FY12 INVITATION
PROPOSAL SUMMARY PAGE

Project Title: Collaborative Data Management and Holistic Synthesis of Impacts and Recovery Status Associated with the Exxon Valdez Oil Spill

Project Period: FY12-FY16

Primary Investigator(s): Matthew B. Jones and Mark Schildhauer, National Center for Ecological Analysis and Synthesis (NCEAS), UC Santa Barbara, Molly McCammon and Rob Bochenek, Alaska Ocean Observing System (AOOS), and W. Scott Pegau (PWSSC).

Study Location: General Spill Affected Area

Abstract: The AOOS-led Long-Term Monitoring (LTM) and the PWSSC-led Herring Research and Monitoring (HRM) programs propose an ambitious monitoring and research agenda over the next five years. These efforts could facilitate a more thorough understanding of the effects of the oil spill if the new data and information on the spill-affected ecosystems are effectively managed and collated along with historical data on these systems, and then used in a comprehensive synthesis effort. We propose a collaboration among NCEAS and the AOOS LTM and HRM teams to help build an effective data management cyberinfrastructure for proposed monitoring efforts and organize these data with historical data, including previous EVOSTC-funded efforts, to prepare for synthesis and ensure all data are organized, documented and available to be used by a wide array of technical and non-technical users. Building on the LTM and HRM syntheses and modeling efforts and the 20-year historical data from EVOSTC projects and any available current data, NCEAS would convene two cross-cutting synthesis working groups to do a full-systems analysis of the effects of the 1989 oil spill on Prince William Sound and the state of recovery of the affected ecosystems.

Estimated Budget:

NCEAS budget described here of $1,590,748 total over 5 years

 Approximate Subtotal for Data Management: $796.2K

 Subtotal for Synthesis: $794.5K

NCEAS Funding Requested:
FY12: 409.7K, FY13: 432.4K, FY14: $335.1K, FY15: $346.6K, FY16: $66.8K

LTM/HRM Data Management, Synthesis, and Modeling (described and budgeted separately in prior proposals at $1,840K)

Date: 9-6-2011

PROJECT PLAN

I. NEED FOR THE PROJECT
A. Statement of Problem

In the two decades following the *Exxon Valdez* oil spill (EVOS), and after extensive restoration, research, and monitoring efforts, it has been recognized that full recovery from the spill will take decades and requires long-term monitoring of both the injured resources and factors other than residual oil that may continue to inhibit recovery or adversely impact resources that have recovered. Monitoring information is valuable for assessing recovery of injured species, managing those resources and the services they provide, and informing the communities who depend on the resources. In addition, long-term, consistent, scientific data is critical to allow us to detect and understand ecosystem changes and shifts that directly or indirectly (e.g. through food web relationships) influence the species and services injured by the spill.

An integrated monitoring program requires information on environmental drivers and pelagic and benthic components of the marine ecosystem. Additionally, while extensive monitoring data has been collected thus far through EVOS Trustee Council-funded projects as well as from other sources and made publicly available, much of that information needs to be collated and assessed holistically to understand factors affecting individual species and the ecosystem as a whole. Interdisciplinary syntheses of historical and ongoing monitoring data are needed to answer remaining questions about the recovery of injured resources and impacts of ecosystem change.

Data collected prior to and in response to the Exxon Valdez oil spill are profoundly heterogeneous. They range from long-term, automated sensing of oceanographic and atmospheric conditions, to short-term, experimental, monitoring, and behavioral studies of biological components of the system. The scientific data to be collected in these studies includes data on population trends, behavior, physiology, disease, and genetics of many species, as well as oceanographic and meteorological data at both regional and local scales. This diversity of data and data collection protocols substantially complicates data management by EVOSTC long-term monitoring projects. In addition, investigators on both the long-term monitoring and herring population studies are affiliated with many different institutions and agencies, each currently collecting data from many sites within the spill region and managing it within the frameworks dispersed among these agencies. Any data management system will necessarily need to accommodate this heterogeneity and dispersion by preserving the original data and providing mechanisms to access, integrate, and analyze the data for crosscutting synthesis. Data management activities for oceanographic information occur in isolated, physically distributed agencies, leading to low cross-agency utilization of data. Technical barriers, complex data formats, a lack of standardization and missing metadata have limited access to data and made the utilization of available scientific information cumbersome and daunting. As a consequence, existing data is underutilized and often has not undergone quality assurance.

