**Transition of the Oculus Glider into Operations for NOAA Ecosystem Research**

(December 6, 2022)

**UxS Data Management Pilot Program**

**2021-07-OMAO |** Transition of the Oculus Glider into Operations for NOAA Ecosystem Research **| RL 7**

[Brief given 2022-02-09](https://docs.google.com/presentation/d/1LX5enKgNQcqA7LhYRg4IuwnT7I5Awxoi/edit?usp=sharing&ouid=105453922989865985777&rtpof=true&sd=true)

*Please provide the following information, and submit to the NOAA DM Plan Repository.12*

**Reference to Master DM Plan (if applicable)**

*As stated in Section IV, Requirement 1.3, DM Plans may be hierarchical. If this DM Plan inherits provisions from a higher-level DM Plan already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DM Plan.*

URL of higher-level DM Plan (if any) as submitted to DM Plan Repository: None

1. **General Description of Data to be Managed**
   1. Name of the Data, data collection Project, or data-producing Program:

Transition of the Oculus Glider into Operations for NOAA Ecosystem Research

* 1. Summary description of the data:

The Oculus Coastal Glider is a novel technology to approach ecosystem research in the shallow depths of the US Arctic and its margins, regions extremely difficult to survey with current observing capabilities. The Glider is intended to increase observing capacity to meet NOAA observing requirements in the US Arctic, specifically for the Ecosystems Fisheries-Oceanography Coordinated Investigations (EcoFOCI) program at NOAA, Pacific Marine Environmental Laboratory (PMEL). NOAA PMEL EcoFOCI will initiate routine operation of the Oculus Glider and integrate these observations with NOAA’s only biophysical mooring array in the US Arctic, and with shipboard operations within the NOAA Arctic Program. These data will be used to support predictive models, document natural variability and provide foundational research to support fisheries resource management decisions in Alaska. Additionally, information will support long-term biophysical observations in the region, will expand partnerships and resources for NOAA Research and the NOAA Arctic Program, and will facilitate international collaboration through the Distributed Biological Observatory.

* 1. Is this a one-time data collection, or an ongoing series of measurements? Ongoing
  2. Actual or planned temporal coverage of the data: Planned missions will be coincident with the EcoFOCI field and cruise schedule, occurring during extended spring and fall seasons.
  3. Actual or planned geographic coverage of the data: US Arctic shallow-water areas including the Bering and Chukchi Seas
  4. Type(s) of data: Digital numeric data.  
     *(e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.).*
  5. Data collection method(s): Autonomous underwater vehicles, glider, Oculus Coastal Glider, biophysical mooring array in the US Arctic, ships from NOAA Arctic Program  
     *(e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)*
  6. If data are from a NOAA Observing System of Record,15 indicate name of system: N/A
     1. If data are from another observing system, please specify:

1. **Point of Contact for this Data Management Plan (author or maintainer)**
   1. Name: Phyllis Stabeno
   2. Title: Physical Oceanographer
   3. Affiliation or facility: NOAA, Pacific Marine Environmental Laboratory
   4. E-mail address: Phyllis.Stabeno@noaa.gov
   5. Phone number: 206.526.6453
2. **Responsible Party for Data Management***Program Managers, or their designee, shall be responsible for ensuring the proper management of the data produced by their Program. Please indicate the responsible party below.*
   1. Name: Shaun Bell
   2. Position Title: Research Scientist/Engineer
   3. Name of current Position holder: Shaun Bell
3. **Resources***Programs must identify resources within their own budget for managing the data they produce.*
   1. Have resources for management of these data been identified? Funding resources for data management are contingent upon approval of the operational implementation of this capability by NOAA Research, and is subject to availability of appropriated funds.
   2. Approximate percentage of the budget for these data devoted to data management (specify percentage or "unknown"): “unknown”
4. **Data Lineage and Quality**

*NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.*

* 1. **Processing workflow of the data from collection or acquisition to making it publicly accessible:**

Near-real-time data (preliminary QC, T/S subset of data) will be delivered from the Oculus Glider to the IOOS (Integrated Ocean Observing System) Glider DAC (Data Assembly Center), making it publicly available within a short timeframe through the Glider DAC ERDDAP interface: https://gliders.ioos.us/erddap/tabledap/. Additionally, data will be findable from a map interface: https://gliders.ioos.us/map/#.

Delayed-mode data will return to the lab with the instrument after field season. The data will be further analyzed and QC’d, using in-house tools, and community supported python glider-data tools. The oceanographic context of the data collection areas will be considered as data are further processed, including water-column signals and data complications that are unique to the Bering and Chukchi Seas. When processing and QC are accomplished, Level 1 data will be made available via a publicly-accessible ERDDAP interface at PMEL.

