Making United States Integrated Ocean Observing System (U.S. IOOS) Inclusive of Marine Biological Resources

Hassan Moustahfid [(1)*],
NOAA/ U.S. Integrated Ocean Observing System
Silver Spring, MD.
hassan.moustahfid@noaa.gov
*corresponding author

Jim Potemra [(2)]
University of Hawaii
Pacific Islands Ocean Observing System,
Honolulu , Hawaii.

Philip Goldstein [(3)]
USGS/Ocean Biogeographic Information System-USA,
Boulder, CO.

Roy Mendelssohn [(4)]
NOAA/Southwest Fisheries Science Center, Pacific Grove, CA.

Annette DesRochers [(5)]
NOAA/Pacific Islands Fisheries Science Center,
Honolulu, Hawaii

***Abstract*— An important Data Management and Communication (DMAC) goal is to enable a multi-disciplinary view of the ocean environment by facilitating discovery and integration of data from various sources, projects and scientific domains. United States Integrated Ocean Observing System (U.S. IOOS) DMAC functional requirements are based upon guidelines for standardized data access services, data formats, metadata, controlled vocabularies, and other conventions. So far, the data integration effort has focused on geophysical U.S. IOOS core variables such as temperature, salinity, ocean currents, etc. The IOOS Biological Observations Project is addressing the DMAC requirements that pertain to biological observations standards and interoperability applicable to U.S. IOOS and to various observing systems. Biological observations are highly heterogeneous and the variety of formats, logical structures, and sampling methods create significant challenges. Here we describe an informatics framework for biological observing data (e.g. species presence/absence and abundance data) that will expand information content and reconcile standards for the representation and integration of these biological observations for users to maximize the value of these observing data. We further propose that the approach described can be applied to other datasets generated in scientific observing surveys and will provide a vehicle for wider dissemination of biological observing data. We propose to employ data definition conventions that are well understood in U.S. IOOS and to combine these with ratified terminologies, policies and guidelines.**

Keywords-component; Biological observations, Observing Systems, Bioinformatics.

I. INTRODUCTION

In recent years there has been a growing need for high quality, high resolution ocean observing datasets to feed models and produce products and services to society. Although, there is many agencies collecting biological observing data from marine ecosystems, their use and dissemination has often been limited to the organizations that collect the data. For researchers outside government agencies, obtaining data often follows an ad hoc process: data are made available under certain conditions and analyses are run using the version received.

Effort has been made to provide informatics services and expose biological data to a wide community (e.g.Ocean Biogeographic Information System (OBIS) [1], and Global Biodiversity Information Facility (GBIF) [2]). Despite accomplishments in technology and communication, their scientific acceptability is often limited by a lack of reproducibility in data analyses and because many applications require richer data than currently supported by community services. For example surveys data are highly heterogeneous and the variety of formats, logical structures, and sampling methods in fishery independent data create significant challenges. OBIS has focuses mostly on exposing species presence information. Species abundance or quantification variable is poorly represented in OBIS because of lack of standardization [3].

To address the data management challenges for biological observing data and to reconcile standards for measured quantities in biology and ecology, U.S. IOOS started in early 2010 a project called IOOS Biological Data Project (IOOS BDP). The project is focusing on a single case-study (i.e. customer-driven) and it's designed to promote biological data standards and interoperability to help the customer access observations from a wide variety of sources and formats. The basic idea is to connect end-users "customer" with rich biological observations databases in a more straightforward way. These databases typically come from disparate organizations, institutions and individuals for differing purposes and have locally-specific structure, contents, methods, and policies. In addition, these data and applications are diverse and could change over time (e.g., data from one source might contain different variables depending on when the data were collected). The objective of IOOS BDP is to help develop an efficient and effective information infrastructure for biological observations, using IOOS Data Management and communications (DMAC) standards [4] and adding components and links as necessary to serve end-users.

The customer selected for this project is the reef fish population assessments (i.e establish reef fish Annual Catch Limits) as defined by the Magnuson-Stevens Fishery Conservation and Management Reauthorization Act (MSRA) [5] and represented by scientists and managers at the NOAA/Pacific Islands Fisheries Science Centers.

