January 15, 2004

OBIS Five-Year Plan

I. Foreword

Because of their area, volume, and diversity of life, the world's oceans are the dominant component of the biosphere. Thus, an assessment of life on Earth must in major part be an assessment of life in the world ocean – a Census of Marine Life. The complexity of the marine ecosystem, and its interactions with social and political systems demand an interdisciplinary and integrated approach. Traditional, discipline-centered research methodologies yield a wealth of snapshots of the complex and ever-changing marine world. The challenges are to fill the gaps in these insights, to synthesize coherent patterns of marine life in space and time, and to develop testable hypotheses and predictive models of the origin and maintenance of these patterns.

Today, as never before, the tools are at hand to meet this challenge – to conduct quantitative, geographically- and temporally-explicit observations and analyses of the living ocean. Taxonomists have new tools to define and identify species through combinations of genomic and morphological analyses, greatly aided by access to worldwide knowledge resources through the Internet. Remotely sensed and in situ observations are increasingly being made available through the Global Ocean Observing System (GOOS), creating an unprecedented amount of geo-referenced environmental and ecosystem data. Computer and communications capabilities permit rapid assembly, and meaningful analysis of immense volumes of diverse data. Moreover, earth and life scientists have developed highly capable systems for planning, coordinating, and executing coherent and effective programs on a global scale.

The Ocean Biogeographic Information System (OBIS), the information component of the Census of Marine Life (CoML), is a rapidly developing international science infrastructure. It will provide access to data content, information infrastructure, and informatics tools, (maps, visualizations, and products from models) through an on-line, dynamic, global 4-D (the three dimensions of space plus time) atlas of biogeographic information (Appendix A, Mission and Goals). The atlas will be used to reveal interesting spatial and temporal patterns, generate new hypotheses about global marine ecosystems, and guide future field expeditions. These hypotheses are unlikely to be tested by experiments but will be supported or rejected by the biogeographic patterns that are the reality of our planet. The scope of OBIS offers new challenges in data management, scientific cooperation and organization, and innovative approaches to data analysis. The on-line, digital atlas developed by OBIS is expected to provide a fundamental basis for societal and governmental decisions on how to harvest and conserve marine life.

II. OBIS Organization and Status

OBIS, under the direction of an International Committee and Secretariat, has been rigorously pursuing research and development in the following directions:

A. Data access: gathering comprehensive, accurate, quality-controlled digital data

Species data are needed to study food webs, population dynamics, evolutionary history, habitat, biogeography, species introductions, and criteria for establishment of marine reserves. Even in the case of broad-scale questions, a "black box" approach to taxonomy is inadequate because particular species are likely to play roles disproportionate to their abundance. Indeed, our ability to address critical ecological questions hinges on our ability to know what organisms we are dealing with. The historical development of taxonomy has brought about multiple views of species definitions, which further complicate the issue of correctly using data from existing collections. Adequate quality control is achieved only through direct involvement of taxonomic authorities for each group. OBIS has mobilized the marine systematic community to digitize and store geo-referenced distribution data on accurately-identified species. The OBIS community is also providing expertise on marine systematics to Species 2000, ITIS (the Integrated Taxonomic Information System) and the Electronic Catalog of Names of Known Organisms Program of GBIF.

B. Data integration: integrating heterogeneous data sources

With the advance of new technologies, digitization of existing records, and active field explorations such as those in the CoML, the production of heterogeneous data with complex interrelationships will increase daily. The volume and diversity of these data constitute an urgent research issue in data integration and interoperation. The major topics include:

1) community-endorsed global data and metadata standards,
2) new tools and algorithms for semantic mapping and integration,
3) efficient algorithms for data aggregation, and specifically, geospatial and temporal data aggregation.

All of these topics are central in building the infrastructure for global biodiversity and environmental information systems and OBIS is playing a leading role in developing research projects and prototypes to solve these problems.

C. Data analysis: providing the ability to discover scientifically important patterns and unique events

The emergence of GOOS and GBIF indicates a paradigm shift in earth system sciences. The ever-increasing volume of ecosystem data and their successful integration pose new challenges to researchers. New, scalable algorithms and tools must be developed to efficiently search for scientifically challenging, temporal-spatial patterns and to identify unique, sometimes disruptive, ecosystem events in large, integrated databases. Data mining techniques are inductive in nature, and the main purpose of scientific data mining is to assist scientists at the initial stage of scientific discovery, i.e., generating hypotheses based on observations and heuristic relationships. Patterns identified with automatic data-mining techniques have to be examined carefully by domain experts and further validated by deduction-based methods. Because of the complexity and our lack of understanding of the ecosystem processes, we need to combine data mining, traditional statistical analysis, and mathematical modeling to understand complex marine ecosystems and to formulate meaningful predictions. OBIS promotes a synthetic and

cooperative approach to ecosystem study and will serve as a global forum for integrated ocean biodiversity study and biogeographical research.

D. Data visualization: developing a new generation of marine Geographical Information System (GIS) and other visualization tools

Data visualization is an important part of the knowledge-discovery process. It is particularly important in marine ecosystem studies because of the spatio-temporal nature of ocean ecosystem data. Current GIS systems cannot deal well with 4-D data so new data structures and algorithms must be developed. Meanwhile, many existing GIS tools cannot meet the user demand for Internet-based mapping services. OBIS is actively working in these two areas and some products have already been offered on the OBIS portal. The goal of OBIS is to become in five years:

- a world-wide information facility serving and archiving data from global marine biodiversity and global ocean observing system studies.
- an international standard body for data modeling, service discovery, and information exchange in the realm of ocean biogeography and marine biodiversity.
- a global forum for integrated ocean biogeography and biodiversity studies.

