

OCEAN DATA PORTAL: A STANDARDS APPROACH TO DATA ACCESS AND DISSEMINATION

Greg Reed⁽¹⁾, Robert Keeley⁽²⁾, Sergey Belov⁽³⁾, Nikolay Mikhailov⁽³⁾

⁽¹⁾ Australian Ocean Data Centre Joint Facility, Level 2, Building 89, Garden Island, Potts Point NSW 2011, Australia.

Email: greg@metoc.gov.au

⁽²⁾ ISDM (Integrated Science Data Management), 1202-200 Kent Street, Ottawa, Ontario K1A 0E6, Canada.

Email: Robert.Keeley@dfo-mpo.gc.ca

⁽³⁾ All-Russian Research Institute of Hydrometeorological Information – World Data Center (RIHMI-WDC), 6, Koroleva Street, Kaluga District, OBNINSK 249035, Russian Federation. Email: nodc@meteo.ru

ABSTRACT

Timely access to quality data is essential for the understanding of marine processes. The International Oceanographic Data and Information Exchange (IODE) programme, through its distributed network of National Oceanographic Data Centres (NODCs), is developing the Ocean Data Portal (ODP) to facilitate seamless access to oceanographic data and to promote the exchange and dissemination of marine data and services. The ODP provides the full range of processes including data discovery, evaluation and access, and delivers a standards-based infrastructure that provides integration of marine data and information across the NODC network.

The key principle behind the ODP is its interoperability with existing systems and resources and the IODE is working closely with the Joint WMO-IOC (World Meteorological Organization-International Oceanographic Commission) Technical Commission for Oceanography and Marine Meteorology (JCOMM) to ensure the ODP is interoperable with the WMO Information System (WIS) that will provide access to marine meteorological and oceanographic data and information to serve a number of applications, including climate. The ODP supports the data access requirements of all IOC programmes areas, including GOOS (Global Ocean Observing System), HAB (Harmful Algal Blooms) and the Tsunami warning system as well as JCOMM.

The diverse data standards and formats that have evolved within the oceanographic community make data exchange complex and the IODE community has recognized standards are critical in defining how and what data is exchanged. To ensure the interoperability of data exchanged between the NODCs and the ODP, the IODE, together with JCOMM, has initiated a standards process that will support the accreditation and adoption of core standards by the marine meteorological and oceanographic communities.

1. DATA SHARING PRINCIPLES

The Earth's oceans form part of an integrated global system and to address global issues, such as climate change, it is essential for scientists to have access to relevant data, information, and products. The full and open sharing of datasets is fundamental to ensure the rapid dissemination of data and information is available to researchers. International policies for the full and open exchange of scientific data and information are advocated by a number of international organizations.

The Intergovernmental Oceanographic Commission (IOC) of UNESCO (United Nations Educational Scientific and Cultural Organization) has adopted a resolution entitled *IOC Oceanographic Data Exchange Policy* (Resolution IOC-XXII-6). This policy recognizes that the timely, free and unrestricted international exchange of oceanographic data is essential for the efficient acquisition, integration and use of ocean observations. These data are gathered for a wide variety of purposes including the prediction of weather and climate, the operational forecasting of the marine environment, the preservation of life, and the mitigation of human-induced changes in the marine and coastal environment. Under this policy, IOC member states agree to provide timely, free and unrestricted access to all data, associated metadata and products generated under the auspices of IOC programmes. In addition, IOC member states are encouraged to provide free and unrestricted access to relevant data and associated metadata from non-IOC programmes that are essential for application to the preservation of life, beneficial public use and protection of the ocean environment, the forecasting of weather, the operational forecasting of the marine environment, the monitoring and modelling of climate and sustainable development in the marine environment [1].

Other international organizations have also adopted similar policies to encourage the sharing of data. The World Meteorological Organization (WMO) has

adopted a policy for the international exchange of meteorological and related data and products. WMO Resolution 40 provides for the free and unrestricted sharing of data [2].

