WORLD METEOROLOGICAL ORGANIZATION

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION (OF UNESCO)

JCOMM/IODE Expert Team on Data Management Practices (ETDMP)

First Session

Hosted by Flanders Marine Institute

Oostende, Belgium
15-18 September 2003

UNESCO

Abstract
The 1st Session of the JCOMM/IODE Expert Team on Data Management Practices (ETDMP) was held in Oostende, Belgium between 15 and 18 September 2003. The ETDMP members discussed the requirements for end-to-end data management (E2EDM); existing and planned data management mechanisms and practices; cooperation with other programmes and expert teams; the strategy for the development of E2EDM; and future cooperation with the Ocean Information Technology (OIT) Pilot Project. The Group agreed on an Action Plan for the intersessional period based on three pilot projects identified by the sessional working groups: metadata management, data assembly, quality control and quality assurance, and the development of an E2EDM Prototype.

TABLE OF CONTENTS

ANNEXES

1. ORGANIZATION OF THE SESSION

1.1 OPENING OF THE SESSION

The First Session of the JCOMM/IODE Expert Team on Data Management Practices (ETDMP) was opened by the Chair of the group, Nick Mikhailov, on 15 September 2003 at 0900 at the offices of the Flanders Marine Institute (VLIZ), Oostende, Belgium. The Chair welcomed the members and noted that Elenor Gowland had replaced Volker Wagner. He also noted that the members of the IODE Group of Experts on the Technical Aspects of Data Exchange (GETADE) had joined ETDMP as a result of the merger of the two groups. The list of Participants is included in this report as Annex II. For the joint Secretariat, Peter Pissierssens welcomed the members. He stated that the Group should agree on concrete activities which could form the basis of future pilot projects.

1.2 ADOPTION OF THE AGENDA

The group adopted the agenda for the session as given in Annex I.

1.3 WORKING ARRANGEMENTS

The group agreed its hours of work and other practical session arrangements.

2. REVIEW OF ETDMP WORK PLAN AND ACTIVITIES

The Chair referred to the Informal Session of ETDMP which was held in Brussels, Belgium, on 28 November, 2002. This Session discussed the JCOMM strategy for end-to-end data management and the document entitled "The Basic Elements of the End to End Data Management Strategy". The ETDMP draft work plan for 2003-2004 was reviewed and the Session discussed an extensive analysis of the tasks and assigned their implementation to members of the Group present during the Session.

Greg Reed from the IOC Secretariat, and former Chair of the IODE Group of Experts on the Technical Aspects of Data Exchange (GETADE) reported on the recommendation of IODE-XVII that the GETADE be merged with the JCOMM Expert Team on Data Management Practices. IODE-XVII also recommended that the funds allocated to the IODE programme for the organization of GETADE sessions be assigned to the organization of ETDMP sessions, thereby assuring annual sessions of ETDMP. The main task of the GETADE work plan is the *"The development of an End-to-End Marine Data Management Framework"*. This mission is in accord with the terms of reference of ETDMP.

3. REQUIREMENTS FOR END-TO-END DATA MANAGEMENT

3.1 GCOS REQUIREMENTS

This agenda item was introduced by Bob Keeley. He showed examples of analyses he had put together examining how well the data reported in real-time met stated sampling requirements for climate. These requirements are contained in the Oceans Observation System Development Panel (OOSDP) final report "Scientific Design for the Common Module of the GOOS and the GCOS: An Ocean Observing System for Climate" and in the results of the Ocean Observations '99 Conference. As an illustration he used the Argo goals of one temperature and salinity profile every 300 km every 10 days in an evaluation of temperature profiles and SST. He showed the contributions to the overall goal through the different instruments and sampling now taking place over the past 12 months. As one component of the observing system, the data system must be able to acquire and move the data to users sufficiently rapidly to meet their needs. He mentioned that it is just such an evaluation that will be used to determine how well the climate observing system is meeting its targets.

3.2 COOP REQUIREMENTS

Catherine Maillard described the data types associated with the different categories of GOOS applications. Observation systems and parameters include coastal buoys, bottom observatories, repeated sections, areas investigation and fisheries monitoring. COOP has a list of recommended common physical, chemical and biological variables to be measured as part of the global coastal system (Reference; GOOS Report No 125. *The Integrated, Strategic Design Plan for the Coastal Ocean Observations Module of the Global Ocean Observing System*).

3.3 MMS REQUIREMENTS

The MMS requirements for end-to-end data management were introduced by Elanor Gowland. The MMS data management rules have been outlined by the WWW data management group and MMS works to these standards and within the regulations set by Commission for Basic Systems (CBS), and other WMO bodies and any end-to-end data management rules will need to consider the integration of the current MMS strategy. Marine Meteorological Services data management covers acquisition, dissemination, quality control, archival and availability of data on differing time scales. MMS provides data acquisition systems (e.g. VOS scheme), data dissemination schemes in real time via GTS and in non real time (MCSS). MMS provides (i) marine meteorological services for the high seas, (ii) marine climatology, (iii) marine meteorological services for coastal and offshore areas, (iv) marine meteorological services for main ports and harbour areas, (v) the WMO Voluntary Observing Ships Scheme, and (vi) training in marine meteorology.

The data management needs of MMS can be summarised as (i) meteorological and oceanographic data and products, (ii) timely data streams of uncorrupted data in real time and non real time, (iii) feed back and follow up action in case of erroneous data or data flow, (iv) metadata with each observation (VOSClim), with each VOS (WMO-No 47), and ODAS, (v) uniform data processing standards, formats and QC principles, and (vi) a reference database

The items which need to be considered with respect to a JCOMM E2EDM strategy are (i) data quality (quality control, site information and software), (ii) safety of data during the transmission process (e.g. corrupted data on GTS, data sinks), (iii) real time transmission additional variables (e.g. VOSClim additional. new parameters), (iv) availability of: real time and non real time data, (v) timeliness of data availability, (vi) improvement of the accounting system, directly affecting the data availability (actual SOT Task Team on Satellite Communication Costs), (vii) unique and complete database at agreed levels, (viii) tagged data (no pseudo duplicates), and (ix) linkage to the global JCOMM E2EDM space. There are also other user groups dealing with marine climatological data, sometimes holding and extracting separate databases for their purposes (e.g. ICOADS to support the marine climatological research community, or other special data holdings from research ships). The needs of those projects should be considered in constructing a global JCOMM E2EDM.

3.4 GCOS/COOP/MMS REQUIREMENTS FOR SATELLITE AND SEA ICE DATA

This item was introduced by Takashi Yoshida who referred to the following documents:

- The Second Report on the Adequacy of the Global Observation System for Climate in Support of the UNFCCC. (GCOS, April 2003)
- The Integrated Strategic Design Plan for Coastal Ocean Observations Module of the Global Ocean Observing System. (UNESCO 2003)
- The reports of CBS/OPAG-IOS Expert Team on Observational Data Requirements and Redesign of the Global Observing System

GCOS Requirements. GCOS required variables are (i) ocean domain variables (SST, sea level, sea ice, ocean colour), and (ii) atmospheric variables over the ocean (surface wind, precipitation). Accuracy, continuity and homogeneity of satellite data should be ensured for the purpose of climate monitoring. Data management issues include (i) long-term archive of original observations during the whole period of each satellite mission with adequate metadata through continual migration to newer storage devices and access software, (ii) access to adequate complementary *in situ* baseline measurements, (iii) use of modern information and communication technology, and (iv) timely delivery of data to the user community and international data centres.

COOP Requirements. The COOP strategic plan identifies the common variables required in order to achieve the goals of the coastal module of GOOS. In the process of identifying the common variables, remotely sensed properties of surface waters were distinguished as the variables to be measured as part of other observing systems and to be shared with the coastal module. Five variables: temperature, salinity, elevation (currents), roughness (wave) and ocean colour (chlorophyll and attenuation of solar radiation) are the COOP requirements for satellite observations. The COOP design plan indicates that existing data management systems function well in the open sea where the required spatial scales of resolution are coarse (> 10 km) relative to coastal water (<1 km). There are serious challenges and gaps in the international institutional structures for data management for coastal waters. The involvement of, or linkage to, the satellite data management systems established by space agencies into the integrated data management system is essential to meet the requirement for much broader spectrum of variability in the coastal waters.

MMS Requirements. Four major application areas which require ocean observations have been identified by the Expert Team on Observational Data Requirements and Redesign of the Global Observing System (ET-ODRRGOS) of the CBS Open Programme Area Group (OPAG). These are (i) Numerical Weather Prediction, NWP, (ii) Ocean Weather Forecast, (iii) Seasonal to Inter-annual Forecast, and (iv) Coastal Marine Services. Satellite observation variables required in those areas are sea/ice surface temperature, sea ice cover, sea-surface wind, wave, ocean salinity, and ocean colour with high temporal and spatial resolutions. Information on coastline change provided with very high-resolution visible/infrared imagers is also a requirement of the Coastal Marine Services. Timeliness of the data distribution from these satellites is particularly important for NWP. All data sources are required to be accompanied good documentation, careful QC, and monitoring. In situ measurements should be distributed in a timely manner to calibrate satellite measurements appropriately.

In summary, the GCOS/COOP/MMS requirements are (i) timely distribution of satellite data in real-time, near-real-time and delayed modes, (ii) to ensure access to long-term archive of satellite data and metadata, (iii) to ensure access to real-time in situ observations data flows and its storage, (iv) use of modern information and communication technology to ease the burden developed from the expanding data amount, (v) good documentation, careful QC, and monitoring, (vi) coding standards, (vii) integration of in situ and remote sensing data, and (viii) timeliness of the data distribution from these satellites is particularly important for NWP.