E. Data utilization: developing outreach initiatives that are supported by the data access, analysis and visualization tools outlined above.

OBIS is creating a central hub for the collection, access, analysis and visualization of data on marine organisms. In the next five years we will build on these strengths to create educational outreach projects for students, teachers and members of the public. Projects will include creating content specifically for teachers and students such as the NSFDL CephSchool project and lesson plans available at http://iobis.org/educatio.shtml.

Below, we lay out the five-year plan in detail. Section III gives an OBIS history and the present Members of the OBIS Federation. Section IV discusses the plans for bringing in OBIS databases. Section V describes some of the major services OBIS will provide, and outlines the timetable for new partnerships and further organizational development of OBIS. Section VI analyzes development costs and funding sources.

III. OBIS history and present federation

OBIS developed from a CoML-sponsored Benthic Census Meeting (http://marine.rutgers.edu/OBIS/origin/MtgRprt.htm) and a prototype OBIS Web site (http://marine.rutgers.edu/OBIS[1]). In 2000, the National Oceanographic Partnership Program (NOPP) requested proposals for OBIS projects and funded eight projects through support from the Alfred P. Sloan Foundation, the Office of Naval Research, and the National Science Foundation (NSF). A more restricted NOPP competition in 2002 resulted in an additional OBIS project on marine mammals, turtles, and birds. An NSF Postdoctoral Fellowship to Karen Stocks at the San Diego Supercomputer Center and

[1] Stocks, K.I., Y. Zhang, C. Flanders, and J. F. Grassle. 2000. OBIS: Ocean Biogeographic Information System. The Institute of Marine and Coastal Science, Rutgers University.

Scripps Institution of Oceanography resulted in an OBIS project on seamounts. Thus, 11 projects have formed the initial components of the OBIS Federation (Table 1).

Table 1. The 11 projects that formed the first components to OBIS.

Title	Principal Investigator	Organisation	Comment
Biotic Database of Indo-Pacific Marine Mollusks	Gary Rosenberg	Academy of Natural Sciences, Philadelphia	
Census of Marine Fishes (CMF)	William Eschmeyer and Rainer Froese	California Academy of Sciences and University of Kiel	Data are from two databases, the Catalog of Fishes and FishBase
CephBase	Phillip Lee and James Wood	University of Texas Medical Branch	
Diel, Seasonal and Interannual Patterns in Zooplankton and Micronekton Species Composition in the Subtropical Atlantic	Deborah Steinberg	Virginia Institute of Marine Sciences	
Digital Archival of Marine Mammal/Bird/Turtle Data for OBIS	Andrew Read	Duke University Marine Laboratory	
FishNet	Edward Wiley	Natural History Museum, University of Kansas	
Gulf of Maine Biogeographic Information System (GMBIS)	Dale Kiefer, Bob Branton and Lou Van Guelpen	Wrigley Institute of Environmental Studies at the University of Southern California, Fisheries and Oceans Canada, and Atlantic Reference Centre (HMSC), Canada	USC provides GIS (EASy) for DFO fishery and ARC museum data
The History of Marine Animal Populations (HMAP)	Neil Ashcroft	University of Hull	
SeamountsOnline	Karen Stocks	San Diego Supercomputing Center/Scripps Institution of Oceanography	
ZooGene	Ann Bucklin	University of New Hampshire	

Each of the seven Census of Marine Life Initial Projects (http://coml.org) will contribute to OBIS with the start of the field programs in 2004-5. The long-term Chesapeake Bay database of the Trophic Interactions in Estuarine Systems (TIES) program and the

Smithsonian Institution's National Marine and Estuarine Invasion Database will also join in 2004. Also in 2004, Australia will be establishing a CoML and OBIS Steering Committee, whose role will include securing access to Australian data through OBIS.

IV. Datasets served by OBIS

A. Biological databases

Among the taxa covered by OBIS are corals, sea anemones, bivalves, gastropods, cephalopods, copepods, euphausids, fishes, turtles, marine mammals, and birds (Table 2). Sponges, nemertean worms, aplacophoran mollusks, and migratory fish will be added in 2003. It is aimed to add seaweeds, amphipods, stomatopods, foraminiferans, coccolithophorid phytoplankton, bryozoans, brachiopods, parasitic copepods, salps and planktonic coelenterates in 2004. The OBIS Secretariat has committed a small amount of funding for data rescue of authoritative global or regional species databases.

Several of the OBIS projects already serve data on-line that is not yet served through the portal, for example GMBIS, SEAMAP, and TOPP. The OBIS portal will focus on serving georeferenced data on species whereas its associated databases may serve significantly more information on species biology and ecology.

Regional studies served by OBIS are the Gulf of Maine Biogeographic Information system (GMBIS), Census of Marine Life in the Gulf of Maine (GoM) Initial Project and projects within HMAP. The deep-sea Atlantic bathypelagic biota have been added and distributions of seamount species are available through SeamountsOnline and distributions of planktonic organisms are available through the World Plankton Database. Data from the Canadian Atlantic and Atlantic Ocean bathypelagic biota have also been added. The Natural Geography in Shore Areas (NaGISA) CoML Initial Project will begin providing data on shallow subtidal benthos on a latitudinal gradient from Alaska and Siberia to New Zealand in 2004. The first Mid-Atlantic Ridge data from Patterns and Processes of Ecosystems in the Northern Mid-Atlantic (MAR-ECO) will be available in 2004. Benthic data from the coastal USA and Europe will be added in 2004. Biodiversity data and oceanographic projects covering regions within Australia's Exclusive Economic Zone from the Southern, Pacific, and Indian oceans will come on-line during 2004.

