OBIS-USA: Establishing The United States Node for the Ocean Biogeographic Information System to Support the Census of Marine Life

USGS-CESU Grant # 04121HS009 to University of Colorado (Award # 04HQAG0121)
Effective Dates: June 11, 2006 through December 11, 2007
Funding Amount: $105,000, augmented $30,000 and extended to June 13, 2008

2007-2008 (Final) Project Report

Reporting Period: June 12, 2007 – June 13, 2008

John J. Kineman, Ph.D.
Senior Research Scientist

Rozita Abdul-Williams
Research Associate Scientist

Cooperative Institute for Research in the Environmental Sciences
Campus Box UCB-216
University of Colorado
Boulder, Colorado 80306
john.kineman@colorado.edu

A Project of the Rocky Mountains Cooperative Ecosystem Studies Unit
and the US Geological Survey
Agency Representative: Dr. Mark Fornwall, USGS/PBIN

Final Reports to be delivered to:

Executive Coordinator, RM-CESU University of Montana School of Forestry Missoula, MT 59812 406-243-5346; 406-243-4845 (fax) rmcesu@forestry.umt.edu	Director, Pacific Basin Information Node 310 W. Kaahumanu Ave., Box 196 Kahului, HI 96732 808-984-3724 mark_fornwall@usgs.gov

Preface

This is the second and final report of the OBIS-USA development project. Please see the first report (2006-2007) for events and accomplishments in the first project-year. The current report covers what occurred in the second year, and provides a final conclusion for the project. These reports cover 12-month reporting periods (the "project-year") from June to June.

Table of Contents

Summary

This report serves as the final project report and 2007-2008 annual report.

The OBIS-USA development project at the University of Colorado completed on June 13, 2008. It produced a pre-operational prototype of the USA Regional OBIS Node (RON), integrated with a new NBII Marine Data Theme that was developed by NBII in conjunction with this project. The data volume reached 1,700,366 quality-checked records from 25 separate data resources contributed by 17 Providers. A total of 10 updates or revisions to these data sets were obtained as a result of provider contact and quality testing procedures shared with them in pre-release reviews. At the time of this report the review and approval process was completed for 15 providers, while the remaining two require additional enhancements to deal with citation credits and proprietary issues. The data continue to be served from a dedicated server at the University of Colorado, jointly operated by CIRES and the CU-Museum. Data services respond to searches initiated from the NBII Data Portal's OBIS-USA search page. Additionally, a "Provider Resources" Website" was established and served from the CU server, to support communications for pre-release development and approval of data and services. This site provides quality-control feedback to providers and data contributors to support an ongoing dialogue about data quality and updates. The final prototype system was delivered as a hard drive image of the OBIS-USA NBII server, which remained functioning at the CU Museum, bioinformatics facility.

OBIS-USA data services were established at:
http://www.nbii.gov/portal/community/Communities/Habitats/Marine/Marine_Data_(OBIS-USA)/.
The OBIS-USA Provider's Resources Site were established at:
http://obis-usa.colorado.edu/dev/.

Besides data and infrastructure deliverables, we produced long-term development plans, poster presentations, standard operating procedures, fact sheets, joint proposals, and experimental analytical data products (see both Annual Project Reports and their appendices). The project team gave informal discussions and presentations at the appropriate venues throughout the project. We advanced the goal of establishing OBIS-USA as a national service by designing, testing, and implementing prototyp systems and services, and by acquiring distributed data resources. The project also served as a catalyst for cross-agency and international discussion in the second year, about integration of physical and biological information and interoperability among data systems of NBII, OBIS, CoML, and IOOS. General benefits to the public include increased access to marine species data and improved quality of those data.

Data acquisition the mechanics of building server and Web infrastructure were equally balanced to ensure effective data acquisition, quality control, and delivery. Initial data acquisition targeted transfer of designated USA Providers from IOBIS to OBIS-USA, and adding new data or corrections from those providers as a result of involving providers in a quality assessment process.. This work was commensurate with Phase I of the three-phase Development Plan produced in the 2nd quarter of the project (see 2006 Annual

Report). The design strategy was iterative, with three phases planned, to ensure adaptability in the face of changing needs, budgets, and technical options. Data volume achieved the full capacity of Phase I architecture.

The project was augmented in the second year to improve the look and feel of the website and to complete the software implementation after UC-Davis's contract for programming support ended in the first year. The project period was also extended to June 13, 2008.

During the two year project period the national and international demand for biological data delivery and integration with other observing systems (notably IOOS) increased steadily. Still, our joint proposal to scale up OBIS-USA services for the Pacific region and to provide interoperability with IOOS (submitted in the first project year) was rejected. The work continued at a reduced level commensurate with remaining funds after January 2008,

The major conclusion of this R&D effort is that meeting national and international desires for a fully operational facility that can handle the estimated volume of USA marine data resources, will require resources to scale up to Phase II/III architecture. Two additional recommendations resulting from this work were: (1) continue to strengthen and expand data services; (2) initiate science components. We recommend pursuing these two recommendations simultaneously so that the development of services and the evolution of needs can adapt to each other; but we recognize the method of implementation depends on funds and various mandates.

Some of our conclusions address the question of overall scientific value; confirming the CoML/OBIS strategy of integrating raw biological occurrence observations, but questioning the IOOS presumption that observations themselves should be combined into one system with physical oceanographic data (Appendix IV). We recommend a higher level integration of analysis and modeling outputs (except for cases where concurrent sampling of physical and biological data is more efficient and in the scope of a given field program).

Technical Development

Providing services through the NBII Data Portal was the greatest challenge in the first project-year because of technical unknowns and features of the Portal in development at the time. A system was implemented whereby searches could be initiated in the Portal and results delivered outside the portal, on the University server. This solution remained in place throughout the second project-year. Development and Provider support services, and operations were temporarily combined in the second year of the project while options for permanent operations were being considered by NBII.

The architectural limit of the initial DiGIR server design (Phase I) was estimated by DiGIR experts early in the project to be less than 1M records and 10-12 resources. We in fact encountered performance problems with 8 Providers comprising 600,000 records. The UC-Davis programmer resolved performance issues and allowed an increase in

capacity by moving temporary operations to the CU server. This provided a temporary solution, still running all services from one machine.

The CU team assumed programming responsibilities through an augmentation to the project in the 3rd quarter of the first project-year, to maintain and improve software produced by UC Davis. This effort was funded at $30K (less 17.5% overhead) and officially began on June 13, 2007. The aim was to complete Phase I implementation and then to conduct a systems analysis for scaling up to Phase II/III (with additional resources). Scaling up would allow us to surpass the volume and performance restrictions of DiGIR and to minimize the need for custom software by taking advantage of standard DBMS functions in readily available software.

Phase I services were completed in August, with the exception of common name search. That function was developed and tested using the Catalog of Life names cross-reference, but its roll-out was deferred to the next design phase as a result of technical issues in activating the additional SQL search engine. A site demonstration and review was conducted August 13, 2007, in preparation for 'going live' with the website in which the NBII review team called for specific revisions (see Appendix V-a). These tasks were accomplished with the remaining funds from the augmentation grant and the site went 'live' with the NBII Portal in November, 2007. A more comprehensive scope for Phase II/III development was prepared and presented to NBII in January, 2007 (Appendix III).

OBIS-USA Site and Resource Status

The OBIS-USA site was activated in a 'pre-operational' mode accessing the CU DiGIR services, at the same time that the NBII Data Portal became public. This was necessarily in trade with continued site development and maintenance of Provider services because development and operations require separate installations which we did not have resources for. At this time the CU team had brought the available data volume to a total of 17 Providers, 25 resources, and 1.7M records. Two of these providers remained in the approval process because of copyright, data use, and citation policy issues. CephBase was in the progress of privatizing from its former University sponsor. Seamap, accounting for over 1M records, was a thematic node and there were issues about how to ensure citation of individual providers that were being aggregated in this one combined resource. A data use policy was drafted (See Data Use Policy, this report) to address some of these concerns. Both of these cases were important from a design perspective. SeaMap and IOBIS were collaborating on a general solution as of this report.

A second augmentation was proposed ($30K, including 17.5% CESU overhead) to implement Phase II/III centralized DBMS architecture and to divide services among multiple machines for increased performance and capacity. Because funds were not available for this upgrade, we were asked instead to continue optimizing the Phase I software, Web site, and data ingest, to support pre-operational demonstrations and other presentations. The primary scope of work on the project was drawn to a close by February 15 except for maintenance and to continue work on five data acquisitions in

progress at the time. The last 6-months of the project were budgeted at a lower level for completion of data acquisitions in progress, for maintenance activities, and for a smooth project closure, maintaining services as long as possible while NBII considered options for the future.

Requirements and Accomplishments: Original Grant ($105,000)

(also see 2006-2007 Report)

General Requirements	Accomplishments
1. Coordinate related activities closely with the Pacific Biodiversity Information Node of NBII, Directed by Mark Fornwall and establish a [United States of America] node for the Ocean Biodiversity Information System (OBIS), to support the 10 year Census of Marine Life (CoML).	Coordination continued among the four 'core' collaborators, CU/CIRES, USGS/NBII/PBIN, UC-Davis, and USGS/FISC. UC-Davis and FISC's involvements completed and an Augmentation to the CU budget was put in place to partially cover those functions. Collaboration was deepened with the CU (University of Colorado) Museum by involving members of their team in the Augmentation grant.
2. Utilize existing capabilities and resources already engaged in various OBIS tasks and leverage on existing capabilities at CIRES and the CU Museum.	See first-year report. In the second year, the CU-Museum continued to host the OBIS-USA server. Affiliates of the Museum were hired under the Augmentation grant (see below) to assist with web design and development.
3. Further interactions with the International OBIS (I-OBIS) and with the Australia/New Zealand node to develop a marine component for the Pacific Biodiversity Information Forum. Expand the US RON team to include a member from IOBIS	See first-year report. In the second year, collaborative research was discussed with IOBIS (Ed Vanden Berghe) and SEAmap (Pat Halprin) regarding species distribution modeling.
4. Locate and acquire species data and metadata, bringing new data into the OBIS system and making it usable through the OBIS portal. Data acquisition and integration of US data will be done initially through the US RON. Serve as the primary point of contact for data acquisition into US RON. Develop decision criteria for selection of databases to add to the US RON	See first-year report. In the second year, data acquisition and servicing of existing providers continued. 5 providers were added and 2 were updated. Data acquisition criteria established in the first year were followed. The data acquisition effort was reduced because: o The budget for the original scope was 50% less than the first year and largely spent by January, except for maintenance of routine operations and project completion tasks.

	o Web development, programming, and server support tasks were an equal priority to data collection. o System performance capacity (Phase I) was limited.
5. Encourage demonstration projects showing or advancing the capabilities of the OBIS node.	See first-year report. In the second year, a demonstration application involving the prediction of species distribution (Great White Shark) using OBIS data was prepared and presented to IOBIS and Seamap, stimulating interest in further collaboration.
6. Outline a long-range strategy for an ongoing US Regional OBIS Node (RON), which will have responsibility for all US OBIS data.	A long-range strategy was developed in the first year (See first-year report). Phase I of that plan was completed in the second year.
7. Establish procedures to ensure quality and that these data are properly entered into the OBIS servers (at Rutgers University) and delivered publicly via the portal.	Procedures were established in the first-year (see first year report). They were maintained in the second year until live operation serving the NBII portal prevented simultaneous use of the server for pre-release data review.
8. **Added tasks (1st year):** Server and Web infrastructure development, including: o Development server (for pre-operational testing) o Web pages (for OBIS-USA info, search within the portal and display of results) o GIS map services and background layers (for results display) o Provider Resources Site (for QC and feedback to Providers) o IT support o Modify development plan	See first-year report. In the second year, the team continued to balance IT, Web, and server development with data acquisition. The required services were maintained to the extent of single server capacity.
9. Collaborate in attaining grants, contracts and other assistance to help with Node operations, research or development.	See first-year report. In the second year we received notification from NOAA that the collaborative IOOS proposal submitted April 17, 2007 was rejected (August, 2007). There were discussions with NOAA about alternatives but no initiatives resulted except the joint CoML/IOOS meeting on OBIS-IOOS interoperability that took place January, 2008 in Washington, DC.

Outcomes with completion dates (reports, publications, workshops, videos, etc.)

(This is a 1-year start-up effort. All outcomes have a 1 year delivery window, June 7 through June 5).

1. Develop and implement an 18 month plan for US RON data access.	Completed December, 2006 (Appendix I).
2. Develop standard operating procedures for adding data to the US RON and the IOBIS portal	Standard Operating Procedures developed in the first year (see first-year report) were unchanged. These covered: Data Ingest and testing Quality assurance Public release Spatial Data Processing
3. Develop NBII Marine Theme including appropriate integration with existing NBII marine activities.	See first-year report. In the second year, significant improvements to the OBIS-USA site were made as a result of NBII reviews.
4. Cooperate with CReefs/CoML project and continue to provide support for data management from the project.	See first-year report. In the second year, Creefs data that were withdrawn by the provider in the first year, were replaced and labeled "CRED."
Set data priorities, identify sources of data, and bring data into the OBIS system. **8 target US data resources were identifies** (plus the 2 already in the prototype)**:** 1. CReefs 2. MMS 3. Exxon Mobile 4. NOAA OE 5. EPA EMAP 6. ArcOD 7. GoMA 8. GoMX	See first-year report. In the second year, **5 new providers were added to OBIS-USA**, Status of original target resources changed: 1. CReefs (re-acquired as CRED) 2. NOAA OE (IOOS-CoML meeting) 3. EPA EMAP (acquired) 4. GoMX (acquired)
5. Develop applications that highlight the utility of the US RON and OBIS.	An experimental application was developed mapping Great White Shark distributions (see Appendix VIII)
6. Publish/present research based on OBIS Data.	See first-year report. A paper on OBIS-USA development will be submitted for publication in 2008.
7. Participate in OBIS or IOOS related meetings as appropriate. Develop presentations, posters and other outreach materials. Collaborate with Oceans.US to coordinate job duties as related to the US RON and IOOS.	See first-year report. In the second year, we attended an IOOS-OBIS interoperability meeting in Washington DC, January 2008 in which overall strategy and OBIS support to IOOS was discussed.
8. Brief USGS and IOOS leads on US RON activities.	This was done through the NBII Project Manager.
9. Enhance the existing Marine Node	See previous table, re: NBII Marine Theme

Web Site and access point to US RON activities.	and OBIS-USA pages.
10. Develop 1 page fact sheet defining the US RON (mission, goals, users, tools, products)	See first-year report.

Requirements and Accomplishments: Augmentation Grant ($30,000)

1. **Define and follow a system concept (technical development plan)**: Develop requirements analysis focusing on critical functional requirements, and produce a well-integrated and coordinated, but parsimonious Technical Development Plan (TDP), with long and short-term objectives.	This item was in anticipation of re-design for Phase II/III. It was partly included in the IOOS proposal, but otherwise canceled because funds were not obtained for implementing a redesign. $4,000 allocated for this task were reprogrammed to provide continuity of existing services past December, 2008.
2. **Enhance web services and user interface:** maintain and improve functionality of the OBIS-USA public and development websites to meet project objectives (collaboratively defined). Critical functions include:	A programmer was hired to continue work after the UC-Davis contract ended. Also a Web Designer was hired. Several site reviews conducted by NBII provided detailed guidance on changes to the website and/or programs.
a. Scientific and common name search and delivery of tabular results with various display and download options for the user.	Scientific name search was implemented. Common name search was developed (but not implemented because of server performance issues), using the Catalogue of Life names. Results are displayed in tabular and map form.
b. Improve spatial mapping functions for search results using data layers provided by CIRES staff, and implementing appropriate tools such as Google Earth to provide regional displays (limited to the OBIS-USA regional window).	A database and GIS data service were developed and integrated with the search results display using an ESRI ArcGIS server. In discussions with NBII we learned that a map service was added to the NBII Portal capabilities in May, 2007. The ESRI spatial data and mapping capability we developed should be compatible, and can be moved into the Portal system in the future if desired.
c. Innovative tools for data display, visualization, and analysis, as available.	Only basic display functions could be supported.
d. Data quality feedback functions (for	A QCAA protocol was executed for

providers) via the CU Server and OBIS-USA development site (such as analytical displays, DiGIR test search capability, data exchange capability, etc.).	each provider's set of resources and a Technical Report prepared on-line for their review. Capabilities include search, metadata, and species name list query, and point maps of all data to verify location.
3. Maintain and develop infrastructure: Work with PBIN and CU CIRES to maintain and improve the infrastructure of OBIS-USA	The CU-museum continued to host the Blade Server as part of their server array. A proposal was submitted to NBII to modernize the architecture and to provide the CU-Museum funds for IT support (Appendix VIII).
a. Perform routine maintenance of the OBIS-USA server environment, ensuring its security and operational status.	We performed routine maintenance and updates of the server.
b. Recommend and implement, as feasible, upgrades and enhancements as they become available, for example in the DiGIR and Web-GIS environment.	Upgrades were not required beyond the initial DiGIR and ESRI ArcGIS installations for Phase I. The major upgrade of both Taxonomic and map services was planned for Phase II/III.
c. Resolve interaction issues between the CU, PBIN, and NBII Servers, Websites, and Services (smooth flow of data and services via the NBII Portal and CU Server development site and/or other mechanisms).	The PBIN server was removed from the OBIS-USA infrastructure plan (see first-year report), and pre-operational functions were transferred to the CU-Server on a 'pre-operational' basis.

Critical Development Priorities

1. Data harvesting, provider and portal issues	
a. Resolve data delivery performance issues via the NBII Portal and associated services, and establish a development schedule to keep pace with anticipated growth in data volume. This may involve development of central DBMS caching functions.	Performance was improved significantly to the design limit of Phase I architecture.
b. Develop a harvesting function to crawl designated providers and update the OBIS-USA data cache (possibly modifying software	An interactive procedure was developed using SOAP request functions. Automation was not implemented because it required central DBMS caching (Phase

provided by IOBIS).	II). Also, many providers were not operating in DiGIR server mode, but were still transferring individual files.
c. Ensure that data can be harvested via DiGIR or its successor, from the OBIS-USA operational server (at PBIN) by IOBIS and GBIF.	IOBIS began harvesting OBIS-USA data from the CU Server in the Fall of 2007.
d. Implement capability for automated update of the PBIN operational server with approved provider resources from the CU Development server.	Options were discussed for establishing an operational server at various locations. NBII did not establish an operational server during the project.
2. Online search and mapping of taxonomic data	
a. Enhance the functionality and appearance of the taxonomic search and mapping functions associated with the NBII Portal / OBIS-USA website.	The site was redesigned to meet detailed specifications from an NBII review.
b. Implement GIS database for visualization of taxonomic distributions	The GIS database was improved and augmented with new data layers.
3. Web design issues/UI	
a. Enhance the functionality and design of the OBIS-USA public site (NBII Portal page), working collaboratively with FISC.	New hires were sent to NBII Portal training and subsequently worked with FISC and CBI staff.
b. Enhance the functionality and design of the OBIS-USA development site (CU Server) working collaboratively with CIRES.	The development site continued to serve Provider interactions adequately and required only routine updates according to SOPs.

Quarterly Activity

1st Project Quarter (June-September)

1. Implemented NBII Augmentation grant (hiring new staff).
2. Initiated Web page re-design and software modifications to complete Phase I system.
3. Attended NBII Data Portal training (Sunny Lu, Leigh Ann McConaughy, Rozita Abdul-Williams)
4. NBII web-site review (Denver in August): List of additional changes to enhance implementation.
5. Continued data acquisition, QCAA, servicing, and update activities.
6. Maintained existing infrastructure services.

2nd Project Quarter (September-December)

1. Completed software revisions and Web re-design. All functions implemented with the exception of common name search.
2. Activated OBIS-USA 'live' via NBII, along with their activation of the NBII Portal (November).
3. A "final list" of desired website changes was communicated from NBII.
4. Continued data acquisition, QCAA, servicing, and update activities.
5. Maintained existing infrastructure services.

3rd Project Quarter (December-March)

1. Reduced budget to maintenance and data set completion levels (as planned)
2. Finalized list of preferred data acquisitions (5 new providers)
3. Attended IOOS-OBIS interoperability meeting in Washington DC (sponsored by CoML), submitted recommendations.
4. Continued data acquisition, QCAA, servicing, and update activities.
5. Maintained existing infrastructure services.

4th Project Quarter (March – June)

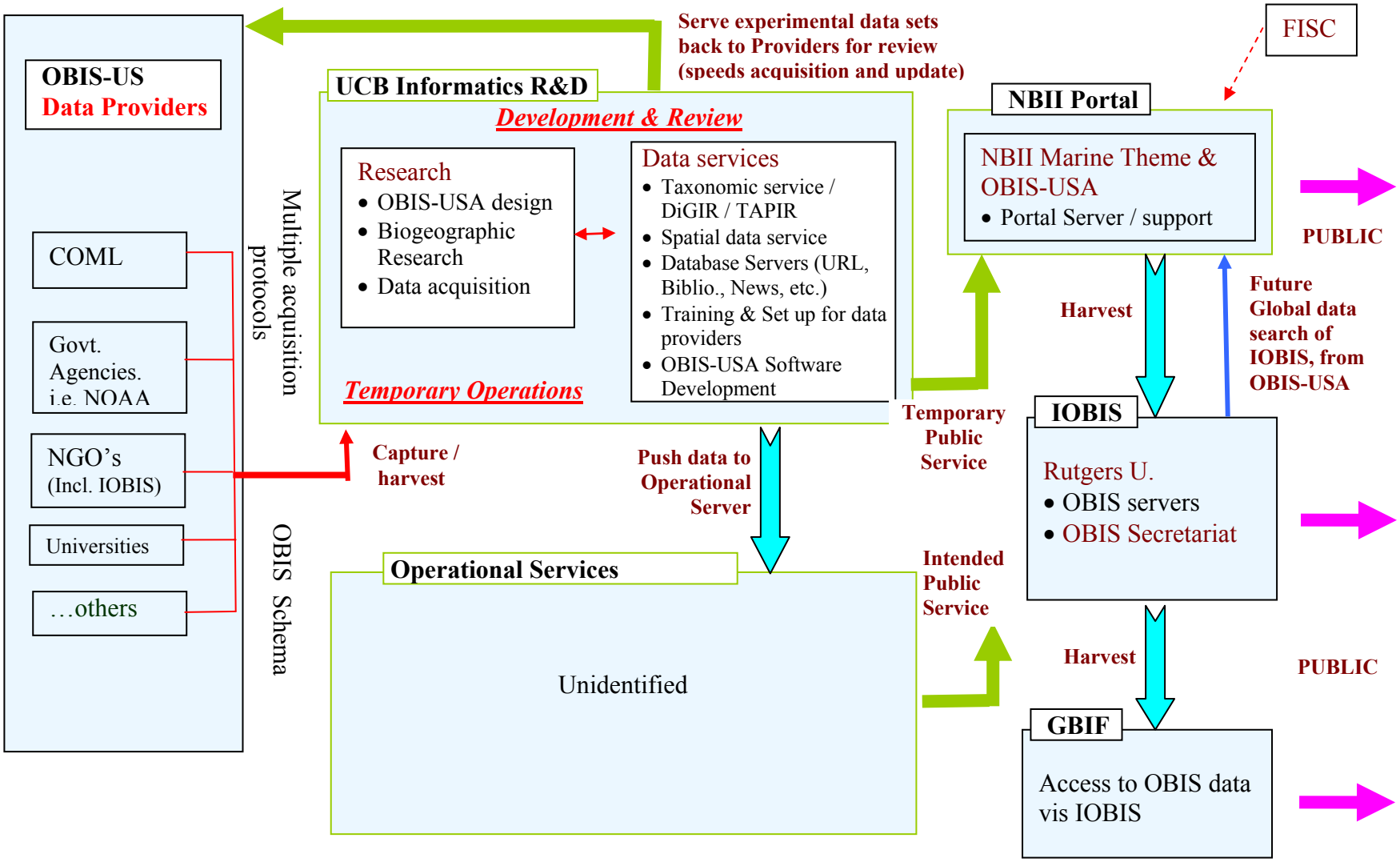
1. Interactions with data providers; finalized 3/5 in progress.
2. Project completion tasks.
3. Continued data acquisition, QCAA, servicing, and update activities.
4. Maintained existing infrastructure services.

Infrastructure Description

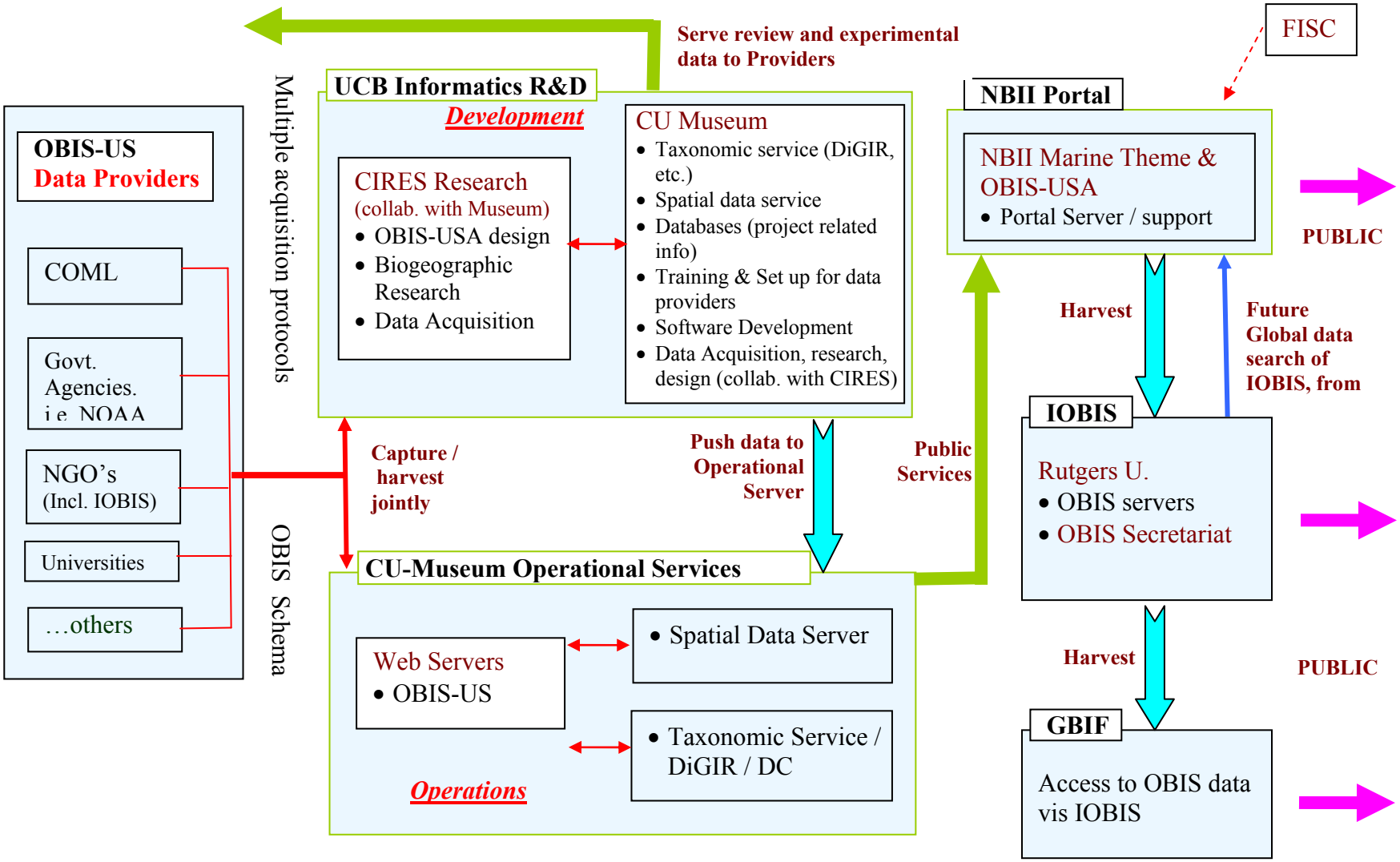
(see system diagrams, below)

- NBII Data Portal
  - NBII Marine Theme Site
  - OBIS-USA information pages
  - OBIS-USA search page
- Primary Data Server (via OBIS-USA services)
  - HTML service
  - DiGIR/Apache service
  - ArcGIS ArcGIS Desktop, Server and IMS for spatial data services
  - IIS and Framework service for software development (UC Davis)
  - Remote access for maintenance and development
  - OBIS-USA results page
- ESRI ArcGIS Data Server (via OBIS-USA services)
  - Selectable GIS raster background layers
  - Selectable GIS vector boundary layers
  - Point plots of OBIS-USA search results
  - Zoom and Pan
  - Window on OBIS-USA sub-regions
- Provider Resource and Development Page
  - QC service for Provider review and approval
  - Facilitate discussions with Providers
  - Encourage new data contributions
  - Solicit application demos
  - Collaborative information and links to OBIS resources
  - Project documents
  - Related resources

