1. PROJECT INFORMATION

GOAIERP Project Number:	G86
Title:	Gulf of Alaska Integrated Ecosystem Research Project (GOAIERP) Data Management
Overall project duration	April 2012 – April 2015
Overall project funding	$499,997
Report period	April 1, 2012 to September 31, 2012
Report submission date	Dec 1st, 2012
Lead Author of Report*	Rob Bochenek

Principal Investigator(s), Co-Principal Investigators and Recipient Organization(s):

Molly McCammon, Alaska Ocean Observing System, 1007 W Third Ave., Suite 100, Anchorage, AK 99501, 907-644-6703, mccammon@aoos.org

Rob Bochenek, Axiom Consulting and Design, 3424 Robin St., Anchorage, AK 99504, 907-230-0304, rob@axiomalaska.com

2. PROJECT OVERVIEW

a. Briefly (4-5 sentences) describe the core purpose of your project, and the underlying need for this research.

This project supplies the NPRB GOAIERP research effort with critical data management support to assist study teams in efficiently meeting their objectives and ensuring data produced or consolidated through the effort is organized, documented and available to be used by GOAIERP investigators and future research efforts. This effort coordinates and shares costs with several existing data management projects that are parallel in scope to the data management needs of the integrated program, as well as leverages the Alaska Ocean Observing System's (AOOS) cyber infrastructure and existing data management capacities. The project team supports data submission, metadata generation and data transfer among study teams by using a stand-alone GOAIERP account through the AOOS Ocean Workspace, a web-based project level data management system. Axiom data analysts and domain experts will review/audit metadata and data structure formats produced from GOAIERP project activities and advise study team members in best practices for data formats and metadata authoring.

b. State the specific GOAIERP hypothesis or hypotheses that your project is addressing.

This project supports all GOAIERP hypotheses by managing the research data produced from GAOIERP efforts.

c. List the specific objective(s) of your research project.

1. Support data management communication, coordination and implementation for GOAIERP project.
2. Provide core data management oversight and services for GOAIERP project, including data ingestion, qa/qc, metadata generation, transfer and sharing of data among project PIs, and final transfer of all data to NPRB and appropriate national data archives.
3. Develop tools to access, query, analyze and visualize information relevant to and produced by GOAIERP PIs that meet the needs primarily of the research team, and secondarily, the needs of the larger scientific community, resource managers and the general public.

3. PROGRESS SUMMARY

a. Provide a table showing the timeline and milestones for the current reporting period only.

FY12 3rd Quarter	(April 2012 – June 2012)
April	Project authorized to begin
April	Release AOOS Ocean Workspace (Project DM System)
April	Meet with NPRB staff and GABI (project kickoff and scoping activities)
April	Set up user profiles and projects for PIs in GOAIERP Workspace
May	Begin transfer of 2011field season datasets into GOAIERP Workspace
FY12 4th Quarter	(July 2012 – Sept 2012)
July	Initiate ingestion of retrospective data sets
September	Audit organizational structure of project information in the workspace
FY13 1st Quarter	(October 2012 - December 2012)
October	Isolate revisions to GOAIERP Workspace for 2013 field/cruise season
November	Deploy first round of Workspace updates
November	Rob Bochenek and Will Koeppen attend GOAIERP planning meeting
November	Submit annual report to NPRB

b. Describe report period progress.

Objective 1)
Investigators have facilitated the data management component for the GOAIERP program since project onset (April 2012), which included participation in PI and GABI meetings, establishment of project profiles within the AOOS Ocean Workspace and several webinars to train GOAIERP PIS on Workspace use. These activities have greatly increased the transparency of data and sampling/analysis efforts for the entire GOAIERP Program.

Objective 2)
Investigators have held several webinars and in person meetings to train and assist GOAIERP researchers in posting and organizing data sets and other electronic information to the AOOS Ocean Workspace. An initial scheme for project organization was established by the GABI early in the project. This initial scheme was revisited in October 2012 and modified to address the need to organize datasets into logical scientific efforts versus GOAIERP components. In addition, a data release schedule was developed by NPRB and Axiom staff to document specific dates for PI submission. This schedule has been enforced by Will Koeppen at Axiom.

