DATA MANAGEMENT PLAN

Our Data Management Plan will follow the policies established by the NSF Directorate for Geosciences and published in the document titled *Division of Ocean Sciences Sample and Data Policy* dated May 2011.

1. Types of data produced: The "REU Site: Ocean Observing for Emerging Ocean Scientists," project implemented by Lisa Campbell and Anthony H. Knap will generate physical, chemical and biological oceanographic data. The data will be produced by electronic sensors, chemical analysis, visual observations made in the lab or field, and numerical models.

2. Data and Metadata Content Standards: To the extent possible we will encode data into forms and formats that conform to community standards, some of which are evolving (e.g., profiling glider data). For physical, chemical data and model output, current community standards include National Ocean Data Center's NetCDF Feature Type Templates, Unidata's: NetCDF Attribute Convention for Dataset Discovery (ACDD), Climate and Forecast (CF) conventions and Standard Names. We will use the ISO 191** Metadata Family as our metadata content standard. For biological data we propose to use FGDC's "Biological Data Profile for the Content Standard for Digital Geospatial Metadata" or Darwin Core when reporting the geographic occurrence of species or taxonomic group. Encodings to be used include netCDF, ASCII, where needed XML.

3. Mechanisms for Access and Sharing: None of data produced by this project will be private, confidential, or require password protection against unauthorized dissemination. The research results from this project will be disseminated through peer reviewed publication in the scientific literature and at national and international scientific conferences. Data produced by this project will reside on computer disk storage at Texas A&M University as stand-alone files or in databases structures. Network accessible catalogs, webpages and "read-me" files will help interested parties discover data. File access will be accomplished through the network transfers using a variety of mechanisms such as FTP, HTTP, or email. We may employ online databases, THREDDS Data Servers, or ERDDAP middleware to improve data discovery and exchange and to provide the capability of remote data subsetting. Large files or data set collections can be written to disk and shipped. Standard best practices for data storage and duplication, and the university's computer network infrastructure will keep data sets secure from accidental loss or malicious deletion or corruption. The Gulf of Mexico Coastal Ocean Observing System's data management system is housed at Texas A&M University. Their systems will be used to supplement and extend data discovery and dissemination services. Data collected by the Imaging FlowCytobot according to the procedures described in the published articles by Campbell et al. in 2010 and 2013. All image data are available via our website: http://toast.tamu.edu/new_data/dashboard.html.

4. Policies and Provisions for Re-use and Re-distribution: We are committed to the open and unrestricted sharing of the full body of data from this project as soon as practical. We respect the data originator's right to first and best use of the data within a reasonable time frame not to exceed two years beyond the date of collection. We will encourage data originators to make data available to the public as soon as practical after basic quality controls have been applied and some use of the data has been made to verify that the data are reasonable and valid. Data originators may freely exchange data within and beyond the project at their discretion. The Data Manager will obtain a data originator's permission before data releases. Through a written Data Policy Agreement, data recipients will be encouraged to contact the data originator and provide attribution if any scientific re-use of the data are intended or made.

5. Plans for Archiving Data: The full body of data and selected model output suitable for preservation will be submitted to the National Ocean Data Center prior to the end of the project for long-term archiving. The data will also be held by the Gulf of Mexico Coastal Ocean Observing System for as long at that entity exists. Data will also be held by the data originators for three years following the end of the project. If required, we will submit biological and chemical oceanographic data to the Biological and Chemical Oceanographic Data Management Office (BCO-DMO) following their formats and procedures.