Current System Status



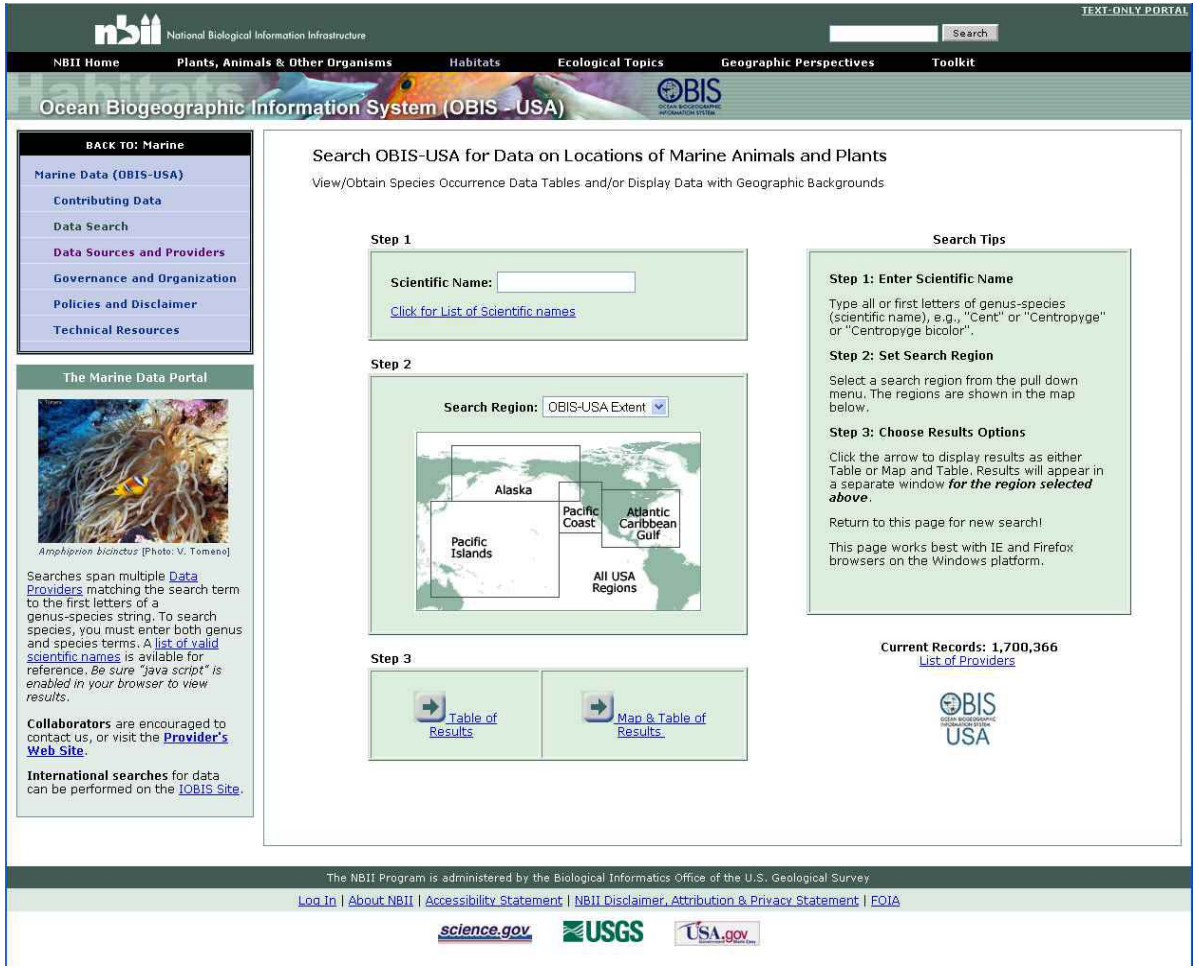
Proposed (December Scope to NBII)



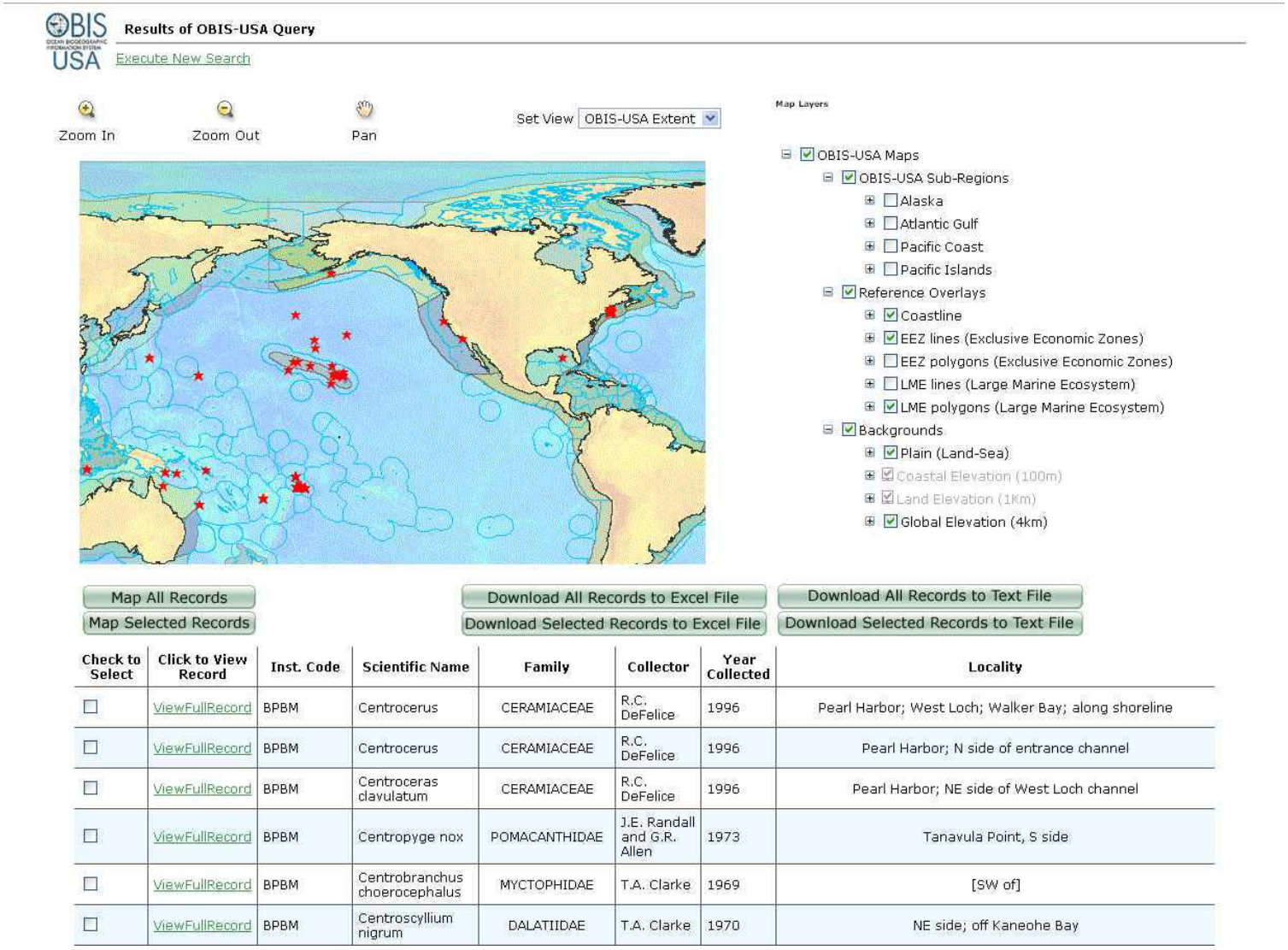
OBIS-USA Web Site (June, 2008)

OBIS-USA Search Page in the NBII Marine Theme

http://www.nbii.gov/portal/community/Communities/Habitats/Marine/Marine_Data_(OBIS-USA)/Data_Search/



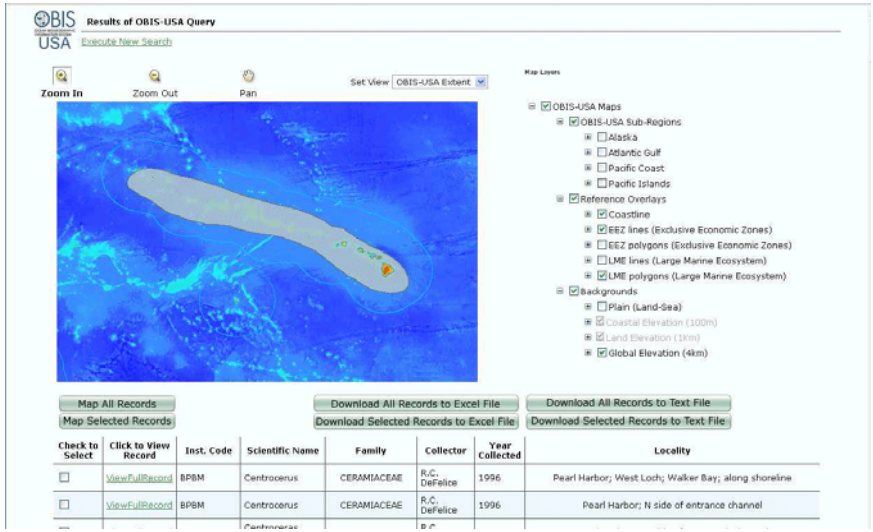
OBIS-USA Results Page served from CU Server



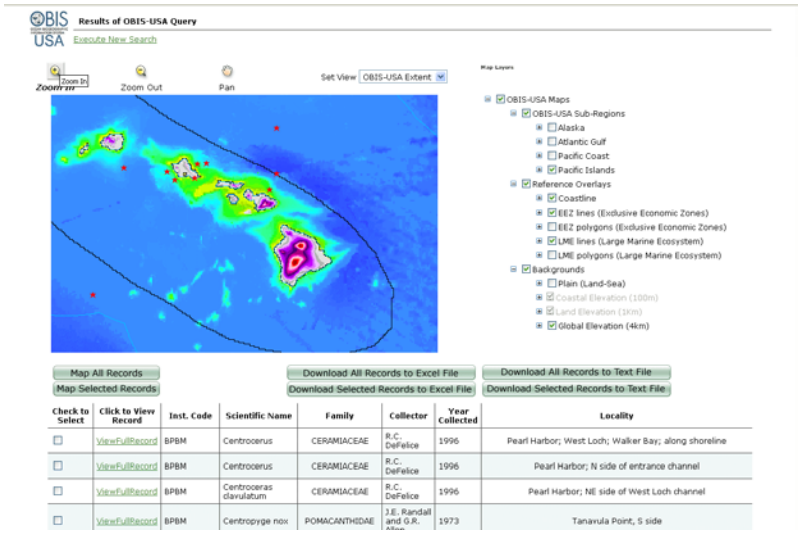
Check to Select	Click to View Record	Inst. Code	Scientific Name	Family	Collector	Year Collected	Locality
☐	ViewFullRecord	BPBM	Centrocerus	CERAMIACEAE	R.C. DeFelice	1996	Pearl Harbor; West Loch; Walker Bay; along shoreline
☐	ViewFullRecord	BPBM	Centrocerus	CERAMIACEAE	R.C. DeFelice	1996	Pearl Harbor; N side of entrance channel
☐	ViewFullRecord	BPBM	Centroceras clavulatum	CERAMIACEAE	R.C. DeFelice	1996	Pearl Harbor; NE side of West Loch channel
☐	ViewFullRecord	BPBM	Centropyge nox	POMACANTHIDAE	J.E. Randall and G.R. Allen	1973	Tanavula Point, S side
☐	ViewFullRecord	BPBM	Centrobranchus choerocephalus	MYCTOPHIDAE	T.A. Clarke	1969	[SW of]
☐	ViewFullRecord	BPBM	Centroscyllium nigrum	DALATIIDAE	T.A. Clarke	1970	NE side; off Kaneohe Bay

Background (raster) and Reference (vector) Layers

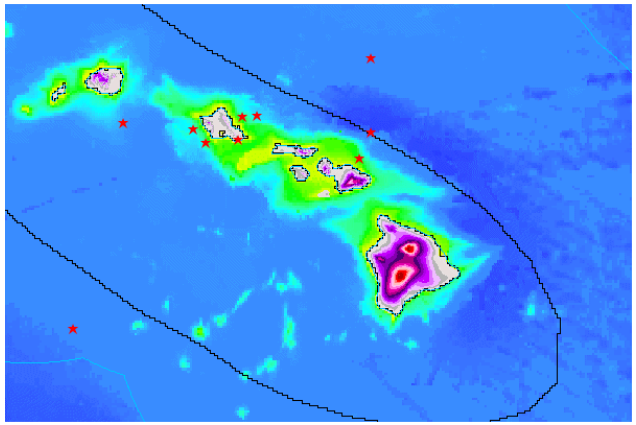
Hawaiian Island Chain: EEZ and Large Marine Ecosystems



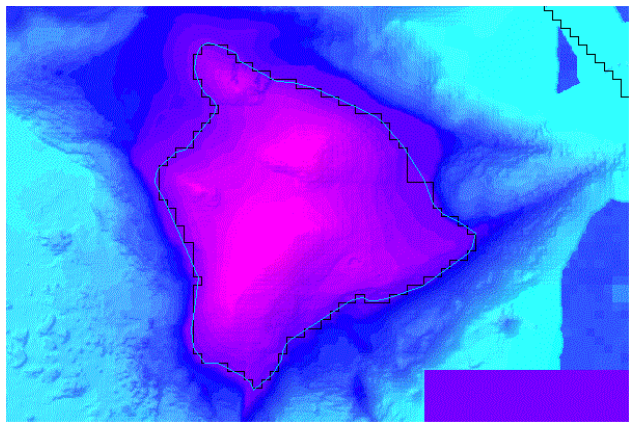
Hawaiian Islands: LME and Medium Res Topo



Medium Res. Topography / Low Res. Bathymetry

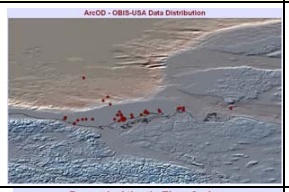
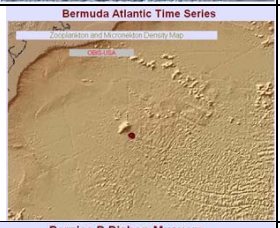
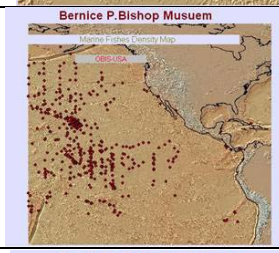
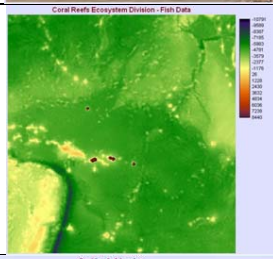


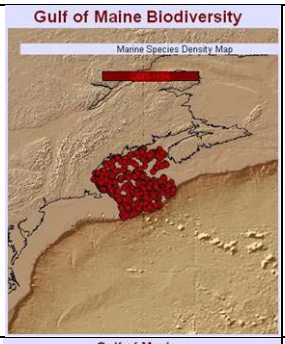
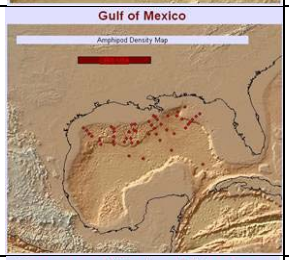
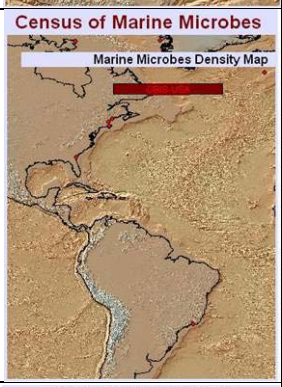
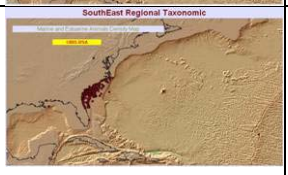
Hi Resolution Topography & bathymetry

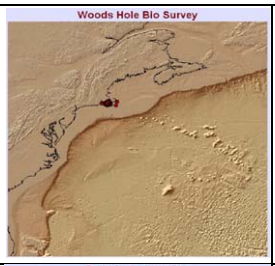
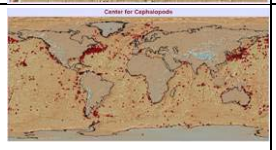
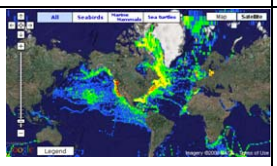


Data Providers as of June 13, 2008

http://obis-usa.colorado.edu/dev

	Provider / resource *(click for QC Report)*	Updates through 06.11.07	Reasons for updates	Volume
1	Arctic Ocean Diversity (ArcOD) QC-Report 	09.22.2006 03.22.2007 01.09.2008	Acquisition. Corrected missing field. New data added.	**177** **23,147**
2	Bermuda Atlantic Time series study Zooplankton Census (BATS) QC-Report 	12.05.2006 09.12.2007	Acquisition. Corrected missing field.	**635**
3	Bernice P. Bishop Museum (BPBM) QC-Report 	09.22.2006	Acquisition.	**7,881**
4	Census of Marine Zooplankton (CMarZ) QC-Report 	10.27.2006 03.22.2007 02.08.2008	Acquisition. Corrected missing field New data added.	**1,467** **3,974**
5	Coral Reef Ecosystem Division (CRED) QC-Report 	03.22.2007 10.10.2007	Acquisition. Replaced by provider.	**2,572** **5,736**
6	Gulf of Alaska (GOA) QC-Report 	03.07.2007	Acquisition.	**422,150**

7	Gulf of Maine Biodiversity Information System (GMBIS) QC-Report		09.22.2006	Acquisition.	**6,155**
8	Gulf of Mexico (GoMex) QC-Report		11.16.2006	Acquisition.	**45,558**
9	Census Of Marine Microbes, International (IcoMM) QC-Report		10.31.2006 11.15.2006 03.17.2008	Acquisition. Corrected missing field. New data added.	**1,467** **5,511**
10	Smithsonian Institution, Census of (SERC) QC-Report		11.02.2006	Acquisition.	**4,808**
11	Southeastern Regional Taxonomic Center (SERTC) QC-Report		11.06.2006 03.22.2007 04.22.2007	Acquisition. Corrected lat-long coding. New data added..	**2,097** **2,856**
12	Zooplankton Genomic (ZooGene) QC-Report		11.16.2006	Acquisition.	**114**

13	A Biological Survey of the Waters of Woods Hole (WoodsHole) QC-Report		01.29.2008	Acquisition.	**23,147**
14	Environmental Monitoring and Assessment Program (EMAP) QC-Report		02.14.2008	Acquisition.	**39,882**
15	University of Washington Fish (UWFish) QC-Report		01.30.2008	Acquisition.	**86,025**
*16	National Resource Center for Cephalopods (CEPH) QC-Report		02.21.2008	Acquisition. (awaiting approval for public release)	**3,172**
*17	Spatial Ecological Analysis of Megavertebrate Population (SEAMap) QC-Report		02.12.2008	Acquisition. (Provider will revise for public release)	**1,033,205**
				Final Volume	**1,700,363**

* In the last year of the project the NBII Portal site went public using the development server at CU because there were insufficient resources to establish the operational server as planned. This terminated the strategy of obtaining and serving preliminary non-public versions of data via the "Providers Resource Site," making providers more cautious about how and when to contribute. CEPHbase has gone private and a special copyright agreement may need to be worked out for them. SEAmap needs to meet their provider needs for individual citation and hence is looking into ways to dis-aggregate their providers (100's), and IOBIS wants to divide the SEAmap providers between OBIS-USA and IOBIS. How to do that without compromising SEAmap's position as a thematic aggregator/node was not worked out as of this report.

Addenda to Standard Operating Procedures

Standard Operating Procedures were developed in the first year of the project and are reported in the 2006-2007 Project Report. As work progressed various aspects of these procedures may have changed or more details may have been added. Here we discuss the changes in procedures in reference to the original SOPs, which are not duplicated here. Most of these changes can be treated as notations or special cases, as the basic principles did not change.

Data Use Policy

A draft Data Use Policy statement (Appendix VII) was transmitted to the Project Officer at NBII for approval, to be added to the OBIS-USA results page display. We recommended that this statement could be made to appear in front of the results display, preventing access but not slowing response of the system. When the user acknowledges this agreement, the display will then clear. Approval was pending at the time of this report.

Data Quality Assessment, Control, and Assurance

These terms and various principles are discussed in the first Project Report. Because of the importance of each aspect of Quality Control, Assessment, and Assurance, we suggest adopting the acronym "QCAA." A good general discussion of data quality issues is given in three GBIF publications:

- **Principles of Data Quality**
- **Principles and Methods of Data Cleaning**
- **Uses of Primary Species-Occurrence Data**

OBIS-USA Data Quality Control Policy

Data published through OBIS is intended to come from credible, authoritative sources. The scientists and institutions responsible for collecting and managing the data should be clearly named in the metadata. Before publication, the data are to pass through a series of technical controls described below, and these should be repeated every time the data are updated from its source. Any errors, such as species name misspellings, names not recognized in OBIS, and possible mapping errors, should be reported to the data provider

for review, and if necessary, correction. Thus the next time the data are published they are more correct, and the source database quality is also improved. Actual use of data in analysis is a very efffective way of finding errors. Users may contact the data source directly or OBIS-USA with such issues.

The OBIS Quality Control protocol is as follows:

1. If the required data fields are not properly filled, notification will be sent to the Data Provider. No further action will be taken until the required fields are filled.
2. If fields have questionable values, notification will be sent to Data Provider. These questionable values will be set as empty in the data published.
3. Data located on land will be reported to the Data Provider but will not be deleted unless instructed by the Data Provider, because they may represent a species in an estuary or the centre point of a location. If a Data Provider changes the values, new values will show up after the next round of crawling.
4. If species names cannot be (a) verified against known valid names in OBIS, or (b) to the OBIS taxonomic hierarchy, the Data Provider will be notified so they can check they are current and correct. Such names will be classified as 'unassigned' on the OBIS portal. People can search on these names but they will be noted as not verified. Some non-verified names may be assigned a position in the taxonomic hierarchy by virtue of their genus.
5. The portal staff will communicate with data providers to inform them of any problems and improve data quality. They will check that the data conforms to the metadata description of the dataset; *i.e.* it should have the correct number of records and species in the right geographic locations. After the data is transferred to the server from where it will be published online, a form email will be sent to the technical person and manager specified, detailing number of records obtained and missing records if applicable, time of next crawling, and any errors identified.

Geo-referencing

Datasets in OBIS vary considerably in their geo-referencing method. This can be an issue of data quality (i.e., errors), precision, or measurement accuracy. Officially the OBIS Schema calls for decimal latitude and longitude coordinates.

Typical problems in geo-referencing are:

1. Formatting errors, typically non-numeric DBMS or Spreadsheet fields or "E," "W," "N," "S" entries to indicate hemisphere.
2. Coding errors, typically blank fields, incorrect signs, or incorrect values.
3. Various conventions, such as representing the centroid of an unknown location that may even fall on land (e.g., a country or island marker).
4. Different or undocumented spheroid and/or datum (e.g., Clarke 1866 and WGS-84)
5. Numerical precision (perhaps rounded)

6. Semantic definitions, for example whether the point represents a species occurrence location, a station location, a lab location, a docking location (e.g., for a fishing boat), etc.
7. Place names without coordinates. This problem introduces subjectivities in identifying the place correctly and inaccuracy in the actual occurrence location.
8. Location by sector, for example broad geographic divisions of a Sea or Gulf. Sometimes such sectors are even oddly shaped (e.g., pie shaped) such that a uniform geospatial uncertainty cannot be specified either.

Procedures are:

1. Identify errors and report them to the Provider for correction or handling.
2. Document errors in the QCAA technical report and maps.
3. Test and offer suggestions to the Provider, as time permits.

Possible solutions:

1. Automated place-name lookup
2. Add error estimates
3. Define polygons to represent location uncertainty where a centroid would not be representative.
4. Search original records for better data

Large database files

If data are delivered by spreadsheet or DBMS file, large files may be difficult to process. They may be divided into smaller files and stored as separate DiGIR resources. Ideally, data links can be established with Provider DBMS software to stream data and updates record by record with an automated crosswalk that eliminates the need for manual processing. When separating files, a logical taxonomic or location classification (or both) should be used.

Disaggregation of data from Thematic Nodes (e.g., Seamap)
Use considerations (e.g., Seamap)

Some OBIS Providers are "Thematic Nodes" which are aggregators of data from other Providers. A problem exists in the OBIS architecture in that it does not allow metadata to be accessed for the separate providers when accessing via the combined thematic resource. There were no procedures agreed or put in place by the end of this project to handle this situation, hence there are no procedures other than to work with IOBIS and the Thematic Node to find a mutually acceptable solution. Some options are:

1. Build metadata navigation to go along with the thematic resource, for example employing the "recordURL" field as a hotlink to metadata, as recommended by SeaMap. This approach preserves and enhances the integrity of the Thematic Node integration, which may offer quality and other enhancements to the data, and their responsibility for the separate providers. It also allows multiple access to the resources, for example through the Thematic Node Website. A simple lookup table could provide navigation to

the separate metadata records. Another advantage of this approach is that a regional Node can process all of a thematic node's data for IOBIS, while limiting its local display to those fitting the regional description, and thus QCAA measures of the regional node can be applied uniformly to the full integrated resource. [This was our recommended approach.]

2. Disaggregate the data into original provider resources, with appropriate metadata using the standard DiGIR architecture, and divide servicing of these separate resources among the appropriate regional nodes. This approach employs the current (limited) DiGIR structure and allows division of resource servicing. A possible disadvantage may result if it diminishes the benefit and role of the thematic Node. If disaggregation is done by the thematic Node the advantages of #1 above can be partly retained. [IOBIS and SeaMap were working on a disaggregation protocol as of this writing.]

Project Documents

Paper documents of the project were maintained in a file system. Aside from various reference materials, a paper folder exists for each provider, containing correspondence, ingest notes, test results, and other information used to register the data resources with OBIS-USA and to construct the Technical Report. These folders should be retained for project documentation, contact information, and general tracking of different versions of data. When automated harvesting is implemented, it is highly recommended to implement an automatic log of updates that can be a digital extension of the project folder, and a sequence of files stored on disk to archive previous versions of the data.. This information can be very important for tracing back to a problem that may be introduced and may go unnoticed until a user discovers it. Being able to recover previous versions of data can then be very important.

Note: These automated procedures were not implemented in the current project, but planned for Phase-II development.

Bibliographic and other References

In the course of working with OBIS and CoML many technical and scientific papers, books and other references become known. Citations should be captured in a bibliographic database using commonly available software. If desired, this information can be ported to the OBIS-USA website as an added information resource. Using database-on-the-web protocols, the bibliography can be made searchable. Depending on resources and activities of the investigators, annotations may be included. Similarly, databases can be set up as additional 'portlets' in the NBII Marine Theme, or on the Provider Resource Site, to present dynamically updated lists of events, web-links, etc.

Note: These additional services were planned for later phases of development, and were not implemented in Phase-I.

Conclusions & Recommendations

The progress of work is reported in two project-year reports, this one serving as the final project report. A vision for the development of OBIS-USA was produced in the 2nd quarter of the first year, in a plan for phased, iterative development (Appendix I, 2006-2007 Annual Report). That plan, as updated in the IOOS proposal, remains valid in most of its recommendations and should be used as a baseline to which these conclusions and recommendations are added.

Operational Strategy

Our mission was to establish a temporary, pre-operational, regional Node of OBIS for the USA and to begin data acquisition and delivery in anticipation of future implementation of a permanent operational activity (at an undetermined location). It was a premise of the project, that a much larger effort was and is needed to establish a permanent operational OBIS Node and infrastructure at a level commensurate with national and international needs. The efforts of this project moved closer to that objective. A much larger effort is needed to keep pace with rising expectations. The budget that NBII managers estimated would be necessary for such a permanent facility was ***$4-5 million per year*** (Appendix I, 2006-2007 Project Report). The current project established a pre-operational prototype and vision for future development. Given what we have learned about this task and rising expectations, the estimated program budget cited above seems appropriate for a full-scale national level facility.

The scope of this effort was to determine the best way to proceed and to design and demonstrate necessary and sufficient components of a future system. All stated tasks of the project were met (see both annual reports), producing an initial-phase prototype system with providers representing each sub-region of the US, operating procedures for processing data and working with providers, a 3-phase infrastructure development plan, a strategic outlook, and recommendations (below) on how such a system can eventually support science. The project could not address transition to Phase II/III, despite rising national and international expectations; but instead optimized the Phase I implementation as much as possible. We do not recommend further optimization of Phase I, but as soon as it is possible to re-design and move to Phase II/III.

We feel strongly that any future development will greatly benefit from the designs, demonstrations, and conclusions produced in this project. The phased, iterative development plan that was produced early in the project should be modernized and implemented. It is clear that infrastructure issues must be addressed up front so as to stay ahead of data and applications development activities. The Standard Operating Procedures should be maintained in future efforts, changing details as appropriate for new processing environments. Finally, the strategy of offering pre-publication provider feedback and support is key to rapid acquisition and quality control of data.

