

Citation:

Costello M.J., Stocks K., Zhang Y., Grassle J.F., Fautin D.G. (March 2007). About the Ocean Biogeographic Information System. First Published on www.iobis.org in April 2007.
url

Preamble

This document is an archive of the text published on the website of the Ocean Biogeographic Information System in April 2007.

Using OBIS	4
What is the aim of this website?	4
What are the benefits of OBIS?	4
What can I do with OBIS at present?.....	5
What improvements is OBIS planning?	5
How can students use OBIS for learning?	5
Who will use OBIS?	5
How do I cite data published through OBIS?	6
What are the limitations of data in OBIS?	7
How can I search and explore data?.....	9
How can I download data?.....	9
Who is using OBIS?.....	10
About the data	11
Where does OBIS get its data?	11
What kinds of data is OBIS interested in?	11
What quality control system is in place for data?	11
Will OBIS pay me for my dataset?	13
Who has intellectual property rights to the data?.....	13
How will my data be cited and credited in OBIS?.....	13
How do I contribute data to OBIS?.....	13
What if I have sensitive data which should not be openly accessible?.....	13
If I want to contribute data to the OBIS system, what should I do next?	13
Data Schema and metadata.....	14
What is the OBIS data schema?	14
What is DiGIR?.....	14
Is the OBIS Data Schema compatible with the Darwin Core standard?.....	14
Isn't it hard to implement all the OBIS Data Schema fields?	14
Do I have to use the same field names in my database as the OBIS Schema?	14
Using an OBIS view	15
Why do you have fields for places that may not apply in the ocean?.....	15
What do you mean by the terms "Collected" and "Observed"?.....	15
OBIS Schema data types.....	15
Can the Schema accommodate tagging data, or multiple sightings of a single individual animal?	15
What is metadata?	15
About databasing	17
How do I start designing a database?.....	17
Do I need to use a relational database?	17
Which relational database should I use, if I use a relational database?	17
What hardware and software do I need to do to serve data through the web?.....	17
OBIS organisation.....	19
What is OBIS?	19
Who can join OBIS?	19
How is OBIS governed?	19
What is the OBIS International Committee?	21
Who is on the OBIS International Committee?	21
What is the OBIS Science (Editorial) Board?.....	23
What is the OBIS Portal?.....	24
What are Regional OBIS Nodes (RON)?	24

What is the OBIS Manager’s Committee?.....	24
What was the OBIS Technical Committee?	26
What are OBIS Scientific Working Groups?.....	26
What is an OBIS contributor?.....	26
What do all the acronyms mean?	26
Who has contributed to OBIS?	27
People.....	27
Data Sources	27
Data Providers.....	27
Tool providers.....	27
Partners.....	27
Sponsors	29
What were the origins of OBIS?.....	29
What has been published related to OBIS?.....	29
How do I cite this web page description of OBIS?.....	29

Using OBIS

What is the aim of this website?

This website provides a portal or gateway to many datasets containing information on where and when marine species have been recorded. The datasets are integrated so you can search them all seamlessly by species name, higher taxonomic level, geographic area, depth, and time; and then map and find environmental data related to the locations (see ‘What can I do with OBIS?’).

This website aims to be (a) authoritative, professional, and credible, (b) concise, easily read and understood, and use minimal jargon, (c) user friendly with a logical consistent design and flow, and (d) regularly updated. It should also (e) have as few clicks as possible, and (f) we should be easily contacted (*e.g.* by email, or see [meetings](#) where OBIS representatives can be met). Please let us know how to improve by emailing [OBISsupport\(at\)marine.rutgers.edu](mailto:OBISsupport@marine.rutgers.edu). We are working to make OBIS as comprehensive as possible (see “What are the limitations of data..?”). .

What are the benefits of OBIS?

The immediate benefit of OBIS is that it publishes data on the spatial and temporal distribution of marine species, with online data exploration software, at no charge to anybody with access to the World Wide Web. The ease of rapid data access then benefits education, research and management of marine biodiversity and ecosystems; and through this use, society in general. OBIS thus saves people and organisations needing primary data on marine species locations both time and money, and fosters new syntheses of data across geographic, taxonomic, and time scales. OBIS can also assist where data generators have a legal or contractual obligation to make data publicly available.

OBIS enables data to be discovered that are otherwise unpublished or difficult to access. OBIS data has scientific authority, in being collected and managed by reputable scientists and institutions. As with paper publications, the publication of data through OBIS increases the visibility of the productivity of these scientists and their organisations.

Participation in OBIS involves working with the latest standards in biodiversity informatics data management. The practice of adopting standards can help data management within organisations and research teams, as well as facilitating data publication.

OBIS keeps data alive and available by publishing it online. OBIS will host data should any data provider be unable to continue to host it in their own system. In addition to having full back-up copies of the data, OBIS is creating several mirror sites so its data will be replicated in different parts of the world.

What can I do with OBIS at present?

From the OBIS portal (the first website page connecting to the data), you can

- (a) search for where a marine genus and/or species is recorded in the data published through OBIS;
- (b) download the data published in OBIS for any species, including location, depth, date and time collected, source datasets, and verified taxonomic name information;
- (c) plot species locations on a range of flat and spherical views of the world, including polar views, using the C-Squares Mapper;
- (d) plot species against background maps of sea temperature, depth, and salinity using the KGS Mapper;
- (e) use environmental data for the locations of these data to predict the species potential range on the KGS Mapper;
- (f) explore relationships between species and environmental data on KGS Mapper to see which parameters best explain a species distribution;
- (g) browse down a taxonomic hierarchy to get lists of all species in OBIS for a phylum, class, order of higher taxonomic group;
- (h) plot maps of all data at a higher taxonomic level;
- (i) search for lists of species recorded in OBIS by country (Exclusive Economic Zone), sea or ocean, Large Marine Ecosystems, FAO and ICES fishery areas, Longhurst's pelagic regions, depth, date, and by entering latitude – longitude coordinates;
- (j) connect to other sources of information on the species, including genetic data, published literature, and images.

What improvements is OBIS planning?

OBIS wishes to

- (1) continue to publish more data and improve data quality. In particular, it intends to publish at least one location for every known marine species before 2010;
- (2) have online tools to show individual animal tracks across the oceans (*e.g.* satellite tagged turtles);
- (3) have online tools to show time-series and depth patterns of species distributions;
- (4) allow users to select different species for mapping together, and improve online editing of data;
- (5) allow users to search OBIS by habitat;
- (6) have maps of the distribution ranges for all species in OBIS;
- (7) expand the OBIS data schema to allow use of abundance data;
- (8) add geographic areas to the “Advanced Search” options of OBIS, so users can search by genus and/or species, time period, depth range, latitude-longitude coordinates, and geographic area.

