

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)

Sixteenth Session of the IOC Committee on International Oceanographic Data and Information
Exchange (IODE), Lisbon, Portugal, 30 October – 9 November 2000

Report on the Medi Pilot Project

Prepared by Greg Reed, Australian Oceanographic Data Centre

1. Background

The Marine Environmental Data Information Referral Catalogue (MEDI) is a directory system for datasets, data catalogues and data inventories within the framework of the IOC's International Oceanographic Data and Information Exchange (IODE) system.

The IOC Working Group on International Oceanographic Data Exchange (now IOC Committee on International Oceanographic Data and Information Exchange - IODE) at its Fifth Session (September 1971) recommended the establishment of a Joint Task Team on Interdisciplinary and Inter-organizational Data and Information Management and Referral (IMAR), with participation by interested international organisations. The team was asked to assess on-going information and data management activities within the respective organisations and to develop a multi disciplinary referral system for providing users with information as to availability and location of environmental data.

The Joint Task Team recommended the establishment of a **Marine Environmental Data Information Referral System (MEDI)** containing technical descriptions of the marine data holdings of the participating organisations in automated-searchable form to enhance the service capabilities of marine data centres. As a first step in this direction the Joint Task Team recommended preparation of a **MEDI Pilot Catalogue**, using input from volunteer centres, which would serve to demonstrate the full scope of the system.

The recommendation of the Joint Task Team was accepted by the Working Committee on International Oceanographic Data Exchange (IODE) at its Eighth Session in 1975 (Rec. IODE-VIII.7), and then approved by the Ninth Session of the IOC Assembly in Resolution IX.30.

In this Resolution, the IOC:

- "Decides, subject to the concurrence of other member organisations of the Joint Task Team i.e., WMO, FAO, ICES, IHO and UNEP that the necessary co-ordinating functions for the Marine Environmental Data Information Referral System be performed by the IOC Secretariat as the UNEP/IRS contact point in the field of marine environmental data and information;
- that the Joint Task Team on IMAR be discharged and its members be thanked for their services; and that an inter-agency mechanism using points-of-contact be established to co-ordinate MEDI development; LI>Approves the establishment of the IODE Group of Experts on MEDI, as recommended by the Eighth session of the Working Committee on International Oceanographic Data Exchange;
- Invites Member States and all interested organisations and agencies to participate in the MEDI Referral System; and
- Invites further the Executive Director of UNEP to support the full implementation of MEDI".

2. The MEDI Catalogue

MEDI, as originally designed and implemented, was a collection of detailed descriptions of marine environmental datasets. It not only contained information about availability and location of individual datasets, but more importantly, specific geographical coordinate information as well as parameters. The MEDI system was designed to operate using a Data Base Management System (DBMS) on the UNESCO mainframe computer.

Since its inception the MEDI catalogue has been published three times:

First Edition

The first edition of the MEDI Catalogue was published in 1979 as IOC Manuals and Guides No. 10 and contained 86 data file descriptions from 40 institutions in 20 countries and international organisations. Although the source file was a mainframe database the product was only available in printed format.

Second Edition

The second edition of the MEDI Catalogue was published in 1985 as IOC Manuals and Guides No. 16. It contained 219 data file descriptions from 64 institutions in 32 countries and international organisations. The product was made available in printed format only.

Third Edition

The third edition of the MEDI Catalogue was published in 1993 as IOC Manuals and Guides No. 16. It contained 247 data file descriptions from 40 institutions in 27 countries and international organisations. The product was available in printed format only. In this edition, designed as a loose leaf set, a 'Product Evaluation Questionnaire' was included. However, only few recipients replied to this questionnaire.

3. The MEDI Pilot Project

The 15th session of IODE recommended that a Pilot Project be undertaken to:

"Test the ways and means of applying modern technology to the further development of the MEDI system and, on the basis of these investigations, to draft a specification for a revised MEDI."

The terms of reference for the Pilot Project included:

- Preparation of a statement on the level and breadth of information to be included in the system;
- Identification of a core set of fields necessary to constitute a MEDI entry;
- Identification of the technical requirements of the database;
- Development of transfer interfaces between existing databases and the agreed-upon system;
- Development of suitable end-user interfaces for the database
- Design a system to ensure the widest possible coverage of data holdings in Member States.
- Ensure the MEDI system is compatible with initiatives developed by other programmes

Appendix A details the IODE-XV recommendations for the development of MEDI including the Terms of Reference for the Pilot Project.

