

Original article

A portal for the Ocean Biogeographic Information System

Un portail pour le Système d'information biogéographique sur l'océan

Yunqing Zhang *, J. Frederick Grassle

Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Rd., New Brunswick, NJ, 089901-8521, USA

Abstract

Since its inception in 1999 the Ocean Biogeographic Information System (OBIS) has developed into an international science program as well as a globally distributed network of biogeographic databases. An OBIS portal at Rutgers University provides the links and functional interoperability among member database systems. Protocols and standards have been established to support effective communication between the portal and these functional units. The portal provides distributed data searching, a taxonomy name service, a GIS with access to relevant environmental data, biological modeling, and education modules for mariners, students, environmental managers, and scientists. The portal will integrate Census of Marine Life field projects, national data archives, and other functional modules, and provides for network-wide analyses and modeling tools.

© 2002 Ifremer/CNRS/IRD/Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

Résumé

Depuis son démarrage en 1999, le Système d'information biogéographique sur l'océan (Obis), est devenu à la fois un programme international de recherche et un réseau de bases de données biogéographiques. Le portail Obis en place à l'Université Rutgers fournit les liens et l'interopérabilité fonctionnelle aux systèmes de bases de données membres du réseau. Ce portail permet une recherche des données dispersées, un service de dénomination taxinomique, un système d'information géographique avec accès aux données environnementales adéquates, de la modélisation biologique et des modules d'enseignement pour les marins, les étudiants, les managers en environnement et les scientifiques. Le portail intégrera les projets de terrain du Recensement de la vie marine, les archives nationales de données et les autres modules fonctionnels ; il fournira également des outils de modélisation et d'analyses en réseau.

© 2002 Ifremer/CNRS/IRD/Éditions scientifiques et médicales Elsevier SAS. Tous droits réservés.

Keywords: OBIS; Portal; Biogeography

Mots clés: OBIS; Portail; Biogéographie

1. Introduction

The Ocean Biogeographic Information System (OBIS) is a user-friendly, web-based provider of global geo-referenced information on accurately identified marine species. Powerful new on-line tools for visualizing relationships among species and their environment will contribute to better understanding of the ecology, evolution, and biogeography of marine organisms. OBIS will assess and integrate geo-referenced ocean data from multiple sources

across disciplines. Multiple approaches to visualization of these data will yield a better understanding of the complex interrelationships among species and the patterns of these relationships with respect to biological, chemical, and physical processes. From this more precise and comprehensive view of the marine environment will come new hypotheses about the origins and maintenance of marine biodiversity and the roles of species in ecosystem function. Users of OBIS, including mariners, students, environmental managers and scientists, will gain a dynamic view of the multi-dimensional oceanic world (Grassle, 2000).

OBIS, as the information component of the Census of Marine Life (CoML) and the marine component of the

* Corresponding author.

Global Biological Information Facility (GBIF), is committed to:

- energize regional, national, and international scale development of ocean biogeographic and systematic databases,
- foster collaboration and interoperability by promoting standards and protocols,
- advance integrated biological and oceanographic research by supporting a multidisciplinary ocean information portal and
- speed the dissemination of, and public access to, ocean biogeographic information.

Some of the initial ideas for OBIS were developed at a CoML meeting on bottom-dwelling ocean life in October 1997. Recommendations from this workshop led to a web site (<http://marine.rutgers.edu/OBIS>) at Rutgers in 1998 to demonstrate the initial OBIS concept (Grassle and Stocks, 1999). The first OBIS International Workshop was held on November 3–4 1999, Washington, DC (Grassle and Decker, 2000). In May 2000, US Government Agencies in the National Oceanographic Partnership Program together with the Alfred P. Sloan Foundation funded eight research projects to initiate OBIS. In May 2001, the US National Science Foundation funded Rutgers University to develop a global portal for OBIS. Also in May, the OBIS International Committee (IC)¹ was formed and its first meeting was held in August 2001. At this meeting, the Committee set the goals for OBIS development and recommended that a technical working group be set up to make ongoing OBIS projects interoperable through the OBIS portal (<http://iobis.org>). In December 2001, the technical working group 2 met in Lawrence, Kansas and protocols and data exchange standards for interoperability were agreed upon. By February 2002, all NOPP-funded OBIS data projects, plus the NSF-funded SeamountsOnline, were made interoperable through the OBIS portal. The International Committee met again in Leiden, the Netherlands, February 2002. It was the consensus of the IC that OBIS will be a federation of functional modules established by its worldwide members and partners. Under the direction of the IC, OBIS members and partners work together with the OBIS secretariat to provide data, GIS, analytical tools, modeling, and education services to users. In the following section, we will review each of the services OBIS provides, including methods, partnerships, current developments, and plans for the immediate future.