In this proposal, we outline the collaboration between the National Center for Ecological Analysis and Synthesis (NCEAS), the Alaska Ocean Observing System (AOOS) and their partner Axiom Consulting, and the investigators of the pending Long Term Monitoring (LTM - proposal submitted by McCammon et al.) and Herring Research and Monitoring (HRM –

proposal submitted by Pegau et al.) programs (see Figure below). This project will augment the expertise in data management and synthesis of these groups to maximize the efficiency of data collection and management for the LTM and HRM programs and expand access to these data, collate additional historical data that are useful for synthesis from the EVOS affected area, and conduct a broad-ranging synthesis of twenty years of EVOSTC funded research data to generate a comprehensive assessment of ecosystem impacts and recovery status for the spill affected area.
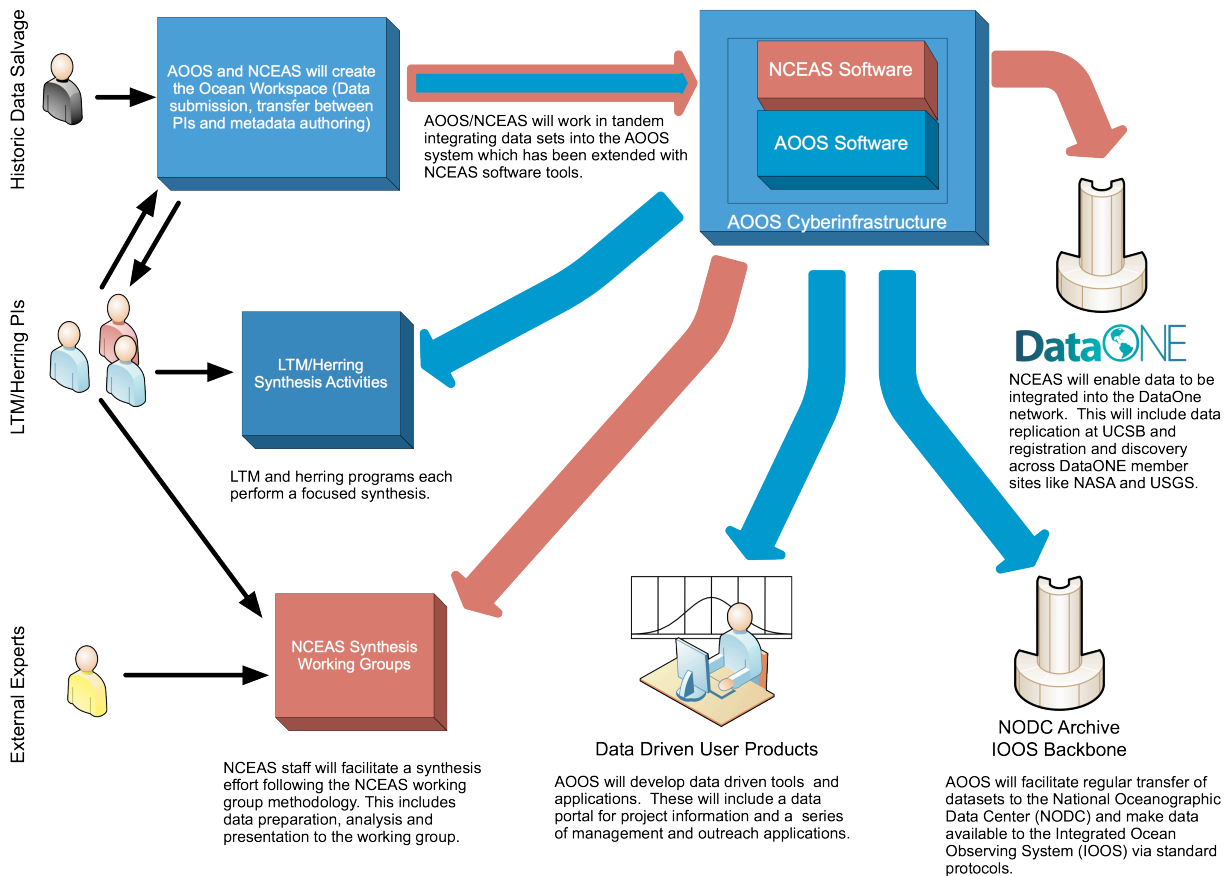

Figure 1. Conceptual description of AOOS/NCEAS/PWSSC collaboration on data management and synthesis activities.

This collaboration document augments the data management, infrastructure development, and synthesis activities previously proposed by the AOOS partners with additional objectives that introduce new technologies from NCEAS to jointly improve the data management infrastructure available to researchers, broaden the scope of data collation and integration, and embark on an ambitious synthesis plan (Figure 1). During the first two years, NCEAS will focus on mining historical data and contributing to development of both the AOOS cyberinfrastructure and the DataONE Federation infrastructure in order to create the necessary data resources for synthesis; during years 3-5, NCEAS will conduct a multi-year working group effort using LTM and HRM principal investigators (PIs) and other internationally renowned researchers to synthesize what is known about spill effects and recovery of ecosystems. These activities will be interwoven with the complementary but distinct data management, technology development, and analysis

activities previously proposed by Axiom and AOOS and which are referenced in the objectives below.

B. Relevance to 1994 Restoration Plan Goals and Scientific Priorities

The LTM and HRM program proposals outline the relevance of the proposed monitoring, data management and syntheses efforts to the EVOSTC 1994 Restoration Plan goals. This project will further support Restoration Plan priorities for "strategies that involve multi-disciplinary, interagency, or collaborative partnerships" and for efforts that will "include a synthesis of findings and results, and will also provide an indication of important remaining issues or gaps in knowledge" (Restoration Plan p. 16). This proposed data management and synthesis collaboration builds on the LTM and HRM programmatic efforts and leverages an additional collaboration with the DataONE federation.

