ATTACHMENT C EVOSTC Annual Project Report Form

Form Rev. 10.3.14

1. **Program Number:** *See*, Reporting Policy at III (C) (1).

15150114-T

2. Project Title: See, Reporting Policy at III (C) (2).

Supplemental Data Management Support for EVOSTC Monitoring Programs

3. Principal Investigator(s) Names: *See,* Reporting Policy at III (C) (3).

Rob Bochenek, Alaska Ocean Observing System (AOOS)

4. Time Period Covered by the Report: See, Reporting Policy at III (C) (4).

February 1, 2015 – January 31, 2016

5. Date of Report: *See*, Reporting Policy at III (C) (5).

March 1, 2016

6. Project Website (if applicable): *See*, Reporting Policy at III (C) (6).

AOOS Workspace Herring Research and Monitoring Program group:

https://workspace.aoos.org/group/3503/projects

AOOS Gulf of Alaska Data Portal:

http://portal.aoos.org/gulf-of-alaska.php#

7. Summary of Work Performed: See, Reporting Policy at III (C) (7).

Deliverable/Milestone	Status
Objective 1. Provide additional,	Ongoing
needed data management support for	
LTM and PWS Herring programs.	
Objective 1, Task 1: Establish data	Completed.
coordinator position to lead the	 Data coordinator Stacey Buckelew was recruited and
PWS Herring program and assist	hired in June 2015
the LTM program.	
Objective 1, Task 2: Help PWS	Ongoing.
Herring program PIs generate	 One-on-one PI meetings held with PIs in December
metadata for existing data, and	2015 and January 2016 to implement best practices for
add NCML metadata to	metadata record creation
preservation-ready LTM and PWS	
Herring data.	

Deliverable/Milestone	Status
Objective 2. Implement technical mechanisms to seamlessly transfer LTM and PWS Herring program data from the AOOS data system to systems maintained by DataONE Network.	Ongoing
Objective 2, Task 1: Extend the LTM (Gulf Watch Alaska) data portal to participate in the DataONE network as a DataOne Member Node.	Ongoing Feasibility and registration for becoming DataOne Member Node completed; implementation planning and development underway

The major focus of this work has been to respond to Exxon Valdez Oil Spill Trustee Council (EVOSTC) and staff feedback by implementing a supplemental data management effort to execute on tasks that have been deemed of high importance, but were not being addressed by previous data management projects supporting EVOSTC programs (Projects 1412011D and 1412011C). Under this effort, the data management support for both Long Term Monitoring (LTM) and Prince William Sound (PWS) Herring programs has been increased by establishing a data coordinator position. Axiom recruited and hired Ms. Stacey Buckelew into this position beginning June 1, 2015. Her responsibilities have targeted improving metadata quality and best practices. As such, the AOOS Herring Research and Monitoring (HRM) program Workspace group was reorganized in fall 2015 to create a cohesive organizational structure to the Gulf Watch Alaska (GWA) program Workspace group. Additionally, one-on-one meetings were scheduled with individual PIs from the LTM and PWS Herring programs during fall 2016 and winter 2016 to provide guidance and support on data submission and metadata authoring. PIs received individual instruction in the use of the AOOS Workspace and exploration of data available in the Gulf of Alaska Data Portal. A metadata process was also established to ease the authoring process by PIs and to help standardize the metadata formats across programs.

Objective 1. Provide additional, needed data management support for LTM and PWS Herring programs.

Task 1: Establish data coordinator position to lead the PWS Herring program and assist the LTM program.

AOOS, through its technical arm at Axiom Data Science, hired Ms. Stacey Buckelew in June 2015 as the data coordinator to lead the PWS Herring program data ingestion effort. Beyond becoming oriented to the data management team, Stacey met with the LTM data coordinator (Tammy Hoem-Neher), the Herring program coordinator (Scott Pegau), and the Program Management team in Homer in July 2015. During these meetings, conventions were laid out to help establish a cohesive organizational scheme between the two programs. Additionally, Stacey attended the annual PI

meeting in Anchorage during November 2015 and the PI meeting in January 2016 at the Alaska Marine Science Symposium, where she became acquainted with the project PIs for both programs.

Task 2: Help PWS Herring program PIs generate metadata for existing data, and add NCML metadata to preservation-ready LTM and PWS Herring data.

The data coordinator has led the PWS Herring program PIs in organizing their project information and generating metadata records, similar to those created by the LTM program. Generating standardized metadata is critical to ensure that the research investment is capitalized in future research efforts in addition to reducing duplication of effort and increasing data discovery and usability.

The AOOS HRM program Workspace group was reorganized in fall 2015 to create a cohesive organizational structure to the GWA program Workspace group. Several meetings were held with HRM Program Manager, Scott Pegau to discuss an agreed-to organizational structure. Workspace folders were then reorganized and retitled according to individual projects in order to clearly establish the association of PIs to project and enhance their sense of 'ownership.' Additionally, data sets were reorganized by projects and tags added by current status, herring age class, and survey type to ease Workspace access by all PIs.

In concert with the Workspace restructure, a data file and metadata inventory by project was completed. The inventory was cross-referenced with project proposals and progress reports to determine which data files had not been submitted to the Workspace (Figure 1). At the PI meeting in November, the data coordinator presented the inventory and discussed a process for meeting the submission benchmarks with the PIs. The process was agreed-to by all PIs present at the meeting to include the PIs collecting content for the metadata record followed by one-on-one meetings to provide guidance and support on data submission and metadata authoring.

From December 2015 to February 2016, the data management team scheduled 24 meetings with over 30 program PIs or researchers to discuss data submissions and metadata authoring (Table 1). Additionally, PIs received written instructional materials about the Workspace metadata editor and hands-on instruction in the AOOS Workspace (refer to Appendix 1), its metadata editor, and linkage to the Gulf of Alaska data portal including exploration of available data sets. A metadata process was also established to ease the authoring process by PIs and to help standardize the metadata formats across programs. The process included the PI completing a metadata questionnaire document before the meeting that included a set of questions about the project research in order to organize content for the metadata record. The questionnaire was adapted from the U.S. Geological Survey (USGS) best management practices to adhere to International Organization of Standards (ISO) metadata standards. For those projects for which a reasonably complete metadata record already existed, the data management team instead utilized the metadata questionnaire as a completeness check. Prior to the meeting, the data management team reviewed the questionnaires and then used the meeting to assist the PIs in walking through creation of the content need to complete or revise the metadata record.

