



United Nations  
Educational, Scientific and  
Cultural Organization



Intergovernmental  
Oceanographic  
Commission

Workshop Report No. 244

# SCOR/IODE/MBLWHOI Library Workshop on Data Publication 4<sup>th</sup> Session

British Oceanographic Data Centre, Liverpool, United Kingdom  
3-4 November 2011



IOC Workshop Report No. 244  
Oostende, 14 November 2011  
English

**Abstract:**

The fourth SCOR/IODE/MBLWHOI Library Workshop on Data Publication meeting was convened by the Scientific Committee on Oceanic Research (SCOR), the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) and the Marine Biological Laboratory/Woods Hole Oceanographic Institution Library (MBLWHOI Library) on 3-4 November 2011 to evaluate progress of the two pilot projects of the activity and to discuss related topics, such as implementation of data repositories in different data centres, cooperation with related national and international efforts, hear about how data publication is being handled in other disciplines, interactions with publishers of scientific journals, economic implications of data publication.

For bibliographic purposes this document should be cited as follows:

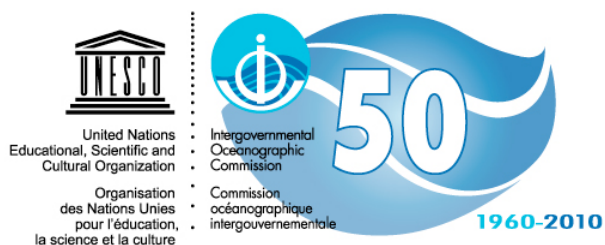
SCOR/IODE/MBLWHOI Library Workshop on Data Publication, 4<sup>th</sup> Session, British Oceanographic Data Centre, Liverpool, United Kingdom, 3-4 November 2011  
Paris, UNESCO, 14 November 2011 (IOC Workshop Report No. 244) (English)

## **TABLE OF CONTENTS**

<b>1. WELCOME AND INTRODUCTION OF THE AGENDA.....</b>	<b>1</b>
<b>2. SESSION 1: PILOT PROJECTS AND ASSOCIATED ACTIVITY REPORTS.....</b>	<b>1</b>
2.1 WOODS HOLE DATA PUBLICATION ACTIVITIES WITH ELSEVIER AND BCO-DMO.....	1
2.2 THE BODC PUBLISHED DATA LIBRARY (PDL) .....	4
2.3 THE PUBLISHED OCEAN DATA REPOSITORY (POD) .....	5
2.4 EXPANDING THE SCIENTIFIC RECORD: DATA CITATION AND PUBLICATION BY NERC'S ENVIRONMENTAL DATA CENTRES.....	5
2.5 OUTREACH ACTIVITIES REPORT. ROUND TABLE DISCUSSION TO CATALOGUE OUTREACH SINCE THE PARIS MEETING (APRIL 2010) AND PROPOSE FUTURE OUTREACH ACTIVITIES.....	7
<b>3. SESSION 2: MANAGEMENT OF DIGITAL OBJECTS IN REPOSITORIES.....</b>	<b>8</b>
3.1 BACK TO THE IT FUTURE: THE CLOUD AND OPEN SOURCE .....	8
<b>4. SESSION 3: THE CODATA TASK GROUP ON DATA CITATION STANDARDS AND PRACTICES .....</b>	<b>9</b>
4.1 CODATA TASK TEAM 2011 MEETING REPORT .....	9
4.2 DISCUSSION ON FUTURE INTERACTIONS BETWEEN THIS GROUP AND THE CODATA TASK TEAM...	10
<b>5. SESSION 4: ECONOMIC ASPECTS OF DATA EXCHANGE.....</b>	<b>11</b>
5.1 THE ECONOMIC CHALLENGES OF OPEN DATA & INFORMATION EXCHANGE.....	11
5.2 DISCUSSION ON ECONOMIC ASPECTS OF DATA AND INFORMATION EXCHANGE.....	11
<b>6. SESSION 5: STIMULATION OF DATA PUBLICATION ACTIVITY IN THE INTERNATIONAL OCEANOGRAPHIC COMMUNITY .....</b>	<b>12</b>
6.1 DATA PUBLICATION ACTIVITY IN AUSTRALIA .....	12
6.2 DATA PUBLICATION ACTIVITY IN NORWAY .....	13
6.3 DISCUSSION ON INITIATION AND CO-ORDINATION OF DATA PUBLICATION ACTIVITIES IN SEADATANET AND THE IODE DATA CENTRE NETWORK. ....	14
<b>7. SESSION 6: ENGAGEMENT OF THE SCIENTIFIC COMMUNITY.....</b>	<b>14</b>
7.1 BODC SECURE FILE ARCHIVE .....	14
7.2 VLIZ MARINE DATA ARCHIVE SYSTEM (MDA) .....	15
7.3 DISCUSSION ON THE RELATIONSHIP BETWEEN RESEARCHERS AND DATA CENTRES IN DATA PUBLICATION (INCLUDES ROLE OF 'SANDBOXES' LIKE MDA AND SFA).....	16
<b>8. SESSION 7: DATA VERSIONING ISSUES .....</b>	<b>17</b>
8.1 VERSIONING ISSUES IN BODC .....	17
<b>9. CONCLUSIONS .....</b>	<b>17</b>
<b>10. CLOSING OF THE MEETING .....</b>	<b>18</b>

**ANNEXES**

[ANNEX I: Agenda of the Meeting](#)  
[ANNEX II: List of Participants](#)



The Intergovernmental Oceanographic Commission (IOC) of UNESCO celebrates its 50th anniversary in 2010. Since taking the lead in coordinating the International Indian Ocean Expedition in 1960, the IOC has worked to promote marine research, protection of the ocean, and international cooperation. Today the Commission is also developing marine services and capacity building, and is instrumental in monitoring the ocean through the Global Ocean Observing System (GOOS) and developing marine-hazards warning systems in vulnerable regions. Recognized as the UN focal point and mechanism for global cooperation in the study of the ocean, a key climate driver, IOC is a key player in the study of climate change. Through promoting international cooperation, the IOC assists Member States in their decisions towards improved management, sustainable development, and protection of the marine environment.



# 1. WELCOME AND INTRODUCTION OF THE AGENDA

Dr Roy Lowry welcomed the participants to the British Oceanographic Data Centre (BODC) and provided information on local arrangements. He then introduced the agenda (attached in Annex I).

## 2. SESSION 1: PILOT PROJECTS AND ASSOCIATED ACTIVITY REPORTS

This Session was chaired by Dr Roy Lowry. He invited all participants to introduce themselves. The list of participants is attached as Annex II. Updates were provided on the two pilot projects: (1) publication of data related to traditional journal articles; and (2) publications created from data previously ingested by data centres.

### 2.1 Woods Hole data publication activities with Elsevier and BCO-DMO

The MBLWHOI Library succeeded in working with an author early enough in the publication process to assign a DOI to the data associated with his article before it was submitted. The author was very excited that it was easy to work with the repository and that the process achieved all his expected outcomes. This author's data was not part of Biological and Chemical Oceanography Data Management Office's (BCO-DMO) mandate, so the data centre referred that author directly to the MBLWHOI Library for creation of metadata and issuance of the DOI.