How can students use OBIS for learning?

OBIS is a source of data and information, and provides links to other authoritative resources. It also allows students to develop skills in using online resources, discriminating between authoritative and less scholarly sources, learn about the strengths and weaknesses of using primary data that are only a sample of the potential data, and using this to infer the distributions of species. Comparison with secondary sources of information is encouraged. An example of one exercise used for several years can be downloaded [here](#).

Who will use OBIS?

The following communities are anticipated to use the OBIS portal (in no order of priority): researchers, fishery scientists and managers, policy maker, educators, amateur naturalists, environmental NGO, consultants, nature conservation organisations, and students.

How do I cite data published through OBIS?

In using data, software, or other information published through OBIS, for example in a publication or presentation of any sort based wholly or in part on material so accessed, OBIS asks that users respect the scientific conventions of citing the data sources. Such citation strengthens the authority of such data analysis and provides due recognition to the data providers.

Users should acknowledge the use of specific records from contributing databases in the form appearing in the 'Citation' field of that dataset and acknowledge its publisher as being OBIS, *i.e.*

Author (or Editor) name. Database title. Retrieved [date accessed] from www.iobis.org

Examples:

(a) Single level

Stocks, K. SeamountsOnline: an online information system for seamount biology. Version 3.1. Retrieved [date] from www.iobis.org.

(b) multiple level (for example, to cite a database published on a Compact Disc and accessed through OBIS)

Picton, B.E., Emblow, C.S., Morrow, C.C., Sides, E.M., Tierney, P., McGrath, D., McGeough, G., McCrea, M., Dinneen, P., Falvey, J., Dempsey, S., Dowse, J. and Costello, M. J. 1999. Marine sites, habitats and species data collected during the BioMar survey of Ireland. In: Picton, B.E. and Costello M. J. (eds), *The BioMar biotope viewer: a guide to marine habitats, fauna and flora in Britain and Ireland*, Environmental Sciences Unit, Trinity College, Dublin. Retrieved [date] from www.iobis.org.

Note that if you use a lot of data, you will have to cite many datasets, just as if you used data from many papers published in a science journal.

A citation for this website text is [here](#).

For information purposes, please provide to OBIS (obissupport@marine.rutgers.edu) the full citation of any publication (printed or electronic) that uses data published through OBIS. This will be listed on the OBIS website [publications list](#) and will thus increase awareness of the study.

What are the limitations of data in OBIS?

OBIS is comparable to a scientific journal that makes data freely available on the internet. Thus the geographic and taxonomic scope, and quantity of data provided, depend on the scientists and organizations that provide data. However, in contrast to data in a journal, the 'reader' can select and combine data in OBIS from a variety of sources. OBIS and its users give feedback on data quality and possible errors to data providers. Because data providers are willing to correct errors, the quality of the data will increase in time. How OBIS provides quality assurance, who is primarily responsible for data published in OBIS (its owners), issues to be considered in using the data, and known gaps in the data, are described below. OBIS also has technical quality control procedures.

Quality assurance

Only data from authoritative scientists and science organizations approved by OBIS are served. All data are subject to quality control procedures before publication, and at regular intervals, with data providers informed of any discrepancies and potential errors (*e.g.* species names spelt incorrectly, mapping errors). OBIS also benefits from user peer-review and feedback to identify technical, geographic, and taxonomic errors in data served. However, although errors will exist as they do in any publication, OBIS is confident that the data are the best available in electronic form. That said, the user needs sufficient knowledge to judge the appropriate use of the data, *i.e.* for what purpose it is fit.

Many of the data published through OBIS have voucher specimens in institutional collections and museums, images of observations, and the original identifier of the specimens is often credited or will be contactable from the data custodian.

Data ownership

Data providers retain ownership of the data provided. OBIS does not own or control or limit the use of any data or products accessible through its website. Accordingly, it does not take responsibility for the quality of such data or products, or the use that people may make of them.

Data use

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

Users should be aware that OBIS is a gateway to a system of databases distributed around the world. More information on OBIS data is available from the data sources websites and contact persons. Users should email any questions concerning OBIS data or tools (*e.g.* maps) to the appropriate contact person and copy this request to OBIS at <obissupport(at)marine.rutgers.edu>.

Data gaps

Major gaps in data and knowledge about the oceans are reflected in OBIS' data coverage. Note the following:

1. Most of the planet is more than 1 km under water: this deep sea is the least surveyed part of our world.
2. Coastal areas have been adequately sampled only for the distribution of most vertebrates (birds, mammals, reptiles, larger fish).
3. The oceans have been better sampled in the northern than the southern hemisphere, as reflected in the distribution of data in OBIS.

4. Most marine species have not yet been recognized or named. A major effort is required to describe marine species, especially invertebrates and deep-sea organisms.
5. Of the marine species that have been described, some have been discovered to be several species, and others combined into single species. Thus, there are changes in the application of species names over time. A checklist of all current marine species names is not available but it is estimated that 230,000 have been described. Only about half of these names have been organized into global species checklists. OBIS includes distribution data on (a) many of these validated names and (b) additional names that remain to be organized into global species checklists. Thus, OBIS has some distribution data for about one third of the known marine species.
6. Some species distribution data are not available in any form, as they have not have been published nor made available for databases.
7. Only some of the recently collected, and less of the older published, data have been entered into databases. Thus databases are incomplete.
8. Of existing databases, many are not connected to OBIS.

You can help address these data gaps by (a) recognizing and encouraging scientists and organizations to make their data available online so they are accessible to OBIS, and (b) advocating for and carrying out field surveys and taxonomic studies designed to fill geographic and taxonomic gaps in knowledge.

How can I search and explore data?

At present, you can search OBIS by a genus and species name, higher taxonomic group (*e.g.* whales, fish, anemones), and geographic area. Some of the mapping tools enable editing of the data.

From the first web page you can (1) enter a genus, species or common name in the search box, or (2) click on an area of the map to get a list of all species in that area. Note you can change the size of the map area searched from 5, to 1 or 0.5 degrees as you wish. The area being searched is shown by a box over the area clicked on the map. Note that common names are only available for vertebrates and a few other species and are not consistently used between countries. Also on the first page are buttons to go to "Advanced Search" and "Browse by taxonomic groups" options.