The Group on Earth Observations (GEO), which is coordinating efforts to build a Global Earth Observation System of Systems, or GEOSS, explicitly acknowledges the importance of data sharing in achieving the GEOSS vision and anticipated societal benefits. GEO is developing a set of high level Data Sharing Principles which call for the "full and open exchange of data, metadata, and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation." These Principles also note that "All shared data, metadata, and products will be made available with minimum time delay and at minimum cost" [3].

2. INTERNATIONAL OCEANOGRAPHIC DATA AND INFORMATION EXCHANGE (IODE) PROGRAMME

The Intergovernmental Oceanographic Commission's IODE programme was established in 1961 *"to enhance marine research, exploitation and development by facilitating the exchange of oceanographic data and information between participating Member States and by meeting the needs of users for data and information products"*. The objectives of the IODE 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.

For nearly 50 years, IOC Member States have been building and contributing to a network of National Oceanographic Data Centres (NODCs) through the IODE Programme. Over this period, IOC Member States have established 80 oceanographic data centres in

IOC Member States (Fig. 1). The IODE network is responsible for the collection, quality control, and archive of many millions of ocean observations, and these data are made available to the Member States. The IODE programme encourages free and open access to, and exchange of, marine scientific and oceanographic data and information among the relevant institutions and agencies in the member states and focuses on all ocean related data including physical, chemical, and biological.

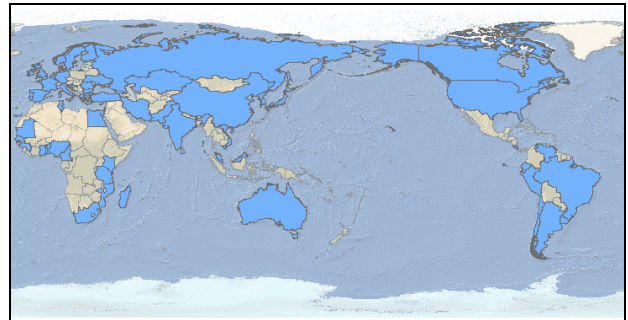


Figure 1. The IODE network of National Oceanographic Data Centres

Many NODCs provide web interfaces to allow users to query and retrieve datasets from localized databases. However, there is currently no focal point where users can go to identify and gain access to all available ocean data and products. To facilitate the timely, consistent and integrated access to the oceanographic data holdings of the NODC network, the IODE has initiated the Ocean Data Portal project which will establish a single point of access to data collections and inventories of marine data to support data discovery and access. 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.

3. DATA STANDARDS

One of the objectives of the IODE is *"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"*. Lack of agreement on common standards, formats and practices means that there is a significant amount of human intervention needed before data downloaded from one site can be used by another. In the past, with slow communications methods and relatively small data volumes, this was not such a significant issue. With the growth of the internet for the exchange of data and information there has been an increased capacity for sending much larger volumes and almost instantly. This has highlighted the need for community wide standards for management and exchange of ocean data simply to improve the efficiency of data exchange. Interoperability between systems can

only be achieved by the use of agreed standards and there are a number of internationally accepted standards which are applicable to oceanographic data. These include those developed by the International Organization for Standardization (ISO), the World Wide Web Consortium (W3C) and the Open Geospatial Consortium (OGC).

In 2008, the IODE, jointly with JCOMM, sponsored a meeting to examine the potential for the development and acceptance of community wide standards for marine data and information management and exchange [4]. The objective of this meeting was to gain general agreement and commitment to adopt standards related to key ocean data management thereby facilitating exchange between oceanographic institutions. The meeting developed a process to accept, evaluate and recommend proposals for community wide standards. Community participation in the process so far has been slow and if this standards development process is to be successful, it is essential for the marine data management community to play an active role.

4. OCEAN DATA PORTAL

The objective of the Ocean Data Portal is to facilitate and promote the exchange and dissemination of marine data and services through the provision of seamless access to collections and inventories of marine data from the NODCs in the IODE network and to allow for discovery, evaluation and access to data via web services. This is achieved through 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 [5].

The key principle of the Ocean Data Portal is interoperability with existing systems and resources. Participating IODE data centres agree to accept and implement a set of interoperable arrangements including the technical specifications and web services for the integration and shared use of the metadata, data and products. This interoperability is achieved through the use of internationally endorsed standards and best practice (such as ISO and OGC) and does not require data centres to change their internal data management systems.