Table 1. A list of the in-person meetings held with HRM and GWS PIs and researchers this reporting period to discuss data submission benchmarks and metadata authoring.

PI & Researcher name (s)	Project	Meeting	Location	Axiom lead
In person PI meetings with dat	a managem	ent team		1
Gorman	HRM	Jan 2016	Prince William Sound	Buckelew
			Science Center	
			(PWSSC), Cordova	
Pegau	HRM	Jan 2016	PWSSC, Cordova	Buckelew
Bishop, Schaefer	HRM,	Jan 2016	PWSSC, Cordova	Buckelew
	GWA			
Bishop, Lewandoski	HRM	Jan 2016	PWSSC, Cordova	Buckelew
Gay	HRM	Jan 2016	PWSSC, Cordova	Buckelew
Campbell	HRM	Jan 2016	PWSSC, Cordova	Buckelew
Heintz, Sewall, Lindeburg	HRM	Jan 2016	Auke Bay, Juneau	Buckelew
Wildes	HRM	Jan 2016	Auke Bay, Juneau	Buckelew
Moran, Lindeberg	GWA	Jan 2016	Auke Bay, Juneau	Buckelew
Arimitsu, Heflin	HRM	Jan 2016	USGS, Juneau	Buckelew
Branch, Trochta	HRM	Jan 2016	Alaska Marine	Buckelew
			Science Symposium	
			(AMSS), Anchorage	
Boswell, Zezone	HRM	Jan 2016	AMSS, Anchorage	Buckelew
Rand	HRM	Jan 2016	AMSS, Anchorage	Buckelew
Moran, Straley, Lindeberg	GWA	Jan 2016	AMSS, Anchorage	Buckelew
Batten	GWA	Jan 2016	AMSS, Anchorage	Buckelew
Iken	GWA	Jan 2016	AMSS, Anchorage	Buckelew
Kaler	GWA	Jan 2016	U.S. Fish and Wildlife	Buckelew
			Service (USFWS),	
			Anchorage	
Olsen, Coletti, Kloeker	GWA	Jan 2016	AMSS, Anchorage	Turner
Olsen (for Matkin)	GWA	Jan 2016	AMSS, Anchorage	Turner
Holderied, Powell	GWA	Feb 2016	Homer	Buckelew
Doroff	GWA	Feb 2016	Homer	Buckelew
Hershberger	HRM	Feb 2016	By phone	Buckelew
Lindeberg	GWA	Feb 2016	By phone	Turner
Danielson	GWA	Feb 2016	By phone	Turner
Hopcroft	GWA	Feb 2016	By phone	Turner
PIs that have not met with data	ı manageme	ent team		
Vollenweider				

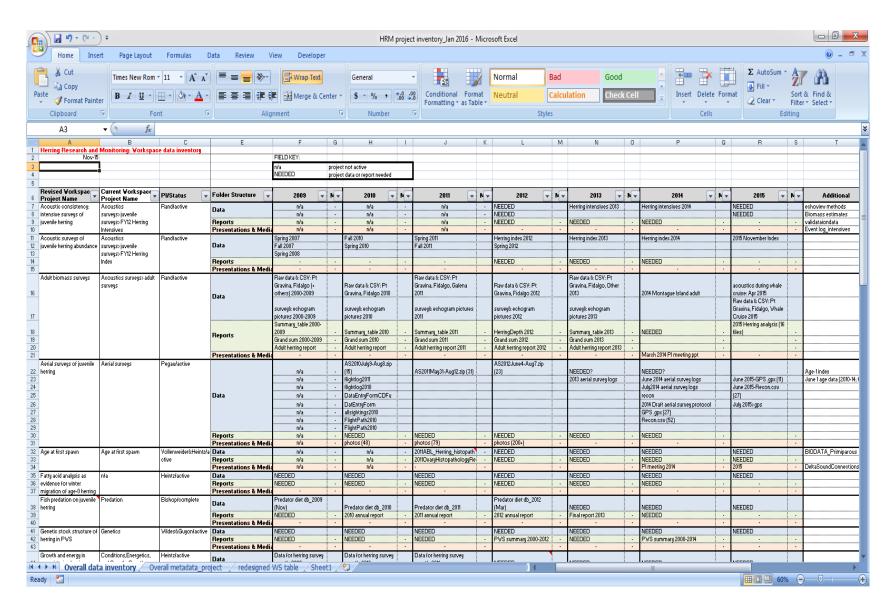


Figure 1. An excerpt of the inventory of submissions made to the HRM Workspace in November 2015. This inventory was presented and discussed with PIs at the Nov 2015 annual program meeting.

The outcome of the meetings was organizing PIs to start writing or make significant progress towards completing a metadata record. In cases, this involved making a plan for how project data files should be structured to assist in describing large data sets or data collections. Information that was already developed for projects, including existing or legacy metadata records, funding proposals, and reports, were utilized to the extent possible. Project titles were adjusted to ensure they were descriptive and included key information for data exploration, including what the data are and where they are located. Additional metadata fields were requested from the PIs to provide details to allow readers to better surmise the data before exploring it further. In cases, this included adding additional metadata fields from the ISO standard format that are not currently recognized in the Workspace metadata editor.

To facilitate continued monitoring of data and metadata submission benchmarks, the Workspace metadata editor was expanded to include a data and metadata file tracking tool for project administrators. This tool eases data management by providing a transparent view of each project's data submissions, metadata record completeness, and data publication to the Portal. The data management team will continue to utilize this tool to monitor the submission progress and maintain regular communications through email, phone, and in-person to assist with metadata authoring.

Objective 2. Implement technical mechanisms to seamlessly transfer LTM and PWS Herring program data from the AOOS data system to systems maintained by DataONE Network.

Task 1: Extend the LTM (Gulf Watch Alaska) data portal to participate in the DataONE network as a DataOne Member Node.

