

## **Data Management Plan**

We will adhere to the principles of FAIR (findability, accessibility, interoperability, and reusability) data. Our management plan will facilitate active communication among all team members and interested stakeholders (e.g., state and federal natural resource management agencies). Also, it will include a work plan to make data and products generated by the project available openly and rapidly to the public. The PIs will be responsible for data collection, quality control, archiving, and reporting following NOAA and NSF policies. The PIs will ensure all project personnel is appropriately trained in data collection methods and in the best practices of data documentation and organization in spreadsheets (e.g., Broman and Woo 2018). At the start of the funding period project, PIs will conduct a video conference with staff, students, and technicians associated with the project to discuss the approach that will be taken for data management. The PIs will lead a discussion of site naming, data processing, and storage conventions that will be used for the project. They will also emphasize the value and importance of metadata.

### 1. Data Types and Formats

Data products from this proposal (e.g., coral demographic data, environmental data, reef community metrics) will be recorded on data sheets and scanned to be stored electronically (.pdf). All data will be entered into digital format as text files. Images files generated by the project will be stored electronically. Also, laboratory data (samples processing for SIA) will be collected in logbooks photographed/scanned daily for duplicity purposes. Field and laboratory-generated data will be entered on a weekly basis into central, shared spreadsheets. Spreadsheets will be designed with data validation tools to reduce data-entry errors. Data obtained from field instruments (temperature loggers) and received from external analyses (e.g., SIA) will be downloaded and stored as Excel spreadsheets and CSV files. This project will generate:

- Benthic digital images and resulting large-image orthomosaics (photomosaics) stored as tiff and 3D model files
- Data from the benthic invertebrate and algae community extracted from photomosaics (species presence, abundance, assemblage, and diversity) stored as CSV and Excel files
- Data of coral species size distributions and spatial pattern metrics derived from the photomosaics
- Urchin counts from field observations and photomosaics stored as CSV and Excel files
- Demographic parameters of coral species and urchin populations stored as CSV files
- R software scripts for data analysis (R open-source format)
- Datasheets and field notes of data collection stored as scanned images (jpg or pdf) and input into excel spreadsheet files
- Presentations using project outputs (Microsoft PowerPoint and pdf format)
- Education materials (pdf format)
- Project papers and reports (pdf format), open access fees paid when possible
- Github repository of R scripts and links to databases

### 2. Standard to be used for metadata

Metadata for this project will be created using the free, web-based CatMDEdit application tool, with details consistent with the Ecological Metadata Language (EML). This tool provides an overview of the project's purpose, sampling design, geographic and temporal extent and details on the specific data collected (data field names, descriptions, units of measure, expected ranges, etc.). Metadata development will comply with the National Science Foundation's Florida Coastal Everglades (FCE) Long-Term Ecological Research (LTER) current information management policy. Project PIs (Santos, James, and Rehage) are an FCE LTER collaborator, [http://fcelter.fiu.edu/research/information\\_management/documents](http://fcelter.fiu.edu/research/information_management/documents). Metadata records for the data to be collected will be available before the data being collected. CatMDEdit allows records to be exported individually as ISO-compliant XML files.

### 3. Policy for data access and sharing

This project will develop and abide by a data release policy. Data will remain within the project for two years after collection or until publication (whichever is first). At that point, data will be available to the broader community for academic use, following NOAA's philosophy of dissemination and sharing of research results and using NSF LTER Data Access Policy. If these requests are within two years of collection, data will be shared unless there is a direct overlap or conflict with achieving the project's aims. All data sharing will be documented in a bespoke database in Microsoft Access, Environmental Data Initiative (EDI) or Github data repository. Data recipients are asked to acknowledge funding agencies and be responsible for errors in subsequent analyses. The project is not anticipated to generate any private or confidential information, but these data will not be released publicly if it does occur. GitHub, the powerful version control, source code management, and sharing platform, will be used to manage and distribute software artifacts.

### 4. Archiving data and other research products, and preservation of access

We will work with the Biological & Chemical Oceanography Data Management Office (BCO-DMO) to create a project page for the long-term archival of all data files and metadata. All data will be deposited and made public within 2 years of collection. This includes data collected from other funded projects that will be leveraged for this project.

### 5. Data backup during project

Google cloud storage (Google Drive) will be used to store and share data among the team, mirrored in second cloud storage (e.g., OneDrive, Dropbox, GitHub). Best version control practices will also be followed when using these cloud storage/sharing services. Laptops and PCs owned by the PI and other project personnel will primarily be used to manage and analyze data sets while they are being assembled. Generated datasets will be secured in PIs Santos/James/Rehage office's computers and external drives, and data at FIU computers/external drives will be subjected to nightly archiving and backup provided by two FIU servers: 1) OneDrive webserver/data cloud and 2) a separate server owned and operated by FIU's Institute of Environment. Data archiving.