The Need for Independent Servers

The technical infrastructure was largely established in the first year (see 2006-2007 report) and few changes were made to it in the second year, except to improve the presentation and functionality on web sites. Delivering services through the NBII Portal was the driving design aim. That aim requires two server installations, one for data acquisition and prototyping, and another independent server for operational delivery of final data approved by providers for public release. Serving both development and public services from the same machine produced a conflict between development and operations. We recommend that a separate operational server be established as soon as possible for the following reasons:
1. Modern IT practice recommends against developing on an operational server; the more that web services can be distributed to different hardware the better it is for system performance, stability, recovery, maintenance, and security.
2. The highly effective strategy for data acquisition, proven in prior work as well, relies on building provider confidence through early participation and contribution of pre-release datasets for internal review. This function must be served independently of the approved public resources.

Data Interoperability and Application – The Case of IOOS

An important goal of this work was to support science applications with regard to the status and future of marine resources in a changing world. Each presentation of goals for the proposed marine species occurrence data system cited this need within national programs. These research questions have been repeated many times as justification for building the needed data resources. Although the current project took on a primarily technical characteristic, its initial scope was very much to evaluate these broader strategic issues. The science that an OBIS data service was intended to support, according to its NBII sponsors, included the following concerns (see Appendix II and the "Strategy" report in the 2006-7 Annual report):

- the capability to make ecological forecasts
- assessing the health and sustainability of our Nation's ocean ecosystem
- to assess and explain the diversity, distribution, and abundance of marine life
- identify 'sentinal' species (i.e. species that are good indicators of change)
- predictive capacity to evaluate changing climate and environmental conditions.
- key biological observations for ocean monitoring
- how to couple scale and resolution of biological and physical phenomena
- support conservation and management decision-making

The scientific community and government agencies are generally unclear about how biological and environmental data should be used together to address these questions. There has been a tendency to treat the problem of "interoperability" and "outreach" to data providers as one of requiring a fully integrated, interdisciplinary data schema. However, it may be a mistaken view that interdisciplinary data need to collected, stored,

or used in a common schema because their interaction is at a higher level of aggregation. For data retrieval purposes, a simple crosswalk between systems may suffice, and for scientific purposes, the combination of models, rather than data, may be more important.

Physical and biological disciplines do not handle location in the same way. One is a measurable aspect of the environment whereas the other is an adaptive distribution with respect to the environment. Thus, while biological and environmental measures need to overlap in order to draw conclusions about their relationship, they do not need to be precisely coincident. Location for biological samples are, in part, functionally determined; and so they may not be driven strictly by physical parameters at that location. The meaning of location information must be determined from analysis, suggesting that integration should occur at the level of modeling and interpretation, not at the level of raw data. To relate biological and environmental data one must model the adaptive strategy. The study of such relationships is the field of ecology, and ecological thinking was noticeably absent from the workshop. The workshop recommended a pilot study to explore hybrid data schemas between OBIS and IOOS, whereas we suggest that normal geo and temporal referencing would be a worthy first step, and quite possibly sufficient for subsequent needs.

In fact, it is likely that environmental and biological measures, if combined in the same data schema, will need to be separated in any case for analysis. We believe that biological standards should be established as needed by the biological community, just as physical environmental standards were in their own discipline. This allows optimizing sampling design and field efficiency for the respective discipline.

Additional conclusions and recommendations are given in Appendix IIIb: *Conclusions and Recommendations from the CoML-IOOS Interoperability Meeting Regarding OBIS-USA Development*.

Final Issues: January, 2008

A list of final issues to discuss or resolve was communicated in January, 2008 because that was the when development ended and the majority of the budget had been spent (as planned), except for routine services, final data set acquisitions, and project closure. The table below indicates how each issue was addressed.

Requirement	**Resolution**
1. Additional changes to OBIS-USA site (see Appendix V-b): respond with feasibility and schedule.	Discussed with Project Manager and determined feasible resolution. Changes were implemented as budgets allowed (see notations in table).
2. Request for detailed report on data sources.	This information is provided on the Provider Resources Site and in this and the previous year's report. (See 2006-2007 Project Report, Appendix XII: List of Data Sources).
3. Resolve formatting	Table format was reviewed and optimized by CIRES

issues for tables and time lag on map: These and other performance issues were addressed within the limits of the present (Phase I) architecture.	programmer. Performance was assessed and determined to be optimal for the current design. Further optimization requires redesign, including implementation of the Phase-II architecture, new hardware, and new software. Some experimentation with the ESRI ArcGIS service could yield improvements in the background and reference layer performances.
4. Description of quality control process today and for the future:	The QC process is described in each annual Project Report. See especially the Standard Operating Procedures section of each report. Critical points about future QCAA are made in various places in these reports, including conclusions from the CoML/IOOS workshop, the "Strategy" long-term development plan (Appendix I of the 2006-2007 Project Report), and other places.
5. Minimum needs for stabilization of website	We performed a comprehensive scoping exercise with the CU-Museum for moving through the phased development to a fully operational status. The results of that exercise were presented to NBII as a joint proposal submitted by CIRES and the CU-Museum for future development. Aspects of that proposal were still being considered as of this report date.
6. Standard operating procedures for adding a database including quality control procedures. (RE: OBIS and GBIF quality procedures).	OBIS has a data quality policy at: Data Quality Control System in Place This policy consists primarily of ensuring operability of the provider data set and informing the provider of problems, but it does not detail internal procedures. GBIF data quality documents are more detailed and they are available on their website at: GBIF Data Quality ○ **Principles of Data Quality** ○ **Principles and Methods of Data Cleaning** ○ **Uses of Primary Species-Occurrence Data** In this project we implemented similar principles in a formal procedure that results in a "Technical Report" to providers. See discussion of Data Quality in this and earlier reports.
7. Review database sources and define characteristics of value (DB importance)	Characteristics of value include: ○ taxonomic representation (broad representation of species) ○ geographic representation (balanced coverage) ○ provider representation (regional coverage, CoML providers) ○ distribution of samples (natural occurrences at location) ○ current status (transfer from IOBIS) ○ operational phase (I = flat files; II = DiGIR service; III = DBMS service) ○ value for science (current topics, researcher specific, species specific)

8. Discuss database updating – iOBIS & OBIS-USA interaction	In the current project IOBIS was updating most of its resources by file transfer, while working toward automated harvesting. OBIS-USA DiGIR services were made available for IOBIS automated harvesting or file transfer if requested.
9. Approach for gaining provider permission to expose data (i.e. John's concern on Seamap)	Provider participation is predicated on trust. This can be established quickly by not exposing data before approval. That can be accomplished by serving data for review before making it public. The process highlights problems privately, builds trust, and gets quick results.
10. SeaMap	The SeaMap provider was trying to resolve criticism from their aggregated providers that users can't see the individual source metadata for giving proper citation. While SeaMap data were already being served by IOBIS in this mode, they did not want to compound the problem by serving it via NBII before this problem was solved. The then public status of our server prevented further experimentation with the SeaMap resource, which had to be removed from the active database.

Summary of SeaMap Concerns

SeaMap Comment	**OBIS-USA Response**
Terms of use are too weak	New data use policy submitted for review, to be added to the site when approved (see Appendix VII)
Data are presented as SeaMap data, but the individual data providers need to be cited and a link is needed to their sites.	In a conference call it was agreed that the Seamap data (and other similarly aggregated data) should be dis-aggregated to the original provider/resource level (primary data authority) so that the individual metadata will be exposed. In doing this, publication responsibilities should be divided between OBIS-USA and IOBIS. Seamap has some Python-based software that may be useful in this task. Additional development may be needed. IOBIS will decide what additional development is needed and how to get it done. OBIS-USA may then implement a common solution based on the IOBIS development. The result will be full exposure of the individual source/resource metadata in a user-friendly hierarchy of some kind. Thus the solution will include proper citation of thematic and regional nodes (as required in the revised data policy above).
No legible metadata for data providers	Aggregated resources must be separated so that detailed metadata can be associated with each provider.
Invalid provider metadata	Frequent updates are required to keep these fields current.
Missing fields	There may be update issues or problems with the crosswalk.
No effort data	This is an issue of the OBIS experimental design.

List of Deliverable Goods and Their Disposition

1. Annual and final reports: sent to Mark Fornwall mark_fornwall@usgs.gov and submitted to CESU in compliance with contractual requirements.
2. Project data and system: Located on the CU "OBIS-USA" development server (Dell Blade) at the CU-Museum (MCOL), contact: Robert Penn Guralnick Robert.Guralnick@Colorado.EDU
    a. Provider Resource Website and HTML service
    b. GIS data and ArcGIS server
    c. DiGIR and all provider resource files
    d. Running system operated jointly by CU-Museum and CIRES
3. Hard drive with backup of all digital files: Located at the CU-Museum (MCOL), contact: Robert Penn Guralnick Robert.Guralnick@Colorado.EDU
4. Project paper files: Transferred to CU-Museum (MCOL), contact: Robert Penn Guralnick Robert.Guralnick@Colorado.EDU
5. OBIS-USA web pages: Located at the Center for Biological Informatics, Denver on the NBII Data Portal. Contact: Janice Gordon janicegordon@usgs.gov
6. List of known Issues (bugs) in the Web and Data Services

Appendix I
Augmentation for Infrastructure Support

Submitted July 14, 2007 (funded)

Proposal / Statement of work: Rocky Mountain Cooperative Ecosystem Studies Units

Augmentation #1
"Establishing a Regional Node of the
Ocean Biogeographic Information System (OBIS)"

Proposal to augment USGS-CESU Grant # 04121HS009 to University of Colorado (Award # 04HQAG0121) to support collaboration with the CU Museum, and to extend the work performance period of the project to 6/13/08

Background and Relevance

The National Biodiversity Information Infrastructure (NBII) has responsibilities for integrating marine biodiversity data into NBII and to establish a United States Regional Node of the Ocean Biogeographic Information System (OBIS-USA). This effort is to provide national coordination for marine biodiversity data, especially for non-commercial species. Work began on pilot implementation in June 2006 through multiple small grants. Overall Program and Technical Coordination was provided by the Pacific Basin Information Node (PBIN). The University of Colorado Cooperative Institute for Research in the Environmental Sciences (CIRES) was funded to design OBIS-USA, contact and network data contributors, help set up providers, collate data, assist with proposals, and coordinate development of the node content (see attached work statement). UC Davis Information Center for the Environment (ICE) was funded to provide technical support for node infrastructure, coordinating with PBIN and CIRES. The USGS Florida Integrated Sciences Center (FISC) was funded to build an NBII Web Portal for a new Marine Theme that would also include the OBIS-USA public delivery site.

Present Status is:

- OBIS-USA has over 12 providers and 500,000 records.
- The NBII Marine Theme webpage has been developed and is nearing public launch.
- A Windows server has been installed at the CU Museum as a development and testing environment.
    - The CU Server is providing the OBIS-USA taxonomic data via a DiGIR/Apache server,
    - Spatial data are being served via an ArcGIS Server.
    - An OBIS-USA development and Provider support website is being served via Apache 2.0 and html services.
- Pilot data services (taxonomic search and mapping) have been implemented via the NBII Portal and Marine Theme via asp.net software provided by ICE. The dynamic content for the NBII portal is currently running outside the Portal, on a

PBIN Server (the DiGIR operational server) and on the CU Server (the ArcGIS server), controlled by forms entry from the Portal's OBIS-USA web site.

Long term intentions are to locate the OBIS-USA node in the Denver-Boulder area, taking advantage of USGS facilities at the Center for Biological Informatics in Denver and existing expertise at the University of Colorado.

This work relates to the Mission of the Rocky Mountain Cooperative Ecosystem Studies Unit by using and improving general expertise and methods in eco-informatics, museum science and biodiversity research at CU, and extending this capacity to national issues and priorities associated with operating a Regional OBIS Node. The work will advance the science of informatics and support for ecosystem management applications.

Proposal / Statement of Work

This proposal is to extend the current project and augment funding to support collaboration with the University of Colorado Museum (CU Museum) to provide technical and scientific expertise in biodiversity informatics and systems development; and to augment CIRES' role regarding IT/programming support and data development. There is currently a critical gap in infrastructure support and development for the OBIS-USA node and this augmentation of the project is to fill that gap. Through this augmentation grant, the CU Museum will provide technical leadership and development for the OBIS-USA server environment and data services on a par with other similar services provided by the CU Museum; in collaboration with CIRES program and technical staff.

Goal

The primary goal of this effort will be to contribute to the development of the USA Regional OBIS Node (RON) and to demonstrate a pilot capability of the NBII to serve as the US marine biodiversity data access system. The key to a successful USA RON will be to have a robust, data-rich, node with applications that facilitate data integration, analysis and application.

Objectives:

4. **Define and follow a system concept (technical development plan)**: Lead or contribute to requirements analysis focusing on critical functional requirements, and produce a well-integrated and coordinated, but parsimonious Technical Development Plan (TDP), with long and short-term objectives, addressing the following critical needs:
    a. DiGIR or equivalent services (maintenance and development) to provide taxonomic search by scientific name and common name, and delivery of results via the OBIS-USA Operational Server (at PBIN) to OBIS and the NBII public Portal site (Marine Theme and OBIS-USA).

   b. "Crawler" and "harvester" functions on the CU Server, to aggregate and re-serve OBIS data from designated USA Providers.
   c. Centralized DBMS caching functionality integrated with the taxonomic server environment (capacity to scale performance with anticipated growth in data volume)
   d. Spatial data services to provide visual display of taxonomic search results with background and reference data layers (data provided by CIRES).
   e. Enhance the public website of OBIS-USA in collaboration with FISC and the NBII Portal
5. **Enhance web services and user interface:** Work with FISC, the NBII Portal system, and CU CIRES to maintain and improve functionality of the OBIS-USA public and development websites to meet project objectives (collaboratively defined). Critical functions include:
   a. Scientific and common name search and delivery of tabular results with various display and download options for the user.
   b. Improve spatial mapping functions for search results using data layers provided by CIRES staff, and implementing appropriate tools such as Google Earth to provide regional displays (limited to the OBIS-USA regional window).
   c. Innovative tools for data display, visualization, and analysis, as available.
   d. Data quality feedback functions (for providers) via the CU Server and OBIS-USA development site (such as analytical displays, DiGIR test search capability, data exchange capability, etc.).
6. **Maintain and develop infrastructure:** Work with PBIN and CU CIRES to maintain and improve the infrastructure of OBIS-USA
   a. Perform routine maintenance of the OBIS-USA server environment, ensuring its security and operational status.
   b. Recommend and implement, as feasible, upgrades and enhancements as they become available, for example in the DiGIR and Web-GIS environment.
   c. Resolve interaction issues between the CU, PBIN, and NBII Servers, Websites, and Services (smooth flow of data and services via the NBII Portal and CU Server development site and/or other mechanisms).

Critical Development Priorities

4. Data harvesting, provider and portal issues
   a. Resolve data delivery performance issues via the NBII Portal and associated services, and establish a development schedule to keep pace with anticipated growth in data volume. This may involve development of central DBMS caching functions.
   b. Develop a harvesting function to crawl designated providers and update the OBIS-USA data cache (possibly modifying software provided by IOBIS).
   c. Ensure that data can be harvested via DiGIR or its successor, from the OBIS-USA operational server (at PBIN) by IOBIS and GBIF.
   d. Implement capability for automated update of the PBIN operational server with approved provider resources from the CU Development server.
5. Online search and mapping of taxonomic data

   a. Enhance the functionality and appearance of the taxonomic search and mapping functions associated with the NBII Portal / OBIS-USA website.
   b. Implement GIS database for visualization of taxonomic distributions
6. Web design issues/UI
   a. Enhance the functionality and design of the OBIS-USA public site (NBII Portal page), working collaboratively with FISC.
   b. Enhance the functionality and design of the OBIS-USA development site (CU Server) working collaboratively with CIRES.

Milestones

Certain tasks will begin immediately in June, or when funds are available. These include IT and programming support at CIRES and the Museum (according to availability of personnel) to resolve current system functionality problems, and Web enhancements to provide a presentable public web presence. These activities will integrate with and coordinate with ongoing project work in building data resources, applications, and networks. This early work will focus on maintaining an operational capacity that can be populated with additional data.

Other tasks will begin in September, according to availability of Museum staff. These include requirements analysis (to the extent needed and unaccomplished by that time), system re-design, and applications development to the extent that resources permit.

Budget Summary

PROPOSED BUDGET DETAILS

Institution: The Regents of the University of Colorado
572 UCB
Boulder, Colorado 80309-0572

Title: OBIS-USA continuing funding,

Principal Investigator: John Kineman
Co-Principal Investigator: Robert Guralnick

Duration: 07/01/07 - 6/31/08

A. Salaries and Wages

Museum Personnel (Co-PI Dr. Robert Guralnik)
Philip Goldstein, Technical Development

20% time, 5 months	4,138
TBD Applications Development	
20% time, 3 months	2,400
Leigh Anne McConnaughey, Web Design	
20% time, 3 months	2,400
CIRES Personnel (PI John Kineman)	
John Kineman, PI (supervisor Carol Wessman)	
Increase from 20% to 25% time, 6 months (Jul-Dec '07)	2,509
Extend at 10%, 5 months (Jan-June, '08)	4,182
Lisa Ho, Programming and DBA support	
20% time, 3 months	4,500
B. Fringe Benefits	
All personnel (20%)	4,026
C. Other Direct Costs	
Software	500
Travel	1,200
Total Other Direct Costs	1,700
E. Total Direct Costs	25,855
F. Indirect Costs	
On Campus: 17% of MTDC, predetermined for the period 7/1/06-6/30/10.	4,395
G. Total Costs 12 months	**$30,250**

Appendix II

Proposal to Upgrade Equipment and to Establish Operational Services

Submitted to NBII July 26, 2007 (not funded)

Proposal / Statement of work: Rocky Mountain Cooperative Ecosystem Studies Units

Augmentation #2
USGS-CESU Grant # 04121HS009 / University of Colorado Award # 04HQAG0121
"Establishing a Regional Node of the Ocean Biogeographic Information System (OBIS)"

Proposal: Architecture upgrade for OBIS/USA
Requested Funds: $30,000

Purpose: To provide operational server capacity and infrastructure/architecture improvements for NBII/OBIS-USA support at the University of Colorado.

Summary: Upgrades to the OBIS-USA operational server environment are needed to accommodate anticipated data throughput and to provide expected services. The current operational server must be replaced and OBIS-USA management has decided to move the operational services that support the NBII Marine Theme via NBII's data portal, to CU Boulder. The objective of this work is to establish an operational data server, secure data processing, and storage infrastructure at the University of Colorado to support the USGS/NBII OBIS-USA data portal.

CESU Mission: This work relates to the Mission of the Rocky Mountain Cooperative Ecosystem Studies Unit by using and improving general expertise and methods in eco-informatics, museum science and biodiversity research at CU, and extending this capacity to national issues and priorities associated with operating a Regional OBIS Node. The work will advance the science of informatics and support for ecosystem management applications.

Background and Relevance

The National Biodiversity Information Infrastructure (NBII) has responsibilities for integrating marine biodiversity data into NBII and to establish a United States Regional Node of the Ocean Biogeographic Information System (OBIS-USA). This effort is to provide national coordination for marine biodiversity data, especially for non-commercial species. Work began on pilot implementation in June 2006 through multiple small grants. Overall Program and Technical Coordination was provided by the Pacific Basin Information Node (PBIN). The University of Colorado Cooperative Institute for Research in the Environmental Sciences (CIRES) was funded to design OBIS-USA, contact and

network data contributors, help set up providers, collate data, assist with proposals, and coordinate development of the node content (see attached work statement).

In anticipation of expanding the CIRES/CU Museum mission to provide operational servers for OBIS-USA, CU has been asked to upgrade its current development architecture to establish an operational infrastructure that can serve OBIS-USA taxonomic and spatial databases, and html services, via the NBII Web Portal. The infrastructure enhancement is needed to establish operational servers for taxonomic, spatial data, and html services (replacing the PBIN server for OBIS-USA data services).

Presently, the infrastructure at CU for the OBIS-USA project provides a data processing (1 workstation) and development server environment (1 server). These facilities are inadequate for providing operational services. Recent tests indicated that a significant improvement in server performance is needed to proceed to Phase II of the OBIS-USA development plan and to accommodate more than a dozen providers or .5MB of data. This improvement can only be obtained by implementing a central RDBMS architecture working either together with or in place of DiGIR or its successor. Furthermore, a strategic decision was made to relocate the operational center from PBIN to CU as feasible. Pilot services now running on the development CU server resolved a major performance bottleneck, thus indicating that it would be best to locate the operational services similarly.

Long term intentions are to locate the OBIS-USA node in the Denver-Boulder area, taking advantage of USGS facilities at the Center for Biological Informatics in Denver and existing expertise at the University of Colorado.

Present OBIS-USA Infrastructure Status:

- Equipment:
  - 1 Dell 1855 Blade Server, hosted at the CU Museum
  - 1 Dell Precision 670 Workstation + monitor, at CIRES office.
- Software (server)
  - Windows Server
  - Arc/GIS Server
  - Dreamweaver MX
  - Visual Studio
  - IIS, Apache/DiGIR
  - MicrosoftSQL
  - Catalog of Life
- Software (workstation)
  - Windows XP
  - Arc/GIS
  - Dreamweaver MX
  - Idrisi Kilamonjaro
  - Apache/DiGIR (test environment)
  - Dreamweaver MX

Statement of Work

We will enhance the OBIS-USA infrastructure in the following manner:

1. Provide a redundant server capability via two additional Dell PowerEdge 1855 servers, to take over operational services previously provided via servers at PBIN. These servers will provide taxonomic data, spatial data, and html services to support the OBIS-USA main website via the NBII portal, and data networking services.
2. Add two workstations for production staff, including student aids.
3. Maintain the current development environment, enhancing functions to provide backup and recovery services for the redundant operational servers and workstations. This will require addition of a data storage unit.
4. Integrate centralized RDBMS services with the taxonomic server environment.

Additional Equipment:

- 2 Dell 1855 Blade Servers (used) [redundant operational server array]
- 2 Dell 670 Workstations (used) + monitors (note: a laptop may be substituted for one of the workstations; TBD)

Additional Software:

- 2 Windows Server (academic pricing)
- 2 Windows XP or Vista (academic pricing)
- 2 Arc/GIS Server (for redundant operational servers)
- 2 Arc/GIS Desktop (for workstations)
- 1 Idrisi Lab Kit (network server for all users)
- MicrosoftSQL Server, and/or Postgress RDBMS (all users)

Operational Server(s): It is necessary to establish operational servers for OBIS-USA to replace the operational server previously at PBIN. This change coincides with other infrastructure upgrades described below. The new servers will be located at the University of Colorado, Museum and housed in a common Dell "Blade" environment. It is best to procure a matched pair of servers for this function, to allow redundant servers to be set up in the future. Initially one operational server will be established. When functioning, an redundant server will be configured for manual switchover on failure. As resources permit, these functions will be automated in the future.

Data Security: As we implement operational services via the CU servers it is essential to ensure data security. Server redundancy (above) is part of that security, but before that we need backup capability for all servers and workstations. The current development server can function also as the backup server, with additional "hot-swap" storage capacity. While OBIS-USA does not intend to operate as a data center for taxonomic data, it will be caching data and should provide backup/recovery of the cache. This can be an informal security backup for data providers as well, or at least a means for recovering the current status of public distribution of their data holdings. In addition, OBIS-USA is building a resident spatial database for visualization and overlay. It will eventually expand that capability for spatial analysis and modeling. These data resources require backup and security.

Central RDBMS: The DiGIR environment and its future enhancements functions best using a central RDBMS. This reduces search time when many resources are being accessed during a query. Also, a central RDBMS will allow many functions to be implemented more easily, including more complex and meaningful searches and displays of data, harvesting capabilities, and quality control tools. The performance limit of the present architecture, relying on DiGIR's built-in resource query capability on flat files or remote resources has been reached. OBIS-USA has orders of magnitude more data to acquire and serve to realize its proper role. This can only be done with a central RDBMS services.

Workstation data processing: The OBIS-USA development team consisted of two part-time employees in the initial year. It has now expanded to include three additional part-time employees and a student aid. While most of these personnel have computing resources available, it is important to do production work locally to ensure proper work flow, backup and recovery of work, overall coordination and the ability to turn over personnel as careers change without interrupting production processes. By providing workstations we can share software and procedures and implement better work flows and quality control.

Work will be in conjunction with ongoing activities, focusing during this transition on:

- Procure hardware and software and implement infrastructure/architecture changes described here. Revise procedures as needed.
- Move provider data from separate disk files into RDBMS tables, one for each provider schema. The OBIS schema will be instantiated in a table, with relational capabilities,
- Relate tables to provide needed crosswalks between Provider schemas and the OBIS Schems. This function is currently accomplished in DiGIR (via XML).
- Ensure proper SOAP request processing, either through DiGIR/TAPIR or separate functionality (requires experimentation and benchmarking)
- Add test and final data as available.
- Revise data ingest/harvesting procedures
- Revise and test implementation via the NBII portal as appropriate.
- Implement improved capabilities for the contributor web-site based on these upgrades.

Note: This proposal does not include funding for ongoing operation of OBIS-USA. This is a one-time upgrade.

Budget Estimate

PROPOSED BUDGET
DETAILS

Title: OBIS-USA continuing funding,
Duration: 07/01/07 - 6/31/08

Institution: The Regents of the
University of Colorado
572 UCB
Boulder, Colorado 80309-0572

Principal Investigator: John Kineman
Co-Principal Investigator: Robert Guralnick

A Salaries and Wages

Museum Personnel (Co-PI Dr. Robert Guralnik)	
System Administrator	
10% time, 5 months	2,069
DiGIR/DBMS Administrator	
10% time, 5 months	2,069
Web Development	
20% time, 3 months	2,400
CIRES Personnel (PI John Kineman)	
Programming	
50% time, 1 month	3,750
Total Personnel	10,288

B Fringe Benefits

All personnel (20%)

C Other Direct Costs

Dell Hard Drive Storage unit (Stallard Tech.) Qn: 1	1,202
Dell 1855 PowerEdge servers (Stallard Tech) Qn: 2	6,006
Dell Precision 680 Workstations Qn: 2	2,358

	Dell Hard Drives (300GB) Qn: 4	880	
	Chemei CMV 221D monitor Qn: 2	650	
	Windows Operating System (Server and XP/Vista) Qn: 4	320	
	ArcGIS (site lisence for CU) Qn: 4	440	
	Idrisi Andes Labkit (15 seats)	1,600	
	Dreamweaver Qn: 1	70	
	ENVI (site lisence for CU) Qn: 2	40	
	Mathematica (free University lisence) Qn: 2	0	
	Total Other Direct Costs		13,566
E.	**Total Direct Costs**		25,912
F.	**Indirect Costs**		
	On Campus: 17% of MTDC, predetermined for the period 7/1/06-6/30/10.		4,405
G.	**Total Costs 12 months**		**$30,317**

Equipment Quotes from Stallard Technologies

Note: We purchased the original OBIS-USA Development server (Dell 1855 Blade) from Stallard and we have had very good experience with the company. They are easy to contact for service and very responsive. By going with all used equipment, under 3-year warrantee from Stallard, we can accomplish this upgrade on a very modest budget.