During this reporting period, progress was made in the planning and early development for the Gulf of Alaska portal to become a DataOne Member Node (MN). The feasibility of the data portal becoming a DataOne MN was assessed. AOOS is considered a strong candidate as the long term availability of data and hosting of metadata documents alongside data products already exists within the portal. As such, AOOS has registered as a DataOne MN to begin the implementation. The approach for implementation is currently being planned, which includes a specific, preservation-oriented repository that uses persistent identifications (i.e., digital object identifiers [DOIs]) and "resource maps" to document the relationship between data products and metadata documents in a data package. The implementation work is ongoing and expected date of completion is end of 2016. As part of the implementation planning Axiom met with the National Center for Ecological Analysis and Synthesis (NCEAS) in fall 2015 to begin collaboration regarding the DataONE member node design for the Gulf of Alaska Data Portal. This collaboration will continue through the implementation phase in 2016.

Work Underway

The data management process will continue through the end of 2016 as additional data sets are submitted. The data coordinator, together with the data management team, will review submitted metadata records for completeness and accuracy. Once metadata records have been validated, they will be published to the portal. Metadata disseminated through the portal will improve the discoverability, access, and reuse of the data by a broader audience. One-on-one meetings with PIs will be scheduled again in fall 2016 to revise the metadata records by reviewing them for clarity and omissions. This quality control of the metadata from PIs will ensure records are both understandable and meet standards requirements. Validation will also involve comparing the metadata output to the Federal Geographic Data Committee

(FGDC)/ISO standard for the DataOne portal to ensure the record conforms with the standardized format structure.

8. Coordination/Collaboration: See, Reporting Policy at III (C) (8).

A. Collaboration and coordination both within your program and between the two programs:

This project is focused on increasing the data management support for both LTM and PWS Herring programs by establishing a data coordinator position to improve metadata quality and best practices. Furthermore, this project also develops a mechanism to transfer and integrate LTM and PWS Herring program data products into DataONE. As such, the data management tools and services provided to the EVOSTC LTM and Herring programs are coordinated and collaborative by their very nature. As users of a central data management system, both programs provide useful feedback that informs the features Axiom develops and implements for the Ocean Workspace and the Gulf of Alaska Data Portal. A data management and metadata authoring process are being implemented uniformly across both programs to create a clear organizational structure and standard format. Additionally, by ingesting, synthesizing, and prioritizing feedback and feature requests from both programs, the project team coordinates the needs of each program into a set of tools useful to both. Similarly, by making data from each program available in the Gulf of Alaska Data Portal, the project team helps the two programs collaborate to provide a comprehensive, holistic portrait of the conditions monitored in the Gulf of Alaska by both programs.

B. Coordination with other EVOSTC funded projects:

Based on feedback acquired from the EVOSTC Science Panel and staff, this project was implemented as a supplemental data management effort to execute on major tasks that have been deemed of high importance but are not being addressed by existing data management projects supporting EVOSTC programs (Projects 1412011D and 1412011C). Therefore, all tasks associated with this project are by nature aligned with tasks from the coordinated projects.

C. Coordination with our trust agencies:

The project team provides data management visualization, and preservation services, including providing access to and facilitating the use of the Ocean Workspace, to a number of other programs that receive funding from or are administered or are overseen by representatives from the trustee agencies. Some of these programs and their associated trustee agencies are given in Table 2 below.

Table 2. Collaborating projects and trust agencies

Collaborating Project	Trust Agency
Arctic Marine Biological Observation Network (AMBON)	Bureau of Ocean Energy Management (BOEM)
Arctic Ecosystem Integrated Synthesis (Arctic EIS)	BOEM
Marine Arctic Ecosystem Study (MARES)	BOEM
Integrated Ocean Observing System (IOOS)	National Oceanic and Atmospheric Administration (NOAA)
Beluga Sightings Database Visualization	National Marine Fisheries Service (NMFS)
Alaska Ocean Observing System (AOOS) Data Management	NOAA
Central and Northern California Ocean Observing System (CeNCOOS) Data Management	NOAA
Gulf of Alaska Integrated Ecological Research Program (GOAIERP)	NMFS
Russian-American Long-term Census of the Arctic (RUSALCA)	NOAA
Spatial Tools for Arctic Mapping and Planning (STAMP)	NOAA
Alaska Data Integration working group (ADIwg)	USGS

9. Information and Data Transfer: See, Reporting Policy at III (C) (9).

A. **Publications produced during the reporting period**: None.

B. Conference and workshop presentations and attendance during the reporting period:

The AOOS data team at Axiom Data Science attended the GWA and HRM PI meetings in November 2015, and the team meetings in January 2016 at the Alaska Marine Science Symposium (AMSS). Presentations were given to PIs at both meeting regarding use of the Workspace, Workspace reorganization, data submission, and metadata authoring process. Additionally, the data coordinator team met with individual PIs of the GWA and HRM programs in Anchorage, Homer, Cordova, and Juneau during December and January 2016. Hands-on demonstrations of the AOOS Workspace and Gulf of Alaska data portal were given at this time. Throughout the year, the project team keeps in contact with the GWA program management team with regular email and phone calls.

10. Response to EVOSTC Review, Recommendations and Comments: *See*, Reporting Policy at III (C) (10).

Science Panel 2015 Comments

The possibility of AOOS joining the DataOne system was discussed at the March 2014 Data Meeting as a way to ensure that the data collected as part of the Programs would be available to the widest audience possible. After reviewing the submitted proposal and the budget clarification provided, we would support the funding of the Data Coordinator position and the tasks associated with becoming a DataOne node. The Data Coordinator position should only be funded for the task of preparing the resource maps for data collected as part of the Council funded Programs. We would recommend that the funding of the NODC and OBIS Submission and associated staff time be considered at a later date.

Data Management Team Response

As was described above, in 2015 AOOS (through Axiom) hired the Data Coordinator position. The Data Coordinator together with Axiom Data Science, has worked to inventory what data has been delivered, which PI is responsible for the dataset, the status of data preparation, processing and metadata development. These are the requisite tasks required to prepare HRM data to be published through the DataOne MN. The implementation work, including preparing the resource maps for data collected, is ongoing and expected date of completion is end of 2016.

11. Budget: See, Reporting Policy at III (C) (11).

Please see program budget work book.

