Exxon Valdez Oil Spill
Restoration Project Final Report

PWS Herring Data Portal

Restoration Project 090822
Final Report

Rob Bochenek
Axiom Consulting & Design
523 West 8th Avenue Suite 104
Anchorage, Alaska 99501

May 2011

Exxon Valdez Oil Spill
Restoration Project Final Report

PWS Herring Data Portal

Restoration Project 090822
Final Report

Rob Bochenek
Axiom Consulting & Design
523 West 8th Avenue, Suite 104
Anchorage, Alaska 99501

May 2011

PWS Herring Data Portal

Restoration Project 090822
Final Report

Study History: EVOSTC project 090822 was originally a one year pilot project (070822) that began in spring 2007 and was extended for the 2008 and 2009 fiscal years. Work performed in 2008 is associated with EVOSTC project number 080822. Steve Moffitt, Alaska Department of Fish & Game Division of Commercial Fisheries, was the PI for the project during 2007 and 2008. In 2009 the PI responsibility was transferred to Rob Bochenek of Axiom Consulting & Design. This project suffered a serious impediment during 2008 when Jim Vansant, lead ADF&G data salvage technician, was killed in a construction accident.

Abstract: This project has developed a web-based data portal and consolidated, documented, and entered data sets, metadata, and other electronic resources into that web portal. The efforts of this project have resulted in a considerable increase in available historic time series data regarding herring ecosystem dynamics in Prince William Sound. The web portal provides public access to information, data and GIS visualizations. Scientist and researchers can use the web portal as a resource to assist in consolidating, accessing and synthesizing herring data. This project has also developed an ArcPad application for collecting herring aerial survey data directly into a GIS format. Also produced by this project are a set of technical Standard Operating Procedures (SOPs) for efficiently ingesting, managing and serving ecological datasets in an open, scalable and standardized way. These SOPs were developed in the last phase of this effort and also document the technical architecture of system implementation. The project was conceived during an EVOS-sponsored workshop in April 2006 that was tasked to identify Prince William Sound herring data gaps and develop restoration or research projects to help herring recovery. Participants indicated that knowledge of the spatial and temporal aspects of herring related data sets (e.g., herring spawn) was necessary to understand how restoration activities might affect herring abundance trajectories. This project has improved data quality and provided easier access and visualization of selected herring data sets and other electronic resources.

Key Words: Pacific Herring *Clupea pallasi*, Prince William Sound, Data Management

Project Data:
Data salvaged from efforts of this project are provided in flattened csv files, MS Access databases and .shp files for storing GIS data types.

Data Custodian Rob Bochenek
E-mail address rob@axiomalaska.com
Telephone (907)-230-0304
Address Axiom Consulting & Design
523 West 8th Ave
Suite 104
Anchorage, Alaska 99501

Citation:

Bochenek, R.J. 2010. PWS herring data portal, *Exxon Valdez* Oil Spill Restoration Project Final Report (Restoration Project 090822), Axiom Consulting & Design, Anchorage, Alaska.

TABLE OF CONTENTS

EXECUTIVE SUMMARY.....1

INTRODUCTION 2

OBJECTIVES.....3

METHODS..... 3

RESULTS..... 10

DISCUSSION.....20

CONCLUSIONS.....23

LIST OF TABLES

Table 1. Summary of salvaged data sets standardized by Prince William Sound Data Portal Project. 10

Table 2. Summary of data sets acquired from regional data providers, integrated into the portal but not salvaged by this project activities…….....16

LIST OF FIGURES

Figure 1. Geospatial OLAP data structure utilized in the project in year one.....4

Figure 2. Screen Captures of 4 dimensional visualizations of data stored in the PWS Herring Portal. 6

Figure 3. Diagram portraying the technical components of phase II PWS Herring Data Portal design …….....7

Figure 4. Asset catalogue data structure…….....9

Figure 5. Aerial biomass observations …….....12

Figure 6. Aerial survey effort data set …….....13

Figure 7. Herring spawn observed during aerial surveys …….....14

Figure 8. Herring Age, Sex, & Length (ASL) data set..…….....15

Figure 9. Commercial harvest data set..….....16

Figure 10. Diagram portraying the components of the user interface of the PWS Herring Data Portal ….....17

Figure 11. Aerial Survey Application Interface.....20

EXECUTIVE SUMMARY

This project successfully rescued a series of key herring data sets collected over three and half decades describing herring biomass trajectories, habitat utilization, commercial fishery impacts and age class composition and health. The rescue effort largely focused on data sets primarily collected by or under the stewardship of the Alaska Department of Fish and Game Division of Commercial Fisheries in Cordova. The rescue effort included digitization of tabular paper data, migration and manipulation of legacy digital data formats and spatial enablement of the data through the creation of geospatial components (points, lines and polygons). Once rescued, the data were documented to create FGDC metadata. Additionally, this project focused on developing a data management framework to host the rescued data so that it could be served out to user groups through web-based visualizations and access tools. Both the data management platform and user tool engineering underwent two cycles of development during the project. This iterative process allowed investigators opportunities to assess approaches and technology and make mid-project corrections to best meet the needs of the user community and the researchers involved in the Integrated Herring Research Program. The Prince William Sound Herring Portal can be accessed online at http://dev.axiomalaska.com/PWSHerringPortal/. This project also developed and supported an Arcpad aerial survey application for use in the ADF&G aerial survey monitoring program. Finally, this effort produced a set of Standard Operating Procedures (SOPs), which provide high-level directives for implementing a scalable and flexible data management framework for ecosystem based research.

Development of the PWS Herring Data Portal was completed in year three (2009) of the project. Although the system is currently running off Axiom Consulting & Design computer systems and network infrastructure, this does not preclude the transfer of the portal to a regional entity for long-term maintenance and support, and plans are underway to do so. Potential organizations for long-term hosting of the system could include the Alaska Department of Fish & Game, Alaska Ocean Observing System (AOOS) or the Prince William Sound Science Center (PWSSC).

INTRODUCTION

Much of the herring data that has been collected in the Prince William has been inaccessible to the broader scientific community and general public for some time. Data management activities for herring information occur in isolated, physically distributed agencies with frequent and serious technical barriers to providing effective access for user groups. Complex and antiquated data formats and lack of standardization have made using these valuable information resources a cumbersome and daunting task. The lack of rapid visualization tools has made data exploration difficult; therefore, data has been underused in addition to lacking thorough quality control and quality assurance. Tools and technology to manage spatial scientific data have not been developed in a robust fashion, resulting in low access to and use of the Gulf of Alaska's spatial information resources as they pertain to species distribution, abundance, habitat and physical and chemical metrics.

Emerging data management protocols and mature open source data management systems provide an effective framework for overcoming these barriers to spatial data management and the development of functional, cost-effective data access and visualization tools. Interoperability protocols and open source data-serving software provide an efficient method for sharing and communicating information that is spatially explicit and complex.

OBJECTIVES

1. Consolidate herring data sets, metadata and other electronic resources to a publicly accessible web portal for herring information.
2. Provide web accessible map-based visualizations of geospatially enabled herring data and provide user interfaces for querying and downloading raw data and metadata.
3. Develop Standard Operating Procedures (SOPs), which document system implementation and provide direction for the absorption of additional herring ecosystem data sets, metadata and information to the centralized herring data system.
4. Develop Arcpad application to automate herring spawn data collection.

METHODS

Objective 1: Consolidate herring data sets, metadata and other electronic resources to publicly accessible web portal for herring information.

Existing herring data sets and electronic resources were assessed, salvaged, documented with metadata and centralized to a common access point for distribution and public access. Much of this work was manual and extremely tedious in nature, required substantial staff resources within ADF&G and required a large amount of the funding provided by this project to spatially enable and digitize these datasets. Data sets were documented via the Federal Geographic Data Committee (FGDC) metadata standard. The metadata document will provide critical contextual information for researchers to use, locate and interpret the data set. Data sets will be stored in aggregated standardized file formats and corresponding FGDC metadata will be authored as an XML document. Both the data set and corresponding metadata document will be available for public download and have been submitted to EVOSTC with this report.

ADF&G technicians were responsible for digitizing and aggregating data sets contained within the Cordova ADF&G office. Staff members at Axiom Consulting & Design were involved in directing digitization efforts at ADF&G and dictated intermediary data formats for prepping data for secondary processing. Once data were digitized and aggregated by ADF&G staff, Axiom staff further processed the digitized data files into flattened relational database tables. The data was then integrated into the PWS Herring Data Portal data management system. Finally, FGDC metadata was created to document the salvaged data resource.

