (ODP).

National Geophysical Data Center (NGDC) NOAA, Code E/GC 325 Broadway Boulder, CO 80303-3328

Phone: (303) 497-6338 Fax: (303) 497-6513

http://www.ngdc.noaa.gov/ngdc.html

D. National Snow & Ice Data Center (NSIDC)

Sea ice and other glaciological data: sea ice, icebergs, ice shelves and associated physical oceanographic and meteorological data.

National Snow & Ice Data Center (NSIDC) CIRES Campus Box 449 University of Colorado Boulder, Colorado 80309 Phone: (303) 492-6199

Fax: (303) 492-2468

http://www-nsidc.colorado.edu/

E. Carbon Dioxide Information Analysis Center (CDIAC)

Carbon dioxide data: archival data for the World Ocean Circulation Experiment (WOCE) and the Joint Global Ocean Flux Study (JGOFS) CO2 measurements.

Carbon Dioxide Information Analysis Center (CDIAC) Oak Ridge National Laboratory

P.O. Box 2008 Oak Ridge, Tennessee 37831-6335 Phone: (423) 574-0390 Fax: (423) 574-2232

http://cdiac.esd.ornl.gov/home.html

Appendix II: Program Specific Requirements

NOTE: The addresses provided (as of May 2011) may change over time. Please check with the cognizant OCE Program Officer if necessary.

A. Biological and Chemical Oceanography Data Management Office

The Biological and Chemical Oceanography Data Management Office (BCO-DMO) serves PIs funded by the OCE Biological and Chemical Oceanography Programs and the NSF Office of Polar Programs (OPP), Division of Antarctic Sciences. BCO-DMO manages and serves oceanographic biogeochemical, ecological, and companion physical data and information developed in the course of scientific research and contributed by the originating investigators. BCO-DMO is a combination of the Data Management Offices formerly created to support the US JGOFS and US GLOBEC programs.

The BCO-DMO data system facilitates data stewardship, dissemination, and storage on short and intermediate time-frames. The office works with PIs on data quality control; maintains an inventory of available data and program thesaurus based on a controlled vocabulary; generates metadata records required by Federal agencies; ensures submission of data to national data centers; supports and encourages data synthesis by providing new, online, web-based display tools; facilitates interoperability between BCO-DMO and distributed data portals; and facilitates regional, national, and international data and information exchange.

The Biological and Chemical Oceanography Programs expect PIs to utilize BCO-DMO as their primary data management source. When awards are initialized investigators should immediately contact BCO-DMO and register their projects by submitting project metadata. This should be followed by timely submission of deployment and dataset metadata and finally the data. Progress on compliance with the data sharing should be addressed in annual and final reports. For projects where data cannot be served by BCO-DMO, or where it is more appropriately served by other community data bases, metadata should still be deposited in BCO-DMO with links to the other data bases.

Forms for contributing information to BCO-DMO can be downloaded at: http://www.bco-dmo.org/resources.

Biological and Chemical Oceanography Data Management Office (BCO-DMO) WHOI MS #36
Woods Hole, MA 02543
http://bco-dmo.org/

Contacts:

Cyndy Chandler (Marine Chemistry and Geochemistry)

e-mail: cchandler@whoi.edu

Robert Groman (Biology) e-mail: rgroman@whoi.edu

B. Marine Geology and Geophysics Data Management Office

The marine geology and geophysics data management office IEDA (Integrated Earth Data Applications) serves PIs funded by the Marine Geology and Geophysics (MGG) Program and some elements of the Office of Polar Programs (OPP), Division of Arctic Sciences and programs in the Earth Sciences Division at NSF. IEDA is a solid-earth science community data management facility that combines and serves solid earth scientists (marine geophysics, petrologic, sediment geochemical, vent fluid, geochronological, and other). It acts as a data repository and data discovery engine for much of the scientific research conducted by MGG-funded PIs. IEDA is a combination of the data repositories and data management/discovery tools developed by the Marine Geosciences Data System (MGDS) and Geoinformatics for Geochemistry (GfG) groups at the Lamont Doherty Earth Observatory of Columbia University.