¹ Neil Ashcroft, Geoff Boxshall, Mark J. Costello, Daphne G. Fautin, Kim Finney, Rainer Froese, Dennis P. Gordon, J. Frederick Grassle, Yoshihisa Shirayama.

² Eli Agbayani, Ann Bucklin, Robert Buddemeier, Joe Cope, Keith Hunsinger, Frank O'Brien, Josephine Rius, Gary Rosenburg, Warren Sass, Karen Stocks, Vardis Tsontos, David Vieglais, James Wood, Yunqing Zhang.

2. OBIS service

2.1. Distributed data service

Global geospatial information on accurately identified marine species provides the foundation for the development of functional components to provide new environmental management tools, educational products, and the common database on marine life needed to stimulate scientific research on marine species and their ecological and evolutionary relationships. To accommodate the diverse data sources, information systems, and products associated with OBIS members/partners, OBIS maintains a network of distributed data systems and a global portal where users can access the whole network through one web-interface. For the initial proof of the concept, the OBIS technical working group selected http as the communication protocol and xml for data exchange within the distributed OBIS network. The Open GIS standard (<http://www.opengis.org>) was also adopted as part of the OBIS vocabulary. The standard is to specify the communication interface between the portal server and member systems with no restrictions on their internal structure, giving each member freedom in system implementation. The seven NOPP-funded OBIS databases and the NSF-funded SeamountsOnline have been able to communicate with the OBIS portal using the defined protocols and standards. As a result, the world's principal databases on fish, octopus, squid, anemones, corals, zooplankton, and seamount fauna are integrated for the first time. Through one single web-based user interface, users can now do one-stop data shopping among more than 400,000 occurrence records covering the oceans of the world. A list of the projects including their URLs can be found in Appendix I.

The success of the OBIS prototype development shows that a distributed system with a portal provides adequate service to users and that open protocols and standards such as http and xml are the right choice for data exchange. Meanwhile, we have also identified directions for further development to support OBIS rapid growth in service and membership.

As more and more functional modules are incorporated into the OBIS portal, OBIS core xml will be expanded to include complex geographical and species-level information. More search options require that nested logical and comparison operators be present in the request to search structured data. In addition, registry and crawling services are needed to maintain an efficient search capability for an increasing number of distributed members. Considering these requirements, we find that Distributed Generic Information Retrieval (DiGIR) can be the solution for OBIS capacity enhancement. Being developed as an open source software, DiGIR is designed to use open protocols and standards (http, xml, uddi). Its most interesting feature is that the protocol, software, and semantics are de-coupled in

order to achieve federation schema pluggability (Blum et al., 2001). The goal is that data federations with distinctively different natures can all use DiGIR as the base framework and then plug in their own federation schema. The implication for OBIS, in the short term, is that if DiGIR is used within OBIS and OBIS upgrades, its xml data core (federation schema), the backbone of the distributed network will not have to be altered. Only the data provider and presentation layer need to be modified. In the long run, with its many functional modules, OBIS will be a federation of data federations, each of which has its own federation schema (i.e. museum collections, surveys, conservation information, genetics) but uses DiGIR as the backbone for its distributed network. Bridges can then be developed between these functional modules for further integration with the OBIS global portal. We expect DiGIR to play an important role in OBIS's development as the marine component of GBIF.

OBIS will work with all CoML field projects (e.g. Tagging of Pacific Pelagics (TOPP), Pacific Ocean Salmon Tracking Program (POST), Natural Geography in Shore Areas (NaGISA), etc.) to make data from expeditions interoperable through the global OBIS portal. A virtual database will be developed to assist the projects to communicate with the OBIS portal using OBIS protocols and data standards. OBIS will also form close alliances with government agencies (e.g. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, National Institute of Water and Atmospheric Research (NIWA), New Zealand, and in the US, National Ocean Data Center (NODC) and National Virtual Ocean Data System (NVODS)) to make national data archives available through OBIS. In particular, OBIS plans to work with major national fishery agencies and FAO/FIGIS to serve interoperable fishery data. This would be a major accomplishment for OBIS, one that has escaped most national agencies for decades.