II. PROJECT DESIGN
A. Objectives

1) Provide data management oversight and services for project team data centric activities that include data structure optimization, metadata generation, and transfer of data between project teams (AOOS lead, with contributions from NCEAS).

2) Consolidate, standardize and provide access to study area data sets that are critical for retrospective analysis, synthesis and model development (AOOS and NCEAS).

3) Develop tools for user groups to access, analyze and visualize information produced or processed by the LTM and Herring Research efforts (AOOS lead, with contributions from NCEAS).

4) Organize, integrate, analyze, and model the 20-year historical data from EVOSTC-funded projects and other monitoring in the spill area in preparation for synthesis (under LTM and HRM programs and in NCEAS working groups) (NCEAS lead with AOOS contributions).

5) Integrate all data, metadata and information products produced from this effort into the AOOS data management system for long-term storage and public use (AOOS lead).

6) Augment AOOS/IOOS preservation and interoperability system with other data systems through integration of DataONE services (NCEAS lead).

7) Conduct additional broad synthesis activities on spill impacts and recovery as part of whole-ecosystem analysis through NCEAS working groups (NCEAS lead with AOOS and PWSSC contributions).

B. Procedural and Scientific Methods

Objective 1. Provide data management oversight and services for EVOS LTM and HRM project team data centric activities that include data structure optimization, metadata generation, and transfer of data among project PIs and between project teams.

Details of these efforts are provided in the individual detailed project descriptions for the data management components included in both the LTM and HRM projects. Because project level data is so heterogeneous in nature and is composed of a wide variety of observational types (see Table 1 in LTM data management proposal, which details an initial effort by the AOOS data management team to assess the characteristics of individual LTM data collection activities), a broad range of data management approaches are needed to manage the data in an automated, standard fashion and to facilitate integration. In addition, the project Principal Investigators (PIs) need both flexible and powerful tools to assist them in sharing, archiving and documenting their research products. AOOS data management staff will provide the primary support for these efforts with the AOOS Ocean Workspace, a web-based platform for PIs to post and share data sets and rapidly author metadata. The system will be enabled with security authentication in order to temporarily limit access to LTM and HRM investigators, project managers and administrators before data are quality controlled; non-sensitive data will be publicly released after quality processing. The system will also provide PIs with tools to generate metadata profiles that comply with national standards. Initially, this system will focus on authoring FGDC metadata formats including tools for authoring the biological extension for taxonomic classifications and measurements.

NCEAS engineers will work with the AOOS data team to extend the AOOS data infrastructure to incorporate additional metadata tools and catalogs that are customized for project-based data management for biological data. The design will include both tools for data access and for data contribution and management by the participating scientific staff. The planned AOOS Ocean Workspace (based on non-proprietary open-source standards endorsed by the national Integrated Ocean Observing System) will be enhanced with more biologically-oriented data management tools in order to enable individuals to describe and deposit all of their heterogeneous data in a uniform data repository. Many tools for biological data management, such as metadata generation tools (e.g., Morpho), data analysis tools (e.g., R, Matlab), and synthesis tools (e.g., Kepler) have been developed in parallel to oceanographic tools in use by IOOS; NCEAS will incorporate these tools as appropriate into AOOS systems such as Ocean Workspace, and where that does not make sense, provide interoperability solutions that allow the appropriate tools to work with the AOOS infrastructure (see Objective 6 below). In addition, the heterogeneous data collected by the LTM and HRM projects necessitates a sophisticated data search and discovery system that is effective across data from historical and current LTM and HRM projects. NCEAS will build on their prior work in this area to create a Smart Semantic Search Service that will be deployed as part of the AOOS infrastructure.

This integration of tools from NCEAS contributors into the AOOS cyberinfrastructure will be conducted after a thorough design review and cyberinfrastructure development plan is jointly assembled by AOOS and NCEAS as part of the initial needs and solutions assessment.

Objective 2. Consolidate, standardize and provide access to related and historic data sets that are critical for retrospective analysis, synthesis and model development within the LTM and HRM programs.

This task will involve isolating and standardizing historic data sets deemed necessary for retrospective analysis by EVOSTC LTM and HRM program synthesis and modeling efforts. Early in the effort the EVOSTC LTM and HRM program researcher teams will be engaged to prioritize sources of relevant data deemed of high value for the synthesis effort. Data will be prioritized by several metrics including its utility to LTM and HRM program syntheses as well as system-wide synthesis efforts (Objective 7), accessibility of the data, length of time series, scientific importance, quality and precision of the data storage format, and the cost of obtaining the data (digitization can be expensive). All data acquired through efforts of this project will be merged into the AOOS data system for long term archival and access.