The ODP is currently deployed in five data centres from the Back Sea region. Further datasets will be contributed to ODP by data centres in the USA, Canada, Australia, UK and the East Asian region by early 2010.

4.1 ODP Architecture

The architecture of the Ocean Data Portal consists of three basic components (Fig. 2):

- (i) **Data Provider.** Provides access to data and metadata of the local data systems. When the wrapper is installed in the local data system, the latter becomes a data source for the distributed data system. A recent addition to the Data Provider software is the Light Data Provider function which offers remote registration of local datasets and allows deployment of the ODP system without the need to install software by the data provider.
- (ii) **Integration Server.** Provides registration and operation status monitoring of the distributed data sources, harvesting of the discovery metadata in coordination with Data Provider, management of the common codes/dictionaries and access to distributed data sources by ODP services. The Integration Server also interacts with other systems (portals) by means of discovery metadata exchange.
- (iii) **ODP Services.** Provides administration, discovery, viewing, analysis and download. The Data Portal includes a GIS-based user interface, metadata and data search, data download and visualisation components. The ODP services include a number of W3C and OGC web-services.

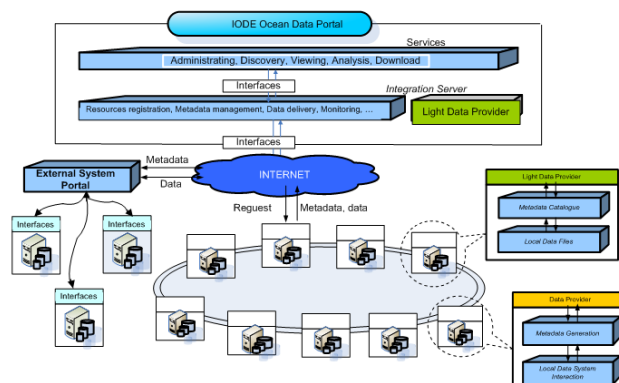


Figure 2. Architecture of the Ocean Data Portal

4.2 Functional and technical requirements

The ODP provides the following functionality:

- distributed marine data infrastructure generation and operation;
- data discovery and access;
- data provision to end-users;

- user management; and
- system monitoring and reporting

The technical aspects include the provision of metadata and services for a distributed marine data infrastructure enabling the interactions among data providers, service providers and the end-users.

The system allows data interaction whilst avoiding data reformatting and delocalization so the data always remains within the data provider's infrastructure. ODP allows adjustment of services invoking online request/response processes and other operations and also allows chaining of services into more complex ones. The system supports "subscription" type services and standing orders (e.g. oil spill monitoring and alerting) and allows the easy identification of, and access to, requested services and data, with progress follow-up until completion. The integration of data and services from multiple domains is possible to facilitate exploitation of main synergies and service and data providers can register, provide and promote their products to other thematic or regional portals.

ODP is built using open standards which provides the capability to interact with other data portals which minimizes the investment required by data and service providers to build on open standards.

Discussions are currently underway between the ODP developers and SeaDatNet, a pan-European infrastructure for marine and ocean data management (<http://www.seadatanet.org>), to ensure interoperability between the two systems, ensuring that data requests and data delivery are coordinated

4.3 Communication requirements

The Ocean Data Portal supports geographically distributed marine data infrastructure operations that can publish and disseminate technical guides and reports to IODE data centres and other participating centres. Dissemination of information about the status of all ODP partners will ensure coordination and cross-communication and provide remote software installations and documentation access. Communication will also allow ODP partner groups to discuss and resolve ongoing development issues.

Communication between Integration Server and Data Provider was improved by web-service creation which provides existing request-response communication both with fault-tolerant processing and error catch, recognition and logging.