The development of the MEDI pilot project included a review of relevant existing national and international data directory systems. A review was undertaken in February 1997 of existing National and International data directories that included a detailed comparison of the data content of the European Directory of Marine Environmental Data (EDMED) and the Marine and Coastal Data Directory of Australia (Blue Pages). The review of these two data directories revealed many similarities in the data structures allowing a direct mapping between most fields. The resulting MEDI structure incorporated many of these fields.

Version 1 of the MEDI Pilot Project software was developed by the Australian Oceanographic Data Centre in 1998 using MS Access as the database management system with an additional spatial search tool to provide a user-friendly interface to the database. A CD version of the software, together with full documentation, was produced in November 1998 and was distributed to data centres in Australia, Thailand, Vietnam and the participating ODINEA countries (Kenya, Tanzania, Mozambique, Madagascar, Seychelles, Mauritius and South Africa). The software was also evaluated by members of the IODE Group of Experts on Marine Data Management (GE-MIM) and at the Sixth Session of GE-MIM in June 1999 suggested amendments to the structure were discussed.

The development of global data directories, in particular the Global Change Master Directory (GCMD), required a review of the MEDI structure to ensure its compatibility with other global directories. The ability to merge metadata collected by MEDI with global data directories is essential if MEDI is to be a viable system for the collection marine metadata. Data centres do not have the resources to maintain multiple data directories so the ability to uploading datasets from MEDI to a global directory is important. At the Eight Session of GE-TADE in March 2000 it was decided to integrate the GCMD DIF structure into MEDI to ensure that MEDI becomes a viable system for collecting marine metadata.

4. Status of the Pilot Project

The first version of the MEDI Pilot Project software was written using Microsoft Access 97. It could run only on PCs and upgrading to the latest version (Access 2000) would require a rewrite of the source code. A major problem with using a proprietary database management system is that IODE is prescribing which software a data centre should use. There is also the issue of backward compatibility (which has occurred with Access 97). It was decided that it would be more appropriate for the MEDI software to be database and platform independent thus providing access to a larger number of users.

To resolve the problem of software and hardware dependency, the software was rewritten using Java with a Java Database Connection (JDBC) allowing the user to connect to the database of their choice ensuring the MEDI software is hardware independent as well as database independent.

A beta version, prepared by AODC, written in Java and using a Java database engine is presented here. The software has been made available to members of GE-TADE, GCMD and other interested people for comment. Feedback from these users has resulted in some minor modification to data entry interface. The software has been tested on PCs (Windows 98, NT) and Unix (Sun Solaris). A Macintosh version will be available after the release of MAC OS10. Online help files, with worked examples, are available to assist the user.

The MEDI Pilot Project database fields are based on the GCMD DIF format. There are a number of GCMD keyword valids and suggested keywords that can be selected from drop-down lists. The software provides easy to use data entry forms, text and spatial search tools and the ability to export metadata records in XML, HTML or text. XML provides the mechanism to transfer metadata to the GCMD.

The MEDI software is available in English, French, Spanish and Japanese.

5. Issues for Discussion

There are a number of issues to be resolved before the Pilot Project software can be used in a production environment.

GCMD issues.

- Location Keywords. GCMD uses Location to describe the name of the place where the data was collected. The GCMD list of Location valids is too general to describe marine datasets. MEDI has adopted the IHO "Limits of Ocean & Seas" to describe Geographic coverage. In some cases the IHO list may be too general if local areas need to be identified. There needs to be a mapping from GCMD Location valids to the IHO list. One solution would be to adopt the GCMD method of describing Parameters. The syntax for parameter is: Category->Topic->Term->Variable->Detailed Variable. The Category, Topic, Term and Variable are all selected from the Parameter valids while Detailed Variable is a free text line of up to 80 characters. The Location could be described in similar way: Location (GCMD valids)->Sea Area (IHO valids)->Detailed Sea Area (free text). An example would be Pacific Ocean->Coral Sea->Gladstone Harbour.
- Updating valid lists. There should be a mechanism in place to request additions to current GCMD valid lists. For example, the GCMD Project valids do not contain many oceanographic projects. Likewise with the Sensor valids and Source valids. All requests to update GCMD valid lists should be filtered through a single source within IODE.
- Mechanisms for uploading data from MEDI to GCMD. MEDI has the ability to upload metadata records to GCMD but the issue of DIF numbering needs to be resolved.
- MEDI DTD definition needs to be finalised. This should be done through discussion with GCMD.
- Set-up a GCMD/MEDI node for metadata creation.

Software enhancements.

