# Division of Ocean Sciences Data and Sample Policy


## National Science Foundation

**NSF 04-004**

# Table of Contents

## I. Purpose

A. The ocean community is experiencing an unprecedented increase in the types and amount of data becoming available and as ocean science has evolved to larger, multi-investigator, interdisciplinary projects, availability and sharing of data and samples in a timely fashion has become increasingly important.

B. This statement updates and revises guidelines to implement Federal data policy by assuring timely submission of quality controlled oceanographic data to the national data centers for secondary use. Guidelines for oceanographic data were first issued by the National Science Foundation's (NSF) Division of Ocean Sciences (OCE) in October 1988 and revised in 1994.

## II. Philosophy

The overall philosophy is embodied in the NSF Grant General Conditions:

> NSF expects significant findings from research and education activities it supports to be promptly submitted for publication, with authorship that accurately reflects the contributions of those involved. It expects investigators to share with other researchers, at no more than incremental costs and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work. It also encourages grantees to share software and inventions or otherwise act to make the innovations they embody widely useful and usable.
>
> Adjustments and, where essential, exceptions may be allowed to safeguard the rights of individuals and subjects, the validity of results, or the integrity of collections or to accommodate legitimate interests of investigators.

## III. General Data Policy

A. Principal Investigators are required to submit all environmental data collected to the designated National Data Centers (Appendix I) as soon as possible, but no later than two (2) years after the data are collected. Inventories (metadata) of all marine environmental data collected should be submitted to the designated National Data Centers within sixty (60) days after the observational period/cruise. For continuing observations, data inventories should be submitted periodically if there is a significant change in location, type or frequency of such observations.

B. Principal Investigators and their institutions, and ship-operating institutions are also 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 principal investigator and institution must determine their legal obligations in this respect, with the assistance of the Department of State and NSF, as necessary.

C. Where no data or sample repository exists for the collected data or samples, metadata must be prepared and made available. The Principal Investigator (PI) is required to address alternative strategies for complying with the general philosophy of sharing research products and data as described above. This must be included in the proposal Project Description. Samples should be curated in a manner that preserves the quality of the samples. The PI is invited to discuss this issue with NSF Program Officers in advance of submitting proposals.

## IV. Proposal Requirements

The NSF Grant Proposal Guide requires that proposal Project Descriptions outline plans for preservation, documentation, and sharing of data, samples, physical collections, curriculum materials and other related research and education products. Plans for the handling of data and other products will be considered in the review process.

## V. Reporting Requirements

Annual reports, required for all projects, should address progress on data and research product sharing. The Division of Ocean Sciences requires that final reports document compliance or explain why it did not occur. In cases where the final report is due before the required data or sample 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.

## VI. More Specific Guidance

A. As discussed in the General Data Policy, for most ocean data there are designated National Data Centers where data must be deposited. 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.

B. Focused programs supported by NSF's Division of Ocean Sciences (OCE) may establish (in consultation with OCE Program Officers) more stringent data submission procedures to meet the needs of such programs. Principal Investigators supported by these programs are required to follow these data submission procedures. Examples of these Program Specific Requirements are listed in Appendix II.

C. 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 Principal Investigator's obligations only if the database submits the data to one of the National Data Centers.

D. Community standards for handling genomic data are still evolving. Principal Investigators 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. Sample Policy

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

A. Marine Geological Samples

Principal Investigators 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-supported repositories 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 must be curated in an institutional facility that has a written and NSF-approved sample distribution policy.
- The facility must be open to all US investigators and metadata on samples must be available through an electronic data base, preferably web based.
- Metadata on samples, including where they are archived must be submitted to the appropriate National Data Center within 60 days post cruise.
- If samples are transferred to a new location for permanent archiving, the metadata at the National Data Center must be updated when the transfer takes place.

NSF supported sample repositories are listed in Appendix IV. The cognizant NSF Program Officer must approve any exceptions to this policy or the two-year exclusive use period 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. PIs 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 environmental 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, 4th 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 April 2003) may change. Please check with relevant Program Officers of the NSF Division of Ocean Sciences if necessary.