1. Dell PowerEdge 1850 Blade Servers (2) = 2 x $3,003 = $6,006

Dell PowerEdge 1855 - Blade Server - PE1855

View larger image

Dell PowerEdge 1855 - Blade Server - PE1855
Pre-Owned, STI Certified Excellent Working Condition
Blade Only - Enclosure Sold Separately (Required to Operate)
2 X Hot Swap HDD Bays | 2 X 1000 Integrated Ethernet
1 Year STI Limited Warranty Included
FREE SHIPPING (within the 48 contiguous US States)
Price: $3,003.00

Product options:

Processors:	Dual 3.6GHz XEON 1MB Cache 800FSB
Memory:	8GB of ECC Memory
Hard Disk Drive 1:	300GB 10K U320 SCSI Hard Drive
Hard Disk Drive 2:	300GB 10K U320 SCSI Hard Drive
Warranty:	3 Year STI Limited Warranty

In stock: 79 items
Amount: 1

Sign in to purchase

2. Dell Storage / Backup Unit (bare) 1 x $1,203

Dell PowerVault 220S (3U Rackmount / U320)

View larger image

Dell PowerVault 220S (3U Rackmount / U320)
Pre-Owned, STI Certified Excellent Working Condition
2 X U320 Modules | 2 X Power Supplies
14 Hot Swap Sleds (Or Select Drives)
Front Bezel Not Included
1 Year STI Limited Warranty Included
FREE SHIPPING (within the 48 contiguous US States)
Price: $1,202.40

Product options:

Included Standard:	2 x Power Supplies / 2 x U320 Modules
Hard Drives:	None (Blank Fillers Only)
RAID:	Perc 4/DC/PCI-E
External SCSI Perc to PowerVault:	3ft Cable / 3' VHDCI68 M to .050 DB68M
Rail Kit:	None
Warranty:	3 Year STI Limited Warranty

In stock: 17 items
Amount: 1

Sign in to purchase

3. Hard Drives for Backup Unit 4 x $220 = $880

SAS Hard Drives - 300GB Maxtor or 146GB 15K Seagate (Special)

View larger image

SAS Hard Drives - 300GB Maxtor or 146GB 15K Seagate (Special)

Capacity: 146GB or 300GB, Depending on Selection
Speed: 15K or 10K, Depending on Selection
Type: SAS U320 SCSI
Part Number: JP620 / G8774
Pre-Owned Hard Drives, Scrubbed & STI Certified Working
Drive Sled: Included
30 Day STI Limited Warranty Included
FREE SHIPPING (within the 48 contiguous US states)

Price: $219.00

Product options:

Select: (•) 300GB 10K (Maxtor / G8774) ( ) 146GB 15K (Seagate ST3146755SS / JP620)

In stock: 45 items
Amount: 4

Sign in to purchase ▸▸

4. Dell Workstations: $1,679 x 2 = $2,358

Dell Precision 670 (Dual 3.6 / 4GB RAM / 2 x 400GB SATA HD)

View larger image

Dell Precision 670 (Dual 3.6 / 4GB RAM / 2 x 400GB SATA HD)
Preconfigured and priced to move - these won't last long
Pre-Owned, STI Certified Excellent Working Condition
650 Watt Power Supply | 3 X Fixed HDD Bays
DVD/CD-RW & Floppy Drive | RAID | OS Not Included
1 Year STI Limited Warranty Included
FREE SHIPPING (within the 48 contiguous US States)

Price: $1,678.80

Product options:

Included Standard: (•) 1 x 650 Watt Power Supply / On-Board RAID / DVD/CD-RW & Floppy
Processors: (•) Dual Intel Xeon 3.6Ghz w/ 1MB Cache / 800MHz FSB
Memory: (•) 4GB of PC-3200 ECC Memory
Hard Drives: (•) 2 x 400GB SATA 7200 Fixed HDD
Video Card: (•) NVidia FX-3400 PCI-E 256MB Video Card
Warranty: 3 Year STI Limited Warranty

In stock: 8 items
Amount: 1

Sign in to purchase ▸▸

Appendix III
Scope of Work for Continuation of OBIS-USA
Submitted to NBII, January 30, 2008 (pending)

Plan Summary for OBIS-USA, 2008

The University of Colorado Museum of Natural History (CUMNH) and the Cooperative Institute for Research in the Environmental Sciences (CIRES) will collaborate to further develop OBIS-USA, contributing the following components:

CUMNH – Technical Development

A. Occurrence Database Description and Implementations

1. The core technology is an RDBMS Species Occurrence Database that harvests data from DiGIR-based providers and other suitable exchange mechanisms according to Provider capabilities. The DBMS will be queryable on the Web and linked to an online mapping application so that users can retrieve map and tabular data. See below for more information on environmental and other contextual GIS data layers associated with the map. Details of the architecture are listed in bulleted form below:
    a. Capacity of ~50 million occurrence records
    b. Store, process and retrieve spatially-referenced occurrence data
    c. Maintain all details at the individual provider dataset level
    d. Enable diverse query capabilities at provider and aggregate level.
    e. Query terms and output formats support direct user queries, web access, DarwinCore2-based access and integration with GIS and other analysis tools.
    f. The Occurrence Database is integrated with an automated data retrieval process (discussed below) that populates the database with participant data on the terms of data exchange arrangements with participants.
    g. The Occurrence Database implementation also provides backup and security, reporting on data integrity and completeness indicators, the ability to track provider metadata and update dates, indexing and performance tuning, and usage statistics tracking.
2. The actual hardware, software and infrastructure for developing OBIS-USA data aggregation and storage functions is as follows:
    a. PostgreSQL open-source relational database manager as a scalable RDBMS platform for occurrence data with the postGIS spatial data extension to store spatial references and perform spatial operations.
    b. We will install one new additional blade server running Windows Server 2003 as the primary database server.
    c. A second existing blade server, that will primarily serve spatial datasets, will be configured as backup to the database server.
    d. Application language tools will be selected as needed, including PL/pgSQL for procedures internal to the database, and PHP for additional database-related programming interfaces.

3. Database implementation will begin immediately with postgreSQL/postGIS installation, data model construction, and data load and query tools.
4. Additional tools for data management, performance tuning and statistics tracking will require a limited requirements definition and prioritization, in order to optimize programming time invested in these tools. These tools include providing up to date information on number of records being provided by OBIS-USA in total, number of records from individual providers, download information in total and by provider, etc. These tools will be delivered by end of the 1 year project.

B. Online Mapping, Spatial/Environmental Datasets, GIS layers and metadata storage and access.

1. We will develop an online map service linked to the RDBMS.
   a. The online map service will allow users to interact with the map or tabular results from the DBMS to view, move or delete records that are in a temporary returned record set.
   b. This service can also call other web services for analysis of species occurrences that have been viewed and validated by the user.
   c. The analysis web services will include a tool for extracting environmental data "underlying" species occurrences, and tools for documenting species sampling quality.
2. The most likely online mapping application we will utilize is Mapserver, given that there is an existing codebase implementing many of the desired features from David Neufeld's previous work that features this online mapping application. Mapserver also has an excellent raster catalog system.
3. We will implement quality control procedures that document each dataset, beyond its own metadata, to check it into OBIS-USA, track its source, status, content and update conditions.
4. We will begin with raster catalog development utilizing tools available in online mapping applications and evaluate more automated solutions as datasets and applications grow.
5. In the second half of year 1, CUMNH will assess the scale of dataset serving challenges, and recommend a scalable solution for future OBIS-USA growth.

C. Web-site design, development, deployment, and maintenance

1. Web site design, deployment and maintenance will be covered by CUMNH staff who manage Blade servers and web servers deployed on those Blades.
2. A web developer will help assure that web site design meets NBII expectations in the NBII portal and to assure consistency across all web pages that are developed for search and retrieval among users and providers. The overall site design will be modeled after the Seamap site (http://seamap.env.duke.edu/).
3. The web developer will also work with database developer to provide web-based user interfaces.

CIRES – Content Evaluation

CIRES staff will as partners in OBIS-USA development, continuing to assist NBII with overall program objectives. Content evaluation will involve data needs, quality, and applications assessment. Two primary areas of support will be:

A. Data Networking and Ingest

1. Maintain and expand provider contacts; improve data acquisition where possible. Work with providers to decide upon remote DiGIR installation for harvesting efficiency or local ingest of data that is then mapped to DarwinCore2 specifications.
2. Identify 5 to 6 new sources of marine biodiversity data (eg. Gulf of Marine dataset http://research.usm.maine.edu/gulfofmaine-census/biodiversity/)
3. Work with Museum staff to ingest new data/metadata into OBIS-USA
4. Identify further needs for key demonstration data (collaboration with CoML and NBII)
5. Participate in IOOS/DMAC and other data standards for to ensure interoperability with IOOS, GCMD, FGDC, and other applicable data and metadata standards.

B. Biogeographical information and applications development

1. Identify sources of environmental and eco-geographical data for reference and analytical use and provide those sources to the development team. Resources likely to include seabed data (http://woodshole.er.usgs.gov/project-pages/aggregates/MARPseabed.htm) that provide physical characteristics of the ocean floor and physical properties of water itself (eg. sea surface temperatures; http://geochange.er.usgs.gov/data/magsst/magsst.html)
2. Explore relationships between taxonomic data and environmental or ecological data and assist the CU Museum with deploying tools that best meet user needs.
3. Identify further research potential for OBIS-USA data (list of science questions).
4. Assess potential OBIS-USA user base.
5. Explore potential applications of OBIS-USA data using workstation tools to produce examples of possible analytical outputs and to help guide on-line tool design and specification.

Team Member Roles

A. CUMNH Staff

1. Dr. Robert Guralnick will manage the implementation team and report to project sponsors. Dr. Guralnick will coordinate with sponsors to define and focus technology deliverables to meet sponsor objectives and will guide the plan, priorities and quality control of team member activities to meet deliverables and objectives.
2. Philip Goldstein will lead the implementation of the relational occurrence database including constructing, populating and maintaining data storage, operations and performance. He will build the tools for query, presentation, download and other database functions on the portal. He will lead the

technology relationship with providers to maintain compatible, complete and current data from each contributor.

3. David Neufeld will lead the geospatial technology to serve maps on the portal for user navigation of environemental themes and geographic extents. Dave will integrate map query and presentation with the occurrence database so query parameters and the presentations of results integrate geography, occurrence, environmental data and time.
4. Leigh Anne McConnaughey will work with other project participants to develop the website, as described above in the "Web-site design, development, deployment, and maintenance" section. She already has NBII portal training and experience with OBIS-USA, and has worked with Philip Goldstein on data query and result returns on other projects.

B. CIRES Staff

1. Dr. John Kineman will manage the content evaluation team, coordinating with the RON manager on project goals, networking new data sources, and directing related research and analytical activities at CIRES aimed at addressing data quality and compatibility issues, including user data and product needs.
2. Dr. Carol Wessman will apply extensive expertise in spatial ecology to considerations of experimental design regarding OBIS-USA data with respect to the applicability of data to science and potential products.
3. Rozita Abdul-Williams will lead data ingest and processing, collaborating with Museum staff to acquire and ingest OBIS-USA data and metadata. She will conduct spatial analysis supporting data quality assessment and control, and applications product development.

Project Timeline (UCMNH Technical Development):

Months 1-6:

a. Develop and test remote DiGIR provider installations and accumulate local records into the RDBMS for targeted data providers. Expect 10 million records at end of Month 6.
b. Develop and test web-based access to the RDBMS to test query and result return capacity for text-based queries. Prototype complete Month 6.
c. Develop map and tabular result functionality so users map and download results from queries. Basic functionalities done at end of Month 6.
d. Develop new website look and feel matching the existing Seamap site.
e. Begin development of tools that use environmental raster data, species occurrence data, and map functionalities. Tools include "environmental extraction" to determine environmental conditions at places where species occur, and potentially tools that return information about quality of sampling in relation to species richness. Not expected to be done by end of Month 6, but will be prototyped at Month 12 (see below).

Month 6-12:

a. Completed infrastructure that is automatically ingesting new records from remote providers with DiGIR installed. Operational at Month 12.
b. Completed map portal allowing users to move between tabular results and map to visually validate and refine occurrence locations prior to analysis. Operational at Month 12.
c. Completed testing performance of query and results returns for large database.
d. Completed development of analysis tool(s) and test performance as prototypes/demonstrations.
e. Completed upload of environmental and contextual ocean GIS layers.

Project Timeline (CIRES Content Evaluation):

Months 1-6:

a. Data content evaluation – selection (in collaboration with Project Manager) of 5-6 key datasets for testing needs and applications.
b. Ancillary data requirements – identify environmental and ecological datasets required for quality assessment and applications development
c. Assess potential of OBIS-USA data and capabilities, producing examples for presentation.

Month 6-12:

a. Final assessment of data potential
b. Final product examples
c. Recommendations for future development

Compliance with standards

The combined team efforts will be represented in IOOS/DMAC and other appropriate data standards fora to ensure interoperability with IOOS, GCMD, FGDC/ISO, and other applicable data and metadata standards. System development and content management will comply with NBII data and metadata standards and applicable OBIS standards, including the OBIS Schema, which is based on Darwin Core II. OBIS is listed in the "Guide for IOOS Data Providers" as representing one of the compliant standards. We expect to contribute to the further development of these standards and vocabularies. Geospatial data incorporated into the OBIS-USA system for public access will be prepared and documented according to the above standards.

Project Deliverables:

The end product will be a demonstration showing how aggregated species occurrence data can be used to determine ocean biodiversity and track how environmental changes may affect that biodiversity. More specifically the end demonstration product includes:

a. A new set of providers delivering data to OBIS-USA with a focus on 5 to 6 new, key data sets that contain high quality data. Targeted datasets include the Gulf of Maine (http://research.usm.maine.edu/gulfofmaine-census/biodiversity/) dataset.
b. A stable infrastructure that includes a federation of providers delivering DarwinCore compliant records to an RDBMS system.
c. A website similar to Seamap (http://seamap.env.duke.edu/) in look and feel that allows users to query the RDBMS system and return both tabular occurrence data

and mapping results simultaneously. Map and tabular results are correctly linked so that map operations can update the local record set and vice versa.

d. A set of environmental and other contextual GIS data that can be linked through the map application to occurrence data. We will focus on sources like seabed (http://woodshole.er.usgs.gov/project-pages/aggregates/MARPseabed.htm) that provide physical characteristics of ocean floor along with the physical properties of water itself (eg. sea surface temperatures; http://geochange.er.usgs.gov/data/magsst/magsst.html).
e. A set of analysis web services for performing operations like calculating species richness, determining environmental conditions for a set of species occurrences, or determining ecologically relevant variables that limit species distributions.

Appendix IV
CoML-IOOS & OBIS Interoperability Workshop(s)

Proposal

Title: ***US Census of Marine Life National Committee "Ocean Observing: Using The Census of Marine Life as the Biological/Biodiversity component of IOOS"***

Mark D. Fornwall
Location TBD
Proposed Date: October, 2007

Recognizing the importance of the oceans, coasts, and Great Lakes to the United States, Congress enacted the Oceans Act of 2000, which created the U.S. Commission on Ocean Policy (USCOP). Subsequently, in 2007, the USCOP, supported by the Interagency Committee on Ocean Science and Resource Management Integration (ICOSRMI), released *Charting the Course for Ocean Science in the United States for the Next Decade : An Ocean Research Priorities Plan and Implementation Strategy (ORPPIS)*. This interagency effort presents the national research priorities that focus on the most compelling issues in key areas of interaction between society and the ocean.

Among the three central elements of the ORPPIS, two, "the capability to forecast key ocean and ocean-influenced processes and phenomena;" and "deploying an ocean-observing system;" directly relate to the development and operation of the Integrated Ocean Observing System (IOOS). Fundamental to the success and effectiveness of IOOS is the integration of not only physical and chemical ocean information but also biological information consisting of both spatial and temporal species data. Biological information and biodiversity data provide a unique view of the oceans and are essential to generate new hypotheses about the overall health and sustainability of our Nation's ocean ecosystem, as well as the entire global marine ecosystem. The "Deployment of a robust ocean-observing system that can describe the actual state of the ocean, coupled with a process to synthesize observational data, will fundamentally alter society's view of the ocean environment."

The Census of Marine Life is a growing global collection of biological information to assess and explain the diversity, distribution, and abundance of marine life in the oceans. The data and information that CoML collects provides a rich context for the biological realm and insight into the biodiversity of the oceans. This data is an indispensable and vital component of an IOOS and identifies sentinel species to enhance the predictive capacity of the system to evaluate changing climate and environmental conditions.

Under the auspices of the U.S. National Committee of the Census of Marine Life (CoML), the Consortium for Oceanographic Research and Education (CORE) is holding a workshop to demonstrate the efficiency and value of incorporating the Census of

Marine Life, and the biological and biodiversity data of their field projects, into IOOS. This undertaking requires a unique collaboration among biological, chemical, and physical oceanographers, decision-makers, and data managers. The workshop will bring together experts from a variety of scientific disciplines, academia, and management agencies to share information with regard to four central questions:

- How can current biological monitoring programs be integrated/ interrelated into IOOS?

- What kind of data would be appropriate and essential to the success of ocean observing; what existing data would be appropriate; and what current data gaps need to be filled to provide an effective IOOS?

- How can the various scales and resolution of the static biological, biodiversity data be matched within the dynamic chemical and physical oceanographic data in an Integrated Ocean Observing System? (e.g. special data vs. temporal data)

- How can these data be used to help develop predictive models to assist decision-makers meet their conservation and management mandates.

The IOOS/CoML Workshop will attempt to determine the best means to integrate and utilize biological data into IOOS. The workshop will develop several pilot programs where existing regional ocean observing system data can be integrated with CoML data and other biological monitoring data to develop predictive data management tools for policy makers and in so doing identify the obstacles and data gaps to such an approach. The ultimate product of this workshop will be to make progress and improve upon the ORPPIS goal that the "deployment of a robust ocean-observing system that can describe the actual state of the ocean, coupled with a process to synthesize observational data, will fundamentally alter society's view of the ocean environment."

Invitation from CoML

Dear Dr. Kineman,

You are cordially invited to attend the workshop entitled *Biological Ocean Observing: Exploring components of IOOS from the perspective of Census of Marine Life*, sponsored by the Census of Marine Life (CoML) within the U.S. The Census of Marine Life (CoML) is an international effort to help assess and explain the diversity, distribution, and abundance of life in the oceans. Toward this goal, CoML focuses on field studies that explore little-known habitats and re-examine familiar areas using innovative technologies. CoML is also providing for development of an integrated biogeographic information system, assessment of historic marine animal populations, and modeling of possible future populations. Potential benefits include enhanced management of fisheries and marine ecosystems, improved capacity to predict and mitigate impacts of human activities, as well as a much richer understanding of the marine world.

A U.S. National Committee for CoML was established in December 2002 to work with the scientific community, funding agencies and policy-makers in the United States. This year the committee elected to focus its workshop on the importance of incorporating biological data, particularly CoML data, into the Integrated Ocean Observing System (IOOS). The workshop aims to accomplish these four objectives:

1. Explore the general types of biological observing data, and identify specific data sets that should be integrated into IOOS;
2. Consider critical data gaps in biological data measurement/collection, analysis, and management that should be addressed by IOOS;
3. Identify specific strategies to advance the interoperability between CoML/OBIS and IOOS;
4. Explore the development of predictive models based on current and future IOOS data efforts to support the move towards ecosystem based management;

Fundamental to IOOS is the integration of not only physical and chemical ocean information, but also biological information consisting of both spatial and temporal species data. Your participation in this workshop will provide valuable input as we explore this topic. CoML has received funding from NOAA and the Moore Foundation for this workshop. The workshop will be held 14-15 January, 2008 at the Latham Hotel in Washington, DC.

Please find a travel announcement attached. Additional information regarding the goals of the workshop and an agenda can be found at http://www.coml.us/Dev2Go.web?Anchor=ocean_observing. We sincerely hope that you will be able to join us. Please RSVP to coml@coreocean.org by December 15, 2007. Feel free to contact the U.S. CoML Program Office at this email if you have any additional questions. We look forward to your participation in this workshop.

Sincerely,
Andrew Rosenberg Ph.D.

Conclusions and Recommendations from the CoML-IOOS Workshop(s)

Held at the Latham Hotel, January 14-15, Washington, DC

John J. Kineman, Ph.D.
January, 2008

Workshop Title: Biological Ocean Observing: Exploring components of IOOS from the perspective of the Census of Marine Life

Note: These conclusions and recommendations are also in response to the ½-day meeting that followed the workshop, focused specifically on the role of OBIS.

Critical needs for integration of OBIS and OBIS-USA with IOOS:

1. Data quality: IOOS expects a much higher level of data quality than currently exists in OBIS because they are used to physical data sets that are more standardized in their original measurements because they have had better support from major international observing programs and systems. This alone makes direct integration of measures difficult. Issues that must be addressed:
    a. Representation of error: points are deceptive unless there are error bars (or circles). This can be added and displayed, however it may require a change in the OBIS schema, which restricts this information to a limited set of error categories.
    b. Some grouping on resolution should be added, so that users can access data of a given quality. Presently location accuracy or knowledge is very poor in many data sets; e.g., placing a point in the middle of Russia because all that is known is that the sample was observed in Russian waters. This requires interpretation on the part of the user. Some will discard it out of hand without even knowing it was intentional. Others will use it without noticing, corrupting their product. The OBIS Schema allows for a rectangular area to specify location, but that is inadequate for most uses, overlapping land and misrepresenting either the true distribution of uncertainty or the true distribution of waters from which the sample may have come.
    c. Feedback on initial contribution data quality is an essential function of OBIS, to provide incentive for contribution, to encourage early contribution, and to facilitate the process of correcting errors.
    d. It may be possible to incorporate polygonal areas into the OBIS and/or other schemas to specify the area of likely occurrence. This especially resolves the problem where only a country or island reference is given and the likely collection area could be on different sides (some OBIS records currently label the center of the land mass in this case).
2. Linking biological and environmental data (integration with IOOS)
    a. The proper link between biological and environmental data is the subject of ecology. It is missing from the data concept and it was largely missing from the disciplinary discussion. Two enhancements are needed to make OBIS and IOOS ecological.
        i. Extend OBIS schema (or develop a second one) to handle ecological factors.
        ii. Collect information on key ecological functions. This information should be collected as close to the data source and context as possible, but is useful in any case.
        iii. For these observing systems to move toward ecology it is critical to develop niche models. Niche models are the relationship between structure (what exists at a given instant) and functions (potentials that drive change). The only function presently captured is presence. Feeding, breeding, decomposer, predation, etc. are also needed.
    b. IOOS has an incorrect concept of biological and environmental data integration. The overwhelming presumption in the meeting was that data

can and should be 'integrated' at the level of observation and raw-data delivery. This is not only unnecessary, it is counter-productive except in the context of a scientifically-driven field sampling design, i.e., a given ecological study. Even in those cases, to support secondary use of the data as presumed by a shared observing system data repository, the biological and environmental components must be handled separately. The reason is that biological entities are not physical objects as such, they are functional adaptations to a set of physical and biological conditions that vary in space and time. As such, a given location has functional not physical association with other measures at that location. Accordingly, the proper way to 'integrate' biological and environmental information is at the level of disciplinary analysis and modeling. Biological models can be associated with environmental models meaningfully, and each is the domain of its respective discipline. *For IOOS to seriously integrate biological measurements, it must seriously represent ecology in the program.*

3. Applications
    a. Applications are needed to demonstrate science using OBIS-USA data. OBIS itself can do only a limited amount of this. We must therefore network with providers and users in a strongly collaborative model, to encourage the contribution of application demonstrations. The incentives that can be offered are:
        i. Linkage with other researchers and sharing of tools
        ii. Visibility and credit in OBIS-USA (and IOBIS and IOOS)
        iii. Peer review and feedback
        iv. Possible assistance with publication (or collaborative publication among those with similar applications).
        v. Enhanced funding opportunities
    b. A limited research development activity is needed in OBIS-USA to:
        i. coordinate the provider network,
        ii. work with applications developers to get applications on the site,
        iii. collaborate in the development of certain applications where research interests overlap.
        iv. work with users to familiarize them with OBIS data and to provide introductory training.
4. Infrastructure
    a. A comprehensive approach to building up the operational capacity of OBIS-USA is required.
    b. Operations need to be under Federal management or established through a cooperative institute as a federally funded operational facility following a well-defined federal requirements specification (with annual or bi-annual reviews). Critical elements are described in more detail in recent proposals. In summary they include:
        i. Stable, robust, redundant, maintained servers at an appropriate location where similar server operations are routine.

   ii. Continuous R&D of server functions, including enhancements to DiGIR (TAPIR and beyond), harvesting and aggregation function, performance evaluation and anticipation of volume requirements.
   iii. Maintenance of the OBIS-USA Web site, including the NBII Portal and remote services. This requires continuous maintenance and update.
- c. A providers/collaborators facility must be maintained to facilitate interaction with providers, both regarding data ingest and QC and applications networking. This site and facility serves as an R&D front end to the operational facility. It can be an accessible site, with restrictions on certain aspects. Data sharing is restricted to providers for review and development. Elements include:
   i. Provider feedback on data quality, content, and compliance
   ii. Provider support in metadata preparation
   iii. Provider support in establishing server functions (if needed)
   iv. Provider training (when we know more than they do)
   v. User networking to solicit applications (users may also be providers)
   vi. User support (helpdesk, training, other assistance)

5. Standards
   a. OBIS-USA needs to be intimately involved in the development of biological and ecological standards. These are critical to building an interoperable system. Standards help integrate data within the boundaries where data integration is appropriate, and they allow models to provide scientific integration outside those boudaries. In particular:
      i. Develop a standard for recording functions, and an associated metadata standard
      ii. Develop a metadata standard for niche models
6. Funding
   a. Funding on the order of $1.7M/yr for 3 years is required to accomplish these tasks. Less than this amount can address individual components, but at the risk that without the whole package it will not be viable or useful to IOOS in the needed time frame.
   b. The most critical aspect of viability and funding for OBIS and OBIS-USA is to make the bridge into ecology. If OBIS remains solely a data service, without partnering with applications developers, it will lose value and funding in the context of IOOS, and in its own right as well, because it will be impossible to demonstrate value and utility. Its uniqueness lies in the fact that it represents and can network with a different community that includes taxonomists, biologists, and ecologists. These groups think and see the world differently than physical oceanographers, and history has proven that the two communities MUST be represented by separate programs so they can pursue their own brand of science, integrating the results.

Appendix V-a
NBII Web Site Review (August, 2007) - Action Items for "going live"

Meeting minutes: Rozita Abdul-Williams, August 13, 2007, USGS/NBII, Denver, CO.

New Changes/updates	Action	Status
A. Development work and staging		
1. Assure version control and approval cycle a. upload new files to development and testing area on server (for team review) b. update NBII active pages (visible through the portal) as approved	Team (Leigh & Sunny to coordinate edits, John to approve upload to active site)	continuous
B. NBII Page		
2. Draft new text box advertising and linking to OBIS-USA (for right column)	John	done
3. Convey changes to Portal pages to FISC a. Change "Marine Data" to "Marine Data (OBIS-USA) b. Insert new text box (OBIS-USA data) above "Dead Zones" c. Change in header: "Ocean Biodiversity Information System" to "Ocean Biogeographic Information System."	John -> Pam	done via FISC
B.OBIS-USA Search page		
4. Make the map smaller, change to dual color with regions labeled (reference for region selection box)	John & Rozita	done
5. Design new layout a. with 3 step search: enter search term, select region	Leigh	done

(from pull-down), execute search (text and map with text options). b. 3-column layout with "Search Help" on the right c. Increase page title size ("Search OBIS-USA") d. Relocate/resize logo		
6. Replace "Science wanted" to Executive Summary of OBIS-USA (summary the about page, including total # of data records & # of providers).	John	done
7. Code and test search functions	Sunny	done
8. Write search help text	John et al.	done
9. Change "Science Wanted" to summary of OBIS-USA (excerpt from About page) a. show #data records and # of providers	John	done
C. OBIS-USA Results Page		
10. Create new layout a. reduce top space as much as possible b. reduce size of logo c. remove unnecessary dividing bars d. center map (left justify legend) e. combine all 5 action buttons in same location (between map and table*) f. change "Select Region" to "Zoom to region" g. delete "Full extent" button, place "Zoom to region" drop box in this location h. Add label to top of legend box "Map Layers (activate/deactivate)" i. Remove wording in red that says please scroll down, if necessary j. Add user instructions on web page for required browsers and settings	Leigh	done

11. Center the map & left justify the legend. 12. Increase size of legend box (wider, same height as map) 13. Display expanded legend as default	Sunny, Rozita	done
14. Re-code the action buttons to work in their new location, test functionality, recommend alternatives as needed. 15. Check/correct table alignment.	Leigh & Sunny	done
16. Map service functions a. Deactivate the map zoom with scroll wheel. b. Set default to plot all points in search	Sunny, Rozita, John (? Gail)	done Unable to set default plot due to code limitations
17. Correct display of error messages	Sunny	done
18. Add/re-label map layers a. Add shaded relief map layer (alt background) b. Add land-sea monochrome map (alt. background) c. Add vector overlay of region boundaries (four) d. Rename map layers appropriately	Rozita	done
D. Add new page for Tabular Data		
19. Design new page for tabular data display (no map)	Leigh	done
20. Code & test functions on new page.	Leigh & Sunny	done
E. Testing and compliance		
21. Check pages at 800X600 and higher resolutions	Team	OK
22. Test in alternate browsers a. IE 6, 7 b. Firefox (recent versions) c. Safari (OS 9 & 10)	Team	IE 6, 7, Firefox OK Safari not reliable
23. Correct code for browser compatibility	Sunny	OK
24. Write user instructions on web page for required browsers and settings	Leigh, John	canceled

25. Test compliance with NBII portal checklist	Team + Janice	OK
26. Schedule & conduct demonstration for NBII reviewers.	John / team	Done
27. Final signoff	Mark Fornwall	Approved, Oct. 10, 2007

Appendix V-b
Web Site Review: Changes to The OBIS Site to Prepare for Briefings

Conf. Call 12/14/07

OBIS Site Improvement suggestions	Comments	Assigned	Status	Resolution
OBIS-USA Home Page				
8. a. Give a brief description to users (top of page) as to the type of info they would be searching for, (i.e. a statement from www.iobis.org) b. Main search box move down–in blank space above box add sentences c. Pop-up with data sources	a. words at top / move box: John *Change subheading to "View/obtain species occurrence data tables and/or display with geographic backgrounds"* b. c. pop-up may require programming: link to providers table provides this information already, so not needed	a. John b. John c. Mark	a. draft b. draft c. decision	a/b Committee for comment - attached See Mike - team for decision
9. a. The introductory description is long, and may be more appropriate as part of the 'About the Project' section. A short phrase, noting that searches span the databases of several providers may suffice. b. The word "providers" can be clickable, linked to the "Current List of Providers" (http://obis-usa.colorado.edu/dev/8-Data%20providers.htm). c. Making this available within the structure of the portal, as a separate page with a consistent look, would reduce user distraction, by following an imbedded link.	a. shorten intro *Need to add numbers of records and sources to current text..* b. 'providers' click to provider list c. if providers.htm is incorporated into the portal, it should be dynamic update from server. Not recommended at this time	a. John b. John c. Mark	a. draft b. draft c. decision	a. Committee - attached b. Committee - attached See Mike - team for decision
3. a. To match user task flow, the list of	a. relocate sci. names link inside search	a. John	a. draft	a. Committee

scientific names should be closer to the search box, such as a look-up tool within 5 degrees of the visual angle, (proven to be the optimal location for a group of interface items) b. Search tip—move closer to scientific name link as we fix default page c. Search by common name, region, ocean… d. move list of sc. Names	box b. unclear, relocate Search Tips – to where? c. search by common name not feasible at this time: not recommended John at one time you had the list from Phoebe? Or am I dreaming? Is there a list? d. unclear, move list where? Redundant with 3b.	 b. Mark c. Mark d. Mark	 b. clarify c. decision d. clarify	- attached b. Committee - attached c.
25. a. Under search tips, clicking on to scientific names opens a gigantic spreadsheet of scientific names with no navigation. The user must scroll both vertically and horizontally. It would be nice to have a linked index so that the user can navigate to the "a" names, "b" names etc. A sorting table (e.g. NBII-BRD) might be useful – html table that has A-D; E-G, etc.	a. create alphabet and html links Note: This is a temporary fix because we need to dynamically query the names, and abandon the static page. The navigation of the dynamic page can be implemented along with its code. What does this mean are you going to do it?	a. John/Rozita	a. ?	
27. a. Move the map image of the search regions so that it next to the step 2 dropdown box b. Use in first page description line; move to background.	a. move map image into Step 2 b. unclear, please clarify	a. John b. Mark	a. draft b. clarify	a. Committee - attached b. Committee - attached
31. There is no message given to the user when there are no results. For example, I accidentally typed in "dcent" instead of "cent" and the results page was blank, except for the download buttons at the top. Add some verbiage such as "No	a. Not recommended. This belongs on the results page, not the search page. Also, it should wait for DBMS structure – could be a chore getting that information from DiGIR.	a. John	a. What are impacts?	