3.5 REGIONAL GRA REQUIREMENTS

This item was introduced by Roger Djiman referring to Document ETDMP-I/12. The GOOS Regional Alliances (GRAs) will focus on observations of common national or regional interest and accordingly, to gain national support, GRAs are composed of National Agencies and Organizations committed to the regional implementation of Global Observation networks. Depending on national and regional priorities, GOOS Regional Alliances may increase the resolution at which common variables are measured, supplement common variables with the measurement of additional variables, and provide data and information products that are tailored to the requirements of the stakeholders in the respective regions. Thus, GRAs both contribute to and benefit from the global network. GRAs could provide the mechanism for coordinating requirements for data and products among the GRAs, LMEs and IOC and WMO regional organizations. Such coordination is essential to reduce duplication of effort and to harmonize the use of observations among various regional users having distinct requirements. GOOS Regional Alliance as a group could complement the activities of JCOMM. A group such as the Regional GOOS Forum could provide the mechanism to establish trans-regional user requirements, specify associated required products, coordinate measurements and sharing of technical information among GRAs, assess the effectiveness of the system and recommend needs to JCOMM.

Variables to be considered by the GRAs will be important for detecting and predicting change in coastal systems, but may not be appropriate for global implementation. These include (i) variables of regional significance that are not global in extent, and (ii) categories of variables that would be defined or measured differently depending on geographic location.

4. REVIEW DATA MANAGEMENT MECHANISMS AND PRACTICES

4.1 METADATA MANAGEMENT SYSTEMS

4.1.1 Review of Marine Metadata Systems

This item was introduced by Ricardo Rojas, referring to Document ETDMP-I/10. A number of marine-related metadata systems were discussed including MEDI, GCMD, EDMED, EDMERP, EDIOS, SEA-SEARCH, ISO 19115 and ODAS. Each of these systems are summarized as follows:

MEDI. The Marine Environmental Data Inventory (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) programme. MEDI has been developed by the IODE Steering Group for MEDI and uses the Directory Interchange Format (DIF) format which includes a well defined Document Type Definition (DTD). Valids are well defined and maintained by NASA/GCMD. MEDI is fully compatible with NASA's GCMD and is ISO 19115 compliant.

GCMD. The Global Change Master Directory (GCMD) is a large directory of descriptive information about data sets relevant to global earth sciences and climate change research. At present there are over 11,000 of metadata descriptions of datasets. GCMD is managed and maintained by the US National Aeronautics and Space Administration (NASA). All metadata descriptions in the GCMD are reviewed by GCMD metadata specialists for conformance with the GCMD controlled vocabulary and structure. Valids are well defined and maintained by NASA/GCMD and there is a well defined DTD.

EDMED. European Directory of Marine Environmental Data (EDMED) is a computer-searchable directory of dataset descriptions relating to the marine environment. It covers a wide range of disciplines including marine meteorology; physical, chemical and biological oceanography; sedimentology; marine biology and fisheries; environmental quality monitoring; coastal and estuarine studies; marine geology and geophysics. EDMED entries can be submitted in a simple free-text format.

EDMERP. The European Directory of Marine Environmental Research Projects (EDMERP) is a directory of European research projects relating to the marine environment. As in EDMED, EDMERP covers a wide range of disciplines including marine meteorology, physical, chemical and biological oceanography, sedimentology, marine biology and fisheries, environmental quality, coastal and estuarine studies, marine geology and geophysics, etc. Research projects are catalogued in EDMERP as fact sheets or abstracts with their most relevant aspects. The primary objective of EDMERP is to support users in identifying interesting research activities and in connecting them to involved research managers and project results like data, models, publications, etc. across Europe.

EDIOS. The European Directory of the Initial Ocean-observing System (EDIOS) is a searchable directory of the ocean observing, measuring, and monitoring systems operating in Europe and is an initiative of EuroGOOS.

Sea Search. The Sea Search website provides an effective navigation tool to data and information sources in Europe, to oceanographic data and information, managed by European centres, and to centres in Europe with expertise and skills in oceanographic and marine data and information management.

ISO 19115. The ISO 19115 metadata standard was developed by the ISO Geographic Information/Geomatic Technical Committee 211 (ISO/TC 211). The ISO 19115 metadata standard provides a structured framework for describing a geospatially dataset for use in search and discovery tools. Virtually all of the descriptive elements found in any of the metadata description systems discussed above can be mapped to one or more of the metadata elements in the ISO 19115 metadata standard. The WMO/CBS Expert Team on Integrated Data Management (ETIDM), in November 2001, recommended the adoption of ISO 19115 by WMO as the WMO metadata standard.

ODAS. The Ocean Data Acquisition System (ODAS) Metadata format had been developed within the framework of the former Commission of Marine Meteorology (CMM) in order to enable the development of a metadata base for ODAS including moored and drifting buoys, offshore platforms, etc. That database will allow an accurate interpretation of the observational data from ODAS that were available in climatological archives. At its first session in Akureyri, Iceland, 19-29 June 2001, JCOMM recommended that the format agreed upon by its sub-group on Marine Climatology be used as the global format for the assembly, exchange and archival of metadata from all types of ODAS, including, in particular, drifting and moored buoys and fixed platforms.

Three alternatives were suggested for developing a common metadata standard for JCOMM:

a. Choose the most-used system by the scientific data management community. However, there is no dominant metadata standard at present in use by either the Oceanography or Meteorology community..
b. Compare all the systems and map their crosswalks and then choose the one that best fits the needs of the entire community. These crosswalks have been created by many individual partners, such as the GCMD which maintains a web page mapping DIF to ISO, FGDC to DIF mapping, Dublin Core to DIF mapping, ANZLIC to DIF, etc. but this has been done just in order to make Directories and databases compatible, not necessarily to integrate or develop a common standard.
c. Start over from the beginning, defining all the elements and attributes necessary to create a new Metadata Directory. This is not a recommended course of action, due to time and resources constraints as well as the availability of several metadata standards that can be used 'as is' or with minor, acceptable modification.

Metadata can be divided in two levels (granularity):

- Level 1. General information describing common features of all data files or databases (title, content, reference system, distribution system and metadata reference information);
- Level 2. Detailed data information describing different datasets (e.g. observational data, integrated data and data products)

All the metadata systems listed above can be described as Level 1 with the exception of ODAS, which is Level 2.

Lesley Rickards raised the question of the role of GOSIC and how this site contributed to the management of metadata of the global programs.

Further discussion on the crosswalks between the different metadata models is detailed under agenda item 4.1.4.

4.1.2 ODAS Metadata

This item was introduced by Lin Shauhua, referring to Document ETDMP-I/11, outlining the implementation plan for the construction and operational running of the ODAS Metadata Management Centre for the receiving, processing, management and service of global surface current data holdings and services.

Prof. Lin recalled that JCOMM-I had recognized the urgent need to identify a centre willing to host the ODAS metadata base. Subsequently, at DMCG-I, the National Marine Data and Information Service, Tianjin, China offered to host the ODAS metadata base.

The ODAS centre will undertake the collection, sorting, processing, management and service of the ODAS metadata from members states, international organizations and cooperative projects and programs and to carry out it long-term operational running, maintenance and service. The practical tasks consist of the followings:

- To extensively conduct the comparative study, based on the ODAS metadata formats, with current metadata standards in use internationally, timely follow up the developing requirements of oceanographic and meteorological metadata information management, maintain and update the ODAS metadata format so as to keep them in present situation;

- To conduct the collection, processing and management of ODAS metadata, develop the program modules for the metadata processing, operation and management, set up the metadata base and its management system, carry out long-time maintenance and updating of it, and regularly provide metadata inventory to the Members and Member States and the relevant organizations;
- To establish and maintain the website for ODAS metadata management centre, develop the release and navigation services of ODAS metadata and provide related information on JCOMM metadata management;
- To follow the development of ODAS, timely complement, perfect, review or work out relevant standards and codes, and provide them to the Members and Member States and the relevant organizations after approved by the expert group;
- To set up and keep the relationships with the JCOMM Members and Member States, international organizations and cooperative projects by the coordination of JCOMM data working group, promptly collect and deal with the demands and suggestions of offers and users of metadata information, and improve and perfect the management and service of the metadata centre;
- To regularly submit the work report of the centre to the DMCG and ETDMP.

The construction of the ODAS metadata management centre will start in October 2003. It is planned to complete the system construction by the end of 2004. The detail work plan is as follows:

- Investigation and collection of ODAS sources will start from October 2003. The investigating targets are mainly the JCOMM members and Member States and the relevant international organizations and cooperative projects. It is suggested that a formal letter will be sent to them with the requirements of the tables designed by the NMDIS, and also suggested that IODE and ETDMP strongly support the work.
- Processing of ODAS metadata and development of relevant operating tools will be completed by April 2004. The operating tools will be developed for the operational running on the basis of the existing format of ODAS metadata.
- In April 2004, the service website for the ODAS metadata management centre will open to provide the network collecting tools and the report on work progress and other relevant information.
- In October 2004, the construction of ODAS metadata base and the development of operational management system will be completed. At the same time, the information release on the website and the navigation testing will be conducted
- In December 2004, basis construction of the operational system will be almost completed. The ODAS metadata centre will try to go into its operational running. The work report will be submitted to the ETDMP and JCOMM Management Committee.