The biological observations used are species presence/absence/abundance for reef fishes in the Hawaiian Archipelago and other locations in the Pacific region. The data providers include NOAA Pacific Island Fisheries Science Center/Coral Reef Ecosystem Division (PIFSC/CRED), National Park Services (NPS) Pacific Island Network (PACN) and the NOAA Papahānaumokuākea Marine National Monument (PMNM) (Figure 1). The data includes independent fishery surveys (diver-based reef fish observations) collected as part of each agency's coral reef monitoring programs.

While a single user-group is identified in the Pacific region for this project, the results from this project should be applicable more broadly to projects utilizing biological observations, and it should be portable to other geographies.

This paper describes the data/metadata standards that this project is adopting or developing to reconcile standards for the representation and integration of biological observations and for users to maximize the value of these observing data.

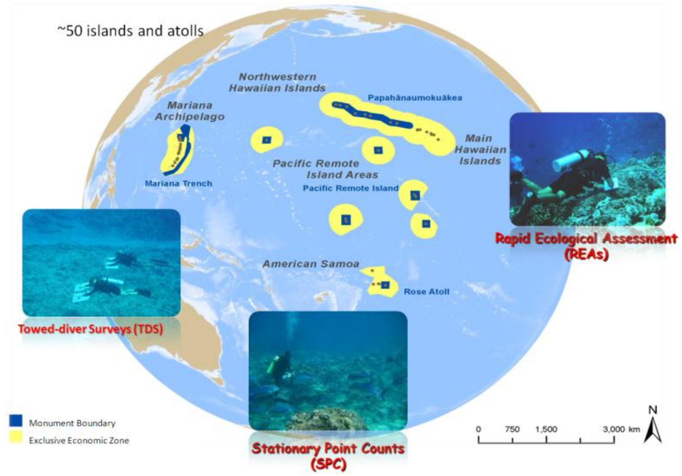

Figure 1. Locations of the reef-areas (i.e., islands and atolls) surveyed by PIFSC/CRED . Pictures represent three main methods used in RAMP program (source NOAA/PIFSC).

II. Project System Concept

The system concept developed for this project is shown in Figure 2. The concept gives different options and formats for accessing the data (web services, database and flat files). The data providers use the format that works best for him to share his data with larger community.

In the case of database connections, we propose to use pull or push approach. Pull– we connect to their database, download data and disconnect; Push – they connect to us and give their data. Web service would access the database, link into this database at this site. Service just searches metadata, not relying on the integration to do a search that spans the contents (part of a query, can be done on the server and some done on remote server).

Data from three different sources doesn't necessarily need to get integrated at the "data assembly" level, data is just assembled there.

For the terminology and configuration we propose to use Darwin Core (dwc) ratified standards [6] and Dublin core [7] that were built for biological services. We also adding some IOOS specific terms ("ioosbds") (details are provided below).

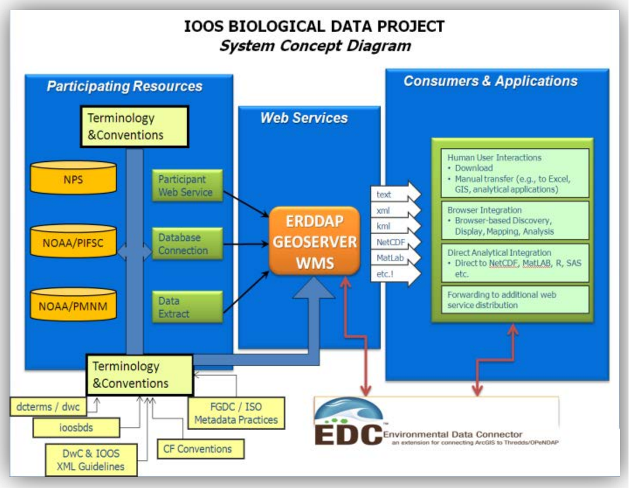

Figure 2. IOOS Biological data project system concept diagram. dcterms- Dublin core, dwc- Darwin Core, "ioosbds" - IOOS Biological Data Service, CF- Climate and Forecast Metadata Convention, FGDC- Federal Geographic Data Committee, ISO 19115-2- Geographic Information – Metadata.