At the top of all other pages except the first, is an option to search for (a) a species by genus, full species or common name ("name search"), and (b) all species in certain geographic areas. These areas are countries Exclusive Economic Zones (EEZ), International Hydrographic Office (IHO) sea and ocean areas, Food and Agricultural Organisation (FAO) fishery areas, International Council for the Exploration of the Sea (ICES) areas, Large Marine Ecosystems (LME), and Longhurst's pelagic regions.

From the Advanced Search page you can refine your search by entering either genus and/or species name, (3) dates, (4) depth range, (5) latitude - longitude coordinates, and/or (6) dataset. To get all data from one dataset, just leave other search options empty and select that dataset.

If you do not know a genus or species name, you can browse down a hierarchy of taxonomic groups (*e.g.* from vertebrate to fish to cod). Note you can also request a map of all locations of records for a higher taxonomic group, such as all whales.

The above searches will provide a page with a list of one or more species names. Note that some names may be spelling variants or synonyms, and that the use of some names may have changed over time. We are planning to add notes on such issues. Click on the species names to produce a "Species Results" page. This shows (a) a map of locations with data in OBIS, (b) several links to more information about the species, including literature, images, and genetic information, (c) a list of the source datasets is at the bottom of the page, (d) buttons to click to view and download the data in .html or .txt formats, and (e) two additional mapping tools.

The KGS Mapper will match the species locations against 52 ocean environmental variables and use this to show where else these conditions have occurred. You can select which environmental variable you feel is most appropriate, and edit them if you have more precise data or you can see that spatial resolution issues may have generated incorrect values. More advanced data editing is also possible.

The ACON Mapper shows data abundance using pie charts on maps, allows editing of data mapped, and can plot maps by the years the data was collected. A range of data editing and download options are available.

All maps can be saved to your desktop by right clicking your PC mouse and clicking "save image".

Note that larger datasets will take longer to search, explore, and map. Consider limiting the amount of data being explored by using Advanced Search to select a limited time, depth or range of datasets. For researchers wanting large amounts of data please do some exploratory work and then contact obissupport@marine.rutgers.edu to email it to you.

How can I download data?

You first need to search OBIS by species, higher taxon, geographic area and/or other options (see above) to get a "Species Results" page. Below the map and mapping options you can view and download data your previously selected in .html or .txt formats.

Who is using OBIS?

In the month of February 2007 OBIS received about over 1 million hits from 50,000 visitors in at least 26 countries. About half of users came from educational, and at least 10% from commercial, institutions. Detailed web statistics and history are available [here](#).

About the data

Where does OBIS get its data?

OBIS publishes data on behalf of scientists from government agencies, museums, universities, commercial companies, and non-governmental organisations. OBIS is always seeking new contributors.

What kinds of data is OBIS interested in?

OBIS is a marine biogeographic information system, meaning that we concentrate on datasets that record particular species (or higher taxonomic group) from particular marine locations, at particular times. At present, we can only publish data where the locations are recorded as latitude and longitude, not as place names. Our focus is on high taxonomic quality, so datasets where organisms have been identified by professional or trained biologists are our priority. In the near future, we will be expanding to take in environmental datasets (*i.e.* coverage of physical, chemical, and geological parameters) that are relevant to understanding the distribution of species. So we are interested in hearing from potential contributors of these datasets, and welcome your contact, but are still in the process building this facility.

What quality control system is in place for data?

Data published through OBIS must come from credible, authoritative sources. The scientists and institutions responsible for collecting and managing the data are clearly named. Before publication, the data must pass through a series of technical controls described below, and these are repeated every time the data may be crawled again from its source. Any errors, such as species name misspellings, names not recognised in OBIS, and possible mapping errors, are reported to the data provider to review, and if necessary, correct. Thus the next time the data are published they are more correct, and the source database quality is also improved. Data use is a very important way of finding actual and possible errors in data. Users may contact the data source directly or OBIS with such issues.

The OBIS Quality Control protocol is as follows:

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

detailing number of records obtained and missing records if applicable, time of next crawling, and any errors identified.

Will OBIS pay me for my dataset?

No, OBIS does not buy data. It is a group of contributors who have agreed to publish their data through a central portal to make it more accessible. However, we may be able to make suggestions for places where you could submit a proposal to fund developing datasets or for digitizing existing datasets.

Who has intellectual property rights to the data?

OBIS claims no ownership nor rights to the data sets it publishes. All rights remain with the data source, whom may at any time decide to remove their data from OBIS. This is true whether you serve the data yourself, or whether you place your dataset at a Regional OBIS Node or the central OBIS portal for serving.

How will my data be cited and credited in OBIS?

All data published through OBIS are labelled with the organization and database from which the data came, and a standard citation is provided. Users are expected to cite the data providers when using data from OBIS as they would cite papers from a conventional print publication (see How do I cite data ...?).

How do I contribute data to OBIS?

There are two models for sharing data through the OBIS system:

1. You become a distributed data contributor. This means that you keep your dataset locally, and set up a server that can respond to OBIS queries. This requires "mapping" your dataset to the OBIS schema (which is not as hard as it sounds!) and installing a free software package called DiGIR to communicate with the portal. There are more details on this below.
2. You provide your data set in electronic form to a Regional OBIS Node, the central data portal, or another existing Data Provider, and it is published from there.

Which choice is right for you depends on whether you are interested in maintaining your own server, and also whether you expect to be making regular updates to the data set. OBIS prefers groups to be distributed data contributors, because we think it is best for the data contributor, who knows the dataset best, to maintain it. That way you can add data and make corrections directly. But if you cannot or do not wish to set up a server, OBIS is happy to host data. In either case the data will be credited to you.

What if I have sensitive data which should not be openly accessible?

The short answer is that OBIS is committed to free, open, public access to data, so if you have sensitive data you probably don't want to publish it through OBIS (or any other publication). However, there are some particular concerns we may be able to help with. For example, if you are concerned about giving the precise location of a rare or commercial species, then we may be able to represent your data at a lower spatial resolution, or to give a bounding box instead of a point location. If you have data that you would like to publish but want to wait until its analysis is published elsewhere, we can help you set up your dataset appropriately now, but agree not to publish it for a certain amount of time.

If I want to contribute data to the OBIS system, what should I do next?