NMDIS will be responsible for the operational running of the ODAS metadata management centre, the long-term collection, processing and management of metadata, and the website maintenance. Users will be able to access all ODAS metadata on the centre's website and the centre will periodically distribute the metadata inventory (by FTP, email or CDs) to users. NMDIS will cooperate with DBCP to acquire samples of the information to be stored in the ODAS to aid their development activities.

The Group thanked Prof Lin and her organisation for their efforts in developing and hosting the ODAS Centre.

4.1.3 IODE activities

This item was introduced by Lesley Rickards, Chair of IODE, who reviewed current IODE activities. The main objectives of the IODE Programme are (i) to facilitate and promote the exchange of oceanographic data and information; (ii) to develop standards, formats and methods for the global exchange of oceanographic data and information; (iii) to assist Member States to acquire the necessary capacity to manage oceanographic data and information and become partners in the IODE network; and (iv) to support international scientific and operational oceanographic programmes of IOC and WMO and their sponsor organizations with advice and data management services. Some important developments include the co-operation with GOOS and JCOMM, the formation of a new Group of Experts on Biological and Chemical Data Management and Exchange Practices (GE-BCDMEP), the new Global Ocean Surface Underway Data (GOSUD) Pilot Project, progress with the Marine Environmental Data Inventory (MEDI) metadata system, and the developments for a marine XML through the EU-funded MarineXML Project and the ICES-IOC Study Group on the Development of Marine Data Exchange Systems using (SGXML).

Cooperation between IODE, JCOMM and GOOS was seen to be an important development. The IODE data centres are an integral part of GOOS (e.g. IODE centres are involved with Argo, GLOSS, etc.) and IODE is a co-sponsors of Ocean Information Technology (OIT) Pilot Project. The merger IODE Group of Experts on the Technical Aspects of Data Exchange (GETADE) with the JCOMM Expert Team on Data Management Practices (ETDMP) will result in more effective use of resources and avoid duplication.

4.1.4 Metadata Integrity Issues

This agenda item was introduced by Don Collins, who discussed three metadata models: ODAS, MEDI/DIF and ISO19115. ODAS Metadata is the most specific and is considered to be the closest to ‘syntactic’ metadata. It is not yet widely adopted or in use. MEDI/DIF is the most generic and contains discovery/descriptive metadata. It is widely used by many communities. ISO19115 is a combination of generic and specific information with descriptive and syntactic elements. This standard is new and not yet widely used.

The granularity of each of the models results in three levels of documentation:

- Discovery - most general level of information: MEDI/DIF or ISO19115
- Request - more specific level of information: ISO19115 or MEDI/DIF
- Use and management - most specific level of information: ODAS Metadata and/or ISO19115

Crosswalks have been constructed between these different models and they reveal the following:

ODAS → MEDI/DIF

- Very little commonality due to different purposes
- Reduce redundant effort when possible
- Aggregate information for collection of data from multiple ODAS to single MEDI/DIF

ODAS → ISO19115

- More commonality due to additional detail found in ISO19115
- Reduce redundant effort when possible
- Aggregate information for collection of data from multiple ODAS to single ISO19115 description

ODAS → other metadata standards

- Initial examination shows closest match with Cruise Summary Report information, perhaps also with EDIOS

Two questions regarding metadata integrity were raised:

1. Do we need another ‘standard’ or should ODAS elements be mapped to ISO19115 with extensions (if needed)?
2. Can current software (based on mappings, e.g. Z39.50) provide enough information management support or should ETDMP recommend investigation into semantic web or other emerging technologies?

The group agreed that the following additional investigation was required:

1. Crosswalks.

   a. Review and revise the ODAS/MEDI crosswalk
   b. Complete and review the ODAS/ISO19115, ODAS/EDIOS, EDIOS/ISO19115 crosswalks
   c. Work with ET-IDM to map ODAS to ISO19115 'core' elements
   d. Document 'optionality' of ODAS elements
2. Controlled vocabularies
   a. Coordinate and map ODAS coded terms to matching GCMD valids terms
   b. Review and assess SGXML Parameter Dictionary structure for applicability to JCOMM requirements
3. Integrity issues
   a. Review and assess how to aggregate ODAS Metadata into MEDI/DIF or ISO19115 metadata descriptions
   b. Develop data flow model from observer → ODAS→ MEDI/DIF→ ISO19115

The Group agreed that the development of a JCOMM metadata model was a high priority and the sessional working group would develop a proposal for a pilot project. The action items arising from the discussions on metadata would be identified in this pilot project.

4.2 OCEANOGRAPHIC DATA MANAGEMENT

4.2.1 GTSPP, Argo and GOSUD data management

This item was introduced by Bob Keeley, referring to Document ETDMP-I/3. He provided an overview of four different programmes: Global Temperature Salinity Profile Project (GTSPP), Surface Drifters (SVP), Argo, and Global Ocean Surface Underway Data (GOSUD) Project. He reviewed the data collection programmes for SVP (data collected from surface drifters), GTSPP (T&S profiles), Argo (data from profiling floats) and GOSUD (surface underway data). All four programs recognize the need for good connections between the data collection, data management and science communities. For each of these programmes, he discussed (i) cooperation between data collectors, data management and the science community; (ii) duplicates management, (iii) data distribution, (iv) QC standards, (v) loss of information, (vi) documentation, and (vii) data system organization. He informed the Group about an experiment underway to use a Cyclical Redundancy Check (CRC) calculation as a way to uniquely tag data and allow matching real-time and delayed mode versions of the data."

4.2.2 Integrity issues of oceanographic data management

This item was introduced by Bob Keeley, referring to Document ETDMP-I/4. The key data management elements required to provide homogeneous data and products to users are (i) documentation, (ii) data version control, (iii) quality control practices, (iv) data transfer, (v) monitoring, and (vi) products.

Documentation. The suggested minimum set of required documentation and Internet facilities to be used for presenting information are (i) data flow from collection to user, (ii) quality control procedures to a level of detail that allows users to assess the suitability to their needs, (iii) guidance about how to determine the version of the data, (iv) techniques and formats used to deliver data to users, and (v) who to contact when questions arise.

Data Version Control. One problem occurs when creating a version of data immediately after collection that is used for real-time distribution. There are two possible solutions: (i) don't make the real-time data any different from the original data, or (ii) if you do create a different version, provide for a unique label that can be connected to the original data. Another problem occurs in the processing of the data. It is important to document the various value adding processes such as quality control, as data is processed. This can be done by (i) recording the processing steps in something like the history structures of GTSPP and Argo, (ii) using a single tag that embodies the significant differences between versions, or (iii) keeping separate copies of the versions. The recommended method is (i) as it is robust, can record a large degree of detail if wanted, and deals with both initial processing and reprocessing. Another significant problem of data versions result from the present practices of data exchange. Each of us physically move data from external sources to our own institutions, reformat the data, add and subtract parts we deem relevant and then send these data to others. Such practices confound the user with multiple copies and differing content. It is recommend that the agency first accepting the data into the data system build a unique identifier and attach that to the data. Thereafter, every institution handling data must preserve this identifier so that this first point of origin always moves with the data.

Quality Control Practices. It is important that there be agreement on standards for recording what has been done and for recording the results. A recommended approach to data flagging is (i) adoption of the original IGOSS quality-flagging scheme. This uses the digits 0-5 to indicate the degree of confidence in the measurement., (ii) every measured value be assigned a quality flag, (ii) tests performed and tests failed information closely accompany the data, and (iv) further study is required and develop a standard.

Data Transfer. Possible solutions to the data transfer process include (i) stop exchanging data between data centres and instead build a system that permits remote data access, or (ii) exploit an existing strategy for data exchange (such as OPeNDAP) or develop a format structure that is self-describing (e.g. BUFR or XML) and sufficiently flexible to accommodate any data and information.

Monitoring. Every data system should implement appropriate monitoring as part of their routine processing. A starting point for what to consider for monitoring could be (i) losses of data during collection, transmission, processing, (ii) incorrect assignments of quality indicators, (iii) faulty processing software, (iv) faulty data entering the data system, (v) slow acquisition of data into the data system, (vi) incomplete capture by the data system, and (vii) insufficient volume of data collection resulting in poor products. Every data system implement appropriate monitoring as part of their routine processing.

Products. Products and users should be tightly coupled. If a product is used by another person or organization in a routine way, its existence should be advertised. Due credit for data collection and value adding should be given with every product.

A study of required standards was suggested. This study should look at:

a) Define a minimum set of documentation that should be available for any data system.
b) Define a universally applicable scheme for uniquely tagging data so that versions are readily identified.
c) Process tracking is a way for the data system to be able to tell a client what has been done to the data. It may be that such level of detail is more than required by a client and a simpler scheme can be employed along the lines used in the satellite community to indicate "levels of processing".
d) It is important to know the original source of data when trying to correct problems. Including such information may be considered part of the unique data identifier, may be considered as part of a process tracking scheme, or a separate issue.
e) A debate is required to determine if there is value in archiving measurements in all cases, or if there are some cases in which it is sensible to delete measurements.
f) A number of schemes for marking the quality of data are now employed. This causes confusion for users. It would be far better if agreement could be reached on what and how this should be done.
g) Although the term "data integration" is used extensively, it is interpreted to mean many different things. It is important to clarify what this term means for data systems.

A number of these actions have similarities or overlaps to the Ocean Information Technology Pilot Project and ETDMP should work with OIT to ensure that we have representation and our requirements are being met.

