Data Management Plan

Introduction

Data management will be coordinated by the Principal Investigators. The plan encompasses three areas: use policies, data preservation and archival, and standards. The PI will leverage existing systems wherever possible. All data will be communicated in a timely fashion following the NSF policy.

Types of data to be collected

This project will generate an interdisciplinary dataset. Laboratory data will include soluble and solid-state measurements during experiments and natural abundance of stable isotopes. Limited field data will include sediment material properties (e.g. operational silica pools) and bottom-water samples (e.g. hydrography, dissolved phase measurements). Measurements of underway physical conditions when taking field samples (e.g. R/V Pelican's underway data) and water-column hydrography (CTD instrumentation to provide context) will be generated also.

Data access and sharing policies

Data collected under the project will be made available to the public with as few restrictions as possible. Under these policies, all PIs plan for publication of most data with metadata, with manuscripts submitted during year 3 of the study and all major results published at a maximum of two years after the completion of the study. Once data are uploaded, PIs will confirm to BCO-DMO convention for making data publicly available within 1 year (or less).

For each sample collected we will generate a suite of data points. Corresponding subsamples will be assigned a unique identifier with associated metadata. This identifier will be used to ensure that data generated in individual labs is associated with the correct sample. All sample data will be logged into shared documents (GoogleDocs or DropBox) that will be accessible to all project personnel. This approach has the added benefit of maintaining a copy of the data in the cloud as a backup. Data will be rigorously checked for quality as it is collected, with raw data maintained as well as processed and analyzed data.

Plans for archiving and preserving data

As is required, metadata and data will be contributed to one or more existing catalogs. PI Krause will be responsible for working with the Biological and Chemical Oceanography Data Management Office (http://www.bco-dmo.org) to establish a project portal. Shipboard underway data generated during multi-corer deployment and CTD hydrocasts will be automatically sent to the Rolling Deck to Repository (R2R) catalog. This will be linked to the final BCO-DMO project. When ready, all PIs will upload their respective data sets to BCO-DMO. Raw data generated from ETH Zürich MS ICP-MS will also be housed locally and maintained indefinitely per institutional regulations.

Standards and formats to be used for metadata and data

The PIs will conform to the metadata standards of BCO-BMO. As much as possible, data will be archived in ASCII format, which is the most flexible and readable over the long term. The PI will archive data in tabular formats that have been proven successful when sharing data. The project-supported laboratory technician and PIs will serve as the lead data managers and ensure data meets quality standards.