Data Management Plan: NGA LTER Site

Data Management System

Data management will be supplied by Axiom Data Science (www.axiomdatascience.com) and leverage the existing data system developed and supported by its partner, the Alaska Ocean Observing System (AOOS). Axiom is an informatics and software engineering firm focused on developing scalable cyberinfrastructure to integrate, curate, and provide access to real-time, modeled, GIS and remote sensing data. Through this data system, PIs will have access to the Research Workspace, which is a web-based project level data management system. The Workspace serves as central platform to store and collaboratively share preliminary data, sampling protocols, and other materials.The Workspace is coupled with an integrated editor for PIs to generate metadata records that comply with ISO 19115 standards. The Workspace and its metadata editor will be supported and cultivated throughout the project with several major modifications and upgrades planned. Data and metadata stored on the Workspace will then be packaged and sent to the DataONE member node maintained by Axiom/AOOS, and linked to the AOOS Gulf of Alaska Data Portal (http://portal.aoos.org/gulf-of-alaska.php) to be made publicly-accessible to broad community of scientists and decision-makers. A large task will be to work on compatibility of data streams from various projects, so that synthesized products and visualizations can be created. A web-based data visualization/integration tool will be deployed, which provides users with an interface to NGA LTER project level data and visualizations. This interface will allow users to query project profiles and metadata by time, parameter, species and spatial location. Significant work has already been undertaken by Axiom staff to design and develop this system as a component of the core AOOS data system. Long-term preservation of data in DataONE and the LTER Network Information System (NIS) Data Portal will be facilitated through the new, automated data-submission feature in development for the Workspace, scheduled for release concurrently with the Axiom/AOOS DataONE member node in November 2016. The NGA LTER website will also inform the public of the project and provide a link to the data visualization interface and archive.

The Axiom team will provide the NGA LTER with critical data management support to assist study teams in efficiently meeting their objectives and ensuring data produced or consolidated through the effort are organized, documented and available to be used by fellow LTER investigators and future research efforts. The data management team proposes to leverage AOOS's existing data system already in use by the Seward Line PIs to support data submission, metadata generation and data transfer to data repositories using the Research Workspace. The data management team will work closely with researchers to track data submissions to ensure they occur within 2 years of collection. It will review/audit metadata and data structure formats produced from NGA LTER projects and advise study team members in best practices for data formats and metadata authoring. Axiom software engineers will support and enhance existing web-based tools designed for the discovery and interactive exploration of visualized NGA LTER project data. The data management team will also guide and support PIs in the generation and update of data management plans, authoring metadata, and archiving project data to the NIS Data Portal and DataONE according to LTER Network policies.

Types of Data

This project will generate and assemble diverse types of data and products from moorings, cruises, and ocean circulation/biogeochemical modeling. Data from shipboard oceanographic sampling will include conductivity, temperature, depth, dissolved organic/inorganic carbon, chlorophyll a, micro- and macronutrients, phyto-, micro-, and zooplankton composition and biomass, primary production via stable isotope analysis, optically-derived particle size/flux, and seabird and marine mammal observations. Moorings will collect oceanographic data at multiple depths, including temperature, conductivity interpreted as salinity, acoustic recordings, acoustic backscatter, photosynthetically active radiation, iron speciation, current velocity, directional wave spectra, particle size spectra, colored dissolved organic matter, optical backscatter, pH, and variety of other data types depending on final instrumentation. Physical-biogeochemical model simulations will produce regional predictions of phyto- and zooplankton community responses to variations in seasonal climatology of freshwater input.

Data and Metadata Standards

Data management staff will assist with the creation of data management plans (DMP) specific to each NGA LTER project at its onset; thereafter, project PIs will update DMPs annually. NGA LTER

projects will use well-defined, community accepted data and metadata formats and appropriate standards. We will use guidelines for data and metadata developed by the LTER Network and consistent with the Division of Ocean Science Sample and Data Policy that are designed to promote broad data access, standardization, and long-term data usability. Tabular data will be delivered to the Workspace as comma-delimited ASCII files (csv) which include header information that is uniformly formatted for each data type. Any geospatial products will be stored in the Workspace in their native formats and converted to shapefiles, geoTIFFs, or netCDF files for long-term preservation and sharing. Oceanographic models (& datasets) will be shared and preserved in netCDF to support a machine-independent format for representing data. Metadata documentation will be generated in either ISO 19115 or Ecological Metadata Language (EML) specifications that are adapted for a variety of data types and used universally in the oceanographic and ecological sciences. Metadata authoring will follow policies that meet the LTER standards for the NIS Data Portal. Data management staff will recommend and provide guidance on the metadata editor tool available through the Research Workspace. The Workspace metadata editor generates ISO 19115 standard metadata, with a roadmap for future developments including the ability to export metadata content as EML. In the interim, investigators that chose the EML format will be required to upload their EML record to the Workspace where it will be packaged, shared, and preserved with the data, and translated into ISO 19115 for discovery through the publicly-accessible portal(s).

Data Submission Policy

The NGA LTER site endorses the LTER Network Data Access Policy, which states that research data must be made available online within 2 years of collection and no later than publication of the main findings. The following guidelines will be followed to ensure the availability of NGA data to a broad research community: 1) Project data management plans must be submitted by the PI at the start of a new project and updated annually thereafter until completion; 2) Data and metadata must be submitted to the Research Workspace within 1 year of collection, and will be freely and publically available in the LTER NIS Data Portal and the AOOS Gulf of Alaska Data Portal not to exceed 2 years after collection. 3) Primary responsibility for data completeness and integrity (quality control) rests with the submitting PI.

Data Use, Archive, and Preservation

Within 2 years of collection, NGA LTER site data will be freely available from the DataONE network, the LTER NIS Data Portal, and from the AOOS Gulf of Alaska Data Portal, along with descriptive metadata and supplemental documentation as appropriate. Data collected during NSF-supported oceanographic research cruises will be submitted to the R2R (Rolling Deck to Repository) program by the vessel operator. Prospective data users will be asked to identify themselves to the LTER PI and/or the PI responsible for the dataset to be used. Metadata for each dataset will include citation information as well as the following use statement: This dataset is released to the public and may be freely reused. Please keep the NGA LTER site and the dataset contact person informed of any plans to use the dataset. Consultation or collaboration with the original investigators is strongly encouraged. Publications and data products that make use of the dataset must include proper acknowledgement. More information on LTER Network data access and use policies is available at: http://www.lternet.edu/data/netpolicy.html.

Milestones and Deliverables

2017 Dec: NGA LTER Workspace group created, PIs have access, trainings scheduled for Jan-Feb 2018.
2018 Oct: Data inventory exists, submissions procedures shared with PIs and tracking by Axiom team.
2018 Oct-2019 Oct: Data from 1st year collection submitted to the Workspace with draft metadata describing projects and datasets; Axiom quality reviews data formats and metadata; repeated annually for data collected in subsequent years.
2019 Oct - 2020 Oct: Data from first collection year has been transferred to LTER NIS with complete metadata; repeated annually for data collected in subsequent years.
2023 Oct: All data and metadata, including model results and documentation, from initial 6 years of NGA LTER has been transferred to LTER NIS and DataONE.