4.3 MARINE METEOROLOGICAL DATA MANAGEMENT

4.3.1 VOS marine meteo data management

This agenda item, introduced by Elanor Gowland, compared the data management processes of the Marine Climatological Summaries Scheme (MCSS) to those of the VOSClim project. The VOS programme provides basic marine meteorological *in situ* data from the world oceans and adjacent waters. It serves the needs of MMS as well as the requirements of the WWW for basic data input to all kinds of weather products and services. It also provides a most important *in situ* database for marine climatological monitoring and research. The basic steps in VOS data management are:

Data generation. Meteorological observations (Obs) are generated through the observers on board recording the basic elements as requested by WMO FM13-XI-Ship Code. Rules for creating, formatting and transmitting observations are laid down in instructions for observers, provided by the National Meteorological Services and disseminated through their national PMO networks. The Ship Code requirements are documented in the Manual on Codes (WMO-No 306) as endorsed by WMO/CBS, and are issued to ships either in paper format (eg logbooks or log sheets) or in computer software format.

Data Transmission. Obs are transmitted via an Inmarsat-C terminal to a land station (LES) and routed onto NMS. Problems can occur when observations are transcribed from hand written journals prior to transmission. New transmission links from ship to shore using email are being developed. All observations are stored on board the ship and are forwarded to the NMSs in delayed mode, through PMOs

Data Dissemination. Satellite data received at LES are sent onto allocated NMS, who put Obs on GTS. Delayed mode data dissemination is regulated by MCSS and data is returned to the NMS which recruited the ship, processed, and sent onto GCCs for quality monitoring and quarterly re-dissemination to Responsible Members.

Data Quality. Reliability of data is dependent on good instrumentation, site, observers and maintenance. Formatting errors and transmission problems occur, but these should be minimal, especially for automatic systems. Delayed mode records are more complete (but handwriting can cause problems). MCSS specify the minimum MQC standards which the data should meet. In case of failures corrections are requested.

Data Archival. Real time data, circulating on the GTS are archived by those NMSs which have the access to the GTS. Delayed mode data is archived by RMs through MCSS. There are no data tagging procedures in place to resolve the problem of duplicates i.e. to provide a clear indication of which was the original observation and what were the changes made to realize any further quality level.

Data Availability. There is no single source that contains all data holdings. There is no commonality of data quality levels between different databases (changes made after MQCS applied) and different data is available on different time scales. GTS data is available “hours” after observation made, however, delayed mode data is not available until “years” after the observation (range: few months ago - 10 years old).

VOSClim is a sub-set of VOS and data management follows the same lines, with some additions, (i) extra metadata to WMO-No 47, now extended to all VOS ships, (ii) additional met information with each observation (delayed mode only), (iii) special monitoring criteria applied to each observation, (iv) same dissemination route as VOS (via GCCs), (v) a dedicated VOSClim project website, run by the DAC, stores all available data from VOSClim ship.

With respect to a JCOMM E2EDM, the following areas could be considered:

- Data generation. Reporting of VOSClim additional data by all VOS.
- Data transmission. Sending extra fields and metadata in real time, new transmission codes (BUFR, XML, etc), worldwide accounting system for fairly sharing transmission costs, security (no corruption).
- Data dissemination. Speed up delayed mode data, better availability of GTS and MCSS data, reducing the time gap between real time and delayed mode.
- Data quality. Agreement on standards (instrumentation / methodology) identified by VOSClim, standardization of quality checks at observing sites, all necessary metadata

available with each observation, reliable maintenance regimes for automated data acquisition systems

- Data Archival. All data readily available (like from VOSClim DAC website) but not necessarily from a central source, data tagging, to avoid problems with duplicate data, transparent quality levels ensuring there is a complete “original “ set of observations.

The data management of the marine meteorological data stream can be summarized in three areas: data flow quality control and access to uncorrupted database. Improvements must be considered in a JCOMM E2EDM context, liasing with ETMC, VOSClim project, SOT and CBS.

4.4 NON-PHYSICAL DATA MANAGEMENT

This agenda item, introduced by Catherine Maillard, referring to Document ETDMP-I/13, which discussed the current conditions of non-physical data management and listed the necessary practical tasks to implement and to meet the needs of GOOS. The data management and communication sub-system requirements for GOOS were reviewed. The Coastal Ocean Observations Panel (COOP) data and information management issues relate to the following themes (i) coastal marine services (e.g., safe and efficient marine operations, coastal hazards), (ii) the health of marine and estuarine ecosystems and its relation to human health, and (iii) living marine resources.

Non-physical observations are in general more complex. The monitoring of non-physical data requires the availability of many more variables including nutrients, chemicals organic compounds and living species in the water column, the suspended matter, the sediment and the biota, and few standards exists for them. The methodologies for data collection and processing are often not compatible from one dataset to another. It was noted that the archiving, management and distribution of non-physical variables, parameters and descriptive information is less advanced than the physical data.

The number of non-physical parameters managed will be limited by the necessity/priority of using methods that provide low-cost measurements at large scale, and these parameters may not correspond exactly to GOOS priorities. As coastal monitoring is based on national and regional priorities, it may be more difficult to impose the IOC and WMO data policy.

The first step in the implementation plan for non-physical parameters would be to agree on a preliminary list of selected parameters that can be collected automatically and released in real time, and for which some first sets of integrated products according to GOOS priorities can be prepared. The selection should consider the following existing systems (i) on going programmes like the Voluntary Observing Ships (VOS) for surface and upper ocean data, important to follow the global distribution of sources and sinks for atmospheric carbon dioxide and the carbon exchanges within the interior of the ocean, (ii) the coastal zone monitoring, (iii) the fisheries and living resources following up. For the first list of selected parameters, standards and first integrated products should be developed as a pre-operational phase. After review of the results, the method could be shifted to an operational phase and extended to more parameters.

4.5 SATELLITE, SEA ICE AND SPATIAL DATA MANAGEMENT

This item was introduced by Takashi Yoshida who reviewed the datasets and services relevant to the Integrated Global Observing Strategy Partners (IGOS-P) ad hoc working group on Data and Information Systems and Services (DISS) as agreed at its meeting in April 2000 (http://ioc.unesco.org/igospartners/). The DISS recommended principle are (i) continuing commitment to data management systems and services, (ii) provision of sufficient, long-term observations, (iii) full and open sharing and exchange of data and products in a timely fashion, (iv) a specific quality and consistency sufficient to meet user requirement, (v) easy and full access to metadata, (vi) preservation of all relevant data and development and implementation of data purging procedures, (vii) readily accessible directories, (viii) internationally-agreed standards in the implementation of these principles, (ix) continuous monitor of the utility and efficiency of the information access and retrieval system, (x) collection, analysis and distribution of information on random errors and systematic biases, and a commitment to ensure the internal consistency of the record, and (xi) provision of comprehensive feedback on problems with data collection or data flow, on the accuracy and usefulness of the products, and on user satisfaction.

As each IGOS Partner and space agency has individual data policy tailored to their individual needs and approved by their governing bodied, an attempt to design an umbrella data policy is not recommended. Each DISS body, such as JPL PO.DAAC for example (http://podaac.jpl.nasa.gov/), provides sophisticated interfaces to users. It is not necessary that JCOMM should establish a new mechanism to handle satellite data as long as each DISS body works properly in accordance with the DISS principles, but a mechanism to provide a linkage to satellite DISS bodies could be developed.

5. COOPERATION WITH OTHER PROGRAMMES

5.1 MARINE XML ACTIVITIES

This item was introduced by Greg Reed, referring to Document ETDMP-I/9, which provided an overview of two current XML activities: the ICES-IOC SGXML and the EU Marine XML Project.

The ICES-IOC Study Group on the Development of Marine Data Exchange using XML (SGXML) was established in 2001. Current activities of the SGXML are focussed on three main areas of interest: parameter dictionaries, point data investigation, and metadata investigation.

- Parameter dictionaries. (i) establish mappings from BODC dictionary commonly used dictionaries, (ii) construct a web interface for accessing the BODC dictionary, (iii) compare and reconcile the parameter dictionary XML structures as defined by the DTD and schema
- Point data investigation. (i) review the 'Keeley brick' concept, determine which parts of the bricks can be substituted with components from OWS, GML and other accepted international standards, (ii) identify and construct the ocean cruise oriented bricks, (iii) apply of the brick / XML structure to 3-d data (e.g., net tow) and identify lacking bricks.
- Metadata investigations. (i) define common terminology for metadata, create a reference model for the abstraction of metadata, (ii) evaluate existing metadata standards by examining ISO19115 to identify elements specific to ocean community needs, complete a comparison mapping of CSR, MEDI, EDMED, USNODC DDF to the ISO 19115, (iii) evaluate the catalogue standard ISO 19110 for application to ocean datasets, and (iv) initiate development of an optimal metadata tag list.

The EU-funded Marine XML project, which commenced in February 2003, aims to demonstrate how XML technology can be used to develop a framework that improves the interoperability of data for the marine community and specifically in support of marine observing systems. The project will develop a prototype of an XML-based Marine Mark-up Language (MML). IODE is responsible for the management of Workpackage 2 – Exploitation and Dissemination, and will be in charge of: (i) disseminating the developments and findings of MarineXML to interested stakeholders and organizations; (ii) developing an Exploitation Plan for identified exploitable project deliverables; and (iii) ensuring the post-project development of MML. The deliverables of the EU Marine XML Project are a working test bed, an outline MML specification and a route-map for development post project.