In order to efficiently integrate and manage the various data sets necessary to performing the data salvage required by this study, the study team adhered to a standard geospatial data framework. The following section details the specific data management methods that enabled the principal investigators to efficiently manage the diverse data sets required to perform the activities of this project.

Geospatial Data Model

Initially, data integration efforts undertaken by this project used a geospatial data structure developed during 2007 and 2008 to store and manage spatially enabled information. In late 2008 and 2009, this initial approach was abandoned in light of emerging open source interoperability systems, standards and formats. The first initial data model was engineered using Online Analytical Processing (OLAP) data warehousing specifications. This type of data structure is inherently designed to store data and information that is in a static and finalized state. This type of database schema is optimized for analysis and rapid data access. Figure 1 (below) provides a diagram of the OLAP data structure which was initially employed by the study team. As more data sets were salvaged, processed and loaded into the below structure, it became clear that this approach was not flexible enough to store and manage the wide array of potentially valuable data sets (raster, imagery and model outputs) requested by users of the PWS Herring Portal. Additionally, as data were imported into the OLAP structure's fact table the performance of the system did not scale well with increasing data volume. By the end of the project's second year it became clear to PIs that this strategy would not be viable for providing high performance access and data visualization and that a new approach would need to be adopted.

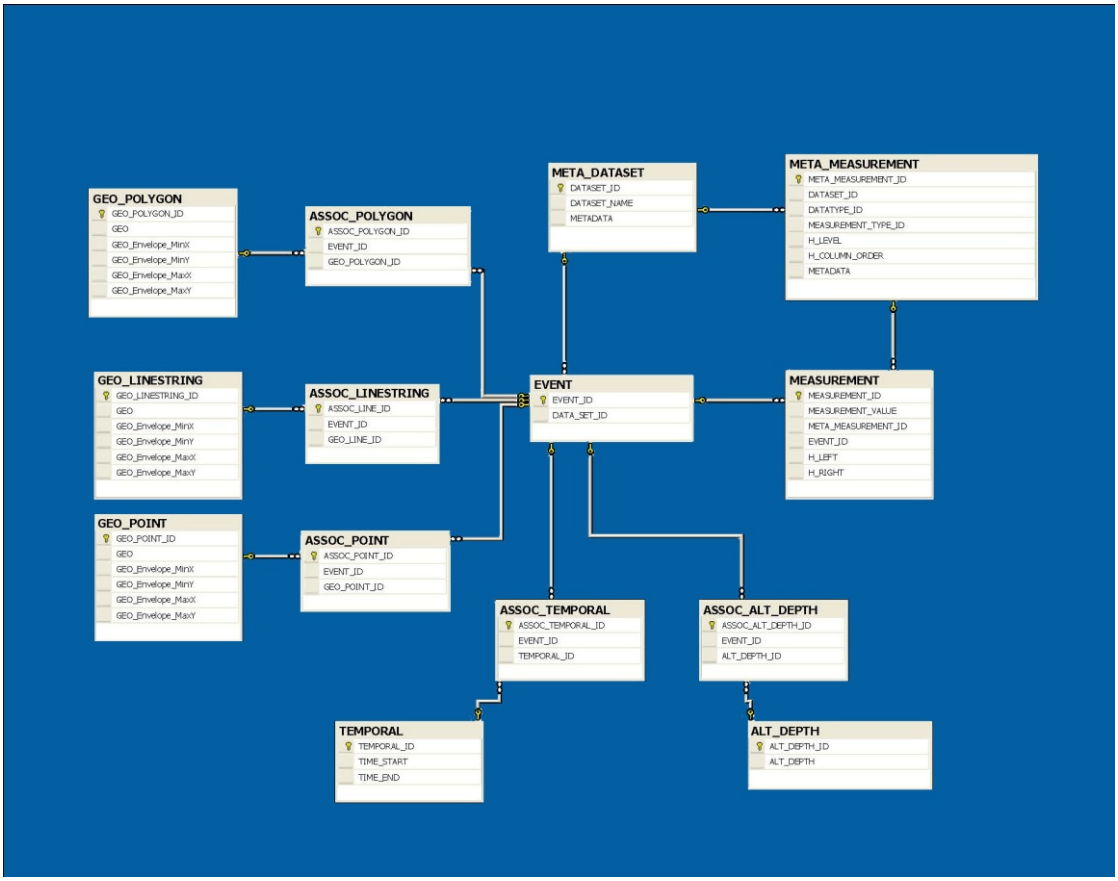


Figure 1: Geospatial OLAP Data Structure abandoned in year two by the study team.

Late in 2008 the principal investigators changed their data management approach to focus on interoperability systems and formats which provided a number of tangible benefits over the previous strategy. The new framework relied upon interoperability concepts to provide an extremely flexible environment for storing a wide variety of data types and formats. The technical foundation of the data management platform relies upon a suite of existing and mature open source software tools and informatics efforts that are actively supported by the open source community and prolifically used by data management experts on a worldwide scale. Consequently, this framework allowed investigators to build upon a constantly maturing set of technologies by simply tapping into the efforts of a large body of programmers and developers who are committed to advancing these technologies. Using open source technologies also enabled the system to scale in a cost-effective manner by allowing additional hardware to be configured to support the growing user base without increases in software licensing costs. Integrating available sources of interoperable data feeds into data access applications and data management systems allowed a variety of data resources to be accessed by the PWS Herring Data Portal at low cost. Large quantities of real-time and historical sensor information, remote sensing satellite information and marine habitat and biological data for the Alaska region are openly available for use through interoperability protocols. For example, NASA Earth Observations (NEO) provides an expansive array of long-term oceanographic, climate and atmospheric remote sensing datasets. Real-time and historical sensor data feeds for the PWS region are available for sensors via SNOTEL, the National Data Buoy Center (NDBC), the Center for Operational Oceanographic Products and Services (CO-OPS) and other NOAA programs. Additional sources of interoperable data include those hosted at NASA's Jet Propulsion Laboratory (JPL), U.S. Geological Survey (USGS) TerraServer and other research organizations.

Data Standards

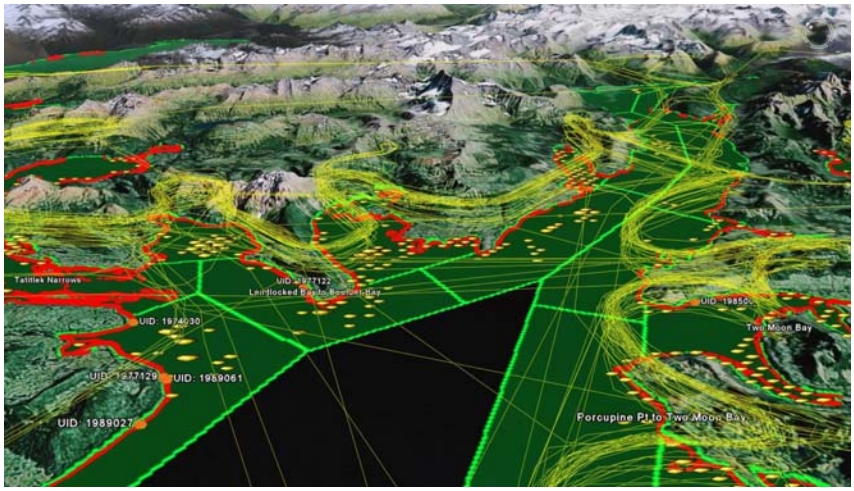
This effort adhered to the following data and metadata standards.

- Open Geospatial Consortium (OGC) Simple Feature Specification
- Federal Geospatial Data Committee (FGDC) metadata standard

This project did not collect measurements; rather, it digitized and geospatially enabled data sets existing within ADF&G in an unorganized state. This involved the creation of FGDC compliant metadata for those resources lacking metadata. Data sets were also up scaled in their data structures to ensure that information contained within the resource could be accessible to and understood by other scientists and the public. Methods such as standard-based naming conventions and normalization of relational database structures were used whenever appropriate. GIS information was structured into a geospatial database following Open GIS Consortium standards to ensure longevity and usability of the geospatial data.

Objective 2: Provide web-accessible map-based visualizations of geospatially enabled herring data and provide user interfaces for querying and downloading raw data and metadata.

