



Ocean Data Interoperability Platform

Deliverable D2.4: Minutes of the 2nd ODIP Workshop

Work Package	WP2	ODIP workshops
Author (s)	Sissy Iona	HCMR
Author (s)	Dick Schaap	MARIS
Author (s)		
Author (s)		
Authorized by		
Reviewer	Adam Leadbetter	NERC / BODC
Doc Id	ODIP_WP2_D2.4	
Dissemination Level	PUBLIC	
Issue		
Date	1 March 2014	



Document History				
Version	Author(s)	Status	Date	Comments
0.1	Sissy Iona (HCMR) Dick Schaap	DRAFT	08 February 2014	First draft
0.2	Dick Schaap	DRAFT	18 February 2014	Editing
0.3	Dick Schaap	DRAFT	01 March 2014	Editing: incorporating feedback from partners
1.0	Helen Glaves	FINAL	23 April 2014	Review
1.1	Adam Leadbetter	FINAL	28 May 2014	Review
2.0	Helen Glaves	FINAL	28 May 2014	Sign-off



Contents

Executive Summary	4
1 Introduction.....	5
2 List of Participants	6
3 Workshop Agenda.....	8
4 Workshop proceedings	14
5 Proposed follow-up and actions.....	39
Appendix A Terminology.....	41

Executive Summary

The 2nd Ocean Data Interoperability Platform (ODIP) Workshop took place on 3 - 6 December 2013 at the Scripps Institution of Oceanography (SIO), University of California, San Diego (UCSD) in La Jolla, USA. Almost 50 people from Europe, USA and Australia attended including representatives from the Intergovernmental Oceanographic Commission's International Oceanographic Data and Information Exchange (IODE). In addition to those attending the meeting in person a number of additional participants joined the meeting remotely. The discussion during the workshop focused on the three ODIP prototype projects that had been formulated as a result of the 1st ODIP Workshop which took place in February 2013. Two further discussion topics were included in the programme for the workshop that had been identified and prioritized by the ODIP project partners.

The main topics addressed during the workshop were:

- ODIP Prototype Project 1: Establishing interoperability between SeaDataNet Common Data Index, US National Oceanographic Data Centre and Australian Integrated Marine Observing System Marine Community Profile data discovery and access services using a brokering solution (including moving towards interacting with the IODE Ocean Data Portal and Global Earth Observation System of Systems, GEOSS, portals)
- ODIP Prototype Project 2: Establishing a common deployment and interoperability between cruise summary reporting systems in Europe, the USA and Australia, including an assessment of the potential for using the open source GeoNetwork solution (and moving towards interacting with the Partnership for Observation of the Global Ocean (POGO) portal)
- ODIP Prototype Project 3: Establishing a prototype for a Sensor Observation Service (SOS) and formulating common Observations & Measurements (O&M) and Sensor Model Language (SensorML) profiles for selected sensors installed on vessels and in real-time monitoring system
- Vocabularies
- Data publication and citation

This deliverable reports on the organisation, proceedings and outcomes of the 2nd ODIP workshop. It concludes with a list of additional actions that have come out of the workshop discussions in addition to the activities planned for the further development of the three ODIP prototype projects and planned deliverables. The development of the prototype interoperability solutions is a joint effort undertaken by ODIP partners leveraging on the activities of current regional projects and initiatives such as SeaDataNet (EU), IMOS (Australia) and Rolling Deck to Repository (R2R,USA). The workshop also identified the potential for a small number of additional ODIP prototype projects which will be explored further by partners.

The 3rd ODIP Workshop is planned to take place at the Australian Institute for Marine Science (AIMS) in August 2014. The focus of this meeting will be on presenting and discussing the progress of the ODIP prototypes under development as well as addressing a number of additional relevant topics such as big data, cloud computing and data quality.



1 Introduction

The Ocean Data Interoperability Platform (ODIP) project aims to establish an EU / USA / Australia/ IOC-IODE coordination platform, the objective of which is to develop interoperability between existing regional marine e-infrastructures to create a global framework for marine and ocean data management, and to demonstrate this coordination through several joint EU-USA-Australia-IOC/IODE prototypes that demonstrate effective sharing of data across scientific domains, organisations and national boundaries.

ODIP will convene four workshops during which the project partners and other invited experts will review and compare existing marine data standards in order to identify major differences between them, and propose how to overcome these through the development of interoperability solutions and/or common standards.

The 2nd ODIP Workshop took place on 3 - 6 December 2013 at the Scripps Institution of Oceanography (SIO), University of California San Diego in La Jolla, USA. The agenda for the meeting has been documented in deliverable D2.3 *2nd ODIP workshop*. The programme for this workshop focused on the three ODIP prototype interoperability development tasks that were initially formulated as a result of the first ODIP workshop which took place in February 2013. Two additional topics for discussion at the workshop were also included in the agenda which were selected in consultation with the ODIP project partners.