By the end of 2004, examples of datasets from most of the major marine habitats will be available through OBIS, including descriptions of highly endemic species from the Australasian region. Datasets will include estuaries, continental shelves, seamounts, canyons, open-ocean pelagic regions, deep-sea hydrothermal vents and cold seeps, and deep-sea sediments. Additional datasets will come from Tagging of Pacific Pelagics (TOPP), Biogeography of Chemosynthetic Ecosystems (ChEss), and Census of Diversity of Abyssal Marine Life (CeDAMar) CoML Initial Projects.

OBIS will include databases of sounds made by marine animals (passive acoustics) and acoustic signatures of marine animals recorded from high-frequency echo sounder or sonar systems (active acoustics). Acoustic databases will be available in OBIS by 2005.

Several of the expected new datasets in 2003 and 2004 are associated with conservation activities. These include the data needed to produce coral hot-spot maps (Conservation International), distributions of invasive species (Smithsonian Institution), and benthic habitat maps (NatureServe, Marine Biological Association UK/MarLIN, World Wildlife Fund, Canada, and World Conservation Monitoring Centre). National environmental protection agency baseline monitoring studies and international petroleum industry impact statements will also be included. Regional marine planning activities under Australia's Oceans Policy will provide access to national scale mapping of fisheries and non fisheries marine uses.

In addition to the Bermuda time-series dataset OBIS will include the long-term continuous planktonic recorder (CPR) data from the Sir Alistar Hardy Foundation for Ocean Science (SAHFOS), and the Tatoosh Island rocky shore studies in the U.S. Pacific Northwest in 2004.

Table 2. Data that has been committed to OBIS. Data recorded prior to 2004 is already served from the OBIS portal. Data attributed to beyond 2003 may be dependant on funding. OBIS is continually seeking scientists willing to make contributions and collaborate with existing contributors.

Year	Data type	Principal Investigator	Project or Source
	Taxon		
2002	Sea anemones & corals	Daphne Fautin	Hexacorallia
2002	Cephalopods	James Wood, Phillip Lee	CephBase
2002	Gastropods & bivalves	Gary Rosenberg, Phillippe Bouchet	
2002	Fishes	Rainer Froese, William Eschmeyer	FishBase, Catalog of Fishes
2002	Fishes, museums (based on museum collections)	Ed Wiley	FishNet
2002	Copepods & euphausids	Ann Bucklin	ZooGene
2003	Marine mammals, turtles, seabirds	Andrew Read	SEAMAP
2003	Bryozoa	Dennis Gordon	NIWA, New Zealand
2004	Nemerteans	Ray Gibson	Liverpool John Moores University
2004	Turbellarians	Seth Tyler	University of Maine, Orono
2004	Trematodes	Tom Cribb	University of Queensland
2004	Serpulid tubeworms	Harry ten Hove	University of Amsterdam
2004	Pycnogonid sea spiders	Roger Bamber	NHM, London
2004	Nemerteans	Jon Norenburg	Smithsonian
2004	Aplacophorans	Amelie Scheltema	Portal
2004	Sponges	Rob van Soest, Shirley Pomponi	
2004	N. Atlantic Deep-sea	Les Watling, Mark Costello	University of Maine,

	corals		Huntsman Mar. Lab.
2004	Seaweeds	Michael Guiry	AlgaeBase
2004	Amphipods	Jim Thomas, Mark Costello	
2004	Stomatopods	Marjorie Reaka-Kudla	Univ. Maryland
2004	Foraminiferans	Columban de Vargas	Rutgers University
2004	Coccolithophorids	Marie-Pierre Aubry	Rutgers University
2004	Bryozoans	Scott Lidgard, Judith Winston, & Dennis Gordon	
2004	Brachiopods	Christian Emig	
2005	Cumaceans	Les Watling	Univ. of Maine
2005	Parasitic copepods	Geoffrey Boxshall	NHM, London
2005	Planktonic coelenterates & salps	Larry Madin	Woods Hole Oceanographic Inst
2006	Polychaeta	James Blake, Kristian Fauchald	ENSR, Smithsonian
	Region		
2002	Gulf of Maine	Dale Kiefer, Bob Branton, Lou Van Guelpen	Gulf of Maine Biogeographic Information System
2002	Fisheries Regions	Neil Ashcroft	HMAP
2003	Atlantic bathypelagic biota	Phillip Pugh	Southampton Oceanography Centre (SOC)
2003	Atlantic Canada	Bob Branton, Lou Van Guelpen	Bedford Institute of Oceanogaphy
2004	Antarctic (mollusks)	Nathan Cunningham	British Antarctic Survey
2004	Gulf of Maine	Evan Richert, Lew Incze, Tom Trott	GoMAP
2004	Western Pacific 0-30m	Yoshihisa Shirayama	NaGISA
2004	Mid-Ocean Ridges	Paul Tyler	ChEss
2004	Coastal U.S.A. benthos	Brian Melzian	EPA-EMAP
2004	Australia EEZ	Kim Finney	National Oceans Office
2004	Gulf of Alaska	Phil Mundy	
2004	Europe	Carlo Heip Mark Costello, Edward Vanden Berghe	MARBEF EurOBIS
	Habitat		
2002	Seamounts	Karen Stocks	SeamountsOnline
2002	World Plankton Database		NODC/NOAA
2004	Canyons	Karen Stocks	SeamountsOnline
2004	Salmon habitat	David Welch	POST
2004	Vents and seeps	Paul Tyler, Cindy Van Dover	ChEss
2004	Estuaries/Chesapeake Bay	Michael Roman, Walter Boynton, Edward Houde, Michael Kemp, William Boicourt	TIES/CEES
2004	Antarctic benthic deep-sea biodiversity	Angelika Brandt	ANDEEP
2004	Abyssal Plain		CeDAMar

