

MARGINS DATA POLICY

An important element of the MARGINS Program is that all data and results be rapidly shared in order to maximize technology transfer across the program, encourage integration of science, coordination of research, and the construction and testing of hypotheses. MARGINS funding carries with it a higher responsibility for data sharing, archiving and dissemination because MARGINS is a time-limited program. The consequence of this time limit is clear — in order to encourage the integration of science within and across the MARGINS initiatives, MARGINS data need to be rapidly released and available to researchers. In accepting NSF support from the NSF-MARGINS program, each Principal Investigator is obligated to meet the following data-management requirements:

- All data collected with MARGINS funding must be archived as soon as practically possible and certainly within the life-time of the grant, along with all relevant metadata, in the institutional archives that are standard for a particular discipline (e.g., IRIS for seismological data, UNAVCO for GPS data, Core repositories for marine geological samples, and NGDC for marine geological and geophysical data per NSF-OCE data guidelines). Data for which no standard archive exists (e.g., MCS, swath data and land geological samples) must be archived by the Principal Investigator and made available (at cost to 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 60 days of ending a field program. In due course and in collaboration with ongoing efforts in MG&G, the MARGINS Office is currently developing tools for preparing and formatting these metadata files.
- All raw data must be made freely available two years after ending a field program, consistent with the data release policies of IRIS/PASSCAL and 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, and preferably within a year, 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 life-time 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, DEMs and other rasters, geological samples and geochemical analyses).

An appropriate plan for data archiving and dissemination is an important element in consideration of proposals for funding. A data archive and dissemination plan is a necessary component of NSF proposals. Investigators who cannot meet the conditions for rapid and complete release of MARGINS data and results should state the justification in their proposed data plan.

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. NSF Program Managers may seek information from the MARGINS Office about compliance with the MARGINS data policy when considering future funding decisions. Proposals funded through the NSF MARGINS program after November 1, 2001, need comply to the letter of this policy. All proposals funded through the NSF MARGINS program prior to November 1, 2001 are required to make field program metadata and data available to investigators as soon as is practically possible.



MARGINS: <http://www.margins.wustl.edu/margins>
National Science Foundation: <http://www.nsf.gov>