

**INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)**

**Nineteenth Session of the IOC Committee on International Oceanographic Data
and Information Exchange (IODE-XIX)
Trieste, Italy, 12-16 March 2007**

The IODE Ocean Data Portal: Concept Paper

(by Greg Reed, Nick Michailov, Peter Pissierssens)

1. PREAMBLE

During the past 46 years, IOC Member States have built a network of National Oceanographic Data Centres (NODCs) (or their smaller Designated National Agency version) through the IODE Programme. The objectives of the IODE (revised at IODE-XVIII in 2005) are:

- (i) to facilitate and promote the exchange of all marine data and information including metadata, products and information in real-time, near real time and delayed mode;
- (ii) to ensure the long term archival, management and services of all marine data and information;
- (iii) to promote the use of international standards, and develop or help in the development of standards and methods for the global exchange of marine data and information, using the most appropriate information management and information technology;
- (iv) to assist Member States to acquire the necessary capacity to manage marine data and information and become partners in the IODE network; and
- (v) to support international scientific and operational marine programmes of IOC and WMO and their sponsor organisations with advice and data management services.

The recurring keyword in the objectives is “exchange” or sharing data between two or more partners. During the pre-internet era (i.e. 1960-1990) of IODE, the NODCs exchanged data with other NODCs using physical media such as tapes, disks, etc. Users were served using the same media. As from the late 1980s the Internet (through email, ftp and WWW) started its spectacular growth and NODCs gradually started exchanging data and information through this new mode of transport. During the past five years many, although not all, NODCs have developed web interfaces enabling users to query and retrieve datasets from the NODC databases. This has been a major improvement to “the old days” but it has not fully satisfied the user community.

2. THE OCEAN DATA PORTAL CONCEPT

2.1 BACKGROUND

If today we visit 50 NODC websites then we will find 50 different user interfaces, different ways to query the databases and different ways (formats) that the data are provided to the user. As users increasingly wish to combine datasets from different data providers this multitude of formats generates does not promote the image of IODE as a coherent and well organized network of data centres. It is with this in mind that the idea of a “Data ATM” (automated teller machine (English), distributeur de billets (French), cajero automático (Spanish), банкомат (Russian)) was born: with a bank card issued in your own country you can travel around the world and, at any time, withdraw cash from the bank account in your own country.

A better analogy can be made with the “electronic repository” or “digital repository” developed in the librarian domain. The following is a definition of a Digital Institutional Repository looks very similar to what we understand as a Data Centre.

What is a Digital Institutional Repository?

“A university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution.”

Clifford A. Lynch,

["Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 \(February 2003\): 1-7.](#)

Many libraries have now established institutional repositories of academic publications. Each document is accompanied by a description (ie metadata) that describe the document (title, author, source,...). At the institutional level the e-repository system can be queried through the description (metadata) and the selected documents retrieved in full text (ie the data). The Librarians then went a step further by networking ex-repositories. A protocol was developed (OAI-PMH: The Open Archives Initiative Protocol for Metadata Harvesting – see <http://www.openarchives.org/OAI/openarchivesprotocol.html>). This enabled the establishment of “portals” that enable searching and retrieving documents across a distributed network of e-repositories. The searching is not done across the entire network in real-time (as this would take too long). Instead the “portal” system regularly visits all “nodes” of the network and “harvests” the metadata. Users will search the network of repositories by searching the harvested metadatabase at the portal location. Once the search is completed and documents selected these are then retrieved in real-time from the e-repository that holds the document.

It is proposed that the IODE Ocean Data Portal uses similar principles for its architecture. The objective of the IODE Ocean Data Portal is to build a distributed network of oceanographic data centres enabling the searching and retrieving of datasets.

There are some fundamental differences between library documents and datasets. We could identify these as follows:

- The Ocean Data Portal should have the capacity to combine datasets from different providers (responding to the same query) and present these in one output (e.g. as tabular output or as a map representation).
- There are many parameters and codes used in oceanography. To enable the combination of data, a user must be certain that the definition of a parameter used by all data providers are the same, and that the unity of measure is the same. This requires the development of a common parameter and code dictionary.
- Metadata. The e-repositories use the Dublin Core metadata standard which is a simple set of 15 elements sufficient to locate a document (author, the year of publishing, etc.). The Ocean Data Portal will provide discovery metadata to enable a user to assess the suitability of marine datasets. Metadata will describe data quality, platforms and equipment, modelling methods. The Ocean Data Portal will have to deal with geospatial data and the ISO 19115/19139 standard for geographic metadata is a suitable standard to describe geospatial data and services. The use of internationally recognised standards will ensure interoperability with other initiatives under development both at the national and international levels. Marine datasets are often dynamic and are updated yearly, monthly or even hourly depending on the activity. The Ocean Data Portal should provide a mechanism to populate and update the discovery metadata from the data provider nodes. This can be done by one of the following methods: (i) direct entry via an online form; (ii) a manual upload of a single record using an XML document conforming to the ISO 19115/19139 specification; or (iii) using a metadata harvester to automatically collect metadata records from contributing data centres.
- Data formats. Marine data is stored in many diverse data formats. The Ocean Data Portal should not require that all submitted data be in a single format, but should provide the tools to identify the local data structures and transform these data into a common transport data format.
- Data granularity. The Ocean Data Portal should recognise the different granularity of marine data and provide appropriate query and transfer of data based on parameter type, geographic area, time period, etc., and should provide functionality to support data granularity, i.e. to describe, search and deliver defined units of data, such as single cruise, single buoy, or single coastal station.