Results Found"				
	Additional changes: a. removed OBIS-USA Regions label (not needed if it is close to the region search box). b. Bold Step 1, 2, 3; Bold "Search Tips"	a. John ?	a. draft ?	a. Committee - attached ?
Search Results Page				
32. a. Results table doesn't align properly in Fire Fox	a. programming	a. Leigh	a. scope task	
33. a. Results page provides no way for the user to execute a new search. There should be a close window button, a link or another search box from the results page.	a. New searches are conducted in the Portal; otherwise there is no need to use the portal. A link back to the portal search page will open a new window. The "search" page in the portal should remain active on the user's desktop, so it can easily be accessed. This is explained in the instructions. Recommendation: Add phrase at top of tabular and map results pages: "***New Search Page (you may also use the original search page)***"	a. John	a. draft	?
40. a. Provide a mechanism so that the user can perform another search. Provide a "close window" link, or a link back to the main search page.	a. same as above	N/A		
36. a. Provide the user with a link to the map view from the table results page. This would allow the user to see the mapping application for this data set without performing a new search.	a. Not recommended. I agree with the idea, but if we do this we should make the two pages fully sequential, where the map page is called only from the tabular page. This might help with the performance problem on the map page,	a. Mark	a. decision	a. Mike – Committee

b. Need to be able to switch between table and map- and add new search	where it has to load the table and GIS and we wait a long time for the GIS. We considered this option but did not experiment with it because of time. It will require re-hiring Sunny Lu to experiment, and there is a chance we will break other features of the site or take it down altogether, because we don't have a development site anymore (it is being used as the temporary operational site). b. Same with switching – the data has to be ported between applications, so this requires new code or significant code modification and testing.	b. Mark	b. decision	b. Mike - Committee
37. a. Polish the look of the mapping application by defining the mapping layers and toolbar with background colors and/or borders. It currently looks like more of a development product, than a production interface.	a. use existing ESRI capabilities to improve display	a. Rozita	a.	
35. a. It would improve readability if the results table used different background colors. Shade every other line	a. This is a change to the program, requires re-hiring Sunny Lu	a. Leigh	a.	a. Mike - Committee
38. a. Open the "release notes and user tips" link in a new window so that you don't need to reload the map and data when going back to the application. "Release notes and user tips" – what does this mean? use "help" and open a new	a. remove "Release notes and user tips" – this was for development. Adding a 'help' button would require a complete help system, not recommended.	a. John	a. draft	?
42. a. The results page borders do not show up in IE 7	a. not tested in IE 7 at time of development. Requires debugging and	a. Leigh	a.	a. Mike - Committee

	testing.			
44. a. Add more descriptive info for the technical resources listed? For example, are these resources that OBIS has created and uses extensively, or simply resources that OBIS recommends?	a. This comment does not apply to the results page, but to the technical resources page. All information on that page are links to IOBIS. Details are provided and maintained there. No action recommended.	a. Mark	a. decision	a. Committee
17. a. Mapping Application pages seem to contain a lot of development info - move to separate area or page?	a. Unclear what this refers to. There is no development information on the Mapping Application page. b. The "Release notes and User Tips" will be removed.	a. Mark b. John	a. clarify b. draft	a. Committee b. Committee
About OBIS and OBIS-USA				
19. a. Misspelling; "to serves" change "to serve" • Second Paragraph	a. correct misspelling	a. Rozita	a. ?	
17. a. Mapping Application pages seem to contain a lot of development info - move to separate area or page? b. NBII's comment?: This is standard for Arc GIS and what is expected by users – we really cannot make this type of modification without a totally building this page.	a. Same comment as above – unclear what this is referring to – need clarification b. If someone is commenting on the look/feel of the GIS layer legend, yes this is out of the box ESRI.	a. Mark b. Mark	a. clarify b. clarify	a. Committee b. ?
41. There is a JavaScript error when using IE 7 • Line 401 Charater 9 b. NBII comment?: Could not replicate	a. I did not get an error.	a. Mark	a. decision – Neither did I.	a. Committee

Appendix VI

OBIS Data Schema (an extension of Darwin Core II)

The OBIS SchemaVersion 1.1
08 July 2005
An extension of DarwinCore 2
Note: only those fields in red are required - all others are optional
Please see the OBIS Schema Field Implementation Notes for more details on usage

The OBIS Schema Version 1.1: Definition of the data standard OBIS uses

Name	Required	Type	Description
Date Last Modified	Optional for OBIS (Required for GBIF/Darwin Core servers)	DateTime	The date and time the record was last modified. Format: ISO 8601 compliant stamp in UTC(GMT) when the record was last modified. Example: "November 5, 1994, 8:15:30 am, US Eastern Standard Time" would be represented as "1994-11-05T13:15:30Z" (see W3C Note on Date and Time Formats - http://www.w3.org/TR/NOTE-datetime). While this field is required by the Darwin Core, OBIS can accommodate datasets without it.
Institution Code	Required	Text	A "standard" code identifier that identifies the institution to which the collection belongs, if there is one. Use the code that is "standard" in your discipline, if there is one (no global registry exists for assigning institutional codes). If not, use a short version of the name of the institution. (e.g. "NMNH" for Smithsonian National Museum of Natural History or "Duke" for Duke University) .
Collection Code	Required	Text	A unique alphanumeric value which identifies the collection within the institution (e.g. "FishBase") .
Catalog Number	Required	Text / Numeric	A unique alphanumeric value which identifies an individual record within the collection, i.e. the key. It is recommended that this value provides a key by which the actual specimen/observation can be identified. If the specimen/observation has several items such as various types of preparation, this value should identify the individual component of the specimen
Record URL	Optional	Text	Gives the web address of the page where more information on this particular record (not on the

			whole dataset) can be found.
Scientific Name	Required	Text	The full name of lowest level taxon the Cataloged Item can be identified as a member of; includes genus, specific epithet, and subspecific epithet (zool.) or infraspecific rank abbreviation, and infraspecific epithet (bot.) Use name of suprageneric taxon (e.g., family name) if Cataloged Item cannot be identified to genus, species, or infraspecific taxon.
Basis of record	Highly Recommended	Text	An abbreviation indicating whether the record represents an observation (O) (this can include a visual observation, a survey catch, a commercial landing record, etc), a collected living organism, such as a tree in a botanical garden (L), a specimen in a collection/museum (S), a collected germplasm/seed (G), a photo (P), or derived from literature, where original basis unknown (D).
Source	Optional	Text	OBIS does not encourage the use of this field - it is a legacy field. Indicates who gave the record to the data provider. Can indicate a literature citation, an electronic dataset, etc. Is used to provide credit.
Citation	Highly Recommended	Text	Indicates how this record should be attributed if used. (e.g. "Jones, T. 2005. Electronic atlas of eel distributions version 3. www.eels.com"). It can contain several layers of credit - e.g. of the original data provider and an intermediate data portal. If all records within a dataset should be credited the same way, the citation field in the dataset metadata can be used instead. It should be <4000 characters long.
Kingdom	Highly Recommended	Text	The kingdom to which the organism belongs
Phylum	Optional	Text	The phylum (or division) to which the organism belongs
Class	Optional	Text	The class name of the organism
Order	Optional	Text	The order name of the organism
Family	Optional	Text	The family name of the organism
Genus	Highly Recommended-when known	Text	The genus name of the organism. While this field is highly recommended when the identification to genus is known, it should not be filled in if the identification cannot be made down to genus with confidence.
Subgenus	Optional	Text	The subgenus name of the organism
Species	Highly	Text	The specific epithet of the organism

	Recommended-when known		
Subspecies	Optional	Text	The sub-specific epithet of the organism
Scientific Name Author	Optional	Text	The author of a scientific name. Author string as applied to the accepted name. Can be more than one author (concatenated string). Should be formatted according to the conventions of the applicable taxonomic discipline. Parentheses should be applied as appropriate for the relevant rules of Nomenclature (ICZN/ICBN) for the name. For example, if the name of an animal has undergone a genus revision, the authority and year should be placed in parentheses. Example: (Hastings, 1986)
Identified By	Optional	Text	The name(s) of the person(s) who applied the Scientific Name to the Cataloged Item.
Year Identified	Optional	Numeric	The year portion of the date when the Collection Item was identified; as four digits [-9999..9999], e.g., 1906, 2002.
Month Identified	Optional	Numeric	The month portion of the date when the Collection Item was identified; as two digits [01..12].
Day Identified	Optional	Numeric	The day portion of the date when the Collection Item was identified; as two digits [01..31].
Type Status	Optional	Text	Indicates the kind of nomenclatural type that a specimen represents, for example holotype, syntype, paratype, lectotype, paralectotype, neotype, schizotype, allotype, hapantotype. OBIS users should select from this list when applicable, but can enter other type categories as needed. In rare cases, a single specimen may be the type of more than one name.
Collector Number	Optional	Text	An identifying "number" (really a string) applied to specimens (in some disciplines) at the time of collection. Establishes a link between different parts/preparations of a single specimen and between field notes and the specimen.
Field Number	Optional	Text	A "number" (really a string) created at collection time to identify all material that resulted from a collecting event, e.g. station or sample numbers
Collector	Optional	Text	The name(s) of the collector(s), people or organisation(s) responsible for collecting the specimen, taking the observation, fishing the catch or doing whatever is the underlying basis of the record.

Year Collected	Highly Recommended	Numeric	The year (expressed as an integer) the sample/observation/record event occurred. The full year should be expressed (e.g. 1972 must be expressed as "1972" not "72"). Must always be a four digit integer. Where the event covers a range of values for year, indicates the mid-point of that range.
Start Year Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the start year of the collecting event. The full year should be expressed (e.g. 1972 must be expressed as "1972" not "72"). Must always be a four digit integer
End Year Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the end year of the collecting event. The full year should be expressed (e.g. 1972 must be expressed as "1972" not "72"). Must always be a four digit integer
Month Collected	Highly Recommended	Numeric	The month of year the sample/observation/record event occurred in the field. Where the event covers a range of values for month, indicates the mid-point of that range. Leave blank if even spans multiple years.
Start Month Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the start month of the collecting event. Possible values range from 01...12 inclusive
End Month Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the end month of the collecting event. Possible values range from 01...12 inclusive
Day Collected	Highly Recommended	Numeric	The day of the month the sample/observation/record event occurred in the field. Possible value ranges from 01..31 inclusive. Where the event covers a range of values for day, indicates the mid-point of that range. Leave blank if event spans multiple months.
Start Day Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the start day of the collecting event. Possible value ranges from 01..31 inclusive
End Day Collected	Optional	Numeric	For samples/observations/record events that were taken over time this gives the end day of the collecting event. Possible value ranges from 01..31 inclusive
Julian Day	Optional	Numeric	The ordinal day of the year for the sample/observation/record event; i.e., the number of

			days since January 1 of the same year. (January 1 is Julian Day 1.). Should be an integer from one to 365, i.e. of the form (([0-3][0-9][0-9)|([0-9][0-9)|([1-9])). Where the event covers a range of values for Julian day, indicates the mid-point of that range. Leave blank if event spans multiple years.
Start Julian Day	Optional	Numeric	For samples/observations/record events that were taken over time this gives the start ordinal day of the year for the collecting event; i.e., the number of days since January 1 of the same year. (January 1 is Julian Day 1.). Should be an integer from one to 365, i.e. of the form (([0-3][0-9][0-9)|([0-9][0-9)|([1-9])).
End Julian Day	Optional	Numeric	For samples/observations/record events that were taken over time this gives the end ordinal day of the year for the collecting event; i.e., the number of days since January 1 of the same year. (January 1 is Julian Day 1.). Should be an integer from one to 365, i.e. of the form (([0-3][0-9][0-9)|([0-9][0-9)|([1-9])).
Time of Day	Highly Recommended	Numeric	The time of day a specimen was collected expressed as decimal hours from midnight (e.g. 12.0 = mid day, 13.5 = 1:30pm)
Start Time of Day	Optional	Numeric	For samples/observations/record events that were taken over time this gives the start time of day of the collecting event expressed as decimal hours from midnight local time (e.g. 12.0 = mid day, 13.5 = 1:30pm)
End Time of Day	Optional	Numeric	For samples/observations/record events that were taken over time this gives the end time of day of the collecting event expressed as decimal hours from midnight local time (e.g. 12.0 = mid day, 13.5 = 1:30pm)
Time Zone	Highly Recommended	Text	Indicates the time zone for the Time of Day measurement, given as +hh:mm or -hh:mm from Coordinate Universal Time (also called Greenwich Mean Time). For example, a local time for Tokyo would have "+09:00" in the Time Zone field.
Continent Ocean	Optional	Text	The continent or ocean from which a specimen was collected or in which the sample/observation/record event occurred. OBIS recommends that ocean names follow the NASA Global Change Master Directory list of Bodies of Water http://gcmd.gsfc.nasa.gov/Data/portals/gcmd/location_search/top.html

Country	Optional	Text	The country or major political unit from which the specimen was collected or in which the sample/observation/record event occurred. ISO 3166-1 values should be used. Full country names are currently in use. A future recommendation is to use ISO3166-1 two letter codes or the full name when searching
State Province	Optional	Text	The state, province or region (i.e. next political region smaller than Country) from which the specimen was collected or in which the sample/observation/record event occurred. There is some suggestion to use the values described in ISO 3166-2, however these values are in a continual state of flux and it appears unlikely that an appropriate mechanism (by ISO) will be in place to manage these changes. Hence it is recommended that where possible, the full, unabbreviated name should be used for storing information. The server should optionally handle abbreviations as an access point. Note: this is a recurring theme (country and state) abbreviations. Check the existence of an attribute type to deal with abbreviations from the bib-1 profile
County	Optional	Text	The county (or shire, or next political region smaller than State / Province) from which the specimen was collected
Locality	Optional	Text	The locality description (place name plus optionally a displacement from the place name) from which the specimen was collected or in which the sample/observation/record event occurred. Where a displacement from a location is provided, it should be in un-projected units of measurement (e.g. "7 miles north of Hawaii"). It is strongly recommended that Locality be used, to allow cross-checking of the latitude and longitude fields
Longitude	Required	Numeric	The longitude of the location from which the specimen was collected or in which the sample/observation/record event occurred. This value should be expressed in decimal degrees (East & North = +; West & South = -). GPS-derived data should be referenced to the WGS/84 datum.
Start Longitude	Optional	Numeric	For samples/observations/record events better represented as line features rather than point features (e.g. extended trawls or transects) this indicates the starting longitude location from which the specimen

			was collected. Express in decimal degrees (East & North = +; West & South = -). GPS-derived data must use the WGS 84 geodetic reference system (http://www.wgs84.com/).
End Longitude	Optional	Numeric	For samples/observations/record events better represented as line features rather than point features (e.g. extended trawls or transects) this indicates the ending longitude location from which the specimen was collected. Express in decimal degrees (East & North = +; West & South = -). GPS-derived data must use the WGS 84 geodetic reference system (http://www.wgs84.com/).
Latitude	Required	Numeric	The latitude of the location from which the specimen was collected. This value should be expressed in decimal degrees (East & North = +; West & South = -). GPS-derived data must use the WGS 84 geodetic reference system (http://www.wgs84.com/).
Start Latitude	Optional	Numeric	For samples/observations/record events better represented as line features rather than point features (e.g. extended trawls or transects) this indicates the starting latitude location from which the specimen was collected or in which the sample/observation/record event occurred. This value should be expressed in decimal degrees (East & North = +; West & South = -). GPS-derived data must use the WGS 84 geodetic reference system (http://www.wgs84.com/).
End Latitude	Optional	Numeric	For samples/observations/record events better represented as line features rather than point features (e.g. extended trawls or transects) this indicates the ending latitude location from which the specimen was collected or in which the sample/observation/record event occurred. This value should be expressed in decimal degrees (East & North = +; West & South = -). GPS-derived data must use the WGS 84 geodetic reference system (http://www.wgs84.com/).
Coordinate Precision	Highly Recommended	Numeric	An estimate of how tightly the locality was specified in the Latitude and Longitude fields; expressed as a distance, in meters, that corresponds to a radius around the latitude-longitude coordinates. Use NULL where precision is unknown, cannot be estimated, or is not applicable.
Start/End Coordinate	Optional	Numeric	An estimate of how tightly the locality was specified in the Start/End Latitude and Longitude fields;

Precision			expressed as a distance, in meters, that corresponds to a radius around the latitude-longitude coordinates. Use NULL where precision is unknown, cannot be estimated, or is not applicable.
Bounding Box	Optional	BOUNDING BOX	This access point provides a mechanism for performing searches using a bounding box. A Bounding Box element is not typically present in the database, but rather is derived from the Latitude and Longitude columns by the data provider.
Minimum Elevation	Optional	Numeric	OBIS does not encourage the use of this field - it is a legacy field. The minimum distance in meters above (positive) or below sea level of the collection/record locality.
Maximum Elevation	Optional	Numeric	OBIS does not encourage the use of this field - it is a legacy field. The maximum distance in meters above (positive) or below sea level of the collection/record locality.
Minimum Depth	Highly Recommended	Numeric	The minimum distance in meters below the surface of the water at which the collection/record was made; all material collected was at least this deep. Positive below the surface, negative above (e.g. collecting above sea level in tidal areas).
Maximum Depth	Highly Recommended	Numeric	The maximum distance in meters below the surface of the water at which the collection/record was made; all material collected was at most this deep. Positive below the surface, negative above (e.g. collecting above sea level in tidal areas).
Depth Range	Optional-not preferred	Text	For data sets that have the depth range expressed in one field (e.g. "150-200 m") it can be entered here as free text. Separate, numeric Minimum and Maximum Depth fields are the preferred format; the Depth Range option is included for legacy data sets.
Temperature	Optional	Numeric	The temperature recorded with the collection/record event. Is assumed to be taken at the collection depth. Expressed in degrees Celsius.
Sex	Optional	Text	The sex of a specimen or collected/observed individual(s). The domain should be a controlled set of terms (codes) based on community consensus. Proposed values: M=Male; F=Female; H=Hermaphrodite; I=Indeterminate (examined but could not be determined; U=Unkown (not examined); T=Transitional (between sexes; useful for sequential hermaphrodites); B = Both Male and Female

Life Stage	Optional	Text	Indicates the life stage present. Will require developing a controlled vocabulary. Can include multiple stages for a lot with multiple individuals.
Preparation Type	Optional	Text	The type of preparation (skin. slide, etc). Probably best to add this as a record element rather than access point. Should be a list of preparations for a single collection record.
Individual Count	Optional	Numeric	The number of individuals present in the lot or container. Not an estimate of abundance or density at the collecting locality.
SampleSize	Highly Recommended	Text	Samplesize: the size of the sample from which the collection/observation was drawn. It can be a volume (e.g. for a phytoplankton sample), a linear distance (e.g. for a visual transect or net haul), a surface area (e.g. for a benthic core), etc. This field must also include the units, e.g. "200 m" for a transect, or "0.25 m^2" for a benthic grab (use ^ to denote a superscript). Note that when multiple collections/observations are reported from the same physical sample, a code identifying the sample can be placed in the Field_Number field to allow all collections/observations from a single sample to be connected.
Observed Individual Count	Highly Recommended	Numeric	The number of individuals (abundance) found in a collection/record event.
Observed Weight	Optional	Numeric	The total biomass found in a collection/record event. Expressed as kg.
Previous Catalog Number	Optional	Text	The previous (fully qualified) catalog number of the Cataloged Item (or collection/record) event if the item earlier identified by another Catalog Number, either in the current catalog or another Institution / catalog. A fully qualified Catalog Number is preceded by Institution Code and Collection Code, with a space separating the each subelement. Referencing a previous Catalog Number does not imply that a record for the referenced item is or is not present in the corresponding catalog, or even that the referenced catalog still exists. This access point is intended to provide a way to retrieve this record by previously used identifier, which may used in the literature. In future versions of this schema this attribute should be set-valued.

Relationship Type	Optional	Text	A named or coded valued that identifies the kind relationship between this Collection Item (or record event) and the referenced Collection Item. Named values include: "parasite of", "epiphyte on", "progeny of", etc. In future versions of this schema this attribute should be set-valued.
Related Catalog Item	Optional	Text	The fully qualified identifier of a related Catalog Item (a reference to another specimen); Institution Code, Collection Code, and Catalog Number of the related Cataloged Item, where a space separates the three subelements.
Notes	Optional	Text	Free text notes attached to the specimen record
OBIS EXPERIMENTAL FIELDS: The following are not part of the current OBIS Schema, but are under consideration for future versions. They represent format recommendations/good data practices.			
GML Feature	Optional	Text	Geographic Markup Language(GML) description of the feature for representing complex shapes such as lines and polygons, per Open GIS Consortium (OGC) standards - http://www.opengis.net/gml/01-029/GML2.html.

Implementation Notes of the OBIS Schema (with annotations for OBIS-USA underlined)

Date Last Modified

it is acceptable to enter a date only if the time is unknown
if the DiGIR data table contains elements with different modification times, enter the most recent time
if the modification date-time is unknown, enter the date-time of first "publication"

Collection Code and Institution Code

The Collection Code and the Institution Code can be the same in the case of Institutions that serve only one Collection
These can be full names instead of codes/abbreviations, if preferred.
The Collection Code (and/or the Field Number) can be hold concatenated Station and Expedition names/codes.

Catalog Number

The Catalog Number should be stable through time. So if a record is deleted, do not re-use the Catalog Number for a new record

Scientific Name

If the record is identified to genus and species level, this field should hold the genus and species epithets with a space between (for a total of 2 words). If subspecific epithet is known, this should be included in the string (for a total of 3 words). If the identification was only to a higher rank than genus, then name of the lowest known rank should be entered (1 word)
do not include the authority for the name here

Scientific Name Author

The year of original publication should be included if known, separated from the author name by a comma and space. If the name has undergone a genus revision, the authority and year should be in parentheses. Valid Examples:

Smith
Jones, 1973
(Hastings, 1986)

Use of the Start/End fields

There are several fields, such as latitude, longitude, day collected, month collected, etc., that have a start and end version. For example, the OBIS schema has "latitude", "Start_Latitude" and "End_latitude." How to fill in these fields is perhaps the most confusing part of the OBIS schema.

Why are all these fields there? They seem redundant?

Yes, they are redundant, but there is a reason for that. The Darwin Core represents all the location and time fields as single fields. But OBIS members thought it was important to be able to express a range of location or time. For example, a trawl might have been taken over a line transect that is better expressed as a start and end latitude and longitude than as a single point. Or an old specimen might only be labeled with the dates of the cruise, and not the day it was sampled, so that all we know is that the sample was taken sometime within a span of several months. For these reasons, OBIS added the start and end fields to the location and time information.

However, the OBIS schema needs to be compliant with the Darwin Core, so we still need to keep the original single-field options. So we end up with a field for, e.g, latitude, one for Start Latitude and one for End Latitude.

Implementing the Start/End Fields.

How you implement the Start/End fields will depend on the kind of data that you have. But regardless of your data structure, you should never have to type the same value into more than one field - you can make the database do this automatically.

Throughout the following directions, we will use latitude as an example. But the same rationale applies to all of the Start/End fields: Year Collected, Month Collected, Day Collected, Time of Day, and Longitude.

Case 1: all of your latitudes are point latitudes; none of them have separate start end latitudes.

In this case, you should have a "Latitude" field in your database into which you enter this information. When you install DiGIR and map your fields to the OBIS Schema, your "Latitude" field will get mapped to the OBIS Schema fields for "Latitude", "Start Latitude" and "End Latitude."

Case 2: You have samples that were taken over space and want to record a start and end latitude for all of them.

You should have "Start Latitude" and "End Latitude" fields in your database. These map to the same fields in the OBIS Schema. You then have a decision. "Latitude" is a required field in the OBIS schema, and a Darwin Core field, so you must map it to something in your database.

Solution A: The best option is to make an OBIS view of your database and create a "Latitude" field that is the average of your "Start Latitude" and your "End Latitude" fields (i.e. sum the fields and divide by 2)

Solution B: If the space covered by the sample is relatively small, you may feel that just using the "Start Latitude" field is good enough.

In either case, though, you must take care that the location precision is accurate (see below).

Case 3: Some of your samples were taken at a point and some were taken over a distance.

In this case, you can use the same method as Case 2 above. For those samples that were taken as a point, the simples approach is to have the "Start Latitude" and "End Latitude" fields be equal. You can fill them in individually by copying and pasting, or by a small script/routing. Alternatively, you can leave the "End Latitude" blank, but remember you'll have to have a way to get out the appropriate precision fields later (see below).

Filling in the Coordinate Precision fields - general comments

This field or fields (see below) indicates the precision with which the latitude/longitude location is given. This is generally a function of the method used (GPS, etc.). While this

is not a required field, it is a very important one and we highly recommend that you include it if at all possible. Note that the unit is meters, while the latitude and longitude fields are reported in decimal degrees. Note that when in doubt it is always better to err on the side of indicating a larger value in this field - it is better to indicate a little too much uncertainty than to report false precision. When in doubt, the number of significant digits in the latitude and longitude may roughly indicate the precision. The precision should never be smaller than the uncertainty created by the number of significant figures in the latitude and longitude (i.e. it doesn't make sense to report that a location is precise to 1 m if the latitude and longitude are only given to the tenth of a degree).

Coordinate Precision versus Start/End Coordinate Precision

The OBIS schema has two location precision fields: "Coordinate Precision" and "Start/End Coordinate Precision." Following the case examples from the "Use of the Start/End fields" notes, this is how they should be filled out.

Case 1: All of your latitudes are point latitudes; none of them have separate start end latitudes. You should have one precision field in your database and use this to estimate the precision with which each sample is measured - this will be dependent on the method used (GPS, etc.). When you map to the OBIS Schema, this field will be mapped to both the "Coordinate Precision" and the "Start/End Coordinate Precision" fields.

Case 2: You have samples that were taken over space and want to record a start and end latitude for all of them. You should have two precision fields in your database. "Start/End Coordinate Precision" should refer to the precision with which the start and end location points are known. "Coordinate Precision" should be a value that is large enough to span the Start and End points from the "Latitude" and "Longitude" fields. An example: say you are recording a 1 km-long trawl and used a GPS to get your start and end points so that you think your lat/lon measurement error is about 10 meters. In this case, your "Start/End Coordinate Precision" is 10. Your "Coordinate precision" will depend on whether you use solution A or solution B above. If you use solution A and report the midpoint of the line for "Latitude" and "Longitude," then the "Coordinate Precision" is 500m. If you use Solution B and report the "Start Latitude" and "Start Longitude" in the "Latitude" and "Longitude" fields, then the "Coordinate Precision" is 1000m.

Minimum and Maximum Elevation versus Depth

Minimum and maximum elevation are included because they are part of the Darwin Core, but for samples below sea level it is synonymous with Depth (except with the opposite sign). OBIS does not query on the elevation fields - it only uses the depth fields.

If all of your data are marine, then you can use just depth in your database. If you want to serve elevation then it can be automatically calculated as -depth. Or vice-versa. Just don't enter the numbers twice!

If you do hold non-marine data, such as data from lakes, then you may need to fill in both fields. In this case, the depth indicates the distance below the water level, while the elevation indicates the height above sea level. So a sample taken 10 meters below the surface of a lake on the top of a mountain that is 3000m high would have a depth = 10 and an elevation = 2990.

Elevation should not be used to indicate height above seafloor for marine sample.