A core strategy of the PWS Herring Data Portal was to create web-based access to spatially explicit visualizations of salvaged data sets. Early in this project, Google Earth was chosen as a platform for visualization of data sets to be served off the PWS Herring Portal. These 4-dimensional visualizations allowed users to view spatial and temporal aspects of the data set. Users could filter by location or time span. The following images (Figure 2) provide examples of these types of visualizations using Google Earth.



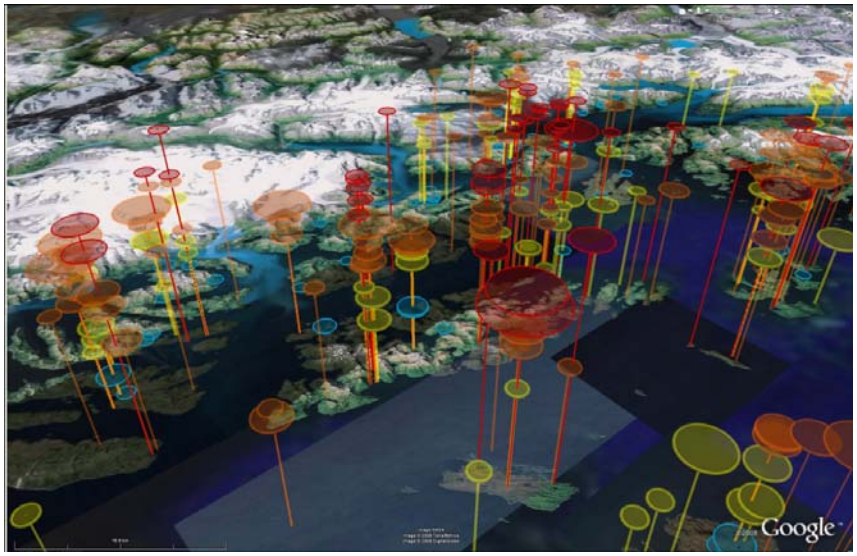


Figure 2: Screen Captures of 4-dimensional visualizations of data stored in the PWS Herring Portal. The first image shows a composite view of the ADF&G Aerial Survey program. The second image displays hydrocarbon densities and sample sizes for oiled mussel populations in PWS.

Data visualizations channeled through Google Earth provide a method for users to rapidly assess and comprehend data sets in their entirety, but this platform only provides data visualization.

After a series of demonstrations to various user groups (scientists, resource managers and the general public), it became apparent that users required both visualization and data access in order for the portal to be truly useful as a data management platform. For this reason, PIs determined in year two of the project that a web-based interface to the underlying data must be developed in order to create tools for users to query and extract data in raw formats. This change in the strategy for developing the PWS Herring Data Portal interface was driven by user feedback to the PWS Herring Portal. The Standard Operating Procedures document included as an appendix to this report goes into greater detail regarding the process of acquiring user feedback, isolating user groups and consolidating feedback into formulating the new technical strategy. The overall response from users was consolidated into the following directives regarding redesign:

1. Data stored in the portal needs to be made available in the following formats when applicable (ESRI Shapefile Output, Excel Files, Microsoft Access Databases).
2. Users need to be able to create and download subsets of information that have been queried by space, time and measurements values.
3. Users will need to have the ability to query across multiple data sets via the same parameters (space, time and measurements).

The data management architecture implemented to power the new user interface is detailed below in figure 3. This new system involved storing spatial data directly in the PostgresSQL PostGIS spatial database.

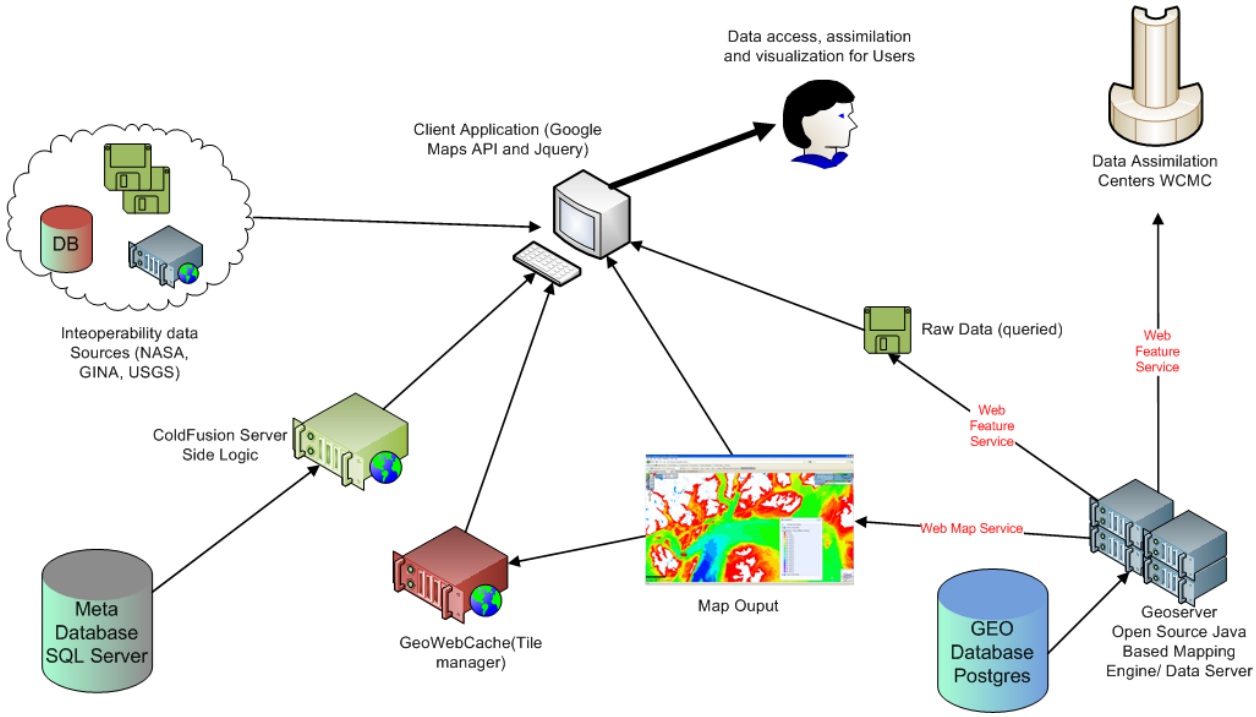


Figure 3: Diagram portraying the technical components of the newly designed PWS Herring Portal.

The following is a description of the technologies referenced in Figure 3 broken into logical information processing tiers. Tier 1 corresponds to data storage (Database), tier 2 corresponds with analysis and aggregation of information (Business Logic) and tier 3 concerns visualization and access tools (Client Interfaces).

Relational Database

- SQL Server 2008 database – SQL Server 2008 is a robust and mature platform for storing complex tabular, relational and spatial information. It is compliant with both ASCI SQL and Open Geospatial Consortium (OGC) specifications, enabling SQL Server to communicate with other enterprise relational databases and commercial and public GIS interfaces such as Geoserver, Mapserver and ESRI products. More information about OGC can be found at http://www.opengeospatial.org/standards. The SQL server was used to store a database of underlying data layers and their associated metadata and dimensional characteristics. This component of the data system is distinct from the GIS data and constitutes the "asset catalogue" which will be further described in this document.
- PostgreSQL – Is an open source relational database which provides advanced functionality to store and analyze spatial information. It is also OGC compliant liker SQL Server. All GIS enabled data layers were stored in the PostgreSQL instance.

Web Services

- OGC Services were handled by Geoserver (www.geoserver.org), an open source java-based server system designed to render web accessible maps and package data into a wide variety of data formats. Geoserver is OGC compliant and works well with many GIS clients (ESRI Arc View, Google Maps, Google Earth and others). The Geoserver instance allows data to be accessed via Web Map Services (WMS) and Web Feature Services (WFS).
- Web Services for the PWS Herring Data Portal user interface were written in Cold Fusion for rapid prototyping. This included the services which powered the user interface tools.

Tier 3 – User Interface

- Data visualization, processing and extraction tools for the users are available via a web-based browser interface using the JQuery javascript library and Google maps code base.