III. Standards Development and Implementation

A. *Webservices*

This project is developing webservices to benefit the identified customers. They are designed to meet the general need of discovery, subsetting, visualization and download of biological data, and in a context that other data sources, such as physical data, can be readily combined with the biological data through the services. The Environmental Research Division's Data Access Program (ERDDAP) [8] is the main webservice for this project. However, other webservices such as Geoserver [9] and WMS [10] are also been considered for this project.

The new ERDDAP version is being improved. It can handle better in situ types of data. It includes for example new output file formats for tabular data, new etopo 1 bathymetry and other new features [5].

B. *Client Tool*

The IOOS BDP project is extending the client tool Environmental Data Connector (EDC) [11] developed by Applied Science Associates, inc (ASA) to access the data using customer application tools such as Geospatial tool ARCGIS.

The EDC is being extended to include the ability to access in situ datasets through ERDDAP. The next release of the EDC will be able to access data through the ERDDAP service. As the EDC starts with a Java front end, there will be a version that integrates these capabilities with ArcGIS, with Matlab, and in the near future with the R statistical language. Rather than just provide access to the biological data, the EDC will integrate into these applications a standard way to access a variety of remote data, both biological and physical, from a wide variety of remote sources.

C. *Schema terminology*

1) *Terminology*

The schema and terminology is based on ratified Darwin Core, Dublin Core and proposed IOOS vocabularies. For the purposes of this paper, the full schema terminology table has been omitted, here we present only some proposed dcterms , dwc and "ioosbds" terms (Table 1). The full schema is available by contacting the corresponding author.

a) Guide to the columns in the terminology table for "ioosbds" terms.

The columns in the table provide the following information:

Name

The term complying with controlled spelling and case. Names come from Dublin Core (dcterms) and Darwin Core (dwc) and from the proposed IOOS vocabulary, "ioosbds" for IOOS Biological Data Services. Some Dublin Core terms are included because they are part of the ratified Darwin Core, so they are inherited by using ratified Darwin Core. Their definitions are maintained by Dublin Core.

Class

This column identifies the Darwin Core class to which each term belongs. The initial benefit of organizing terms by class is for schema clarity. Organizing by class will enhance future benefits, particularly the extensibility of the Darwin Core / IOOS terms for data services and applications.

Classes represented in this design include Location, Event, Occurrence, Identification, Taxon, and “all".

NS – “namespace”

According to XML practices, a namespace identifies an authority for definition of terms or “names”. In this design, there are three such authorities, including “dcterms” for Dublin Core and “dwc” for ratified Darwin Core. The third namespace authority will be for IOOS-defined terms, and is indicated in the table as “ioosbds” for “IOOS Biological Data Service”. The proposed acronym for this namespace, “ioosbds”, can be revised as desired while services are in development. The namespace must be finalized before publicly releasing any services. A publicly-accessible web document will be established as a reference URL for the ioosbds namespace.

Note that definition and other information for all the Darwin Core terms - those in the namespaces "dcterms" and "dwc"- can be found in the Biodiversity Information Standards (tdwg) [6].

Req. – “Requirement”

This column will reference each specific requirement(s) that relate to and justify the term. Such requirements also specify the format, values, functions and usage of each term.

Discussion

This column details the meaning and practices for each term and identifies known issues related to the term that may require attention during implementation.

