

**INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION  
(of UNESCO)  
Twentieth Session of the IOC Committee on International Oceanographic Data  
and Information Exchange (IODE-XIX)  
Beijing, China, 4-8 May 2009**

**IODE Arrangements for the Long-Term Secure Archival of Data  
and Information**

**(Mr. Sydney Levitus, Mr Robert Keeley)**

---

Appendices:

A. *Full background*

**DRAFT TEXT FOR INCLUSION IN THE SUMMARY REPORT**

*This agenda item will be introduced by the Co-Chairs. It will address the future cooperation between ICSU's WDC system (or what it will become) and IODE and whether this new system will serve the needs of IODE.*

*The Committee will be reminded that historically IODE has had strong cooperation with the ICSU WDC system for the long-term archival of oceanographic data and information. Recent developments in technology and the announcement by ICSU that it is replacing the WDC system with a World Data System (WDS) suggests that the IOC/IODE should reexamine its requirements for the long-term archival of oceanographic data and information. In addition, it is more critical than ever that regional and global ocean scientific data sets and products based on those data sets be available in a timely manner for scientific assessments that guide national and international policy makers such as the IPCC.*

*The Committee will also be informed about discussions held during the World Data System Transition Team (WDSTT) at ICSU, Paris, 10 March 2009, to which IODE was invited.*

*The Committee will be requested to discuss and agree on the needs of the IODE network of data centres for long-term secure archival of data and information and identify ways and means to cover these needs.*

## **Appendix A: Full Background**

### **1. Introduction**

Historically IODE has had strong cooperation with the ICSU WDC system for the long-term archival of oceanographic data and information. Recent developments in technology and the announcement by ICSU that it is replacing the WDC system with a World Data System (WDS) suggests that the IOC/IODE should reexamine its requirements for the long-term archival of oceanographic data and information. In addition, it is more critical than ever that regional and global ocean scientific data sets and products based on those data sets be available in a timely manner for scientific assessments that guide national and international policy makers such as the IPCC. This document is intended to stimulate discussion.

### **2. Changes in the ICSU World Data Center system**

ICSU now has a new Strategic Goal which is “to facilitate a new, coordinated global approach to scientific data and information that ensures equitable access to quality data and information for research, education and informed decision making”. ICSU is now progressing from a collection of World Data Centers to a World Data System (WDS). Two of the immediate ICSU goals (among others) are:

- i) “Develop and implementation of a single data portal for locating and accessing all data and services [2009];
- ii) Accreditation process for all existing data centres and services to be completed [end 2009]”.

The quoted material above is from the “ICSU Ad hoc Strategic Committee on Information and Data: Final Report to the ICSU Committee on Scientific Planning and Review”.

### **3. The WMO Integrated Global Observing System**

The WMO Integrated Global Observing Systems (WIGOS) is “a concept for a comprehensive, coordinated and sustainable system of observing systems. WIGOS is based on all WMO Programmes’ observational requirements. It ensures availability of required data, products and information and facilitates access through the WMO Information System (WIS) according to identified requirements.

Benefits of WIGOS to Members and partner organizations will be improved services, increased quality, consistency and access to multi disciplinary observations, more efficient use of resources, better preparedness to incorporate new observing systems.”

([http://www.wmo.int/pages/prog/www/wigos/index\\_en.html](http://www.wmo.int/pages/prog/www/wigos/index_en.html)).

IOC/IODE is involved with WIGOS via a JCOMM Pilot project

([http://www.wmo.int/pages/prog/www/wigos/documents/PP\\_JCOMM.pdf](http://www.wmo.int/pages/prog/www/wigos/documents/PP_JCOMM.pdf)).

The ODP will become a Data Collection Production Center (DCPC) for WIS. The functions of a DCPC are to collect, disseminate, add value to, and archive regional or programme-specific data and products. DCPCs also produce regional or specialized information, and provide related data and products for international exchange. DCPCs maintain catalogues of their holdings and services, and appropriate parts of the catalogues also update a comprehensive catalogue of WIS holdings. The ODP will provide the ocean component to WIS.

### **4. New Technology**

The new technology referred to earlier to is the emergence of the Internet, the World Wide Web, and relatively inexpensive high-speed computers and data storage devices that are both faster and capable of storing more data than previously possible and at lower costs. These technologies allow for the possibility of distributed data centers. In fact:

- i) numerous countries (*e.g.*, Japan, Republic of Korea, Russia, United States among others);
- ii) international organizations and consortia (*e.g.*, ICES, SeaDataNet);
- iii) international research projects (*e.g.*, WOCE, JGOFS)

already distribute, or have distributed in the past, all or some of their data online.

A question though is how well these data sites will be maintained into the long-term future, particularly for research project databases.

## 5. Issues

The oceanographic data management community is developing distributed ocean data systems, as noted above. These include developing global distributed data systems enabling perhaps hundreds of nodes of small to large size to contribute data to the "global commons". This will help serve our user community but there are also inherent risks:

- Who can assure that each node remains alive forever?
- What happens if a node shuts down either willingly or through a disaster?
- How do our community provide a "long-term secure archival" service for these nodes even though some existing NODCs already have a system in place for "long-term secure archival".?
- If we have a distributed system like ODP or WIGOS/WIS then will this also provide data products or should this be a niche for a new WDC: producing global data products addressing global issues?