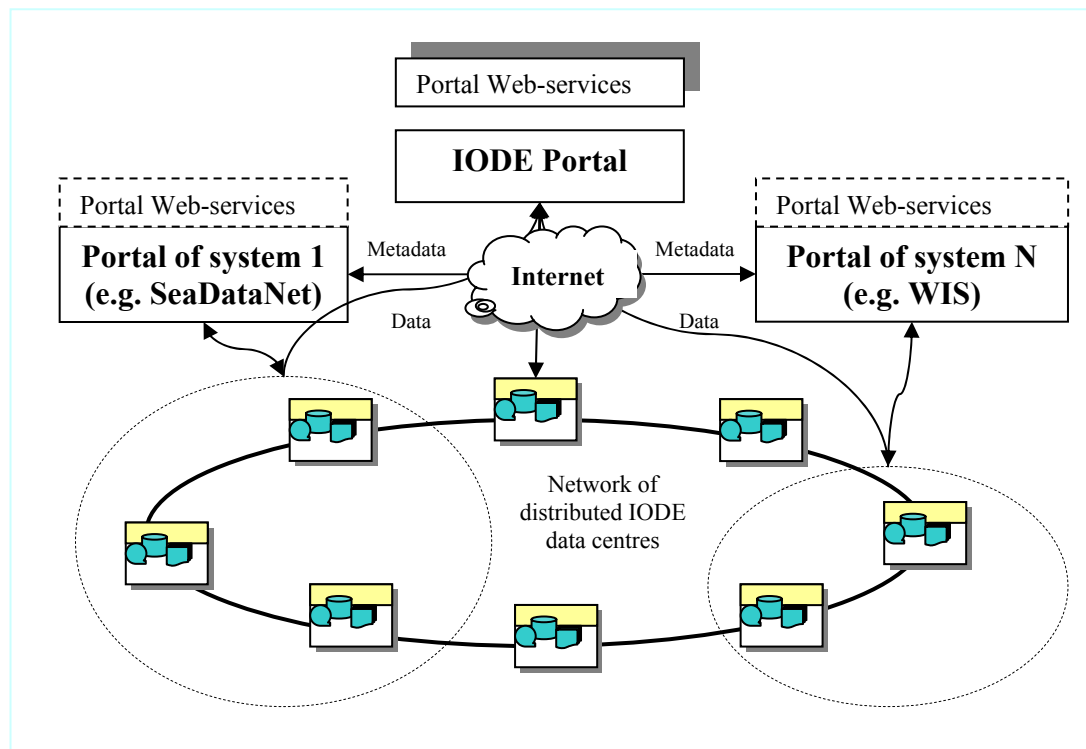
It is important that the IODE Ocean Data Portal supports the data access requirements of all IOC programme areas, including GOOS, JCOMM, HAB and the Tsunami warning system.

2.2 OBJECTIVE

The objective of the Ocean Data Portal is to facilitate and promote the exchange and dissemination of marine data and services. The Ocean Data Portal will provide seamless access to collections and inventories of marine data from the NODCs in the IODE network and will allow for the discovery, evaluation (through visualisation and metadata review) and access to data via web services. The system architecture will use Web-oriented information technologies to access non-homogeneous and geographically distributed marine data and information.

2.3 VISION

The Ocean Data Portal will deliver a standards-based infrastructure that provides the integration of marine data and information from a network of distributed IODE NODCs as well as the resources from other participating systems, as shown on the diagram below.



The Ocean Data Portal will provide on-line access to the marine data and information resources of the participating data centres including:

- operational and delay-mode data
- data and services from the oceanographic and marine meteorological domains
- from multiple data source formats and local data systems (DBMS, data files, GIS, electronic documents)
- from multiple data providers in different geographic regions

The Ocean Data Portal will not create a new data system. The key principles of the Portal will be interoperability with existing systems and resources. Participating IODE data centres will need to accept and implement a set of agreed interoperable arrangements including the technical specifications and Web-services for the integration and shared use of the metadata, data and products. This interoperability will be achieved through the use of internationally endorsed standards (such as SOA, ISO and OGC) and it will not be a requirement for data centres to change their internal data management systems. The interoperability arrangements will be developed in close cooperation with existing and developing systems (see section 3) and will follow international standards and best practices.

The Ocean Data Portal will provide the full range of processes including data discovery, access, and visualization. The following functionality is envisaged:

- IODE data centres will generate discovery metadata about their datasets for distributed data search and retrieving.
- The Portal will periodically harvest these metadata, monitor the accessibility of a remote data sources and update the portal metadata catalogue.
- Users can access the system via a web browser and search for single or multiple data types from a distributed set of sources.
- Data request will be sourced from the appropriate data centre and returned to the Portal.
- Portal tools will fuse the aggregated data and services in real time to produce a new product or service of value to the user.

The Ocean Data Portal will conform to open standards service-oriented architecture (SOA) and will include the following components:

- (i) Data providers Participating NODCs will form nodes on the IODE data centres network and will provide data and metadata to the Portal. Discovery metadata will be harvested from the contributing nodes to populate the Catalogue.
- (ii) Data Portal. An interface to provide the ability to search for data from contributing data centres which will make available data discovery, visualization, content management and administration tools.
- (iii) Catalogue. The metadata catalogue will provide a registry of existing data and services available from the data providers.

The final list of Portal components and their implementation are not final and will be determined when the Portal architectural design is developed.