Communication across the ODP consists of two processes:

- Metadata harvesting by the Integration Server from the Data Providers. These resource descriptions are exposed to a harvester, which is part of the Integration Server. This software will regularly (at any set frequency) check all data centres for new resource descriptions and download these as necessary. These descriptions are added to a central repository that covers all data centres connected to the distributed system.
- Request on data. Requests are transmitted by HTTP in encoded form. HTTP-connection between the Integration Server and Data Provider are active during processing and the Data Provider requests acceptance, validation, execution and response return. Communication based on web-services provides transactional and fault-tolerant mode. If errors occur the Integration Server will immediately receive a message with an error code and description and all errors will be published by the specific web-service by means of *mostError* method. If a request for data was executed successfully the Data Provider will invoke *postResult* method with a response-message in the transport protocol structure.

4.4 Access to distributed marine data

The Ocean Data Portal provides data and product dissemination services, which are divided into discovery, viewing, analysis and download services.

Discovery service disseminates a data source catalogue with descriptions of resources in the form of XML files. The metadata record is based on ISO 19115. The ODP service provides user interfaces for data and product search supported by the catalogue (Fig. 3). The data source catalogue can be accessed from external systems directly or alternatively by reformatting into other metadata structures.

- data search that defines the sampling criteria using a spatial region, time period, phenomena, platform, etc.
- access to remote data sources via the Integration Server including request status monitoring; and
- processing of transport data files and tabular-graphic and map visualization of data using standard forms.

Analysis service has been developed to provide near real-time GIS-layer generation from distributed datasets both with interactive and fast presentation of multidisciplinary data and products on a map. It also includes Web Map Services (WMS) as a viewing service for data representation on a map. The user can adjust the composition of the map layers, the number of maps for viewing and other specifications. The mapping service enables a joint analysis of data to provide a view of the spatial variability of marine processes. ODP

renders maps generated by the analysis service using Open Layers (Fig. 4) and MapServer (Fig. 5).

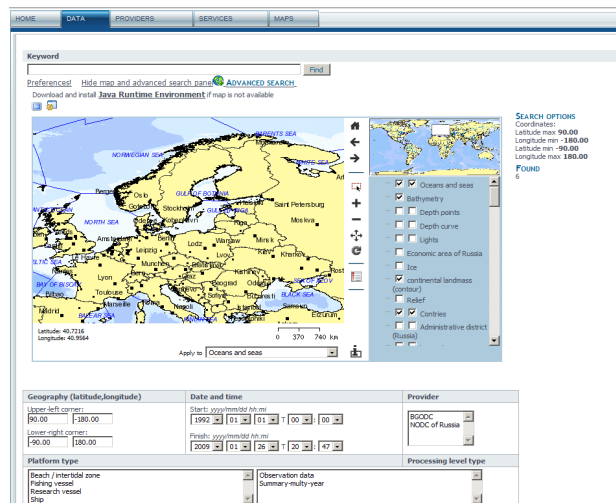


Figure 3. Discovery service user interface provides the ability to search the metadata catalogue.

Viewing service is based on web-based applications accessible via the web browser. The services provided include:

Download service allows the user to download selected data to the local computer after viewing. If time scheduling is required to download data, the user can register the site for downloading, the list of required datasets and the sampling criteria (Fig. 6). The transport data file formats that are available are:

- NetCDF - E2E convention [6]
- ASCII
- XML

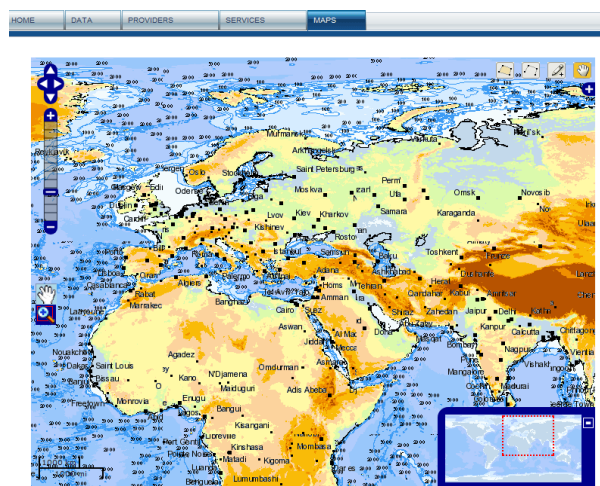


Figure 4. OpenLayers is used to display thematic layers generated by the ODP WMS service from near real-time Data Provider datasets.