TABLE I. IOOSBDS TERMINOLOGY TABLE

Term	Class	NS	Req.	Discussion
scientificName	Taxon	dwc		The lowest level taxonomic name to which the occurrence is identified, in scientific (Latin) form. If the traditional Genus-and-species scientific name is available, use both words, and if identified to subspecies, use all three. If identified at a higher level (e.g., family) use only the single word that expresses the lowest level name available.
recordedBy	Occurrence	dwc		The name of the person or institution that made and recorded the observation. Can be a delimited list.
sampleWidthInMeters	Event	ioosbds		The width in meters of a sample from which the observation/collection was drawn.
waterTemperatureCelsius	Event	ioosbds		The temperature of the water, expressed in degrees Celsius. Question: Is there a consistent method for measuring temperature associated with an event (the sampling event) or will it be necessary to capture more data or metadata about the temperature?
Additional Dublin Core terms: type, language, rights, rightsHolder, accessRights, bibliographicCitation	all	dcterms		IOOS will evaluate these terms to determine their applicability to IOOS Biological Data Services. Each of the terms listed in this row is a Dublin Core term that is included in the ratified Darwin Core terminology. The functional meaning of each one may be relevant, but IOOS may address the requirement by other means, for example through FGDC, ISO, or CF metadata. It has not been determined if Dublin Core is a desirable practice to apply generally. One Dublin Core term, "modified" (see above), is recommended for IOOS in its Dublin Core form because of its definite and distinct functional usefulness.

D. IOOS BDS XML Guide

1) Purpose of the XML Guide

XML guide is also based on Darwin core XML guidance.

This section is the beginning of an "ioosbds XML Guide". The first component of this guide is a schema, shown below using XML. The provided xml is an "xsd" or "XML Schema Definition" document, that lists the complete "ioosbds" terminology. The xsd depiction of the schema terms can be further developed to define more details about the terms, if needed. The xsd included here is a start.

For further use of terminology, in xml form and others, ratified Darwin Core provides an XML Guide, found at [6]. This Darwin Core XML Guide is recommended for "ioosbds" wherever it applies.

XML is not the only way to depict or to use "ioosbds" terms. Database definitions, for example, a mysql "CREATE TABLE" statement, or web service configurations such as those for ERDDAP may use the proposed terms without relying on the xsd in exactly the form provided here.

2) Namespace Policy

One feature of the Darwin Core XML Guide is retained and recommended for this project. This is to use namespace references in all cases where IOOS terms are used. The xsd below shows this. Namespaces are declared at the start of the document, including "dcterms", "dwc" and "ioosbds". Further, when in use, each term carries its namespace as a prefix, for example, ioosbds:sampleLengthInMeters is a term from the IOOS namespace. The Darwin Core namespace policy is at [6].

3) Sample Schema Definition ("xsd")

The sample xml document called a "schema definition" or ".xsd" is provided in the Appendix. As a stand-alone file this document would have a file extension ".xsd". The xsd in appendix is a format for listing terms in "ioosbds". Note the xsd also provides urls that are namespace references. For example, the namespace reference for ratified Darwin Core is xmlns:dwc="http://rs.tdwg.org/dwc/terms/". This means that the url "http://rs.tdwg.org/dwc/terms/" links to definitions for all the terms prefixed with "dwc:".

E. CF Conventions

We are also applying CF Conventions at the field level. This is to assure that IOOS Biological Project will be compatible with other geophysical datasets.

CF Conventions refer to Climate and Forecast Metadata Conventions and are associated with both field-level and dataset-level metadata protocols for use with NetCDF [12].

An essential focus of CF Conventions at the field level is to enforce data-type definitions and standards. This is especially important for data types related to georeferences and temporal references.

Observing CF Conventions is also compatible with serving data through ERDDAP the webservice chosen for this project to expose biological data. Essentially, if CF Conventions are neglected during data development, they will have to be revisited to make best use of data when it does become configured in ERDDAP. Therefore, a policy of building CF Conventions into schema definition and data development will assure ERDDAP and IOOS physical data compatibility.

Beyond field-level metadata, CF Conventions apply at the dataset level as well. The IOOS BDP has determined that FGDC or ISO-19115-2 will be the primary protocol for dataset-level metadata, and applicable elements of FGDC and/or ISO is used to populate CF dataset-level metadata where applicable. There may be practices in CF dataset-level metadata that are not directly paralleled in FGDC or ISO, in which case CF dataset-level metadata may have some additional detail that will not also be in FGDC or ISO.

F. Metadata standards

FGDC is the Metadata standards adopted for IOOS BDP .