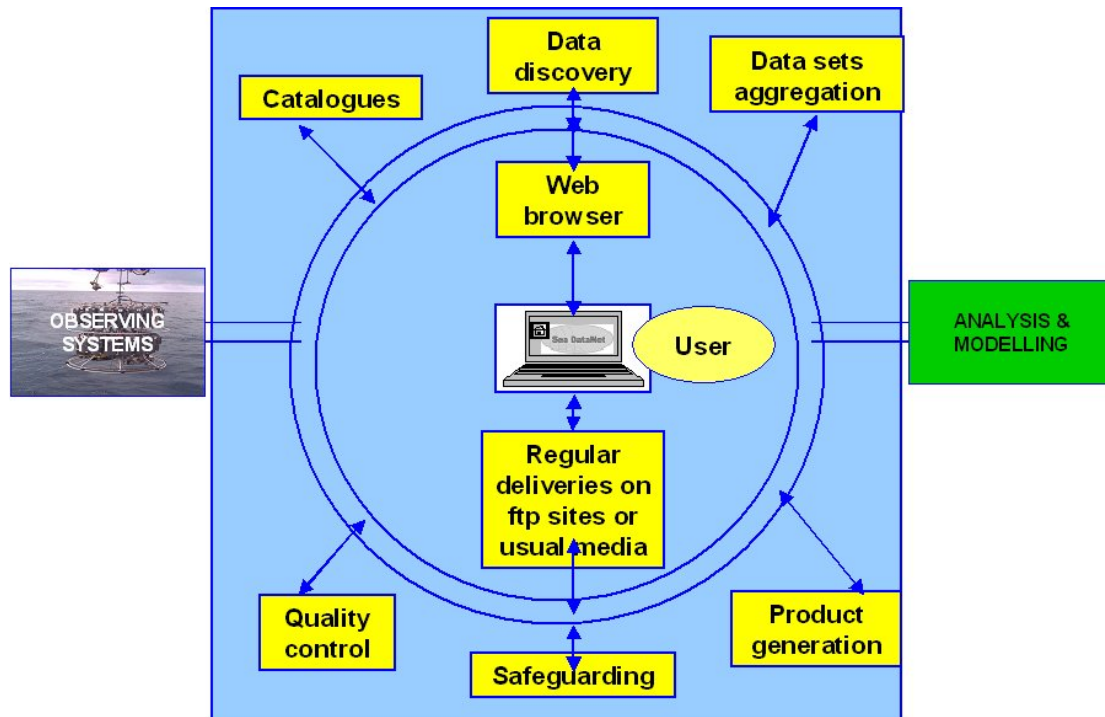
3. SIMILAR INITIATIVES

A number of similar initiatives are planned or have started. These include both national and regional initiatives. Some examples are described below. The IODE Ocean Data Portal should not compete with existing projects but rather strive towards compatibility between these initiatives and possibly set up a global harvester/portal system, thereby complying with IODE's global coordination and exchange mandate.

3.1 SEADATANET EU Project

SeaDataNet (a Pan-European infrastructure for Ocean and Marine Data Management, 2006-2010) is a major Pan-European and EU-funded project, undertaken by the National Oceanographic Data Centres (NODCs) and marine information services of major national institutes from coastal states bordering the European seas (<http://www.seadatanet.org/>). SeaDataNet provides the data management infrastructure, to insure long term archiving and access to the large number of multidisciplinary data (i.e. temperature, salinity current, sea level, chemical, physical and biological properties): by many different sensors installed on board of research vessels, satellite and the various platforms of the marine observing system.

SeaDataNet objective is to construct a standardized system for managing the large and diverse datasets collected by the oceanographic fleets and the new automatic observation systems. The objective is to network and enhance the currently existing infrastructures, which are the national oceanographic data centres and satellite data centres of 35 countries, active in data collection. The networking of these professional data centres, in a unique virtual data management system will provide integrated data sets of standardized quality on-line.



Main Functional Components and Links of the SeaDataNet Data Management System

SeaDataNet Virtual Ocean Data Centre is a distributed infrastructure that provides transnational access to marine data, metadata, products and services through 40 interconnected data platforms from countries around the Black Sea, Mediterranean, North East Atlantic, North Sea, Baltic and Arctic regions:

- National Oceanographic Data Centres (NODC),
- Designated National Agencies for International Exchange of Marine Data and Information,
- Operational Data and Products Thematic Centres,
- Satellite Data Centres.

SeaDataNet is providing access to metadata, data and added value products through a central portal. The SeaDataNet portal is providing information services through the following catalogues:

- EDMED - search for Marine Environmental Data;
- EDMERP - search for Marine Environmental Programmes;
- EDMO - search for Marine Environmental Organisations;
- CSR - search for Cruise Reports;
- EDIOS - search for Operational Activities;

These directories, available for consultation on the SeaDataNet website

<http://www.seadatanet.org>, represent a key tool for data search and retrieval and an EU contribution to larger international systems.

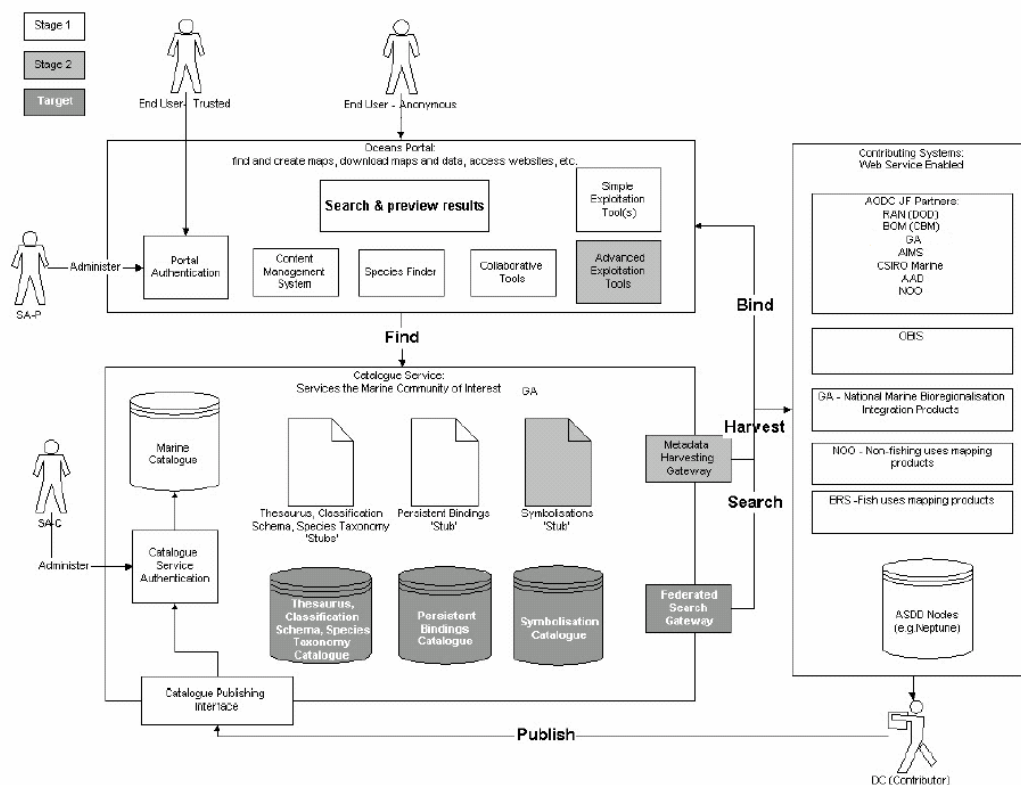
The common standards for vocabularies, discovery services, definition and adoption of formats and protocols for data checking are developed in the project Technical Task Team in strong cooperation with international experts. They follow international ISO basic standards and general practices in data management for set-up of web services and content governance structures, transformation services, downloading services and viewing services. Following this necessary prerequisite, the interoperable system software is developed in order to have 11 test platforms connected in version 1 (2008), and the 40 platforms connected by steps at the end of the project (2010).