2.2. Taxonomy name service

Accurate taxonomic information is a central feature of the OBIS data service. Species' distribution patterns cannot be accurately identified without complete knowledge of synonyms and the history of name changes. Common names are the key to providing user-friendly service to resource managers, pre-collegiate students and the general public. Partnered with Species 2000 and FishBase (<http://www.fishbase.org>), OBIS has developed a prototype name service to provide common name/scientific name and synonym translation. OBIS is currently working on integrating the name service with its distributed data search, to build a foundation for its educational program.

2.3. GIS and environmental data service

With the development of a global ocean observing system and advancements in physical, chemical, and bio-

logical ocean modeling, synoptic ocean environmental data products have proliferated. However, these products are not available in forms readily accessible for marine biogeographic studies. This is, in part, due to a lack of open Internet access and data interoperability. One of the goals of OBIS is to make these data available on its portal as part of its web-based GIS service. In the distributed data search module, OBIS is providing a C-square mapping service to users (Rees, 2002). The Easy Netviewer (<http://netviewer.usc.edu/web/index.html>) will also be incorporated into the OBIS portal so that users can have access to customized GIS services offered by this versatile tool. OBIS is also developing an alliance with ESRI to provide an Arc Internet Map Server-based mapping service for automatic overlay of geo-referenced species data with environmental data layers. Most importantly, OBIS will be working with NVODS, Ocean.US and NODC to make environmental datasets under their management available to OBIS users. With such collaborations, OBIS will in effect establish a bridge between elements of the Global Ocean Observing System (GOOS) and GBIF, bringing the oceanographic and biology communities together in a joint effort to study environmental change and its impact on marine habitats.

2.4. Biogeographic modeling service

As an extension to its key data service, OBIS is working with its partners at the San Diego Super Computer Center and the Kansas Geological Survey to make Biodiversity Species Workshop (http://biodi.sdsc.edu/bsw_home.html) and LoiczView (<http://www.palantir.swarthmore.edu/loicz/help>) operational on the OBIS portal. Biodiversity Species Workshop is a set of analysis tools based on the GARP modeling algorithm. It allows researchers to perform a variety of analyses using biodiversity data. The LoiczView web-based geospatial cluster analysis, statistical, and visualization tool (Maxwell and Buddemeier, 2002) is available to OBIS users via the partnering arrangements between the Hexacorallia project, IGBP-LOICZ, and its developers. The tool is scale-independent, and works with any formatted geo-referenced data; it has a user-upload option as well as direct-transfer capability from the Hexacorallia database; all data, images, results, and intermediate files and tables are downloadable.

OBIS is working with CoML's History of Marine Animal Populations (HMAP), and Future of Marine Animal Populations (FMAP), two other major components of CoML to provide the temporal context for charting fluctuations of species' distributions past and present, and for building the foundation for modeling and prediction of future oceanic communities.

2.5. Education programs

Disseminating ocean biogeographic information to the public is one of OBIS' major goals and OBIS has been

developing education programs for pre-collegiate students. The OBIS portal and the Institute of Marine and Coastal Sciences at Rutgers University are developing a series of hands-on classroom and Internet-based activities using OBIS data to illustrate core scientific principles taught as part of schools' existing curricula. These integrated thematic modules will be used by students in grades 6–12. The modules will include lessons on defining and quantifying biodiversity, building an ecologically balanced living space, extinction, and the value of biodiversity. Collaborative projects that will allow students to exchange ecological information and become amateur taxonomists using web-based lesson plans and the OBIS-portal are planned. A minimum of four classroom activities and two Internet applications will be developed and field-tested by the participants, and the draft modules will be posted on the OBIS portal by September 2002 (McDonnell, 2001).

2.6. Partnership with the UN Atlas of the Oceans

The CoML is collaborating with the UN Atlas of the Oceans, an international initiative of UN institutions and their partners. The Atlas provides users with up-to-date information relevant to sustainable development of the oceans and ocean science. CoML serves as Editor of the Biology topic in the About the Oceans part of the Atlas, developing and maintaining information and links on biology and providing direct access to OBIS. This partnership integrates two major international efforts designed to inventory and make available ocean information. Needless to say, it avoids possible duplication of efforts. The Atlas is web-based (<http://www.oceansatlas.org/>) and freely accessible to everyone.

3. Conclusion

OBIS is developing and promoting protocols, standards, and software for managing and disseminating ocean biogeographic information. Through a global portal and a worldwide network of distributed data systems, OBIS provides geo-referenced species data, a taxonomy name service, GIS analysis with environmental data, biogeographic modeling, and education services to researchers, students, conservation organizations, resource managers, and the general public. OBIS provides an information integration service to all CoML field projects and, together with HMAP and FMAP, works to answer the questions: what lived in the oceans, what lives there now, and what will live there in the future. OBIS has formed strategic alliances with major international organizations and national and regional governments to build a unique platform for collaboration among biologists, ecologists, and oceanographers. With its members and partners, OBIS is building a global infrastructure for education and research in ocean biogeography.