Please read the information about OBIS on this website, and contact the Executive Director or Chair of the OBIS International Committee.

Data Schema and metadata

What is the OBIS data schema?

The OBIS schema is a list of data fields with names, descriptions, and format notes.. It is an extension to the Darwin Core Version 2 standard.. When the OBIS portal sends queries out to its distributed data contributors, the portal will request data using these fields and needs to have data returned using these fields. The DiGIR software provides the programming to turn an OBIS query into a search on your particular database, but in order to install DiGIR you need to "map" the OBIS schema fields to the fields in your database. Download the following files for details.

- [The OBIS Schema General Questions](#)
- [The OBIS Schema Version 1.1: Definition of the Data Standard](#)
- [Implementation Notes](#)
- [OBIS Schema 1.1 \(xsd file\)](#)

What is DiGIR?

DiGIR is the software through which OBIS communicates with its distributed data contributors – defining how data is exchanged. When a user of the OBIS portal inputs in a query (such as 'show me all the locations where the fish, *Beryx splendens*, has been found'), DiGIR allows the portal to send that query to the data contributors, for the data contributors to translate that query into a search on their local database, and to send the data back to OBIS. More information on DiGIR can be found on digir.sourceforge.net. Please contact OBIS before installing the DiGIR software to ensure that you are using a compatible version and have the OBIS configuration details. You can download full details here: [OBIS DiGIR Installation Software and Instructions](#)

Is the OBIS Data Schema compatible with the Darwin Core standard?

Yes. The OBIS Data Schema was built as an extension to the Darwin Core version 2 standard (<http://speciesanalyst.net/docs/dwc/index.html>). The Darwin Core is a standard that is used by the Global Biodiversity Information Facility and others. The OBIS Schema also contains some additional fields for holding information that the Darwin Core does not handle. So, if you implement the OBIS Schema in your database, you will be compliant with both the Darwin Core standard and the OBIS standard.

Isn't it hard to implement all the OBIS Data Schema fields?

Think of the OBIS schema as a menu of options. There are only four fields that are required in order to be compatible (latitude, longitude, taxonomic name, and date/time of last modification). For all of the other fields you only need to include them if you want to have that information in your database. If you don't plan to hold a particular type of information, you can leave it out of your database. However, if you do include a type of information covered by a field in the OBIS schema, you should represent it as described in the OBIS schema.

Do I have to use the same field names in my database as the OBIS Schema?

No, you can use any field names you like. When you implement DiGIR, it will ask you to tell it which fields relate to which OBIS Schema field. You should keep track of this and make sure that there is a one-to-one mapping of fields in the OBIS schema to your database and that you use the required format for the field. For example, because the OBIS Schema has separate fields for day, month, and year of the record, it is best to hold these in three separate fields and not in a single date field (or at least have a plan for how to separate the pieces for serving to OBIS).

Note that most database software will allow you to do automatic operations on fields. You may prefer to enter your location information as degrees and minutes for latitude and longitude instead of decimal degrees. That's fine, because it will be easy for you to create a database view with a latitude field in decimal degrees calculated from (latitude degrees + (latitude minutes/60)).

Using an OBIS view

It may be easiest for you to implement DiGIR by creating a "view" or query in your database that has all the OBIS schema fields in one table. You may have separate tables for "species names", "observations" etc., but may create one virtual view that does all the joins necessary for the OBIS query. You can also have it do any reformatting (such as the latitude degrees and minutes calculation mentioned in the above question) required. Then it will be easy during DiGIR installation to map onto the OBIS schema.

Why do you have fields for places that may not apply in the ocean?

Because these are Darwin Core fields. To be compliant with the Darwin Core, OBIS must allow all Darwin Core fields to be entered. But remember, these are all optional fields. Don't even put them in your database if you don't need them - most OBIS datasets won't. OBIS operates off latitude and longitude locations, which are why these two are the only required locality information.

What do you mean by the terms "Collected" and "Observed"?

The OBIS databases hold information on the locations where different species have been found. The act of finding a species at a place is called a "collection" or an "observation" throughout the schema documentation. This term is meant to apply very broadly, and includes cases where species were literally seen during a visual search, were collected in a sample of any kind (research survey, fisheries catch date, etc.), where a specimen in a museum indicates the location where it is from, etc.

OBIS Schema data types

The schema indicates the data type for each field. These are general categories, and your particular database software may use different terms. Where there are additional restrictions placed on the data format, this will be indicated in the Description.

Can the Schema accommodate tagging data, or multiple sightings of a single individual animal?

Yes, the OBIS Schema can accommodate data from multiple observations of a single individual organism, such as data produced by tagging studies. To implement this, the user should 1) create a record in their database for the individual and use the Catalog Number as a unique identifier for that individual; and 2) each observation for that individual should be entered as a separate record in the database and tied to the individual by setting the Related Catalog Item field equal to the Institution Code, Collection Code, and Catalog Number of the individual record, and the Relationship Type field to 'point observation for tagged animal of.' (see the Technical Resources page for more on the OBIS Schema fields).

What is metadata?

Metadata are information about data records. Some metadata are included in the OBIS data schema, and describe features of individual records (*e.g.* sampling method). Other "discovery" metadata describe a dataset. OBIS requires data providers to provide Discovery metadata in advance of data publication to allow comparison of the data published with what was anticipated, and to be aware of how planned data publication will contribute to OBIS. A standard citation is important to enable users to cite the dataset correctly. Other metadata, such as sampling, taxonomic, geographic, and habitat information enables OBIS to identify data gaps, and further data exploration features, such as a filter that would allow users to select datasets with planktonic data.

OBIS discovery metadata fields are: Dataset name, Citation (so data users can cite the dataset in standard format of author (or editors), title (descriptive), host institution), Taxonomic coverage,

Geographic coverage, Temporal coverage, Habitat coverage, Total distribution records, Total number of taxa, Collection method, Data source, Abstract (describes dataset), Scientific Contact (responsible for data collection and accuracy), Technical contact (responsible for data management), Website, Date this form completed, Publications from this data (so users can read these for more details about the data origins and its uses). Where standards exist for metadata these fields comply with them, but no standards yet exist for taxonomic, habitat and sampling metadata. OBIS is collaborating with international initiatives (including MEDI of IODE) to develop marine ecological metadata standards that build on ISO, FGDC, GCMD and others. Collaboration between OBIS and the United States National Aeronautics and Space Administration (NASA) Global Change Master Directory (GCMD) has developed the OBIS Master Directory at GCMD: <http://gcmd.gsfc.nasa.gov/KeywordSearch/Home.do?Portal=OBIS&MetadataType=0>

About databasing

How do I start designing a database?