3.2 AUSTRALIA'S OCEAN PORTAL

The Australian Ocean Data Centre Joint Facility (AODCJF), the national body responsible for coordinating the management and distribution of ocean data, is developing a multi-agency distributed data management system to manage the ocean data resources of the partner agencies. The AODCJF will provide a national infrastructure flexible enough to permit on-line access to the distributed archives of data held in the partner institutions across Australia. To achieve this infrastructure, AODCJF is developing a set of standards, business rules and technologies to improve the online exchange of marine data and information that utilises a service-oriented-architecture. The Oceans Portal is the content and service discovery component of this architecture and provides the primary portal to share and exchange data and products at the national level.

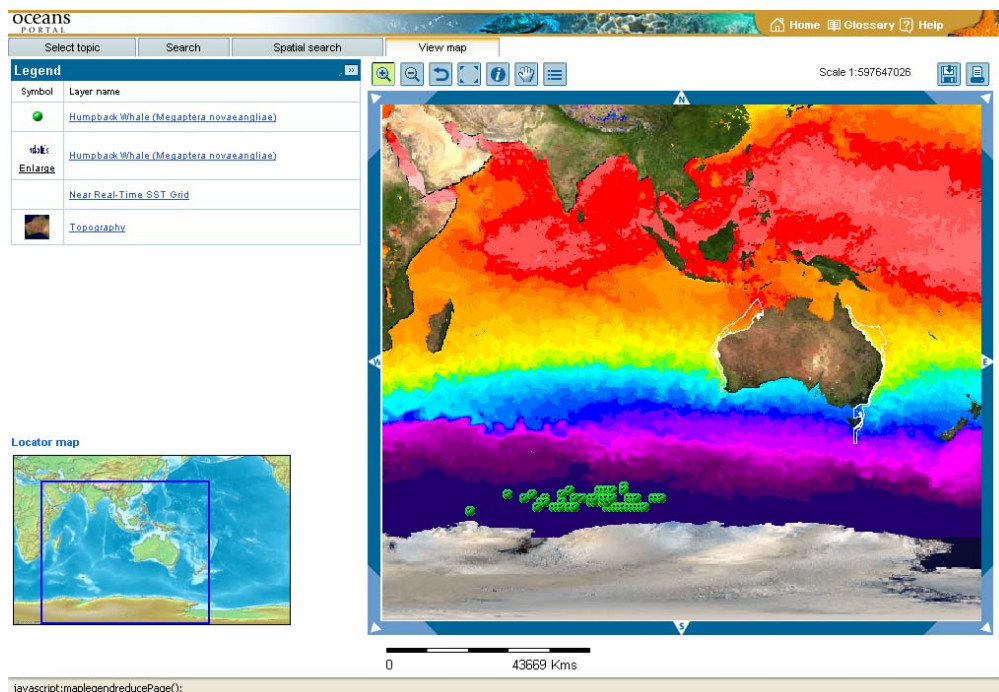
The Oceans Portal is modelled on services oriented enterprise architecture with three main components:

- The portal - a web-based interface including discovery and visualisation functions to discover, access and utilise information resources registered in the marine catalogue.
- The marine catalogue - a web-accessible repository used to register and manage metadata describing discoverable content (data and services). The marine catalogue content can be discovered through the portal or any other client that supports the OGC Catalogue Services interface.
- Web-services that provide the discoverable content (data and services) that are described and registered in the marine catalogue and can be discovered through the portal.



The key principles of the Oceans Portal include:

- (i) Separation of components - the key system components of the Oceans Portal (Portal, Catalogue and Service and Content providers) are logically and physically separated.
- (ii) Single point of truth – the Oceans Portal promotes access to authoritative sources of data and information services.
- (iii) Metadata driven – the Oceans Portal is metadata driven. Any data or information service must have a metadata record. The integrity of the Portal is based on the quality of the metadata.
- (iv) Discovery of any data or service – any data or information service that has its metadata record published will be discoverable by the user.
- (v) Multiple methods of access – the Oceans Portal allows for multiple methods of access. Points of access will be through client applications that adhere to the relevant standards.
- (vi) Open standards – the Oceans Portal conforms to open standards based on a service-oriented architecture (SOA). SOA uses the “publish, find, bind” paradigm.



The Oceans Portal application provides:

- The ability to search for marine data and metadata available from a network of linked data providers;
- A range of data portrayal services such as web mapping;
- The ability to download data from a network of interoperable data providers;
- Links to relevant national and international oceans-based activities and initiatives.

By using open standards the Oceans Portal will be capable of accessing data and services from similar system being developed internationally.

3.3 OBIS

The Ocean Biogeographic Information System (OBIS) is an on-line, open-access, globally-distributed network (1000 researchers in 73 countries) of systematic, ecological, and

environmental information systems (12.9 million records of 200 databases). OBIS is based on DiGIR (Distributed Generic Information Retrieval). The key DiGIR components are:

- Provider - an application that makes data available to compliant portals;
- Portal - an application that communicates with multiple providers, performs operations to retrieve and integrate data (and metadata), and serves as the point of access for users;
- Protocol - the definition of message formats (i.e. requests and responses) by which components must communicate.

DiGIR supports only one type of the local data storage system, DBMS, which restricts the scope of the DiGIR application. The DiGIR technology can be used as the basic technology for access to distributed databases (see, IODE/JCOMM E2EDM technology prototype).