* + 1. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan:
  1. **Quality control procedures employed:** When these instruments are operational, QC for near-real-time data will occur as the data are collected using built-in backend software designed for gliders. Automated quality checks will occur when data land at the NOAA DAC repository. For delayed-mode data QC, data will be plotted and checked for gross outliers. Python coding and existing open-source python glider tools will be used to vet data. The software will be tailored to address region-specific and investigation-specific issues, e.g. sharp thermocline signals unique to the Bering Sea shelf.

1. **Data Documentation***The EDMC Data Documentation Procedural Directive requires that NOAA data be well documented, specifies the use of ISO 19115 and related standards for documentation of new data, and provides links to resources and tools for metadata creation and validation.*
   1. Does metadata comply with EDMC Data Documentation directive? Yes.
      1. If metadata are non-existent or non-compliant, please explain:
   2. Name of organization or facility providing metadata hosting: PMEL internal servers, PMEL external server and NOAA/NCEI. Metadata are combined with data and no additional hosting service is needed.
      1. If service is needed for metadata hosting, please indicate:
   3. URL of metadata folder or data catalog, if known: unknown
   4. Process for producing and maintaining metadata *(describe or provide URL of description)*: The Oculus Gliders have back-end software developed by the University of Washington for post-processing, engineering-level checks, and point-by-point QC of data. Metadata from near-real-time data from the instrument is auto-compiled and bound to the data in CF-compliant (CF-1.6) NetCDF formatted files. For delayed-mode data, the ERDDAP interface stores metadata in XML format, with information culled from near-real-time NetCDF metadata and enhanced in-house via glider tools. Archiving to NCEI requires ISO-compliant metadata that will be written and submitted as part of each data contribution.
2. **Data Access***NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements.  The EDMC Data Access Procedural Directive contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, provides information about resources and tools to enable data access, and includes a Waiver to be submitted to justify any approach other than full, unrestricted public access.*
   1. Do these data comply with the Data Access directive? Yes
      1. If the data are not to be made available to the public at all, or with limitations, has a Waiver (Appendix A of Data Access directive) been filed? N/A
      2. If there are limitations to public data access, describe how data are protected from unauthorized access or disclosure: N/A
   2. Name of organization of facility providing data access: NOAA, Pacific Marine Environmental Laboratory ERDDAP (delayed-mode data).
      1. If data hosting service is needed, please indicate: N/A
      2. URL of data access service, if known: https://ferret.pmel.noaa.gov/pmel/erddap/info/index.html
   3. Data access methods or services offered:  ERDDAP web-based interface, includes tools to subset, plot and map data. We plan to serve these data through the PMEL web-based ERDDAP data interface.
   4. Approximate delay between data collection and dissemination*:*
      1. If delay is longer than latency of automated processing, indicate under what authority data access is delayed:

Once the units are in operation and scheduled, near-real-time-data will go directly to the Glider DAC, with perfunctory, automated QC via backend glider software. Data will be available in minimal time, based on the DAC system of automated data-checking before posting. Delayed-mode data will be returned to PMEL for further processing and QC after the field season. Data will be made available within one year of the return of the data to the office setting.

1. **Data Preservation and Protection***The NOAA Procedure for Scientific Records Appraisal and Archive Approval14 describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.*
   1. Actual or planned long-term data archive location:  
      *(Specify NODC, NCDC, NGDC, World Data Center (WDC) facility, Other, To Be Determined, Unable to Archive, or No Archiving Intended)*

PMEL public ERDDAP server and NOAA/NCEI data repository. One set of field demonstration data has been made available at NCEI (Oculus Accession Number 0211129, [https://data.nodc.noaa.gov/cgi‐bin/iso?id=gov.noaa.nodc:0211129](https://data.nodc.noaa.gov/cgi%E2%80%90bin/iso?id=gov.noaa.nodc:0211129))

* + 1. If World Data Center or Other, specify:
    2. If To Be Determined, Unable to Archive or No Archiving Intended, explain:
  1. Data storage facility prior to being sent to an archive facility (if any):

Internal PMEL servers and RAID array, both with built-in redundancy. Public-facing server and ERDDAP front end.

* 1. Approximate delay between data collection and submission to an archive facility:

Once the units are in operation and scheduled, delayed-mode data will be returned for further processing and QC after the field season. Data will be made available within one year of the return of the data to the office setting.

* 1. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive? Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection:

Once the gliders are operational, near-real-time data will be directed to the off-site Glider DAC. Delayed-mode data will also arrive at the lab for further inspection and processing. These data are stored locally at PMEL on a large, internal multi-disk system with built-in redundancy. The storage is not accessible to the public, is behind secure firewalls, and is compliant with NOAA/OAR IT security standards. Cold cloud storage will be established to prevent loss and provide disaster recovery.

1. **Additional Line Office or Staff Office Questions***Line and Staff Offices may extend this template by inserting additional questions in this section.*