BCO-DMO has deposited data several data sets associated with articles after publication. So although the published article does not contain a DOI linking to the data, the Woods Hole Open Access Server (WHOAS) records link the datasets to the article and the article record to the datasets.

dc.format.mimetype	text/csv	
dc.format.mimetype	application/pdf	
dc.format.mimetype	application/vnd.google-earth	
dc.publisher	Biological and Chemical Oceanography Data Management Office (BCO-DMO). Contact: bco-dmo-data@whoi.edu	
dc.relation	http://osprey.bcodmo.org/dataset.cfm?id=13552&flag=view	
dc.relation.ispartof	http://hdl.handle.net/1912/4852	
dc.subject	New Production During Winter Convective Mixing Events Project	en_US
dc.subject	CTD Sea-Bird 911 data	en_US
dc.subject	Niskin Bottle samples	en_US
dc.title	Biogeochemistry Data collected from the R/V Oceanus cruises : OC399-03, OC408-01 and OC408-02 from the Northwestern Sargasso Sea roughly 35-28°N and 58-68°W, in water depths exceeding 4200 meters, from February 14, 2004 to March 14, 2005 (New Production project)	en_US
dc.type	Dataset	en_US
dc.identifier.doi	10.1575/1912/4853	

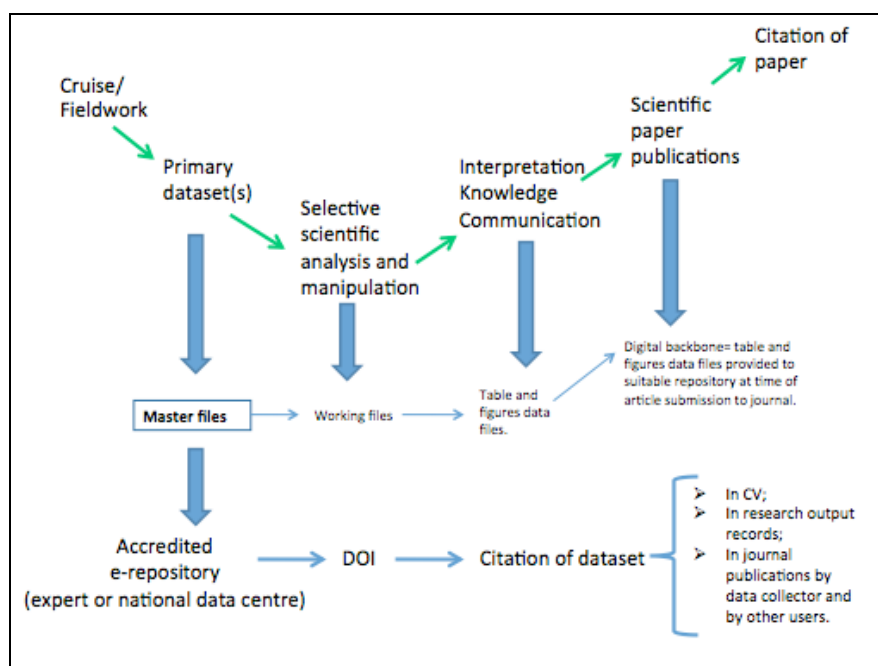
Files in this item				
Files	Size	Format	View	Description
<a href="#">Biogeochemistry.csv</a>	434.8Kb	CSV file	<a href="#">View/</a> <a href="#">Open</a>	Biogeochemistry data
<a href="#">Dataset_description.pdf</a>	16.21Kb	PDF	<a href="#">View/</a> <a href="#">Open</a>	Additional file information: Dataset description
<a href="#">Field_names.pdf</a>	25.35Kb	PDF	<a href="#">View/</a> <a href="#">Open</a>	Additional file information: Field names
<a href="#">OC399-03_Biogeochemistry.kmz</a>	95.95Kb	application/vnd.google-earth	<a href="#">View/</a> <a href="#">Open</a>	Biogeochemistry data (zip-compressed KML files to be viewed in Google Earth)

Figure 1: Dataset record

dc.contributor.author	Lomas, Michael W.	
dc.contributor.author	Lipschultz, Fredric	
dc.contributor.author	Nelson, David M.	
dc.contributor.author	Krause, Jeffrey W.	
dc.contributor.author	Bates, Nicholas R.	
dc.date.accessioned	2011-10-17T17:43:22Z	
dc.date.available	2011-10-17T17:43:22Z	
dc.date.issued	2008-09-17	
dc.identifier.citation	Deep Sea Research Part II: Topical Studies in Oceanography 56 (2009): 843-860	
dc.identifier.uri	http://hdl.handle.net/1912/4852	
dc.description	Citation only. Published in Deep Sea Research Part II: Topical Studies in Oceanography 56: 843-860, doi: 10.1016/j.dsr.2008.09.002	
dc.description.provenance	Made available in DSpace on 2011-10-17T17:43:22Z (GMT). No. of bitstreams: 0 Previous issue date: 2008-09-17	en
dc.publisher	Elsevier B.V.	en_US
dc.relation.haspart	http://hdl.handle.net/1912/4853	
dc.relation.uri	http://dx.doi.org/10.1016/j.dsr.2008.09.002	
dc.subject	New production	
dc.subject	Export production	
dc.subject	Nitrate uptake	
dc.title	Biogeochemical responses to late-winter storms in the Sargasso Sea, I - Pulses of primary and new production	en_US
dc.type	Article	en_US

**Figure 2: Article record**

The Library Repository has also become a resource for researchers looking for a place to deposit data to meet requirements of funding agencies. The Library uses procedures and metadata for these datasets similar to those for datasets associated with articles. To date, DOIs have not been assigned to these data.



**Figure 3: Data publication pathway**

Ms Cynthia Chandler is a Co-Principal Investigator (PI) for the NSF-funded Biological and Chemical Oceanography Data Management Office (BCO-DMO, <http://bco-dmo.org/>). The office is an intermediate data assembly centre created to serve PIs funded by the U.S. NSF Geosciences Directorate's Division of Ocean Sciences (OCE) Biological and Chemical Oceanography Sections and the Office of Polar Programs' (OPP) Antarctic Sciences (ANT) and Organisms & Ecosystems Programs. BCO-DMO is not a permanent data archive facility, and therefore the partnership with the MBLWHOI Library provides a mechanism for BCO-DMO to provide DOIs for "published" data sets.



Over the past year staff members of BCO-DMO and the WHOI Data Library and Archive have worked together to develop a semi-automated process by which data sets managed at BCO-DMO and available from the BCO-DMO data catalogue can be 'packaged' and assigned a DOI on demand. The process includes (1) export of selected data as a comma separated (CSV), plaintext file (or other non-proprietary format, if CSV is not appropriate); (2) export of essential metadata fields that are tagged with Dublin Core identifiers; and (3) preparation of a Metadata Encoding and Transmission Standard (METS) package that contains a 'mets.xml' file, comma-separated value CSV data file(s), and supplemental documents (usually as PDF files), for example, methods or cruise reports. Future METS packages will also include a Keyhole Markup Language (KML) file to provide geospatial context and support a map display. When the METS package is accessioned by the Data Library, a DOI is returned and the data set record at BCO-DMO is updated with the DOI information. Having the data set and published article DOIs completes the connection between data repository, data library and publisher.

It was noted that for each version of a data set a new DOI is created. Provenance is included in the metadata. The publication has a link to the version that was used for that publication.

The question was raised about file types and the fact that they become obsolete after some time. It was noted that BCO-DMO is not a long-term archive. So this issue should be addressed by the US-NODC. This issue exists for libraries, as well, of course.

Ms Chandler attended EGU 2011 and presented a poster for our project. This led to conversations with an Elsevier representative who was very interested in our work. Cyndy Chandler and Lisa Raymond followed up with a phone conversation. Elsevier is partnering with individual data repositories that publish data associated with articles. They have developed a system that points to data from ScienceDirect. When someone locates an article in ScienceDirect that has data in a repository, a banner is displayed with a link to the repository (see Figure 2). The MBLWHOI Library consulted with its DSpace vendor and decided to proceed with the collaboration. A Web application and banner have been created and the MBLWHOI Library is in the final stages of implementing the service. This collaboration required a "Linking Agreement" from Elsevier. The Library found no proprietary language to prevent them from approving this agreement, but they won't be able to keep track of ScienceDirect statistics because of their confidentiality agreement with journals.