### A. U.S. CLIVAR

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 Global Hydrographic Survey, require Principal Investigators to submit data collected to a Data Assembly Center (DAC) for the purposes of quality control and data synthesis within shorter time periods. 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:

Dr. David Legler
U.S. CLIVAR Office
1717 Pennsylvania Avenue, NW
Suite 250
Washington, DC 20006
Phone: (202) 223-6262
Fax: (202) 232-3065
http://www.usclivar.org/

### B. U.S. GLOBEC

In addition to the data submission requirements mentioned in this document, the U.S. GLOBEC Scientific Steering Committee (SSC) requires all Principal Investigators to submit plans for the collection of data to the U.S. GLOBEC Data Management Office (DMO) at least ***three (3) months*** prior to execution of a sampling program. Specifics to be included in the data collection plan are detailed in U.S. GLOBEC Data Policy, Report Number 10, February 1994, available from:

U.S. GLOBEC National Coordinating Office
P.O. Box 1459
Leonardtown, MD 20650
Phone: (301) 997-0853
Fax: (301) 997-0854
http://www.usglobec.org/

Principal Investigators are responsible for documenting measurement and analysis techniques used to produce data sets and estimating accuracy and precision of these measurements. Specific physical measurements must be acquired along with all biological measurements and must meet pre-defined standards (see Report No. 10). In addition, the report specifies requirements for preservation of biological samples, including for the purpose of subsequent genetic analysis.

Data from measurements which do not involve manual analysis and which would be useful to the scientific community must be submitted by the principal investigator to the DMO within ***six (6) months*** after collection. All other measurements and any standard analyses of these measurements must be available to the community within ***one (1) year*** after collection. PIs will submit data either directly to the DMO or by placing it on-line as a U.S. GLOBEC distributed database. Format standards for submission of data and development of the database will be specified by the DMO. The DMO will serve as an intermediate archival location and data source and will transfer data to the NODC and prepare necessary documentation for data collected in foreign waters.

## C. U.S. JGOFS

U.S. JGOFS chief scientists are required to submit all data to the Data Management Office (DMO) within **one (1) year** after the sampling date. However, data derived from long analytical procedures (e.g. $^{228}Ra$) which prevent the researcher from being able to readily analyze/publish can be exempted from this ***one (1) year*** requirement. In addition, final versions of Basic Core Measurements (i.e. temperature, salinity, dissolved oxygen) must be received by the DMO within ***six (6) months*** after the sampling date. Again, some exceptions can be made for data requiring extensive analyses. However, all principal investigators making core measurements are urged to make their data available as quickly as possible. When submitting data to the JGOFS Distributed Data Management System (DDMS), Program principal investigators have two options: 1) to store data locally serving as a host node on the DDMS, or 2) to submit data to the U.S. JGOFS Data Management Office (DMO) and they will serve the data. $CO_2$ measurements should be submitted to the WOCE World Hydrographic Programme (WHP). More detailed information on the U.S. JGOFS requirements for data submission is available from:

U.S. JGOFS Planning and Data Management Office
GEOSECS Building
MS 43
Woods Hole Oceanographic Institution
Woods Hole, MA 02543-1535
Phone: (508) 289-2497
Fax: (508) 457-2161
http://www1.whoi.edu/jgofs.html

## D. Ocean Drilling Program

The Ocean Drilling Program supports regional geological and geophysical field studies which can be used to develop mature drilling proposals in the Joint Oceanographic Institutions for Deep Earth Sampling (JOIDES) system. The geological and geophysical data from these projects are a primary source of information in planning drilling and should be available for review by the Site Survey and Pollution Prevention and Safety panels of JOIDES. Site survey data requirements for mature drilling proposals are identified in the JOIDES Journal issue titled, "Guide to the Ocean Drilling Program." Additionally, such data can be important in interpreting the results of a drilling leg and should be available to cruise participants.

Successful applicants are expected to deposit data from their cruises in the Ocean Drilling Program Site Survey Data Bank at Lamont-Doherty Earth Observatory, in addition to other data archiving requirements described in this document (see Appendix I.C.). The address is the following:

ODP Site Survey Data Bank
Lamont-Doherty Earth Observatory
Palisades, New York 10964
Phone: (845) 365-8542
Fax: (845) 365-8159
Email: odp@ldeo.columbia.edu
http://www.ldeo.Columbia.edu/databank/

At the earliest possible date, the chairperson of the JOIDES Site Survey Panel, the manager of the Data Bank, and the representative of the appropriate national data center should be notified of the data types and schedule for submission.

The Ocean Drilling Program also supports more limited data collection activities through the U.S. Science Support Program administered by the Joint Oceanographic Institutions (JOI). Data reporting requirements under this program are the same as those identified above.