First, figure out what information you have or plan to have. If you already have datasets, either in electronic or paper formats, look at the data that are included. Make a list of these fields (*i.e.* the column headings in your data table). Then go through the OBIS schema. If there is one or more fields in the OBIS schema that cover information that you want to hold, then use that field name and the suggested format. If there are one or more fields in the OBIS schema that do not apply to your data, just leave them out. If there is additional information that you want to keep in your database that is not covered by the OBIS schema, then you can add additional fields.

Do I need to use a relational database?

A relational database is a class of software that allows you to hold data in linked tables. You do not need to use a relational database, as you can hold your information in a "flat file" such as an Excel worksheet. However, a relational database offers great advantages. First, it lets you enter information more efficiently. For example, you can enter a scientific name once into a table of names. Then for every record or observation you have for that species, you won't need to type that name in again, you can just pick it from a drop-down list. So it is faster and you don't have to worry about making typing or spelling mistakes. In addition, relational databases can be queried in more complex ways than a spreadsheet. For example, you can ask for "all the records for species X that were caught north of 30° north, shallower than 300m, and between 1980 and 1985."

A note on text files. Delimited text files are good for archiving data (saving it in such a way that someone years from now will probably still be able to get to it), but not very good if you actually want to do things with the dataset, like extract certain data of interest, update it, or serve it.

Which relational database should I use, if I use a relational database?

Several products are available, and many are quite similar and have similar functionality, so this isn't a critical decision. Microsoft Access is common desk-top commercial software packages. Microsoft SQL Server, Oracle, Sybase and PostgreSQL are common "industrial-strength" databases. This means that they are designed to be efficient with large volumes of data. Generally, if you expect a dataset with hundreds of thousands of records, then you should consider one of the industrial databases, or one of several available free-ware solutions.. If you have tens of thousands of records or fewer, then Access should be fine. The trade-off is that the industrial packages tend to cost more and also be a little less user-friendly. PostgreSQL and MySQL are free, open-source relational database packages that are quite good. You can find more information on MySQL at <http://www.mysql.com> (note that companies may sell packages that include extra documentation, etc., but the core software is free) and PostgreSQL <http://www.postgresql.org/>. PostgreSQL is more powerful than MySQL; especially its geographic features which are why many people prefer it to MySQL. But power comes at the expense of more complexity. Whichever one you choose, just make sure it is "ODBC compliant" – this means that it can communicate with other sources (for exporting, for serving data, etc.). Most of the relational database packages are ODBC complaint, but some "home-grown" systems are not.

What hardware and software do I need to do to serve data through the web?

If you want to publish data directly to OBIS, you will need a computer with an operating system that has your database, server software, and DiGIR installed. If you would like to set up your own web page, you will also have to program to create your web pages and the search functions users will access. HTML is the language that web pages are built in; several languages such as Perl/CGI can be used to create search forms for users to enter data into, and these will need to include the SQL or other commands that actually search your database. Software such as Dreamweaver helps to

make programming web pages easier. For those on a budget, there are some good freeware options. Linux is a free operating system, Apache is a server, and MySQL and PostgreSQL are free relational databases.

OBIS organisation

What is OBIS?

OBIS was established by the Census of Marine Life program (www.coml.org). It is an evolving strategic alliance of people and organizations sharing a vision to make marine biogeographic data, from all over the world, freely available over the World Wide Web. It is not a project or program, and is not limited to data from CoML-related projects. Any organization, consortium, project or individual may contribute to OBIS. OBIS provides, on an 'open access' basis through the World Wide Web:

- taxonomically and geographically resolved data on marine life and the ocean environment;
- interoperability with similar databases;
- software tools for data exploration and analysis.

OBIS was one of the earliest Associate Members of the Global Biodiversity Information Facility (www.gbif.org) which publishes data on all species. OBIS is a very active participant in GBIF activities, and one of the largest publishers of data to GBIF, reflecting its role as a specialist network for marine species. GBIF recommends that marine data are first published through OBIS, because OBIS can add special value (*e.g.* depth) and will manage the subsequent publication of data through GBIF. This also avoids duplication of data being separately published to GBIF and OBIS.

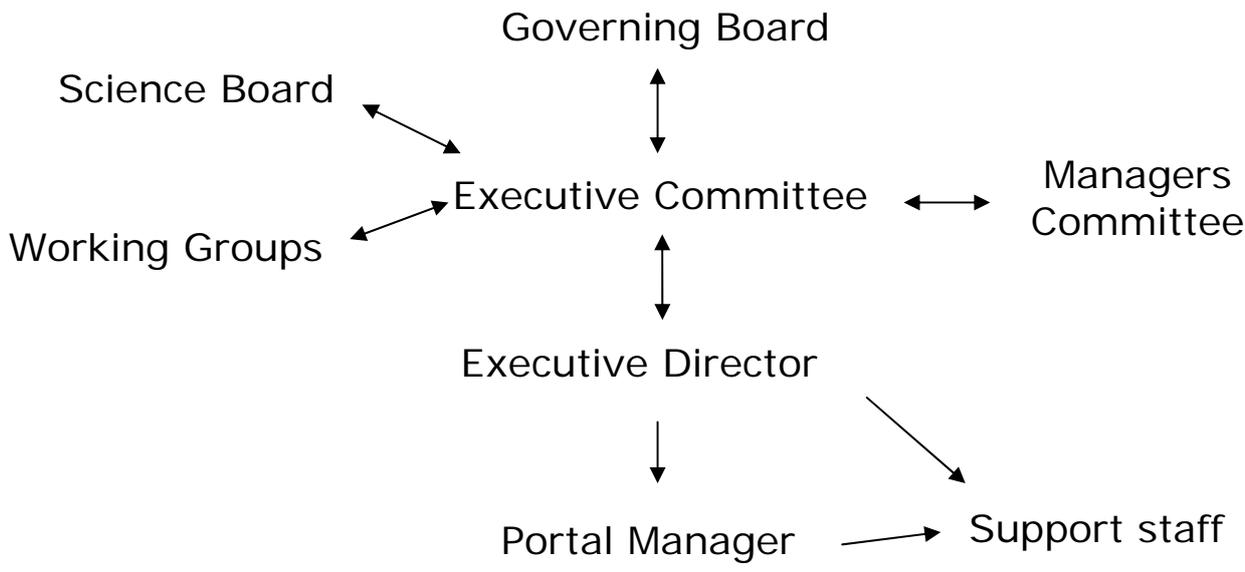
OBIS does not intend to be a special resource for molecular or environmental data, species names, literature, or all information on marine species. However, it does link out to authoritative resources in these and other areas, and can provide a species distribution data service to other species information systems.