In the year ahead the Library expects to continue work with BCO-DMO to streamline data publication processes. They expect to continue outreach and believe that there will be interest among authors to link to data from ScienceDirect.

## **CONCLUSIONS**

**Meeting participants agreed that the project could be considered a success when several NODCs reach the current implementation level of the MBLWHOI Library system for data publication. Meeting participants recommended that documentation should be prepared on the WHOAS-BCO-DMO system (including both installation instructions and end-user instructions).**

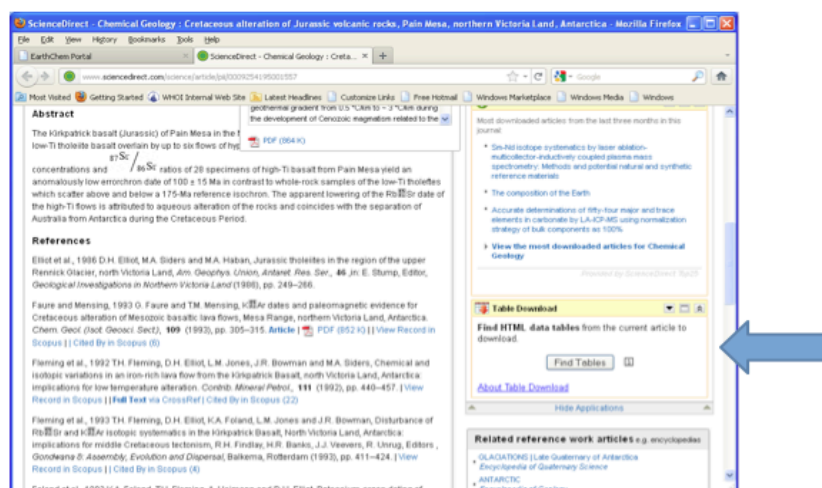


Figure 4: linking to data set from ScienceDirect

## 2.2 The BODC Published Data Library (PDL)

This agenda item was introduced by Dr Roy Lowry.

BODC expects researchers to provide data as part of the data publication process in standard formats that will eventually allow ingestion into the BODC data system. Researchers can also use exports of data from the BODC data system as “database output snapshots”. Such data are stored as physical versioned files. This will be an on-demand service, but the question “What is a data set?” remains. It could be data from an entire project, a group of cruises, a single cruise etc. Answering this question requires negotiations between researchers and the data centre. BODC will not deal with synthesized data (data behind the graph) which are a patchwork of data from various sources, because there is no potential for future data ingestion.

The planned implementation will include the following:

- Static pages and an example landing page are already available on development server
- Oracle back end needs designing and implementation
- Data storage needs to be finalized
- Deployment on live servers
- At least four data sets in the queue for publication
- Delivery of at least three datasets has been promised by the end of November 2011

BODC will identify a number of formats they will accept (e.g., txt, html,...) or reject (e.g. xls, zip, doc, mdb,...) for the PDL, which will conform to acceptability criteria established across the NERC designated data centres.

Dr Lowry recalled that the plan was to have a relationship between BODC and IODE (POD). Unfortunately progress has been slower than planned due to Marc Goovaerts’ (U Hasselt) limited availability. Consequently, having POD fully operational by the end of November 2011 is unlikely. BODC will use its own data vault in the short term, but will undertake further trials with POD.

### Conclusions

**The meeting recommended identifying under “best practices” the recommended (and discouraged) formats, taking into account the longevity of formats. This information**

**should also be disseminated through OceanTeacher. The Digital Curation community should be considered as a source of information on the technical stability of formats.**

Further discussion on POD was held under 2.3

## **2.3 The Published Ocean Data repository (POD)**

This Agenda Item was introduced by Mr Marc Goovaerts (participated by video conferencing).

He reported on the current status of POD. This approach was discussed last year and also this year during the IODE 50<sup>th</sup> anniversary conference. A test server is now available at the University of Hasselt. It is an adapted DSpace installation. Dr Gwen Moncoiffé reported satisfaction with the metadata structure.

Dr Moncoiffé noted that the main issue with the current online interface is that only one file can be uploaded per record. This is a major problem for situations in which many files constitute the dataset, such as for example a series of CTDs profiles. The interface should allow batch loading of multiple files for a single record. Submitting a zip file or any other compressed formats is an option but it is not an acceptable format for long-term archiving and should be avoided. Mr Goovaerts responded that it was a DSpace requirement that for each file there should be one metadata record. However, the possibility of uploading multiple files can be investigated. Ms Lisa Raymond noted that administrators can do batch uploads. Mr Goovaerts noted that the University of Antwerp has developed a system that accepts zip files and unpacks the files in the zip folder immediately. This application might be linked to the POD DSpace system, so Mr Goovaerts will experiment with this.

The possibility of allowing batch loading of multiple files with accompanying records was also discussed and will be investigated by Mr Goovaerts as a future enhancement to the current version of POD.

Mr Goovaerts informed the meeting that the installation will shortly be transferred to the IOC project Office for IODE, Oostende.

### **Conclusions**

**Meeting participants noted that the experimental installation currently hosted by UHasselt will be migrated to the IOC Project Office for IODE, Oostende by the end of November 2011. The meeting requested that Mr Goovaerts further investigate the issue of single file submission per metadata record and also consider enabling batch loading of multiple metadata records with accompanying files to POD.**

## **2.4 Expanding the scientific record: data citation and publication by NERC's environmental data centres**

This agenda item was introduced by Dr Sarah Callaghan.

NERC funds research projects, which produce data. It is essential that these data are properly managed to ensure their long-term availability. NERC's network of data centres provides support and guidance in data management to scientists funded by NERC; the data centres are responsible for the long-term curation of data and to provide access to NERC's data holdings. NERC has a new data policy (<http://www.nerc.ac.uk/research/sites/data/policy.asp>), which came into force in January 2011. The Data Policy details its commitment to support the long-term management of data and also outlines the roles and responsibilities of all those involved in the collection and management of data. The policy's requirements for identifying data of long-term value and for developing data management plans have yet to be formally implemented, because it is still necessary to allow time for NERC to develop the necessary processes and mechanisms for reviewing and managing data management plans, and to

work with the NERC community to develop criteria to help identify data of long-term value. The British Atmospheric Data Centre (BADC) has a model data policy ([http://badc.nerc.ac.uk/data/BADC\\_Model\\_Data\\_Policy.pdf](http://badc.nerc.ac.uk/data/BADC_Model_Data_Policy.pdf)) that gives criteria for selecting simulated/model data for management.

Dr Callaghan reviewed why we want to cite and publish data:

- (i) Pressure from the UK government to make all data from publicly funded research available to the public for free, while still giving scientists attribution and credit for their work. The general public want to know what the scientists are doing (c.f. Climategate);
- (ii) Research funders want reassurance that they're getting value for money from their funding. Value is determined, in part, on peer-review of science publications (well established) and of data (not yet tradition);
- (iii) Allows the wider research community to find and use datasets outside their immediate domain, confident that the data are of reasonable quality; and
- (iv) From a strict data-centric point of view, citation and publication provides an extra incentive for scientists to submit their data to data centres in appropriate formats and with full metadata.

Publishing data for the scholarly record: Scientific journal publication mainly focuses on the analysis, interpretation and conclusions drawn from specific datasets. Peer-review and publishing the raw data that form the dataset is more difficult, as datasets are usually stored in digital media, in a variety of (proprietary or non-standard) formats. Peer review is generally only applied to the methodology and final conclusions of research, and not to the underlying data. But if the research conclusions are to withstand the test of time, the data must be of good quality. A process of data publication, involving peer review of datasets, would be of benefit to many sectors of the academic community.