3.4 OPeNDAP/DODS

The Distributed Oceanographic Data System (DODS) is a system that provides access to large distributed datasets (over 340 datasets, 1000s of data files, from more than 15 partners around USA) through data analysis and visualization software packages (such as IDL, Ferret, ODV, Grads, etc.) and Application Program Interfaces (APIs). The DODS technology is based on the OPeNDAP software and protocol tools. The OPeNDAP provides functionality to serve supporting binary data file collections and client functions to access data over the internet. The OPeNDAP core library supports different data access API - netCDF, JGOFS, HDF, DSP (satellite AVHRR/ CZCS data), GRIB (WMO) and BUFR (WMO). API is also declared for the FreeForm format using the format definition language for conversion of data points and data grid models. OPeNDAP/DODS supports only data file collections in these formats.

3.5 NERC DATA GRID

The Natural Environment Research Council (NERC) Data Grid aims to provide a framework for the discovery and use of data required by NERC. The NERC Data Grid concept is based on the federation of integrated provider resources, each of which can comprise one or more sources (data systems). Key elements are: D-metadata (discovery), B-metadata (security) and A-metadata (format and usage). A source interface has three components: anonymous access to D-metadata via OAI and A-metadata and D-metadata interfaces. To provide user access three additional services are expected to be developed: NERC DataGrid Version 0.2 release was declared for discovery and delivery data of many aspects of science.

3.6 WMO INFORMATION SYSTEM

WMO information system (WIS) provides a single entry point for any meteorological data request, be it on a routine basis by dissemination of certain user defined information sets or on an ad-hoc basis for a special dataset. The WIS components are:

- Metadata catalogues;
- Internet portal;
- Data acquisition and discovery service;
- Data distribution service: push and pull;
- Monitoring.

From a functional view, the WIS consists of three major components, (i) National Centres (NC), (ii) Data Collection or Product Centres (DCPC), and (iii) Global Information System Centres (GIS) together with a data communication network connecting the

components. The WIS software component is developed under the SIMDAT EU project (based OAI) and DWD (Germany, WebWerDis software). The WMO ISO 19115 metadata profile describes the WMO datasets and this standard is used for GISC portal function implementations to harvest metadata from DCPC/NC nodes and to publish data and products. The WIS is under development and is focussed on the following prototype projects:

- WMO Core Profile version of metadata and reference implementation;
- RA VI VGISC project (Exeter, Offenbach, Toulouse) as a GISC prototype;
- DCPCs prototypes including the ECMWF and EUMETSAT DCPC projects associated with the VGISC project, also including the SIMDAT project;
- DCPC prototype (World Data Centre, Obninsk) for JCOMM related data (IODE/JCOMM E2EDM pilot project);
- NCAR DCPC prototype.

3.7 GEOSS

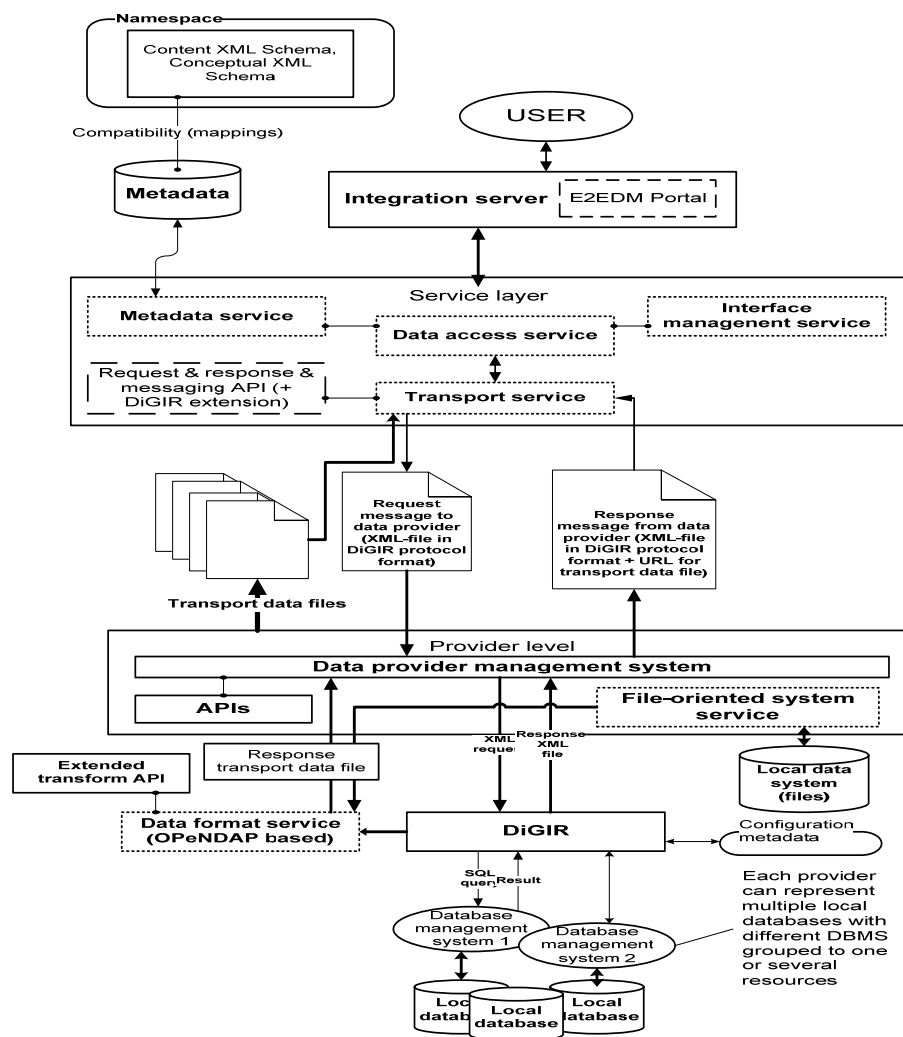
The intergovernmental Group on Earth Observations, GEO, is working to establish a Global Earth Observation System of Systems (GEOSS). The Architecture and Data Committee has been established to support GEO in all architecture and data management aspects of the design, coordination, and implementation of GEOSS. A GEO system task, DA-06-07, has been established to

Define a model web portal system for access to all Earth observation data, based on existing portals and systems, designed to increase use, quality, and accessibility of existing information, tools, and networks. Particular attention will be given to the coordination of networks in specific societal benefit areas to enable reuse thereby to achieve synergy and leverage

A portal prototype solutions will be sought to provide seamless common access to collections and inventories of geographic information, including imagery, vector, and sensor data allowing for the discovery, evaluation (visualisation, metadata review), and access or order of data via Web Services. The portal solutions will be based on open standards for catalogue services (OGC Geospatial Portal Reference Architecture) and will provide timely access to distributed geographic catalogues maintained by participating organisations. It is expected that a demonstration version of the GEO portal will be available by September 2007.

