SOFTWARE SYSTEMS TO MANAGE INTERDISCIPLINATY MULTIVARIABLE DATA AND METADATA OF LARGE ENVIRONMENTAL DATA SETS

V. VLADYMYROV[1], V. LYUBARTSEV[1], V. MYROSHNYCHENKO[2] and A. KONOPLYANIKOV[1]

[1] Marine Hydrophysical Institute, 2, Kapitanskaya St., Sevastopol, UKRAINE

[2] Institute of Marine Sciences, Middle East Technical University, Erdemli-Icel, TURKEY

ABSTRACT

The recent oceanographic and environmental projects operate usually with large interdisciplinary multivariable data sets. These data sets have to be quality checked, reliably stored, and analysed. It is a very complicated and time-consuming task to combine these interdisciplinary data and to provide the possibility to store and analyse them jointly. The industry standard database management systems are not often suitable enough to store and to maintain such data and to perform this task. Taking into account that data exchange is often strictly limited, it is necessary also to have the possibility to manage and distribute both data themselves and only metadata carrying all information on the entire data set.

Two special software tools were developed at the Marine Hydrophysical Institute (Sevastopol, Ukraine) to manage this kind of data and metadata. They are: OceanBase system - for data and Oceanographic Data Inventory – for metadata.

OceanBase is a generic name of a family of the comprehensive software system to work with large multivariable interdisciplinary marine and environmental databases. The main objective of this system is to provide easy, quick and effective work with interdisciplinary ocean and environmental data in oceanography and ecology. First version of OceanBase was developed by the Database Laboratory of the Marine Hydrophysical Institute (MHI, Sevastopol, Ukraine) in early 90-th on a basis of the FoxPro system and the own original package of software.

Second generation of the system was developed using Borland Delphi in 1996-1997 jointly with IMS-METU primarily for the purposes of the NATO TU-Black Sea project. This version of OceanBase works under Windows-9*, has a customized user-friendly multi-windows interface, and provides quick and comfortable access to the entire database. System allows to load, view, sort, select, process, combine, and export all necessary data and metadata. User can make a selection on institutions, data sets, time (year, month, day, hour), region, sea depth, measured variables, etc. It is possible to make data profiles smoothing and interpolating, to calculate mean profiles, spatial distributions, time variability, various statistics, etc. Data and information in the different windows are cross-linked and synchronized. Many graphic tools (plots, maps, and histograms) are embedded into the system. User can customize the plots to obtain the presentation quality pictures for scientific papers or reports. The system generates various reports as for selected data and the results of their processing as for the metadata related to the selected data.

The system permits to store a quality code with each data value and to use the quality criteria in the data selection. OceanBase can manage not only the standard type of oceanographic data (a value of a variable at the defined space-time point), but also the layer type of data that are obtained in a layer, generally, biological data, for example, zooplankton net tows. It gives the possibility to export selected data to files in different common ASCII formats to process and present them further in the user preferable software. Besides the NATO TU-Black Sea project, this version of OceanBase was also used in a number of national and international projects.

The recent version of OceanBase consists of three main applications: OceanBase Administrator, OceanBase Data Manager, and OceanBase Data Explorer (former OceanBase itself). The OceanBase Administrator is the major application of the OceanBase and it permits to run all main modules and tasks of the system, except the OceanBase Explorer. It is the main instrument of the database Administrator. The OceanBase

Data Manager module permits to organise work of the Administrator and the Data Managers on the data entry, editing, and quality control. The OceanBase Data Explorer is the main module for the end-users of the system. It provides the possibilities for data selection, access, and processing, but it does not give the possibility to change data loaded into the database.

The Data Managers can create, edit and delete records in the database. Almost all records keep the information on their creation and modification. The only Data Manager, who entered the data, is permitted to modify or delete them. Only the Administrator can change this reference. By default, Data Manager who created the record is responsible for its content.

The most valuable feature of OceanBase is the integration of many service tools that allow to the user to carry out almost all data processing without any additional software. This feature makes it very attractive for the end user.

One of the modifications of the OceanBase (version 2.02.PLANKTON) allows to manage marine biological data at the species level. Zooplankton and phytoplankton taxonomic classifications are stored in the database and used for selection of biological characteristics. Taxonomic classification is based on the *Integrated Taxonomic Information System* (ITIS Project, US). The OceanBase system allows one to select not only the biomass and abundance of individual plankton species but also to calculate summaries for taxonomic groups: from Genus and Family to total Plankton as the whole. To simplify this task, the special Plankton Query interface was designed, which represents taxonomic classification in form of hierarchical taxonomic tree. Each node of the taxonomic tree contains check box, which, when checked, results to calculation of summaries on the sub-tree. After selection, plankton data can be processed jointly with other parameters using full set of possibilities of OceanBase tools, such as mapping, plotting, histogram calculation, etc.

The recent version of the Oceanographic Data Inventory was developed in the framework of the Black Sea Ecosystem Processes and Forecasting/Operational Database Management System Project of the NATO Science for Peace Programme and based on the first version of inventory developed for the similar project in the framework of the NATO Science for Stability Programme.

The main aim of the Oceanographic Data Inventory is a search for metadata on various complex criteria, display found metadata on the map, and generate some reports on these metadata. It is possible to construct rather complicate search queries using coordinates, date, time, measured parameters, organizations, etc.

The Oceanographic Data Inventory has user-friendly interface, and can be run from CD-ROMs without any installation procedure. There are no any reasonable limitations on amount of stations and datasets stored in the Inventory. In contrast of the old version, the new one is well documented and supplied with a set of utilities allowing to tune it up to different regions of the World Ocean. The data import utility provides a user with a possibility to import metadata in different formats, check and edit them. The Inventory Map window has some simple GIS possibilities and permits to present effectively the selected metadata.

Both the OceanBase and Oceanographic Data Inventory systems are the very useful tools for the data and metadata users. They are now in use in many oceanographic and environmental institutions and projects in different countries around the World and they can be used in any oceanographic or environmental project or organization dealing with the interdisciplinary multivariable data.