The PWS Herring Data Portal is powered by an internal database (asset catalogue) of data sources and their characteristics. The asset catalogue implemented for the data management framework is a simple system that stores information about available data resources. At the most granular level, the system tracks information about individual data layers including name, description, time span, access methods, service urls and symbology. These layers are grouped into various layer groups and further organized into specific application configurations. Time period information is also stored for data resources with discrete time coverages. A web service library developed by Axiom software engineers provides the application with access to information in the asset catalog through simplified methods such as GetDataCatalog, GetLayers and GetTimePeriods. The use of an intermediate web service library allows multiple applications running in different domains to access the asset catalogue while simultaneously concealing its inner complexities. Several automated import tools have been developed which periodically update the asset catalogue with information about data resources on remote systems.

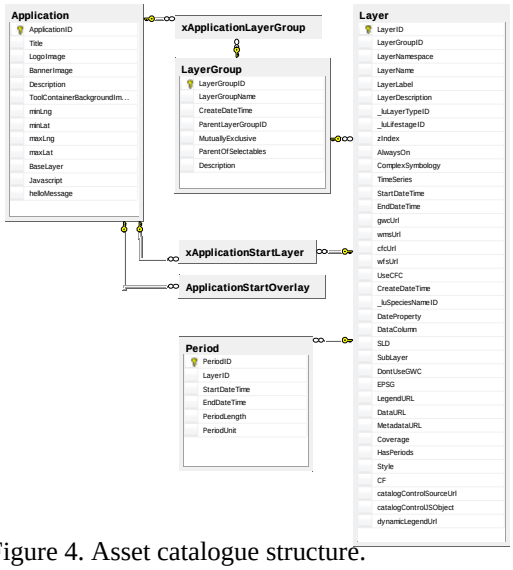


Figure 4. Asset catalogue structure.

Objective 3: Develop Standard Operating Procedures (SOPs) which document system implementation and provide direction for the absorption of additional herring ecosystem data sets, metadata and information to the centralized herring data system.

The Standard Operating Procedures (SOPs) technical document is attached as an appendix to this final report and provides a strategic approach to developing data management tools for ecological and physical data sets as they pertain to an ecosystem based herring monitoring program. The document is driven by the best practices and conceptual understanding that was developed during the project. Additionally, this document details technical underpinnings of the PWS Herring Data Portal (technical documentation). The ultimate goal of this document is to direct future data management efforts to power ecosystem approaches to research. More regarding this topic can be reviewed in Appendix 1.

Objective 4: Develop Arcpad application to automate herring spawn data collection.

PIs were also tasked with developing and supporting an ESRI Arcpad application to facilitate automated data collection of herring aerial survey spawn and biomass observations in addition to ancillary species observations. This primarily involved software engineering and data preparation for GIS base layers to power the Aerial Survey application. The tool underwent

several seasonal revisions (three separate releases over three years) based upon feedback and functionality requests of biologists using the application. The application was updated and overhauled at the beginning of each season based upon functional limitations isolated from the previous season's aerial survey program. ADF&G staff members beta tested the application and were tasked with providing comments which were prioritized and then implemented by Axiom software engineers.

RESULTS

Objective 1: Consolidate herring data sets, metadata and other electronic resources to publicly accessible web portal for herring information.

The following data sets were digitized, salvaged, restructured, documented and loaded into the PWS Herring Portal. The focus of this effort was to standardize data stewarded at the Cordova ADF&G office which has sponsored a continuous 35 year herring monitoring effort. The following data sets should be viewed as a package of data that describe the effort of monitoring activities and what was observed to inform and monitor the long-term health of the herring fishery.

Data Set Name	Start Year	End Year	# of Records Salvaged
Aerial Survey Biomass	1974	2007	5,751
Aerial Survey Effort	1974	2005	19,217
Aerial Survey Spawn	1973	2007	5,348
Herring Age Sex Length (ASL)	1973	2007	187,358
Herring Commercial Harvest	1972	1999	202

Table 1: Summary of salvaged data sets standardized by PWS Herring Data Portal project.

During year one of the project the herring spawn and biomass observational data sets were salvaged and structured into modern data formats. In years two and three the ASL, aerial survey effort and commercial harvest data sets were salvaged, structured into modern data formats and ingested into the PWS Herring Portal. Additional data sets were brought to light during the project as potential candidates for salvage efforts. These additional data sets were acquired in their existing states and assessed based upon criteria that were developed by engaging members of the Integrated Herring Ecosystem Program and ADF&G Division of Commercial Fisheries research staff. Prioritization of effort was based upon the following two criteria:

1. Length of Time Series – Data sets which were continuously collected several years or decades using the same standard operating procedures were deemed to be much more important and valuable than those which had been collected during a single season or sporadically over several years.

2. Quality of Source Data – Data sets missing key information such as spatial/temporal associations and other key information were assessed to be of much lower value than those with explicit well-defined fields and spatial/temporal components. These data sets were also considered to be potentially misleading to synthesis efforts due to their lack of concrete contextual information.

During the project additional datasets were brought to light as potential candidates for salvage activities by the Integrated Herring Research Program investigators. These datasets included information gathered during the EVOSTC sponsored Sound Ecosystem Assessment (SEA) project, EVOSTC APEX project and additional short time series data sets produced by individual projects. Though these projects could provide potential data which would be a value to long term herring monitoring in Prince William Sound, they were prioritized lower for salvage activities than the 35 year time series stewarded by the Alaska Department of Fish & Game. Also, because the ADF&G data was used for management purposes, it resided in a much higher state of quality and standard operating procedures were much better documented than these other datasets. For these reasons, in addition to the funding constraints of the project, PIs focused their efforts on the core ADF&G herring monitoring datasets.

The following provides a basic description of each data set salvaged from this project. More in-depth information describing the data sets is contained in the FGDC metadata record associated with each data set.

Aerial Survey Biomass Observations - This data set describes herring biomass observations in PWS observed during ADF&G Aerial Surveys occurring between the 1974 and the 2007 field seasons. Approximately 5,800 observations of herring school aggregations are represented in this data set. Many records are compound observations, meaning that they represent multiple schools. School observations are based upon an internal Standard Operating Procedure developed by the Cordova ADF&G which classifies school sizes into small, medium, large and super aggregations which correspond to the following short ton measurements respectively (10 tons, 40 tons, 60 tons and 350 tons). Data were digitized using ESRI Arc View Software.

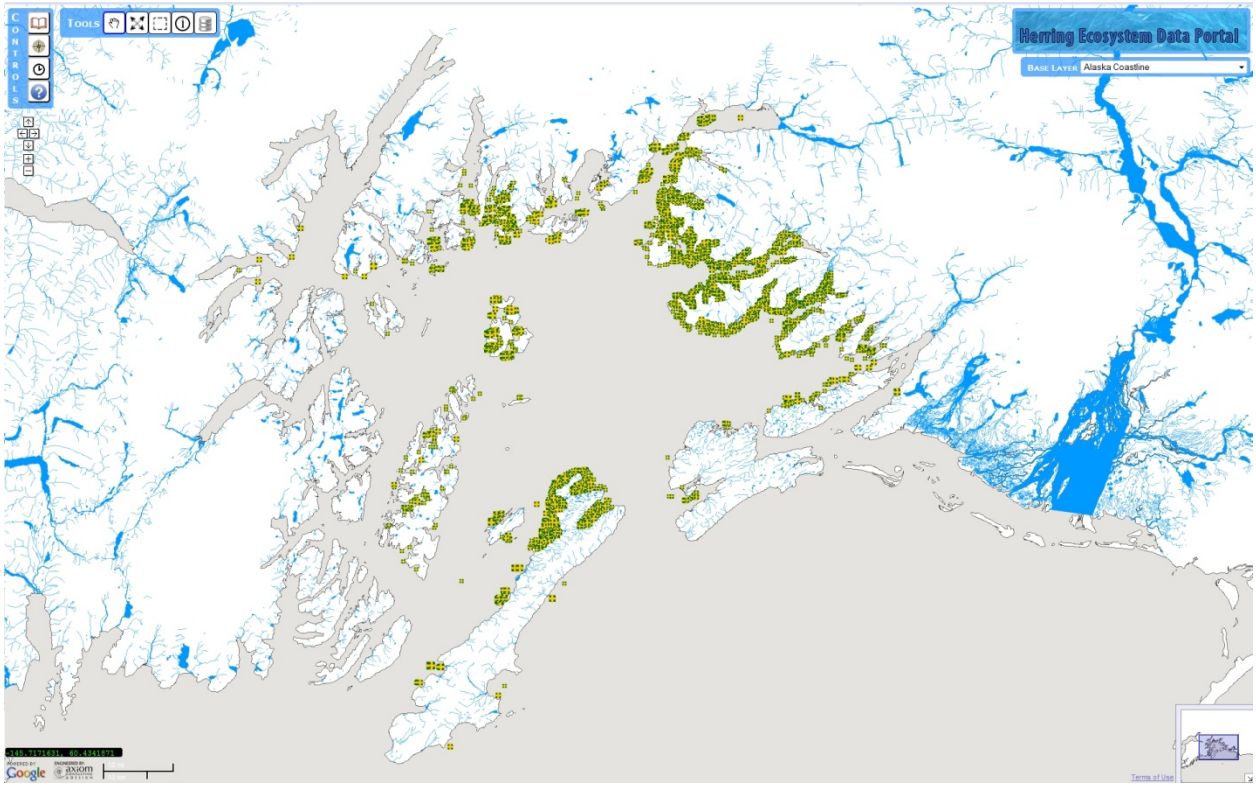


Figure 5: Aerial observations of biomass herring schools visualized through the PWS Herring Data Portal user interface.