1) Dataset-Level Metadata Practices (FGDC/ISO)

The description provided here relies mainly on FGDC specifications and tools for metadata. However, ISO-19115-2 is also under consideration for IOOS BDS. The focus of the following section is on the information content of metadata. Comparable information content can be addressed by both FGDC and ISO. Therefore, much of the discussion here is independent of implementation in either FGDC or ISO. Still, the primary examples, where given, are for FGDC.

a) Goals of Metadata

The following discussion of goals of metadata focuses on several functional themes that help organize the goals. The table that follows shows FGDC examples of how metadata serves these goals.

Metadata can often be time-consuming to prepare and difficult to resolve. This can place a burden on participants. The design recommendation for IOOS BDP is to prioritize metadata decisions based on what best serves participants and users of the data. The themes used below to organize metadata functions are also intended to assist prioritization, so that functions most relevant to IOOS and its participants and users can receive the most attention.

Identification

This function of metadata identifies the dataset, the researchers, the data gatherers, collaborators, funders, and associated programs. Identification gives the dataset a reliable and useful name. A name such as JK34_IomL_95 is not likely to promote usability. A name such as "Near-shore Invertebrate Survey" could be more specific and helpful.

Attribution

Institutions that provide data can expect to get credit when their data are used. Attribution means that metadata documents precise citation requirements, made easy for users to understand and apply. There may be additional use constraints or disclaimers that the data originator wants users to understand and consent. Metadata will contain text for users to directly obtain attribution information.

Referral

Metadata contains information to refer users to the originators or other holders of the data. Referral information should include the name and contact information for principal investigators and other staff. Referral information can also include URLs for an institution or individual's home page, special project pages, associated metadata or data search or download pages.

Dataset History and Quality Information

Data resource history documents relevant steps and details of data gathering, processing and curation. All datasets go through some processing where choices are made concerning, for example, observing and recording methods, data classification, vocabulary, and data qualification criteria. Scientific datasets can potentially go through many very complex steps, all important to document if users are to understand and use data properly.

Quality information can include statements of quality standards that were applied during data gathering and subsequent processing. Quality information can also include formal or incidental quality reviews and follow-ups that occurred over time as a data resource has been used and maintained.

Discovery

Data discovery refers to tools and practices designed to help users locate resources. Resource locating scenarios include automated searches using local or internet-wide search engines. Other scenarios include users manually browsing summary and descriptive information. Key features for discovery include keywords, such as representative taxonomic keywords (scientific- or common-name form) or thematic keywords.

Additional features that assist discovery include statements of scope and range in metadata. For example, date range, geographic extent, expressions of depth range or included habitats or other boundary definitions can be placed in metadata to assist both human and automated search and review.

Understandability

Understandability refers to a suite of metadata features that aid the user in understanding the full scientific background of data in relation to original purposes and most importantly, in relation to the application for which the current user desires the data.

Features that assist understandability include broad narrative components such as an abstract, and description of purpose and methods. These components of metadata will reflect the original context of the program that gathered and developed the data. For best utility, these narratives may be oriented to anticipated users who may want to apply the data to new types of research and application. Goals for this function may include alerting users to caveats and special conditions in the dataset that may make use more complicated or require special attention.

There are more specific ways metadata can support proper use and understanding of a data resource. For example, in FGDC, the "Entities and Attributes" section provides reference to define how all terms in a dataset are defined and used by the data originator. This can get very specific. For example, metadata can contain a list of all the values used in the dataset to identify sampling protocol, and provide a narrative and/or quantitative description of each method. There is no inherent limit on the level of detail metadata can capture this way, although users may observe practical limits.

A relevant example for IOOS BDP is the use of metadata to explain sampling protocol. Darwin Core and IOOS terminology provide for data originators to record some brief text to represent sampling protocol, most likely an abbreviation or code, but also possibly a short descriptive phrase. For further information, the "Entities and Attributes" section can include special detail dedicated to sampling protocol. In "Entities and Attributes", metadata developers will record each distinct value used for sampling protocol, and provide a more in-depth explanation of each value used for sampling protocol.

Metadata Goals in FGDC Example

The following table shows how some components of FGDC metadata serve goals as described above.