Depth Range

The preferred method is to use the "Minimum Depth" and "Maximum Depth" fields, with both fields being equal when a collection was made at a single depth point, and not to use the Depth Range field. All new data entry projects should follow this format. However, we recognize that there are some legacy databases that have a single depth range field and where the data contributors can't take the time to individually split them up. If the depth is always recorded in meters using numerals (i.e. "10" not "ten"), then one of the OBIS contributors, SEAMAP, has developed a nice routine for pulling out the minimum and maximum automatically - you can contact Ben Best for details: bbest@duke.edu. But for those of you with fields that look like "from one to 10 fathoms" and don't have the time to convert them one by one, you can use the "Depth Range" field for free text information on depth. Note that there should be no cases in which all three are filled out for an individual record: if you have the Minimum and Maximum, then the range can be calculated and it should not be entered.

Individual Count versus Observed Individual Count

The Darwin Core developed from the museum community, so "Individual Count" refers to the number of specimens that were saved, not the number of individuals that were caught. OBIS has added the "Observed Individual Count" to indicate the total number per species that were caught. So if a fisheries survey caught 100 squid of a certain species and preserved 10 for a museum collection, then Individual count = 10 and Observed Individual Count = 100. Most databases will only have one or the other of these pieces of information saved.

Related Catalog Item

The Relationship Type and Related Catalog Item can be used to express tagging data following an individual through time (i.e. a later sighting is related to an earlier sighting). A special "relationship type" term should be defined for this.

Appendix VII: Proposed Data Policy

User Acknowledgment

By using data, software, or other information accessed through the OBIS-USA Portal, I agree that, in any publication or presentation of any sort based wholly or in part on material so accessed, I will:

1. Acknowledge the use of specific records, **citing the primary data authority** (i.e. precise source responsible for the scientific content and delivery of the data: the 'originator', 'author', 'original provider', or 'source' – all terms that may be used interchangeably) from contributing databases in a scientific citation format (Author (responsible for the data), date (of publication), database (or 'resource') name, source institution, primary publication outlet (publisher, but also add reference to OBIS as a publisher even if not considered primary), additional tech info as available, such as volumes, media descriptors, etc). Information about the primary author/source can be found in the original provider metadata. This information will be made available on-line via the OBIS regional and international search result service. Institution codes, primary contacts, and other information are available in the OBIS Data Schema for reference to more detailed information that may be made available on the web-site via additional services. In any case, correct citation of the original source is essential and it is the responsibility of the user.

2. **Acknowledge the use of the OBIS facility** in an appropriate form. For example a complete citation might be:

Example: "Richer de Forges, B. 2001. Electronic Database of ORSTOM Sampling on the Norfolk Ridge. In K. Stocks, 2003, SeamountsOnline Version 3.1. http://seamounts.sdsc.edu. Accessed via [*reference to any intermediate thematic node, such as SeaMap*] from [*name of regional Node, e.g., OBIS-USA at the National Biological Information Infrastructure Data Portal*] and the Ocean Biogeographic Information System, OBIS Distributed Data Search (OBIS-USA World Wide Web electronic publication http://obis.nbii.gov))

Example:

3. For information purposes, **provide to the OBIS Webmaster the full citation of any publication** (printed or electronic) produced from these data, citing OBIS or any constituent part.

Users beware

The information provided was prepared by contributors accessible through dynamic on-line publication. The data presented did not receive peer-review, except as performed by

the original provider. Intermediate thematic and regional nodes provide additional quality control procedures to ensure the original providers approve public distribution and that they remain responsible for content as the primary data authority.\

OBIS and the OBIS regional Nodes do not own or try to control or limit the use of any data or products accessible through its website(s). Accordingly, they do not take responsibility for the quality of such data or products, or the use that people may make of them, although they may act as agents for communication of quality and usability issues.

Users must recognize that the analysis and interpretation of data requires background knowledge and expertise about marine biodiversity (including ecosystems and taxonomy). Users should be aware of possible errors in the use of species names, georeferencing, data handing and mapping. They should crosscheck their results for possible errors, and qualify their interpretation of any results accordingly. Appropriate caution is thus necessary in the interpretation of results derived from OBIS-USA.

Users must be aware that OBIS is a federation and that **it is necessary to refer to the websites of individual contributors for appropriate contextual, explanatory and interpretive information, and for relevant metadata.** Users should also refer questions that they have concerning such issues to individual contributors though the procedures established in their respective websites. Questions about the portal function itself should be directed to the OBIS webmaster."

OBIS Data Use and Citation Agreement

By using data, software, or other information accessed through this OBIS Portal and its affiliated providers and portals, I agree to the terms stated in the OBIS data use policy (above). The correct references to data and OBIS nodes are provided via Web services on the regional and/or international portals.

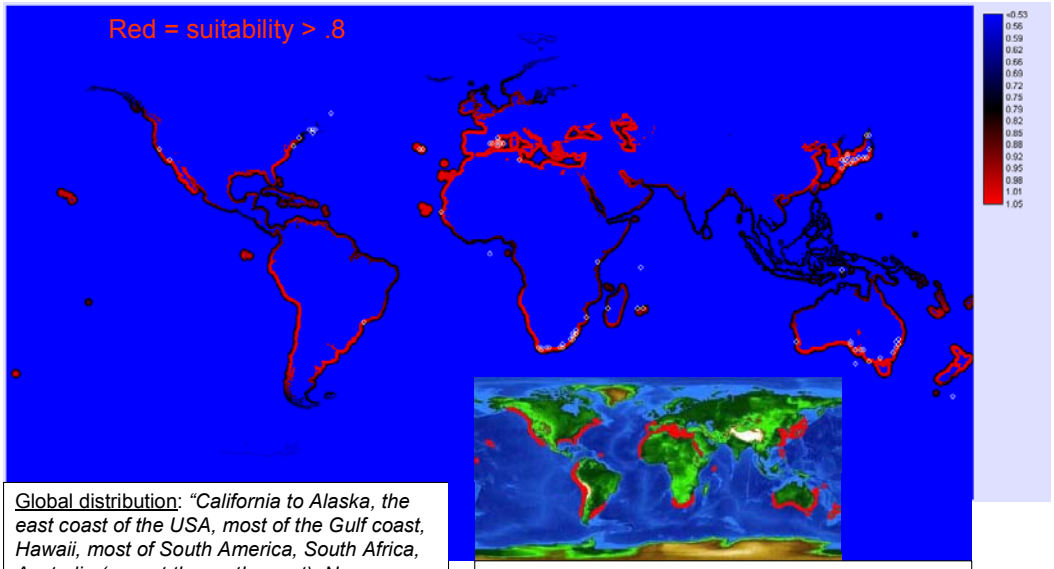
□ I acknowledge having read the above User Agreement and I agree to comply with all of its terms.

[clicking on the above will return the user to the data search results screen]

Appendix VIII

Species Distribution Model (Great White Shark)

Global Niche model Carcharodon carcharia (Great White Shark)

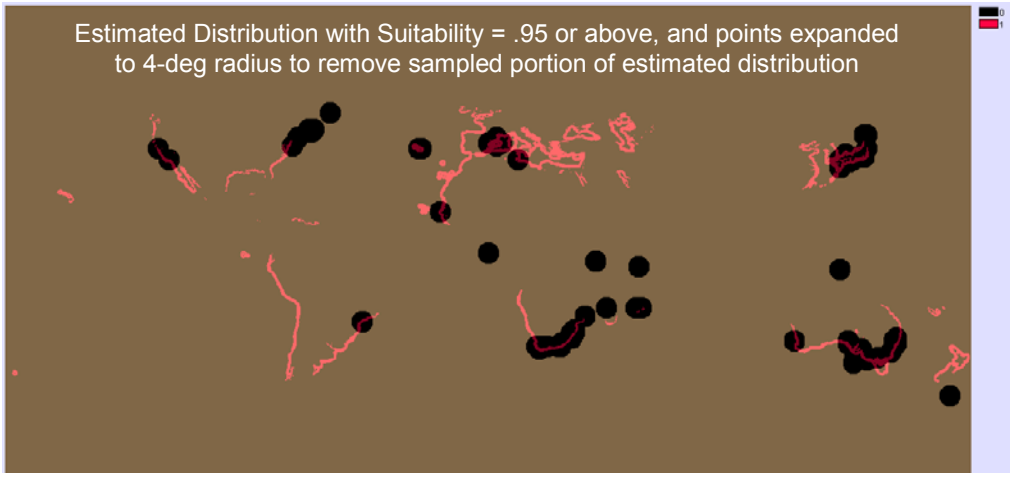


Global distribution: *"California to Alaska, the east coast of the USA, most of the Gulf coast, Hawaii, most of South America, South Africa, Australia (except the north coast), New Zealand, Mediterranean Sea, West Africa to Scandinavia, Japan, and the eastern coastline of China to Russia."*

(From Robin Dale from: www.geocities.com/CapeCanaveral/2774/grtwite.html ; www.ozemail.com.au/~bilsons/Sharks2.htm#greatwhite ; www.EnchantedLearning.com/subject/sharks.species/GreatWhite.shtml)

Global distribution according to the Florida Museum of Natural History
http://www.flmnh.ufl.edu/fish/Gallery/Descript/Whiteshark/whiteshark.html

Area With Data vs. Area of Model Prediction




```
Area on file: \\Krishna\f-drive\data\GreatWhite\global_model_rcl.rst
Category            Cells
      0          9186193
      1           145007

Area on file: \\Krishna\f-drive\data\GreatWhite\model-points.rst
Category            Cells
      0          9223576
      1           107624
```

107624 / 145007 = .74

Thus, 74% of the estimated range is unsampled!

Appendix IX: Software Implementation

Gail Lampenin, Integration Center for the Environment, UC-Davis

General Orientation

There are 2 directories – PbinMap2, which contains the user interface (ASPX pages) and the code for handling them, and DiGIRQuery, which contains the XML Web Service that builds a request and queries DiGIR.

The main page under PBinMap 2, essentially takes a common or scientific name then first calls the DiGIRQuery Web Service to build a request and query DiGIR to dynamically find out what resources (ie Bishop Museum data, Arctic Ocean Diversity, etc.) are available. It uses an XmlReader to extract what is needed from the SOAP response from DiGIR. As it walks through the XML, it binds each resource name (and human readable name) to a Repeater element which appears in the ASPX page. (Each Repeater element has a nested GridView element under it. As it binds the name, it calls the DiGIRQuery Web Service again, this time to build a request and query DiGIR to obtain records from that particular resource (like Bishop Museum) for the scientific or common name. As it receives the SOAP response back from DiGIR , it uses an XML reader then an XSL transformation to transform the XML into a format that can be used for display by GridView (this makes the table that you see on the page). The code also handles all the mapping applications, displaying records as point features, converting the GridView tables into Excel spreadsheet or text file, as requested.

The DigirQuery Web Service builds either a request for resources or a search request. For handling common names, it has to query the MySql database twice to build up a dynamic list of scientific names matching that common name. It then sends the SOAP request to DiGIR.

Please, please take a little time to read the comments (marked in Green) for the code in Attachment A, just for the code behind pages DigirQuery.Aspx.cs and QueryService.cs. This will really help you understand the project. You do not have to understand the code, simply read the green comments.

I also have the ASPX page the user sees and a sample XSL stylesheet which strips out qualified names and converts the Darwin elements to attributes as needed by GridView. You do not have to look at these.

Please also take a second to see what the request to Digir for a common name looks like (and my comment at the end).

Why Programming takes time

…
In addition to the current code that you see, there have been many stages of interim code developed which was not necessarily used in the end, which took time. In part, this was because of changes in what was desired for the project.

For example, originally, we were only returning a few fields for display, so I had to do a separate DiGIR request (and multiple xsl transformations) to produce the "view record" window with the information on all fields for a particular record. When we switched to downloading all the fields with the original search, the "view record" became extremely simple, eliminating a lot of code and 2 xsl stylesheets.

Also, in January, Derek had logically asked me to use AJAX (Asynchronous Javascript with XML) in order to update part of the page (the records returned from the database) without updating the whole page and the map, saving waiting time for the user. I did this successfully (as you remember, there was also no way to tell the page had been updated with the records, you just had to scroll down). However, I ended up taking all the AJAX code out of this project recently for two reasons. First, the map needed to have all prior mapped records (red stars from a previous search) removed when a new name search was performed. In addition, the map image was requested to be returned to the full extent of the selected geographic region, in case the user had zoomed way in or out. Therefore, the map image needed refreshing every time a search was performed, so a full page refresh was required. Secondly, the AJAX code conflicted with the JavaScript required to capture a user hitting the enter key for a name search instead of pressing the "Go" button. It took me a little time to figure this out. I was wondering why the JavaScript code would work in one area and not the other until I realized the conflict between the AJAX code and the JavaScript needed for capturing the return key…There are many more examples that I could provide.

Then there was the NBII portal system. (Derek, please show Mark the programming part of the Portal Developer's Workshop notebook so he can really appreciate this. I have read that thing from cover to cover). I originally tried just creating remote portlets, and their underlying objects. Next, I tried building the projects using the Plumtree API and their EDK development kit. This definitely took some time and a lot more code, but still the map image did not work. Finally, I returned to just using remote portlets. (Even obtaining the correct EDK for my .NET version took some time, because I could not download it directly from the Aqualogic website, as I did not have an account or permission. I had to get someone at NBII to download it, then make it available to me).

Furthermore, sometimes something that appears simple may take a lot of code to produce and one runs into unexpected problems. For one example, the GridView which displays the records requires the XML to be in a different format from what is produced by DiGIR. First, the transformation engine for .NET would not recognize the root element in the XML produced by DiGIR. I used an XML Reader to get a subset of the XML and rebuild a new XML document in memory to get around this. Secondly, the XML needs to

have fields as attributes, not elements, in order to display them in a GridView and I needed to strip out the qualified namespace. This required a XSL transformation. (Then, I use the XPath expression to let the GridView know where the fields are in the XML…). These are not DiGIR limitations, but limitations in .NET. It really should recognize qualified names…Discovering ways to work around these limitations, however, took time, trial and error, and creativity. See the commented code.

As another example, creating the Excel file should have been easy. It should be fairly straightforward to produce one from a GridView. However, the process chokes on certain fields, such as those containing checkboxes, hyperlinks, etc. I found my own way around this, by hiding specific problematic fields and one that I did not want to show. (You could also use this method to hide lat and long, if desired for sensitive species). However, I left in some code that I found on the web that automatically looks for problematic fields and deals with them. For my purposes, that code was insufficient (long explanation), but it could be useful later if someone adds one of these fields and forgets to add code to deal with it, or has a simpler GridView for another project.

In addition, I had to determine the structure and foreign keys of the Catalogue of Life by exploratory querying of the database. To check the common name search, I have to do some underlying querying to find common names that would return only one scientific name record, versus those that would return more than one, because I had to test both situations (since the underlying Digir queries are different – they would not be in any decent database).

Finally, the geographic query in DiGIR poses several problems. First, Digir does not support standard SQL syntax. Moreover, one cannot simply link multiple "and" conditions (such as latitude > -40 and latitude < 25 and longitude > -240 and longitude < -140). It does permit a bizarre nesting arrangement, though. Since digir is inadequately documented, I had to determine the correct syntax in large part through trial and error. Furthermore, the data in the databases are in standard longitude and latitude. But everything else is in the custom Pacific centered "Oceanic" projection. There is no way to tell digir to change the data, nor does it recognize mathematical operations. Therefore a geographic region that goes from longitude -140 to -240 must be searched in two areas – longitudes 140 to -180 and 120 to 180, as well as defining the latitude boundaries.

These are a few brief examples, but I hope they give you some sense of why things take time.

Programming Notes

Working with the NBII Portal System:

Short summary: 1. First Login and select the administration tab. Navigate to the Gail Lampinen folder in the Development Area.

2. I have already created the Remote Server Object for you. The Remote Server name is OBIS-USA Colorado server.
3. I have already created the Web Service for you. It is called Live Pacific Map Service. Note: if you create a new web service, increase the portlet and gateway timeout. Also, for the http configuration settings, make sure the "Transform JavaScript Files" and "Transform CSS files" are checked and uncheck "Use Hosted Display mode on Gatewayed Pages." Important: set the cached timed to 0 seconds. We do not want to used cached content ever.
4. I have already created the Remote Portlet for you. It is called Live Pacific Map Portlet. Look in the NBII Developers notebook for extensive documentation on how to do this. Note: if you create new objects, you need to specify who has access to them, or they cannot be used.

Updating the Project – Please get Visual Studio:

No development should ever, ever happen directly on the server. Once you have made the changes, package up the project. I have a setup projects already established for Lisa, so she does not need to know how to create one yet. It will do the work of bundling everything together into a msi installer that can install the updated project on the blade. First, rebuild the setup project. Then either zip or simply copy all the files under the debug folder for the setup project. (The files will probably be found under her "…My Documents\Visual Studio 2005\Projects\PBINMap2\PBinMapSetup1\Debug." The debug folder will have 2 files - a msi file and a steup.exe file. I always check the times and dates to make sure they are correct). On the blade, simply go to the contol panel and select Add/Remove programs. Both PBinMap2 and DigirQuery will be listed. Remove whatever project you updated, then click on the msi to start the installation of the updated project. You can accept the defaults on all options.

Attatchment A: Source code (example versions)

Note: the current working code is on the OBIS-USA server and backup devices. All code that was developed or altered was done so by Gail Lampenin, ICE/UC-Davis, and later by Sunny Ho, CIRES/UC-Boulder. Various layout and design features may have been modified by Leigh Ann McConnaughy and (to a much lesser extent) by John J. Kineman.

