Limited Distribution **IOC/IODE-XVII/19**
Paris, 18 December 2002
Original: English

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)

Seventeenth Session of the IOC Committee on International Oceanographic Data and Information Exchange (IODE), Paris, France, 3-7 March 2003

Development of a marine XML: ICES-IOC Study Group on the Development of Marine Data Exchange Systems Using XML (SGXML)

(co-chairs: A.Isenor and R. Gelfeld)

1. Terms of Reference

The SGXML Terms of Reference (TOR) for the 2002/2003 period are as follows:

TOR 1. Create, evaluate and discuss intersessional work on SGXML parameter dictionary including the population of the dictionary for distribution via a defined XML structure.

The XML web distribution of the parameter dictionaries should be completed and the usefulness of the exercise for cross mapping of parameter dictionaries needs to be assessed. The applicability of the XML structure for other dictionaries should also be determined.

TOR 2. Evaluate and discuss intersessional work on point data structure. Evaluate the usefulness of the generalised Keeley brick approach with application to various point data types.

The generalised point data structure needs to be critically evaluated from the perspective of the international data centres. The applicability of the abstract Keeley bricks needs to be evaluated.

TOR 3. Report on the investigation into other available existing standards (e.g., geographers through the Open GIS consortium, taxonomy, ISO standards, metadata standards (MEDI, GFDC, EDMED, etc), utilising what has already been built.

The metadata problem is common to many organisations and considerable effort has been made by these other organisations. The usefulness of these efforts needs to be evaluated within the context of ocean data transfer.

TOR 4. Evaluate and discuss intersessional work on metadata. Evaluate the usefulness of linkages to other metadata standards and on the implications of a generalised metadata model to existing models.

Progress on the generalisation of the metadata model needs to be evaluated. The generalised model needs to be considered within the context of existing models.

2. Short Summary Core Elements of Last Session

The first meeting of the SGXML resulted in the initial development of a plan to guide an investigation into how XML technology might best be used in an oceanographic context. From an IOC/IODE perspective, the requirement was to design a framework for an XML structure that data centres can use. It is thought that a mutually acceptable structure will solve many problems.

A series of short presentations were given to illustrate XML developments in the various organizations represented. These presentations, predominantly served to initiate a series of discussions. The main requirements for an XML based exchange system involved the development of a document type definition (DTD) and standardised syntax for common content (code tables). With these, it should be possible to exchange data knowing the details of the content. Deriving a common data model is the key to XML developments. Related to this are various data relationships (internal and external), normalization, common dictionary development, common code tables and general data model description. Then, the description could be mapped to an XML syntax and thus produce a marine XML.

It was recognized that the definition of tags, etc. was important. However, the importance of a common data dictionary stems from the distributed data system approach. If such a system sends out a request for data (e.g. using spatial-temporal-parameter query) and collects together data from different databases, and then pulls all collected data together, it would presently result in many different names for the same parameter. However the Group considered that all dictionaries must be considered equal and a system for mapping between the different dictionaries is necessary.

An essential component of the development is a general data model containing a common structure. The Group plans to instigate development of a common data model on the basis of the full application domain (marine data in general). Russia will be leading this issue at a national level, to produce a general data model as the basis of a DTD for the widest possible domain.

IOC has registered the domain name marinexml.net, which it established as a community portal for marine XML discussion.

3. Proposed Major lines of Action by Last Session

Parameter Dictionaries

Action 1: *D. Collins will provide the definitions for the above elements and tags.*

Action 2: *A total of 11 internal dictionaries will be mapped to the XML structure defined in this document. The mappings will be conducted by: K. Manni, R. Gelfeld, A. Isenor, N. Mikhailov, G. Reed, F. Nast, J. Szaron, P. Alenius, R. Lowry, J. Gagnon, T. Carval.*

Action 3: *E. Vanden Berghe will provide a DTD for the above structure.*

Action 4: *P. Alenius will provide a schema for the above structure.*

Point Data Investigation

Action 5: *A. Isenor will investigate applying Keeley bricks to point data as a test case.*

Action 6: *E. Vanden Berghe will provide biological and taxonomic input to the Keeley bricks. This will probably result in the taxonomic brick being completely reformed.*

Action 7: *The draft point data definition should be commented on by others in the group. Please send comments to A. Isenor who will produce version two of the point data structure by July. Identified reviewers were P. Alenius, F.*

Nast, T. Carval, K. Manni, E. Vanden Berghe, and R. Lowry. Then at the next meeting we will discuss how to take the point data structure further.

Metadata Investigation

Action 8: *N. Mikhailov will attempt to construct a general metadata model including the definition of EDMED, MEDI, CSR,* etc. *Mikhailov will look to the GETADE work with EDMED/ROSCOP and version descriptions. The hope is that next year there will be something tangible to work with. CSR is only one visualisation of the metadata.*

Action 9: *Reviewers of this general metadata model were identified as: P. Alenius, R. Gelfeld, D. Collins.*

Action 10: *The Group will attempt a mapping between MEDI, EDMED, and CSR and produce a description as for dictionaries. Metadata elements for MEDI are well established. After mapping, design new tags for non-mappable fields. Supply the results of the mapping exercise to A. Isenor for incorporation into Keeley bricks and point data structure.*

Action 11: *Reviewers were identified to review the mapping: H. Dooley, A. Isenor, F. Nast, P. Alenius, and D. Collins. The review should be complete by July. Then A. Isenor will incorporate into point data structure by September.*

Other Items

Action 12: *G. Reed will establish a SGXML communication site under the marine XML community portal.*

The Session Report available on the MarineXML home page at MarineXML.net.

4. Proposed Work Plan, Timing and Budget

The above action items are intended for completion in the 2002/2003 intersessional period. In most cases the Institutions of group members will cover all funding however some funding is required to provide assistance to participants to attend professional and sessional meetings. The next session of the Group is planned for May/June 2003.

Funding required for 2003-2005 is $15,000.

5. Source of Funding

Funds will come from IODE extra budgetary funds.

6. Requested Actions from the Committee

The Committee is requested to:

- Adopt the summary report of the 1st Session of the ICES-IOC Study Group on the Development of Marine Data Exchange Systems Using XML (SGXML);
- Approve funding for the concerned actions: US$ 15,000 for the period 2003-2005.

[end]