Selected data can be either downloaded in a single zip-file or viewed using the ODP result viewing service.



Figure 5. MapServer is used to display thematic layers (shape files) generated by the ODP WMS service from near real-time Data Provider datasets.

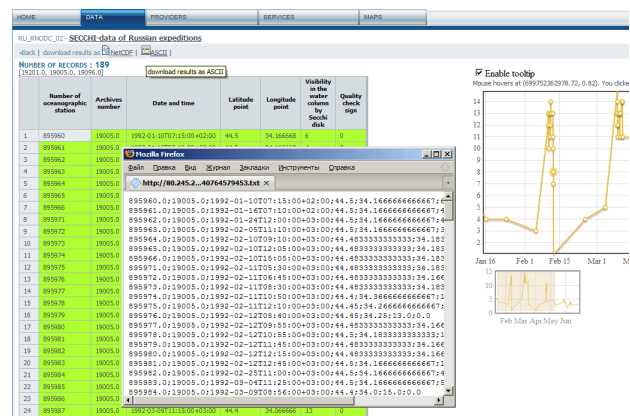


Figure 6. The user is able to download selected data in NetCDF or ASCII format.

5. INTEROPERABILITY WITH OTHER OCEAN DATA SYSTEMS

Interoperability arrangements for the Ocean Data Portal are being developed in close cooperation with existing and developing systems such as the WIS and the GEOSS.

The WIS provides a single coordinated global infrastructure for the collection and sharing of information in support of all WMO and related international programmes. The WIS consists of three major components: National Centres (NC), Data Collection or Product Centres (DCPC), and Global Information System Centres (GISC). The Ocean Data Portal will contribute to the WIS as a DCPC. The WIS will also be the core component of the GEOSS for weather, water, climate and disaster societal benefit areas.

GEOSS will be based on existing observing, data processing, data exchange and dissemination systems and the implementation of GEOSS will facilitate the development and availability of shared data, metadata, and products. GEOSS interoperability will be based on non-proprietary standards, with preference to formal international standards. Interoperability will be focused on interfaces, defining only how system components interface with each other and thereby minimizing any impact on affected systems other than where such affected systems have interfaces to the shared architecture.

6. SUMMARY

The Ocean Data Portal will deliver a standards-based infrastructure to provide integration of marine data and information from a network of distributed IODE NODCs as well as the resources from other participating systems. It will serve to coordinate the view of ocean data resources with other developing systems such as WIS and GEOSS. This interoperability will be achieved through the use of internationally endorsed standards and it will not be a requirement for data centres to change their internal data management systems.

7. REFERENCES

1. Intergovernmental Oceanographic Commission. 2003. IOC Oceanographic Data Exchange Policy.
http://www.iode.org/index.php?option=com_content&task=view&id=51&Itemid=95. Accessed: 2009-04-19.
2. World Meteorological Organization. 1995. Resolution 40 (Cg-XII).
http://www.wmo.int/pages/about/Resolution40_en.html. Accessed: 2009-04-19.
3. Group on Earth Observations. 2009. GEO Data Sharing Principles Implementation,
http://www.earthobservations.org/geoss_dsp.shtml. Accessed: 2009-4-19.
4. Intergovernmental Oceanographic Commission. 2008. IODE/JCOMM Forum on Oceanographic Data Management and Exchange Standards, IOC Project Office for IODE, Oostende, Belgium 21-25 January 2008. Oostende, Belgium: IOC/IODE Project Office, 45pp. (IOC Workshop Report No. 206) (English)
5. International Oceanographic Data and Information Exchange. 2009. Ocean Data Portal.
<http://www.oceandataportal.org/>. Accessed: 2009-04-19.
6. JCOMM (Joint WMO-IOC Technical Commission on Oceanography and Marine Meteorology). 2008. Overview of End-to-End Data Management Technology.
http://data.meteo.ru/e2edm/files/resourcesmodule/@random4240448bb69c9/1237217263_1_E2EDM_overview_bk_gr_mn.pdf