## E. MARGINS

An important element of the MARGINS Program is that all data and results be rapidly shared in order to encourage integration of science, coordination of research, and the construction and testing of hypotheses. All data collected with MARGINS funding must be archived as soon as practically possible, along with all relevant metadata, in the institutional archives that are standard for a particular discipline (e.g., Incorporated Research Institutions for Seismology (IRIS) for seismological data, University NAVSTAR Consortium (UNAVCO) for Global Positioning System (GPS) data, Core repositories for marine geological samples, and NGDC for marine geological and geophysical data per Appendix I.C). Data for which no standard archive exists (e.g., Multi Channel Seismic (MCS) data, swath data and land geological samples) must be archived by the Principal Investigator and made available (with the cost of copying paid by the recipient) to researchers upon request.

Basic metadata (e.g., data types, sample locations, cruise tracklines, etc.) must be provided to the MARGINS Office within ***sixty (60) days*** of ending a field program. In due course and in collaboration with ongoing efforts in the Marine Geology and Geophysics Program, the MARGINS Office is currently developing tools for preparing and formatting these metadata files. It is the responsibility of the Principal Investigator to provide to the MARGINS Office, for publication on the MARGINS Office web site, details of and links to all datasets acquired or generated with MARGINS funding. Contact information for the MARGINS office is provided below.

MARGINS Office
Lamont-Doherty Earth Observatory
P.O. Box 1000, 61 Route 9W
Palisades, New York, 10964 USA

Phone: (845) 365-8665
Fax: (845) 365-8156
http://www.ldeo.columbia.edu/margins/Home.html

All raw data must be made freely available ***two (2) years*** after ending a field program, consistent with the data release policies of the NSF Division of Ocean Sciences and other national and program centers (IRIS, Program for Array Seismic Studies of Continental Lithosphere (PASSCAL), and US National Ocean Bottom Seismography Instrument Pool (OBSIP)). In the case of datasets that are not available to the investigators at completion of the field-season/cruise, for example, because they are assembled by the relevant data-center before distribution, the two year moratorium period begins on the date that the complete dataset is made available to investigators. However, Principal Investigators are encouraged to release data to other Focus Site investigators as soon as possible following the end of a field season or completion of dataset processing.

Processed, derived and interpreted datasets must be made publicly available as soon as possible, certainly within the lifetime of the grant. This policy applies even to those data and results that Principal Investigators have traditionally not been required to make publicly available (e.g. stacked and migrated seismic sections, geochemical analyses, Digital Elevation Models (DEMs) and other rasters, geological samples and geochemical analyses).

## F. Ridge 2000

1. Introduction

The data management strategy for the Ridge 2000 (R2K) Program is designed to address the needs of the program, individual R2K investigators, and the larger scientific community. Central to this strategy is timely submission and sharing of all metadata and data collected in both Integrated Study Site (ISS) field programs and Time Critical Study (TCS) rapid response cruises as well as sharing of all relevant historical data. Rapid dissemination of the metadata and data will maximize information transfer across the program, facilitate proposal preparation by investigators new to the program, and encourage integration of science, coordination of research, and the construction and testing of hypotheses. In keeping with this philosophy, all data used in R2K proposals should be in the public domain, or at least metadata identifying the location, data types, and contact person should be in the public domain at least 30 days before a grant proposal is submitted. R2K is a time limited program, thus all data collected should be rapidly released for maximum benefit to all. A strong commitment to data management is required of each participating PI. In requesting and accepting NSF support within the NSF-R2K program, each PI is obligated to meet the data management and disclosure requirements as an integral aspect of their participation in the program.

To facilitate data management, a data management system (DMS) will be implemented, maintained and operated by a data management office (DMO). The field data from R2K Time Critical Studies, which are currently limited to the Northeast Pacific, will be included in the Endeavour ISS database. The mission of the DMO will be to ensure that all R2K data sets are readily accessible by all R2K investigators on a common time base and within a common spatial framework.

While recognizing the legitimate rights of data originators and collaborating PIs to the first use of the data they collect, the R2K Program encourages the oceanographic community to use data collected by the program, and in particular, believes that data availability should be restricted only in exceptional cases.

2. Data Policy

The R2K Data Management Policy is predicated on guidelines that encourage openness and sharing of data for the mutual benefit of the scientific community. This policy sets responsibilities for release of data with the understanding that some measurements will require long analytical or data reduction procedures that prevent early release after collection.