Aerial Survey Effort – Documentation of observation effort for the historic ADF&G aerial survey program is separated into two distinct data files corresponding to sampling periods for 1974-1999 and 2000-2005. The more current set of data describes actual GPS routes of aerial surveys. The older data is an extrapolation of survey effort based upon the ADF&G Index Areas. ADF&G biologists have divided PWS into approximately 250 index areas which are distinct polygons denoting specific areas of the sound. The older extrapolated data set was constructed by associating the spatial bounds of the index area with documentation from the ADF&G aerial survey logs listing the index area as being surveyed during that flight. Data salvage activities for this information were extremely tedious and manual. In 2000, Geographic Positioning Systems (GPS) were employed to record the exact flight paths of aerial survey efforts. These flight paths were stored as line strings and associated with relevant survey header information (pilot, date, biologists, start time, end time and comments). Each aerial survey path is unique and documents the non-uniform sampling methodology employed by ADF&G. The modern, non-extrapolated data set contains approximately 65 GPS routes, and the older extrapolated data set contains approximately 19,000 survey effort polygons. Figure 5 below provides spatial visualizations of both data sets.

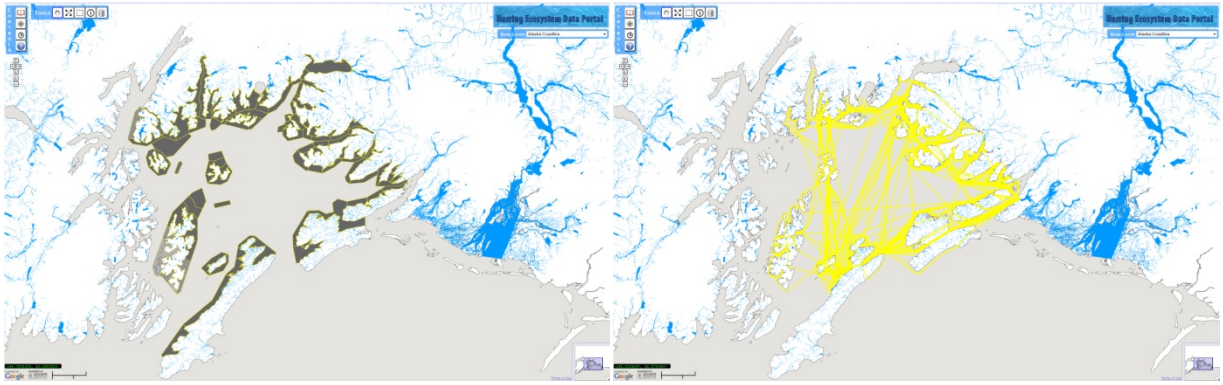


Figure 6: Visualizations of aerial survey effort for two regimes. Early routes (1974-1999) were extrapolated based upon extents of PWS index areas and are portrayed on the left. Modern survey efforts (2000 – 2005) have been captured more exactly using GPS tracks.

Aerial Survey Herring Spawn – Herring spawn activity has been monitored by ADF&G research and management staff since 1973 and provides a metric for biomass estimation and reproductive activity. Surveys are flown and sketches are drawn upon index area maps that show the extent of coastal area impacted by spawn. The total mile coverage of spawn is then calculated and used to drive internal in-season fishery management formulas. The spawn observations were digitized from hand-drawn sketches using Arcview software. Additional information describing spawn character is also included in the data set. There are approximately 6,000 spawn observations included with the resource. Figure 6 below provides a spatial visualization of these spawn observations. Once data were transferred from techs at ADF&G, Axiom data managers further processed the raw data into combined logical data sets.

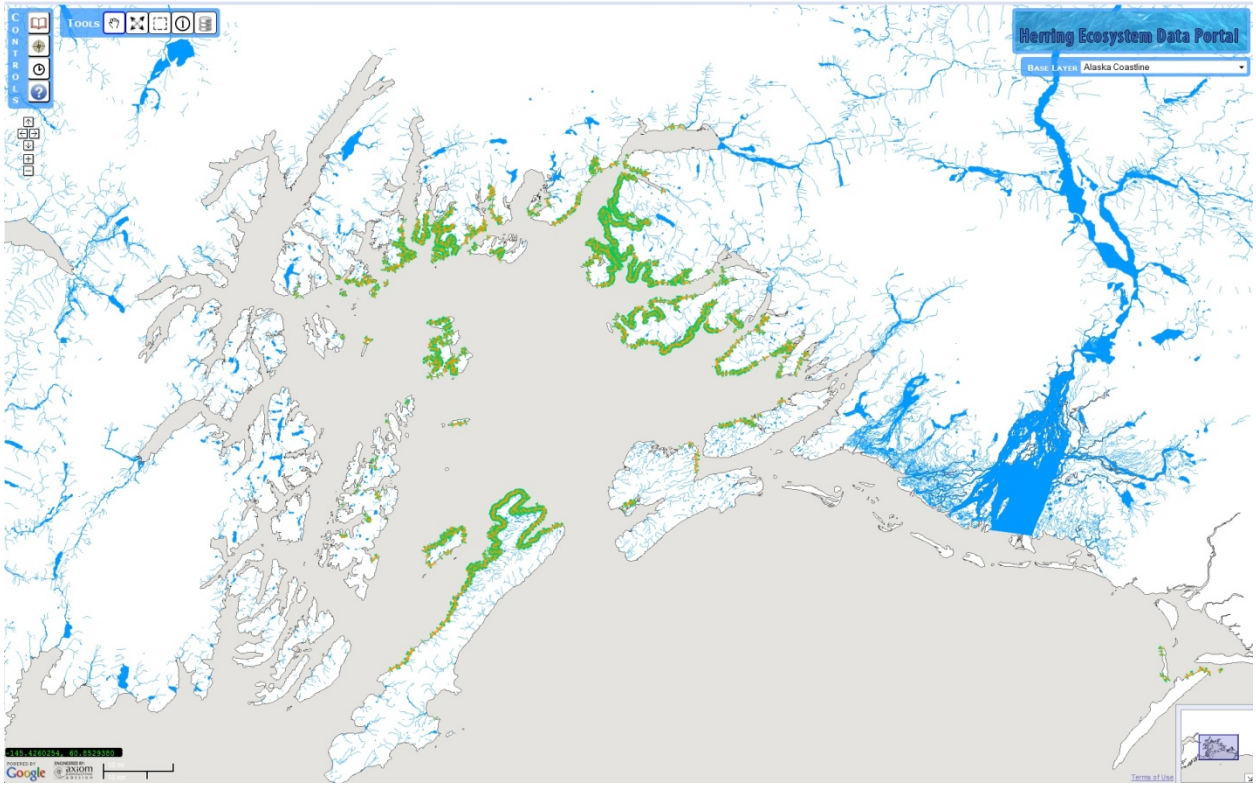


Figure 7: Herring spawn observational density of ADF&G aerial survey effort.

Herring Age Sex Length - This dataset contains approximately 200,000 individual herring age, sex, length and weight (ASLW) measurements sampled between 1973 and 2007. The information contained in this data set describes the age class of an individual fish based upon scale analysis, its length in millimeters and its weight in grams. Some records also contain gonad maturity and gonad weight. Sampling mechanism (type of trap or net) is also documented so that capture bias for size can also be considered during future analysis. Data were aggregated from various sources at ADF&G including spreadsheets, disparate databases and paper forms that were often hand keyed by technicians. In many cases inexact spatial information and location associated fish with a PWS Herring Index Area instead of a lat/long. In these cases, a point source location for the sample was extrapolated from creating a centroid of the index area polygon. Spatial information that was extrapolated is noted in the dataset.