Metadata Goal	FGDC Section(s)
Identification, Attribution and Referral	<idinfo><citation> - include title, author, publisher, publication date and location. Together these components construct citation text. In addition, the <citation> element contains reference to URLs for the dataset and explanation of the URLs if necessary. <idinfo><ptcontct> - identifies and provides contact information for a current representative for the dataset <idinfo><datacred> - additional information about developers of the data and processes, collaborators, sponsors <idinfo><useconst> and <idinfo><accconst> - use constraints and access constraints <distinfo> - details contacts, URLs, procedures and instructions as well as disclaimers for users interested in obtaining data.
History and Quality	<idinfo><status> - for recording status of the dataset <dataqual> - includes description of data development factors that may effect data quality. Summarizes georeferencing method. <dataqual><lineage><method> - provides for explaining research and data gathering methods in detail. Note this also serves the goal of "Understanding" (below). <dataqual><lineage><procstep> - records steps in processing data from original form gathered to form offered to users. <metainfo> - records the protocol used to develop metadata and identifies the metadata developer name or institution. Includes contact information
Discovery	<idinfo><timeperd> - identifies the date range or temporal coverage of the dataset. <idinfo><spdom> - identifies the geographic extent of the data, supporting both descriptive geography and bounding coordinates <idinfo><keywords><theme> - provides for thematic keywords that describe high-level characteristics of the study data, such as "coral reef decline" or "fish abundance". These can be specific to data originator and data user needs, though use or development of a thesaurus may be recommended. <idinfo><keywords><place> - geographic keywords for local or internet search <idinfo><taxonomy> - provides for a list of representative taxonomic keywords for search
Understanding	<idinfo><descript> - contains sections such as abstract, purpose and supplemental information to document features of the dataset. <eainfo> (entity and attribute information) - provides reference to general data definition, such as a general reference to Darwin Core or IOOS term definitions used in the dataset. Also provides for detailed explanation of specific terms used in any particular dataset. Suitable for explaining important elements in detail, such as sampling protocol.

IV. Extensibility of IOOS DBP project

Much of the extensibility established in IOOS biological data solutions in this project comes from the use of standards for data, metadata and web services.

The technology developed by IOOS is extensible at three levels. Below is a list of the three levels with brief description.

1) At a general level, a wide range of more biological data applications are now possible. A foundation is now in place for more applications beside presence-absence-abundance. Examples of general biology/ecology applications that can now be addressed include species geospatial distribution, abundance. This is possible because the current IOOS project established standards and examples of major data definitions and implementations. Features such as standards selection for data and metadata, georeferencing and temporal referencing, taxonomic resolution, capture of sampling events and methods, integration with CF Conventions, serving with web service technology - these features have now been established and serve as a valuable precedent and investment for extending to biology topics beyond presence-absence-abundance. There are no constraints on geographical (i.e., regional) applicability of these features, so at extensibility options include geography.

2) At the method-specific survey level, this project has demonstrated a specific application of a presence-absence abundance data service. This service is extensible to other surveys that use similar details and any geography. If other surveys vary in some details, such as sampling method, additional terms may be required. It is anticipated that if methods differ, additional terms are a normal expectation. However, a large core of basic terms, such as those for general biological data (#1) as well as many sampling and measurement details, will be common to additional survey types, therefore directly extensible. In cases where new terms are required, the same standards should be consulted for guidelines. Darwin Core and FGDC already address many more scientific, geospatial, biological and methodological topics than those included in the current IOOS solution. It is possible that for extensions for IOOS, the standards, Darwin Core and FGDC already offer solutions. If such solutions are not found in ready form in the standards, then the standards should be used as guidelines for developing new terms as needed.

3) General quantification standards: This IOOS project dealt with quantifying biological life, also referred to as presence-absence-abundance, where the "absence" and "abundance" distinctly mean that a numeric measure of quantity of life forms is expressed. This project did address quantification in a very method-specific way. It is most applicable to surveys that use similar methods to quantify life. For additional features, it is also possible to quantify life in a more general way, a way that is less method-specific, and can be applied to results of more different kinds of surveys and expressions of quantity. Standards for general quantification can be found in Darwin Core in the practices defined in the category "MeasurementOrFact", found in the Ratified Darwin Core [6]. This is the recommended set of terms related to measurement: type, value, accuracy, units, determinedDate, determinedBy, method, and remarks. This Darwin Core guideline is recommended for future use by IOOS for quantification data services because of its consistency with the now-established IOOS standard, its functional extensibility, its compatibility and integration with both Darwin Core and

FGDC as established by IOOS and its widespread acceptability as part of the widely-accepted Darwin Core.