Who can join OBIS?

Any organization, consortium, project or individual may contribute to OBIS. Contributions typically take the form of publishing data through OBIS, or providing software tools. The data may be published through an on-line connection from another database or be located on the OBIS server at a Regional OBIS Node or at Rutgers, The State University of New Jersey. To join OBIS contact the Chair of the OBIS International Committee or the Executive Director in the first instance. Technical aspects of connecting to the OBIS website will then be planned.

How is OBIS governed?

OBIS is managed by an International Committee with advice from the CoML Steering Committee. The Chair of the OBIS International Committee, Dr Mark J. Costello, was also the OBIS Executive Officer until 2007, when a full-time Executive Director was appointed, Dr Edward Vanden Berghe. An OBIS Governing Board will be established during 2007. This will replace some of the roles of the International Committee, and will work closely with the OBIS Manager's Committee and Chairs of the OBIS Working Groups. The Executive Director reports to the International Committee (and in time to the Governing Board), coordinates OBIS activities, and manages the Secretariat and Portal. The Executive Committee is a sub-group of the International Committee (or Governing Board) and oversees implementation of activities between meetings.



What is the OBIS International Committee?

The OBIS International Committee (IC) is composed of invited expert scientists. The IC decides OBIS policy and strategy. It actively fosters OBIS development through data sharing and exchange, making software available, and supervising the OBIS portal development. It may establish working groups, for example to deal with technical data management issues. The IC is responsible for ensuring that contributors have expertise in their field, and have taken reasonable steps to ensure the quality of their data and/or software.

Members of the IC are selected as individuals to represent the constituencies of OBIS, including data providers, software developers, and regions of the world. They bring their expertise, knowledge, and connections to OBIS, and it is anticipated that they would be in regular contact with people and organizations who are potential users of, and contributors to, OBIS. Members do not represent their institutions. The IC elects the OBIS Chair from its members.

Who is on the OBIS International Committee?

Dr Mark J. Costello (Chair OBIS International Committee)
Leigh Marine Laboratory, University of Auckland
Warkworth, New Zealand

Dr Ann Bucklin
Department of Marine Sciences, Marine Sciences and Technology Center,
University of Connecticut, Groton, Connecticut, USA

Dr Daphne G. Fautin
Department of Ecology and Evolutionary Biology
Natural History Museum and Biodiversity Research Center
University of Kansas
Lawrence, Kansas, USA

Dr Patrick N. Halpin
Geospatial Analysis Program, Nicholas School of the Environment and Earth Sciences
Duke University
Durham, North Carolina, USA

Dr Gary C. B. Poore
Museum Victoria
Melbourne, Victoria, Australia

Dr Tony Rees
Divisional Data Centre
Commonwealth Scientific and Industrial Research Organisation (CSIRO) Marine Research
Hobart, Tasmania, Australia

Dr Karen Stocks
San Diego Supercomputer Center
University of California San Diego
La Jolla, California, USA

Ex officio

Mr Robert M. Branton (as Chair of the OBIS Manager's Committee)
Centre of Marine Biodiversity, Bedford Institute of Oceanography

Dartmouth, Nova Scotia, Canada

Dr Edward Vanden Berghe (Executive Director)
Rutgers, The State University of New Jersey
New Brunswick, New Jersey, USA

Dr J. Frederick Grassle (Director OBIS Secretariat, Chair CoML Scientific Steering Committee)
Institute for Marine and Coastal Sciences
Rutgers, The State University of New Jersey
New Brunswick, New Jersey, USA

Dr Yunqing (Phoebe) Zhang (Portal Manager)
Rutgers, The State University of New Jersey
New Brunswick, New Jersey, USA

Former OBIS IC members were:

- Dr Neil Ashcroft, University of Hull, UK;
- Dr Geoff Boxshall, The Natural History Museum, London, UK;
- Dr Dennis P. Gordon, National Institute of Water & Atmospheric Research (NIWA),
Wellington, New Zealand;
- Ms Kim Finney, National Ocean Office, Hobart, Australia;
- Dr Rainer Froese, Leibniz-Institut für Meereswissenschaften an der Universität Kiel, Kiel,
Germany
- Dr Yoshihisa Shirayama, Seto Marine Laboratory, Shirahama, Japan;
- Dr John Wilkin, Rutgers - The State University of New Jersey, USA.

What is the OBIS Science (Editorial) Board?

The OBIS Science Board is analogous to the Editorial or Advisory Board of a science journal in that members play an important role as advisors, such as in ensuring quality in the data publication process. This is achieved through members recommending good quality datasets for publication, acting as ambassadors within their specialist community for data publication through OBIS, and providing feedback and suggestions for improvements to OBIS. Where Board members were the scientists responsible for datasets published through OBIS, they maintain a role in correcting any errors apparent in the data and providing additional information to data users when requested. Where Board members have special expertise in taxonomy, oceanography, ecology, or information technologies, they advise OBIS on how to improve its quality and functionality. Members also add authority to OBIS by virtue of their scientific reputations. Members are:

Bailly, Nicolas Fish taxonomy, FishBase; Phillipines	Froese, Rainer FishBase; SpeciesBase Leibniz-Institut für Meereswissenschaften, Kiel, Germany	Research & Education, Washington, D.C. USA
Bouchet, Philippe Mollusca; taxonomy Museum National d'Histoire Naturelle, Paris, France	Gordon, Dennis Bryozoa, New Zealand; National Institute of Water and Atmospheric Research Ltd., Wellington, New Zealand	Rees, Tony Databases, Data Quality, Georeferencing, Web Mapping Tools; Commonwealth Scientific and Industrial Research Organisation (CSIRO) - Marine and Atmospheric Research, Tasmania, Australia
Boxshall, Geoff Copepoda, taxonomy The National History Museum, London, UK	Grassle, J. Frederick Chair Census of Marine Life Institute of Marine and Coastal Sciences, Rutgers - the State University of New Jersey, USA	Starkey, David HMAP University of Hull, United Kingdom
Buddemeier, Bob Hexacorallia, environmental data and characterization as it relates to habitat; Kansas Geological Survey, Lawrence, USA	Holm, Poul HMAP; Centre for Maritime and Regional Studies, University of Southern Denmark, Denmark	Stefansson, Gunnar; FMAP; Marine Research Institute of Iceland
Costello, Mark J. European Register of Marine Species, BioMar (Ireland) dataset, University of Auckland, New Zealand	O'Dor, Ron Cephalopoda, telemetry/animal tracking; Dalhousie University, Canada and Consortium for Oceanographic	Stocks, Karen; CenSeam; San Diego Supercomputer Centre, La Jolla, USA
Fautin, Daphne Hexacorallia; Natural History and Biodiversity Research Center, University of Kansas, USA		Wood, James; CephBase; Bermuda Biological Station, Bermuda

What is the OBIS Portal?

