

DATA MANAGEMENT PLAN

We will produce five types of data during each year of the project (Table 1): (1) georeferenced photomosaics with (2) GIS annotations for coral colony and environmental logger positions, (3) 3D surface meshes of habitat, (4) biodiversity and trait data for fish, corals, and potentially other organisms such as urchins and crustaceans, and (5) data from experimental studies of settlement structures. Part of our proposed work is to coordinate with other groups using photogrammetry to map habitat structure and produce a common standard and repository for photo-mosaics and associated 3D products (see “Broader Impacts” in the Project Description).

All data will be collected, compiled and managed in accordance with Biological and Chemical Oceanography Data Management Office (BCO-DMO) guidelines. For instance, all data will be deposited within a public repository within two years of collection; although, to protect a student’s thesis work, we may embargo data until publication. Our data management plan will be housed at the BCO-DMO and we will ensure that all data products are discoverable through the Office’s metadata system. We may house raw data in other repositories (for examples, see Table 1); however, we will ensure the BCO-DMO’s system has permanent DOI links to those other repositories.

Table 1. Summary of standard and repositories for data types collected during the proposed project.

	Data type	Metadata standard	Repository	Permanent DOI
1)	Photomosaics	Geotiff	HIMB and UH Storage, and Squidle*	BCO-DMO, Zenodo, Figshare, Dryad or equivalent
2)	GIS species annotations	GIS Shapefiles	HIMB and UH Storage	
3)	3D meshes	Polygon File Format	HIMB and UH Storage, see “Broader Impacts”	
4)	Biodiversity data	Darwin Core	BIOTime Coral Traits	
5)	Experimental data	Ecological Metadata Languages	Knowledge Network for Biocomplexity (DataOne)	

The project investigators are well-versed in managing large data compilations, including analysis code, images, maps, and environmental and ecological data. PI J Madin co-developed and directs the Coral Trait Database, an open data initiative to collate all Scleractinian coral functional and life history trait data (<https://coraltraits.org>). SP Dornelas runs and directs the BIOTime database that collates and store temporal biodiversity database globally (<http://biotime.st-andrews.ac.uk>). All investigators understand and practice open, reproducible science and use a combination of R, markdown/latex, git and GitHub to create workflows and coordinate projects. PI Madin and SP Dornelas have experience writing and publishing data paper in journals such as *Ecological Archives*, *Scientific Data* and *Global Ecology and Biogeography*. PI Madin and SP Dornelas have used pre-print servers to speed up dissemination of important results and release of data products. We will make all data and publications open access, and have requested publication funds to help do so (see “Budget”).

*Squidle: <http://squidle.acfr.usyd.edu.au>