LTM PIs have already developed a preliminary list of historical data sources under their stewardship which could be of potential value to the LTM program and synthesis effort (see Table 2 in LTM data management proposal), as well as those data PIs would be interested in getting access to are currently unaware of sources (Table 3 in LTM proposal). AOOS funding leverages numerous data sets available through the AOOS website and data system, including the herring and PWS ecosystem data sets that were standardized and made available through the actions of the PWS Herring Portal Project (EVOS Project 070822, 080822 and 090822).

Although data capture will be a collaborative effort, we expect to roughly divide activities into three focal sets of data: 1) LTM and HRM data sets that are newly collected under these projects (AOOS focus); 2) Other EVOSTC project data sets, both current and historical, that lay outside of the LTM and HRM projects (NCEAS focus); and 3) external data sets from other funding groups (joint NCEAS and AOOS focus depending on source).

Objective 3. Develop tools for user groups to access, analyze and visualize information produced or processed by the LTM and HRM efforts.

AOOS will take the lead on these efforts, as described in the data management DPDs for the LTM and HRM programs. The AOOS data team will work with project investigators to develop web-based data driven tools based upon prioritization and direction from agency managers, outreach staff and user groups. Effective data summarization and visualization exposes problems, manifests trends, and allows for high-level comparisons with other sources of information. Data visualization products are also ideal tools to communicate information to audiences with varying degrees of familiarity in meaningful and easily understandable ways. NCEAS will provide input and expertise into development of these tools.

Objective 4. Organize, integrate, analyze, and model the 20-year historical data from EVOSTC-funded projects in the spill area in preparation for LTM and HRM program and NCEAS working group synthesis efforts.

The current AOOS plan is to emphasize the capture of historical data from previous studies related to the Exxon Valdez oil spill during the first two years of the project, as well as to prepare the system to receive the monitoring data generated during this project. NCEAS will collaborate with the AOOS team in order to collate, summarize, visualize, and integrate these historical data

in order to prepare them for synthesis and analysis. NCEAS has developed a group of scientific programmers who specialize in assisting in cross-cutting analysis and modeling, and we will employ one of these scientific programming specialists along with a graduate student assistant to collate, standardize, integrate, summarize, and visualize the data needed for synthesis activities. Digital, graphical and visualization products generated by NCEAS from the 20-year historical datasets will be used for the cross-cutting synthesis activities of the year three EVOSTC joint workshop between the LTM and HRM programs and for the broader EVOS impact syntheses described in Objective 7. Products from these activities will include: data summaries and visualizations from each of the prioritized EVOSTC data sets; quality assurance analyses on input data to resolve issues prior to analysis; integrated data products that resolve methodological differences to combine multiple related primary data sets into long-term, cross-scale derived data products; and analyses of these derived products that illustrate long-term, cross scale aspects of spill impacts and recovery. These activities will build upon the LTM and HRM program synthesis and conceptual ecological modeling efforts focused on the monitoring program data. Please see the detailed project descriptions on LTM synthesis (Holderied), LTM ecological modeling (Hollmen), HRM synthesis (Pegau), and HRM modeling for additional information, as well as the synthesis activities in Objective 7 regarding cross-cutting synthesis efforts.

Objective 5. Integrate all data, metadata and information products produced from this effort into the AOOS data management system for long-term storage and public use.

The ultimate goal of this project is to provide services to assist in the organization, documentation and structuring of data collected and made available via EVOS LTM and HRM project activities so that it can be transferred efficiently to long term data archive and storage centers and made available for future use by researchers and other user groups. This task will leverage the AOOS cyberinfrastructure, long-term funding and other active data management projects being undertaken by that organization. Data sets produced from the integrated research effort will be served to users by extending existing data access, analysis and visualization interfaces currently supported and under development by the AOOS data management team. AOOS systems have the capabilities to ingest, archive and serve model output, remote sensing and real time/archived sensor data streams, and, as of fall 2011, ingest and archive GIS and project level data. AOOS is currently developing a mirror site in Portland, OR to ensure long-term security of its data and software. In addition, AOOS has prioritized working with state and federal agencies to ensure long-term access and archiving of agency data and information products.

Objective 6. Augment AOOS preservation and interoperability system with other non-IOOS data systems through integration of DataONE services.

NCEAS will augment the capabilities of the AOOS data system by incorporating the services that are part of the DataONE data federation[1]. These include open services for writing data and metadata, controlling access to data products as they are populated in the system, and services for replication and preservation of data. By using the DataONE service framework, this will also link the AOOS and IOOS system to the DataONE federation, which includes partners such as the U.S. Geological Survey, Knowledge Network for Biocomplexity and NASA Distributed Active

---

[1] http://dataone.org

Archive Centers. This broader federation will be critical in other stages of the project, especially for access to satellite data during synthesis and analysis activities.

In addition, DataONE services include a comprehensive, cross-institutional data preservation model that involves mirroring of data at multiple DataONE participating institutions and continuous active monitoring to ensure data remain valid and that adequate replication is present even in the event of institutional failures. In this project, we will establish the AOOS Asset Catalog as a Member Node in the DataONE network, and thus be able to replicate all EVOS data to DataONE partner institutions to ensure longevity, accessibility, and validity of EVOS data. Funding for these replicas will largely be supported through storage already available on the DataONE network (approximately 1.2 petabytes available for replication), although exceedingly large data sets (above ten terabytes) will need to be discussed.