- Data validation is yet to be implemented. For example, there is no check on date values, numeric values are not checked for validity and string length is not checked.
- Save as default. This function allows certain fields to be entered once only and these values are then saved as default values. A decision on which fields should be set as defaults must be made before this feature can be implemented.
- Translation of Help files. The MEDI data entry interface is available in English, French, Spanish and Japanese. However, the help files are only available in English. If there is a demand from other language groups for the software then the help files will have to be translated.

6. Conclusion

The MEDI Pilot Project has resulted in the development of user-friendly software for the creation of metadata records in a standard format, ensuring compatibility or convertibility with the emerging ISO standard for metadata and other major national standards.

The philosophy behind the design of MEDI structure has been to develop a system to allow users to adequately describe marine datasets to a level of detail sufficient for a dataset to be identified and for a judgment of the quality of the dataset to be made. This requires the collection of only the minimum amount of metadata to describe the dataset. This approach allows the user to easily populate and maintain the directory and will thus encourage its use.

The ability to transfer metadata descriptions from MEDI to global data directories, in particular the Global Change Master Directory, will ensure that MEDI becomes a viable system for collecting marine metadata. Data centres do not have the resources to maintain multiple data directories so the ability to upload datasets from MEDI to a global directory is important. In addition, major global data collection programs, such as GOOS, can use MEDI to describe their datasets, with MEDI providing the mechanism to transfer metadata to global directories.

Appendix A

RECOMMENDATIONS
Recommendation IODE-XV.1
DEVELOPMENT OF MEDI

The IOC Committee on International Oceanographic Data and Information Exchange,

Recognizing the value of a top level directory of marine data holdings within the IOC Member States to a broad user community, including the IOC programmes and related activities within other global programmes such as IGBP, WCRP and GCOS,

Recognizing further the significance of such projects as GODAR for locating and identifying datasets not already exchanged as a fundamental requirement to support national, regional and global data requirements arising from GOOS,

Noting its longstanding involvement in the development of the MEDI system and that past efforts have been inhibited by the lack of appropriate supporting technology,

Further noting that recent developments in computer technology, particularly the World Wide Web, provide new opportunities for the proper development of MEDI;

Recommends that a pilot project be developed to test the ways and means of applying modern technology to the further development of the MEDI system and, on the basis of these investigations, to draft a specification for a revised MEDI,

Further recommends that the pilot project should include:

- (i) a review of relevant existing national and international data directory systems;
- (ii) an assessment of the potential application of existing bibliographic systems to MEDI;
- (iii) consideration of the implications of inter-operability with similar systems on the international scene such as those being developed by GCOS and the ICSU Panel on WDCs;

Supports the Terms of Reference of the Pilot Project, as shown in the Annex to this Recommendation;

Recommends that a 2-3 day Technical Workshop of experts should be held at the end of the pilot project (between May/August 1997) to review the results of investigations carried out, to agree on a specification for a revised MEDI system and to prepare plans for its implementation;

Invites the IOC Executive Secretary, Member States and relevant agencies to support the pilot project and the Technical Workshop.

Annex to Recommendation IODE-XV.1
TERMS OF REFERENCE FOR THE PILOT PROJECT ON THE REVISION OF MEDI

Objectives of MEDI:

To develop a global database of information on data holdings held in the IOC Member States and agencies with the following specifications:

- (i) the database will be a compilation of input assembled by IOC Member States and relevant agencies;
- (ii) the database will allow the end-user to search, as a minimum, on location, data type, temporal resolution and organizational parameters;
- (iii) the database will provide the end-user with information describing the selected data holdings and their sources;
- (iv) the system will be designed in such a manner as to ensure the widest possible coverage of data holdings in Member States.

Objectives of the Pilot Project:

The Pilot Project will:

- (i) prepare a clear statement on the level and breadth of information to be included in the system;
- (ii) identify a core set of fields necessary to constitute a MEDI entry;
- (iii) identify the technical requirements of the database;
- (iv) develop transfer interfaces between existing databases and the agreed-upon system;
- (v) develop suitable end-user interfaces for the database also bearing in mind the needs of users not connected to the Internet;
- (vi) define necessary structures for the maintenance and further development of the system;
- (vii) make the MEDI system compatible developed with initiatives developed by other programmes, e.g., GCOS, WCP, INFOTERRA.

The Participants in the Pilot Project:

The Pilot Project will be carried out by a Working Group composed of:

BODC (UK);
NODC (Russian Federation)
ISMARE (Ireland);
AODC (Australia);
Chair GEMIM;
Chair GETADE;
IOC Secretariat.

The activities of the Working Group will be co-ordinated by Dr. M.T. Jones, Director, British Oceanographic Data Centre.

(end of document)