Objective 3)
The capabilities of the AOOS Ocean Workspace have been augmented throughout this 6-month reporting period through the direct efforts of this project. The user interface has been overhauled several times for usability. Project tagging has been implemented to assist in project organization. File-level metadata was implemented near the end of this reporting period. Additionally, investigators facilitated an exercise to isolate functional improvements to the Workspace by establishing and engaging the GOAIERP Data Management Working Group.

c. Describe preliminary results.
The primary results produced by this project include the further cultivation and enhancement of the AOOS Ocean Workspace's capabilities and the acquisition and documentation of GOAIERP PI-produced data sets.

The Workspace is a centrally managed web-based data management system built on open-source software and focused on addressing the acute needs of scientists working together across organizational and disciplinary boundaries. This approach has four main benefits: first, it ensures that multiple copies or versions of software will not have to be maintained; second, it is operating-system and platform independent, which is beneficial for both maintainers and users; third, improving the Workspace for one group of users improves it for everyone; and fourth, there are no licensing or subscription fees. The Workspace is built upon several horizontally-scaling technologies to provide high availability, redundancy, and performance to support and scale to a growing user base. During this 6-month period the Workspace has been developed to have the following capabilities:

Secure group, user, and project profiles — Users of the Workspace have a password protected user profile which is associated with one or more disciplinary groups or research programs. The current interface allows users to navigate between various groups through a simple drop down control. Transfer of data and information occurs over Secure Socket Layer (SSL) encryption for all interactions with the Workspace. Since the Workspace is already enabled to utilize the Google applications authentication system, users who are already logged into their Google applications account (e.g., Gmail, Google Docs, etc.) can launch the Workspace without the need to re-authenticate.

Metadata Authoring — The Workspace provides investigators with an intuitive interface to author descriptive and rights-management metadata for projects in the Workspace. The metadata elements currently available to researchers in the Workspace come from the Federal Geographic Data Committee (FGDC) designed Content Standard for Digital Geospatial Metadata (CSDGM) and the FGDC endorsed Biological Data Profile of the CSDGM. AOOS engineers also developed an integrated FGDC biological profile extension editor that allows users to search the ~500,000 taxonomic entities of the Integrated Taxonomic Information System (ITIS) and rapidly generate the biological profile component of the FGDC metadata record. Most importantly, the Workspace operates as a cloud-based service and is accessed through a web browser interface, allowing researchers to move between computers during the metadata generation process in addition to allowing team members and administrators to simultaneously review and edit metadata in real time. Late in this reporting period, Axiom engineers implemented file level metadata authoring compliant with the ISO 19115 standard for geospatial metadata, extended to include the metadata content from CSDGM Biological Profile.

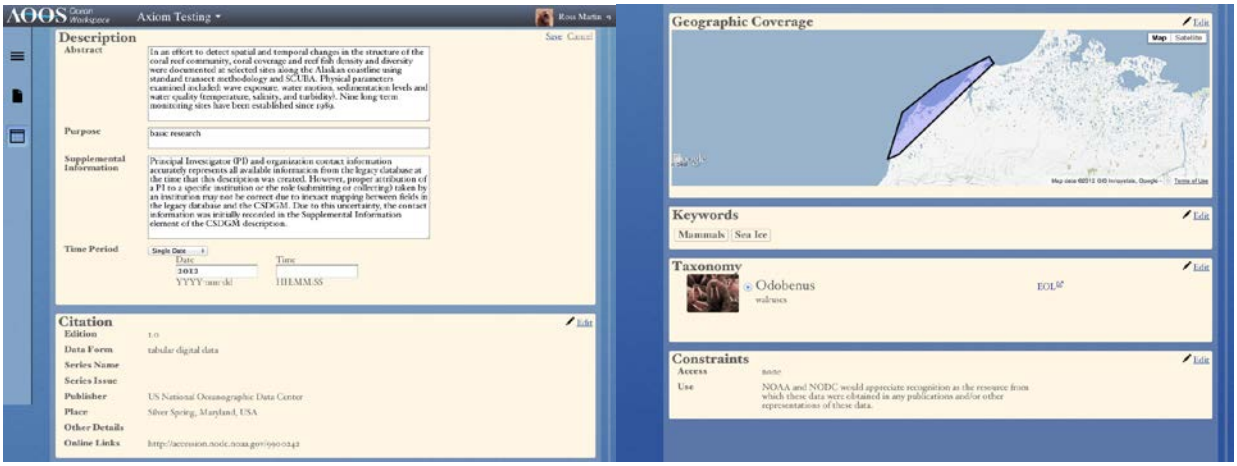

Figure 1. Screen captures of the Workspace metadata interface. The left capture shows the interface to author basic descriptive and citation metadata fields. The right capture displays a tool that allows researchers to define the geographic extent of the project, keywords, FGDC biological profile information and data constraints.