	Applications		
2004	BP Deep-Sea Data	Mark Collins	WCMC
2004	Hot-spot maps	Michael Smith	Conservation Int.
2004	Invasive species	Ross Simon	Smithsonian
2004	Status & trends benthic data		NOAA
2004	Benthic habitat maps	Mark Schaefer	NatureServe
	Time Series		
2002	Bermuda Time Series	Deborah Steinberg	BATS
2003	Fish stocks	Neil Ashcroft	HMAP
2004	Continuous Plankton Recorder data	Chris Read	SAHFOS
2004	Tatoosh Island Northwest U.S	Robert Paine	

B. Environmental data and overlays

Satellite remote sensing provides global synoptic coverage of biogeographically-important geophysical variables. These include sea-surface temperature, chlorophyll a, surface turbidity, and surface roughness. Maps of primary production are calculated on the basis of additional oceanographic information (http://marine.rutgers.edu/opp/index.html). The data from Rutgers' Ocean Primary Productivity Study will be part of OBIS in 2004. Databases of optical signatures of phytoplankton, both from satellites and shipboard flow cytometry, will be sought for inclusion in OBIS as they become available.

The global array of ARGO drifter floats provides conductivity, temperature, and depth data throughout the world oceans and these data are assimilated into ocean models to produce maps of ocean currents. The Global Ocean Data Assimilation Experiment leads the effort to provide these ocean circulation products. A good example of some of the readily available data can be found on the MERCATOR Web site: http://www.mercator.com.fr/en/.

In coastal areas, hourly sea-surface currents are provided for the New Jersey Continental Shelf through the New Jersey Shelf Observing System: http://marine.rutgers.edu/cool/codar/real-time/raw_lr.html. These and other coastal oceanographic data from the observing system are assimilated by ocean circulation models. Model output can be analyzed to determine biogeographically important ocean circulation events and estimate rates of larval dispersal of marine organisms.

OBIS has adopted the OpenDAP/DODS (Distributed Ocean Data System) protocol for handling physical data, and will become interoperable with the developing Ocean.US (operational organizing body for U.S. GOOS) data system by 2006.

C. History of Marine Animal Populations (HMAP)

HMAP, the historical component of CoML, is designed to illuminate the dynamic interplay of anthropogenic and natural factors in the process of long-term change in the diversity, distribution, and abundance of life in marine ecosystems. In order to enter the historical data generated from a vast range of archival sources by each HMAP project,

the University of Hull has made the contents of its database accessible to OBIS through a dedicated server. As part of the OBIS plan HMAP has:

- completed the connectivity process in the near future, so that the HMAP Database will become a full participant in OBIS.
- worked with OBIS to perfect structures for integrating and representing historical data.
- continued to develop HMAP, so that data from all present and future HMAP projects are available to OBIS.

D. Growth of OBIS datasets

Based on addition of the datasets outlined in Table 2, OBIS will make at least 10 million georeferenced species records available on line by 2007 (Figure 1).

V. OBIS Services

A primary mission of the OBIS Portal is to develop a variety of tools for use through the Portal. Priority areas for portal development are: integration and mining services, links to genetic databases, taxonomic name services, and visualization services. Members of the Federation are also developing services available to other Members and the Portal.

A. Data integration and mining service

By the end of 2004 the Portal will develop a quality-control and data-cleaning service. A semi-automated database integration service and a data-standard service will be available by the end of 2005. We expect to have a system-wide, Internet-based, spatial-temporal data-mining service by 2007.

B. Genetic information service

ZooGene already has links to GenBank, and we expect to link other databases to GenBank through the OBIS Portal by 2004. Genetic sequence data is used to separate species and indicate evolutionary relationships. As a result of a workshop early in 2003, this approach is extended to other groups of marine species, and a DNA "barcoding" Protocol for DNA sequence analysis of a uniform set of target genes in conjunction with taxonomic and specimen information has been developed for the Census of Marine Life.

C. Taxonomy name service (in collaboration with Species 2000)

The Portal and some of the OBIS Federation Members have taxonomic name servers. These will continue to develop in conjunction with Species 2000, ITIS, and the GBIF Electronic Catalog of Names of Known Organisms. By the end of July 2004, the latest version of species name server will be integrated with the OBIS data service.

D. Visualization service

The OBIS Portal is presently using C-square Mapper from CSIRO and the Dynamic Specimen Mapper from the Kansas Geological Survey. In 2003 the Portal contracted with for implementation of the EASy 4-D GIS system already in use by some of the OBIS initial projects. Working with the marine mammal, turtle, and bird (SEAMAP) component of OBIS the Portal has a long-term plan to develop an Open Source Map Server.

VI. Organization/Partnership Development

The OBIS internal organization and external partnerships provide the backbone for OBIS services and are the key to long-term, sustainable growth.

A. Organizational development

All CoML projects are automatically committed to serving their data through the OBIS Portal. Other organizations/individuals can contribute by providing data and other services to the OBIS portal that can be accessed by all OBIS users. The major data and service sources described in Table 2 will all contribute to OBIS in the next five years. In 2003, OBIS will examine intellectual property considerations for all contributors and establish mechanisms for automatic transfer of rights and responsibilities into the future.