Mr Reed informed the group of the MarineXML community portal site that has been established by IODE to provide a discussion forum for Marine XML activities and includes details of the SGXML and EU Marine XML programmes. The MarineXML site is at http://marinexml.net.

5.2 THE FUTURE WMO INFORMATION SYSTEM

This item was introduced by Steve Forman, who provided an overview of the Future WMO Information System (FWIS). It is envisioned that FWIS will be used for the collection and sharing of information for all WMO and related international programmes. The FWIS vision provides a common roadmap to guide the orderly evolution of these systems into an integrated system that efficiently meets all of the international environmental information requirements of Members. The FWIS will provide a single point of contact for obtaining data to encourage inter-disciplinary collaboration. The FWIS will provide an integrated approach to meeting the requirements of (i) routine collection of observed data, (ii) automatic dissemination of scheduled products, both real- and non-real-time, (iii) ad hoc, non-routine applications (e.g. requests for non-routine data and products), and (iv) different user groups and access policies. The functional components of FWIS are: National Centres (NC), Data Collection or Product Centres (DCPC) and Global Information System Centres (GISC).

There are a number of similarities between FWIS and E2EDM. Both rely on internet technology, though the techniques still being developed. FWIS is concentrating on technical solutions, not requirements. The key issues are data catalogues and technologies to support the system.

5.3 DATA MANAGEMENT AND COMMUNICATIONS SYSTEM OF US IOOS

This item was introduced by Steve Hankin, referring to the Data Management and Communications Plan documents. He provided an overview of the Data Management and Communications subsystem (DMAC), which is being developed under the U.S. Integrated Ocean Observing System (IOOS) as a contribution to the international GOOS. The DMAC will knit together the distributed components of IOOS, and function as a unifying component within the international GOOS framework. The IOOS includes four subsystems: the Observing subsystem; the Data Management and Communications subsystem; the Modelling and Analysis subsystem; and the Information Product subsystem. The DMAC subsystem consists of a data communications infrastructure, an archive capability and administration functions. The data communications infrastructure includes standards, protocols, and tools to support metadata management, data discovery, data transport; and on-line browse

The DMAC Steering Committee is composed of 4 expert teams (Metadata/Discovery, Transport, Archive, and Products/Applications) and two outreach teams (User Outreach and Facilities Managers). DMAC is not truly a "data management" system but might be thought of as a Data Communications Infrastructure. The technical components of the DMAC Plan are (i) data discovery (metadata), (ii) data transport, (iii) on-line browse, (iv) archive, and (v) metrics, feedback and fault correction.

There are opportunities for "internationalizing" DMAC through cooperation with ETDMP. For example, joint work on metadata standards (ISO 19115), joint work on semantic data model, joint development of Data Transport tools, joint development of Data Archive plan, cross-membership on DMAC Steering Committee, and joint pilot projects. One possibility would be for JCOMM to sponsor the "standards process".

The group agreed that cooperation between DMAC and ETDMP would benefit both groups. Dr Hankin would explore the possibility of a member of ETDMP joining the DMAC Steering Committee, either as an observer or as a full member.

6. JCOMM STRATEGY FOR END-TO-END DATA MANAGEMENT

6.1 BASIC ELEMENTS OF THE END TO END DATA MANAGEMENT STRATEGY

This item was introduced by the chair, Nickolay Mikhailov, referring to Document ETDMP-I/6. The basic principles of the JCOMM E2EDM Strategy defines the overall vision of the end-to-end data management process and contains the general proposals and decisions on various (technological, institutional and other) aspects of the E2EDM establishment including the measures which are necessary for the E2EDM design and implementation. The objectives of E2EDM are (i) to ensure the quality, completeness and comparability of operational and delayed marine data collected from different sources, as well as of forecast, analysis and climate products generated by various organizations and groups; (ii) to organize the full and continuous marine data and information cycle from data collection to product generation; and (iii) to provide the timely delivery of marine data and products for scientific, forecasting, industrial and environmental needs. E2EDM should not replace, but should build on the existing infrastructure of marine data acquisition and management, e.g. the infrastructure developed under major national and international programmes of agencies such as IOC, WMO, ICSU, GOOS, etc. The establishment of E2EDM system will be provided through (i) the improvement of the existing data management practices for operational observed data, marine diagnostic and forecast information, delayed data and climate products; and the transfer and sharing of the best DM practices, experience and knowledge at mono-disciplinary and multi-disciplinary levels; (ii) the development of the new information technology enabling the integration of various DM components and coordinated management and use of marine information resources with the full interaction between data sources on regional/global scales; and (iii) the development of the E2EDM scheme to meet the GCOS/COOP/MMS (as external forces) needs, and the mechanism of this integrated DM scheme adopted and implemented by all participants.

The chair outline the proposed implementation of E2EDM which would include an implementation planning phase, implementation of specific activities through pilot projects, and the continuous assessment and improvement of the individual aspects of the E2EDM system to ensure the appropriate performance of the system.

6.2 JCOMM E2EDM INTEGRATION TECHNOLOGY

This item was introduced by the chair, Nickolay Mikhailov, referring to Document ETDMP-I/7. which discussed the approaches and primary design decisions for the E2EDM integration technology (E2EDM IT). The E2EDM IT model is an abstraction of the architectural software/information elements required to establish a web-based end-to-end data management system.

The integration of data management practices can be integrated by implementing the two main tasks (i) developing the integration technology in the form of a technological “umbrella” over the existing data sources. This will allow us to make the transparent exchange between DM blocks and provide user access to numerous data flows/sets/bases in the unified information space (E2EDM UIS) in a “single stop shopping” manner, and (ii) mapping this technology on the data centre network, when each data centre will support specific data and products. This provides the interface to local data and products according to the standards of the integration technology.

An integration technology is a Web-based technology and development should be based on the solutions and standards of the World Wide Web. An integration technology is a set of rules, standards and tools to support a Web-based, distributed, marine information resource system. The key object of an integration technology is the “information resource” which has two connected views (i) a low technological view – the software and transport/protocol used in the context of web-technologies (WSDL, SOAP, XML-RPC, JMS, application and original servers and others), and (ii) a high semantic (problem-oriented) view - the common dictionaries, structures and formats, links, software and other tools which reflect the specificity of the marine information resources and serve for the achievement of required functionality.

7. OCEAN INFORMATION PILOT PROJECT

This item was introduced by Greg Reed, referring to Document IOC/INF-178, Steering Team of the Ocean Information Technology Pilot Project (ST-OIT), First Session. He provided the background and rationale for the OIT Pilot Project. The First Session of the OIT Steering Team was held in Brussels, 29 November 2002 and the objectives of the first session were to (i) share information on new approaches to applying information technology to managing and exchanging of ocean data and information, (ii) identify OIT pilot project components, (iii) agree on the project management structure and team, and (iv) agree on a way forward: action plan and team leaders. The Steering Team agreed that the OIT PP should include the following components:

- Improved telemetry
- Metadata Management
- Data assembly, data set integrity, quality control
- Data circulation and transport
- Archives and archaeology
- Applications and user interfaces
- Capacity enhancement, training

- Governance, oversight, metrics

An Action Plan was developed based on the following priorities: (i) metadata systems, (ii) data circulation and communication, and (iii) data assembly, quality control and quality assurance.

The ETDMP Group agreed to work closely with the OIT Pilot Project as the priorities of OIT are the same as those for ETDMP. The sessional working groups will focus on the OIT priority areas when discussing possible pilot projects.

8. SESSIONAL WORKING GROUPS

The previous agenda items had discussed a wide range of information relating to metadata and data systems. A number of approaches to the various issues were reported and discussed. The OIT identified three priority areas, namely (i) metadata systems, (ii) data circulation and communication, and (iii) data assembly, quality control and quality assurance. The Group decided to split into three sessional groups to hold discussions in the areas defined by the OIT. The goal of each group was to explore possible pilot projects that can show progress in each of the areas."

The following guidance for the Sessional Working Groups, as suggested by Bob Keeley, was adopted by the Group:

1. Accept DMAC description of the goals as the same as the one we all have and the one to which we will all point to explain what we are striving to achieve.
2. Accept that we will all work towards this goal after due review of proposed solutions and agreements.
3. Look at problems to be explored and pick out those on which we think we can make progress, and more than one of us has to work on at home.
4. Publish interim and final results in a way that everyone can review and approve.
5. Choose at least one project that has a concrete result that we can demonstrate at JCOMM-2 in June 2005 that has high positive impact on members.
6. Chose pilot projects that are complementary to other initiatives.
7. Form OIT project teams and approach people to take part.

The following Working Groups were formed:

- WG1. Metadata Management (Rojas, Collins, Lin, Rickards, Djiman)
- WG2. Data assembly, quality control and quality assurance (Keeley, Maillard, Vanden Berghe, Gowland, Yoshida)
- WG3. E2EDM Prototype. (Mikhailov, Rees, Hankin, Foreman)

On the final day of the session, each WG presented their outline for a pilot project.

8.1 PILOT PROJECT 1. METADATA MANAGEMENT

Objective: To develop and provide practical testing of a comprehensive metadata model that takes into account existing and planned initiatives in the metadata management field.