Objective 7. Conduct broad synthesis activities on EVOS impacts and recovery as part of whole-ecosystem analysis through NCEAS working groups.

Since 1995, the National Center for Ecological Analysis and Synthesis (NCEAS)[2] has been advancing the state of ecological and environmental knowledge through synthetic and collaborative research that aims to discover general patterns and principles based on existing data. The premise at NCEAS is that many decades of data have been collected that can be synthesized to produce novel insights into important scientific and societal issues, and that the expertise and information resources necessary to accomplish these syntheses are latent but distributed throughout the science community. To promote effective synthesis of environmental data, NCEAS has sponsored and executed more than 450 working groups over 15 years, many of which have had major scientific and policy impacts (e.g., changes in habitat conservation plans for endangered species, and creation of marine reserve initiatives based on scientific principles). Sociological studies of the working groups in action at NCEAS have demonstrated major shifts in the culture of synthesis in ecology and gains in collaborative productivity via the working group model at NCEAS (Hackett et al. 2008).

 Despite decades of monitoring and analysis of EVOS-affected systems, there is still a major lack of understanding of oil spill impacts and recovery at a holistic level. Many of the studies to date have been at the single species level, and recovery status is tracked on a case-by-case basis. In addition, because all of the historical data have never been fully integrated, it has been impossible to conduct a holistic analysis of the effects of the oil spill and recovery of impacted regions. Such a holistic view is critical to guide future monitoring and recovery initiatives, which are expected to continue for decades. NCEAS and PIs from the LTM and HRM programs will conduct two holistic synthesis activities aimed at understanding the long-term, ecosystem-wide consequences of EVOS and the effectiveness of recovery initiatives:

- Synthesis Working Group: Assessing Ecosystem-wide, Long-Term Impacts from the Exxon Valdez Oil Spill
- Synthesis Working Group: Understanding Ecosystem Recovery following the Exxon Valdez Oil Spill

---

[2] http://www.nceas.ucsb.edu

The first will address system-wide impacts from EVOS, and the second will specifically focus on an assessment of recovery of affected systems and reasons for recovery successes and failures that will assist in future recovery initiatives. As detailed below in methods, the products from these syntheses will include a series of reports and academic papers supported by synthesized data, archived models and analyses, and archived model outputs.

These syntheses will build upon the more focused efforts to be conducted by the LTM and HRM programs. For example the working group on *Understanding Ecosystem Recovery* will benefit from the efforts to understand the recovery of an individual species (herring), but expand upon that to include other species including those in the LTM program. It will also provide an opportunity to further explore the connections between environmental variables to the recovery of herring and other species. Because the working group approach takes a more holistic approach than the individual species approach proposed by the HRM program we expect that in answering the question of *Understanding Ecosystem Recovery* we will provide new findings that will guide the LTM and HRM programs in the future.

C.   Data Analysis and Statistical Methods

Data Management and Infrastructure Methods
The overarching strategic plan for the AOOS data system is described in detail in both the LTM and HRM data management detailed project descriptions. It involves implementing an end-to-end technological solution which allows data and information to be channeled and distilled into user-friendly products while simultaneously enabling the underlying data to be assimilated and used by the emerging external data assembly systems. AOOS will lead the development of this system, with NCEAS contributing to the design and implementation, particularly in areas where dealing with data heterogeneity is paramount, such as semantic search. The system has four tiers: 1) data, models and metadata; 2) interoperability systems which facilitate data search, query and delivery; 3) an asset catalogue and Smart Semantic Search Services; and 4) user applications that are web-based. The intended result is the facilitation of rapid data discovery, improved data access, understanding, and the development of knowledge about the physical and biological marine environment. This system meets all the standards of the national Integrated Ocean Observing System.

The asset catalog developed by AOOS will provide an index of all project data and provide direct connections to other Alaska data systems as well as those of the national Integrated Ocean Observing System and Global Ocean Observing Systems. The analysis and synthesis activities described in this proposal however, will also need access to a much broader set of data available not only from AOOS and IOOS, but also from other federated data systems such as NASA's Earth Science Data Information System (ESDIS) and the Earth Observing System Clearinghouse (ECHO). NCEAS engineers will work with the AOOS data team to enhance the AOOS asset catalog, in particular by linking it to the DataONE federated catalog, thereby providing access to non-IOOS data, such as MODIS and other satellite data managed by DataONE Member Nodes. This linkage will require NCEAS to extend AOOS data systems to be compatible with the interoperable web services framework used by DataONE. Current and emerging AOOS web services will be harmonized with DataONE services to allow applications to connect to the asset catalogue and get access to the underlying descriptions of all known data sources. Thus,

EVOSTC data will be directly incorporated at the national and global scales into both the IOOS oceanographic data network as well as other data federations via DataONE, thereby greatly expanding agency and public access. When complete, all data deposited in the AOOS system will also be replicated to participating DataONE member nodes, which are continuously monitored for availability and integrity to enable long-term data preservation.