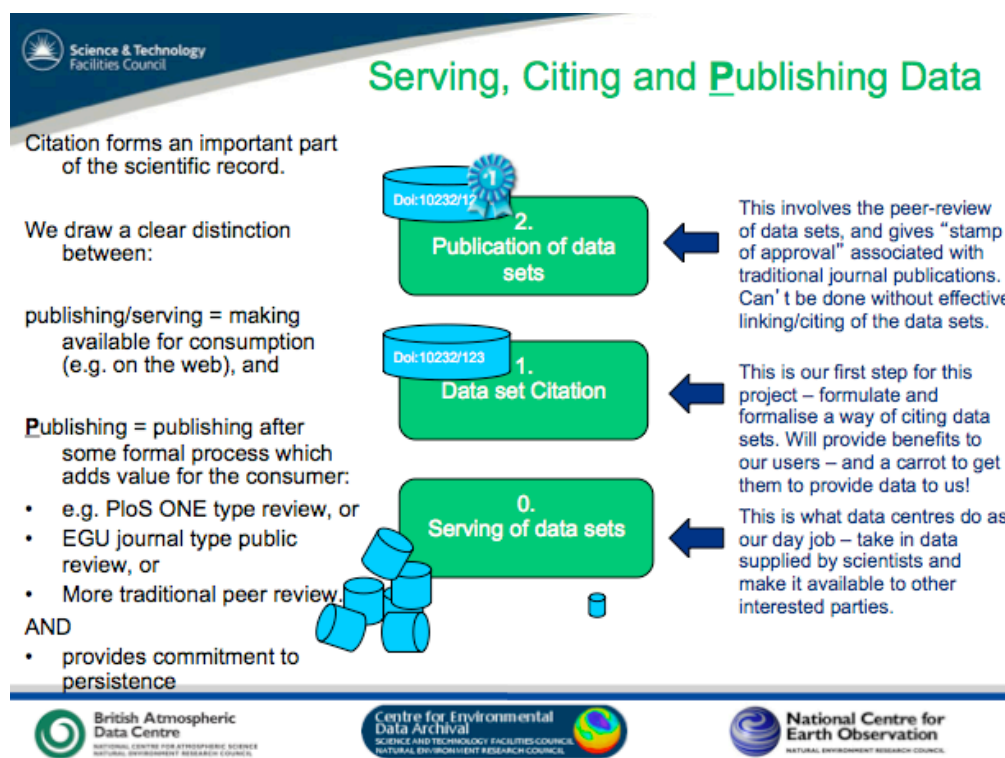


Figure 5: Serving, Citing and Publishing Data

The difference between levels 1 and 2 is analogous to the difference between grey literature and peer-reviewed journal articles.

There is much discussion in the data management community about “Open Data.” Some scientists are quite wary of it, for several reasons, mainly revolving around not getting credit for the work involved in creating the dataset. Scientists are accustomed to journal publications, so piggybacking on these for data publications could bridge the gap between open and closed data.

#### Data Citation and Publication Project Aims:

- To implement publication and citation of datasets held within the NERC data centres.
- To increase NERC’s influence on work to provide and cite data outputs from scientific work in similar ways to scientific papers.
- To demonstrate to the NERC community that data citation and publication is both personally and scientifically advantageous.
- To form partnerships with other organisations with the same goal of data publication to exploit common activities and achieve a wider community buy-in. To this end, project team members are involved with both the SCOR/IODE/MBL WHOI Library Data Publication Working Group, the CODATA-ICSTI Task Group on Data Citation Standards and Practices and the DataCite Working Group on Criteria for Datacentres.

#### What sort of data can we/will we cite?

A dataset must be:

- Stable (i.e. it will not be modified)
- Complete (i.e. it will not be updated)
- Permanent – by assigning a DOI, the data centre/library is committing to make the dataset available for posterity
- Good quality – by assigning a DOI the data centre/library is giving it their stamp of approval, saying that it’s complete and all the metadata is available

When a dataset is cited that means:

- There will be bitwise fixity
- With no additions or deletions of files
- No changes to the directory structure in the dataset “bundle”

A DOI should point to an *html representation* of some *record* which describes a *data object*.

Upgrades to versions of data formats will result in new editions of datasets.

The NERC Data Citation and Publication Project has been running for one year. The project is now entering phase 2 (which will take two years). The end of phase 2 will result in guidelines for the NERC data centres on what is an appropriate dataset to cite, as well as guidelines for data providers about data citation and the sort of datasets that can be will cited.

## **2.5 Outreach activities report. Round table discussion to catalogue outreach since the Paris meeting (April 2010) and propose future outreach activities**

This agenda item was introduced by Ms Lisa Raymond. In order to promote the data publication initiative, presentations and/or posters were presented at the SAIL Annual Meeting, Galveston, TX, 2010; IAMSLIC Annual Conference, Argentina, 2010; AGU Annual Fall Meeting, 2010; EGU Annual Meeting, 2011; ACM/IEEE JCDL, Ottawa, 2011, CODATA Annual Meeting, 2011, WDS Meeting Kyoto, 2011.

In response to the presentation at SAIL, John Cruickshank, Librarian, Skidaway Institute of Oceanography, visited Woods Hole to learn more about the data publication project. He is very interested in implementing data publication at his institution and has reported on his progress with interviews with researchers.

As a result of a conversation at the JCDL in Ottawa, Lisa Raymond has had correspondence with librarians at the University Corporation for Atmospheric Research (UCAR). They are also interested in data publication, having recently hired a person to lead digital initiatives. Lisa shared information about metadata, DOIs and related readings.

A lecture on data publication was also included in the OceanTeacher course “Data Curation for Information Managers and in-depth Digitization Practicum”. This exposed students from around the world to the data publication concept.

Dr Ed Urban suggested that after the Oceanographic Data Publication Cookbook is completed it could be advantageous to work through the U.S. Consortium for Ocean Leadership for implementation in the United States. This organization represents the largest U.S. oceanographic institutions. Roy Lowry asked for other ideas for future outreach. Urban suggested that another meeting with journal editors (as at the 2008 AGU meeting) could be helpful, particularly after the Cookbook is completed. The Cookbook is expected to be completed by the end of 2012.

### **Conclusions**

**The meeting recommended that a “Cookbook” should be prepared providing full documentation (including WHOAS documentation) and guidelines for all elements and applications of data publication. The meeting recommended that this should subsequently be provided to the Consortium for Ocean Leadership, which could assist with its promotion. In terms of future opportunities for outreach, the meeting recommended a new meeting with publishers, as well as participation in the next EGU conference. In this regard, it was decided that the WHOAS portion of the Cookbook should be ready by March/April 2012 to be available for the EGU Conference.**

## **3. SESSION 2: MANAGEMENT OF DIGITAL OBJECTS IN REPOSITORIES**

### **3.1 Back to the IT Future: The Cloud and Open Source**

This agenda item was introduced by Dr Michael J. Ackerman. Dr Ackerman’s presentation focused on Openness, Interoperability and The Cloud.

At the U.S. National Library of Medicine (NLM) they developed the Insight Tool Kit (ITK) to manage medical imagery. ITK enables re-creating a 3D model based upon 2D serial images. ITK is an application with 400 C++ classes, 300,000 lines of code and required \$7M to develop over a period of 4 years. It costs \$250K per year to maintain. This demonstrates that open source software is not inexpensive to develop or maintain.

Open access data: Data have value, for example, in providing a way to a new grant. For a doctor, it is also a way to retain patients, for providers who will not release data to alternative providers. An alternate tenure system might include the question, how many papers were published using your data? For open access data, who will ensure continued availability? Will it be the publishers or the data centres/libraries?



