

Data Management Plan: AMOR Project

Introduction: Data management of chemical sensor data will be coordinated by PI C.S. Martens and Engineer H. Mendlovitz at UNC-Chapel Hill. Microbial community data will be managed by Co-PI Karen Lloyd at U. Tennessee (UT). All project participants will have access to and participate in data analysis and interpretation. The plan encompasses three areas: use policies, standards, and data preservation and archival. We will leverage existing systems wherever possible. All data will be communicated in a timely fashion following the NSF policy.

Several types of field data will be generated in the proposed activities. These include chemical and physical data collected during oceanographic cruises including mini-lander chemical sensor time-series measurements, microbial community meta-data, water chemistry data from CTD casts, shipboard wet lab chemical dissolved inorganic nitrogen, 16S rRNA gene amplicon libraries, cell counts, metagenomes, and metatranscriptomes.

Data access and sharing policies: We are committed to making all data types publicly available through peer-reviewed publications and public databases with as few restrictions as possible. As data is processed and analyzed, both raw and processed data will be uploaded to networked servers maintained at UNC-Chapel Hill and the UT. All servers are backed up regularly and data can be made available to collaborators needing access. Relatively rapid (within months) sharing of data sets between UNC and UT will be crucial for the AMOR project to fully succeed. Data will be transferred to Biological and Chemical Oceanography Data Management Office (BCO-DMO) following processing, and public access will be granted to data following its publication or at most, within two years after its collection.

Formats to be used for metadata and data: We will conform to the metadata standards established by the BCO-DMO. As much as possible, data will be archived in ASCII format, which is the most flexible and readable over the long term, though long time series data may be transferred to BCO-DMO in more native formats, e.g. matlab. In the case of chemical sensor data concentrations will be reported to BCO-DMO along with detailed information on blank corrections and calibration between sensors including laser methane and oxygen optode sensors.

Plans for archiving and preserving data: Data files will also be stored in long-term storage space provided to UNC and UT researchers by campus support facilities. In all our efforts we will work with the BCO-DMO to archive the data and to ensure our metadata conform to their standards. Raw DNA reads will be submitted to the Short Read Archive at the NCBI. Metagenomes and their metagenome-assembled genomes, as well as metatranscriptomes, will also be deposited at the IMG/M database at the Joint Genome Institute, since this is a more searchable database than NCBI.