Appendix 1. Ocean Workspace- Metadata Quick Reference Guide

This guide is designed as a quick-reference to assist metadata writers in the production of metadata using the Ocean Workspace by providing definitions and examples for the metadata elements.

General notes:

- This version of the Workspace supports two levels of metadata creation: project and file level metadata.
 - O Project level metadata is created for the parent folder (in which associated data files are nested). Project-level metadata is recorded at a broad level for an entire project (irrespective of the techniques used) and covers general project elements, such as project overview, dates, keywords, study species, project details, and geographic location.
 - O The file-level metadata is created for individual files within a project. File-level metadata is typically very specific and metadata provides technical information about the associated dataset, including methods for data collection, instrumentation, data processing, etc. The file-level metadata also provides descriptive information about the encoded dataset.
 - O This version of the Workspace does not support the creation of metadata for nested subfolders.

PROJECT-LEVEL METADATA DEFINITIONS

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
Identification	Title	The name given to the project.	YES	The title should include descriptive elements of: where, what, when.	1.1
	Abstract	An abstract describing the project content (e.g., data). It contains a concise and significant summary of the project data, and is generally intended to serve as a stand-alone description. Coupled with pertinent bibliographic information, it provides users with supportive information for evaluation when conducting a data search. Be sure to include:	YES	The abstract contains generalized statements to convey to the user what the project data are about. It is brief and does not contain specific findings. Its purpose is to acquaint users with the subject content of the data and to help them decide whether or not to consult the original source. In other words, the abstract may be the only text that users search and consult (if they choose not to retrieve the original data). Remembering this may help the writer focus on the key elements and	1.2

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
		 overview statement summarizing why the study was conducted; a short description of data parameters (e.g., summary of what the data are); a brief description of how the data were created; (continued) a general timeline for when the data were collected (by season or month and year); a general location where data were collected. 		select terminology to be included.	
	Purpose	A brief summary of the intentions with which the data or information resource was developed. This statement describes the "why" aspects of the data set (e.g., Why were the data collected?). The Purpose differs from the Abstract in that the latter describes the "what" aspects of the data (e.g., What information is in the data set? etc).	YES	The Purpose should include a summary narrative about: i) what motivated the question (or focus) of the study or the relevancy of data collected; ii) the focal ecosystem; and iii) how the project is associated or contributes to an overarching effort (i.e., assuming project is part of an integrated ecosystem study).	1.3
	Supplemental Information	A comment field where information that is not elsewhere covered can be placed. This item describes information which is deemed unnecessary to include in the abstract, but which is important to further explain the pertinent usage of the	NO	Information relevant and important to the project may be included, such as related studies: additional taxonomic or keywords that are not listed in the respective dictionaries; citations to associated project reports and publications; ancillary files to the dataset (e.g., ReadMe or	1.4

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
		data.		Species List, refer to Taxonomy below); information about specialty equipment used; credit to funders, partners, and/or research affiliates.	
	Time Period	The time period(s) to which the project or dataset corresponds.	YES	The time period should represent the temporal bounds (beginning and end) of the project. This can be a single date, multiple dates, a range of dates, or multiple data ranges. For a project that includes a series of data collection efforts (e.g., cruises) the exact dates of each of these efforts should be specified in the Lineage section.	1.5
	Online Links	The Online Link is intended to provide reference or additional information about the project or data set.	NO	Links are not intended to replace fields in the metadata record because long-term preservation or archiving of the website cannot be assumed. Links that point to archived contextual documents (e.g., those with an associated digital object identifier [DOI] or persistent identifier) would be the exception and are the preferred types of online links to provide.	1.6
	Contact	The individual(s) primarily responsible for creating the project content and who should be contacted with questions. Contact includes name, address, email, and organizational affiliation for those listed.	YES	The contact(s) is typically the lead principal investigator (PI) or author associated with the project. Profile information: first and last name; job title; group or organization names in full. Contact information includes mailing address, phone number, email, and website (if applicable). Information should be completed for all contact	1.7

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
				fields. This will ensure appropriate contacts are made for data inquiry and use. This is particularly important considering PIs may retire or change organizational affiliation over time.	
	Geographic Coverage	The geographic area domain for the project or dataset.	YES	At a minimum, the location name should be provided in the text box and/or a description of the study area bounding box (e.g., 5 transects with 20 sampling points). To provide finer spatial resolution, the bounding coordinates of the study area can be included. These can be created manually using the polygon tool, or exact coordinates imported to create a bounding box of the study area.	1.8
	Keywords	Keywords are words or phrases summarizing the project or dataset. There are two categories of keywords: controlled and uncontrolled. Controlled keywords are terms taken from an established authoritative list (thesaurus) of indexing terms. Uncontrolled keywords are terms applied as free text and are not derived from an established authoritative list. Keywords are used to describe data themes, strata, places, and/or temporality.	YES	In this version of the Workspace, the keyword text box only accepts keywords from the 'science keywords' controlled vocabulary from National Aeronautics and Space Administration's (NASA's) Global Change Master Directory (GCMD). The GCMD science keywords are a standardized, hierarchical set of Earth science keywords. While globally recognized, this directory is not as precise for biological and ocean sciences. We still recommend users to select keywords from the GCMD controlled vocabulary. If keywords must be used to describe data that do not exist in the GCMD, then provide these additional keywords in the Supplemental Information. Insert specialized keywords under a Keywords header using a	1.9

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
				comma-separated list.	
	Taxonomy	Taxonomy is the common-use or Latin name used to describe the subject of the data.	YES, if applicable	In this version of the Workspace, the Integrated Taxonomic Information System (ITIS) is used to search common or scientific names. Results from this directory are reasonably comprehensive, although species recognition limitations may exist for some marine species or taxonomic groups. It is encouraged to utilize the controlled vocabulary for assigning taxonomy to the extent possible. If a taxonomy is not recognized by ITIS, then use the Supplemental Information. Insert taxonomy (using common and Latin names) under a Taxonomy header using a comma-separated list. If there are a large number of taxa (> 30) referenced, consider including a separate csv with the data files that list the common and scientific names for all the species. Reference this csv in the Supplemental Information section of the metadata.	1.10
Lineage lineage of a dataset consists of its entire processing history. This includes its origin (e.g., the source data set, the recording	Statement	This element describes how the data were created (akin to data collection methods) and any data sources used. Lineage is narrative information about the data collection events, parameters, and source data which was used to construct the dataset, and information about the responsible parties.	YES	This summary statement provides information about the events or source data used in constructing the dataset. This information may include: • Events or transformation in the life of a dataset • Information about equipment used • Source data used in creating the data • Spatial reference system used by the source	2.1