3.8 IODE/JCOMM E2EDM TECHNOLOGY PROTOTYPE

The prototype of the “end-to-end data management” technology (E2EDM) was developed under the JCOMM/IODE ETDMP Pilot Project. The E2EDM components are:



- Technical specifications for metadata and data records is based on a unified data model and ISO 19115 with extensions (description of a data source, common codes and elements dictionaries, local data storage system, response/replay protocol).
- Integration Server – (i) harvesting and synchronization of metadata from data providers; (ii) interaction with external applications on response/replay protocol, (iii) connection to data sources and delivery of transport data files from data sources.
- Data Provider– (i) supporting local metadata; (ii), connection to the data source (DBMS, data file, electronic document types) on requests from Integration Server, (iii) data selection, reformatting local codes to system codes, preparing response message and transport data file (netCDF, ASCII, XML) to Integration Server.

The development of the E2EDM technology components used existing software and tools (e.g. DiGIR, OPeNDAP, and WMO WIS Metadata Profile). The E2EDM technology has been tested using the following marine data sources:

- MetOffice, UK. IMMT data files with historical marine meteorological data.
- VLIZ, Belgium. DBMS, historical ocean cruise data.
- RIHMI-WDC, Russia. DBMS, electronic documents, real-time GTS marine meteorological (SHIP), ocean (BATHY and TESAC), monthly climatic fields, marine weather charts (imagery).
- IFREMER, France. DBMS; real-time GTS ocean (TESAC/ARGO) data.

The E2EDM technology prototype began operations in 2005. The current E2EDM status is:

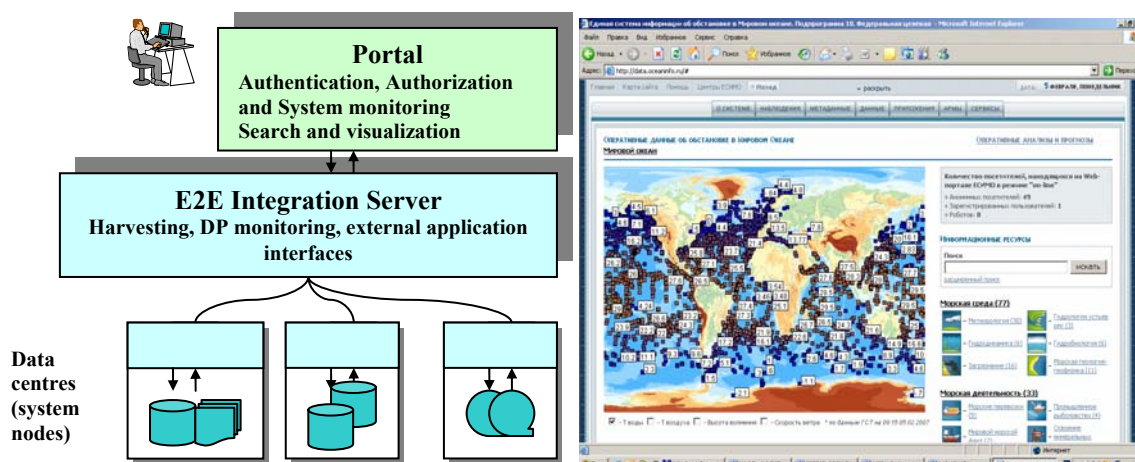
- Integration Server at VLIZ (RIHM_WDC – mirror);
- 3 data sources and 18 information resources - Data Providers at VLIZ, IFREMER, MetOffice and RIHMI-WDC;
- External application as Ocean Data Portal prototype at Integration Server at VLIZ (RIHMI-WDC mirror).

3.9 RUSSIAN ESIMO

The ESIMO (Unified System of Information for World Ocean conditions) has been developed as part of the “World Ocean” federal programme to integrate marine data and products (operational/non-operational, ocean, meteo, ice, geology, fishery, marine transport, economy, etc.) stored by the participating agencies and to provide the user with access to these data and products.

The ESIMO development is coordinated by RIHMI-WDC/Russian NODC. The ESIMO system is made up of the following components:

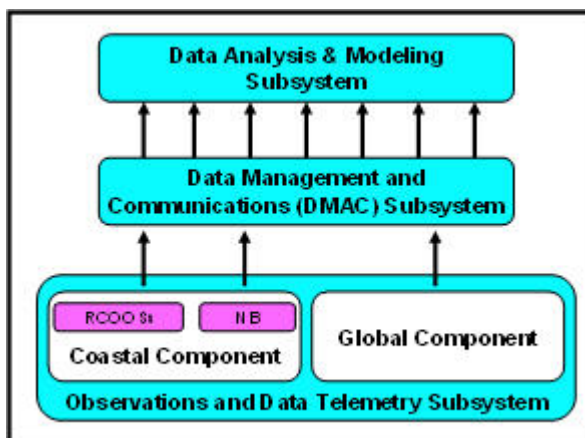
- Virtual telecommunication network connecting the data sources form the data centres.
- Distributed information resource system integrating of the geographic distributed data sources (databases, data file, electronic documents, and applications). The E2E technology is used to manage this distributed data system.
- Portal – a single entrance into the system providing authentication, authorization, system administration, and data publishing.
- Centralized metadata base to manage common codes and classifiers, information about platforms, organizations, experts, projects, etc.