Open access literature also has value: license fee for copies, subscription fee for journals or access. New publisher models refer to the publisher as a certifier rather than distributor (“good housekeeping seal”). In this model, the authors pay for the certification process. Ed Urban noted that this is essentially the model being used for some journals used by ocean scientists (e.g., the journals sponsored by EGU).

PubMed Central is an open repository for medical literature. It holds the original submitted manuscript rather than the published paper (which may be different) to avoid the infringement of journals’ copyrights. This is causing some problems as referring authors may cite information contained in the submitted paper, rather than the published version.

NLM did an experiment with the Optical Society of America, in which they created an online version of an existing journal, enabling free download of the PDFs of the articles and supplementary information. The PDFs included clickable images that included the computerised axial tomography (CAT) scans needed to create three-dimensional images. From the producer’s viewpoint, the experiment was not entirely successful because authors did not want to participate because they did not want to share their data. Reviewers were hesitant to participate because they did not know how to peer review the raw data associated with images. However, from the reader’s viewpoint the experiment was very successful in that they were excited by being able to interact with the 3D image data.

Openness implies interoperability and interoperability requires standards. But there are too many standards. How do you assure interoperability without creating new standards? What are the consequences of interoperability? There is a need for greater openness and interoperability, but this leads to less privacy.

The concept of “Cloud computing” can be thought of as the return of the mainframe. There are different applications of the Cloud. There is the computational cloud in which software resides and is executed in the cloud, not on individual computers. The use of the software becomes a billable service. There is the storage cloud, in which users rent space for storage of files. There is the network cloud, which ensures delivery through a given amount of bandwidth. The cloud should make openness and interoperability easier to attain because the cloud imposes standards.

## **4. SESSION 3: THE CODATA TASK GROUP ON DATA CITATION STANDARDS AND PRACTICES**

### **4.1 CODATA Task team 2011 meeting report**

This agenda item was introduced by Dr Sarah Callaghan. She reported on the “Developing Data Attribution and Citation Practices and Standards An International Symposium and Workshop” held in Berkeley, California, USA, on August 22-23, 2011.

Slides on-line at [http://sites.nationalacademies.org/PGA/brdi/PGA\\_064019](http://sites.nationalacademies.org/PGA/brdi/PGA_064019)

The proceedings were recorded and the aim is to publish a full report of the symposium in April 2012.

There were five topics in the meeting, reflecting the terms of reference of the group:

1. What are the major technical issues that need to be considered in developing and implementing scientific data citation standards and practices? (Session IIA <http://ietherpad.com/Session-IIA-2> )

2. What are the major scientific issues that need to be considered in developing and implementing scientific data citation standards and practices? Which ones are universal for all types of research and which ones are field- or context-specific? (Session IIB <http://sync.in/DataCitationIIB> )
3. What are the major institutional, financial, legal, and socio-cultural issues that need to be considered in developing and implementing scientific data citation standards and practices? Which ones are universal for all types of research and which ones are field- or context-specific? (<http://ietherpad.com/Session-III>)
4. What is the status of data attribution and citation practices in individual fields in the natural and social (economic and political) sciences in United States and internationally? Case Studies. (Session IV <http://ietherpad.com/Session-IV> ) (One of the case studies concerns ocean science, providing an opportunity for input from our project. Two participants in the Liverpool meeting are members of the CODATA group.)
5. Institutional Roles and Perspectives: What are the respective roles and approaches of the main actors in the research enterprise and what are the similarities and differences in disciplines and countries? The roles of research funders, universities, data centres, libraries, scientific societies, and publishers will be explored. (Session V <http://ietherpad.com/DataAttrib20110719>)

The next meeting of the CODATA committee will be held in June 2012 in Copenhagen. Another meeting of the committee, to work on the white paper and discuss the interim results, will be held at the CODATA Conference in Taipei in October 2012, with the final publication of the white paper scheduled for early 2013.

Dr Callaghan also attended the DataCite meeting; slides from this meeting are available on-line at <http://www.datacite.org/node/41>; Blog post by Jan Brase summarising the event at <http://datacite.wordpress.com/2011/08/27/datacite-summer-meeting-recap/>

## 4.2 Discussion on future interactions between this group and the CODATA Task Team.

This agenda item was introduced by Dr Roy Lowry. He recalled that the question was whether the CODATA committee is doing the same as we are and the conclusion was that the groups have different, but intersecting, tasks. We are trying to establish data publication practices specific to the oceanographic community and for NODCs and other data centres that specialise in oceanographic data. CODATA is more general and covers all domains and disciplines. So, we have a mission to feed expertise into the IODE data centre community, but the CODATA committee can recommend standards and best practices. Oceanography is relatively mature in data sharing thanks to 50 years of IODE. We can feed the results of our discussions and pilot projects into the CODATA project through Mr Helge Sagen and Dr Sarah Callaghan. We also need to brief the IODE community on what is happening in CODATA.

### Conclusions

**The meeting recommended that information on the activities of our group be transmitted to the CODATA committee through Mr Helge Sagen and Dr Sarah Callaghan. The meeting noted that it is not duplicating the work of CODATA. The meeting noted that its work is strongly related to the objectives of SeaDataNet-2**



## 5. SESSION 4: ECONOMIC ASPECTS OF DATA EXCHANGE

### 5.1 The Economic Challenges of Open Data & Information Exchange

This agenda item was introduced by Mr Mark Kurtz.

Knowledge, including data and information products, is often asserted as an inherent public good. In fact, IODE products and services do not possess the attributes of pure public goods: they are demonstrably excludable (individuals can be excluded from use of some IODE products), and almost certainly rivalrous (competitive) in several respects. Moreover, nothing in published IODE Objectives or the IOC Oceanographic Data Exchange Policy appears to *require* provision of IODE products and services on a public-good basis. Nevertheless, IOC's Strategic Plan for Oceanographic Data and Information Management asserts a vision of a "global public commons" for oceanographic data and information services. That is admirable, but creates a "free-rider"<sup>1</sup> problem that, unless overcome, will inhibit IODE's ability to provide its products and services. Neoclassical economics has traditionally held that it falls to governments to provide public goods via compulsory taxation, as these goods by their nature cannot be provided by the private market. In the context of global public goods, taxation is not an option (the "Westphalian Dilemma"). Therefore, several other mechanisms for overcoming free-ridership, identified over time by economists and social scientists, were evaluated by Mr Kurtz: assurance contracts, Coasian solution, joint products, introduction of an exclusion mechanism. In addition, aggregation technologies (the manner in which contributions to a public good determine the total quantity of the good available for consumption) were briefly reviewed in the context of IODE products and services: summation, weakest-link, best shot, and weighted sum.

### 5.2 Discussion on economic aspects of data and information exchange

This Agenda Item was introduced by Dr Roy Lowry.

Dr Lowry asked the group to consider how we can make our data publication project economically sustainable. Research is largely government, university, project or grant funded. Based upon this model, how will the data publication process serve those who pay for the research? Mr Pissierssens mentioned the considerable cost (100K\$/day) of operating a research vessel. If the researcher loses a laptop on which the collected data were stored then 100K\$ (or more) is lost. If the data are put into the data repository, then that cannot happen.

Dr Lowry informed the group that in all NERC projects there has to be a data management plan, and those preparing research proposals come to BODC and ask for an estimate of the cost of management they expect to result from the research. If the proposal is successful, a proportion of the budget will go to data management.

Ms Chandler pointed out that publication costs are allowed on NSF grants.

---

<sup>1</sup> Free riders are individuals who benefit from a public good to a greater extent than they pay for it.

## 6. SESSION 5: STIMULATION OF DATA PUBLICATION ACTIVITY IN THE INTERNATIONAL OCEANOGRAPHIC COMMUNITY

### 6.1 Data publication activity in Australia

This Agenda Item was introduced by Dr Roger Proctor.