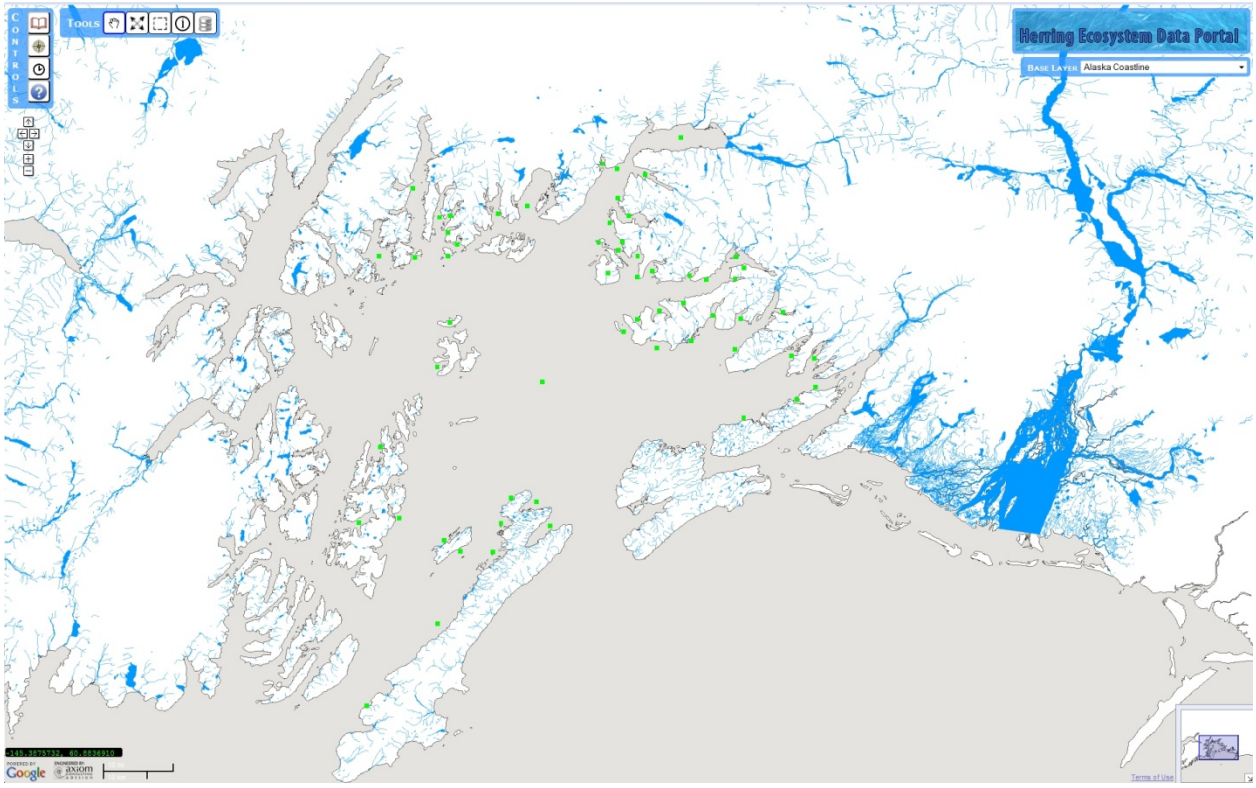


Figure 8: PWS Herring Age Sex Length data observations portrayed as green dots within the PWS Herring Data Portal interface.

Herring Commercial Harvest – During years of active herring fisheries in PWS, commercial harvests of herring were taken out of the ecosystem. This data set describes the details of all know herring commercial fishery activity within PWS. Included in this data set is information which details the biomass harvest, spatial/temporal aspects of the fishery and contextual information regarding the fishery opening. Spatial data is associated with polygons which can sometimes be large and cover large portions of PWS. These fisheries are also referenced to their ADF&G Emergency Order (EO). There are about 200 records in this data set. Two versions of the data set are included with the deliverables from the PWS Herring Portal. One of these data sets is composed of polygons which detail the extent of each fishery bounds and another data set provides a polygon centroid point location for each commercial fishery activity.

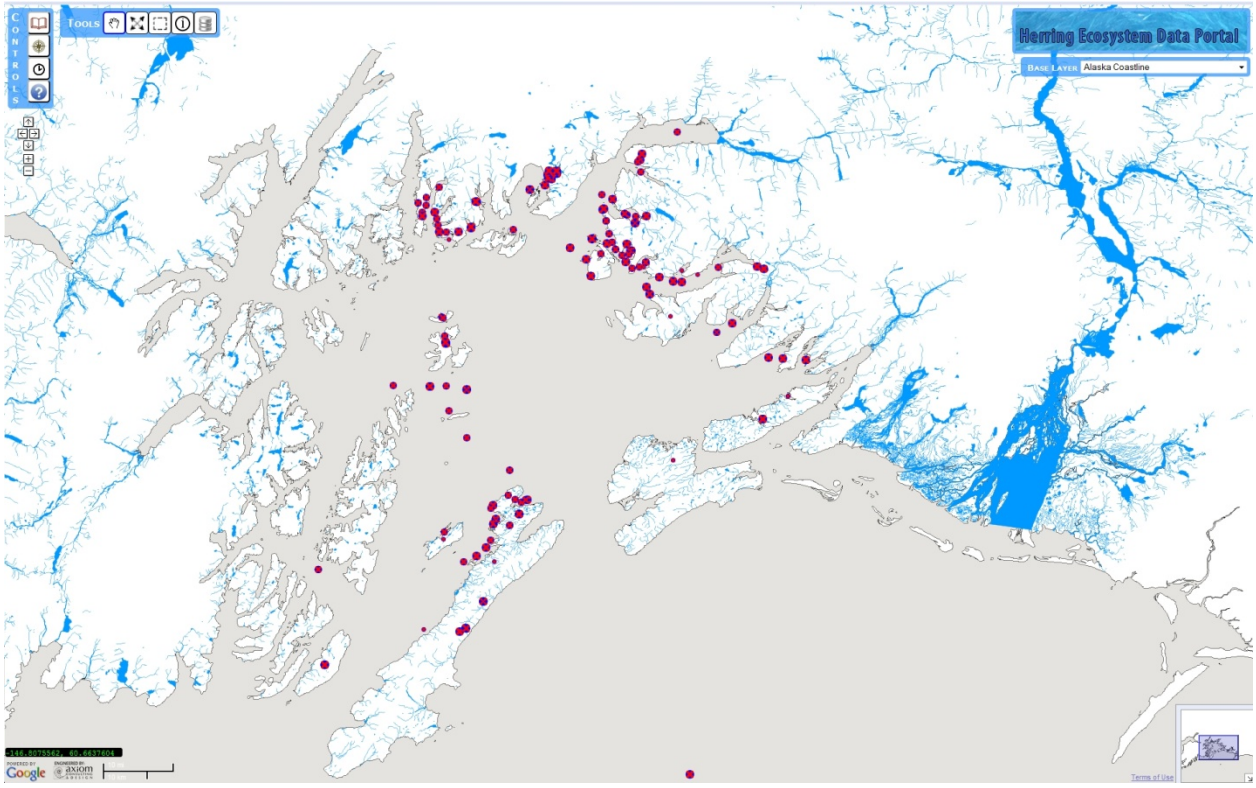


Figure 9: Herring commercial harvest centroids.

Additional data sets available from regional agencies which did not require salvage activities by investigators were also integrated into the PWS Herring Data Portal. These data sets were selected by their ability to provide contextual information regarding herring ecosystem in additional to detailing the distribution of herring predators and other potential GIS data layers which were of relevance to the PWS area. Table 2. Below denotes these datasets integrated into the portal.

Data Set Name	Start Year	End Year	Source
EVOS Oil Spill Surveys	1989	1991	ASGDC
Seabird Colony and Abundance Data	NA	NA	USFWS
Pelagic Mammal Surveys (NPPSDB)	1974	2002	USGS
Pelagic Seabird Surveys (NPPSDB)	1974	2002	USGS
Kamishak Herring Data	1978	2002	ADF&G/OSRI
Freshwater Data (streams, rives, lakes)	NA	NA	ASGDC

Table 2: Summary of data sets acquired from regional data providers, integrated into the portal but not salvaged by this project activities.