Acknowledgements

We thank all OBIS participants for their contributions and unfailing support.

Appendix A: Current OBIS projects:

- *The Fishnet Distributed Biodiversity Information System*
URL: <http://habanero.nhm.ku.edu/fishnet>
Dr. Edward Wiley, Natural History Museum, University of Kansas.
- *Development of a Dynamic Biogeographic Information System: A Pilot Application for the Gulf of Maine*
URL: <http://netviewer.usc.edu/web/index.html>
Dr. Dale Kiefer, Wrigley Institute of Environmental Studies, University of Southern California.
- *Biogeoinformatics of Hexacorallia (Corals, Sea Anemones, and their Allies): Interfacing Geospatial, Taxonomic, and Environmental Data for a Group of Marine Invertebrates*
URL: <http://www.kgs.ukans.edu/Hexacoral>
Dr. Daphne Fautin, University of Kansas and Dr. Bob Buddemeier, Kansas Geological Survey.
- *Expansion of CephBase as a Biological Prototype for OBIS*
URL: <http://www.cephbase.utmb.edu/>
Drs. Phillip Lee and James Wood, University of Texas Medical Branch.
- *A Biotic Database of Indo-Pacific Marine Mollusks*
URL: <http://errato.acnatsci.org/obis>
Dr. Gary Rosenberg, The Academy of Natural Sciences, Philadelphia.
- *ZooGene, a DNA Sequence Database for Calanoid Copepods and Euphausiids: An OBIS Tool for Uniform Standards of Species Identification*
URL: <http://www.zoogene.org>
Ann Bucklin, University of New Hampshire Durham, NH; Bruce W. Frost, University of Washington, Seattle, WA; Peter H. Wiebe, Woods Hole Oceanographic Institution, Woods Hole, MA; Michael J. Fogarty, NOAA/NMFS Northeast Fisheries Science Center, Woods Hole, MA.
- *Diel, Seasonal, and Interannual Patterns in Zooplankton and Micronekton Species Composition in the Sub-tropical Atlantic*
URL: <http://www.vims.edu/bio/zooplankton/bats.html>
Dr. Deborah Steinberg, Virginia Institute of Marine Sciences.
- *Census of Marine Fishes (CMF): Definitive List of Species and Online Biodiversity Database*
URL: <http://www.fishbase.org>
Dr. William Eschmeyer California Academy of Sciences and Dr. Rainer Froese, FishBase Coordinator, Institut für Meereskunde.
- *SeamountsOnline*
URL: <http://seamounts.sdsc.edu/>

Dr. Karen Stocks, San Diego Supercomputing Center and Scripps Institute of Oceanography.

References

- Blum, S., Vieglais, D., Schwartz, P.J., 2001. DiGIR — Distributed Generic Information Retrieval. Presented at 2001 TDWG Meeting, Sydney, Australia. DiGIR.sourceforge.net/events/20011106/DiGIR.ppt.
- Grassle, J.F., 2000. The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multi-dimensional geographic context. *Oceanography* 13 (3), 5–9.
- Grassle, J.F., Decker, C.J., 2000. Feature: The Ocean Biogeographic Information System (OBIS). *ASC Newsletter* 28 (6), 8–9.
- Grassle, J.F., Stocks, K.I., 1999. A Global Ocean Biogeographic Information System (OBIS) for the Census of Marine Life. *Oceanography* 12 (3), 12–14.
- Maxwell, B.A., Buddemeier, R.W., 2002. Coastal typology developments with heterogeneous data sets. *Reg. Environ. Change* (in press).
- McDonnell, J.D., 2001. Best practices in marine and coastal science education: lessons learned from a national estuarine research reserve. In: Fedler, A.J. (Ed.), *Defining Best Practices in Boating, Fishing, and Stewardship Education*. Prepared for the Recreational Boating and Fishing Foundation under Contract #RBFF-00-C-004, 31 July 2001.
- Rees, T., 2002. “c-squares” – a new metadata element for improved spatial querying and representation of spatial dataset coverage in metadata records (abstract). *EOGEO 2002 Proceedings*, Ispra, Italy, 2002. World Wide Web publication, available via the EOGEO website at: eogeo.jrc.it/.