The Integrated Marine Observing System (IMOS) is a research infrastructure initiative of the Australian Government (funded to 2013). Its principles include the following:

- A national system providing a service
- A multi-platform, multi-disciplinary integrated system
- Delivering repeated observations in an enduring way
- Data are available at no charge, available to all, and delivered in a timely manner

IMOS ([www.imos.org.au](http://www.imos.org.au)) is establishing Australia's Sustainable Marine Observing System. The Australian climate is strongly influenced by surrounding ocean and coastal circulation is strongly influenced by boundary currents, so the benefits of such an observing system are obvious.

IMOS is delivered by institutions operating 10 National Facilities. The IMOS portal provides information on deployments, ISO-standard metadata, access to actual data, as well as plots, maps, visualizations, etc.

Australia is also setting up the AODN (Australian Ocean Data Network), marine and coastal data resources: <http://portal.aodn.org.au>.

This originated from the AODCJF (Australian Ocean Data Centre Joint Facility) (<http://www.aodc.gov.au/>), which was established 10 years ago. The AODN is composed of the following organizations:

- Australian Antarctic Division
- Australian Institute of Marine Science
- Bureau of Meteorology
- CSIRO Marine and Atmospheric Research
- Geoscience Australia
- Royal Australian Navy – Hydrography and Metoc Branch

The AODN is Australia's interface to large amounts of otherwise undiscoverable marine data. Its infrastructure and visualization tools establish free access to data and metadata. The AODN is neither *generating* nor *storing* data. There are two challenges:

1. Publishing interoperable data from multiple disciplines
2. The AODN can only work if people agree to share the data

#### The AODN data policy:

The AODN would like data that are:

- unencumbered – free, no charge to the end user
- with metadata that is ISO19115 MCP compliant – currently the only identifier
- well curated
- licensed appropriately

The AODN expects providers to be responsible for data curation, data accuracy, data storage, and creating metadata records for datasets on behalf of individuals. They will apply the Creative Commons BY Attribution License

Priority datasets will be received from Commonwealth Agencies, National programmes (e.g. CERF, NERP), Marine National Facility, State Governments, other NCRIS projects, fisheries, offshore operators, and international projects with significant Australian content (e.g. OBIS).

There is also the Australian National Data Service (ANDS). <http://www.ands.org.au>. This will be a portal of portals in Australia. ANDS is also linked to DataCite and covers the cost of purchasing DOIs. ANDS is conducting a pilot program for assigning DOIs to datasets and IMOS is participating in this pilot.

## 6.2 Data publication activity in Norway

This Agenda Item was introduced by Mr Helge Sagen.

The Norwegian Marine Data (NMD) centre acts as national coordinator for marine data and manages research data from the Institute of Marine Research and other institutions. It has been a National Oceanographic Data Centre since 1971.

### 3 MAIN ACTIVITIES OF NMD:

#### 1. DATA CENTRE

- Data catalogue - Metadata
- Long-term stewardship of research data
- Operational data management

#### 2. APPLICATION DEVELOPMENT

- Database design
- Quality control
- Project specific developments
- Web Map Service client for MAREANO

#### 3. DATA EXCHANGE

- Queries
- Web Map Services (WMS)
- Visualisation on Internet
- MAREANO portal

The Institute of Marine Research adopted its data policy on 15 May 2007: *“The research data originated by the Institute of Marine Research are freely available for use by own researchers, national and international researchers, government officials and society in general if nothing else is devised. The institute manage data by adopting best data management practices in accordance with established data type guidelines, quality assurance routines and data availability.”*

The Norwegian Research Council published three data catalogues in 2003: (i) important climatic data series; (iii) important terrestrial and limnological data series; and (iii) important marine data series.

During the International Polar Year (IPY) a project was implemented on “meta data harvesting and publishing, DOKIPY”. This involved metadata from 30 Norway funded IPY projects. Norway also participates in SeaDataNet.

There is also a data catalogue at the IMR: <http://www.imr.no/datacatalog> (about 500 data sets/time series). There is also a catalogue that includes also 16 other Norwegian institutions. This will become the Norwegian Marine Data Centre (NMDC). It will provide seamless access to marine data. This will start in March 2012.

Norway has established a license for Norwegian data. This will be similar to the CC BY license (although it is not mentioned). So it may be necessary to apply a double license approach.

The IMR Library has developed an IMPR Publication database that uses DSpace.

Mr. Sagen is a member of the Committee on Data for Science and Technology (CODATA)-International Council for Science and Technical Information (ICSTI) Task Group on Data Citation Standards and Practices.

### **6.3 Discussion on initiation and co-ordination of data publication activities in SeaDataNet and the IODE data centre network.**

This Agenda Item was introduced by Dr Roy Lowry.

Dr Lowry noted that, if there is already a data centre then why not set up a data publication service rather than send the information to PANGAEA? SeaDataNet-2 (SDN-2) will include geochemical and biological data. Dr Lowry noted that SDN-2 will also need to resolve the versioning issue. The Common Data Index (CDI) should include versioning information and a mechanism is needed to ensure that previous versions remain accessible. The SDN-2 Technical Team will meet at BODC on 22-23 November 2011. The versioning issue will need to be discussed there. Versioning should also be addressed by the IODE SG-ODP. Each version of a dataset should have its own DOI.

#### **Conclusions**

**The meeting called on the SDN-2 Technical Team and IODE SG-ODP to address versioning. The meeting also called for SDN, ODP and IMOS (Australia) to start data publication activities within their systems.**

## **7. SESSION 6: ENGAGEMENT OF THE SCIENTIFIC COMMUNITY**

### **7.1 BODC Secure File Archive**

This Agenda Item was introduced by Dr Ray Cramer.

So far files can be uploaded to BODC only through ftp which is somewhat cumbersome to use. Email is another option for file upload but here the size of attachments is an issue.

The aim is a secure file archive and a system is being developed with the following attributes:

- Project participants shared area
- Secure access through BODC website login to authenticate user
  - User registration/roles
  - Automated email to responsible data scientist
  - Automatic virus checking
- Hierarchical approach uses Teams with Datasets which contain files
  - Keep it simple approach

### **Users perspective:**

- A user is presented with a list of Teams to which they are affiliated
  - Sign Team read/write licenses
  - New Teams – created by BODC
  - Registration as a Team member
  - Publicity of Teams (private or joinable)
- Select a Team – displays Datasets
- Select a Dataset – displays Files
  - Select a file for upload
  - Download a file

### **Data Model**

- A user's Role must match the Team
- Each Team has
  - Read and Write policy
  - Responsible Data Scientist
  - Maximum Team Space limit
  - Maximum File Upload limit
  - Each Dataset has
  - Brief description
  - Initiators reference
- Each File has
  - File name
  - File size
  - Initiators reference
  - Upload status
    - New
    - Updated, Accessioned, Approved, Closed
  - Download status
    - Private
    - Metadata, Download, Accession

## **7.2 VLIZ Marine Data Archive System (MDA)**

The MDA system is an online interface system for archiving of data files. The purpose is to offer scientists, scientific research groups and project participants the possibility to archive their data files and, at the same time, share their data files with a selected group of people. The MDA makes it possible to store data files in a fully documented way. The uploaded data files are stored together with their 'technical' or 'use' metadata, ensuring future usability and interpretation of the files. Files deposited in the archive receive a unique identifier URL code that allows linking from corresponding publications and discovery dataset descriptions. The files are backed-up following specific archival procedures. Each data file is stored together with an xml version of the metadata elements on a data storage server. There is a daily backup of this data storage server on tape. A monthly backup each week and a permanent backup each year are stored at different locations.