Objective 2: Provide web accessible map-based visualizations of geospatially enabled herring data and provide user interfaces for querying and downloading raw data and metadata.

The data management system and web-based portal produced from this project builds upon mature, open source technology that supports data management in a distributed environment. Our work relies upon the PostgreSQL database platform for data storage, the Geoserver platform for geospatial rendering and data access and the jQuery Javascript library and Google Maps API for the client user interface. All of the previous technologies have been in production and actively developed for a substantial period of time and support for these technologies has been secured from industry, ensuring that they will persist into the foreseeable future. The portal is available online at http://dev.axiomalaska.com/PWSHerringPortal/ and can be transferred to any host site desired by the EVOS TC.

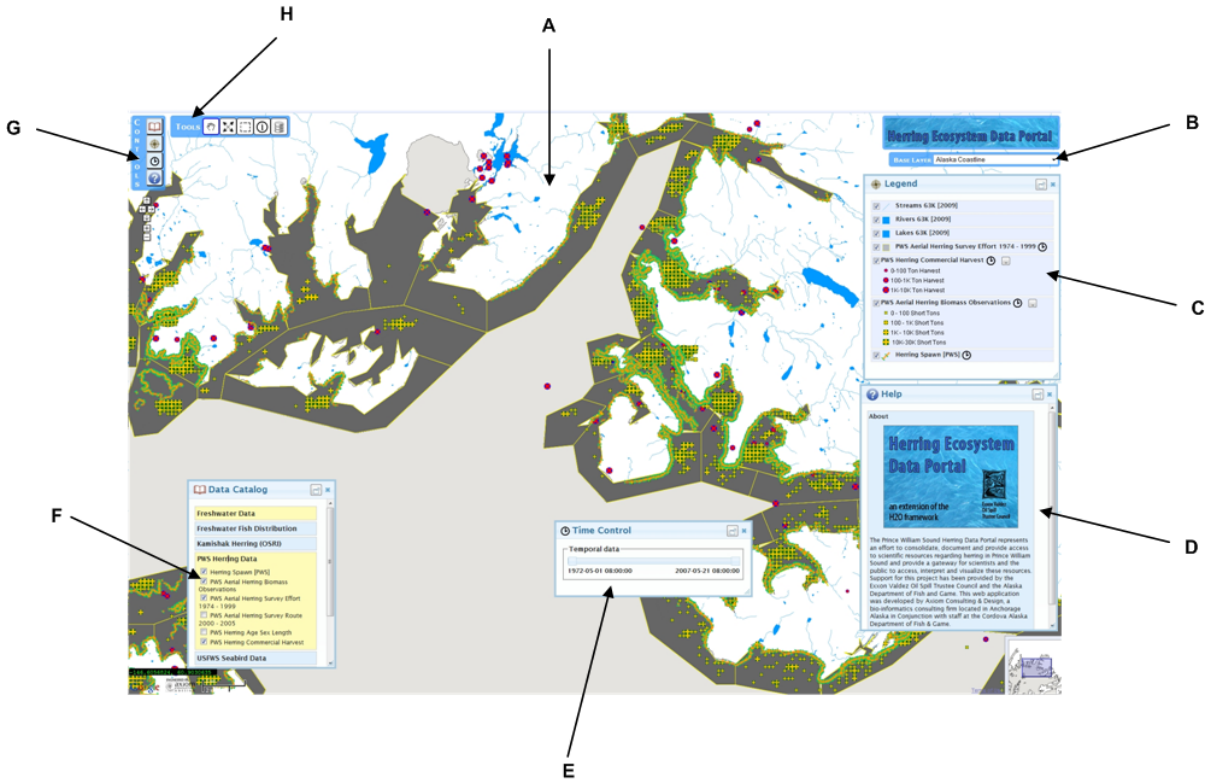


Figure 10: Diagram portraying the user interface of the PWS Herring Data Portal with many of the data sets salvaged from this effort shown upon the mapping interface.

Figure 10 (above) diagrams and details the various components of the PWS Herring Data Portal user interface. These components are listed and described below:

A. **Spatial Data Viewer -** This is the primary pane of the user interface for the PWS Herring Data Portal and provides the user with spatial visualizations of datasets. This pane is

manipulated by the user interacting with the other configuration tools of the interface. These interactions could include loading or unloading data layers, panning and zooming the map, filtering by time or interacting with GIS information.

B. **Base Map Selector** – Provides the user with options to select base maps based upon preference. Options include the various Google Maps base layers (satellite, default, physical features), two base layers served out of USGS (USGS topographic, USGS orthophotos quads), Geographic Information Network of Alaska Best Data Layer for satellite imagery and a polygon Alaska Coastline vector layer which has been designated as the default base layer for the application.
C. **Data Legend –** Provides users with details regarding the symbology and characteristics of the data loaded on the map. The legend also allows users to manipulate the order in which those data layers are stacked upon each other in the map. Data which has a temporal component is noted in the legend with a watch symbol. Layers loaded onto the map can be turned on or off on the map pane from the legend control.
D. **Info Help Modal –** The info help modal provides users with contextual information regarding the project and also with a help menu which details how to navigate the applications and use the various data and interface tools.
E. **Time Control –** Provides users with a tool to filter data layers through manipulation of a time period slider. As data sets are added to the map pane, the time controller adjusts its bounds based upon the summary bounds of all temporal data layers loaded on the map.
F. **Data Catalogue –** The data catalogue provides the user with an interface to load or remove layers from the legend. Layers are hierarchically organized into logical groupings of data. Layers are added to the map by selecting from a drop down list or checking a check box from various expanding accordion style interfaces. The catalogue has been set up to only allow logical combinations of data to be loaded onto the map by users.
G. **Modal Controller -** The modal controller allows the user to reload the various application control tools once the user has closed them from the interface. These control tools include the time control, data catalogue, legend and help modal.
H. **Data Tool Palate –** This palate allows users to interact with the map and data displayed on the interface. The pan tool is symbolized with a hand and allows users to pan across the map interface. The zoom to full extent tool is represented by a set of four arrows expanding in all directions. This tool allows users to zoom back out to the full extent of the PWS Herring Data Portal data resources. The marquee zoom tool allows users to zoom into areas on the map by clicking and dragging a marquee box around the area of interest. The feature info tool, symbolized with a circular (i), allows users to extract underlying tabular information out of GIS data loaded on the map interface. Finally, the data download tool symbolized with a database (drum) symbol allows users to draw a box around spatial features and extract the underlying raw data in a variety of formats including excel, csv and shapefiles.

Objective 3: Develop Standard Operating Procedures (SOPs) which document system implementation and provide direction for the absorption of additional herring ecosystem data sets, metadata and information to the centralized herring data system.

This intent of the Standard Operating Procedures (SOPs) document is to provide a high level directive for implementation of data management practices to provide a foundation for ecosystem based long term monitoring in addition to providing strategic guidance for addressing the needs of all potential users of the evolving PWS Herring Ecosystem Portal. Documenting system requirements from the perspective of user groups provides a solid base for the development of use cases, measurable goals, data flow diagrams and other software design methodologies. These methodologies will enable data managers to more effectively plan, design, code, implement and maintain the PWS Herring Ecosystem Portal and continue to integrate additional sources of data into a data management framework. The SOPs developed during this project are formalized in an appendix of this report.

Objective 4: Develop Arcpad application to automate herring spawn data collection.

The Prince William Sound Arcpad Aerial/Boat Survey application provides an automated and standard way to document herring-specific physical and biological observations. The Arcpad application is still used each season by ADF&G for Aerial surveys. A copy of the application with system documentation will be issued to EVOSTC as a deliverable for this project. The following screenshot (Figure 11) provides a view of what the biologist interfaces with when using the application on a laptop computer while flying the aerial survey.