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
instrument, the instrument's operating parameters) as well as all subsequent processing steps (algorithms and respective parameters) applied to it.				 Published references for the source data 	
	Processing Steps	Describe the general steps or process used for the data set. For example, How was the data entered? Digitized? Scanned?	YES, if applicable	Processing steps can be a single collective description or individual process steps based upon: stages of processing, incorporation of sources, or project milestone. For each processing step, provide (where possible) the name, date and scale of the source data, a description of the processing steps performed, scanning or digitizing specs, equipment calibrations, software used, tolerances set, etc. This information is hard to remember later, so it's best written down as the data is created.	2.2
Constraints	Access	Any restrictions or legal prerequisites to accessing the actual data set. Commonly applies to data sets that are exempt from public records laws such as endangered species, personal health, and intellectual properties.	YES	Access restraints to data are not common. Examples may include the need to protect the exact location of threatened or endangered species. When there are no access constraints, the words "none" or something equivalent should be written; blank fields are ambiguous.	3.1

Group	Metadata Element	Definition	Mandatory	Notes	Example Reference
	Use	Any restrictions or legal prerequisites to using the data set.	YES	Common constraints include: • Must read and fully comprehend the metadata prior to data use • Acknowledgement of the Contact/Originator when using the data set as a source • Sharing of data products developed using the source data set with the Contact/Originator • Data should not be used beyond the limits of the source scale • The data set is NOT a survey document and should not be utilized as such. When there are no access constraints, the words "none" or something equivalent should be written; blank fields are ambiguous.	3.2

FILE-LEVEL METADATA DEFINITIONS

General notes:

- File-level metadata includes the identical metadata elements for Project-level metadata plus the addition of these below elements.
- File-level metadata is encouraged to be created for data records within the Workspace. At a minimum, file-level metadata is required to be completed for data that will be shared publically in an online portal or published to a national data repository.
- File-level metadata should provide granular information that is specific to the associated file(s) in the collection.
- If file-level metadata applies to more than one file, metadata records can be easily copied and pasted to another file. To do this, click on the file of interest to expand the metadata record view. In the upper right hand corner of the metadata view, select the gear icon. Click "Copy metadata". Then, select the file to where you would like to paste the metadata. Expand the metadata record view. Select the gear icon and click "Paste metadata".

Grou	Metadata Element	Definition	Mandator	Notes	Example
p			y		Reference
Dat a Att rib ute s	Name	The name of the attribute encoded in the file.	YES	This is the name for the associated data column.	4.1
	Definition	A brief narrative description of the attribute. Provide this if your database is not documented in another form such as a data dictionary or data specification manual.	YES, if applicable	Provide a narrative descriptive of what the data is this field describes. If data within this field are encoded provide a list/dictionary for the encoded values here.	4.2
	Measurement type	A categorical descriptor for the	YES	Select from the drop down list the type that	4.3

	measurement scale or type of data: unordered, ordered, relative, absolute, or data/time.		 best describes the data for that attribute: unordered= unordered categories or text ordered= ordered categories relative= values from a scale with equidistant points but no meaningful zero point (e.g. temperature, dollars, etc.) absolute= measurement scale with a meaningful zero point (e.g. wind velocity, wave height, age) date/time= date or time values 	
Reference url	If this attribute is described in detail elsewhere, provide a link to the attribute definition, data specification manual or some other format that describe your data, if applicable. Otherwise leave this element blank.	YES, if applicable	A website link to the document.	-

The purpose of this section is to provide examples for metadata fields found in the Workspace metadata editor. Many of these are examples from research and management agencies that have been collected across different scientific disciplines. If you have any specific questions about what information you should provide for any metadata field or element, contact your data management administrator.

1. Introduction

1.1 Title

Example 1 Title: Conductivity, temperature and depth data for 12 northwestern Gulf of Mexico locations, May to July 2012

Example 2 Title: Bicknell's Thrush Habitat in the Gulf of Maine (BICKHAB83), 1983-1999

Example 3 Title: Northeast Fisheries Science Center 2001 Fall Bottom Trawl Survey

Example 4 Title: Aerial survey data for the assessment of the distribution of cownose rays (*Rhinoptera bonasus*) in the Eastern Gulf of Mexico, from May to October 2008

1.2 Abstract

Example 1 Abstract: Gulf of Maine Habitat Mapping Project for Bicknell's Thrush

The Gulf of Maine Habitat Mapping Project used occurrence information and species/habitat models to map potential habitat for 64 species of primary concern to the U.S. Fish and Wildlife Service between 1983 to 1993. These species include migratory birds, anadromous and estuarine fishes, and threatened or endangered species. The habitat models are based on published literature, agency reports, and knowledge of experts working with the species.

For Bicknell's thrush, the model considered elevation, cover type, and sites of known occurrence. We selected as 'potential habitat' areas with elevations at or over 3000'. Bicknell's thrush occurrences in the Northeast were digitized as point locations. Where 'potential habitat' coincided with observed use, these areas were scored according to the cover type. Point occurrences at lower elevations were buffered 100 m, and these areas also were scored according to cover type. Other areas having elevations at or over 3000' and appropriate cover type, but not known to be used, were scored at 0.5 times the nominal values. (source: US FWS)

Example 2 Abstract: Northeast Fisheries Science Center 2001 Fall Bottom Trawl Survey
The NEFSC bottom trawl survey is a fisheries independent, multi-species survey that provides the
primary scientific data for fisheries assessments in the U.S. mid-Atlantic and New England regions. Two
bottom trawl surveys are conducted each year, one in the spring and one in the autumn. The survey is a
standardized, stratified random design, with stratification based on bathymetry and multiple trawl sites
within each stratum. Trawl sites are selected randomly, but the overall ship path is south to north. The
survey covers the continental shelf and U.S. exclusive economic zone (EEZ) from Cape Hatteras, North
Carolina into the Canadian EEZ. The primary gear is a bottom trawl, with CTD, multifrequency
echosounder, and a host of other scientific sensor data collected ancillary to the bottom trawl catches.
(Source: DOC/NOAA/NOS/OCS > Office of Coast Survey, National Ocean Service, NOAA, U.S.
Department of Commerce)