Each file in the MDA is stored with metadata describing the data file. This documentation guarantees that the data in the files can be interpreted correctly. The required documentation is dependent on the particular data type (publication; CTD; ADCP; meteorology data; image tracking set; biotic trawl data), but generally includes information on:

- who (collected, sampled, processed, wrote);
- when (date or start and end date);
- where (geographical location and coordinates);

- what (biological, chemical, publication, photos);
- how (instruments, serial number);
- conditions of use;
- objectives.

The MDA provides users with three workspaces:

- private workspace: files and metadata are only accessible by the individual user;
- public workspace: files and metadata in this workspace are accessible (at least 'read only' access rights) to everyone who has access to the MDA;
- shared workspace: files and data shared within a scientific research group, project,...

So, depending on a user's personal profile and the folder and workspace they are working in, the type of access rights can be 'read only', 'write' or 'admin'. This means the user can respectively access, edit or manage the files and metadata.

The MDA can be accessed online at <http://mda.vliz.be>.

Dr Hernandez noted that MDA would need some reprogramming before it can be made available for other centres to install. He informed the group that this work could not be scheduled before 2013. Documentation on MDA (a user guide) exists but may need some updating. Dr Hernandez noted that the use of MDA would not be appropriate in all circumstances as most data centres already have a data management system. In these cases a modification of the existing system may be more appropriate.

Dr Hernandez noted that the current maximum file size is 400MB.

### **Conclusions**

**The meeting recommended a survey amongst all IODE NODCs to assess the need for an MDA-type service. VLIZ offered to host a service for interested NODCs on the VLIZ server. This will enable the NODCs to try out and assess the system. Depending on the feedback reprogramming of the application can then be scheduled in 2013 to enable later release of the application to all interested NODCs.**

## **7.3 Discussion on the relationship between researchers and data centres in data publication (includes role of 'sandboxes' like MDA and SFA)**

### **Conclusions**

**The meeting noted that having a tool that the researchers can use (MDA, SFA) and be proactive would be a giant step forward. They can start assembling their data sets and delivering them.**

**The meeting stressed the importance of effective communication between the researchers and the data centres.**

**The meeting concluded that tools such as MDA or SFA can be an effective mechanism to promote the data flow to NODCs: researchers can safely store their data elements in MDA or SFA first. Subsequently, when writing their journal article they can make the data suitable to attaching to the article. Assistance of a data manager will be essential during this stage when also the metadata record has to be written and DOI requested. At this stage the data can also be sent to the NODC for further processing.**

## 8. SESSION 7: DATA VERSIONING ISSUES

### 8.1 Versioning issues in BODC

This Agenda Item was introduced by Dr Roy Lowry.

Data are held in two ways:

1. **Series model:** a physical file holds the data (but not metadata) for a feature instance (profile, time series, trajectory). Two file formats can be used with the series model: (i) NetCDF: data with only parameter names and units of measure as metadata; (ii) ASCII: data plus more complete metadata, including spatio-temporal co-ordinates. URL to complete usage metadata (Oracle report as XHTML). The metadata are stored in an Oracle database.
2. **Samples model:** Both data and metadata are held in an Oracle relational schema. Data are served by formulating an SQL query and supplying the resultant output.

#### Versioning

Physical NetCDF files are formally versioned, and previous versions are preserved. If the NetCDF or any metadata included in an ASCII file is changed then the file is regenerated, its version number is incremented and the previous version is preserved. The usage metadata report isn't versioned, but there are automatically maintained timestamps and audit trails. Changes in the usage metadata report do not trigger regeneration of the ASCII file. Samples data aren't versioned, but there are automatically-maintained timestamps and audit trails

#### Publication Standard Versioning

- Series Model
  - Introduction of version management for usage metadata
  - Introduction of versioned addressing (URLs) for usage metadata
  - Entire ASCII stock would need republishing with versioned addressing links
  - Currently believed feasible
  - Do BODC need to do this?
- Samples Model
  - Automated publication triggered by changes in the relational schema consigned to the 'too hard' basket.
  - Snapshot generation and preservation on demand is one viable option
  - Another (already operational) option is to expose samples model data through the series model

#### Conclusions

**The meeting concluded that modifications of usage metadata constitute a reason to republish the entire dataset because there is a risk that changes in usage metadata can change the interpretation of the data set.**

## 9. CONCLUSIONS

The meeting agreed on the following work plan:

- **Cookbook:** to be prepared by Roy Lowry within 12 months (**November 2012**). This will be included in OceanTeacher.

- **Surveying IODE network:** it is necessary to investigate (i) which IODE NODCs are working on data publication; (ii) if there is interest in data publication in the IODE NODC community; (iii) whether data versioning requirements are satisfied by the IODE NODCs (can previous versions of a data set be retrieved). Mr Helge Sagen will collect this information (in collaboration with the IODE Secretariat) and feed it back to the group. The information could then also be provided to CODATA (**deadline: end of April 2012**)
- **Data versioning:** Roy Lowry/Adam Leadbetter will follow-up this matter with SDN-2. Peter Pissierssens will talk to Sergey Belov for ODP. (**deadline: end of January 2012**)
- **MBLWHOI (WHOAS) repository documentation:** Lisa Raymond/Cynthia Chandler (**deadline: by EGU in April 2012**)
- **Making available MBLWHOI (WHOAS):** see previous item
- **POD: installation on IODE server:** Marc Goovaerts (deadline end November 2011) / Roy Lowry: making available data vault for 3 data sets (**deadline: end of November 2011**)
- **Hosting service by MDA+user guide:** VLIZ (**deadline: throughout 2012**) / **promotion of the service:** IODE project office (**deadline: starting January 2012**)
- **BODC SFA system:** Roy Lowry needs to investigate this further (identify data set for publication).

It was noted that Roy Lowry will retire and he invited someone else to take over the “driving” of the activity. Ed Urban and Lisa Raymond volunteered to coordinate the Data Publication activity with support from the current members of the group. It was agreed to ask Sarah Callaghan to continue involvement.

The meeting agreed to meet again **8-10 October 2012**, possibly in Woods Hole. The participants will include those who are involved in the work plan but we may also invite other experts (for example someone who tried out the applications/services).

## 10. CLOSING OF THE MEETING

The meeting was closed on Friday 4 November 2011 at 12h10. The meeting requested the Secretariat to distribute the draft report for corrections during the next week and to publish the report not later than 14 November 2011.



## **ANNEX I**

### **AGENDA OF THE MEETING**

1. WELCOME AND INTRODUCTION OF THE AGENDA
2. SESSION 1: PILOT PROJECTS AND ASSOCIATED ACTIVITY REPORTS
  - 2.1 WOODS HOLE DATA PUBLICATION ACTIVITIES WITH ELSEVIER AND BCO-DMO
  - 2.2 THE BODC PUBLISHED DATA LIBRARY
  - 2.3 THE PUBLISHED OCEAN DATA REPOSITORY
  - 2.4 EXPANDING THE SCIENTIFIC RECORD: DATA CITATION AND PUBLICATION BY NERC'S ENVIRONMENTAL DATA CENTRES
  - 2.5 OUTREACH ACTIVITIES REPORT. ROUND TABLE DISCUSSION TO CATALOGUE OUTREACH SINCE THE PARIS MEETING (APRIL 2010) AND PROPOSE FUTURE OUTREACH ACTIVITIES
3. SESSION 2: MANAGEMENT OF DIGITAL OBJECTS IN REPOSITORIES
  - 3.1 BACK TO THE IT FUTURE: THE CLOUD AND OPEN SOURCE
4. SESSION 3: THE CODATA TASK GROUP ON DATA CITATION STANDARDS AND PRACTICES
  - 4.1 CODATA TASK TEAM 2011 MEETING REPORT
  - 4.2 DISCUSSION ON FUTURE INTERACTIONS BETWEEN THIS GROUP AND THE CODATA TASK TEAM.
5. SESSION 4: ECONOMIC ASPECTS OF DATA EXCHANGE
  - 5.1 THE ECONOMIC CHALLENGES OF OPEN DATA & INFORMATION EXCHANGE
  - 5.2 DISCUSSION ON ECONOMIC ASPECTS OF DATA AND INFORMATION EXCHANGE
6. SESSION 5: STIMULATION OF DATA PUBLICATION ACTIVITY IN THE INTERNATIONAL OCEANOGRAPHIC COMMUNITY
  - 6.1 DATA PUBLICATION ACTIVITY IN AUSTRALIA
  - 6.2 DATA PUBLICATION ACTIVITY IN NORWAY
  - 6.3 DISCUSSION ON INITIATION AND CO-ORDINATION OF DATA PUBLICATION ACTIVITIES IN SEADATANET AND THE IODE DATA CENTRE NETWORK.
7. SESSION 6: ENGAGEMENT OF THE SCIENTIFIC COMMUNITY
  - 7.1 BODC SECURE FILE ARCHIVE
  - 7.2 VLIZ MARINE DATA ARCHIVE SYSTEM
  - 7.3 DISCUSSION ON THE RELATIONSHIP BETWEEN RESEARCHERS AND DATA CENTRES IN DATA PUBLICATION (INCLUDES ROLE OF 'SANDBOXES' LIKE MDA AND SFA)
8. SESSION 7: DATA VERSIONING ISSUES
  - 8.1 VERSIONING ISSUES IN BODC
9. CONCLUSIONS
10. CLOSING OF THE MEETING