ESIMO has been tested under a number of pilot projects and the system integrates more 800 information resources of 29 data sources (including 50 databases, data file system, document repository). It is planned to begin the system operation at the beginning of 2008.

3.10 IOOS DATA MANAGEMENT & COMMUNICATIONS SUBSYSTEM (DMAC)

Central to the success of [Integrated Ocean Observing System \(IOOS\)](#) and other regional, national, and international ocean and coastal observing systems, is the presence of a Data Management and Communications (DMAC) Subsystem capable of delivering: real-time and delayed-mode observations to a wide variety of users: data streams to modeling centers; model-generated forecasts to users; distributed biological measurements to scientists, educators, and planners; and all forms of data to and from secure archive facilities. The needs of end users must be a part of the implementation and operation of the subsystem, both as sources of specifications for subsystem design, and as agents of change to keep the delivery of products from IOOS-relevant to national interests. The Figure shows

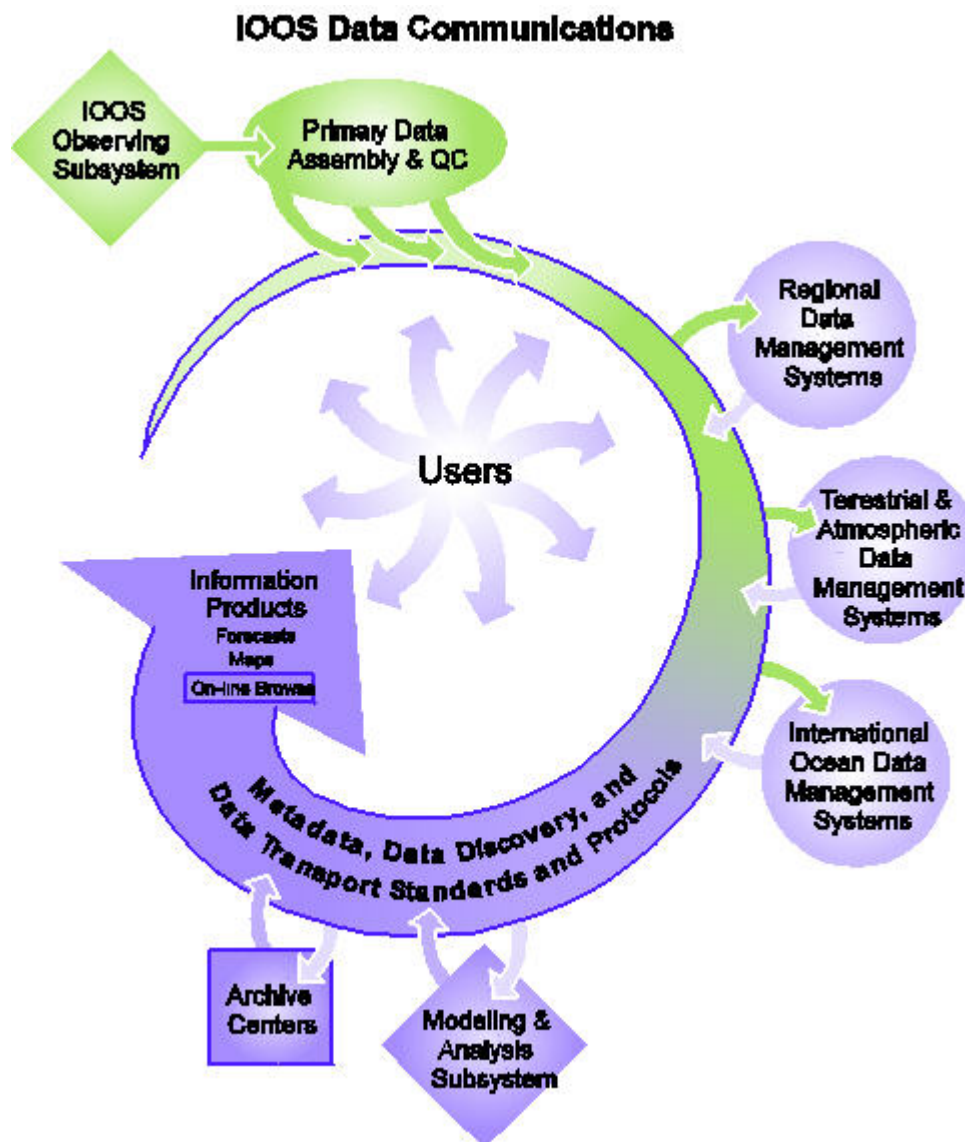


The needs of end users must be a part of the implementation and operation of the subsystem, both as sources of specifications for subsystem design, and as agents of change to keep the delivery of products from IOOS-relevant to national interests. The Figure shows

Conceptually, the IOOS consists of 3 efficiently linked subsystems (see Figure): Observing System (in situ measurements, remote sensing, and data telemetry), a Data Management and Communications (DMAC) subsystem and a Data Analysis and Modeling (DAM) subsystem

The relationship among the DMAC Subsystem and other IOOS components and partners is depicted in the figure below. Data flow within IOOS begins with the Observing Subsystem. Raw measurements from the Observing Subsystem elements are processed at various Primary Data Assembly and Quality Control sites to make them available to the uniform DMAC Data Communications Infrastructure. The infrastructure consists of standards and protocols to support:

- IOOS-wide descriptions of data sets (Metadata);
- the ability to search for and find data sets of interest (Data Discovery);
- the ability to access the data in an interoperable manner from client applications (Data Transport)
- the ability to evaluate the character of the data through common Web Browsers (Uniform On-line Browse); and
- the ability to securely archive data and metadata and retrieve them on demand (Data Archive).



Schematic of the DMAC subsystem. Solid outlines indicate the elements of the IOOS Data Communications framework, which are detailed in the [DMAC Implementation Plan](#). The arrows flowing outward from the users indicate the feedback and control mechanisms through which users ultimately direct the functioning of all parts of the system. Note that the National Data Management Systems are included in the concept of Primary Data Assembly and Quality Control.

