

## DATA MANAGEMENT PLAN

Data and information management for this project will be done in close collaboration with the Moorea Coral Reef LTER program at UC Santa Barbara, which has agreed to provide data documentation as well as archival and web access services to this project (see attached letter of commitment from MCR LTER Lead PI Schmitt). The MCR LTER Information Management System (IMS) will facilitate the archival cataloging of our data for long term preservation, enable the discovery of our data and enhance their suitability for synthesis by us and others.

The MCR Information Management System meets or exceeds all current LTER criteria for Information Management as set forth by the LTER Information Managers Committee and LTER Network Information Systems Advisory Committee. In brief, UCSB's Marine Science Institute provides infrastructure and IT resources, including a web server, database server, and filesystem server with a Storage Area Network (SAN) supporting snapshots. Purposeful redundancy in backup systems provides for disaster recovery with off-site copies stored in a separate location, and for file restores from more frequent on-site backups. The entire data catalog inventory is cached annually as a DVD archive off campus. With respect to data documentation, metadata documents in EML 2.1 dynamically derived from Metabase will be produced and maintained for our data (see below), which will make them readily adaptable to future enhancements in metadata documentation. Data accessibility will be through the MCR LTER website (<http://mcr.lternet.edu/>), which meets or exceeds the Guidelines for LTER Website Design and Content (version 1.1 2009). The website uses hierarchical navigation to provide access to the data catalog, publications and research foci. The MCR provides similar data and IM services to a coral reef LTREB program in the Caribbean, and the description of the LTREB program and metadata are publicly accessed from the front page of the MCR website.

### **Data Access Policy and Data Distribution**

Our databases and model codes will be subjected to the same data access policy and data distribution schema as the MCR. MCR data use policy and data release policy conform to LTER Network policies and use "Type I versus Type II" terminology. With some exceptions that are consistent with LTER data policies, MCR data are Type I (publicly available). There is no delay in releasing MCR data to the public once Quality Assurance has been verified. In accordance with LTER policy, graduate student thesis data are archived and cataloged as Type II (not released until thesis publication). In addition, some types of data are subject to guidelines for use of Human Subjects. For example, to protect informants' identities, some data will be presented in a coded format that hides all identifying characteristics of the informants. MCR collects data use information upon acknowledgement of the data access policy.

### **Metadata**

Data packages will conform to the most recent (August 2011) version of Best Practices for LTER dataset EML. Metadata features will include embedded or online links to methods and protocols, full temporal, spatial, and taxonomic coverage, keywords from the MCR vocabulary, the NBII thesaurus, and/or the LTER Controlled Vocabulary, and units registered in the LTER Unit Dictionary. All EML will be version 2.1.0. All data tables will be congruent as far as the EML Congruency Checker will be able to check. Beyond the required elements, some datasets may provide explicit indexing keys and table-joining keys to facilitate cross-dataset synthesis.

### **Dataset Management (Data Life Cycle)**

To assure immediate and long term usefulness of project data requires metadata of the highest quality. Datasets will vary in update frequency. Data reach the IM office in different stages of maturity depending on their type. Quality control (QC) is done first automatically within the database, flagging further QC for human inspection. The MCR Information Manager (M. Gastil-Buhl) will work with us regarding naming conventions, guidance for file organization and format, and designation of space on the server to back up raw files from the field. Co-PI Rassweiler will be responsible for ensuring data and model codes are added to the catalog in a timely fashion.

### **Types of Data, Samples, and Other Materials to be Obtained**

The following data will be collected: livelihood strategies, dietary diversity, migration, remittances, gardening, time allocation dedicated to marine and land resource harvesting strategies, consumption patterns, perceptions of environmental changes, fishing behavior and effort, and social network information that records the names of people an informant mentions in their social network. All data will be collected in the communities of Ha'apiti, Ma'atea, and Paopao on Mo'orea, French Polynesia. All data will be collected in the form of interviews, surveys, and questionnaires. All surveys and questionnaires will be back-translated by Tahitian research assistants to ensure accuracy of translations. We will store interview notes, recordings, transcriptions and questionnaire data to both a personal laptop and an external hard drive where they will be organized using the software programs UCINET, Microsoft Excel, and ArcGIS. Upon return to the US, we will use computers at UCSB and SDSU and additional software, like STATA, for statistical analysis of questionnaire data. We will also keep and store all handwritten field notes. In addition to the data just described, historical catch data for fishers on Mo'orea will be obtained from archives. Both during fieldwork and once back in the US, we will back up all data files on the MCR server to ensure against damage or loss of the laptop and/or external hard drive. All quantitative data will be archived in fully documented MCR databases and other pertinent materials will have metadata documentation that also is accessible on the MCR website.