Participants:

- Ricardo Rojas, Don Collins (co-leaders)
- IODE/JCOMM representative
- SGXML representative
- GCMD representative
- WMO representative
- OpenDAP representative
- EU Marine XML representative
- ETDMP representatives

- Argo data management group representative
- ODAS Metadata Centre representative
- DMAC representative

Activities:

- Define thematic and system metadata requirements (elements) to fully describe any data that is collected by the JCOMM community in relation to the ISO19115 metadata standard (END DATE = June 2004)
- Create additional columns in the metadata system crosswalk for existing systems as needed, e.g., ODAS, ARGO metadata (START DATE = Oct 2003; END DATE = April 2004)
- Request review comments from scientists for additional input regarding missing or needed metadata elements (START DATE = April 2004; END DATE = May 2004)
- Analyze and comment on GOSIC as an alternative single point of entry to other metadata (END DATE = October 2003).
- Analyze appropriate keyword lists (GCMD, BODC, ODAS, WMO Pub. 47, others?) to determine where most terms of interest are available and develop 'master list' for use in JCOMM metadata descriptions (END DATE = August 2004)
- Identify which elements in the metadata model require using keywords (START DATE = January 2004; END DATE = March 2004)
- Match new ODAS terms to GCMD valids (START DATE = January 2004; END DATE = March 2004)
- Decide on 'master list' of keywords for use in JCOMM metadata descriptions.
- Investigate possibilities of multilingual keyword management and maintenance, e.g, GCMD valids translated to and maintained in non-english languages (Lesley to investigate contact with EU Mermaid project) (START DATE = October 2003; END DATE = August 2004)
- Develop/review test software to support online access to various metadata systems
- Investigate emerging technologies, such as Distributed Generic Information Retrieval (DiGIR), Search/Retrieve Web (S/RW, from Library of Congress), semantic web and/or others for efficacy and usefulness in supporting web-portal oriented single-user simultaneous online access multiple metadata systems (END DATE = September 2004, ETDMP meeting) [VLIZ, GCMD]
- Review and evaluate DiGIR and at least one other similar technology (e.g., S/RW) (START DATE = ? END DATE = ?)
- Develop appropriate XML schema to enable distributed metadata search and discovery
- Recommend how to aggregate ODAS Metadata into MEDI/DIF and/or ISO19115 descriptions (END DATE = ?)
- Devise software to create MEDI/DIF and/or ISO19115 descriptions from aggregated ODAS metadata
- Investigate C-Squares software (from CSIRO) as alternative to SVG/GIS capabilities to create online browse images of data described in a metadata description.
- Test software identified above that provides online access to ODAS and/or JCOMM IODE/WMO metadata in distributed metadata databases or metadata systems. (END DATE = December 2004)
- Demonstrate links between a number of (4-6?) centres of metadata
- Compare these findings with GOSIC as an alternative single point of entry to other metadata.

Expected results:

1. The Pilot Project team will propose a JCOMM metadata model based on essential elements identified from crosswalk development efforts to map significant metadata systems to each other. The metadata model must be fully compatible with the ISO19115 Metadata Standard.

2. The Pilot Project team will propose a preliminary (or 'version 1') 'master list' of JCOMM keywords that adequately characterize data holdings of the JCOMM community, recognizing that the preliminary list will be updated as needed. Recommend additional keywords for inclusion in other lists, such as GCMD or BODC. A report that defines the keyword lists analyzed, mapping of keywords between lists and recommendations about a standard keyword list, including status report on multilingual keywords will also be prepared.
3. The Pilot Project team will propose the appropriate software to provide online access to distributed metadata systems.
4. The Pilot Project team will demonstrate access to metadata using the proposed software (#3) from at least 3 centres.

Budget:

1. Contract support: $20,000
2. Travel for meeting: $20,000

8.2 PILOT PROJECT 2. DATA ASSEMBLY, QUALITY CONTROL AND QUALITY ASSURANCE

The Working Group proposed two projects:

- Unique tags for original data
- Data Quality Assessment and Flagging

8.2.1 Unique Tags for Original Data

Objective: To examine a scheme for assigning a unique tag to original oceanographic, meteorological and biological data.

Activities:

1. Request GTSPP to provide documentation of the details of how the tag scheme works including how the tag is created, when in the data assembly process it is applied, what problem it solves, and what changes are required to implement it.
2. Assess the success of the experiment being carried out under GTSPP to apply a unique tag to data collected by the SEAS program
3. Request GTSPP to provide a document for JCOMM-2 that explains the experiment and demonstrates the success.
4. Evaluate how to broaden the application of the tag so that it can also be used for other types of data including meteorological and biological data. This will include consideration of the impacts on data system operations.

Deliverables:

1. A description of how the unique tag is created, and appropriate software if required.
2. An assessment and presentation of the success of the experiment.
3. A white paper on how the tagging scheme can be extended to other types of physical, meteorological and biological data.

Participants

- Chair of GTSPP (Keeley) - Lead
- Edward Vanden Berghe, Elanor Gowland, Catherine Maillard, Takashi Yoshida
- Volunteers

Timetable of deliverables:

1. December 2003. Keeley to circulate draft of issues to address
2. April 2004. Circulate discussion papers. Keeley (ocean), Takashi (met), Vanden Berghe (biology)
3. April 2004. Interim assessment presentation

Budget: One day meeting of participants to decide final draft, Brussels (ICES MDM) May 2004 ($5,000)

8.2.2 Data Quality Assessment and Flagging

Objective: A number of procedures are used to assess the quality of data and there are a number of schemes used to mark data with the results. This causes confusion for users. This project will compare quality control procedures applied throughout the world for a subset of variables, and examine the variations in both procedures and the way quality flags are assigned.

Activities:

1. Assemble inventory of documents describing quality control procedures that are applied to ocean and marine meteorology variables.
2. Compare the procedures variable by variable and recommend a standard
3. Compare the procedures for recording the results of data quality testing with an assessment of the merits and problems of each.
4. Recommend a standard procedure

Deliverables:

1. A web page that contains the inventory (links or documents) of documentation of QC procedures and flagging schemes.
2. A white paper that compares the QC procedures for a limited number of variables and recommends standards (propose: some COOP variable).
3. A white paper comparing QC flagging schemes and recommending a standard

Participants:

- Catherine Maillard and Takashi Yoshida (or other) to lead
- Keeley, Gowland, Vanden Berghe
- International GODAE Steering Team - Jim Cummings / Neville Smith
- US CLIVAR - David Legler
- COOP representative - Nadia Pinardi

Timetable of deliverables:

1. January 2004. Keeley to undertake with MEDS resources then pass results to ICES MDM; ICES MDM to continue activity and publish interim report by April 2004
2. January 2004. Keeley to prepare draft, ICES MDM to extend by April, 2004,
3. May 2004. Catherine Maillard to prepare interim report and report to ETDMP

Budget: Meeting to discuss standards in conjunction with MDM (in conjunction with Project 1)

8.3 PILOT PROJECT 3. E2EDM PROTOTYPE

Objective: To build and demonstrate a prototype system which can undertake real-time data fusion from distributed sources into sample products of interest to a JCOMM user.

Overall time frame: 18 months from project start (January 2004)

Participants:

- Nick Mickhailov (Project leader)
- Tony Rees (Co-leader)
- Steve Hankin - PMEL
- Steve Foreman or delegate – UK Met Office/WMO Task Team
- Edward Vanden Berghe – VLIZ
- Catherine Maillard (or delegate) – IFREMER
- Observers: Bob Keeley, Don Collins, Ricardo Rojas

Vision statement:

- Pilot should demonstrate real-time access to, and fusion of, data:
  - at operational time scale
  - across multiple disciplines
  - preferably non-traditional variables
  - from multiple source formats
  - from multiple providers in different geographic regions
  - of utility to some user group
- The pilot should demonstrate the full range of processes including data discovery, access, and visualization
- It should utilize pre-existing components where possible and be achievable with modest incremental effort

High-level functionality:

The following functionality is envisaged:

1. A user can enter the system, either via a web browser or a dedicated client, and request data of a single or multiple types, from a distributed set of sources, over a single (or possibly multiple) space-time region(s)
2. Appropriate data to the user's request will be automatically sourced from wherever it resides, and returned to the requesting machine (which may be the user's machine, or an intermediate portal providing value-added services)
3. Tools will exist (again either on a dedicated client, or on an intermediate portal) to fuse the aggregated data in real time to produce a newly created data product of value to the user.

Conceptual components required:

The pilot E2EDM system will require the following components, some of which currently exist, some of which do not:

1. Data sources, with data of potential interest to the system, and the technological means for such data to be accessed
2. A master list of such sources – which could be generated as a virtual list by querying one or multiple sources, or reside as an independent entity
3. "System search" metadata for each source, which describes at a high level, in a machine-readable structured way, at least the following:
   a. Data class – according to an agreed semantic model yet to be defined (e.g. satellite data, *in situ* oceanographic data, biological data …)
   b. Parameter list (according to agreed semantic model)
   c. Overall space, time footprint (according to ISO metadata standard)
   d. Location of, and access protocol for remote requests to connect to the data
4. For complex data providers, e.g. sources of data on multiple parameters with discontinuous distributions in time and/or space, more detailed search metadata describing the individual space-time footprints of every parameter (e.g. different biological species distributions)
5. One or more "request brokers" capable of querying first the search metadata, then the relevant data sources, to retrieve data relevant to the user's request. (Such a "request broker" could either be client software installed on the user's machine, or a dedicated portal to which the user connects via a standard web browser)
6. One or more user interfaces which permit the user to formulate an appropriate request
7. One or more applications capable of generation of real-time data products from the data returned as a result of the distributed query
8. Relevant software and hardware to connect the various components of the system, and
9. Relevant data and metadata models to ensure that requests can be formulated by the request broker, and responded to, in a consistent manner.