OBIS will work with national CoML Programs to bring national databases into OBIS. Representatives from countries and regions may wish to set up their own OBIS Portals and campaign for internal recognition and support. For example, in Australia a joint CoML/OBIS Steering Committee is being established to bring together Australian research that is easily aligned with Census goals. The Australian national OBIS component will support the data management requirements of these projects through web enabling datasets and with data modeling and data visualization services. Existing national datasets that can be brought on-line rapidly will also be used to establish an Australian OBIS portal. The protocols and standards used will ensure that the portal and its services form an important element of the international OBIS federation, possibly providing a model for future OBIS expansion in other regions.

OBIS is forming national and regional committees, some of which are combined OBIS and CoML committees. By the end of 2004, OBIS will have formed coalitions with Australia, Canada, China, European Union, India, Japan, New Zealand, South America, United Kingdom, and U.S.A. By the end 2006, all regional nodes will be fully interoperable with OBIS.

In 2004, the Secretariat will seek support for the OBIS Secretariat at The Huntsman Marine Science Centre in Canada and the OBIS Portal at Rutgers University. Some of the present Secretariat functions may be distributed among national CoML program offices. Already, the number of meetings, potential data sources to be developed, and number of members and partners to be communicated with, is growing rapidly. A full-time OBIS Program Manager will be required to coordinate activities.

B. Partnership development

A broad range of affiliations has been formed through OBIS Members, the International Committee (IC, Table 3), and CoML. The major affiliations include but are not limited to GBIF, Intergovernmental Oceanographic Commission (IOC), ITIS and Species 2000, Scientific Committee on Ocean Research (SCOR), DIVERSITAS, International Council for the Exploration of the Seas (ICES), International Association of Biological Oceanographers (IABO), and the International Union of Biology (IUBS). In 2003, the United Nations Environmental Program (UNEP), World Conservation Monitoring Centre

(WCMC), and Global Ocean Observing System (GOOS—through Ocean.US) have become OBIS partners. We are building a relationship with the Ocean Research Interactive Observatory Networks (ORION) through the Program Office.

Table 4. Members of the OBIS International Committee.
* Ex-officio Members. ** Committee chair.

Name	Position	Organisation	Location	Country
Neil Ashcroft	HMAP Research Fellow	University of Hull	Hull	UK
Geoff Boxshall	Merit Researcher	The Natural History Museum	London	UK
** Mark J. Costello	Executive Director	The Huntsman Marine Science Centre	St. Andrews	Canada
Daphne G. Fautin	Curator	The University of Kansas Natural History Museum	Lawrence	USA
Rainer Froese	FishBase-Coordinator	Institut für Meereskunde	Kiel	Germany
J. Frederick Grassle	Director	Rutgers University	New Brunswick	USA
Tony Rees	Data Centre Manager	CSIRO	Hobart	Australia
Yoshihisa Shirayama	Director and Professor	Seto Marine Laboratory	Sharahama	Japan
Edward Vanden Berghe	Manager	Flanders Marine Data Information Centre	Ostend	Belgium
John Wilkin	Assistant Professor	Rutgers University	New Brunswick	USA
* Yunqing (Phoebe) Zhang	Portal development	Rutgers University	New Brunswick	USA
* James Wood	Outreach and website development	University of Texas		USA
* Karen Stocks	Data capture development	San Diego Supercomputing Centre	California	USA

1. OBIS GBIF development

OBIS has become an Associate Member of GBIF and will be GBIF's primary marine component. We propose the following mutual path of development:

2002 OBIS represented on multiple GBIF committees
2003 Establish an OBIS GBIF Node
Participate in Electronic Catalogue of Names of Known Organisms
Participate in Digitization of Natural History Collection Data
Participate in data access and infrastructure building
2004 OBIS-developed ecological and fishery information system becomes a GBIF prototype
2005 OBIS-developed genetic information service, linking to GenBank, becomes a GBIF prototype
2006 OBIS-developed environmental data service adopted by GBIF
2007 OBIS-developed geospatial data integration and mining tools adopted by GBIF

OBIS will continue to actively participate in GBIF short-term development efforts, which are likely to focus on data provision and system adaptation. As the only member specializing in marine information, OBIS plays an important role in providing and integrating marine data. In the short term, OBIS is adopting existing products developed/endorsed by GBIF. In this way, OBIS is establishing a strong identity in the GBIF community that will ensure that OBIS programs can be supported through GBIF funds.

2. OBIS GOOS Development

Projection of a mutual development path

GOOS is in its planning phase. At the data level, OBIS should work closely with Ocean.US to improve the usability of the DODS/OpenDAP standard for biological data. OBIS will play a leading role in standard-making in the information component of GOOS, bridging GBIF and GOOS, two prominent "Mega Science" programs. As in GBIF, there are no major players yet in the area of geospatial marine biological data integration and mining in the GOOS community and little has been done in the research community either. OBIS will seize the opportunity and become a leader in these areas.