V. CONCLUSION

The informatics framework for biological data described in this work can be applied to other datasets generated in scientific surveys, provide a vehicle for wider dissemination of biological data and reconcile standards for the representation of species presence/absence/abundance data. We recommend employing data definition conventions that are well understood in Integrated Ocean Observing System (IOOS) and combine these with ratified Darwin Core terminology, policies and guidelines.

ACKNOWLEDGMENT

This workshop was supported by U.S. IOOS. We thank the IOOS BDP team who all contributed to the success of this work.

REFERENCES

[1] www.iobis.org

[2] www.gbif.org

[3] E. Vanden Berghe, P.P. Halpin, F. Lang da Silveira, K. Stocks and F. Grassle. Integrating Biological Data Into Ocean Observing Systems: The Future Role of OBIS. OceanObs 09.

[4] http://www.ioos.gov/dif/

[5] Magnuson-Stevens Fishery Conservation and Management Reauthorization Act (P.L. 109-479)

[6] http://dublincore.org/

[7] http://rs.tdwg.org/dwc/terms/)

[8] http://coastwatch.pfeg.noaa.gov/erddap/

[9]http://geoserver.org/display/GEOS/Welcome

[10] http://www.opengeospatial.org/standards/wms

[11] http://www.pfeg.noaa.gov/products/edc/

[12] http://cf-pcmdi.llnl.gov

Appendix. XML schema definition

```
<?xml version="1.0"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:dwc="http://rs.tdwg.org/dwc/terms/"
xmlns:ioosbds="http://ioosdomain/to-be-defined-
page/">
<!-- record level (class="all") terms -->
<xs:element name="dcterms:modified"
type="xs:dateTime"/>
<xs:element name="dcterms:type" type="xs:string"/>
<xs:element name="dcterms:language"
type="xs:string"/>
<xs:element name="dcterms:rights"
type="xs:string"/>
<xs:element name="dcterms:rightsHolder"
type="xs:string"/>
<xs:element name="dcterms:accessRights"
type="xs:string"/>
<xs:element name="dcterms:bibliographicCitation"
type="xs:string"/>
<xs:element name="dwc:datasetID"
type="xs:string"/>
<xs:element name="dwc:institutionCode"
type="xs:string"/>
<xs:element name="dwc:collectionCode"
type="xs:string"/>
<xs:element name="dwc:datasetName"
type="xs:string"/>
<xs:element name="dwc:institutionID"
type="xs:string"/>
<xs:element name="dwc:collectionID"
type="xs:string"/>
<xs:element name="dwc:ownerInstitutionCode"
type="xs:string"/>
<xs:element name="dwc:basisOfRecord"
type="xs:string"/>
<!-- location terms -->
<xs:element name="dwc:decimalLatitude"
type="xs:string"/>
<xs:element name="dwc:decimalLatitude"
type="xs:string"/>
<xs:element name="dwc:geodeticDatum"
type="xs:string"/>
<xs:element
name="dwc:coordinateUncertaintyInMeters"
type="xs:float"/>
<xs:element name="dwc:coordinatePrecision"
type="xs:string"/>
<xs:element name="dwc:minimumDepthInMeters"
type="xs:float"/>
<xs:element name="dwc:maximumDepthInMeters"
type="xs:float"/>
<xs:element name="dwc:verbatimCoordinates"
type="xs:string"/>
<xs:element name="dwc:verbatimCoordinateSystem"
type="xs:string"/>
<xs:element name="dwc:verbatimSRS"
type="xs:string"/>
<xs:element name="dwc:georeferencedBy"
type="xs:string"/>
<xs:element name="dwc:georeferenceProtocol"
type="xs:string"/>
<xs:element name="dwc:waterBody"