Example 3 Abstract: Physical Oceanographic Surveys of DeSoto Canyon, Gulf of Mexico 2012 Forty three conductivity, temperature and depth (CTD) casts made from the RV Walton Smith in the northern Gulf of Mexico near DeSoto Canyon in July-August 2012 as part of the Grand Lagrangian

Deployment (GLAD) experiment. These CTD casts were made to determine the hydrography of the upper water column as one way to characterize the meso- and submesoscale variability in the region where 297 CODE-type ocean drifters were launched in an attempt to measure multi-scale near surface dispersion. The drifters were deployed at 1 meter and most of them were launched in triplets (initially separated by roughly 100 meters). This dataset was created by the Consortium for Advanced Research on Transport of Hydrocarbon in the Environment (CARTHE). This research was made possible by a grant from BP/The Gulf of Mexico Research Initiative.

(source: Gulf of Mexico Research Initiative)

1.3 Purpose

Example 1-Purpose: The data provide consultants, planners, and resource managers with information on wetland location and type. The data were collected to meet U.S. Fish and Wildlife Service's mandate to map the wetland and deep water habitats of the United States.

(source: US FWS)

Example 2-Purpose: The NEFSC bottom trawl survey provides fisheries independent abundance and biological data for fisheries assessments along the U.S. east coast.

(Source: DOC/NOAA/NOS/OCS > Office of Coast Survey, National Ocean Service, NOAA, U.S. Department of Commerce)

Example 3- Purpose: These CTD casts were made as part of the GLAD experiment to determine the hydrography of the upper water column as one way to characterize the meso and submesoscale variability in the GLAD experiment region.

(source: Gulf of Mexico Research Initiative)

Example 4- Purpose: This dataset was developed as part of a research project investigating the effects of the Deepwater Horizon oil spill on salt marsh biogeochemistry. This particular project was directed to determine: 1) if the marsh's ability to cycle reactive nitrogen was inhibited (nitrification potential); 2) if there was a significant impact on AOB and/or AOA; and 3) if there were spatial (regional, within marsh) or temporal patterns in nitrification potential.

(source: Gulf of Mexico Research Initiative)

1.4 Supplemental Information

Example 1- Supplemental Information: Chum Salmon Stock Discrimination using Microchemistry

Keywords, theme: Otolith element analysis, chum salmon

Keywords, place: Arctic, Bering Sea, Chukchi Sea, U.S. Exclusive Economic Zone

These data contributed to the following manuscript:

Sutton, T. M., and K. L. Pangle. Regional discrimination of chum salmon in Alaskan waters of the Bering and Chukchi seas using otolith elemental analysis. Deep-Sea Research Part II: Topical Studies in Oceanography.

This is the first version of the Global Subnational Infant Mortality Rates dataset. If you discover any errors or have any issues with the data, please let us know at ciesin.info@ciesin.columbia.edu.

Example 2- Supplemental Information: Data of field activity of 03008 in Puerto Rice trench, Caribbean Sea, 0201802993 to 03-08-2003

Equipment Used - tempsalinometer

Notes - Vessel from NOAA. Related Web Sites: http://oceanexplorer.noaa.gov/explorations/03trench/explorers/explorers.html

Publications -

http://oceanexplorer.noaa.gov/explorations/03trench/explorers/explorers.html

ten Brink, Uri, Danforth, W., Polloni, C.F., Parker, C.E., Uozumi, T., Williams, G.F., 2004, Project PROBE Leg II - Final Report and Archive of Swath Bathymetric Sonar, CTD/XBT and GPS Navigation Data Collected During USGS Cruise 03008 (NOAA Cruise RB0303) Puerto Rico Trench 18 February - 7 March, 2003, U.S. Geological Survey Open-File Report 2004-1400, available on line at: http://pubs.usgs.gov/of/2004/1400/data/oceanography/ctd/ctd.htm

Similar information is available for thousands of other USGS/CMG-related Activities.

If known, available are Activity-specific navigation, gravity, magnetics, bathymetry, seismic, and sampling data; track maps; and equipment information; as well as summary overviews, crew lists, and information about analog materials.

If available, access to physical samples is described in the "WR CMG Sample Distribution Policy" at: http://walrus.wr.usgs.gov/infobank/programs/html/main/sample-dist-policy.html

Primary access to the USGS/CMG Information Bank's digital data, analog data, and metadata is provided through: http://walrus.wr.usgs.gov/infobank/

This page accommodates a variety of search approaches (e.g., by platform, by region, by scientist, by equipment type, etc.).

1.5 Time Period

2011-02-15: single date

2013-06-01; 2013-07-15; 2014-06-10; 2014-07-12, etc: multiple dates

2012-07-01 to 2015-06-30: date range

2012-07-01 to 2015-06-30; 2009-09-05 to 2015-12-03: multiple date ranges

1.6 Online Links

Example 1 (preferred link to persistent identifier): https://search.dataone.org/#view/doi:10.5063/F1TB14

Example 2 (relevant link but not preferred as website may not be archived in the long-term): www. Gulfwatchalaska.org

1.7 Contact

Jane Researcher

Professor of High Esteem University of Somewhere 123456 Apple Road Somewhere, AK 98765 (111) 111-1111 jane.researcher@somewhere.edu www.somewhereuniveristy.com

1.8 Geographic Coverage

West and East Coordinates must be between -180.0 and 180.0 North and South Coordinates must be between -90.0 and 90.0

1.9 Keywords

Aquatic ecosystems, marine habitat, plankton, phytoplankton, zooplankton, primary production, oil spill, ocean temperature chlorophyll

1.10 Taxonomy

Pacific herring (*Clupea pallasii*) Common murre (*Uria aalge*)

2. Lineage

2.1 Statement

Example 1- Lineage Statement: The list of evaluation species was created in a series of steps, starting with a comprehensive survey of species of high national importance occurring within USFWS Region 5. This was developed by combining lists of all federally listed Threatened and Endangered species, 'nongame birds of Management Concern', and waterfowl, shorebirds, anadromous and interjurisdictional fishes (inshore species of concern to the USFWS and National Oceanic and Atmospheric Administration, NOAA), which have significantly and persistently declined in abundance.

