

DATA MANAGEMENT PLAN

We are committed strongly to ensure the long-term use and availability of the data from this project. We have a particularly keen interest in doing so given our ongoing efforts in meta-analysis, which (at the present time) are severely constrained by: 1) limited data availability; 2) the highly summarized nature of many publicly available datasets; and 3) poor metadata and documentation associated with existing data.

As part of this project we will not only strive to ensure the availability of our new data, but we also will make available data that were collected as part of our past NSF-funded research in Moorea. These older data were archived without regard to rigorous metadata protocols that are now commonplace (see below) and are, therefore, not currently available to the broader scientific community. In this proposal, we have budgeted at least one graduate student throughout every year of the project (to assist not only with fieldwork in French Polynesia, but also with other activities on main campus, such as database management). One key responsibility of the students during the academic year will be to assist with data processing, proper archival and data sharing.

Data Collection/Dissemination/Sharing: Our proposed research will yield ecological and physiochemical data sets including spatial distributions and dynamics of vermetids and corals, demographic responses of corals to vermetids, responses of other coral reef fauna to long-term manipulations of vermetids, and physical descriptions of the micro-environment around corals and vermetids. Most of these data will also be spatially referenced. We will retain hard copies of all field and lab data (typically recorded on waterproof paper) in a central repository at the University of Florida, and preliminary electronic copies will be entered into spreadsheets within 48h of collection, with electronic copies maintained at both the University of Florida and Victoria University of Wellington. Metadata will be made publicly available within 6 months of completion of a particular survey or experiment. QA/QC-ed data files will be made publicly available within 2 years of completion of a particular survey or experiment (or at the time of publication, if sooner).

We also will have an extensive compilation of data from the proposed meta-analysis, which will consist of extracted data summaries from the literature (effect sizes, variances, sample sizes), along with ancillary information (e.g., taxonomic status, location, experimental design, physical description of the site, etc.). These data also will be made available. In cases, where data are obtained from authors and not publications, we are obligated to obtain the permission of the original authors before making the provided data available publicly.

Data/Metadata standards/Archival: Data archival is rapidly evolving with a range of groups and initiatives. To ensure the most accessible approach, we will coordinate with personnel associated with NCEAS/EcolInformatics/KNB, DataOne, and the Moorea Coral Reef LTER. For now, we anticipate that we will fully document our data sets using EML (Ecological Metadata Language: <http://knb.ecoinformatics.org/software/eml/>). QA/QC-ed data sets will be archived (and made available to users) as 'flat files' (e.g., text files in .csv format) maintained at the Knowledge Network for Biocomplexity (<http://knb.ecoinformatics.org>). Some data sets may also

be published as 'Data Papers' in Ecological Archives (Ecological Society of America: <http://www.esapubs.org/archive/default.htm>).

Access/Use policies: All data will be made publicly available (e.g., at Knowledge Network for Biocomplexity or MCR-LTER sites) coincident with publication in the peer-reviewed literature or at most, within 2 years of completion of the study. Users of data will be required to provide name and contact information prior to downloading online data to facilitate tracking and notification in case changes are made to data (this is a standard feature available at KNB and LTER). Users will also agree to cite the data set author(s) and the NSF award number in all publications in which the data are used. Users will be prohibited from selling/redistributing data without prior permission from the data set author. Users will be asked to provide copies of any resultant publications to the data set author. In some cases, more detailed collaboration may be possible, and in those cases, we will encourage users to contact the data set author(s) and/or PI to facilitate collaborative projects.