All data sets must contain a uniform suite of mandatory metadata that conforms with the policies to be developed for the R2K DMS. It is likely that the minimum requirements for each station or observation will include: cruise ID, time and date (UTC), position (lat/long and if available, xy coordinates with system origin), and event/operation number. For sub-samples from a bottle or other bulk sample, each data record must contain: cruise ID, event, dive or cast number, and sample or bottle number. For each data set, the metadata should include: descriptions of standards used for measurement of time and position, shipboard sampling procedures, sample treatment and preparation, analytical procedures, equipment calibrations, data reduction techniques, computation algorithms, analyses of standards or other data suitable for quality control and inter-laboratory comparison, citations, and any other useful information.

It is essential that PIs use standard digital forms to submit metadata to the DMO at the conclusion of each field program to facilitate effective and efficient use. Several levels of metadata exist, each defining a particular stage in the data acquisition to publication process. The levels are defined as follows: Level 1 - Basic description of the field program including: cruise ID and dates, participating scientists, operation logs, navigation files and corrections, data types, and available underway data. Level 2 - A final cruise report with complete data inventory in R2K standard format. Level 3 - Data access information including: data formats, data quality assessments, details of processing procedures, and information on ongoing data processing and experimental studies. Level 4 - Models and publications derived from the data.

A suite of basic environmental data is essential to enable interpretation of many data sets in the context of the ISS. Basic field data include tide data, pressure sensor data, current meter data, bathymetry, vent field maps, and CTD (conductivity, temperature, depth) or comparable data on water column temperature and chemistry. All basic environmental data and metadata should be submitted to the DMO
for inclusion in the DMS ***within 6 months of collection***. ***PIs may place reasonable, time-limited restrictions on data use (less than two years)***. In some cases, it may be appropriate to provide metadata that describe derived data or analyses that are currently in progress. It is essential that all investigators using data from the DMS cite the originators of the data, even if no restrictions apply to its use.

All other data should be submitted to the DMO for inclusion in the DMS ***within 12 months of data acquisition***. Data sets and collections that require lengthy analytical and/or processing procedures should be submitted as they are completed. In these cases metadata describing the work in progress are expected to be included in the DMS. For laboratory or theoretical

studies, (meta)data to be submitted to the DMS include procedures, techniques, model parameters and computer codes. Historical data that would increase the value of the DMS should also be submitted promptly.

3. Responsibilities of Principal Investigators and Chief Scientists

The principles outlined above impose a series of responsibilities on Principal Investigators, Chief Scientists and the Data Management Office (DMO). Chief Scientists, in particular, have an ongoing responsibility to ensure that data are submitted and updated in timely and user-friendly fashion.

a. The Chief Scientist of each R2K cruise must submit Level 1 metadata as soon as the field program is complete. The Chief Scientist should ensure that a uniform, detailed operations log records at least the following information for every sampling operation: dive/operation number, station number, date, time, position, sampling device, and other comments. Standard digital forms will be provided by the DMO.
b. Digital cruise reports in DMO-standardized format, including the detailed operations log and cross-referenced detailed sample inventories, will be submitted to the DMO within 60 days of the end of the cruise.
c. Consistent with shipboard processing capabilities, basic data (e.g. bathymetric maps) should be available in preliminary form at the end of each cruise. The Chief Scientists should distribute these data, labeled as preliminary, to the DMO at the end of the cruise.
d. Final versions of basic environmental data should be submitted to the DMO for inclusion in the DMS as soon as possible and no later than 6 months after sample collection or instrument retrieval. Where this is not possible due to the nature of the analytical or data reduction process, Level 3 metadata indicating the existence and status of data-in-progress must be submitted and updated every 6 months. Data being used for Masters or PhD theses should be identified within the metadata and investigators wishing to use such data should first discuss their use with the PI.
e. Within one year of each cruise, PIs must submit all available data to the DMO, accompanied by Level 3 metadata. For data-in-progress, metadata indicating the existence and status of data-in-progress must be submitted and subsequently updated every 6 months. PIs making delayed measurements should strive to meet a timely release date. Unless authorized for early release by the responsible PI, all data will be on "restricted release" until 2 years post-cruise, after which time they will be freely available. Requests for data on restricted release will be referred by the DMO to the responsible PI.
f. Principal Investigators are responsible for the quality and correctness of data submitted to the DMS and should interact with the DMO to ensure that: (1) data comply with R2K DMS standards; (2) data subject to revision are updated promptly in the DMS; and (3) queries and criticisms from other users are promptly resolved.