DigirQuery.Aspx.cs:

```
using System;
using System.Data;
using System.Configuration;
using System.Web;
using System.Web.Security;
using System.Web.UI;
using System.Web.UI.WebControls;
using System.Web.UI.WebControls.WebParts;
using System.Web.UI.HtmlControls;
using System.Net;
using System.IO;
```

```
using System.Xml;
using System.Xml.Xsl;
using System.Text;
using System.Xml.XPath;
using System.Collections;
using System.Collections.Generic;
using ESRI.ArcGIS.ADF.Resources;
using ESRI.ArcGIS.ADF.Web.Display.Graphics;
using ESRI.ArcGIS.ADF.Web.DataSources.Graphics;
using ESRI.ArcGIS.ADF.Web.Geometry;
using ESRI.ArcGIS.ADF.Web.DataSources;
using MySql.Data.MySqlClient;

public partial class _Default : System.Web.UI.Page
{
    // Global variables
    int intDB = 0;
    string serverName = "obis-usa.colorado.edu:8080";
    string searchType = "ScientificName";
    string searchTerm = "";
    string regIndx = "";
    string strConn =
ConfigurationManager.AppSettings["mySqlConn"].ToString();
    string sysAdmin =
ConfigurationManager.AppSettings["SysAdmin"].ToString();
    double currMinx = -240, currMaxx = -40, currMiny = -35, currMaxy =
90;
    Boolean mapAll = false, textAll = false;

    protected void Page_Load(object sender, EventArgs e)
    {
        if (!IsPostBack)
        {
            // if the request is sent from SciNameSearch, it will have
these 3 terms
            // however, this page can be loaded on its own with no
parameters, hence the check
            if (HttpContext.Current.Request.QueryString.Keys.Count > 0)
            {
                // get regIndx, the geographic region, and the search
terms
                regIndx =
HttpContext.Current.Request.QueryString["RegIndx"].Trim();
                txtComName.Text =
HttpContext.Current.Request.QueryString["ComName"].Trim();
                if (txtComName.Text.Trim() != "")
                {
                    searchType = "ComName";
                    searchTerm = txtComName.Text.Trim();
                }
                txtSciName.Text =
HttpContext.Current.Request.QueryString["SciName"].Trim();
                if (txtSciName.Text.Trim() != "")
                {
                    searchType = "ScientificName";
                    searchTerm = txtSciName.Text.Trim();
                }
```

```
            }
            // set the dropdownlist to a particular geographic region
or the default region, if none supplied
            switch (regIndx)
            {
                // Pacific Islands: Lon: -230 to -140, Lat: -25 to 42
                case "1":
                // Alaska Lon: -245 to -115  Lat: 42 to 80
                case "2":
                // Gulf, Atlantic, and Puerto Rico:  Lon: -110 to -55,
Lat: 10 to 50
                case "3":
                // Pacific Coast:  Lon: -140 to -110, Lat: 20 to 55
                case "4":
                // OBIS-USA
                case "P":
                // Whole Area
                case "W":
                    ddlGeoRegion.SelectedValue = regIndx;
                    break;
                default:
                    ddlGeoRegion.SelectedValue = "P";
                    break;
            }
            // For a new page loading, set the geographic parameters
needed for setting
            // the map extent and for limiting a search to a particular
geographic region
            getGeoRegion(regIndx);
        }
        else
        {
            // for postback, set the geographic parameters based on the
selected value in the dropdownlist
            getGeoRegion(ddlGeoRegion.SelectedValue);
        }
        // associate JavaScript function, named clickButton, with each
particular textbox.
        // The JavaScript function recognizes if the user pressed the
return/enter key and implementing the correct search
        // by forcing the proper buttonclick event.
        txtComName.Attributes.Add("onkeypress", "return
clickButton(event,'" + btnGetByCommon.ClientID + "')");
        txtSciName.Attributes.Add("onkeypress", "return
clickButton(event,'" + btnGetSpp.ClientID + "')");
    }
    protected void Page_PreRender(object sender, EventArgs ea)
    {
        // set the map extent to the desired geographic region the
first time the page loads
        if (!Page.IsPostBack)
        {
            ESRI.ArcGIS.ADF.Web.Geometry.Envelope initextent = new
ESRI.ArcGIS.ADF.Web.Geometry.Envelope(currMinx, currMiny, currMaxx,
currMaxy);
            Map1.Extent = initextent;
        }
```

```
    }

    protected void Page_LoadComplete(object sender, EventArgs e)
    {
        // first time page loads, call proper search
        if (!IsPostBack)
        {
            if (txtSciName.Text.Trim() != "")
            {
                btnGetSpp_Click(sender, e);
            }
            if (txtComName.Text.Trim() != "")
            {
                btnGetByCommon_Click(sender, e);
            }
        }
    }
    // Scientific name search
    protected void btnGetSpp_Click(object sender, EventArgs e)
    {
        // hide Common Name Search limitations warning and clear
general error message
        lblComLimits.Visible = false;
        lblCommonError.Visible = false;
        lblError.Visible = false;
        lblError.Text = "";
        // Call procedure to reset map extent and clear any mapped
records
        mapClear();
        // Obvious
        if (txtSciName.Text.Trim() != "")
        {
            searchType = "ScientificName";
            searchTerm = txtSciName.Text.Trim();
            getRecords(searchType,searchTerm);
        }
        else
        {
            lblError.Text = "Please enter a species name into the text
box before pressing the 'Go' button";
            lblError.Visible = true;
        }
    }
    // Now for Common Name Search
    protected void btnGetByCommon_Click(object sender, EventArgs e)
    {
        int sppCount = 0;
        lblError.Visible = false;
        // only show warning if able to search
        lblCommonError.Visible = false;
        mapClear();
        // make sure the user entered something
        if (txtComName.Text.Trim() != "")
        {
            // ok able to search, so show warning about limitations of
common name searches
            lblComLimits.Visible = true;
```

```
            searchType = "ComName";
            searchTerm = txtComName.Text.Trim();
            // check if any scientific names are returned for this
particular common name search
            using (MySqlConnection mySqlConn = new
MySqlConnection(strConn))
                try
                {
                    string myScalarQuery = "Select
count(distinct(name_code)) from common_names where common_name like '"
+ searchTerm + "%'";
                    MySqlCommand myCommand = new
MySqlCommand(myScalarQuery, mySqlConn);
                    mySqlConn.Open();
                    // Make sure there are some records for this common
name search
                    sppCount =
Convert.ToInt16(myCommand.ExecuteScalar());
                    mySqlConn.Close();

                    // if scientific names are returned, go ahead with
search
                    if (sppCount > 0)
                    {
                        getRecords(searchType, searchTerm);
                    }
                    // no scientific names found, tell user to try
another name search
                    else
                    {
                        lblCommonError.Text = "No records for Species
were found with this Common Name Search. Please try another search
term";
                        lblCommonError.Visible = true;
                    }
                }
                // Catch MySql errors
                catch (MySqlException sqlError)
                {
                    lblError.Text = "Please email the System
Administrator " + sysAdmin + " with this SQL error message: " +
sqlError.Message;
                    lblError.Visible = Boolean.Parse("True");
                }
                // Grabbag of any other errors
                catch (Exception ex)
                {
                    lblError.Text = "Please email the System
Administrator " + sysAdmin + " with this general error message: " +
ex.Message;
                    lblError.Visible = Boolean.Parse("True");
                }
                finally
                {
                    mySqlConn.Close();
                }
        }
```

```
        else
        {
            lblError.Text = "Please enter a common name into the text
box before pressing the 'Go' button";
            lblError.Visible = true;
        }

    }
    // Called by both Scientific and Common Name Search
    protected void getRecords(string searchType, string searchTerm)
    {
        lblScroll.Visible = true;
        intDB = 0;
        //lblTimer.Text = "Please wait. Retrieving records from ";
        //PanelTimer.Visible = true;
        StringBuilder sbError = new StringBuilder();
        // Create new QueryService (calls web service)to dynamically
discover resources
        // This call will return the names of the database resources
available to query
        DigirQuery.QueryService qs = new DigirQuery.QueryService();
        // The following was used for debugging purposes only. If you
have trouble with accessing resources
        // StringBuilder sbResults = new StringBuilder();

        // Get Resources - Create an XmlReader to read XML returned
from DiGIR
        // XmlReader Settings
        XmlReaderSettings settings = new XmlReaderSettings();
        settings.ConformanceLevel = ConformanceLevel.Auto;
        settings.ValidationType = ValidationType.None;
        settings.CloseInput = true;
        settings.IgnoreComments = true;
        try
        {
            // Always use using syntax - see optimization techniques
            using (XmlReader resourceRdr = XmlReader.Create(new
StringReader(qs.GetResources(serverName)), settings))
            {
                // move to provider element, if it exists, or show
error
                if (resourceRdr.ReadToDescendant("provider"))
                {
                    string digirAccess;
                    // create and ArrayList to hold the resource names
                    // which will be used by the Repeater Conrol
                    ArrayList resList = new ArrayList();
                    if (resourceRdr.ReadToDescendant("accessPoint"))
                    {
                        resourceRdr.Read();
                        digirAccess = resourceRdr.Value;
                        while (resourceRdr.Read())
                        {
                            if (resourceRdr.IsStartElement("resource"))
                            {
                                if
(resourceRdr.ReadToDescendant("name"))
```

```
                                    {
                                        resourceRdr.Read();
                                        // get name to show to the world
                                        string resFullName =
(resourceRdr.Value);
                                        resourceRdr.MoveToElement();
                                        // get the resource name used by
DiGIR

resourceRdr.ReadToFollowing("code");
                                        resourceRdr.Read();
                                        string resCode =
(resourceRdr.Value);
                                        // custom ResData class defined at
end
                                        // concatenate names and add to
ArrayList
                                        resList.Add(new
ResData(string.Concat(resCode, ":", resFullName)));
                                    }
                                    else
                                    {
                                        sbError.Append("Error reading
resources. ");
                                    }

                                }
                            }
                            // Bind the resources to the Repeater Control.
As they are bound,
                            // populate the underlying GridView in
rprData1_onItemDataBound procedure below
                            rprData1.DataSource = resList;
                            rprData1.DataBind();
                        }
                        else
                        {
                            sbError.Append("Error reading access points.
");
                        }
                    }
                    else
                    {
                        sbError.Append("Fatal Error reading Metadata file.
Unable to reach data providers. Please contact the system administrator
" + sysAdmin + " and attempt your query later. Error:");
                    }
                }
            }
            // Catch any SOAP errors raised by the DiGIR web service and
display error message
            // This happens mainly when Apache is down. It does not catch
if the server is down or the IP address or Port are wrong, yet
            catch (XmlException xel)
            {
```

```
            lblError.Text = "Fatal Error - Unable to reach the main
data provider. Please contact the system administrator " + sysAdmin + "
and attempt your query later. " + xel.Message;
            lblError.Visible = Boolean.Parse("True");
            lblTimer.Text = "";
            PanelTimer.Visible = false;
        }
        // not using these right now. I never had time to finish this
        lblTimer.Text = "";
        PanelTimer.Visible = false;
    }
    // As each item in the Repeater control is bound (each resource
database name), nested GridView control
    public void rprData1_onItemDataBound(Object Sender,
RepeaterItemEventArgs e)
    {
        // XMLReader is needed to take a subset of the XML produced by
DiGIR, since the root element
        // produced by DiGIR is not recognized by the Microsoft XSL
Transformer engine
        StringBuilder sbError = new StringBuilder();
        DigirQuery.QueryService qs = new DigirQuery.QueryService();
        // XMLReaderSettings
        XmlReaderSettings settings = new XmlReaderSettings();
        settings.ConformanceLevel = ConformanceLevel.Auto;
        settings.ValidationType = ValidationType.None;
        settings.CloseInput = true;
        settings.IgnoreComments = true;
        // XmlWriterSettings
        XmlWriterSettings xws = new XmlWriterSettings();
        xws.ConformanceLevel = ConformanceLevel.Auto;
        // XSLTSettings
        XsltSettings xslSettings = new XsltSettings();
        XsltArgumentList argList = new XsltArgumentList();
        xslSettings.EnableDocumentFunction = true;
        // Path to xslt document
        string xslPath = Server.MapPath("~/Elem2Att.xsl");
        // Execute the following logic for Items and Alternating Items.
        if (e.Item.ItemType == ListItemType.Item || e.Item.ItemType ==
ListItemType.AlternatingItem)
        {
            // StringBuilder sb is used to accept output created by the
XmlWriter, which applies
            //    an XSL transformation to the subset of the XML
created by the XmlReader
            // StringBuilder sb2 is used to build the new XML with a
simple root element that GridView can handle
            StringBuilder sb = new StringBuilder();
            StringBuilder sb2 = new StringBuilder("<?xml
version='1.0'?><response>");
            // extract both the name of the resource needed for the
Digir Request - ResourceName
            // and the name used for display - longName
            string longName = ((ResData)(e.Item.DataItem)).Name;
            string ResourceName = longName.Substring(0,
longName.IndexOf(":"));
            sbError.Append(ResourceName);
```

```
            argList.AddParam("ResourceName", "", ResourceName);
            try
            {
                // Always use using whenever possible - see .NET
optimization techniques
                // Create QueryService and call DiGIRQuery web service
to get XML from DiGIR
                using (XmlReader reader = XmlReader.Create(new
StringReader(qs.GetResults(ResourceName, searchType, searchTerm,
100000, 0, currMinx, currMiny, currMaxx, currMaxy)), settings))
                {
                    // Get subset of XML returned below content
                    if (reader.ReadToDescendant("content"))
                    {
                        // Create XmlWriter with settings and
stringbuilder
                        using (XmlWriter xw = XmlWriter.Create(sb,
xws))
                        {
                            // XSL transformation
                            XslCompiledTransform transform = new
XslCompiledTransform();
                            transform.Load(xslPath, xslSettings, null);
                            transform.Transform(reader, argList, xw);
                            // remove garbage and close <response>
element
                            sb2.Append(sb.ToString());
                            sb2.Remove(40, 214);
                            sb2.Append("</response>");
                            // Add to XmlDSDiGIR, the XmlDataSource
underlying each GridView
                            XmlDataSource xds =
(XmlDataSource)e.Item.FindControl("XmlDSDiGIR");
                            // EnableCaching must be set to false.
Otherwise the results of a new search are not updated in the GridView
                            xds.EnableCaching = false;
                            xds.Data = sb2.ToString();
                            GridView gv =
(GridView)e.Item.FindControl("GridView1");
                            // Set GridView datasource
                            gv.DataSourceID = "XmlDSDiGIR";
                        }
                    }
                    else
                    {
                        // if no content element was returned, then
DiGIR must have returned an error for this resource
                        sbError.Append("Error reading resource: " +
((ResData)(e.Item.DataItem)).Name + ". Attempting to process another
resource. ");
                    }
                }
            }
                // Soap exceptions raised by DiGIRQuery web service
            catch (XmlException xel)
            {
```

```
                lblError.Text = "Error reading resource: " +
((ResData)(e.Item.DataItem)).Name + ". Attempting to process another
resource. Please contact the system administrator " + sysAdmin + " and
attempt your query later. " + xel.Message;
                lblError.Visible = Boolean.Parse("True");
            }
        }
    }
    protected void GetTextFile(Boolean textAll)
    {
        string contentType = "application/vnd.txt";
        StringBuilder str = new StringBuilder();
        // iterate through Repeater
        for (int r = 0; r < rprData1.Items.Count; r++)
        {
            int RowNum = 0;
            // get individual GridView for this particular resource
            GridView gv =
(GridView)rprData1.Items[r].FindControl("GridView1");
            for (int i = 0; i < gv.Rows.Count; i++)
            {
                // check if record is selected. If so, include in
output
                CheckBox cb =
(CheckBox)gv.Rows[i].FindControl("cbxGvw");
                if (textAll || cb.Checked)
                {
                    for (int j = 0; j < gv.Rows[i].Cells.Count - 1;
j++)
                    {
                        // Skip first 2 columns - the checkbox and
"view record" ones
                        if (j == 2)
                            // add value from 3rd column
                            str.Append(gv.Rows[i].Cells[j].Text);
                        else if (j > 2)
                            // add a comma separator with the value for
all other columns
                            str.Append("," +
(gv.Rows[i].Cells[j].Text));
                    }
                    // Add a line break
                    str.Append("\r\n");
                    RowNum++;
                }
            }
        }
        System.IO.StringWriter sw = new System.IO.StringWriter();
        System.Web.UI.HtmlTextWriter hw = new HtmlTextWriter(sw);
        // Output the file to the stream
        Response.Clear();
        Response.AppendHeader("content-disposition",
"attachment;filename=SelectedRecords.txt");
        Response.ContentType = contentType;
        Response.Write(str.ToString());
        Response.End();
    }
```

```
    // Now for all the !$@*% render problems to produce an Excel file
    protected void btnExcel_Click(object sender, EventArgs e)
    {
        string contentType = "application/ms-excel";
        StringWriter sw = new StringWriter();
        HtmlTextWriter hw = new HtmlTextWriter(sw);
        // Output the file to the stream
        Response.Clear();
        Response.ClearContent();
        Response.ClearHeaders();
        Response.AppendHeader("content-disposition",
"attachment;filename=AllRecords.xls");
        Response.ContentType = contentType;
        // iterate through repeater
        for (int r = 0; r < rprData1.Items.Count; r++)
        {
            // get GridView for this particular resource
            GridView gv =
(GridView)rprData1.Items[r].FindControl("GridView1");
            // check if it has any records
            if (gv.Rows.Count > 0)
            {
                foreach (DataControlField field in gv.Columns)
                {
                    // Don't render these fields - select box, click to
view record, or hidden DiGIR resource name
                    if (field.HeaderText.Contains("Click to View
Record") || field.HeaderText.Contains("Select") ||
field.HeaderText.Contains("ResourceName"))
                    {
                        field.Visible = false;
                    }
                }
                // Deal with other rendering problems, if any, - see
procedure below
                // At this point, you should not need this. However, I
am leaving this in, so you know how to deal with them
                PrepareGridViewForExport(gv);
                gv.RenderControl(hw);
            }
        }
        Response.Write(sw.ToString());
        Response.End();

    }
    // code modified from c-sharpcorner...not currently needed unless
you add one of these types
    private void PrepareGridViewForExport(Control gv)
    {
        LinkButton lb = new LinkButton();
        Literal l = new Literal();
        string name = String.Empty;
        // Note: the gv.Controls.AddAt(i,1) would add a text
placeholder to the output
        // I would not want any text or placeholder in the output, just
the values from the fields
        for (int i = 0; i < gv.Controls.Count; i++)
```

```
        {
            if (gv.Controls[i].GetType() == typeof(LinkButton))
            {
                //l.Text = (gv.Controls[i] as LinkButton).Text;
                gv.Controls.Remove(gv.Controls[i]);
                //gv.Controls.AddAt(i, l);
            }
            else if (gv.Controls[i].GetType() ==
typeof(HyperLinkField))
            {
                //l.Text = (gv.Controls[i] as HyperLinkField).Text;
                gv.Controls.Remove(gv.Controls[i]);
                //gv.Controls.AddAt(i, l);
            }
            else if (gv.Controls[i].GetType() == typeof(DropDownList))
            {
                l.Text = (gv.Controls[i] as
DropDownList).SelectedItem.Text;
                gv.Controls.Remove(gv.Controls[i]);
                gv.Controls.AddAt(i, l);
            }
            else if (gv.Controls[i].GetType() == typeof(CheckBox))
            {
                l.Text = (gv.Controls[i] as CheckBox).Checked ? "True"
: "False";
                gv.Controls.Remove(gv.Controls[i]);
                //gv.Controls.AddAt(i, l);
            }
            if (gv.Controls[i].HasControls())
            {
                PrepareGridViewForExport(gv.Controls[i]);
            }
        }
    }
    // Avoids another problem - leave this in
    public override void VerifyRenderingInServerForm(Control control)
    {
    }
    // Code to produce a text file of all the fields of the SELECTED
records
    // Calls GetTextFile procedure, which actually does the work
    protected void btnText_Click(object sender, EventArgs e)
    {
        textAll = false;
        GetTextFile(textAll);
    }
    // Same as above, but takes values from all rows, not just selected
ones
    protected void btnTextAll_Click(object sender, EventArgs e)
    {
        textAll = true;
        GetTextFile(textAll);
    }
    // Code to map records on the map image
    protected void MapPoints(Boolean mapAll)
    {
        lblScroll.Visible = false;
```

```
        int RowNum = 0;
        // set extent
        ESRI.ArcGIS.ADF.Web.Geometry.Envelope initextent = new
ESRI.ArcGIS.ADF.Web.Geometry.Envelope(currMinx, currMiny, currMaxx,
currMaxy);
        ESRI.ArcGIS.ADF.Web.UI.WebControls.Map map =
(ESRI.ArcGIS.ADF.Web.UI.WebControls.Map)form1.FindControl("Map1");
        map.Extent = initextent;

        // Get graphics resources
        map.InitializeFunctionalities();
        IEnumerable gfc = map.GetFunctionalities();
        ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource gResource
= null;
        foreach (IGISFunctionality gfunc in gfc)
        {
            if (gfunc.Resource is
ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource)
            {
                gResource =
(ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource)gfunc.Resource;
                break;
            }
        }

        if (gResource != null)
        {
            for (int i = 0; i < gResource.Graphics.Tables.Count; i++)
            {
                // ElementGraphicsLayer
                if (gResource.Graphics.Tables[i].GetType() ==
typeof(ElementGraphicsLayer))
                {
                    gResource.Graphics.Tables[i].Rows.Clear();
                }
            }
        }
        // Note: Elements are in an ElementsGraphicsLayer. You cannot
use an Acetate Layer
        // for this many elements - the performance would be way too
slow
        // ESRI recommends no more than one or two items in an Acetate
layer

        // Create ElementsGraphicsLayer
        ESRI.ArcGIS.ADF.Web.Display.Graphics.ElementGraphicsLayer
eglayer =
          new
ESRI.ArcGIS.ADF.Web.Display.Graphics.ElementGraphicsLayer();
        // Create the Marker Symbol - a red star size 10
        ESRI.ArcGIS.ADF.Web.Display.Symbol.SimpleMarkerSymbol sms =
          new ESRI.ArcGIS.ADF.Web.Display.Symbol.SimpleMarkerSymbol();
        sms.Color = System.Drawing.Color.Red;
        sms.Width = 10;
        sms.Type =
ESRI.ArcGIS.ADF.Web.Display.Symbol.MarkerSymbolType.Star;
```

```
        // iterate over Repeater
        for (int r = 0; r < rprData1.Items.Count; r++)
        {
            // Cast the Repeater item to a GridView
            GridView gv =
(GridView)rprData1.Items[r].FindControl("GridView1");
            // Get the underlying XmlDataSource
            XmlDataSource xds =
(XmlDataSource)rprData1.Items[r].FindControl("XmlDSDiGIR");
            for (int i = 0; i < gv.Rows.Count; i++)
            {
                CheckBox cb =
(CheckBox)gv.Rows[i].FindControl("cbxGvw");
                // if Map All records is chosen or if the checkbox is
checked
                if (mapAll || cb.Checked)
                {
                    DataKey data = gv.DataKeys[i];
                    if (data.Values["Latitude"].ToString() != "" &&
data.Values["Longitude"].ToString() != "")
                    {
                        double tmpLong =
System.Convert.ToDouble(data.Values["Longitude"].ToString());
                        // Convert Longitude to Oceanic coords, if
needed
                        if (tmpLong > 20) tmpLong -= 360;
                        // create point
                        Point point = new Point(tmpLong,
System.Convert.ToDouble(data.Values["Latitude"].ToString()));

ESRI.ArcGIS.ADF.Web.Display.Graphics.GraphicElement ge =
                        new
ESRI.ArcGIS.ADF.Web.Display.Graphics.GraphicElement(point, sms);
                        eglayer.Add(ge);
                        RowNum++;
                    }
                }
            }
        }
        gResource.Graphics.Tables.Add(eglayer);
        if (RowNum == 0)
        {
            lblError.Text = "Error - No Records Selected. Either select
some records using the check boxes or choose the 'Map All Points'
button";
            lblError.Visible = Boolean.Parse("True");
        }
        else
        {
            map.Refresh();
        }
    }
    protected void btnMap_Click(object sender, EventArgs e)
    {
        mapAll = false;
        MapPoints(mapAll);
    }
```

```
    protected void btnMapAll_Click(object sender, EventArgs e)
    {
        mapAll = true;
        MapPoints(mapAll);
    }
    // Clears mapped records and resets extent
    protected void mapClear()
    {
        // reset geographic boundaries
        ESRI.ArcGIS.ADF.Web.Geometry.Envelope initextent = new
ESRI.ArcGIS.ADF.Web.Geometry.Envelope(currMinx, currMiny, currMaxx,
currMaxy);
        ESRI.ArcGIS.ADF.Web.UI.WebControls.Map map =
(ESRI.ArcGIS.ADF.Web.UI.WebControls.Map)form1.FindControl("Map1");
        map.Extent = initextent;
        // get graphic resources
        map.InitializeFunctionalities();
        IEnumerable gfc = map.GetFunctionalities();
        ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource gResource
= null;
        foreach (IGISFunctionality gfunc in gfc)
        {
            if (gfunc.Resource is
ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource)
            {
                gResource =
(ESRI.ArcGIS.ADF.Web.DataSources.Graphics.MapResource)gfunc.Resource;
                break;
            }
        }
        if (gResource != null)
        {
            if (gResource.Graphics != null)
            {
                for (int i = 0; i < gResource.Graphics.Tables.Count;
i++)
                {
                    // ElementGraphicsLayer
                    if (gResource.Graphics.Tables[i].GetType() ==
typeof(ElementGraphicsLayer))
                    {
                        // clear the geometry for all points
                        gResource.Graphics.Tables[i].Rows.Clear();
                    }
                }
            }
        }
        map.Refresh();
    }

    protected void ddlGeoRegion_SelectedIndexChanged(object sender,
EventArgs e)
    {
        // call procedure to change the geographic extent in global
variables
        getGeoRegion(ddlGeoRegion.SelectedValue);
    }
```

```
    // set the geographic region parameters used in setting the map
extent and
    // limiting searches to a particular geographic region
    protected void getGeoRegion(string regIndx)
    {
        switch (regIndx)
        {
            // Pacific Islands: Lon: -230 to -140, Lat: -25 to 42
            case "1":
                currMinx = -230; currMaxx = -140; currMiny = -25;
currMaxy = 42;
                break;
            // Alaska Lon: -245 to -115  Lat: 42 to 80
            case "2":
                currMinx = -245; currMaxx = -115; currMiny = 42;
currMaxy = 80;
                break;
            // Gulf, Atlantic, and Puerto Rico:  Lon: -110 to -55, Lat:
10 to 50
            case "3":
                currMinx = -110; currMaxx = -55; currMiny = 10;
currMaxy = 50;
                break;
            // Pacific Coast:  Lon: -140 to -110, Lat: 20 to 55
            case "4":
                currMinx = -140; currMaxx = -110; currMiny = 20;
currMaxy = 55;
                break;
            // OBIS-USA
            case "P":
                currMinx = -240; currMaxx = -40; currMiny = -35;
currMaxy = 90;
                break;
            // Whole Area
            case "W":
                currMinx = -340; currMaxx = 20; currMiny = -35;
currMaxy = 90;
                break;
            default:

                currMinx = -240; currMaxx = -40; currMiny = -35;
currMaxy = 90;
                break;
        }
    }
    // ResData defined
    public class ResData
    {

        private string name;

        public ResData(string name)
        {
            this.name = name;
        }
```

```
        public string Name
        {
            get
            {
                return name;
            }
        }
    }
}
```

QueryService.cs:

```
using System;
using System.Web;
using System.Web.Services;
using System.Web.Services.Protocols;
using System.Net;
using System.IO;
using System.Text;
using System.Xml;
using System.Xml.Serialization;
using System.Web.Script.Services;
using System.Collections;
using MySql.Data.MySqlClient;
using System.Configuration;

[WebService(Namespace = "http://pbin.usgs.gov/")]
[WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
[ScriptService]
public class QueryService : System.Web.Services.WebService
{
    string strConn =
ConfigurationManager.AppSettings["mySqlConn"].ToString();
    string sysAdmin =
ConfigurationManager.AppSettings["SysAdmin"].ToString();
    //int RecordLimit = 5000;
    public QueryService()
    {

        //Uncomment the following line if using designed components
        //InitializeComponent();
    }
    [WebMethod]
    public string GetResources(string ServerName)
    {
        string Request;
        Request = BuildResourceReq();
        string Results = DigirQry(ServerName, Request);
        return Results;
    }
    private string BuildResourceReq()
    {
        string dt = DateTime.Now.ToString();
        // get requestig address
        string reqUrl =
HttpContext.Current.Request.UserHostAddress.ToString();
        // start building the xml request
        string theReq = "<request> " +
        " xmlns='http://digir.net/schema/protocol/2003/1.0' " +
        " xmlns:xsd='http://www.w3.org/2001/XMLSchema' " +
        "
xmlns:darwin='http://digir.net/schema/conceptual/darwin/2003/1.0' " +
        " xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance' " +
        " xsi:schemaLocation='http://digir.net/schema/protocol/2003/1.0
http://digir.sourceforge.net/schema/protocol/2003/1.0/digir.xsd
http://digir.net/schema/conceptual/darwin/2003/1.0
```

```
http://digir.sourceforge.net/schema/conceptual/darwin/2003/1.0/darwin2.
xsd'>" +
        "<header>";
        theReq += "<version>1.0</version>'";
        theReq += "<sendTime>" + dt + "</sendTime>";
        theReq += "<source>" + reqUrl + "</source>";
        theReq += "<destination
resource='test'>http://localhost/DiGIR/DiGIR.php</destination>";
        theReq += "<type>metadata</type>" +
        "</header>" +
         "</request>";
        return theReq;
    }
    [WebMethod]
    public string GetResults(string ResourceName, string SearchType,
string SearchTerm, int RecordLimit, int StartRecNum, double CurrMinx,
double CurrMiny, double CurrMaxx, double CurrMaxy)
    {
        string Request;
        Request = BuildReq("localhost", ResourceName, SearchType,
SearchTerm, RecordLimit, StartRecNum, CurrMinx, CurrMiny, CurrMaxx,
CurrMaxy);
        string Results = DigirQry("localhost", Request);
        return Results;
    }
    private string BuildReq(string ServerName, string ResourceName,
string SearchType, string SearchTerm, int RecordLimit, int StartRecNum,
double CurrMinx, double CurrMiny, double CurrMaxx, double CurrMaxy)
    {
        string dt = DateTime.Now.ToString();
        string reqUrl =
HttpContext.Current.Request.UserHostAddress.ToString();
        string theReq = "<request " +
        " xmlns='http://digir.net/schema/protocol/2003/1.0' " +
        " xmlns:xsd='http://www.w3.org/2001/XMLSchema' " +
        "
xmlns:darwin='http://digir.net/schema/conceptual/darwin/2003/1.0' " +
        " xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance' " +
        " xsi:schemaLocation='http://digir.net/schema/protocol/2003/1.0
http://digir.sourceforge.net/schema/protocol/2003/1.0/digir.xsd
http://digir.net/schema/conceptual/darwin/2003/1.0
http://digir.sourceforge.net/schema/conceptual/darwin/2003/1.0/darwin2.
xsd'>" +
        "<header><version>1.0</version>";
        theReq += "<sendTime>" + dt + "</sendTime><source>" + reqUrl +
"</source>";
        theReq += "<source>" + reqUrl + "</source>";
        theReq += "<destination resource='" + ResourceName +
"'>http://localhost/DiGIR/DiGIR.php</destination>";
        theReq += "<type>search</type></header><search><filter><and>";
        // Scientific name search with like matching
        if (SearchType == "ScientificName")
        {
            theReq += "<like>" + "<darwin:ScientificName>" + SearchTerm
+ "%</darwin:ScientificName>" + "</like>";
        }
        // Common Name Search - using specific scientific name matching
```

```
        else if (SearchType == "ComName")
        {
            // Get Scientific names for Common Name search from MySql
            using (MySqlConnection mySqlConn = new
MySqlConnection(strConn))
                try
                {

                    ArrayList sciList = new ArrayList();
                    ArrayList darwinList = new ArrayList();
                    // get the name_code from the common_names table
which links to name_code in
                    // scientific_names table.  Add to ArrayList
                    string mySqlQuery = "Select distinct(name_code)
from common_names where common_name like '" + SearchTerm + "%'";
                    MySqlCommand myCommand = new
MySqlCommand(mySqlQuery, mySqlConn);
                    mySqlConn.Open();
                    MySqlDataReader myReader;
                    myReader = myCommand.ExecuteReader();
                    // Always call Read before accessing data.
                    if (myReader.HasRows)
                    {
                        while (myReader.Read())
                        {
                            sciList.Add(myReader.GetString(0));
                        }
                        // always call Close when done reading.
                        myReader.Close();
                        // Now get the genus & species to make a
scientific name from the scientific_names
                        // table in MySql. Add each name_code to the
"in (" clause
                        string mySciNameQry = "Select genus, species
from scientific_names where name_code in (";
                        foreach (string codeName in sciList)
                        {
                            mySciNameQry +=  "'" + codeName + "',";
                        }
                        // remove final , and add trailing close
parenthesis
                        mySciNameQry =
mySciNameQry.Remove(mySciNameQry.Length-1) + ")";
                        MySqlDataReader mySciRdr;
                        // place sci names query to MySql
                        MySqlCommand mySciNameCmd = new
MySqlCommand(mySciNameQry, mySqlConn);
                        mySciRdr = mySciNameCmd.ExecuteReader();
                        // build list
                        while (mySciRdr.Read())
                        {
                            darwinList.Add("<darwin:ScientificName>" +
mySciRdr.GetString(0) + " " + mySciRdr.GetString(1) +
"</darwin:ScientificName>");

                        // The following is only used with <like>
searches - note the % term is added
```

```
                            // darwinList.Add("<darwin:ScientificName>"
+ mySciRdr.GetString(0) + " " + mySciRdr.GetString(1) +
"%</darwin:ScientificName>");
                        }
                        mySciRdr.Close();
                        mySqlConn.Close();
                        if (darwinList.Count > 1)
                        {
                            // works for exact name search only
                            theReq += "<in><list>";
                            foreach (string darwinName in darwinList)
                                theReq += darwinName;
                            theReq += "</list></in>";
                            // end works

                            // The following was one of the many
attempts to use like in a list of terms - none worked
                            //theReq += "<in><list>";
                            //foreach (string darwinName in darwinList)
                            //    theReq += "<like>" + darwinName +
"</like>";
                            //theReq += "</list></in>";

                            // The following was one of my attempts for
a nested like using <or> syntax.
                            // The request seems to be formed properly,
but does not work correctly
                            //for (int i = 0; i < darwinList.Count-2;
i++)
                            //{
                            //    theReq += "<or><like>" +
darwinList[i] + "</like>";
                            //}
                            //theReq += "<like>" +
darwinList[darwinList.Count-1] + "</like>";
                            //for (int i = 0; i < darwinList.Count - 2;
i++)
                            //{
                            //    theReq += "</or>";
                            //}
                        }
                        // one species is simple. However the
<in><list> syntax cannot handle
                        // a single item, it needs more than one. So,
we need to handle that case here
                        else
                        {
                            theReq += "<like>";
                            foreach (string darwinName in darwinList)
                                theReq += darwinName;
                            theReq += "</like>";
                        }
                    }
                    // should never get here
                    else
                    {
```

```
                        throw new
System.Web.Services.Protocols.SoapException("Error retrieving
Scientific Names from the Catalogue of Life database. Please contact
the system administrator " + sysAdmin + " with this message.", new
XmlQualifiedName("QueryService"));
                    }
                }
                catch (MySqlException sqlError)
                {
                    throw new
System.Web.Services.Protocols.SoapException("Error connecting to the
Catalogue of Life database. Please contact the system administrator " +
sysAdmin + " with this message: " + sqlError.ErrorCode + " - " +
sqlError.Message, new XmlQualifiedName("QueryService"));
                }
                catch (Exception ex)
                {
                    throw new
System.Web.Services.Protocols.SoapException("Error connecting to the
Catalogue of Life database. Please contact the system administrator " +
sysAdmin + " with this message:" + ex.Message, new
XmlQualifiedName("QueryService"));
                }
                finally
                {
                    mySqlConn.Close();
                }

        }
        theReq += "<and>";
        // restrict latitudes
        theReq += "<greaterThanOrEquals>";
        theReq += "<darwin:Latitude>" + CurrMiny +
"</darwin:Latitude>";
        theReq += "</greaterThanOrEquals>";
        theReq += "<and>";
        theReq += "<lessThanOrEquals>";
        theReq += "<darwin:Latitude>" + CurrMaxy +
"</darwin:Latitude>";
        theReq += "</lessThanOrEquals>";
        // adjust for longitude from the database which is not in
"Oceanic" projection but standard lat-long
        // for instance, something that goes from -140 to -240 must be
searched in two areas -140 to -180 and 120 to 180
        if (CurrMinx < -180 && CurrMaxx >= -180)
        {
            Double TempMinx = CurrMinx + 360.0;
        theReq += "<or>";
        theReq += "<lessThanOrEquals>";
        theReq += "<darwin:Longitude>" + CurrMaxx +
"</darwin:Longitude>";
        theReq += "</lessThanOrEquals>";
        theReq += "<greaterThanOrEquals>";
        theReq += "<darwin:Longitude>" + TempMinx.ToString() +
"</darwin:Longitude>";
        theReq += "</greaterThanOrEquals>";
        theReq += "</or>";
```

```
        }
        // Adjust longitudes if both are below -180
        else if (CurrMinx < -180 && CurrMaxx < -180)
        {
            Double TempMinx = CurrMinx + 360.0;
            Double TempMaxx = CurrMaxx + 360.0;
            theReq += "<and>";
            theReq += "<greaterThanOrEquals>";
            theReq += "<darwin:Longitude>" + TempMinx.ToString() +
"</darwin:Longitude>";
            theReq += "</greaterThanOrEquals>";
            theReq += "<lessThanOrEquals>";
            theReq += "<darwin:Longitude>" + TempMaxx.ToString() +
"</darwin:Longitude>";
            theReq += "</lessThanOrEquals>";
            theReq += "</and>";
        }
        else
        {
            theReq += "<and>";
            theReq += "<greaterThanOrEquals>";
            theReq += "<darwin:Longitude>" + CurrMinx +
"</darwin:Longitude>";
            theReq += "</greaterThanOrEquals>";
            theReq += "<lessThanOrEquals>";
            theReq += "<darwin:Longitude>" + CurrMaxx +
"</darwin:Longitude>";
            theReq += "</lessThanOrEquals>";
            theReq += "</and>";
        }
        theReq += "</and>";
        theReq += "</and>";
        theReq += "</and></filter>";
        theReq += "<records limit='" + RecordLimit + "' start='" +
StartRecNum + "'>";
        theReq += "<structure
schemaLocation='http://localhost/digir/ObisFull.xsd'/>" +
        "</records>" +
        "<count>true</count>" +
        "</search>" +
        "</request>";
        //debugging only - if you need to look at the request that gets
created this will throw back to calling page
        //throw new
System.Web.Services.Protocols.SoapException(theReq.ToString(), new
XmlQualifiedName("QueryService"));
        return theReq;
    }
    #region Post Query to Digir
    private string DigirQry(string ServerName, string theReq)
    {

        try
        {
            // posting a request to the DiGIR service
            string theUrl =
"http://localhost/digir/digir.php?doc="+theReq;
```

```
            WebRequest req = WebRequest.Create(theUrl);
            WebResponse result = req.GetResponse();
            Stream receiveStream = result.GetResponseStream();
            StreamReader rs = new StreamReader(receiveStream);
            string response = rs.ReadToEnd();
            rs.Close();
            return response;
        }
        catch (Exception e)
        {
            return e.ToString();
        }
    }
#endregion

}
```