Commentary: The above list attempts to identify the components which will be required, but makes no final decision as to whether they may exist as real or distributed entities, or where they should reside. For example, the system search metadata described in point (3) has an obvious overlap with the conventional thematic metadata directories (GCMD, EDMED, MEDI, others) and could conceptually reside there in "distributed" form, alternatively it could reside in a separate "registry" more directly under control of the "owners" of the distributed JCOMM system (one could even start with one model, and migrate to another model over time).

Similarly, the more detailed search metadata described in (4) could reside in an intermediate registry or cache, or be generated on demand from the data sources in real time, or simply be ignored for the purpose of the pilot project.

Proposed methodology:

1. Agree on a single high-level architecture for the prototype system. Questions to be decided here will include:
    a. Will the "master list" of accessible data providers described above in (2) be generated on demand from another source (e.g. GCMD, or distributed metadata query), or maintained as a separate entity, for the purposes of this pilot project.
    b. Will the "system search metadata" required for this project will reside in such metadata directories along with the thematic metadata, or in a dedicated registry
    c. Will there be a need for more detailed information to be stored, on specific space-time footprints by individual parameter as described above at E(4),or whether it is sufficient to generate such information via real-time request to the data sources
    d. Whether the "request broker" described above at (5) will comprise client software to be installed on user's machines, or whether there will be a single portal (or replicated portals) providing such functions, accessed via a user's web browser (or both)

    Comment: Such decisions should include an analysis of the strengths and weaknesses of existing architectures of similar systems, e.g. DODS/NVODS, OBIS, others.

    Estimated duration: 3 months

    Target date: March 2004

2. Identify a limited, but challenging suite of parameters to be accessible via the pilot system. For example such parameters might include one to a few oceanographic *in situ* measurements (temperature, salinity); satellite imagery (e.g. ocean colour); marine meteorological and biological observations (e.g. accessible via OBIS or independent source)

    Estimated duration: 1-3 months

    Target date: March 2004

3. Identify a set of data providers who are agreeable to becoming test initial "JCOMM E2EDM data sources" for the purpose of this pilot project. (Target: a suite of between 5 and 10 data sources, representing a range of themes and geographic locations of potential value to a JCOMM user).

    Estimated duration: 1-3 months

    Target date: March 2004

4. Construct semantic data models for:

a. Required "system search" metadata
b. The syntax for "system search metadata" requests and responses
c. The syntax for data requests and responses

Comment: Point 4a should be considered with reference to existing metadata standards (e.g. ISO 19115), the JCOMM community profile of same (under development via parallel metadata pilot project), and additional specific fields required for this project

Estimated duration: 3 months

Target date: June 2004

5. Investigate viability of using existing software components (e.g. OpenDAP, DiGIR) for interfacing with providers' data systems using schemas and syntaxes developed under (4) above, and test install in at least 2 locations, to discover and overcome any problems

   Estimated duration: 3 months

   Target date: September 2004

6. Develop a specification for the user interface (e.g. web interface, personal client software, automated system), construct and refine prototype

   Estimated duration: 3 months initial, plus ongoing refinement

   Target date: September 2004, plus ongoing refinement

7. Develop a specification for the real-time data fusion and visualisation software, construct and refine prototype, or investigate currently available products – e.g. OceanDataView

   Estimated duration: 3 months initial, plus ongoing refinement

   Target date: September 2004, plus ongoing refinement

8. Connect the various components, assess performance of the system, and refine as necessary

   Estimated duration: 3 months

   Target date: December 2004

9. Consider ongoing maintenance and management requirements of the system – e.g. automating repetitive functions, refreshing registry content as required, manual oversight of system functionality, method for extending range of either parameters covered or data sources, method for generating and disseminating system metrics

   Estimated duration: 3 months

   Target date: December 2004

10. Document the system in its "version 1" form, communicate results to JCOMM and/or other interested parties

    Estimated duration: 1-3 months

    Target date: March 2005

Deliverables:

1. Working prototype system, demonstrating the achievement of the project objective
2. Written report to JCOMM
3. Presentation of results at a significant technical workshop or scientific conference

Linkages with other proposed Pilot Projects:

1. Require output for JCOMM metadata model, for use and/or possible extension for this project
2. If GCMD etc. are adopted as the repository for the "system search metadata", may require the distributed search mechanism which is also a planned output from this project

Budget:

1. Contract software development – 3-6 months?
2. Face-to-face meetings of development team – 1-3 meetings?
3. Site visits by team members if required, to install and/or evaluate software/hardware issues
4. Written report preparation and production costs
5. Communication/reporting costs in person (e.g. attendance at conference for presenter)

Total Budget: $60,000

9. ACTION PLAN FOR 2003-2004

The Action Plan for the intersessional period would be based on the three pilot projects identified. Members of each project team would review the pilot project proposals and send comments to the respective project leaders. The project leaders will review and submit the final updated proposals to the Chair by 15 October 2003. Funding sources for the projects will need to be identified and it was agreed to request assistance from the Secretariat in attempting to identify these sources.

The Group discussed the mechanism for adoption of pilot project recommendations. The recommendations would first be circulated to ETDMP members then sent to the DMCG for endorsement. DMCG will distribute the documents to the chairs of other PA, then to IODE, WDCs and other programmes such as CLIVAR, GOOS. The secretariat will document the procedures for a distribution strategy.

10. DATE AND PLACE OF NEXT SESSION

The Group agreed that the next session of ETDMP should be held in 2004 and noted that this would be last meeting of the Group before JCOMM-2. The next session will be held in Geneva in September 2004, with the exact date to be confirmed by the Secretariat.

11. CLOSURE OF THE SESSION

The Chair thanked the Director of the Flanders Marine Institute for hosting the meeting. The first session of ETDMP closed at 1430 on 18 September 2003.

ANNEX I

AGENDA

1. OPENING OF THE SESSION
    - 1.1 OPENING
    - 1.2 ADOPTION OF THE AGENDA
    - 1.3 WORKING ARRANGEMENTS

2. REVIEW OF ETDMP WORK PLAN, DECISIONS RELEVANT TO ETDMP AND ETDMP ACTIVITY IN 2002-2003

3. REQUIREMENTS FOR END-TO-END DATA MANAGEMENT
    - 3.1 GCOS REQUIREMENTS
    - 3.2 COOP REQUIREMENTS
    - 3.3 MMS REQUIREMENTS
    - 3.4 GCOS/COOP/MMS REQUIREMENTS FOR SATELLITE AND SEA ICE DATA
    - 3.5 REGIONAL GRA REQUIREMENTS

4. REVIEW OF EXISTING AND PLANNED DATA MANAGEMENT MECHANISMS AND PRACTICES
    - 4.1 METADATA MANAGEMENT SYSTEMS
        - 4.1.1 REVIEW OF THE MARINE METADATA SYSTEMS
        - 4.1.2 ODAS METADATA
        - 4.1.3 EXISTING WMO/IOC(IODE) OCEAN & MARINE METEO DATA HOLDINGS
        - 4.1.4 INTEGRITY ISSUES OF METADATA MANAGEMENT
    - 4.2 OCEANOGRAPHIC DATA MANAGEMENT
        - 4.2.1 THE GTSPP, ARGO DM, GOSUD PROJECT DATA MANAGEMENT
        - 4.2.2 DATA MANAGEMENT OF IODE SYSTEM, GODAR AND WOCE PROJECTS
        - 4.2.3 OCEANOGRAPHIC OPERATIONAL AND CLIMATE PRODUCT MANAGEMENT ISSUES
        - 4.2.4 INTEGRITY ISSUES OF OCEANOGRAPHIC DATA MANAGEMENT
    - 4.3 MARINE METEOROLOGICAL DATA MANAGEMENT
        - 4.3.1 THE VOS MARINE METEOROLOGICAL DATA MANAGEMENT ISSUES
        - 4.3.2 MARINE METEO OPERATIONAL AND CLIMATE PRODUCT MANAGEMENT
    - 4.4 NON-PHYSICAL DATA MANAGEMENT
    - 4.5 SATELLITE, SEA ICE AND SPATIAL DATA MANAGEMENT

5. COOPERATION WITH OTHER PROGRAMMES AND EXPERT TEAMS
    - 5.1 THE ICES/IOC SGXML - INITIAL VIEW ON MARINE XML IN CONCERNING WITH E2EDM
    - 5.2 THE FUTURE WMO INFORMATION SYSTEM
    - 5.3 DATA MANAGEMENT AND COMMUNICATIONS SYSTEM OF U.S. INTEGRATED OCEAN OBSERVING SYSTEM

6. JCOMM STRATEGY FOR END-TO-END DATA MANAGEMENT
    - 6.1 BASIC ELEMENTS OF JCOMM E2EDM STRATEGY
    - 6.2 JCOMM E2EDM INTEGRATION TECHNOLOGY
        - 6.2.1 CONCEPTUAL DECISIONS ON JCOMM E2EDM INTEGRATION TECHNOLOGY

6.2.2 APPROACHES AND INITIAL DESIGN DECISIONS ON INTEGRATED METADATA MODEL

7. OCEAN INFORMATION TECHNOLOGY PILOT PROJECT (OIT))
   7.1 GENERAL DESIGN DECISIONS ON OIT PROJECT
   7.2 ETDMP ACTIVITY ISSUES IN FRAMEWORK OF THE OIT PROJECT

