

## 8. Data management plan

All data collected during this project will be managed by the Biological and Chemical Oceanography Data Management Office (BCO-DMO) located at WHOI. The BCO-DMO will also handle submission of the data to NODC for final archiving at the end of the project. We will meet with BCO-DMO staff during the first six months of our project to discuss details associated with each of our data types and define protocols for producing appropriate data format, documentation of quality control, and metadata. BCO-DMO staff will also provide guidance on best practices for cruise data management (cruise reports and sampling event logs) and facilitate the publication of our results after the cruise. Underway data are critical to this project, and we are pleased to contribute standard underway ship-based measurements as well as our own measurements as part of the UNOLS central data repository at <http://www.rvdata.us/catalog/>, managed by the Rolling Deck to Repository (R2R) project.

While access to data will be limited to the participating investigators for an initial period of time, public access to all data and supporting documentation (metadata) will be granted within two years. Data and metadata from this project will be available as part of the larger BCO-DMO data system. The ability to integrate results from this project with those from prior research will greatly enhance the value of the data to be collected and ensure its central maintenance and accessibility into the future.

Dr. McGillicuddy has extensive experience with BCO-DMO staff, having worked closely with them in management of data from the NSF-sponsored EDDIES project (OCE-0241310), Quantification of *Trichodesmium* spp. vertical and horizontal abundance patterns and nitrogen fixation in the western North Atlantic (OCE-0925284), as well as a collection of data sets dealing with harmful algal blooms in the Gulf of Maine:

<http://www.bco-dmo.org/project/2048>

<http://www.bco-dmo.org/project/2104>

<http://www.bco-dmo.org/project/2118>

Results from three of Dr. McGillicuddy's modeling projects are also archived there:

<http://www.bco-dmo.org/dataset/3198>

<http://www.bco-dmo.org/dataset/3195>

<http://www.bco-dmo.org/project/473687>

and we will do the same for the modeling results generated by the present project. We will also supply model output (or links thereto) to the two relevant regional components of NOAA's Integrated Ocean Observing System: the Northeastern Regional Association of Coastal and Ocean Observing Systems (NERACOOS), and the Mid-Atlantic Regional Coastal Ocean Observing System (MARCOOS).

*Image and image product data sets* – Because of the unique challenges in effectively sharing the large image data sets and the analysis products associated with them, we will also provide specialized access to these observations. This will be accomplished with mechanisms already established in the Sosik laboratory for rapid and easy web-based access to these types of data. This access includes not only raw images (~several GBytes day<sup>-1</sup>), but also the metadata associated with each image (e.g., date/time, as well as fluorescence, light scattering, location in camera field, etc.) and the routine image products produced from our analysis pipeline (masks to identify target pixels, extracted features and biomass metrics, taxonomic classification results for each image). This will follow the pattern currently implemented for IFCB imagery collected at MVCO; see <http://ifcb-data.whoi.edu/> where images and associated metadata are openly accessible through html, PNG, JPG, RDF, XML, and other standard formats via web services.