

Data Management Plan

1) Products of the Research: The primary data products for this project will include the evaluation of fluid structures in a custom designed turbulence apparatus and the behaviors of copepods in response to these signals. *Digital video and photomicrographs* resulting from this work will serve as a permanent visual record of the zooplankton used in our experiments and their behavioral responses to characterized fluid structures.

2) Standards to be used for data and metadata formatting and content: As samples are processed and data are generated they will be loaded into spreadsheets; tabular data and metadata will be kept primarily on Excel spread sheets (for all tabular data); SEM and light microscope images will be stored as Tif files; video files will be saved as uncompressed avi files. Spreadsheets are backed up in near real time using TimeMachine or a similar program. *Lab notebooks* containing the raw data generated by this project, detailed descriptions of procedures and methodological approaches, deviations from protocols, specific equipment, and chemical reagents utilized for this project will be cataloged by all project participants. Notebooks and digital copies thereof will serve as permanent records of the project and will be available upon request from NSF program managers.

3) Mechanisms for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements: For sharing files between the laboratories we will use in-house storage facilities which are password protected. Distribution of data to BCO-DMO will be done within two years of their collection via password-secure FTP. All data generated as a result of research performed at the Laboratory (or Institute), or with Laboratory (or Institute) funds, shall be the intellectual property of the Laboratory (or Institute) and the Senior Research Scientists overseeing the research. As such, the Laboratory (or Institute) will retain an implied copyright for these data.

4) Policies and provisions for re-use, re-distribution, and the production of derivatives: Data from these laboratory experiments will be submitted to BCO-DMO where they will be publically available. Peer-reviewed papers will be written by Fields, Yen and Webster which present these experimental data, along with details of their collection and data interpretation.

5) Plans for archiving data, samples, and other research products, and for preservation of access to them: Bigelow has a dozen large freezers, which are used to archive physical samples after initial analysis until the complete dataset is assembled. These freezers are attached to dedicated plugs in line with a backup generator for security. Samples will be archived at -80 °C and divided among two freezers for safe-keeping at Bigelow. We are presently working to implement a lab-wide digital tracking program for logging samples collected during research programs.

We will work with staff at the Biological and Chemical Oceanography Data Management Office (BCO-DMO) to insure that all of our data are archived. Additionally, our data sets will be available online from the BCO-DMO data system (<http://bco-dmo.org/data/>) where they will be managed. We will work with BCO-DMO personnel to ensure that all data components are linked among the different databases. Images will be submitted to Microbe Library <http://www.microbelibrary.org/>, a site created and hosted by ASM to generate a database of microbial images for use in undergraduate education and will also be submitted to Micro*Scope <http://starcentral.mbl.edu/microscope/>. All data will be archived as soon as it becomes available on storage devices located in each of the PIs labs and on their institutional servers, which are backed up both locally and remotely to ensure data security and prevent loss. In addition to on-line repositories, project results will be presented at scientific meetings and in peer reviewed publications.

Short-term storage and data management

Data volumes. Data storage shall be appropriately and redundantly stored using the computer hardware and software that is available to the Fields laboratory, and may include in-house and/or offsite resources. The PIs are responsible for due diligence with respect to short-term storage of data. Additionally, all data shall be retrievable from primary media or back-ups, as well as reasonably protected from accidental loss due to corruption, power loss, or failure of computer hardware.

Data security. Data shall be stored on either off-network mobile devices (external hard drives) or offsite cloud resources. Password protection will be utilized. This data (and associated passwords) must be made available to senior institution officials in the case that any institutional liability issues should arise.

Data backup. Stored data shall be backed-up weekly on redundant in-house data backup systems.

Deposit and long-term preservation

Long-term strategy. Within two years of data collection, data will be transferred to BCO-DMO for public access and long-term storage. After the project has been completed, arrangements shall be made to transfer data at Bigelow Laboratory from short-term storage to a long-term archival system.

Length of archival. Data will be kept in long-term storage for at least five years, or until it has been successfully uploaded to and made publically available through a nationally or internationally funded database specific to that data.

6) Documentation and Sharing of Data and Samples: As indicated in the 'Products of the Research' section above, all aspects of our research will be documented in both physical (Lab Notebooks) and digital formats. The project will be highlighted on the websites of the PIs at Georgia Tech and Bigelow, where we will also provide links to digital data repositories following the NSF guidelines for the timeline of both data and sample release specified in the '2004 NSF Division of Ocean Sciences Data and Sample Policy'. Logs will be kept by PIs and project participants for all samples collected and analyses performed. The PIs will collaborate closely on NSF Annual Reports to provide updates to NSF program managers, and to the general public. The PIs will collaborate on data management in their respective laboratories. Archived and frozen (-80 °C) sub-samples backup samples will be available to interested parties for their independent analysis following publication of results.

7) Curriculum Materials: Large amount of teaching and public lecture materials will result from the proposed research, primarily in the form of digital presentations and large-format research posters. These materials will be made available on the websites of each PI for educational use following the guidelines from NSF on data release and fair use. The PIs expect to incorporate results from the proposed research into their undergraduate course lecture materials.