Data Management plan

i.   Types of data

This project will generate six types of data, which lend themselves to archiving and sharing:

(1) physical data (seawater flow, temperature)
(2) community structure (abundance, size, and growth rates)
(3) lesson plans and outreach data
(4) high resolution color images of benthic quadrats
(5) publications & supporting data, particularly as they related to process-oriented studies
(6) genetic data from sampled coral colonies (~ 160 samples)

ii.   Standards to be used for data

This proposal will build on the strong history of sharing data through BCO-DMO, MCR-LTER, NCEAS, USGS (through the Powell Center), and most recently through the Environmental Data Initiative (EDI) to which much of the legacy project data are being transferred. EDI offers many advantages in this regard, notably in the user interface and the ability to more readily access and reference data sets through unique DOIs. Through these efforts, I have considerable experience in developing relational databases (with metadata) that can be readily utilized by others. Additionally, we have designed QA/QC procedures that flag data based on the likelihood they are incorrect (for example, because abundance data changes to an unrealistic extent between years) and marks them for further evaluation.

iii.   Policies for access to and sharing data

Following NSF policies, project-related data available will be made publically accessible servers within 12 months of collection. The exceptions to this rule will be data related to graduate thesis projects, which will not be made available until 12 months following graduation.

Through on going work with the EDI, data archived for the current project will be accessed through end-user consideration of a revised Intellectual Rights statement:

 "This data package is released under the Creative Commons license Attribution 4.0 International (CC BY 4.0. This license states that consumers ("Data Users" herein) may distribute, adapt, reuse, remix, and build upon this work, as long as they give appropriate credit, provide a link to the license, and indicate if changes were made. If redistributed, a Data User may not apply additional restrictions or technological measures that prevent access."

iv.   Polices and provisions for re-use, redistribution, and production of derivatives

All users will have open and free access to our data within 12 mo of collection, unless otherwise embargoed to meet the needs of graduate thesis preparation. Although not required for access, users will be encouraged to acknowledgement access to project data and make contact with the Edmunds in the spirit of effective collaboration.

v.   Archiving and access to data

Over the last 12 y, the Edmunds lab has had a strong and well-developed policy of sharing data through web-accessible systems. I have spent considerable time working with BCO-DMO to make data available from the expired LTREB award (see BCO-DMO project 734983), and we are making data from published manuscripts available within months of publication. We are working to close this temporal gap by assigning DOI values as manuscripts are submitted, and we are working within the EDI framework to further close this gap.

We have worked to leverage the IM support inherent in the Moorea Coral Reef LTER (on which Edmunds is a co-PI) to promote the management of data from St. John. The objective is to provide wider access to

thematically cohesive data, and leverage access and exposure to coral reef time series data at the MCR-LTER site. A key role of local data management (through H. Ake as an Independent Contractor charged with managing the project web site) is to provide wider access to data and deliverables from St. John research that exceeds the services provided by BCO-DMO and now, EDI. The project web site is hosted on WordPress through CSUN, but project data remain archived in an NSF-approved framework at the MCR-LTER site (with much of this transitioning to EDI). The objective is to serve project deliverables to resource managers in Caribbean nations, and provide a data and learning portal suitable for outreach activities at schools where we work in California. The last decade of the LTREB award has benefited greatly from the MCR system, and data and deliverables have been made available for multiple years. Currently I am working on an Advisory Committee to BCO-DMO with the objective of refining their services to better meet the needs of the research and educational community. An important recent (and on-going) development is the initiation of the ERDDAP portal that greatly increases the ease with which data can be sorted, winnowed and downloaded. While this remains a work in progress, it demonstrates the progress that has been made in making 32 y of project data available through local and centralized (BCO-DMO, EDI) resources.
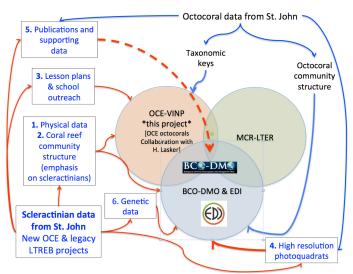

Fig. 1. Schematic showing relationships among the proposed project (with 6 data streams, 1-6), the Moorea Coral Reef (MCR) LTER, a collaborative OCE project to study octocorals (OCE 17-56678), and **three** IM systems (**circles**). Currently we are transitioning much of the time series data from the present project to EDI (as is MCR-LTER), and together the four systems are complementary to serve a broad end-user community with data, graphically rich deliverables, and outreach products. Core raw data – high-resolution images – currently are archived on the MCR-LTER, and we are developing means by which these very large data sets (~50 GB/y) can be served to the community; iDIGBIO offers a promising way to store and access this graphically rich data. MCR-IM and EDI meets the NSF-mandated needs of the LTER community, and the coral reef theme of this system makes it a focus for other coral reef projects such as the OCE project proposed herein. BCO-DMO meets the broadest needs of data hosting for the OCE community (and related projects). Published papers are the key deliverable of all projects and are shown in the common overlap among all three IM systems. H. Ake serves as our contract IM profession who manages the project web site, and will continue in this role in the new award. Her experience working for BCO-DMO provides a unique opportunity to interface the OCE-VINP data needs. Project data are not hosted locally at CSUN. Instead they currently are hosted at MCR, BCO-DMO or EDI, and moving forward, all data will be transitioned to the EDI platform. The project website provides a "one stop shop" for all of the USVI data in a format that has met the sharing requirements for multiple branches of NSF over the last 33 years.