DigirQuery.Aspx: (What the user sees, not the code)

```
<%@ Page Language="C#" AutoEventWireup="true"
CodeFile="DigirQuery.aspx.cs" Inherits="_Default" %>

<%@ Register Assembly="ESRI.ArcGIS.ADF.Web.UI.WebControls,
Version=9.2.0.1324, Culture=neutral, PublicKeyToken=8fc3cc631e44ad86"
    Namespace="ESRI.ArcGIS.ADF.Web.UI.WebControls" TagPrefix="esri" %>

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head runat="server">
            <title>OBIS-USA Mapping Application</title>
            <!--THIS IS HEADER INFORMATION. ONLY THE TITLE AND META
NAME "KEYWORDS" ARE SUBJECT TO CHANGE//-->
            <meta http-equiv="PICS-Label" content='(PICS-1.1
"http://www.weburbia.com/safe/ratings.htm" l r (s 0))'>
            <meta http-equiv="Content-Type" content="text/html;
charset=iso-8859-1">
            <meta content="30 days" name="revisit-after">
            <meta content="government science" name="classification">
            <meta content="The Pacific Basin Information Node of the
NBII is a collaborative effort among federal, state, local, and
international agencies to provide data and information concerning the
biota of the Pacific Basin."
                  name="description">
            <meta content="NBII, national, biological, information,
infrastructure, biology, biodiversity, reference, education, U.S.,
United States, U.S. Programs, U.S. Activities, States ,U.S.
government,USA, US, federal, national, local, government agencies,
international, global, world, multinational, Hawaii, Hawai'i, Hawaiian,
Hawaiian Islands, islands, Pacific, Pacific Ocean, Pacific Islands,
Pacific Basin, invasive, invasive species, invaders, alien, aliens,
alien species, alien invaders, non-native species, non-indigenous
species, exotic species, coral reefs, Hawaii Natural Heritage Program,
Bishop Museum, USGS, PIERC, ecosystem, ecosystems, ecology, ecological,
vegetation mapping, gap analysis program, GAP, geographical,
geospatial, conservation, biota, organisms, plants, botany, animals,
zoology, amphibians, reptiles, herpetology, mammals, mammalogy, birds,
ornithology, fish, icthyology, invertebrates, malacology "
                  name="keywords">
            <meta content="Pacific Basin Information Node"
name="organization">
            <meta content="general, safe for kids" name="rating">
            <meta content="PBIN" name="author">
            <meta content="en-us" name="language">
            <meta content="Geographic Perspectives" name="nbii theme">
            <meta content="Regional Information Nodes - Pacific Basin"
name="nbii category">
            <!--THIS IS THE LINK TO THE PBIN STYLESHEET. MAKE SURE YOU
HAVE AN UP-TO-DATE COPY OF THE STYLESHEET//--><LINK
href="nbiistyle.css" type="text/css" rel="StyleSheet">
        <style type="text/css">
        .hiddenCol { display: none; }
        </style>
```

```
     <script language="javascript" type="text/javascript">
   function clickButton(e, buttonid){
          if(e && e.which){ //if which property of event object is
supported (NN4)
              e = e
              charCode = e.which //character code is contained in NN4's
which property
           }
           else{
              e = window.event
              charCode = e.keyCode //character code is contained in
IE's keyCode property
           }
      var bt = document.getElementById(buttonid);
      if (bt){
          if (charCode == 13){
                bt.click();
                return false;
          }
      }
}

    </script>
</head>
<body>
    <form id="form1" runat="server">
        <asp:ScriptManager ID="ScriptManager1" runat="server"
AsyncPostBackTimeout="900" />
        <noscript>The mapping function and examining a full record of a
species in a popup window will <strong>not</strong> work without
enabling javascript in your browser</noscript>
        <div id="SciNameDiv">
                  <table cellSpacing="0" cellPadding="0" align="center"
width="100%">
                        <tr>
                              <td align="left" rowspan="4" colSpan="1"
width="130px"><img src="OBIS-USA_Logo.JPG"  alt="OBIS(Ocean
Biogeographic Information System)USA" height="114"/>
                        </td>

                  <td align="center" colspan="2">
                    <asp:label ID=lblQryResults runat="server" Font-
Names="Verdana" Font-Bold="true" Font-Size="150%">Results of OBIS-USA
Query</asp:label><br /><br /></td>
                </tr>
                        <tr>
                              <td align="center"  colSpan="2"
style="height: 29px"><asp:Label ID="lblError" runat="server" Font-
Names="Verdana" Font-Size="12pt" ForeColor="Red"></asp:Label>
                        <asp:Label ID="lblCommonError" runat="server"
Font-Names="Verdana" Font-Size="12pt" ForeColor="Red"></asp:Label>
                          <br />
                        <asp:Label ID="lblScroll" runat="server" Font-
Names="Verdana" Font-Size="12pt" ForeColor="Red"
                              Text="Please scroll down below the map to
see the search results." Visible="False"></asp:Label><br />
```

```
                        <asp:Label ID="lblComLimits" runat="server"
Font-Names="Verdana" Font-Size="12pt"
                            ForeColor="Red" Text='Note: If the Common
Name search does not return any records, use a more specific term (eg.
use "Angelfish" instead of "Angel")'
                            Visible="False"></asp:Label></td>
                        </tr>
                        <tr>
                              <td class="cell" align="center"
colspan="2" >
                                    <p class="center">Common Name:
                                          <asp:TextBox ID="txtComName"
runat="server"></asp:TextBox><asp:Button id="btnGetByCommon"
runat="server" Text="Go" OnClick="btnGetByCommon_Click"
Visible="True"></asp:Button>
                                          <br />Scientific Name:
                                          <asp:TextBox ID="txtSciName"
runat="server"></asp:TextBox><asp:Button id="btnGetSpp" runat="server"
Text="Go" OnClick="btnGetSpp_Click"  Visible="True"></asp:Button></p>
                              </td>
                        </tr>
                        <tr>
                              <td class="cell" align="center"
colspan="2" ><p class="center"></p>
                              </td>
                        </tr>
                        <tr>
                    </tr></table>
        </div>
       <asp:Panel ID="PanelTimer" runat="server" Height="50px"
Width="100%" Visible="false" >
           <asp:Label ID="lblTimer" runat="server" Text="" Font-
Names="Verdana" Font-Size="12pt" ForeColor="Red"></asp:Label>
                            </asp:Panel>
                                  <div id="divMapOptions"><table
cellSpacing="0" cellPadding="0"  align="center" width="100%">
                                                <tr>
                              <td colSpan="3" align="center"
class="head">
                                    <p class="center"><font
color="white">Mapping Options</font></p>
                              </td>
                        </tr>
                        <tr>
                              <td class="cell" colspan="3"
style="height: 24px"><p class="center">
                              <asp:Button ID="btnMapAll" runat="server"
Font-Bold="True" Font-Names="Verdana" OnClick="btnMapAll_Click"
Text="Map All Records" /> 
                        <asp:Button ID="btnMap" runat="server"
OnClick="btnMap_Click" Text="Map Selected Records" Font-Bold="True"
Font-Names="Verdana" /> <asp:Label
                            ID="lblGeoRegion" runat="server" Font-
Bold="True" Font-Names="Verdana"
                            Text="Select Region:"></asp:Label>
                        <asp:DropDownList ID="ddlGeoRegion"
runat="server" Font-Names="Verdana"
```

```
OnSelectedIndexChanged="ddlGeoRegion_SelectedIndexChanged">
                            <asp:ListItem Value="1">Pacific
Islands</asp:ListItem>
                            <asp:ListItem
Value="2">Alaska</asp:ListItem>
                            <asp:ListItem Value="3">Atlantic and Puerto
Rico</asp:ListItem>
                            <asp:ListItem Value="4">Pacific
Coast</asp:ListItem>
                            <asp:ListItem Value="P">OBIS-USA Pacific
Wide</asp:ListItem>
                        </asp:DropDownList>
                    </td>
                </tr>
                </table>
</div>

            <esri:Toolbar ID="Toolbar1" runat="server"
BuddyControlType="Map" Group="Toolbar1_Group"
            Height="50px" ToolbarItemDefaultStyle-BackColor="White"
ToolbarItemDefaultStyle-Font-Names="Verdana"
            ToolbarItemDefaultStyle-Font-Size="Smaller"
ToolbarItemDisabledStyle-BackColor="White"
            ToolbarItemDisabledStyle-Font-Names="Verdana"
ToolbarItemDisabledStyle-Font-Size="Smaller"
            ToolbarItemDisabledStyle-ForeColor="Gray"
ToolbarItemHoverStyle-BackColor="White"
            ToolbarItemHoverStyle-Font-Bold="True"
ToolbarItemHoverStyle-Font-Italic="True"
            ToolbarItemHoverStyle-Font-Names="Verdana"
ToolbarItemHoverStyle-Font-Size="Smaller"
            ToolbarItemSelectedStyle-BackColor="White"
ToolbarItemSelectedStyle-Font-Bold="True"
            ToolbarItemSelectedStyle-Font-Names="Verdana"
ToolbarItemSelectedStyle-Font-Size="Smaller"
            WebResourceLocation="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/" Width="400px">
            <ToolbarItems>
                <esri:Tool ClientAction="DragRectangle"
DefaultImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oomin.GIF"
                    HoverImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oominU.gif" JavaScriptFile=""
                    Name="MapZoomIn" SelectedImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oominD.gif"

ServerActionAssembly="ESRI.ArcGIS.ADF.Web.UI.WebControls"
ServerActionClass="ESRI.ArcGIS.ADF.Web.UI.WebControls.Tools.MapZoomIn"
                    Text="Zoom In" ToolTip="Zoom In" />
                <esri:Tool ClientAction="DragRectangle"
DefaultImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oomout.GIF"
```

```
                    HoverImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oomoutU.gif" JavaScriptFile=""
                    Name="MapZoomOut" SelectedImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/z
oomoutD.gif"

ServerActionAssembly="ESRI.ArcGIS.ADF.Web.UI.WebControls"
ServerActionClass="ESRI.ArcGIS.ADF.Web.UI.WebControls.Tools.MapZoomOut"
                    Text="Zoom Out" ToolTip="Zoom Out" />
                <esri:Tool ClientAction="DragImage"
DefaultImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/p
an.gif"
                    HoverImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/p
anU.gif" JavaScriptFile=""
                    Name="MapPan" SelectedImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/p
anD.gif"

ServerActionAssembly="ESRI.ArcGIS.ADF.Web.UI.WebControls"
ServerActionClass="ESRI.ArcGIS.ADF.Web.UI.WebControls.Tools.MapPan"
                    Text="Pan" ToolTip="Pan" />
                <esri:Command ClientAction=""
DefaultImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/f
ullext.gif"
                    HoverImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/f
ullextU.gif" JavaScriptFile=""
                    Name="MapFullExtent" SelectedImage="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/Toolbar/ArcGIS/Images/f
ullextD.gif"

ServerActionAssembly="ESRI.ArcGIS.ADF.Web.UI.WebControls"
ServerActionClass="ESRI.ArcGIS.ADF.Web.UI.WebControls.Tools.MapFullExte
nt"
                    Text="Full Extent" ToolTip="Full Extent" />
            </ToolbarItems>
            <BuddyControls>
                <esri:BuddyControl Name="Map1" />
            </BuddyControls>
        </esri:Toolbar>
       <div id="World_Map">
       <table><tr><td>
        <esri:Map ID="Map1" runat="server"
MapResourceManager="MapResourceManager1" Height="400px" Width="600px"
WebResourceLocation="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/">
        </esri:Map>
       </td>
       <td>
            <esri:Toc ID="Toc1" BuddyControl="Map1" runat="server" />
             
       </td></tr></table>
        </div>
```

```
        <esri:MapResourceManager ID="MapResourceManager1" runat="server"
WebResourceLocation="http://obis-
usa.colorado.edu:8080/aspnet_client/ESRI/WebADF/">
            <ResourceItems>
                <esri:MapResourceItem Definition="&lt;Definition
DataSourceDefinition=&quot;In Memory&quot;
DataSourceType=&quot;GraphicsLayer&quot; Identity=&quot;&quot;
ResourceDefinition=&quot;&quot; DataSourceShared=&quot;True&quot;
/&gt;"

DisplaySettings="visible=True:transparency=0:mime=True:imgFormat=PNG8:h
eight=100:width=100:dpi=96:color=:transbg=False:displayInToc=True"
                    Name="Overlay" />
                <esri:MapResourceItem Definition="&lt;Definition
DataSourceDefinition=&quot;http://obis-
usa.colorado.edu:8080/arcgis/services&quot; DataSourceType=&quot;ArcGIS
Server Internet&quot;
Identity=&quot;QVUMQNHBSBWBJNENYNWBENUMEMHMQNENYMPZUMAEPRPZPKSBOPPZ&quo
t; ResourceDefinition=&quot;(default)@OBIS-
USA_Detailed_Mapservice&quot; DataSourceShared=&quot;True&quot; /&gt;"

DisplaySettings="visible=True:transparency=0:mime=True:imgFormat=PNG8:h
eight=100:width=100:dpi=96:color=:transbg=False:displayInToc=True"
                    Name="OBIS-USA" />
            </ResourceItems>
        </esri:MapResourceManager>
        <div id="RecordsDiv">
                <table cellSpacing="0" cellPadding="0" align="center"
width="100%">
                    <!--<tr><td><asp:Label ID="Label3"
runat="server" Text="Label"></asp:Label></td></tr>-->
                    <tr>
                        <td class="head">
                            <p class="center"><font
color="white">Species Records</font></p>
                        </td>
                    </tr>
                    <tr>
                        <td class="cell" align="center">
                            <asp:Button ID="btnExcel"
runat="server" Text="Download Excel File of All Records"
OnClick="btnExcel_Click" Font-Bold="True" Font-Names="Verdana" /> 
                            <asp:Button ID="btnText" runat="server"
OnClick="btnText_Click" Text="Download Text File of Selected Records"
Font-Bold="True" Font-Names="Verdana" /> 
                            <asp:Button ID="btnTextAll" runat="server"
OnClick="btnTextAll_Click" Text="Download All Records to Text File"
Font-Bold="True" Font-Names="Verdana" /><br />
                        </td>
                </tr>
            </table>
        </div>
       <asp:UpdatePanel ID="updRecordActions1"
UpdateMode="Conditional" runat="server" Visible="true">
        <ContentTemplate>
                <table cellSpacing="0" cellPadding="0" align="center"
width="100%">
```

```
                    <!--<tr><td><asp:Label ID="Label4" runat="server"
Text="Label"/></td></tr>-->
<tr><td style="height: 16px">

          <asp:Repeater ID="rprData1" runat="server"
OnItemDataBound="rprData1_onItemDataBound">
         <ItemTemplate><b>Results from<%#
DataBinder.Eval(Container.DataItem,"Name").ToString().Substring(DataBin
der.Eval(Container.DataItem,"Name").ToString().IndexOf(":")) %></b>
<asp:GridView ID="GridView1"  Width="610px"  CellPadding="3"
EmptyDataText="No records found in this database within this geographic
region." EmptyDataRowStyle-Font-Names="Verdana" CellSpacing="1"
BorderWidth="2px" runat="server"  AllowPaging=false AllowSorting=false
DataKeyNames="CatalogNumber,Latitude,Longitude"
AutoGenerateColumns="False">
                 <Columns>
                     <asp:TemplateField HeaderText="Check to Select"
HeaderStyle-BackColor="#CEEFBD">
                         <ItemTemplate>
                             <asp:CheckBox runat="server" ID="cbxGvw" />
                         </ItemTemplate>
                    </asp:TemplateField>
                     <asp:HyperLinkField HeaderText="Click to View
Record"  Text="ViewFullRecord" HeaderStyle-BackColor="#CEEFBD"
DataNavigateUrlFields="InstitutionCode,CollectionCode,CatalogNumber,Sci
entificName,BasisOfRecord,DateLastModified,Kingdom,Phylum,Class,Order,F
amily,Genus,Species,Subspecies,ScientificNameAuthor,IdentifiedBy,YearId
entified,MonthIdentified,DayIdentified,TypeStatus,CollectorNumber,Field
Number,Collector,YearCollected,MonthCollected,DayCollected,JulianDay,Ti
meOfDay,ContinentOcean,Country,StateProvince,County,Locality,Longitude,
Latitude,CoordinatePrecision,MinimumElevation,MaximumElevation,MinimumD
epth,MaximumDepth,Sex,PreparationType,IndividualCount,PreviousCatalogNu
mber,RelationshipType,RelatedCatalogItem,Notes"
DataNavigateUrlFormatString="http://obis-
usa.colorado.edu:8080/PbinMap2/DisplaySingleRecord.aspx?Institution
Code={0}&Collection Code={1}&Catalog Number={2}&Scientific
Name={3}&Basis Of Record={4}&Date Last
Modified={5}&Kingdom={6}&Phylum={7}&Class={8}&Order={9}&Family={10}&Gen
us={11}&Species={12}&Subspecies={13}&Scientific Name
Author={14}&Identified By={15}&Year Identified={16}&Month
Identified={17}&Day Identified={18}&Type Status={19}&Collector
Number={20}&Field Number={21}&Collector={22}&Year Collected={23}&Month
Collected={24}&Day Collected={25}&Julian Day={26}&Time Of
Day={27}&Continent Ocean={28}&Country={29}&State
Province={30}&County={31}&Locality={32}&Longitude={33}&Latitude={34}&Co
ordinate Precision={35}&Minimum Elevation={36}&Maximum
Elevation={37}&Minimum Depth={38}&Maximum
Depth={39}&Sex={40}&Preparation Type={41}&Individual
Count={42}&Previous Catalog Number={43}&Relationship Type={44}&Related
Catalog Item={45}&Notes={46}" Target="_blank" />
      <asp:BoundField DataField="InstitutionCode" ItemStyle-
Wrap="false" HeaderText="Institution Code" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="CollectionCode" ItemStyle-
CssClass="hiddenCol" ItemStyle-Wrap="false" HeaderStyle-
CssClass="hiddenCol" HeaderText="Collection Code" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
```

```
      <asp:BoundField DataField="CatalogNumber" ItemStyle-
CssClass="hiddenCol" ItemStyle-Wrap="false" HeaderStyle-
CssClass="hiddenCol" HeaderText="CatalogNumber Name" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="ScientificName" ItemStyle-Wrap="false"
HeaderText="Scientific Name" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="YearIdentified" ItemStyle-
CssClass="hiddenCol" ItemStyle-Wrap="false" HeaderStyle-
CssClass="hiddenCol" HeaderText="Year Identified" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Latitude" ItemStyle-
CssClass="hiddenCol" ItemStyle-Wrap="false" HeaderStyle-
CssClass="hiddenCol" HeaderText="Latitude" HeaderStyle-
BackColor="#CEEFBD"  ReadOnly="True" />
      <asp:BoundField DataField="Longitude" ItemStyle-
CssClass="hiddenCol" ItemStyle-Wrap="false" HeaderStyle-
CssClass="hiddenCol" HeaderText="Longitude" HeaderStyle-
BackColor="#CEEFBD"  ReadOnly="True" />
      <asp:BoundField DataField="ResourceName" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="ResourceName" HeaderStyle-BackColor="#CEEFBD"
ReadOnly="True" />
      <asp:BoundField DataField="DateLastModified"
ConvertEmptyStringToNull="True" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Date Last Modified"
HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True"
/>
      <asp:BoundField DataField="BasisOfRecord"
ConvertEmptyStringToNull="True" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Basis Of Record"
HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True"
/>
      <asp:BoundField DataField="Kingdom" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Kingdom" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Phylum" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Phylum" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Class" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Class" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Order" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Order" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Family" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Family" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Genus" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Genus" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Species" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Species" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
```

```
      <asp:BoundField DataField="Subspecies" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Subspecies" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="ScientificNameAuthor" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Scientific Name Author" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="IdentifiedBy" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Identified By" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="YearIdentified" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Year
Identified" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="MonthIdentified" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Month
Identified" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="DayIdentified" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Day
Identified" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="TypeStatus" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Type
Status" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="CollectorNumber" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Collector Number" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="FieldNumber" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Field
Number" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="Collector" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Collector" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="YearCollected" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Year
Collected" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="MonthCollected" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Month
Collected" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="DayCollected" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Day
Collected" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="JulianDay" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Julian Day" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
```

```
      <asp:BoundField DataField="TimeOfDay" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="Time
Of Day" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="ContinentOcean" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Continent Ocean" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Country" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Country" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="StateProvince" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol" HeaderText="State
Province" HeaderStyle-BackColor="#CEEFBD" NullDisplayText="NULL"
ReadOnly="True" />
      <asp:BoundField DataField="County" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="County" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Locality" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Locality" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="CoordinatePrecision" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Coordinate Precision" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="MinimumElevation" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Minimum Elevation" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="MaximumElevation" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Maximum Elevation" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="MinimumDepth" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Minimum Depth" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="MaximumDepth" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Maximum Depth" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Sex" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Sex" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="PreparationType" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Preparation Type" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="IndividualCount" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Individual Count" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="PreviousCatalogNumber" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
```

```
HeaderText="Previous Catalog Number" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="RelationshipType" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Relationship Type" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="RelatedCatalogItem" ItemStyle-
CssClass="hiddenCol" HeaderStyle-CssClass="hiddenCol"
HeaderText="Related Catalog Item" HeaderStyle-BackColor="#CEEFBD"
NullDisplayText="NULL" ReadOnly="True" />
      <asp:BoundField DataField="Notes" ItemStyle-CssClass="hiddenCol"
HeaderStyle-CssClass="hiddenCol" HeaderText="Notes" HeaderStyle-
BackColor="#CEEFBD" NullDisplayText="NULL" ReadOnly="True" />
                </Columns>
        </asp:GridView>
            <asp:XmlDataSource ID="XmlDSDiGIR" runat="server"
XPath="//response/content/record" ></asp:XmlDataSource>
        </ItemTemplate>
        </asp:Repeater>
 </td>
                </tr>
                  </table>
         </ContentTemplate>
         <Triggers></Triggers>
       </asp:UpdatePanel>
                  <table>
                        <tr><td><asp:Label ID="lblResults"
runat="server"></asp:Label>
                                  </td>
                        </tr>
                  </table>
    </form>
</body>
</html>
```

Elem2Att.xsl:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet
xmlns:darwin="http://digir.net/schema/conceptual/darwin/2003/1.0"
xmlns:my="my:my" xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
version="1.0" exclude-result-prefixes="darwin my">
  <xsl:output method="xml" />
 <xsl:output omit-xml-declaration="yes" indent="yes"/>
 <xsl:strip-space elements="*"/>
 <xsl:param name="ResourceName" />

 <my:elNames>
   <name>ScientificName</name>
   <name>DateLastModified</name>
   <name>InstitutionCode</name>
   <name>CollectionCode</name>
   <name>CatalogNumber</name>
    <name>Longitude</name>
    <name>Latitude</name>
    <name>BasisOfRecord</name>
    <name>Kingdom</name>
    <name>Phylum</name>
    <name>Class</name>
    <name>Order</name>
    <name>Family</name>
    <name>Genus</name>
    <name>Species</name>
    <name>Subspecies</name>
    <name>ScientificNameAuthor</name>
    <name>IdentifiedBy</name>
    <name>YearIdentified</name>
    <name>MonthIdentified</name>
    <name>DayIdentified</name>
    <name>TypeStatus</name>
    <name>CollectorNumber</name>
    <name>FieldNumber</name>
    <name>Collector</name>
    <name>YearCollected</name>
    <name>MonthCollected</name>
    <name>DayCollected</name>
    <name>JulianDay</name>
    <name>TimeOfDay</name>
    <name>ContinentOcean</name>
    <name>Country</name>
    <name>StateProvince</name>
    <name>County</name>
    <name>Locality</name>
    <name>CoordinatePrecision</name>
    <name>BoundingBox</name>
    <name>MinimumElevation</name>
    <name>MaximumElevation</name>
    <name>MinimumDepth</name>
    <name>MaximumDepth</name>
    <name>Sex</name>
    <name>PreparationType</name>
```

```
    <name>IndividualCount</name>
    <name>PreviousCatalogNumber</name>
    <name>RelationshipType</name>
    <name>RelatedCatalogItem</name>
    <name>Notes</name>
 </my:elNames>

 <xsl:variable name="elNames" select="document('')/*/my:elNames/name"/>

  <xsl:template match="@* | node()">
    <xsl:copy>
      <xsl:apply-templates select="@* | node()"/>
    </xsl:copy>
  </xsl:template>

  <xsl:template match="*">
    <xsl:choose>
      <xsl:when test='not(local-name() = $elNames)'>
        <xsl:copy>
          <xsl:copy-of select="@*"/>
            <xsl:apply-templates select="*[local-name() = $elNames]"/>
            <xsl:apply-templates/>
        </xsl:copy>
      </xsl:when>
      <xsl:otherwise>
      <xsl:if test="position()=1">
          <xsl:attribute name="ResourceName">
          <xsl:value-of select="$ResourceName" />
          </xsl:attribute>
        </xsl:if>
        <xsl:attribute name="{local-name()}">
        <xsl:value-of select="."/>
        </xsl:attribute>
      </xsl:otherwise>
    </xsl:choose>
  </xsl:template>

</xsl:stylesheet>
```

Resources.xml

Example <provider.resource>.xml

Goa.xml

Example DiGIR Request

This is what a DiGIR request looks like for the common name search “Bicolor”:

```
<request
   xmlns='http://digir.net/schema/protocol/2003/1.0'
   xmlns:xsd='http://www.w3.org/2001/XMLSchema'
   xmlns:darwin='http://digir.net/schema/conceptual/darwin/2003/1.0'
   xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
   xsi:schemaLocation='http://digir.net/schema/protocol/2003/1.0
http://digir.sourceforge.net/schema/protocol/2003/1.0/digir.xsd
http://digir.net/schema/conceptual/darwin/2003/1.0
http://digir.sourceforge.net/schema/conceptual/darwin/2003/1.0/darwin2.xsd'>
        <header>
<version>1.0</version>
        <sendTime>7/7/2007 10:16:03 AM</sendTime>
<source>127.0.0.1</source>
        <source>127.0.0.1</source>
        <destination
resource='ArcticOceanDiversity'>http://localhost/DiGIR/DiGIR.php</destination>
        <type>search</type>
</header>
<search>
<filter>
<and>
<in>
<list>
<darwin:ScientificName>Chlopsis bicollaris</darwin:ScientificName>
<darwin:ScientificName>Opsanus dichrostomus</darwin:ScientificName>
<darwin:ScientificName>Hemicaranx bicolor</darwin:ScientificName>
<darwin:ScientificName>Stegastes partitus</darwin:ScientificName>
<darwin:ScientificName>Centropyge bicolor</darwin:ScientificName>
<darwin:ScientificName>Cetoscarus bicolor</darwin:ScientificName>
<darwin:ScientificName>Scarus rubroviolaceus</darwin:ScientificName>
<darwin:ScientificName>Parupeneus barberinoides</darwin:ScientificName>
<darwin:ScientificName>Ecsenius bicolor</darwin:ScientificName>
<darwin:ScientificName>Pseudanthias bicolor</darwin:ScientificName>
```

```
<darwin:ScientificName>Chromis dimidiata</darwin:ScientificName>
<darwin:ScientificName>Anguilla bicolor</darwin:ScientificName>
<darwin:ScientificName>Siganus uspi</darwin:ScientificName>
<darwin:ScientificName>Labroides bicolor</darwin:ScientificName>
<darwin:ScientificName>Dionda dichroma</darwin:ScientificName>
<darwin:ScientificName>Cyclothone pallida</darwin:ScientificName>
<darwin:ScientificName>Chlopsis bicolor</darwin:ScientificName>
<darwin:ScientificName>Poecilia catemaconis</darwin:ScientificName>
<darwin:ScientificName>Lipogramma klayi</darwin:ScientificName>
<darwin:ScientificName>Parma bicolor</darwin:ScientificName>
<darwin:ScientificName>Canthigaster smithae</darwin:ScientificName>
<darwin:ScientificName>Parapercis multifasciata</darwin:ScientificName>
<darwin:ScientificName>Prognathodes dichrous</darwin:ScientificName>
<darwin:ScientificName>Chromis margaritifer</darwin:ScientificName>
<darwin:ScientificName>Baptisia bicolor</darwin:ScientificName>
<darwin:ScientificName>Parkia bicolor</darwin:ScientificName>
<darwin:ScientificName>Lespedeza bicolor</darwin:ScientificName>
<darwin:ScientificName>Lupinus bicolor</darwin:ScientificName>
<darwin:ScientificName>Asplenium heterochroum</darwin:ScientificName>
<darwin:ScientificName>Nerisyrenia camporum</darwin:ScientificName>
<darwin:ScientificName>Platanthera bicolor</darwin:ScientificName>
<darwin:ScientificName>Hipposideros bicolor</darwin:ScientificName>
<darwin:ScientificName>Oecomys bicolor</darwin:ScientificName>
<darwin:ScientificName>Crocidura fuscomurina</darwin:ScientificName>
<darwin:ScientificName>Crocidura leucodon</darwin:ScientificName>
<darwin:ScientificName>Sonorella animasensis</darwin:ScientificName>
<darwin:ScientificName>Xerarionta tryoni</darwin:ScientificName>
<darwin:ScientificName>Isognomon bicolor</darwin:ScientificName>
</list>
</in>
        <and>
          <greaterThanOrEquals>
            <darwin:Latitude>-35</darwin:Latitude>
          </greaterThanOrEquals>
        <and>
          <lessThanOrEquals>
            <darwin:Latitude>90</darwin:Latitude>
          </lessThanOrEquals>
        <or>
          <lessThanOrEquals>
            <darwin:Longitude>-40</darwin:Longitude>
          </lessThanOrEquals>
          <greaterThanOrEquals>
            <darwin:Longitude>20</darwin:Longitude>
          </greaterThanOrEquals>
        </or>
```

```
        </and>
        </and>
        </and>
</filter>
        <records limit='100000' start='0'>
           <structure schemaLocation='http://localhost/digir/ObisFull.xsd'/>
        </records>
        <count>true</count>
        </search>
</request>
```