Advanced and Secure File Management — The core functionality of the Workspace is the ability to securely manage and share project-level digital resources in real-time with version control. The Workspace functions as a working platform for sharing preliminary data sets and other digital resources among researchers and study teams. Users of the Workspace are provided with tools that allow them to bulk upload files, organize those documents into folders or collections, and organize projects with predefined and user-created context tags. The Workspace also has the ability to track file versions: if a user re-uploads a file of the same name, the Workspace displays the most current version, but stores and provides access to past versions as well. Ongoing development work will allow researchers to fine-tune read-and-write permissions for individuals within projects, beyond the group level.

Figure 2. Screen captures of project and file management in the Workspace. The left screenshot shows a list of projects to which the example user has access rights. The right screenshot displays the interface a PI would use to organize independent files into collections or folders, and the versioning capabilities of the Workspace: if a file is uploaded with the same name into the same folder, the Workspace tracks the individual versions.

d. Describe integration activity.

The technology being developed and utilized for data management of the GOAIERP program is designed to help facilitate the integration of data sets across disciplines and researchers within the GAOIERP program. Teams and investigators are able to access each other's datasets in a seamless fashion.

e. Describe any concerns you may have about your project's progress.

Software engineering is an involved progress and extremely labor intensive. A concern of the investigators regarding this project stems from being able to keep up expectations from the GOAIERP researchers and programmatic staff.

f. Poster and oral presentations at scientific conferences or seminars

Several demonstrations of the workspace have been given to a wide variety of users including GOAIERP PIs.

g. Education and outreach

N/A

4. PROGRESS STATUS

Investigators feel they have made headway and progress on all objectives of the GOAIERP DM Project. The workspace has proven itself to be a platform that can be easily utilized by GOAIERP researchers and something that they find useful and effective. Compared to the legacy Microsoft Sharepoint system, usability has greatly increased in addition to the flow of information from and to project teams.

In October 2012, GOAIERP DM staff and the Data Working Group composed of GOAIERP PIs developed the "GOAIERP Workspace Functional Improvement Plan" to isolate functional improvements to the workspace and to develop a preliminary timeline for the implementation of such functionality. As of this writing (December 18th, 2012) several of the improvements detailed in the plan have been implemented according to schedule but several have proved to be more complex than initially assessed. Specifically, systems to capture and manage contact information across projects and files have proved to be more complex than originally envisioned. As a result, the milestone has been pushed back and has also had an effect on the implementation schedule for other functional improvements.

5. FUTURE WORKPLAN and DATA DELIVERY

Current through end of project Tasks, Assignments, Timeline

What	*Who*	*Start and end dates*	*Other key dates*
Continue to support the transfer and documentation of GOAIERP data sets. Auditing and restructuring/reorganizing.	**Rob Bochenek William Koeppen**	October 1 2012 – April 2013	
Continue to cultivate and support the functional capabilities of the AOOS	**Ross Martin Lance Finfrock**	October 1 2012 – April 2013	

Ocean Workspace to address GOAIERP researcher needs			
Ability to see who uploaded each file in the Workspace	**Ross Martin**	October 1 2012 – October 15 2012	
Associate contact information projects and files	**Ross Martin**	October 1 2012 – December 31 2012	
Ability to move folders between projects	**Ross Martin**	November 30 2012 – Jan 21, 2013	
Ability to modify file names within the Workspace	**Ross Martin**	November 30 2012 – Jan 21, 2013	
Ability to create file-level metadata (November 30th, 2012)	**Ross Martin**	October 1 2012 – November 30th, 2012	
Ability to tag files and datasets	**Ross Martin, Jim Mitchener**	November 30 2012 – February 15, 2013	
Ability to search for files across projects	**Ross Martin, Jim Mitchener**	December 30 2012 – March 1 2013	
Have more projects viewable on small screens	**Ross Martin,**	October 1 2012 – October 15 2012	
Ability to flag issues with individual files (comments)	**Ross Martin, Jim Mitchener**	Jan 15 2013 – March15 2013	
Ability to have folders within folders	**Ross Martin, Jim Mitchener**	Jan 15 2013 – April 15 2013	

DATA DELIVERY:

This section is not applicable to this project. The project supports the data delivery component of the GOAIERP program.