Due to data heterogeneity, data discovery is difficult for complex, multidimensional and cross-disciplinary data that will be collected by the LTM and HRM program research teams. The AOOS system incorporates a metadata authoring tool that includes extensions for biological metadata. In this project, NCEAS and AOOS will expand on that system and build **Smart Semantic Search Services** that understand the scientific content of data to improve the effectiveness of data searches. The NCEAS team has pioneered a semantic scientific observations model that allows scientists to precisely discover measurements of interest and subset data to only include observations relevant to their studies. NCEAS developed the Extensible Observations Ontology (OBOE; Madin et al. 2008) to enable semantic search and access services that facilitate much higher precision and recall than have been possible with traditional metadata-driven systems. We will incorporate these semantic search services into the AOOS Tier 3 asset catalog, and help to develop the catalog so that semantic markup of data on ingest is easily accomplished. Thus, in addition to managing information about data availability and access methods, the asset catalogue will also contain ontologies that map source data descriptions and metadata to a common set of internally stored terms with strict definitions. This mapping will allow users to easily locate related sets of information without having explicit knowledge of the internal naming conventions of each data-providing agency. The development of an internal ontology will also enable future endeavors to connect the asset catalogue to global ontologies in the semantic web. Because the asset catalogue contains a semantic definition of data sources and maps all known data sources to a common definition, applications can be developed which connect users to vast arrays of data through simple but powerful interfaces.

Collaborative Synthesis and Analysis Methods
Two working groups consisting of LTM and HRM program PIs with additional nationally renowned scientists will undertake a broad synthesis of the 20-year data set from EVOSTC-funded projects and other spill area monitoring to improve our assessment of impacts and recovery associated with the EVOS:
- Synthesis Working Group: Assessing Ecosystem-wide, Long-Term Impacts from the Exxon Valdez Oil Spill
- Synthesis Working Group: Understanding Ecosystem Recovery following the Exxon Valdez Oil Spill

The working group syntheses will build on and expand programmatic syntheses conducted under the proposed LTM and HRM programs.

NCEAS has an extensive history of convening highly productive synthesis activities through its use of a working group model, involving face-to-face meetings and ongoing virtual collaboration supported by the Center (Hackett et al. 2008). Under this successful NCEAS model, committed working group participants conduct relevant analysis and modeling on a continuous basis for approximately two years, punctuated by periodic working meetings to come to consensus and drive further work by participants.   The momentum of the group is maintained by postdoctoral

fellows, funded by this proposal, that reside at NCEAS, working on the group's analysis, modeling, and other synthesis tasks while being able to take advantage of the computational and analytical support services available at NCEAS. Working Groups are composed to represent a wide variety of scientific expertise, including both scientists that are closely involved in the problem at hand, as well as researchers from adjoining disciplines that help broaden the scientific perspective of the group. In addition, Working Groups typically include a mix of more senior scientists and younger scientists that are eager to dive into the required analysis and modeling activities. Although all travel expenses are paid for by the project, Working Group participants serve voluntarily on these working groups, making the activities especially cost effective.

To initiate these Working Group activities, NCEAS will organize and constitute the groups during year two, and working group activities will commence in year 3. Working Group leaders will be selected for their knowledge of the issues at hand as well as their ability to effectively motivate a group of up to 14 other working group participants. We would expect that many of the PIs from the LTM and HRM programs would be participants in the synthesis working groups along with nationally renowned experts in population and community modeling, ecosystem modeling, and coupled whole-system analysis. In addition, because NCEAS is already running a working group on ecotoxicology associated with the BP Deepwater Horizon spill[3], we would expect significant coordination and cross-pollination with these new EVOS synthesis groups.

Based on the preparatory data analysis and modeling conducted to assemble and integrate the 20-year historical data set with available current data from the LTM and HRM program syntheses (see Objective 4), NCEAS will work with the leaders of the "Assessing Impacts" and the "Understanding Recovery" working groups to outline an initial set of goals and deliverables for each of the two working groups. At a minimum, each group will produce a comprehensive synopsis report of results from analysis and modeling of the impacts and recovery in the historical and current data that will be written into a series of papers targeting both the science and management communities. The groups will also provide input to the LTM and HRM program teams on recommendations for evolution of the EVOSTC-funded monitoring efforts beyond the initial 5-year programs. All analyses, models, results, and data backing these conclusions will be published alongside these papers in the spirit of open science and to maximize reproducibility of the results (see the previous NCEAS Global Marine Impacts[4] synthesis for an example of this type of output). The actual synthesis activities and products will be selected by working group participants and driven by the data analysis and modeling to maximize working group effectiveness and the relevance of their products. However, example synthesis activities might include cross-scale analysis of the relationship between oceanographic processes and the recovery of forage fish; meta-analysis of the relationship between extent of injury and extent of recovery for organisms crossing taxonomic groups (e.g., mammals, birds, fish, plankton); and, performance of forecasting of cross-trophic recovery scenarios in light of observed population trends.