The OBIS portal is the software that enables data search and retrieval through www.iobis.org. It is physically located at the Institute of Marine and Coastal Sciences at Rutgers – The State University of New Jersey.

What are Regional OBIS Nodes (RON)?

Regional OBIS Nodes (RON) are organizations that have committed to a continued support of OBIS within a geographic and/or national region using resources they have obtained. This will include serving data online and developing a data provider and end-user community. Some RON will provide tools, different language versions of the OBIS website, and/or provide mirror sites for the OBIS portal.

What is the OBIS Manager's Committee?

The OBIS MC is composed of the Manager of each Regional OBIS Node. Members include

Australia

Cath Frampton, Department of the Environment and Water Resources, National Oceans Office, Commonwealth Scientific and Industrial Research Organisation (CSIRO)

Canada

Tana Worcester, Centre of Marine Biodiversity, Bedford Institute of Oceanography

China

Xiaoxia Sun, Institute of Oceanology, Qingdao

Europe

Ward Appeltans, Vlaams Instituut voor de Zee (VLIZ) Belgium

Indian Ocean

Vishwas Chavan, National Chemical Laboratory, National Institute of Oceanography India

Japan

Junko Shimura, National Institute for Environmental Studies

South-West Pacific

Don Robertson, National Institute of Water & Atmospheric Research New Zealand

South America

Mirtha Lewis, Centro Nacional Patagónico (CENPAT) CONICET Argentina

Fábio L. da Silveira and Rubens M. Lopes, University of São Paulo, Brazil

Ruben Escribano, FONDAP COPAS, Chile

Sub-Saharan Africa

Marten Grundlingh, Southern African Data Centre for Oceanography, South Africa

USA

Mark Fornwall, National Biological Information Infrastructure (NBII), Pacific Basin Information Node (PBIN)

Former OBIS MC members were:

Robert Branton, Canada

Alicja Mosbauer, Australia

Song Sun, China,
Edward Vanden Berghe, Europe

What was the OBIS Technical Committee?

This committee was a sub-committee of the OBIS International Committee from 2001-2006. It focused in detail on technical issues and reported to the OBIS IC.

What are OBIS Scientific Working Groups?

Scientific Working Groups are small groups of one to four people who conduct a well-defined task for OBIS for a specific period. They replaced the former OBIS Technical Committee. Current groups are listed below. Additional groups to consider linking of OBIS to molecular data systems, and to develop Educational tutorials are under consideration.

Taxon names Working Group, Tony Rees (Chair), Gary Poore, Edward Vanden Berghe, Mark Costello. This develops the (a) informal hierarchy of taxonomic groups used on the OBIS portal to enable users to browse down to their group of interest, (b) a world list of marine genus names to enable OBIS to find data on marine species in other data sources (*e.g.* GBIF), (c) mechanisms for labelling species by items of interest (*e.g.* if invasive, threatened), and (d) coordinates with other initiatives to make an authoritative World Register of Marine Species names available to the scientific community and as a quality control service on the OBIS portal.

Visualisation tools Working Group, Bob Branton (Chair), Pat Halpin, Jerry Black, Dan Ricard. This group is developing tools to show on the OBIS portal (a) individual animal tracks as lines, (b) time series data, and (c) abundance and absence data.

Habitat classification Working Group, Mark Costello (Chair), Karen Stocks. This group is preparing approaches to enable OBIS data to be searched by habitats, including pre-defined geographic regions (*e.g.* sea areas), physiographically defined seascapes (*e.g.* seamounts), and biologically defined biotopes.

Fishery data Working Group, Bob Branton (Chair), Dan Ricard. This group is reviewing the needs of fishery scientists with a view to publishing additional types of data through OBIS, and/or alternative ways of expressing or displaying data.

Discovery Metadata Working Group, Bob Branton (Chair), Melanie Meaux. This group is implementing discovery metadata for all datasets published on OBIS in conjunction with the GCMD metadata portal. During 2005, this group defined the metadata fields required for OBIS and included Mark Costello (Chair), Melanie Meaux, Edward Vanden Berghe, David Watts, Klaas Deneudt, and Simon Claus.

What is an OBIS contributor?

People, organizations and projects that provide data, software, expertise, or other resources to OBIS can be recognized as “Contributors to OBIS”.

What do all the acronyms mean?

Please see [Acronyms](#)

Who has contributed to OBIS?

People

Many scientists have contributed to the development of OBIS, including those on its International Committee, Manager's Committee, Technical Committee, Working Groups, Science Board, and working in its Regional Nodes, Portal and Secretariat. Others have contributed by providing data, software tools and know-how to OBIS; while more have provided invaluable feedback on their experiences and expectations using OBIS. This feedback has been very helpful in planning improvements to the OBIS portal and website.