type="xs:string"/>
<xs:element name="dwc:islandGroup"
type="xs:string"/>
<xs:element name="dwc:island" type="xs:string"/>
<xs:element name="dwc:country      "
type="xs:string"/>
<xs:element name="dwc:stateProvince"
type="xs:string"/>
<xs:element name="dwc:county" type="xs:string"/>
<xs:element name="dwc:municipality"
type="xs:string"/>
<xs:element name="dwc:locality" type="xs:string"/>
<!-- event terms -->
<xs:element name="dwc:observationDate"
type="xs:dateTime"/>
<xs:element name="dwc:verbatimObservationDate "
type="xs:string"/>
<xs:element name="dwc:samplingProtocol"
type="xs:string"/>
<xs:element name="ioosbds:sampleWidthInMeters"
type="xs:float"/>
<xs:element name="ioosbds:sampleLengthInMeters"
type="xs:float"/>
<xs:element name="ioosbds:sampleHeightInMeters"
type="xs:float"/>
<xs:element name="ioosbds:sampleShape"
type="xs:string"/>
<xs:element
name="ioosbds:sampleAreaInSquareMeters"
type="xs:float"/>
<xs:element
name="ioosbds:sampleVolumeInCubicMeters"
type="xs:float"/>
<xs:element name="dwc:habitat" type="xs:string"/>
<xs:element name="ioosbds:bottomType"
type="xs:string"/>
<xs:element name="ioosbds:visibilityInMeters"
type="xs:float"/>
<xs:element name="ioosbds:visibilityType"
type="xs:string"/>
<xs:element name="ioosbds:waterTemperatureCelsius"
type="xs:float"/>
<!-- occurrence terms -->
<xs:element name="dwc:catalogNumber"
type="xs:string"/>
<xs:element name="dwc:recordedBy"
type="xs:string"/>
<xs:element name="dwc:individualCount"
type="xs:integer"/>
<xs:element name="dwc:sex" type="xs:string"/>
<xs:element name="dwc:lifestage"
type="xs:string"/>
<xs:element
name="ioosbds:observedIndividualLengthInCm"
type="xs:float"/>
<xs:element name="ioosbds:observedMeanLengthInCm"
type="xs:float"/>
<xs:element name="ioosbds:observedMaxLengthInCm"
type="xs:float"/>
<xs:element name="ioosbds:observedMinLengthInCm"
type="xs:float"/>
<xs:element name="ioosbds:lengthType"
type="xs:string"/>
<xs:element name="ioosbds:aphiaID"
type="xs:string"/>
<xs:element name="ioosbds:tsn" type="xs:string"/>
<!-- identification terms -->
<xs:element name="dwc:identifiedBy"
type="xs:string"/>
<xs:element name="dwc:identifcationDate"
type="xs:dateTime"/>
<xs:element name="dwc:identificationQualifier"
type="xs:string"/>
<xs:element name="dwc:identificationRemarks"
type="xs:string"/>
<!-- taxon terms -->
<xs:element name="dwc:scientificName"
type="xs:string"/>
<xs:element name="dwc:kingdom" type="xs:string"/>
<xs:element name="dwc:phylum" type="xs:string"/>
<xs:element name="dwc:class" type="xs:string"/>
<xs:element name="dwc:order" type="xs:string"/>
<xs:element name="dwc:family" type="xs:string"/>
<xs:element name="dwc:genus" type="xs:string"/>
<xs:element name="dwc:subgenus" type="xs:string"/>
<xs:element name="dwc:specificEpithet"
type="xs:string"/>
<xs:element name="dwc:infraspecificEpithet"
type="xs:string"/>
<xs:element name="dwc:taxonRank"
type="xs:string"/>
<xs:element name="dwc:nomenclaturalCode"
type="xs:string"/>
<xs:element name="dwc:scientificNameAuthorship"
type="xs:string"/>
<xs:element name="dwc:vernacularName"
type="xs:string"/>
```

```
   <xs:element name="dwc:taxonRemarks"
type="xs:string"/>
   &lt;/xs:schema&gt;
```