The watershed boundary was constructed by selecting the outer boundaries of all smaller watersheds in Maine, New Hampshire, and Massachusetts that flow into the Gulf of Maine. These watersheds were identified using U.S. Geological Survey (USGS) 1:24,000- and 1:100,000-scale hydrology coverages. The boundary was extended into the Gulf at Cape Cod and eastern Maine.

Atwood et al. (1996) developed a habitat model for Bicknell's thrush and found that occurrences were associated with vegetation, elevation, and latitude. Our model applied the same elevation and, as far as possible, cover type parameters. We also integrated sites of known past occurrences. We selected as 'potential habitat' areas with elevations at or over 3000', based on digital contour maps (Maine) or digital raster graphics (New Hampshire). Lower elevations were regarded as likely to be unsuitable. Bicknell's thrush occurrences in the Northeast are listed by mountain name in Atwood and Rimmer (1994), supplemented with information from Tom Hodgman, Maine Department of Inland Fisheries and Wildlife. These were digitized as point locations. Where 'potential habitat' coincided with observed use, these areas

were scored according to their cover type (see

http://r5gomp.fws.gov/gom/habitatstudy/metadata/Bicknell's thrush model.htm). Point occurrences at lower elevations were buffered 100 m, and these areas also were scored according to cover type. Other areas having elevations lower than 3000' and appropriate cover type, but not known to be used, were scored at 0.5 times the tabulated values.

(source: US FWS)

Example 2- Lineage Statement: The source data set provided 467 weekly images of each of the nine regions of the world oceans; these weekly files were averaged in the present data set to provide monthly composite images.

Source citation: Smith, E. 1991 NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data set. Temperature data set produced by the Univ of Miama/Rosenstiel School of Marine Atmospheric Science.

Example 3- Lineage Statement: The geographic area boundaries, names, codes, and the relationships among the various geographic levels are found on Statistics Canada's Spatial Data Infrastructure. These data for administrative areas are updated using information from provincial and territorial sources. These data for statistical areas are updated using the results of the previous census and input from users."

Source citation: The Spatial Data Infrastructure (SDI) is the source for all 2006 Digital Boundary File products.

Example 4- Lineage Statement:

Source Data:

- 1. National High Altitude Program (NHAP) color infrared and black and white aerial photography, 6/1979 - 5/1988, 1: 58000 and 1:80000.
- 2. National Aerial Photography Program (NAPP) black and white aerial photography, 1990-1996, 1:40000.
- 3. Topographic maps, U.S. Geological Survey, 1955-1996, 1:24,000, stable-base material.
- 4. National Wetlands Inventory maps, U.S. Fish & Wildlife Service, 1988-1992, 1:24,000, stablebase material.

U.S. Fish & Wildlife Service Processing Steps:

NWI maps are compiled through manual photointerpretation of NHAP aerial photography supplemented by Soil Surveys and field checking of wetland photo signatures. Delineated wetland boundaries are manually transferred from interpreted photos to USGS 7.5 minute topographic quadrangle maps and then manually labelled. Quality control steps occur throughout the photointerpretation, map compilation, and map reproduction processes.

Digital wetlands data are either manually digitized or scanned from stable-base copies of the 1:24,000 scale wetlands overlays registered to the standard U.S. Geological Survey (USGS) 7.5 minute quadrangles into

topologically correct data files using Wetlands Analytical Mapping System (WAMS) software. Files contain ground planimetric coordinates and wetland attributes. The quadrangles were referenced to the North American Datum of 1927 (NAD27) horizontal datum. The scanning process captured the digital data at a scanning resolution of at least 0.001 inches; the resulting raster data were vectorized and then attributed on an interactive editing station. Manual digitizing used a digitizing table to capture the digital data at a resolution of at least 0.005 inches; attribution was performed as the data were digitized. The determination of scanning versus manual digitizing production method was based on feature density, source map quality, feature symbology, and availability of production systems. The data were checked for position by comparing plots of the digital data to the source material.

(source: USFWS National Wetland Inventory)

2.2 Processing Steps

Example 1- Processing Steps:

LG data for all of Minnesota was downloaded from the NWI ftp server. Wetland codes for each 7.5 minute quadrangle were loaded into a statewide NWI code list from which a unique code number was assigned for each wetland type. Wetland code data from FWS were incorrectly coded into all capital letters on the following 100K sheets: Bigfork, Duluth, Ely, Milbank, Pokegama Lake, Vermilion Lake and Willmar. These data were changed into the correct upper and lower case codes. The DLG files were translated into ARC/INFO double precision net (polygon/line) and point coverages and the Minnesota unique wetland code number was moved into the data set. (nwi2arc.aml). Labelerrors in the net covers were cleaned up if any existed.

The coverages were then snapped to an existing 7.5 minute quadrangle coverage and corner tics were added to create a seamless data base. Additional locational attributes were added and projection information copied into each coverage (nwiproc.aml). Coding and positional discrepancies between 7.5 minute quadrangles were identified and fixed (nwiatt.aml).

7.5 minute quadrangles in Iowa and Canada that contain small areas of Minnesota NWI data were merged into adjacent 7.5 minute quadrangles. The following 100K sheets have such data: Austin, Albert Lea, Hallock and Cavalier. The quads from Charles City and Mason City that were merged into Albert Lea and Austin quads are 4827-4836.

Final NWI data was summarized by type (point, line and polygon), projection information added and the files were exported for archive purposes (nwiexp.aml). The data was transformed into single precision shifted NAD27 coverages for use in PC ARC/INFO. The data was also projected into double precision NAD83 coordinates, but the 7.5 minute quadrangle frame still has the NAD27 boundary.

Staff either at LMIC or DNR converted the data to shapefile format for posting on the DNR Data Deli.

Example 2- Processing Steps: The source data set (Smith 1991) provides 467 weekly images of each of the nine regions of the world oceans; these weekly files were averages in the present dataset to provide monthly composite images.