8. ACTION PLAN FOR 2003 –2004
9. DATA AND PLACE OF NEXT SESSION
10. CLOSURE OF THE SESSION

ANNEX II

LIST OF PARTICIPANTS

I. ETDMP MEMBERS

Donald W. COLLINS
US National Oceanographic Data Center
SSMC3, Room 4618
1315 East West Highway, 4th Floor
Silver Spring, MD 20910-3282
UNITED STATES OF AMERICA
Tel: +1 (301) 713 3275
Fax: +1 (301) 713 3302
E-mail: Donald.Collins@nodc.noaa.gov

Roger DJIMAN
Centre National Océanographique
Centre Beninois de la Recherche Scientifique et Technique (CNO/CBRST)
BP 03-1665 Cotonou
BENIN
Tel: +229 321263, 32-62-14
Fax : +229 32 36 71 / 32 62 14
Email: r.djiman@odinafrica.net
rodjiman@yahoo.fr

Elanor GOWLAND
Met Office
S9, Saughton House
Broomhouse Drive
Edinburgh, EH11 3XQ
Scotland
UNITED KINGDOM
Tel: +44-131 528 7313
Fax: +44-131 528 7345
Email: elanor.gowland@metoffice.com

Robert KEELEY
Marine Environmental Data Service
W082, 12th floor
200 Kent Street
Ottawa, Ontario
CANADA K1A 0E6
Tel: +1-613 990 0246
Fax: +1-613 993 4658
Email: keeley@meds-sdmm.dfo-mpo.gc.ca

Catherine MAILLARD
SISMER
Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER)
Centre de Brest, B.P. 70
29280 Plouzané
FRANCE
Tel: +33 (0)2 98 22 42 79
Fax: +33 (0)2 98 22 46 44
Email: catherine.maillard@ifremer.fr

Nickolay MIKHAILOV (Chair)
Russian National Oceanographic Data Centre/RIHMI-WDC
6, Korolev St.
Obninsk, Kaluga Region 024020
RUSSIAN FEDERATION
Tel: +7 08439 74907
Fax: 7 09525 52225
E-mail: nodc@meteo.ru

Tony REES
CSIRO Marine Research
Castray Esplanade
Hobart, Tasmania 7000
AUSTRALIA
Tel: +61 3 6232 5318
Fax: +61 3 6232 5000
Email: tony.rees@csiro.au

Ricardo ROJAS
CENDOC-SHOA
Servicio Hidrográfico y Oceanográfico de la Armada de Chile
Errazuriz 232, Playa Ancha
Valparaíso
CHILE
Tel: +56 32 266674
Fax: +56 32 266542
Email: rrojas@shoa.cl

Edward VANDEN BERGHE
Flanders Marine Institute
Vismijn, Pakhuizen 45-52
B-8400 Oostende
BELGIUM
Tel +32 59 342130
Fax +32 59 342131
Email: ward.vandenberghe@vliz.be

Takashi YOSHIDA
Office of Marine Prediction
Japan Meteorological Agency
1-3-4, Otemachi, Chiyoda-ku,
Tokyo, 100-8122
JAPAN
Tel: (81 3) 3212 8341 ext. 5128

Fax: (81 3) 3211 3047
Email: tyoshida@met.kishou.go.jp

II. OBSERVERS

Steve FOREMAN
CBS/OPAG-IOS
Met Office
Fitzroy Road
Exeter EX1 8PB
UNITED KINGDOM
Tel: +44 1392 886 230
Email: steve.foreman@metoffice.com

Steve HANKIN
Chair, IOOS/DMAC Steering Committee
NOAA/PMEL
7600 Sand Point Way NE
Seattle, WA 98115
UNITED STATES OF AMERICA
Tel: +1 (206) 526-4809
Fax: +1 (206) 526-6744
Email: steve.c.hankin@noaa.gov

Shaohua LIN
Chairman, JCOMM Data Management Coordination Group
National Marine Data & Information Service
93 Liuwei Road, Hedong District
Tianjin 300171
CHINA
Tel: +86 22 2401 0803
Fax: +86 22 2401 0920
Email: shlin@mail.nmdis.gov.cn

Lesley RICKARDS
Chairperson, IODE
British Oceanographic Data Centre
Bidston Observatory
Bidston Hill, Prenton
Merseyside CH43 7RA
UNITED KINGDOM
Tel: +44151 653 1514
Fax: +44 151 652 3950
Email: ljr@bodc.ac.uk

III. JOINT SECRETARIAT

Peter PISSIERSSENS
Intergovernmental Oceanographic Commission (IOC) of UNESCO
1 rue Miollis
75732 Paris Cedex 15
FRANCE
Tel: +33(1) 45 68 40 46
Fax: +33(1) 45 68 58 12
E-mail: p.pissierssens@unesco.org

Greg REED
Intergovernmental Oceanographic Commission (IOC) of UNESCO
1 rue Miollis
75732 Paris Cedex 15
FRANCE
Tel: +33(1) 45 68 39 60
Fax: +33(1) 45 68 58 12
E-mail: g.reed@unesco.org

Bruce SUMNER
World Meteorological Organization (WMO)
7bis rue de la Paix
Case postale No 2300
CH-1211 Geneve 2
SWITZERLAND
Tel : +41 22 730 80 04
Fax : +41 22 730 80 21
Email: brucesumner@hotmail.com

ANNEX III

LIST OF ACRONYMS

BODC	British Oceanographic Data Centre
BUFR	Binary Universal Form for Representation of meteorological data
CBS	Commission for Basic Systems (WMO)
CMM	Commission of Marine Meteorology
COOP	Coastal Ocean Observations Panel (GOOS)
CRC	Cyclical Redundancy Check
CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
DAC	Data Assembly Centre
DBCP	Data Buoy Cooperation Panel
DIF	Directory Interchange Format
DISS	Data and Information Systems and Services
DMAC	Data Management and Communication
DMCG	Data Management Coordination Group (JCOMM)
DODS	Distributed Oceanographic Data System
DTD	Document Type Definition
EDIOS	European Directory of the Initial Ocean-observing Systems
EDMED	European Directory of Marine Environmental Data
EDMERP	European Directory of Marine Environmental Research Projects
ETDMP	Expert Team on Data Management Practices
ETIDM	Expert Team on Integrated Data Management
ETMC	Expert Team on Marine Climatology
ETODRRGOS	Expert Team on Observational Data Requirements and Redesign of the Global Observing System (CBS)
E2EDM	End-to-End Data Management
FGDC	Federal Geographic Data Committee
FTP	File Transfer Protocol
FWIS	Future WMO Information System
GETADE	Group of Experts on the Technical Aspects of Data Exchange (IODE)
GEBCDMEP	Group of Experts on Biological and Chemical Data Management and Exchange Practices (IODE)
GCC	Global Collecting Centre (VOS)
GCMD	Global Change Master Directory
GCOS	Global Climate Observing System
GLOSS	Global Sea-Level Observing System
GML	Geography Markup Language
GOOS	Global Ocean Observing System
GOSIC	Global Observing Systems Information Centre
GOSUD	Global Ocean Surface Underway Data Pilot Project
GRA	GOOS Regional Alliance
GTS	Global Telecommunication System
GTSPP	Global Temperature Salinity Profile Programme
ICES	International Council for the Exploration of the Sea
ICSU	International Council of Scientific Unions
ICOADS	International Comprehensive Ocean Atmosphere Data Set
IFREMER	Institut Français de Recherche pour l'Exploitation de la Mer
IGOS-P	Integrated Global Observing Strategy Partners
IGOSS	Integrated Global Ocean Services System (IOC-WMO) [*superseded by JCOMM*]
IOC	Intergovernmental Oceanographic Commission (of UNESCO)
IODE	International Oceanographic Data and Information Exchange Programme (IOC)
IOOS	Integrated Ocean Observing System (US)

ISO	International Organization for Standardization
JCOMM	Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology
JMS	Java Message Service
LME	Large Marine Ecosystem
MCSS	Marine Climatological Summaries Scheme
MDM	Marine Data Management working group (ICES)
MEDI	Marine Environmental Data Inventory
MEDS	Marine Environmental Data Service (Canada)
MML	Marine Markup Language
MMS	Marine Meteorological Services
NASA	National Aeronautics and Space Administration
NMDIS	National Marine Data and Information Service, China
NMS	National Meteorological Service
NODC	National Oceanographic Data Centre
NVODS	National Virtual Ocean Data System
NWP	Numerical Weather Prediction
OBIS	Ocean Biogeographic Information System
ODAS	Ocean Data Acquisition System
OIT	Ocean Information Technology Pilot Project
OOSDP	Ocean Observing System Development Panel
OPAG-IOS	Open Programme Area Group on Integrated Observing Systems (CBS)
OPeNDAP	Open-source Project for a Network Data Access Protocol
OWS	Ocean Weather Station
PMO	Port Meteorological Officer
PODAAC	Physical Oceanography Distributed Active Archive Centre
QC	Quality Control
SGXML	ICES-IOC Study Group on the Development of Marine Data Exchange Systems
SOAP	Simple Object Access Protocol
SOT	Ship Observations Team (OPA)
SST	Sea Surface Temperature
SVG	Scalable Vector Graphiccs
SVP	Surface Velocity Programme (WOCE)
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFCCC	United Nations Framework Convention on Climate Change
VLIZ	Vlaams Instituut voor de Zee/Flanders Marine Institute
VOS	Voluntary Observing Ships
VOSClim	VOS Climate (project)
WMO	World Meteorological Organization
WOCE	World Ocean Circulation Experiment
WSDL	Web Services Description Language
WWW	World Weather Watch (WMO)
XML	Extensible Markup Language