2002-2003	OBIS working closely with the OCEAN.US Data Management and Communication System (DMACS—the information component of U.S. GOOS) Steering Committee and Subcommittees
2003-2005	Collaborate with Distributed Ocean Data System (DODS) and National Virtual Ocean Data System (NVODS) on data integration Collaborate with NOAA on data discovery standards Supplement GBIF and GOOS standards Collaborate with the Universities Center for Atmospheric Research (UCAR) on Data Archiving
2006	Lead GOOS biogeospatial data modeling/integration development
2007	Lead GOOS biological data mining initiative

OBIS/OpenDAP collaboration initiative

The National Ocean Partnership Program funded projects under two topics in the year 2000: One was to design and implement an infrastructure for the National Virtual Ocean Data System (NVODS). Its underlying framework is DODS, which has evolved into the OpenDAP of today. The other was to establish and populate OBIS. Both OpenDAP and OBIS have made remarkable progress in the last two years, resulting in an unprecedented amount of oceanic data freely accessible over the Internet. Because the systems have focused on different types of ocean data and different aspects of integration and interoperability, it is now time for them to collaborate on system integration; together they can provide the best data service to GOOS. The collaboration effort will:

- enable data to flow freely between systems so the user can access more data.

- develop integration and interoperability so that users can get data and data services that reconstruct, aggregate, and manipulate the data based on users' needs.
- provide the knowledge base for analyzing and modeling the changing marine environment.

3. Partners in Technology Development and Database Publication

OBIS is developing partnerships with several proprietary technologies. Software development partners include EASy and open-source code partners such as DiGIR and DODS.

OBIS will examine outlets for publication of CDs such as the U.N. Atlas of the Oceans and the World Ocean Atlas. CDs may also be an important outreach tool and the most effective way of distributing information to developing countries and schools in which Internet access is limited.

4. Regional OBIS nodes

At present, there is one OBIS portal, although several components of OBIS serve significantly more information from their own databases. A network of OBIS "nodes" may have several strategic and practical advantages in growing and sustaining OBIS.

The most important function of OBIS nodes is to foster the on-line provision of marine biogeographic data from their region and areas of expertise. For example, a European node should actively arrange for data held within European databases to be made available on-line through a European OBIS portal. Each node would thus populate OBIS with data and aid development of on-line data analysis and presentation tools. It would also probably have its own user community, some of which would be more interested in local and regional, than global data sets.

OBIS nodes would be an integral part of OBIS, and nodes would actively support each other's development through provision of advice and technology. OBIS nodes would be independently motivated by the vision to 'publish' quality data on-line, so as to increase the availability of data to researchers and educators around the world in the interest of good science.

Each OBIS node would seek its own regional and national funding independently of other nodes. However, nodes would support each other's applications, emphasizing the added value of additional nodes. It is envisaged that each node will involve two to three staff that may be involved in related projects and activities.

The modest overlap in capabilities and services between nodes would not be redundant, but would provide backup services to the global user community should other nodes malfunction. Thus for example, the same data mapping tools may be available from most of the nodes.

Varying regional circumstances will bring unique features to each node. Some may have particular strengths in certain data types or software resources, or both. Levels of activity and productivity will vary between and within nodes depending on funding levels. Regional and national pride will encourage nodes to at least match other nodes in their state of development.

OBIS nodes will differ in the scope of their data, both geographic and data type. For example, some may focus on governmental fishery or oceanographic data, and others on museum and academic sourced data. Some may cover a geographic region; others selected taxa, and most will probably cover a mix of these data types.

The development of OBIS nodes should be organic. At present, regional nodes are developing in Antarctica, Australia (Southern Pacific Equatorial), Northern Western Hemisphere (Canada), Europe, Indian Ocean, Japan, Southern Pacific (New Zealand) South America, and U.S.A. They will be connected to the central node at Rutgers University, USA. The geographic, taxonomic and tools provided by these nodes will not be restricted to these regions. Cooperation with other countries, especially those with fewer resources to establish independent nodes, is anticipated. It is anticipated that there will be five to eight nodes globally. The number of nodes will be limited by the ability of each node to be confident of being able to secure funding to remain operational in the long-term. Nodes should have arrangements for the maintenance of their databases and software by other nodes, in case their funding becomes insufficient to keep the node up to date.

VII. Cost analysis and funding sources

Approximately $13 M has been committed to OBIS since 2000, mostly from U.S.A. sources (Table 1). These funds have supported the development of 12 OBIS projects. The future of OBIS depends on each projects scientists writing proposals to maintain their support. Each one acts as an individual entrepreneur adding to the databases, conducting quality control and quality assurance, forming new coalitions, fostering software development, developing new Web functions, reaching out to the public, and developing new curricula. Similarly, the Portal and Secretariat are exploring opportunities for joint funding. In addition to Sloan Foundation support, OBIS has received approximately $5 million in grant support.

Table 1. The history of OBIS funding. NOPP = National Oceanographic Partnership Program (USA); NSF = National Science Foundation (USA); Sloan = Alfred P. Sloan Foundation; Lounsbery = Lounsbery Foundation; GeoConnections = GeoConnections Canada.

Period	Budget	Projects	Source funding
2000-2002	$3.7 M	8 OBIS projects	NOPP
2000-2002	$0.1 M	SeamountsOnline	NSF
2000-2002	$0.4 M	OBIS Portal	NSF
2000-2002	$1.73 M	9 CoML Initial Field projects	Sloan
2001-2003	$0.75 M	HMAP	Sloan
2001-2003	$0.15 M	Chesapeake Bay database	Sloan
2001-2003	$0.35 M	Marine Invasion database	Lounsbery
2002-2003	$0.6 M	OBIS Secretariat	Sloan
2002-2003	$0.14 M	SOC Bathypelagic database	Sloan
2002-2004	$0.9 M	CephBase for curriculum development	NSF
2002-2006	$1.9 M	for marine mammals, turtles, and birds	Sloan, NOPP
2003-2004	$0.0? M	Eastern Mediterranean and Black Sea database	Niarchos
2003-2005	$0.2 M	Atlantic Canada on-line database	GeoConnections
2003-2005	$0.9 M	HMAP	Sloan
2003-2006	$0.5 M	OBIS Portal develoment	NSF