The IEDA group facilitates data stewardship, dissemination, and storage on short and intermediate time-frames. The office works with PIs on data quality control; maintains an inventory of available data and program thesaurus based on a controlled vocabulary; generates metadata records required by Federal agencies; ensures submission of data to national data centers; supports and encourages data synthesis by providing new, online, web-based display tools; facilitates interoperability between IEDA and distributed data portals; and facilitates interoperability between IEDA and distributed data portals, and facilitates regional, national, and international data and information exchange.

For most cases, the Marine Geology and Geophysics Program expects PIs to utilize IEDA systems as their primary data management source. When awards are initialized investigators should immediately contact IEDA and register their projects by submitting project metadata. This should be followed by timely submission of deployment and dataset metadata and finally the data. Progress on compliance with the data sharing should be addressed in annual and final reports.

In cases where data are more appropriately archived through other community data bases, the Data Management Plan should specify where project data will be made available.

Forms for generating Data Management Plans that are compatible with the new NSF requirement can be found at: http://www.iedadata.org/compliance/plan. Information on data submission formats and conventions can be found on the IEDA website http://www.iedadata.org/ or by calling and consulting with the data managers (see contact information below).

Integrated Earth Data Applications (IEDA)

Lamont-Doherty Earth Observatory of Columbia University 61 Route 9W Palisades, NY, 10964 USA

Phone: +1 (845) 365-8506 Fax: +1 (845) 365-8156 http://www.iedadata.org

C. Ocean Drilling Program

The Ocean Drilling Program provides support for participation and drilling-related research performed by US scientists. This support focuses on the following areas: Investigations of potential drilling regions, especially by means of regional geological and geophysical field studies (site surveys), the feasibility and initial development of downhole instruments and techniques, downhole geophysical and geochemical experiments, and support for Expedition Objective Research (EOR) proposals to meet the scientific objectives of specific drilling expeditions.

Regional geological and geophysical data from investigations of potential drilling regions are a primary source of information in planning Integrated Ocean Drilling Program (IODP) drilling proposals. The Site Survey Data Bank (SSDB) at the Scripps Institution of Oceanography is the long-term repository for these geophysical data submitted in support of IODP drilling proposals and expeditions. Site survey data requirements for mature drilling proposals, submission deadlines and other details are posted on the IODP web site (http://ssdb.iodp.org).

Archiving requirements for geological and geophysical data collected from the other ODP-funded areas listed above (i.e., Expedition Objective Research and downhole geophysical and geochemical experiments) are described in Appendix I.C (National Data Centers – NGDC) of this document.

D. U.S. CLIVAR – Climate Variability and Predictability

All CLIVAR data shall be made available no later than two (2) years after collection, unless specifically waived by the international CLIVAR Scientific Steering Group (SSG). However, several CLIVAR activities, like the Carbon/Global Hydrographic Survey, require PIs to submit data collected to the CLIVAR and Carbon Hydrographic Data Office (CCHDO) at SIO within 6 months of collection. CCHDO is a Data Assembly Center (DAC) set up for the purposes of quality control and data synthesis that will submit the final data sets to the national archive (NODC) on behalf of the PIs. In general, the CLIVAR program requirements for data submission are similar to those found in WOCE Report No.104/93, WOCE Data Management. For more information contact:

CHDO

UCSD/Scripps Institution of Oceanography 9500 Gilman Drive 0214

La Jolla, CA 92093-0214 Phone: 858-822-1770 Fax: 858-534-7383

Email: cchdo@ucsd.edu
http://www.cchdo.ucsd.edu

or

Cathy Stephens U.S. CLIVAR Office 1717 Pennsylvania Avenue, NW Suite 250 Washington, DC 20006 Phone: (202) 419-3482

cstephens@usclivar.org http://www.usclivar.org/

Appendix III: Other Database Activities

A. Ocean Biogeographic Information System (OBIS)

The Ocean Biogeographic Information System (OBIS) is a community effort to develop a database of global marine animal and plant distributions. OBIS allows participants to develop distributed databases and serve their own data. This may be a particularly appropriate venue to provide certain classes of biological data (i.e., species level data) and PIs may benefit from data management tools being developed as part of this program.

Further information on OBIS is available at http://www.iobis.org/

B. The National Center for Biotechnology Information (NCBI)

The National Center for Biotechnology Information (NCBI) is a facility of the National Institutes of Health (NIH) that hosts GenBank and other molecular databases. Additional information is available at: http://www.ncbi.nlm.nih.gov/.