There is interest in whether IODE should have its own centers for the long-term archiving of oceanographic data including one or more WDCs for Oceanography and/or also use existing NODCs for this purpose, and/or virtual WDCs. Also, there are types of oceanographic data that do not presently have a long-term archive, in particular some chemical and biological data.

A WDS for Oceanography could be a network of distributed IODE centers making use of the Ocean Data Portal (ODP) technology to contribute data through a distributed network.

It is still not clear as to exactly what ICSU's intentions are. Given this uncertainty one view is that it is appropriate for IODE to have at least one WDC for Oceanography. It is not clear that IODE needs more than one WDC for Oceanography though there may be data types (*e.g.*, biological) or functions that only a specialized NODC with interest in a particular data type would be able to maintain. Thus, some long-term archiving might be performed by centers with assigned responsibilities and might be designated as "Responsible NODCs" similar to the structure that IODE had in the past. Such a center could be a WDC for specialized type of data or could be part of an IODE "Virtual WDC system".

## 5. Stewardship

Who will provide stewardship of ocean data archives? Stewardship is much more than just storing data. For example, as modern or near-real-time data are distributed to users via NODCs as part of a portal system or via a WDC, problems are often found with the data or associated metadata which are reported back to the submitting organization for correction and resubmission. For example the IODE/GTSPP serves this purpose for real-time data. This activity needs to continue with historical, archived data with the originator's submission being maintained and a corrected version of the data being made available as well. The originator's submission should always be maintained in case a "correction" to a historical data set turns out to be incorrect.

But not all data storing nodes and facilities on a distributed system have a long-term perspective. Having a base of hundreds of archiving and distributing nodes increases complexity and the chance of failure. Only a few nodes may be prepared for such a task.

## 6. Products

At present, WDC for Oceanography, Silver Spring produces the *World Ocean Database* (WOD) and the *World Ocean Atlas* (WOA) series. WOD is a global collection of historical and modern ocean profile data of mainly physical and chemical data and includes some plankton data. The WOA is a set of objectively analyzed gridded fields of some of the variables in WOD. These products have proven to be very useful to the scientific community. For example scientific studies based on these products were a prominent part of the chapter on "Chapter 5: Observations: Oceanic Climate Change and Sea Level" in the Fourth IPCC Assessment. Enabling studies of the role of the ocean as part of earth's climate system depends on the availability of the most comprehensive ocean data sets possible.

In the case of WOD and WOA there is offsite backup of the products every month. In addition, when these products are published on DVD every 3-4 years, a few hundred sets of these DVDs are distributed world-wide. This represents another backup. Measurements of ocean data are priceless because we obviously can not go back in time to make more measurements. A simple computation shows that in today's dollars, the cost of ship time alone to have made these measurements is several billion U.S. dollars. This fact also justifies the relatively small cost of providing long-term archiving of these data.

Who will prepare crucial products for scientific research and operational forecasting groups and for international assessments such as those produced by the IPCC.

## 7. A Suggested Direction for IODE

The WDCs-Oceanography have been partners with IODE for many years to secure the marine data collected by individual nations for use by future generations. The relationship between WDCs and IODE was established many years ago. With the changes taking place in ICSU this provides an opportunity to redefine mutually supportive roles and responsibilities. Any changes will have to operate within the borders of new ICSU structures and support activities of IODE member states.

This is a complex task and not one that can be decided quickly. Equally, the exact nature of the changes in ICSU operations is not yet clear. It is recommended that the IODE Committee take the following action.

1. Form a working group composed of one or both IODE co-chairs, up to 3 other volunteers from NODCs, and invite participation of the Directors of the 3 WDCs-Oceanography and the chair of GEMIM to provide their perspective.
2. The WG is instructed to produce a report with recommendations to IODE on a new relationship between NODCs and WDCs-Oceanography taking into consideration such issues as:
  - The changing ICSU structure
  - Changes in technology for managing data and information (e.g. distributed data provision, versioning, discovery)
  - A consensus view of the complimentary and value-adding roles of NODCs and WDCs.
  - Securing valuable data and information into the distant future
  - The contributions of research project and other data sets assembled outside of NODC or WDC influence.
  - Managing the present and increasing volumes and diversity of marine observations.

A draft of the report should be ready for discussion by the end of 2009 and be circulated to IODE comment.

**ACRONYMS:**

DCPC: Data Collection Production Center

ICES: International Council for Exploration of the Sea

ICSU: International Council of Science

IPCC: Intergovernmental Program on Climate Change

ODP: Ocean Data Portal

WDC: World Data Center

WDS: World Data System

WIGOS: WMO Integrated Global Observing Systems

WIS: WMO Information System

WMO: World Meteorological Organization

WOD: World Ocean Database

WOA: World Ocean Atlas

WOCE: World Ocean Circulation Experiment

JGOFS: Joint Global Ocean Flux Study