The contributions of OBIS will approximately double in the next two years (Figure 2). A minimal level of maintenance averages about $0.2 million per year per Member. With continued growth, plus the cost of the Secretariat and Portal, the annual cost of maintaining OBIS in five years will be about $10 million per year. For OBIS to be successful, a large part of this support must come from government agencies responsible for the development of national information infrastructures and ocean observing systems. OBIS will provide an international context for national databases and enable government agencies to achieve their missions more effectively. Biological data at the species level are an important component of these systems and OBIS is needed to provide authoritative biological information, create and enforce standards and protocols, and to assist in setting priorities.

As the international custodian and developer of new ocean biological information infrastructure OBIS will seek funds, in coordination with GBIF and GOOS, to:

- develop basic species-referenced and geo-referenced databases
- serve as international authorities for marine taxa
- establish and enforce standards and protocols for marine biological databases
- link marine biological subsystems through a central Portal
- provide vision and focus for international cooperation
- set research priorities for developing information infrastructure
- coordinate initiatives for funding from a broad spectrum of government agencies

Several NSF (U.S.A.) programs are potential sources of support for OBIS:

- Information Technology Research,

- Biodiversity Surveys and Inventories,
- Digital Government,
- Digital Libraries,
- education, and
- areas of research and education requiring biological database components such as ORION.

Potential NSF sources could support a significant portion of OBIS research. The European Commission Union 6th Framework Program for research is another likely source of support along with national initiatives in several European countries. CoML Projects in each of the countries with National CoML Programs (Australia, Canada, Europe, Japan, U.S.A., South America, Southern Africa) will also contribute to OBIS.

The data management component of the Global Ocean Observing System (GOOS) could become a major source of support for OBIS. The Data Management and Communications System (DMACS) component of U.S. GOOS is expected to have an initial budget of $18 M or about 13% of the national initiative. Nearly half of this will be for systems integration and interoperability through international, national, and regional pilot projects. One such project will be integration and interoperability of OBIS with the rest of the system.

The formation of a comprehensive business plan for OBIS is an ongoing important task for OBIS. OBIS will serve both providers and users of biological knowledge about the oceans. Principal clients will include:

- Research Organizations
- Fisheries agencies
- Environmental protection agencies
- Navies
- Database infrastructure organizations
- Museums
- Libraries
- Marine protected areas
- Conservation and environmental organizations
- Precollegiate, undergraduate, and public education
- International standards organizations
- Industries (biotechnology, pharmaceuticals, transportation, resource extraction, software developers, publishers, education services, etc.)
- University researchers

OBIS will need to raise approximately $100 million to become a fully developed and sustainable international program by 2012. The present success of OBIS and immediate opportunities indicate that it is on track to accomplish this goal (Figure 3).

Appendix A.
OBIS Mission and Goals

OBIS is an on-line, open-access, globally-distributed network of systematic, ecological, and environmental information systems. Collectively, these systems operate as a dynamic, global digital atlas to communicate biological information about the ocean and serve as a platform for further study of biogeographical relationships in the marine environment. Emphasis is on accurately-identified, species-level, geo-referenced abundance data. Through use of Internet-enabled GIS and other Web-based analytical tools, biological data can readily be integrated with environmental data, maps, visualizations, and model outputs for a broadly based community of users.

OBIS, and its diverse community of researchers, will elicit the needs and concerns of user groups, including resource managers, navies, industries, and environmental and educational groups, to balance their needs in a utility of broad value. OBIS' primary concern will be to make basic data sets accessible and interoperable. In most cases, it is expected that additional value, in terms of customization, advanced visualization, and applications, will be carried out by, or nearer to, the end users.

OBIS will build coalitions with national and international database systems 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
- speed the dissemination of and public access to ocean biogeographic information while appropriately addressing intellectual property rights issues.

OBIS is developing as both a major international research program and a federation of databases made interoperable through the OBIS Portal maintained by the Secretariat.

OBIS will:

- achieve more uniform and higher quality standards for marine biological data (data will be presented using consistent nomenclature in formats conducive to analysis and comparison)
- bring relevant geo-referenced and species-referenced biological and environmental data together in an interoperable system
- bring the marine biology community together through coalitions among existing regional, national, and international information systems and research programs
- take the lead in species-relevant database development (including linkages to genetic databases)
- develop a procedure for timely provision of synoptic environmental datasets and map products for the analysis of biological and environmental data at multiple spatial and temporal scales to reveal ecological and biogeographic patterns
- support the objectives and principles of the Global Biodiversity Information System (GBIF)
- provide, through the OBIS Portal, interoperability, analytical tools, and access to all components of the system

Membership in the OBIS federation is open to any interested individual, country, or organization committed to the long-term maintenance of an accessible, relevant, biological database. Present members of the federation include CoML, NOPP-funded OBIS programs, and national databases interested in developing ties with the OBIS international system of databases. The OBIS portal is responsible for making the entire system interoperable and each member of the Federation will, in addition to maintaining their own database systems, be committed to provide data through the OBIS Portal and assist in its support.