Appendix IV. Sample Repositories

A. Institutional Repositories for Sediment and Rock Samples

Sediment cores and dredged rocks from the seafloor are collected at great cost and are often of benefit to the research community beyond the projects for which they were originally collected. In recognition of the value and use of these samples, the Marine Geosciences Section (MGS) provides partial support for a limited number of institutional repositories. These repositories ensure that samples of sediment and rock collected from the seafloor are properly curated, preserved, and disseminated to qualified researchers. If a PI places samples in any of these or other NSF-approved repositories, they are still responsible for providing metadata for samples to the appropriate National Data Center and, if applicable, to the Marine Geology and Geophysics funded Data Management Group, IEDA, (see Appendix II). The following are some samples of NSF-approved institutional marine sample repositories:

- Lamont-Doherty Earth Observatory <u>http://www.ldeo.columbia.edu/CORE_REPOSITORY/RHP</u>1.html
- Oregon State University http://corelab-www.oce.orst.edu
- Scripps Institution of Oceanography http://gc.ucsd.edu
- University of Rhode Island <u>http://www.gso.uri.edu/MGSLsite/mgsl_homepage.htm</u>
- Woods Hole Oceanographic Institution http://www.whoi.edu/corelab/

The Marine Geosciences Section also provides support for maintenance and access to cores collected by the Deep Sea Drilling Project (DSDP), the Ocean Drilling Program (ODP) and the Integrated Ocean Drilling Program.

See http://www.iodp.org/repositories/2/ for more details about IODP repositories. Locations presently supported by NSF are:

I. The Bremen Core Repository (BCR)

Located at Bremen University in Bremen, Germany, stores DSDP, ODP and IODP cores collected in the Arctic Ocean (north of the Bering Strait), the Atlantic Ocean, and the Mediterranean and Black Seas.

BCR Superintendent
Marum – Bremen University

Leobener Str D-28359 Bremen Germany

Tel: +49 421 218-65566 Fax +49 421 218-9865566

http://www.marum.de/en/IODP Core Repository.html

II. The Gulf Coast Repository (GCR), located at Texas A&M University in College Station, Texas, U.S.A., stores DSDP, ODP and IODP cores collected the Pacific Ocean (Pacific plate east of the western boundary), the Caribbean Sea and Gulf of Mexico, and the Southern Ocean (south of 60° except the Kerguelan Plateau).

GCR Superintendent Gulf Coast Repository 1000 Discovery Drive Texas A&M University College Station, TX 77845

Tel: 979-845-8490 Fax: 979-845-1303

http://iodp.tamu.edu/curation/gcr/

III. The Kochi Core Center (KCC), located at Kochi University, Kochi, Japan, stores cores from the Pacific Ocean (west of western boundary of the Pacific Plate), the Indian Ocean (north of 60°S), all of the Kerguelan Plateau, and the Bering Sea.

Kochi Core Center 200 Monobe Otsu Nankoku, Kochi 783-0802 Japan

Fax: 088-86400040

http://www.kochi-core.jp/en/iodp-curation/index.html

IV. The Rutgers/New Jersey Geological Survey, located at Rutgers University, Piscataway, New Jersey, USA stores land-based New Jersey and Delaware cores associated with ODP Legs 150X and 174AX.

Curator
Earth & Planetary Sciences
Rutgers University
Wright-Rieman Labs
610 Taylor Road

Piscataway, NJ 08854 Phone: (732) 445-2044 FAX: (732) 445-337

B. Living Culture Facilities

The Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP) accepts strains for deposition if the Director or Curator agrees that the strains are a valuable addition to the collection. Examples include strains that are referred to in publications, contain interesting molecular, biochemical or physiological properties, are the basis for taxonomic descriptions, are important for aquaculture, or are from an unusual geographical location or ecological habitat. Strains should be identified to the species level, and they should be unialgal, preferably clonal (if possible) and axenic. Depositions should include information regarding collection (site, environmental/ecological data, date and collector), identification (who identified the strain), known properties (e.g., toxicity, bioluminescence, pigments), culture medium (media) in which it grows, cell size range, temperature range for growth, and other information that may be of value to others. Contact information for the CCMP is provided below:

Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP)
Bigelow Laboratory for Ocean Sciences
P.O. Box 475
West Boothbay Harbor, Maine 04575
https://ccmp.bigelow.org/