In addition to people directly involved in OBIS and the Census of Marine Life, detailed user feedback was provided by students at the University of Auckland in 2005 (Alana Mary Alexander, Phillip Ian Baker, Jennifer Marie Blair, Bevan Ceran, Jennifer Elizabeth Davison, Catherine Anne De Silva, Sarah Victoria Hearne, Dianne Julie Hicks, Silje Hogner, Caroline Holt, Jacob Dylan Hore, Lee Jenks, Carl Johan Martin Karlsson, Inigo Macey Koefoed, Claudius Matthias Kroenke, Ta-Yuan Lee, Stefan Andreas Vallon Nord, Kylie Jane Park, Kirsten Jayne Redmond, Breeanne Louise Sandall, Kelly Joanne Siddle, Janice Mary Stanborough) and 2006 (Haman Amir-Shahpari, Brenda Fumiko Asuncion, Dane Attwood, Natalie Louise Beaton, Stephanie Jane Behrens, Dhruvi Bell, Marie Alice Bjorland, Cara Ann Blomfield, Blair Alexander Cardno, Rohan Fernandes, Ola Mea Gausen, Nicola Claire Gearing, David Gray Clutton Gilbert, Kimberley Anita Greenwell, Liv Marie Gustavson, Hanna Habte Selassie, Oliver Desmond Hannaford, Sarah Kay Harrison, Tania Dianne Hurley, Nicola-Anne Hawkins Jack, Joann Elizabeth Kelly, Emma Gail Lucas, Jenna Lee Martin, Tristan McArley, Stephen Paul Gifford Moore, Quentin O'Brien, Shane Edward Oswald, Mate Ki Tawhiti Chloe Nadia Peni, Jia Ren, Ben Richter, Elizabeth Jane Ross, Sarah-Jayne Rynbeck, Andrea June Seymour, Jenni Anne Stanley, Katrina Lee Subedar, Linda Carina Jeanette (Jeanette) Svensson, Nicola Mary Tarbutt, Elis Carl Petter Tibblin, Kari Torp, Rebecca Katherine Trigg, Kelly Patrick Withers, and Amie Wolken).

Data Sources

OBIS Data Sources are the authors, editors, and/or organisations that have published one or more datasets through OBIS. They would be the owners or custodians of the data. The data are served through an OBIS Data Provider who may or may not be the same organization or person.

Data Providers

OBIS Data Providers are organizations and persons that serve data through the OBIS portal. These data maybe held by themselves or served on behalf of others.

Tool providers

OBIS Tools are software that operates through the OBIS portal (*e.g.* mapping tools), is used in OBIS data management, or that can be downloaded by users. Examples include: CMR C-squares Mapper; Kansas Geological Survey (KGS) Mapper; ACON Mapper; VLIZ & IODE maps of Countries (Exclusive Economic Zones), and of Sea and Ocean areas.

Partners

OBIS Partners are organisations with whom OBIS cooperates to mutually support each others goals and activities. They include:

- Intergovernment Oceanographic Commission (IOC) and its International Oceanographic Data and Information Exchange (IODE) and Global Ocean Observing System (GOOS)
- Global Biodiversity Information Facility (GBIF)

- Pacific Biodiversity Information Network (PBIN)
- Inter-American Biodiversity Information Network (IABIN)
- Biodiversity Information Standards (Taxonomic Data Working Group, TDWG)

- Food and Agriculture Organization (FAO)
- International Council for the Exploration of the Sea (ICES)

- National Center for Biotechnology Information (NCBI)
- Consortium for the Barcode of Life (CBOL)

- Scientific Committee on Ocean Research (SCOR)
- International Association for Biological Oceanography (IABO)
- Global Earth Observation System of Systems (GEOSS)
- Ocean.US and Integrated Ocean Observing System (IOOS)
- Marine Metadata Interoperability project (MMI)
- Global Change Master Directory (GCMD)

- World Conservation Monitoring Centre (WCMC)
- International Union for the Conservation of Nature (IUCN)
- The Nature Conservancy (TNC)
- NatureServe
- Conservation International (CI)
- Diversitas

- Species 2000
- Integrated Taxonomic Information System (ITIS)
- Society for the Management of European Biodiversity Data and its European Register of Marine Species

Sponsors

OBIS sponsors provided significant funding that has contributed to OBIS development. This may be direct funding, such as from Alfred P. Sloan Foundation, European Commission, Gordon and Betty Moore Foundation, National Science Foundation (NSF), National Oceanographic Partnership Program (NOPP), and the National Oceanic and Atmospheric Administration (NOAA).

Indirect financial support is equally significant. It includes support from the institutions supporting the Portal and Secretariat, Regional OBIS Nodes, other major Data Providers, and salaries of committee members. These include: Rutgers (The State University of New Jersey, USA), the EU network of excellence 'Marine Biodiversity and Ecosystem Functioning (MarBEF)', Vlaams Instituut voor de Zee (VLIZ, Flanders Marine Institute, Belgium), San Diego Supercomputer Centre (USA), Centre of Marine Biodiversity (Canada), Bedford Institute of Oceanography (Canada), Commonwealth Scientific and Industrial Research Organisation (CSIRO) Marine Research (Australia), Duke University (USA), Natural History Museum (London, UK), National Institute of Water and Atmospheric Research (NIWA, New Zealand), Museum Victoria (Australia), Institute of Oceanology (China), National Chemical Laboratory and National Institute of Oceanography (India), South African Data Centre for Oceanography, National Biological Information Infrastructure (NBII, USA), and the Universities of Sao Paulo (Brazil), Concepcion (Chile), Kansas (USA), Kiel (Germany), Auckland (New Zealand).

What were the origins of OBIS?

The initial idea for OBIS developed from a CoML-sponsored Benthic Census Meeting held in October 1997. Recommendations from this meeting led to the establishment of a prototype OBIS Web site by J. F. Grassle, K. Stocks and Y. Zhang at Rutgers in 1998 to demonstrate the initial OBIS concept. The first OBIS International Workshop was held in November 1999 in Washington, D.C., and in 2000 the International Committee was formed to govern its global development. In 2000, the National Oceanographic Partnership Program (NOPP) requested proposals for OBIS projects and funded eight through support from the Alfred P. Sloan Foundation, Office of Naval Research (ONR) and National Science Foundation (NSF). A more restricted NOPP competition in 2002 resulted in an additional OBIS project on Marine Mammals, Turtles, and Birds (called OBIS-SEAMAP), and an NSF Postdoctoral Fellowship to Karen Stocks at the San Diego Supercomputing Center resulted in the OBIS linked database SeamountsOnline. Subsequently many additional datasets were published through OBIS, including the Census of Marine Life Initial Projects. See OBIS News webpage for further developments.

What has been published related to OBIS?

Please send additions to the list of OBIS related publications to OBISsupport@marine.rutgers.edu.

How do I cite this web page description of OBIS?

The above and other text on the OBIS website has been developed over several years by several people, and continues to be edited with updates and to answer questions we receive. It may be cited as: Costello M.J., Stocks K., Zhang Y., Grassle J.F., Fautin D.G. (March 2007). About the Ocean Biogeographic Information System. Retrieved www.iobis.org on [date].