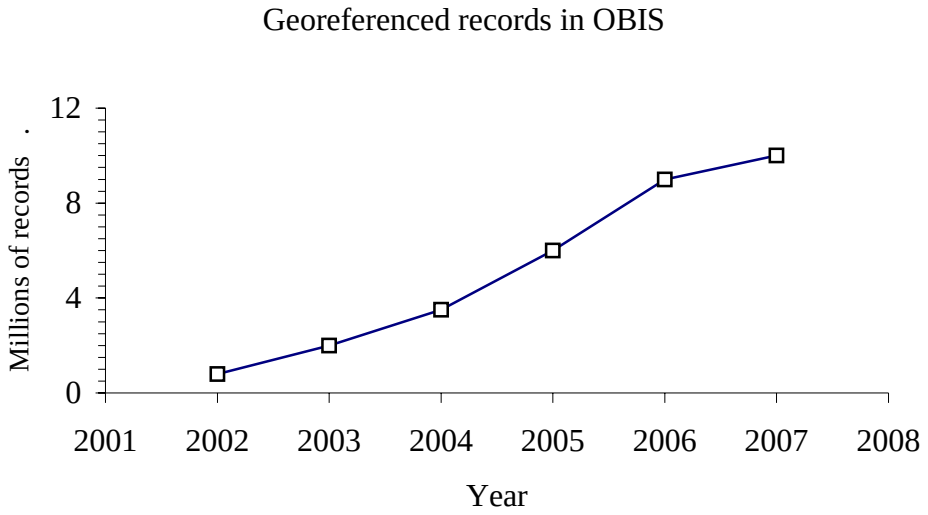

Figure 1. The number of georeferenced records of marine species available through the OBIS portal, or through OBIS related websites that are in the process of connecting to the portal.

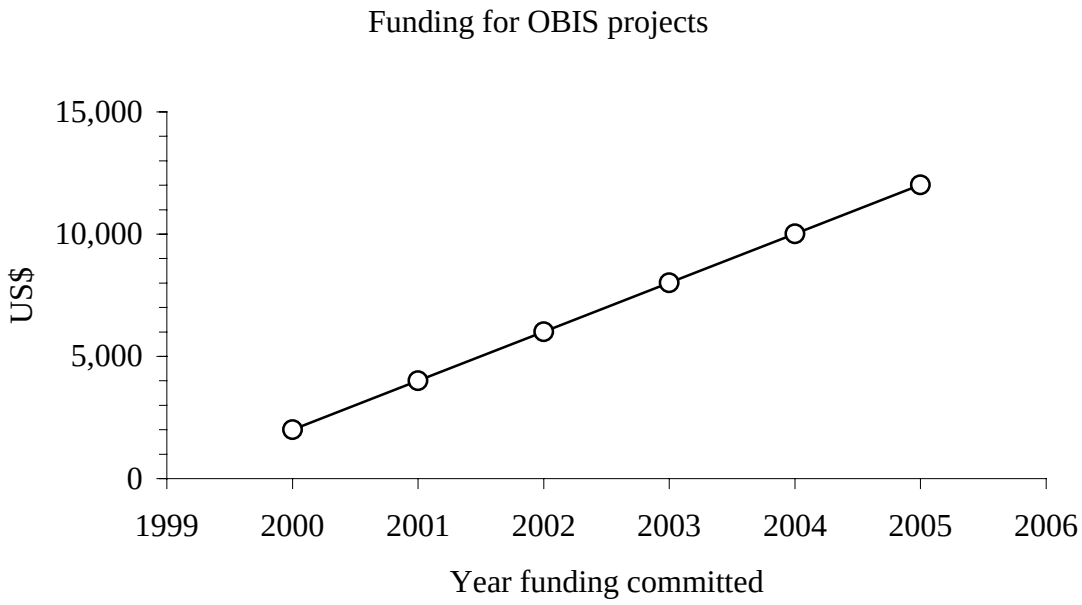

Figure 2. The accumulated funding for OBIS related projects. Details in text.

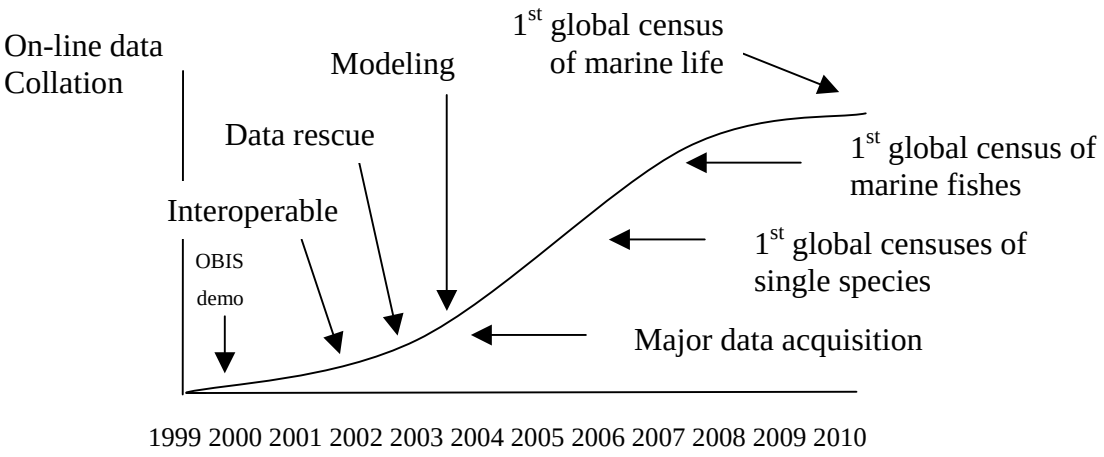

Figure 3. A diagrammatic presentation of activities required to produce a census of marine life in relation to the rate of data acquisition by OBIS.