D. Description of Study Area
The study area for this project will include the entire EVOS spill affected area. The north, east, south, and west bounding coordinates of this area are 59.767, -145.837, 61.834, and -154.334

---

[3] Anderson, Cherr, and Peterson; Ecotoxicology of the Gulf Oil Spill: A holistic Framework for Assessing Impacts
[4] http://www.nceas.ucsb.edu/globalmarine

E. Coordination and Collaboration with Other Efforts

We propose to integrate the efforts in this project as an additional part of the multi-disciplinary "Long-Term Monitoring of Marine Conditions and Injured Resources and Services" program proposal submitted by McCammon et al. to the EVOSTC. The project represents a collaboration among AOOS, NCEAS, and the other LTM and HRM science project PIs both for individual program data management and in developing syntheses that connect individual project results.

Regarding the data management aspect, AOOS brings extensive experience with creation, collation, and access to extensive oceanographic (physical, chemical and biological) data throughout Alaska, as well as a variety of visualization tools and products for resource managers and marine stakeholders. Its initial focus has been on serving up real-time sensor and remote sensing data and forecast models. A new application in October 2011 will include the ability to query, discover and access project level and GIS data sets. In addition, AOOS brings a significant level of leveraged resources, regional data management projects and partnerships to this effort, which could not be accomplished for the budgeted amount without these leveraged resources. These include funded projects for the Alaska Ocean Observing System's Ocean Data Portal, the Prince William Sound Science Center, Northern Forum/USFWS Seabird Data System, the Alaska Department of Fish and Game, and the Cook Inlet Regional Citizens Advisory Council.

NCEAS brings complementary skills to the data management efforts. They have extensive expertise in cyberinfrastructure systems for synthetic environmental science (c.f., Reichman et al. 2011, Jones and Gries 2010, Jones et al. 2006). NCEAS has developed software systems supporting long-term data preservation and sharing, is a leader in metadata systems for science data, and is a progenitor of the DataONE[5] interoperability framework to create a global data federation for open access to scientific data. NCEAS' focus on project-level data management for highly heterogeneous data allow the management of current and legacy data that are critical to synthesis but that often are not captured by large-scale agency data systems, such as the EOSDIS program or the IOOS program. Thus, the initiatives at NCEAS for capturing complex but smaller-scale biological and physical data will be an effective complement to the ocean observatory data management systems that are provided by AOOS.

The syntheses efforts of the LTM, HRM, and NCEAS programs are to be synergistic. The syntheses of the LTM and HRM programs are expected to be program focused. The NCEAS working group syntheses efforts will build upon and augment the programmatic syntheses of the LTM and HRM programs by using a larger-scale synthesis of historical and current monitoring data to provide an assessment of the overall ecosystem impacts of and recovery from the EVOS. There is coupling between the LTM and HRM programs in that the environmental factors important to herring survival are primarily collected in the LTM program and herring represent an important factor in controlling the upper trophic level observations of the LTM. However, the collaboration with NCEAS will allow a more holistic view of how the findings of these programs are connected not only to each other, but with other types of research being conducted. The LTM and HRM program syntheses and personnel are expected to be an important resource for the NCEAS efforts to build upon. In turn the NCEAS led efforts will provide new perspectives

---

[5] http://dataone.org

to help guide future LTM and HRM efforts. It should be noted that the success of the NCEAS efforts depends on the participation of members of the LTM and HRM programs because of their intimate knowledge of the ecosystem within the EVOS affected region.

III. SCHEDULE
A. Project Milestones focused on NCEAS Activities

Objective 1. Provide data management oversight and services for EVOS LTM project team data centric activities that include data structure optimization, metadata generation, and transfer of data between project teams.

This objective will be addressed by AOOS and NCEAS throughout the entire span of the project and will follow the annual cycle of field data collection and analysis by principal investigators. NCEAS milestones will include incorporation of project-specific data management tools into the Ocean Workspace and development of Smart Semantic Search Services for data discovery.

Objective 2. Consolidate, standardize and provide access to study area data sets that are critical for retrospective analysis, synthesis and model development.

This objective will be primarily met by AOOS and NCEAS by the fourth quarter of year two of the effort (September 2013). However, AOOS will continue to add data to the system throughout the entire life of the project, and NCEAS will continue to add data as needed by synthesis efforts through year 4.

Objective 3. Develop tools for user groups to access, analyze and visualize information produced or processed by the LTM and HRM efforts.

For AOOS, see milestones in LTM and HRM detailed project descriptions. For NCEAS, analysis and visualization tools that are incorporated into the system will be available at the end of year 2 when other software deliverables are produced.

Objective 4. Integrate all data, metadata and information products produced from this effort into the AOOS data management system for long-term storage and public use.

This objective will be addressed throughout the entire span of the project. The AOOS data system is to serve as the vessel to capture all project level data produced through this effort in addition to those datasets salvaged to inform the historic synthesis effort. This task will be ongoing as long as the program is producing or acquiring additional data.