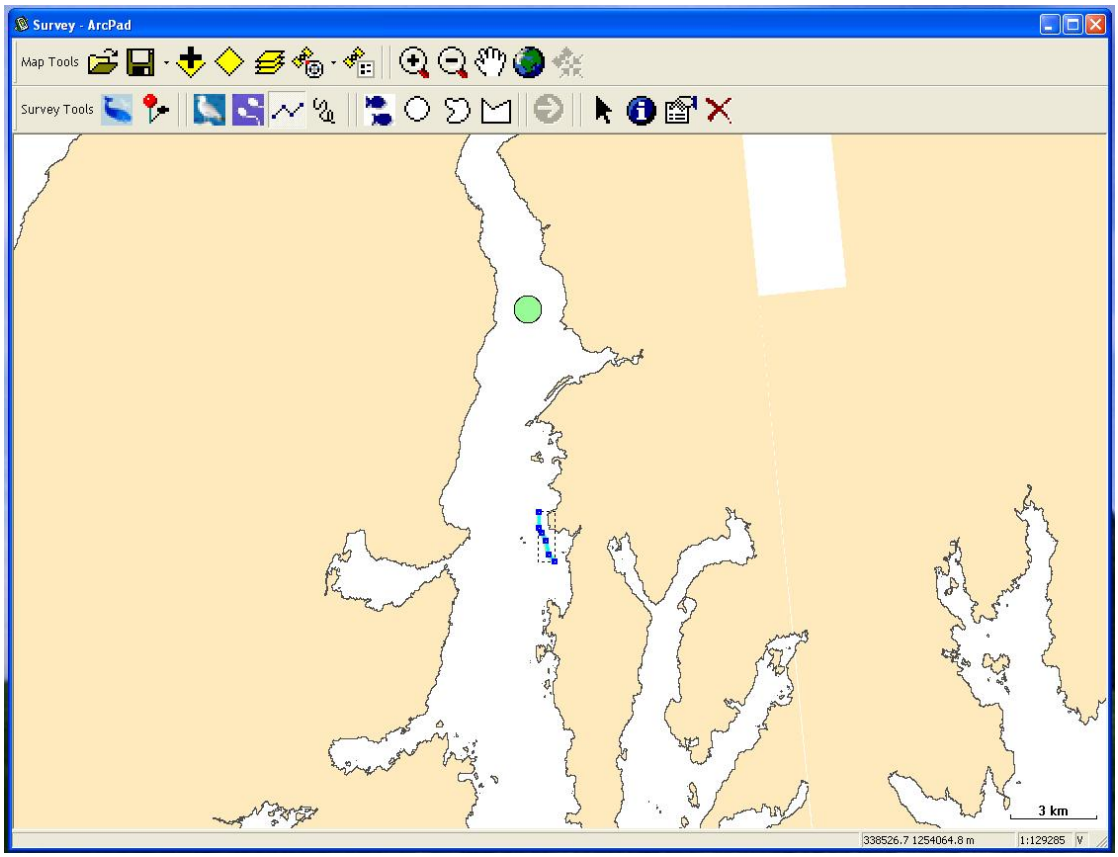


Figure 11: Aerial Survey Application Interface.

The survey application provides automated and standard data collection for herring spawn extent (spawn class, polyline location), marine mammal observations (species, number, point location), Seabirds (species, number, polyline location), fish schools (species, biomass, polygon location), and other periphery observations. The application also stores GPS tracklog information for flight/cruise routes. This application has undergone three design iterations.

DISCUSSION

Objective 1: Consolidate herring data sets, metadata and other electronic resources to publicly accessible web portal for herring information.

Data salvage activities are extremely time consuming and costly. It is clear from the labors of this project that implementing correct data management procedures during or immediately after information is collected is orders of magnitude more efficient than attempting to document data years after the fact. This project focused much of its effort on the tedious and extremely labor intensive activity of salvaging, processing and staging historic data sets. The PIs severely underestimated the amount of effort it would take to standardize, digitize, document and spatially enable the resources that are detailed in this report. Up-front implementation of correct data

management protocols during data collection can drastically reduce the amount of resources required to stage data for potential synthesis efforts.

Objective 2: Provide web accessible map-based visualizations of geospatially enabled herring data and provide user interfaces for querying and downloading raw data and metadata.

The data access application deployed during this effort was driven by the need to design a prototype multidisciplinary spatial management tool and data management framework. The PWS Herring Data Portal was rapidly developed by Axiom to explore spatial management tool technology, develop a web-based map application which served data from various disciplines (marine and terrestrial) and provide a mechanism to engage user groups and acquire feedback for future development iterations. The tool provides access to a large number of different types of spatial data layers via a common interface for integrating these data visually. This functionality is coupled with access to underlying raw data and the ability to filter multiple disparate data sources simultaneously by space and time. The system uses Geoserver to serve raw data and create visualizations in addition to an asset catalogue (ontological metadata) tier which powers the user interface controls and allows the implementation of simple user interfaces to operate on multiple data layers.

Once the prototype was developed, users were engaged in an informal requirements analysis exercise to isolate successes and failures of the platform and guide future development iterations. Several users were engaged from the following demographics to acquire feedback regarding the functionality of the prototype:

- Scientific research community
- Resource management
- Education (teachers and students)
- General public

Although the overall response to the tool was very positive, users routinely voiced a common set of issues regarding their interaction with the prototype and its existing functionality. The review of this system is relevant to this proposal and summarized below:

1. *Non-uniform performance on different browser systems* – This issue is due to the reliance upon JavaScript for client application services (Google Maps and OpenLayers). Browsers such as Mozilla Firefox and Google Chrome are much more efficient at supporting JavaScript based systems than Internet Explorer. Many users employ Internet Explorer as their browser either by default or as a requirement by internal organizational IT security policies. Those users experience application sluggishness when loading large amounts of data layers to the application; the resulting performance can be significantly slower than that experienced by users with Firefox or Google Chrome browsers.

2. *Need to improve access to more types of data and improved metadata* – Though a large number of data layers and spatially enabled data sets are available via the prototype, users voiced a need for access to data which were currently not available through the tool. Users commonly requested real-time data streams, improved metadata quality and access

to more advanced types of data such as climate models and other four dimensional data sets describing the atmosphere, water column or geological structure.

3. *Need to develop both general and targeted tools* – Users also consistently voiced the concern that the interface may be too complex for many users. Expert users who are familiar with GIS software and other types of complex software systems would find the interface familiar, but those unaccustomed to such systems may find the prototype interface intimidating. Users requested that there be multiple interfaces developed, including tools for expert users and simplified topic-focused applications developed for less technically oriented users.

Objective 3: Develop Standard Operating Procedures (SOPs) which document system implementation and provide direction for the absorption of additional herring ecosystem data sets, metadata and information to the centralized herring data system.

The Standard Operating Procedure (SOP) document included with this report as Appendix 1 details the technical facets of the PWS Herring Data Portal and provides a framework for future ingestion of datasets into the information management system. This document lists a series of methodologies and technologies that can be used to successfully manage multidisciplinary data sets which are collected or aggregated for use in an integrated ecosystem research program. Information technology is inherently a constantly evolving landscape and as a result the best methods and technologies for managing data may change rapidly. It is vital that data managers constantly assess and explore new approaches/technologies to meet their goals. A core set of strategies for data management which are clearly here for the long term are the concepts which are fundamental to interoperability systems. These systems provide an extremely flexible and scalable environment for storing, integrating and sharing a wide array of data types and metadata.

Objective 4: Develop Arcpad application to automate herring spawn data collection.

This application is still used today by ADF&G staff in Cordova to facilitate the automated collection of aerial survey observational data. Even though this application still proves to be useful, it is clear that the ESRI Arcpad technology has limitations. New emerging open source client based software (such as Adobe AIR) can provide a technical environment that better addresses the needs of the Cordova office. Adobe AIR provides a much more flexible and richer Application Programming Interface (API) for developers. Using Adobe AIR over ESRI Arcpad will make system development significantly more efficient and provide access to functionality not possible through the current ESRI programming environment. If future efforts to revamp the aerial survey application are undertaken, it would be highly effective to redesign the application employing the Adobe AIR programming environment.

CONCLUSIONS

This project explored methodologies to address the data management needs of a multidisciplinary integrated ecosystem research program. Focusing on an iterative approach to project implementation, investigators were able to adapt strategies employed to best meet the voiced needs of the entire user base. Products from this effort include a standardized and salvaged 35 year Prince William Sound herring monitoring dataset, an advanced spatially/temporally explicit data management and visualization system and a best practices guide to direct future data management efforts for ecosystem approaches to research. It is clear that data managers must take a pragmatic approach to implementing solutions for the management of scientific information. This approach must constantly evolve based upon lessons learned and the adoption of emerging technologies and methodologies.