4. Responsibilities of the Data Management Office

a. The DMO will provide a secure, web-based data retrieval system. The DMO will catalog submitted data and documentation such that they can be retrieved using criteria such as time, location, keyword, and/or sample identifier. Moreover, with input from the

community, the DMO will define the data formats to be used for all types of data and provide PIs with digital forms on which to record their Level 1 and 2 metadata.

b. While PIs have primary responsibility for data quality, the DMO will provide basic assessment of all data for compliance with R2K DMS standards. The DMO will notify investigators of problems identified in their data sets by the DMO or by other users and work with investigators to resolve such problems. The DMS will be a circular system that responds to feedback from users and providers of data and metadata.
c. The DMO will ensure that Level 1 and 2 data and metadata are compiled and submitted to appropriate national data repositories in a timely fashion following a cruise.
d. The DMO will release all data to the public domain two years after sample collection or instrument retrieval. Where appropriate, the DMO will ensure that R2K metadata and data sets are transferred to NODC and NGDC or other national databases. This release/submission will fulfill the obligation of the PIs as defined in the OCE data policy, but will not shift responsibility from the PI.
e. The DMO will liaise with PIs, the ISS coordinators, the R2K database Working Group and the R2K Office to encourage and evaluate community feedback, to ensure that community needs are being met and to ensure that all levels of metadata are available in the appropriate time frame.

## Appendix III: Other Database Activities

### A. Ocean Biogeographic Information System (OBIS)

The Ocean Biogeographic Information System (OBIS) is a nascent 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 and PIs may benefit from data management tools being developed as part of this program. This may also provide an alternative way for PIs to conveniently comply with data sharing requirements, particularly for data that is not now being accepted by NODC or other National Data Centers.

Further information on OBIS is available at http://marine.rutgers.edu/OBIS/.

### 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/.

### C. RIDGE Multibeam Synthesis

The RIDGE Multibeam Synthesis accepts multibeam bathymetry and amplitude data from the world's ocean basins. Although mid-ocean ridges are the focus of the synthesis, datasets from other settings will be incorporated. The synthesis also accepts side-scan sonar data and derived data products such as magnetic and gravity anomaly grids. Data submitted prior to the two-year proprietary hold period are incorporated into the database with password-only access.

Access and information are available at http://ocean-ridge.ldeo.columbia.edu/. Investigators wishing to contribute data should contact S. Carbotte, W.B.F. Ryan, S. O'Hara or B. Arko at Lamont-Doherty Earth Observatory (LDEO) for information and assistance on how to submit data.

### D. The Petrological Database of the Ocean Floor (PetDB)

PetDB is a web-based data management system for geochemical and petrological sample-based data. In its current application, PetDB contains and provides on-line access to data for sub-oceanic igneous and metamorphic rocks generated at mid-ocean ridges including back-arc basins, young seamounts, and old oceanic crust.

The chemical data that are currently stored in the database comprise major oxides, trace elements, radiogenic and stable isotopes, and analytical age determinations measured on whole rock samples, volcanic glasses, minerals, and melt inclusions. A wide range of supplementary information (metadata) describes samples (e.g. by rock type, texture, age, modal composition, alteration), sample locations (e.g. by geographical coordinates, location names, tectonic setting),

sampling process (e.g. by technique, date, cruise), archive, analytical procedures (e.g. by method, precision, standard measurements), and the source of the data (reference, author(s)). These metadata are not only essential for selecting, sorting, and evaluating data properly, but they are fundamental for current and future integration with other data types, especially geospatial data, and interoperability with other databases.

PetDB's web interface allows users to select data of their choice using a large variety of search criteria such as location, rock type, and reference. Users can view these data and download them in customized data tables. Summaries of all published data for individual samples can be obtained as all samples in the database can be uniquely identified.

Data entry is currently based on flat file forms (spreadsheets) from which information is uploaded into the database. The use of these forms is not required for data submission. For data to be included in PetDB, a set of mandatory metadata such as geographical coordinates for the sample location, rock type, sampling cruise or date and PI (essential to assign unique sample identifiers), units for the analytical values, and analytical method need to be provided. Complete sets of other metadata are strongly encouraged.