Calculate monthly averages and composite monthly averages. The included C-language programs sum.c and combine.c were used to calculate the monthly and weekly average SST files. For each grid cell in the images, sum.c calculates the arithmetic average of the corresponding cell in the input files for each month or week of

the year. Results are written to a set of intermediate files which are interpreted by combine c. The program combine decodes the intermediate files written by sum and writes each average image into a new files.

Create GIF and PICT images of month and weekly averages. The C-language program mrletoppm.c converts a monthly or weekly average file into a portable pixmap. GIF and PICT images were derived from these pixmpas using the freely available PbmPlus toolkit developed by Jeff Poskanzer.

Source reference Smith E. 1991. A user's guide to the NOAA Advance High Resolution Radiometer Multichannel Sea Surface Temperate data set. Internal report, 10 p.

Example 3- Processing Steps: This dataset has a simple version of observed velocities, useful for most purposes. The complete data sets with all configuration and processing details and diagnostic data (e.g. error velocity, AGC, spectral width) are available from the NODC Joint Archive for Shipboard ADCP, or by request from S. Pierce. Processing steps included: editing of the data using various diagnostics, calibration of the phase and amplitude errors of the ADCP/navigation/gyrocompass system by covariability analysis between currents and ship velocity, reference layer velocity smoothing, and final production of earth-referenced velocities. For more details regarding methods, see: Pierce et al. (2000), DSR II 47, 811-829. (Source: DOC/NOAA/NOS/OCS > Office of Coast Survey, National Ocean Service, NOAA, U.S. Department of Commerce)

Example 4- Processing Steps: Octopus were collected from commercial fishers during regular fishing operations targeting Pacific cod using pot gear. Octopus were assessed for condition and placed in tanks on board the fishing vessels. After a period of seventy two hours or less they were transported to the Kodiak Laboratory either via a tender vessel or the fishing vessel. Octopus were placed in individual tanks upon arrival at the laboratory. Within a 48 hour period a detailed assessment of the condition of each octopus was conducted, the gender of the octopus was determined, and each octopus was weighed. To weigh individual octopus, they were removed from their tanks, excess water was released from the mantle, and the octopus were weighed using standard bench top scales. Octopus were held for twenty one days; during this period they were fed herring to satiation two times per week. After 21 days, another detailed assessment was conducted.

(Source: Discard mortality for the giant Pacific octopus in the Gulf of Alaska, 2014-15, NPRB Project 1203)

Example 5- Processing Steps: AVHRR Binary Flat Files were loaded into SeaDAS. SeaDAS is a comprehensive image analysis package for the processing, display, analysis and quality control of all SeaWIFS data products. It also displays AVHRR Binary Flat Files and many other data products. The file, mbari.lut, is used as the colormap for the images. Processing scripts include: musesst.pro for bulk processing, musebit1.pro, musebit2.pro, and musebit3.pro for bit shifting.

(source: MBARI Upper Water Column Science Experiment)

Example 6: Sediment samples were washed on a 0.062 mm sieve to separate the foraminifera from the silt and clay. Foraminifera were picked from the fraction retained on the sieve and individually identified and counted with a binocular microscope using reflected light.

(Source: USGS Benthic Foraminifera samples)

3. Constraints

3.1 Access

Example 1- Access Constraints: None

Example 2-Access Constraints: CIESIN offers unrestricted access and use of data without charge, unless specified in the documentation for particular data. All other rights are reserved.

Example 3- Access Constraints: While every effort has been made to ensure that these data are accurate and reliable within the limits of the current state of the art, NOAA cannot assume liability for any damages caused by any errors or omissions in the data, nor as a result of the failure of the data to function on a particular system.

NOAA makes no warranty, expressed or implied, nor does the fact of distribution constitute such a warranty.

Example 4- Access Constraints: Not to be used for navigation. Although these data are of high quality and useful for planning and modeling purposes, they are not suitable for navigation. For navigation, please refer to the NOS nautical chart series.

3.2 Use

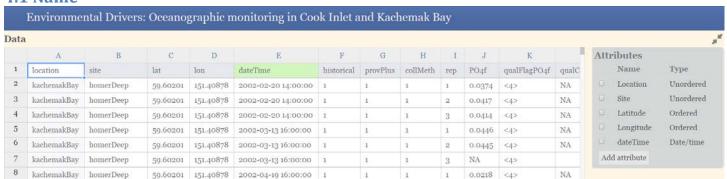
Example 1-Use Constraints: None

Example 2-Use Constraints: The Wildlife Conservation Society (WCS) and Trustees of Columbia University in the City of New York hold the copyright of this dataset. Users are prohibited from any commercial, non-free resale, or redistribution without explicit written permission from WCS or CIESIN. Users should acknowledge WCS and CIESIN as the source used in the creation of any reports, publications, new datasets, derived products, or services resulting from the use of this dataset. WCS or CIESIN also request reprints of any publications and notification of any redistributing efforts.

Example 3- Use Constraints: There are no restrictions on the use of this data. However, secondary distribution must be accompanied by this documentation. Credit should always be given to the data source when this data is transferred or printed.

4. Data Attributes

4.1 Name



In the above example, the attribute name for column A is 'location, column B 'site', column C 'lat', etc.

4.2 Definition

Example 1- Name= Location; Definition= region where sampling occurred

Example 2- Name= Site; Definition= Homer or Seldovia, surface or deep mooring

Example 3- Name= Length; Definition= The fish length from the tip of the nose to the tip of the longer lobe of the caudal fin. Measured in mm.

Example 4- Name= Otter Behavior; The predominant behavior of the animal at the time of observation.

R= resting

T=traveling

G= grooming

F= foraging

4.3 Measurement Type

Unordered= unordered categories or text (statistically nominal)

Examples: Male/Female; Homer/Seldovia; Site A/Site B

Ordered= ordered categories (statistically ordinal)

Examples: Low/High; Surface/Mid-water/Bottom

Relative= values from a scale with equidistant points (statistically interval)

Examples: 12.2 degrees Celsius;

Absolute= measurement scale with a meaningful zero point (statistically ratio)

Examples: 273 Klein; 5.4 kg; 217 mm

Date/time= date or time values from the Gregorian calendar

Examples: 2012-10-24

2015-04-23