Objective 5. Provide preservation and interoperability with other non-IOOS data systems through integration of DataONE services.

Initial integration with DataONE will occur in year 1 with a prototype release in Quarter 4, and a final release of DataONE services in year 2 Quarter 4. Once operational, data will continue to be replicated to DataONE as they are produced throughout the span of the project.

Objective 6. Organize, integrate, analyze, and model the 20-year historical data from EVOSTC-funded projects and other monitoring in the spill area in preparation for LTM and HRM program and NCEAS working group synthesis efforts

Historical and newly generated data will be collated throughout years 1 and 2, with integration and modeling of these occurring as they are collated. Data and modeling summaries will be posted in Quarter 4 of year 1, and the complete historical data set will be available in Quarter 4 of year 2. NCEAS working groups will continue to integrate the data used in their synthesis activities with new data from LTM and HRM projects as it becomes available during years 3 and 4.

Objective 7. Conduct broad synthesis activities on spill impacts and recovery as part of whole-ecosystem analysis through NCEAS working groups.

Organization of synthesis activities will begin in year 2, with working group meetings and synthesis activities occurring throughout years 3 and 4. Publications and final analyses and conclusions of working groups will be produced in year 5, but we expect some of the publications in earlier years.

B. Measurable Project Tasks by NCEAS

FY12 1st Quarter (October 1, 11 to December 31, 11)

October	Project authorized by trustee council
October	NCEAS staff hiring and reallocation when funds become available
November	Collaborate with AOOS to initiate historic data aggregation effort
November	Attend LTM program PI meeting
December	Draft historic data set manifest
December	Design and begin implementation of DataONE integration

FY12 2nd Quarter

January	Prioritize historic datasets for inclusion into synthesis efforts
January	Design and begin implementation of data discovery and management tools
February	Begin historic data aggregation effort and integration into AOOS
March	Begin ongoing integration, analysis, and modeling (throughout year)

FY12 3rd Quarter

May	Attend HRM program PI meeting

FY12 4th Quarter

August	Submit input for LTM program annual report
September	Initial analysis results drive FY13 data salvage and integration
September	DataONE integration prototype demonstration

FY13 1st Quarter (October 1, 12 to December 31, 12)

October	Assess/Validate year 1 datasets and metadata submitted through AOOS and NCEAS

November	Participate in LTM program PI meeting and support first LTM conceptual modeling workshop
December	Prototype data discovery and management tools demonstration

FY13 2nd Quarter

FY13 3rd Quarter

May	Participate annual HRM program PI meeting
June	Complete integration of data salvaged into AOOS DM System
July	Full release of data discovery and management tools

FY13 4th Quarter

September	Select synthesis working group leaders, organize WG activities
September	DataONE Integration services released

FY14 1st Quarter (October 1, 13 to December 31, 13)

October	Assess year 2 datasets and metadata submitted to AOOS
October	Finalize user access tool work plan version 1 and initiate development
November	Participate in LTM program PI meeting

FY14 2nd Quarter

Winter	EVOSTC workshop with LTM and HRM programs supported by LTM and HRM synthesis reports and NCEAS historical data synthesis

FY14 3rd Quarter

May	Participate in annual HRM program PI meeting

FY14 4th Quarter

September	Create synopsis of FY14 synthesis WG meetings, draft publications

FY15 1st Quarter (October 1, 14 to December 31, 14)

October	Assess year 3 datasets and metadata submitted through AOOS
November	Participate in LTM program PI meeting

FY15 3rd Quarter

May	Participate in annual HRM program PI meeting
May	Submit input for five-year plan for FY17-22

FY15 4th Quarter

September	Create synopsis of FY15 synthesis WG meetings, draft and submit publications

FY16 1st Quarter (October 1, 15 to December 31, 15)

October	Assess year 4 datasets and metadata submitted through AOOS
November	Continue working on acceptance of synthesis group publications
November	Participate in LTM program PI meeting

FY16 4th Quarter
September   Finalize all synthesis group papers and products

References

Hackett EJ, Parker JN, Conz D, Rhoten D, Parker A. 2008. Ecology transformed: The National Center for Ecological Analysis and Synthesis and the changing patterns of ecological research. Pages 277–296 in Olson GM, Zimmerman A, Bos N, eds. Scientific Collaboration on the Internet. MIT Press.

Jones MB and Gries C. 2010. Advances in environmental information management. Ecological Informatics 5: 1–2. doi:10.1016/j.ecoinf.2010.01.001

Jones MB, Schildhauer M, Reichman OJ, and Bowers S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. Annual Review of Ecology, Evolution, and Systematics. 2006. 37:519–544.

Madin JS, Bowers S, Schildhauer M, and Jones MB. 2008. Advancing ecological research with ontologies. Trends in Ecology and Evolution 23 (3): 159-168. doi:10.1016/j.tree.2007.11.007

Reichman, OJ, Jones MB, and Schildhauer MP. 2011. Challenges and Opportunities of Open Data in Ecology. Science 11 February 2011: 703-705. doi:10.1126/science.1197962