A community-based process is underway by the [DMAC Steering Team](#) (appointed by Ocean.US in spring 2005) to identify, evaluate, and adopt existing data and metadata standards (protocols, interfaces, formats,) and best practices relevant to IOOS DMAC, and to propose development of new standards and tools. A description of this process, and the latest version of the resulting document, Guidance to IOOS Data Providers, can be found at <http://dmac.ocean.us/dacsc/guidance02.jsp>.

The information technology required to meet most of the needs of DMAC, while challenging, can be developed from existing capabilities through relatively straightforward software engineering. The greatest challenge facing DMAC is one of coordination and cooperation among IOOS partners and user communities. DMAC can succeed only if the participants actively use the data and metadata standards, communications protocols, software and policies that will knit IOOS into an integrated whole. The creation of a successful IOOS

DMAC will require a sustained effort, a commitment across the U.S. marine community, and continual coordination with our international counterparts.

Additional points to make:

- IOOS is organizing framework for US contribution to GOOS
- DMAC guide has been updated in June 2006
- DMAC principles involve adopting existing standards whenever possible, and looks to international standards promulgated under IOC and WMO as a source
- DMAC Steering Team has included an "International Caucus" in its structure to facilitate collaboration with the international community. The present Caucus membership includes several IODE and JCOMM participants (see http://dmac.ocean.us/dacsc/group_international.jsp)

DMAC Guiding Principles	
<u>Interoperability:</u>	DMAC will serve as a framework for interoperability among heterogeneous cooperating systems. The cooperating systems will be free to evolve independently to address the needs of their target users. Software and standards needed to participate in DMAC will be available directly to partners, or provided through commercial and non-commercial sources. DMAC will also be interoperable with systems outside of the marine community that manage atmospheric and terrestrial data.
<u>Open, easy access and discovery:</u>	DMAC will enable users from all over the globe to easily locate, access, and use the diverse distributed forms of marine data and their associated metadata and documentation in a variety of computer applications (e.g., Geographic Information Systems [GIS], and scientific analysis applications). Users will be unencumbered by traditional barriers such as data formats, volumes and distributed locations. DMAC will integrate cooperating systems so that data discovery will be seamless, and multiple versions will be easily tracked. There will be a "free market" of ocean sciences information, including officially sanctioned IOOS data sets, as well as data and products from other sources.
<u>Reliable, sustained, efficient operations:</u>	DMAC will provide high reliability with 24/7 delivery of real-time data streams from measurement subsystems to operational modeling centers and users with time-critical requirements. It will provide high reliability in the delivery of computer-generated forecasts, estimates of state, and delayed-mode and real-time data to end users. DMAC will require sufficient bandwidth and adequate carrying capacity to support large exchanges of raw data and model outputs among high-volume users. DMAC will offer techniques that reduce the need for large data transfers, such as server-side subsetting and computation, to allow users with limited bandwidth to enjoy the benefits of IOOS. Feedback mechanisms will be built into the technical design of DMAC to ensure that problems are detected and rapidly addressed.
<u>Effective user feedback:</u>	IOOS will provide a continuous, vigorous outreach process addressing all levels of users of marine data, emphasizing the benefits of

	participation in IOOS/DMAC, and helping to identify and remedy difficulties encountered by those who are participating. In addition, this process will identify and address changing user requirements that drive the development and growth of DMAC.
<u>Open design and standards process:</u>	DMAC will commit to an open software design. All standards and protocol definitions will be openly published so that participating organizations may create functioning DMAC components based on these specifications. The standards development process will be open and inclusive, so that it fosters buy-in by all stakeholders. Existing information technology and scientific standards will be used in preference to development of new solutions, whenever suitable standards exist. The standards and protocols will be of sufficient breadth and quality to guarantee interoperability of all observations and products. Institutions participating in IOOS will ensure that the data they contribute comply with these standards and protocols.
<u>Preservation of data and products:</u>	Irreplaceable observations, data products of lasting value, and associated metadata will be archived for posterity in an efficient and automated manner.
To learn about the DMAC planning and coordination activities and access DMAC plans and other guidance documents, go to: http://dmac.ocean.us/index.jsp	

4. THE IODE OCEAN DATA PORTAL PROJECT

The Ocean Data Portal Project will have the following work packages:

1. Project coordination and management

This package will assemble the main players from existing national or regional initiatives (data nodes) with the objective to ensure compatibility between the existing systems and the proposed IODE Portal.

2. Standards development package

This package will develop and maintain data standards and commonly associated terms, definitions and schemas to enable the efficient and effective exchange and use of data. Issues of interoperability, harvesting, web services, data dictionaries, etc will be addressed. Appropriate standards, best available practices and technical solutions for data interoperability will be identified and, wherever possible, existing standards will be used. The work package will work closely with other IODE projects, such as ETDMP, MarineXML and MEDI, to ensure existing standards and protocols are used.

3. Portal implementation package

This package will define the data portal model to access to marine data, based on existing portals and systems. It will focus on implementing a system to harvest metadata from the participating data providers and to set up a global portal through which users will be able to query the global catalogue and retrieve datasets from the participating data provider nodes.

5. WORKPLAN AND BUDGET

At this early conceptual stage we are not submitting a detailed work plan and budget. It is instead proposed to start with the composition of the project working group (work package 1), followed by work package 2. The project working group will be responsible for the drafting of a detailed work plan and budget that will be strongly guided by the aforementioned similar initiatives to avoid duplication of effort. It is recommended that the project working group meets once in 2007 and once or twice in 2008, depending on need. The working group will work by email as much as possible.

Timeline:

June – December 2007	work package 1 startup
January 2008 – June 2009	work package 2 and work package 3

Costing

2007:	US\$ 20,000
1/2008 – 6/2009:	US\$ 100,000 (preliminary estimate)

The Committee is requested to:

- review and comment on the Document
- establish the project working group and define its terms of reference, taking into consideration the information included in the Document
- consider the establishment of an IODE Ocean Data Portal Pilot Project

[DRAFT RECOMMENDATION WILL BE PREPARED]

[end]