Contact information:
Home page: http://petdb.ldeo.columbia.edu/petdb
Technical support, error report: lehnert@ldeo.columbia.edu
petdb@ciesin.columbia.edu

Data submission: lehnert@ldeo.columbia.edu
General information: lehnert@ldeo.columbia.edu
langmuir@eps.harvard.edu

Kerstin Lehnert
Lamont-Doherty Earth Observatory
Office of the Director, Monell Building
61 Route 9W
Palisades, NY 10964
Phone: (845) 365-8506, Fax: (845) 365-8162

## Appendix IV. Sample Repositories

### A. Institutional repositories for sediment and rock samples

Sediment cores and dredged rocks 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 institutional collections, the Marine Geosciences Section (MGS) provides partial support for the curatorial requirements. It is the responsibility of the sample repositories to ensure that samples of sediment and rock collected from the sea floor are properly curated, preserved, and disseminated to qualified researchers. Support for these collections is evaluated on a regular basis. In addition to standard NSF criteria, proposals are reviewed for the extent and costs of services provided by the repository to the overall community, the usage of material in the repository, the availability of information about the materials and services and the procedures and policies for access to the materials. Only basic curatorial functions are supported. Descriptions of samples must be provided to the National Geophysical Data Center for use in the "Index to Marine Geology Samples" to assist potential users. MGS anticipates continued support of repositories that demonstrate cost-effective service to the marine geoscience community.

MGS provides partial support to the following institutional repositories:

Lamont-Doherty Earth Observatory
www.ldeo.columbia.edu/CORE_REPOSITORY/RHP1.html

Oregon State University
http://corelab-www.oce.orst.edu

Scripps Institution of Oceanography
http://gc.ucsd.edu

University of Rhode Island
http://www.gso.uri.edu/MGSLsite/mgsl_homepage.htm

Woods Hole Oceanographic Institution
http://www.whoi.edu/corelab/

MGS also provides support for maintenance and access to cores collected by the Deep Sea Drilling Project and the Ocean Drilling Program at the following locations (for more information including links to the repository web sites, see http://www-odp.tamu.edu/curation/repos.htm ):

The Bremen Core Repository (BCR), located at Bremen University in Bremen, Germany, stores all cores from the Atlantic and Southern Oceans, and the Black, Caribbean and Mediterranean Seas beginning with Leg 151 (since Sept. 1993).
To contact the staff of this repository:
**E-mail** bcr@odpemail.tamu.edu
**Telephone:** 49-421-396-6336, **fax:** 49-421-396-6684; or
**Write to:** Walter Hale

BCR Superintendent
Universitat Bremen
Ocean Drilling Program (BCR)
Konsul-Smidt Str 30,Schuppen 3
D-28217 Bremen, Germany

The East Coast Repository (ECR), located at the Lamont-Doherty Earth Observatory in Palisades, New York, U.S.A., houses the cores collected by both the ODP and the Deep Sea Drilling Project (DSDP) from the Atlantic and Southern Oceans and the Black, Caribbean, and Mediterranean Seas.
To contact the staff of this repository:
**E-mail** ecr@odpemail.tamu.edu
**Telephone:** 845-365-8446, **fax:** 845-365-8178, or
**Write to:** Gar Esmay
ECR Superintendent
Ocean Drilling Program
East Coast Repository
Geosciences Building, Room 104
Lamont-Doherty Earth Observatory
Palisades, NY 10964, U.S.A.

The Gulf Coast Repository (GCR), located at Texas A&M University in College Station, Texas, U.S.A., houses the cores collected by the ODP from the Pacific and Indian Oceans since 1984.
To contact the staff of this repository:
**E-mail** gcr@odpemail.tamu.edu ;
**Telephone:** 979-845-5056, **fax:** 979-845-0876; or
**Write to:** Phil Rumford
GCR Superintendent
Ocean Drilling Program
1000 Discovery Dr.
College Station, TX 77845-3469, USA.

The West Coast Repository (WCR), located at Scripps Institution of Oceanography in La Jolla, California, USA, houses the cores collected by the DSDP from the Pacific and Indian Oceans and the Red Sea.
To contact the staff of this repository:
**E-mail** wcr@odpemail.tamu.edu
**Telephone:** 858-534-1657, **fax:** 858-534-4555; or
**Write to:** Jerry Bode
WCR Superintendent
Ocean Drilling Program
West Coast Repository
Scripps Institution of Oceanography
La Jolla, CA 92093-0231, USA

## B. Living Culture Facilities

General Information: Deposition of Cultures

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
http://b250.bigelow.org/
http://B250.bigelow.org/GI/GI_00.html