## ANNEX II

### LIST OF PARTICIPANTS

Dr Michael J. ACKERMAN  
Assistant Director  
US National Library of Medicine  
8600 Rockville Pike  
Bethesda MD 20894  
United States  
Tel: +1-301-402-4100  
Fax: +1-301-402-4080  
Email: [mackerman@mail.nih.gov](mailto:mackerman@mail.nih.gov)

Dr Sarah CALLAGHAN  
British Atmospheric Data Centre  
STFC - Rutherford Appleton  
Laboratory  
Chilton  
Didcot  
Oxfordshire  
OX11 0QX  
United Kingdom  
Tel: +44-1235-445770  
Email: [sarah.callaghan@stfc.ac.uk](mailto:sarah.callaghan@stfc.ac.uk)

Ms Cynthia CHANDLER  
Informatics Specialist  
MS 36, Woods Hole, MA, 02543  
United States  
Tel: +1-508-289-2765  
Fax: +1-508-457-2169  
Email: [cchandler@whoi.edu](mailto:cchandler@whoi.edu)

Mr Marc GOOVAERTS  
Director  
Universiteit Hasselt  
Campus Diepenbeek  
Agoralaan  
Gebouw D  
3590 Diepenbeek  
Belgium  
Tel: +32-11-26 81 24  
Fax: +32-11-26 81 24  
Email: [marc.goovaerts@uhasselt.be](mailto:marc.goovaerts@uhasselt.be)

Dr Francisco HERNANDEZ  
Manager Datacentre  
(VLIZ) Flanders Marine Institute  
Vlaams Instituut voor de Zee  
Wandelaarkaai 7  
8400 Oostende

Belgium  
Tel: +32-59- 342130  
Fax: +32-59-342131  
Email: [francher@vliz.be](mailto:francher@vliz.be)

Mr Mark KURTZ  
Director of Business Development  
BioOne  
21 Dupont Circle NW  
Suite 800  
Washington DC 20036  
United States  
Tel: +1-202-296-1605 x5  
Email: [mkurtz@arl.org](mailto:mkurtz@arl.org)

Dr Adam LEADBETTER  
Data Scientist  
British Oceanographic Data Centre  
Joseph Proudman Building  
6 Brownlow Street  
Liverpool  
L3 5DA  
United Kingdom  
Tel: +44-151-7954878  
Email: [alead@bodc.ac.uk](mailto:alead@bodc.ac.uk)

Dr Roy LOWRY  
British Oceanographic Data Centre  
Joseph Proudman Building  
6 Brownlow Street  
Liverpool  
L3 5DA  
United Kingdom  
Email: [rkl@bodc.ac.uk](mailto:rkl@bodc.ac.uk)

Dr Holly MILLER  
Library Director  
MBLWHOI Library  
Marine Biological Laboratory  
7 MBL Street, Woods Hole, MA  
02540  
United States  
Tel: +1-508-289-7632  
Email: [hmiller@mbi.edu](mailto:hmillers@mbi.edu)

Dr Gwenaëlle MONCOIFFE  
Marine Data Scientist  
British Oceanographic Data Centre

Joseph Proudman Building  
6 Brownlow Street  
Liverpool  
L3 5DA  
United Kingdom  
Email: [gmon@bodc.ac.uk](mailto:gmon@bodc.ac.uk)

Ms Cathy NORTON  
Senior Library Scholar  
MBLWHOI Library  
7 MBL ST  
Woods Hole MA 02543  
United States  
Tel: +1-508-5600617  
Email: [cnorton@mbi.edu](mailto:cnorton@mbi.edu)

Ms Linda PIKULA  
Chair IODE GE-MIM  
Regional Librarian  
NOAA Central and Regional Libraries  
4301 Rickenbacker Causeway  
Miami Florida 33149  
United States  
Tel: +1-305-3614429  
Fax: +1-305-3614552  
Email: [linda.pikula@noaa.gov](mailto:linda.pikula@noaa.gov)

Dr Roger PROCTOR  
Director  
Integrated Marine Observing System,  
eMarine Information Infrastructure  
University of Tasmania  
Private Bag 21  
Hobart Tasmania 7001  
Australia  
Tel: +61-3-62261977  
Fax: +61-3-62262997  
Email: [roger.proctor@utas.edu.au](mailto:roger.proctor@utas.edu.au)

Ms Lisa RAYMOND  
Associate Library Director  
MBLWHOI Library  
Woods Hole Oceanographic Institution  
Woods Hole MA 02543  
United States  
Tel: 1 508 289 3557  
Fax: 1 508 457 2156  
Email: [lraymond@whoi.edu](mailto:lraymond@whoi.edu)

Mr Helge SAGEN  
Head of Norwegian Marine Datacentre  
(NODC)  
Institute of Marine Research, Bergen  
Nordnesgaten 33  
Postboks 1870 Nordnes

5817 Bergen  
Norway  
Tel: +4755238447  
Fax: +47-55238531  
Email: [Helge.Sagen@imr.no](mailto:Helge.Sagen@imr.no)

Dr Ed URBAN  
Executive Director  
Scientific Committee on Oceanic  
Research  
College of Marine and Earth Studies  
Robinson Hall  
University of Delaware  
Newark DE 19716  
United States  
Tel: +1-302-831-7011  
Fax: +1-302-831-7012  
Email: [ed.urban@scor-int.org](mailto:ed.urban@scor-int.org)

### **IOC Secretariat**

Mr Peter PISSIERSSENS  
Head, IOC Project Office for IODE,  
Oostende, Belgium  
UNESCO/IOC Project Office for  
IODC  
Wandelaarkaai 7 - Pakhuis 61  
B-8400 Oostende  
Belgium  
Tel: +32-59-340158  
Fax: +32-59-79 5220  
Email: [p.pissierssens@unesco.org](mailto:p.pissierssens@unesco.org)





## **Intergovernmental Oceanographic Commission (IOC)**

United Nations Educational, Scientific and Cultural Organization (UNESCO)

1, rue Miollis, 75732 Paris Cedex 15, France

Tel: +33 1 45 68 39 83

Fax: +33 1 45 68 58 12

<http://ioc.unesco.org>

IOC Project Office for IODE

Wandelaarkaai 7

8400 Oostende, Belgium

Tel: +32 59 34 21 34

Fax: + 32 59 34 01 52

<http://www.iode.org>