2 List of Participants

As part of the community building for the ODIP project an extensive mailing list of experts representing the consortium partners and their associated projects and initiatives is being maintained. This list was used in conjunction with the ODIP website to invite participants for the second ODIP workshop. A draft agenda for the workshop was circulated and updated in consultation with partners. As a result 54 attendees from 11 countries took part in the 2nd ODIP Workshop (7 of them participated remotely by WebEx). They were:

Robert ARKO	LDEO, United States
Irina BASTRAKOVA	Geoscience Australia, Australia
Rita BAUER	SIO-UCSD, USA
Sergey BELOV	RIHMI-WDC, Russian Federation
Justin BUCK	NERC-BODC, UK
Thierry CARVAL	IFREMER, France
Cyndy CHANDLER	WHOI, USA
Yanning CHEN	RPI/TWC, USA
Dru CLARK	SIO-UCSD, USA
Simon COX	CSIRO, Australia
Karien de CAUWER	RBINS-MUMM, Belgium
Steve DIGGS	SIO-UCSD, USA
Paolo DIVIACCO	OGS, Italy
Nkem DOCKERY	FSU COAPS, USA
Ben DOMENICO	UNIDATA, USA
Renata FERREIRA	SIO-UCSD, USA
Linyun FU	RPI/TWC, USA
Jeff GEE	SIO-UCSD, USA
Helen GLAVES	NERC-BGS, UK
John GRAYBEAL	Marine Explore, USA
Sissy IONA	HCMR, Greece
Matthias LANKHORST	SIO-UCSD, USA
Thomas LOUBRIEU	IFREMER, France
Roy LOWRY	BODC, UK
Angelos LYKIARDOPOULOS	HCMR, Greece
Sebastian MANCINI	UTAS, Australia
Steve MILLER	UCSD, USA
Stefano NATIVI	CNR-IAA, Italy
Mohamed OUBERDOUS	ULG, Belgium



Elena PARTESCANO	OGS, Italy
Jay PEARLMAN	IEEE, USA
Francoise PEARLMAN	IEEE, USA
Anna POTTER	Geoscience Australia, Australia
Gael QUEMENER	IFREMER, France
Dick SCHAAP	MARIS, Netherlands
Serge SCORY	MUMM, Belgium
Adam SHEPHERD	WHOI, USA
Jean-Marc SINQUIN	IFREMER, France
Shawn SMITH	FSU COAPS, USA
Derrick SNOWDEN	NOAA/US IOOS, USA
Jordi SORRIBAS	CSIC, Spain
Karen STOCKS	SIO-UCSD, USA
Aaron SWEENEY	SIO-UCSD, USA
Tobias SPEARS	DFO-BIO, Canada
Rob VAN EDE	TNO, Netherlands
Sytze VAN HETEREN	TNO, Netherlands
Matteo VINCI	OGS, Italy
Lesley WYBORN	Geoscience Australia, Australia

The participants in the 2nd ODIP workshop represented a good cross-section of the relevant EU, USA and Australian regional infrastructure projects and initiatives that are stakeholders in the ODIP project. There was also representation from the international IOC-IODE Ocean Data Portal (ODP) project.

3 Workshop Agenda

The main focus of the 2nd ODIP workshop was the progress being made on the three prototype interoperability solution development tasks that are currently being undertaken by the ODIP partners. The meeting programme included a plenary session followed by a discussion/working session for each task. The ODIP prototype development task leaders prepared an agenda for their dedicated session which included relevant presentations and speakers giving an overview of the individual prototype project objectives and the current state of the art in the three regions (Europe, USA and Australia).

Two additional discussion topics were included in the meeting programme. These topics were identified and prioritized by the ODIP project partners prior to the meeting. A leader was also selected for each of these topics who were responsible for compiling a programme of relevant presentations and speakers from the three regions involved in the project.

The workshop agenda was circulated to all ODIP partners by e-mail before the workshop and also published on the ODIP website.

Workshop Topics

Session	Title	Leader
1	ODIP Prototype project 1	Dick Schaap
2	ODIP Prototype project 2	Bob Arko
3	ODIP Prototype project 3	Sebastian Mancini
4	Vocabularies	Roy Lowry, Simon Cox, John Graybeal
5	Data publishing and citation	Cyndy Chandler, Justin Buck, Lesley Wyborn
6	Feedback on sessions	Session leaders
7	Workshop wrap-up	Helen Glaves



11.45 - 12.00 Potential activities and work plan for ODIP prototype project 1
Dick Schaap (EU)

12.00 - 13.00 Discussion *Led by Dick Schaap (EU)*

Session 2

ODIP Prototype Development Project 2

14.00 - 14.15 Introduction *Bob Arko (USA)*

14.15 - 16.00 Presentations and discussion:

- SeaDataNet CSR, POGO and Eurofleets
Sissy Iona (EU)
- Use of GeoNetwork for harvesting CSR
Thomas Loubrieu (EU)
- German data portal *Ana Macario (EU)*
- Cruise metadata in Australia (MNF)
Sebastien Mancini (Australia)
- R2R - Cruise reports in USA *Bob Arko (USA)*

16.00 - 17.00 Discussion *Led by Bob Arko (USA)*

Wednesday, 4 December 2013

Session 3

ODIP Prototype Development Project 3

- 9.00 - 9.30 Introduction *Sebastien Mancini (Australia)*
- 9.30 - 12.00 Presentations and discussion:
- What's happening in Australia
Sebastien Mancini (Australia)
 - What's happening in the USA
Derrick Snowden (USA)
 - What's happening in Europe
Jordi Sorribas (EU)
 - SOS and 52North *Simon Jirka (52North)*
 - O&M and SensorML update
Simon Cox (Australia)
- 12.00 - 13.00 Discussion *Led by Sebastien Mancini (Australia)*
- 13.00 - 14.00 Lunch

Session 4

Vocabularies

- 14.00 - 15.30 Presentations and discussion
Led by Roy Lowry (EU), Simon Cox (Australia) and John Graybeal (USA)
- What's happening in Europe
Roy Lowry (EU)
 - What's happening in Australia
Simon Cox (Australia)
 - What's happening in USA
John Graybeal (USA)



-
- 15.30 - 16.00 Checking the status of actions from ODIP 1st Workshop
Led by Roy Lowry (EU)
- Harmonisation for event logging on research vessels
Karien de Cauwer (EU) and Cyndy Chandler (USA)
 - Possible use of GMX:Anchor
Thomas Loubrieu (EU)
- 16.00 - 17.00 Discussion
*Led by Roy Lowry (EU), Simon Cox (Australia)
and John Graybeal (USA)*

Thursday, 5 December 2013

Session 5

Data publication and citation

- 9.00 - 13.00 Presentations and discussion *Led by Cyndy Chandler (USA), Justin Buck (EU)
and Lesley Wyborn (Australia)*
- ANDS Research Data Citation
Lesley Wyborn (Australia)
 - ESRI developments *Dawn Wright (ESRI)*
 - What's happening in Europe
Justin Buck (EU)
 - Data Citation in USA *Cyndy Chandler (USA)*
- 12.00 - 13.00 Discussion *Led by Cyndy Chandler (USA), Justin Buck (EU)
and Lesley Wyborn (Australia)*
- 13.00 - 14.00 Lunch

Vocabularies (continued)

- 14.00 - 17.00 Discussion *Led by Roy Lowry (EU), Simon Cox (Australia) and
John Graybeal (USA)*

Friday, 6 December 2013

Session 6

Feedback from sessions

9:00 – 10:00 ODIP prototype development projects

Feedback from each group on activities during the workshop:

- ODIP 1 *Dick Schaap (EU)*
- ODIP 2 *Bob Arko (USA)*
- ODIP 3 *Sebastien Mancini (Australia)*

10:00 – 10:30 Session 4 report:

- Vocabularies *Simon Cox (Australia) and Roy Lowry (EU)*

Session 7

Workshop wrap-up

10:30 – 11:00 Further potential prototype projects (D3.4)

Helen Glaves (ODIP co-ordinator)

11:30 – 12:00 Workshop summary

Dick Schaap (ODIP Technical co-ordinator)

12:00 – 12:30 Plans for next 8 months

Helen Glaves (Co-ordinator)

12:30 – 12:45 3rd ODIP workshop

Sissy Iona (EU)

12:45 – 13:00 Closing remarks

Helen Glaves (Co-ordinator)

13:00 Closure of the 2nd ODIP Workshop

4 Workshop proceedings

4.1 Introduction

The 2nd ODIP Workshop was held on the 3 - 6 December 2013 at the Scripps Institution of Oceanography. The meeting was opened by Jeff Gee, Deputy Director of SIO-UCSD who welcomed everyone to the meeting and gave background about SIO, its organisation and activities, including involvement in international oceanography projects and research infrastructures. This was followed by a welcome from Karen Stocks of SIO-UCSD, who together with Rita Bauer (SIO-UCSD) explained the local logistics for the workshop and the arrangements regarding remote WebEx participation and recording of the presentations during the meeting. (The recordings made during the workshop have been made available on-line via the IODE portal http://www.iode.org/index.php?option=com_oe&task=viewEventRecord&eventID=1420)

Helen Graves (HG), the ODIP project coordinator, welcomed all participants to the meeting. HG noted that there was a good cross section of attendees from across the project consortium including representatives from IODE. Participants present in the room and also joining on-line were then invited to introduce themselves and give a brief statement regarding their role in the project.

4.2 Overview of the ODIP project

Helen Graves (HG) presented an overview of the ODIP project, its concept, objectives, activities and structure, as well as progress to date. HG highlighted the need for extensive dissemination of the project tasks and objectives to the wider marine community. This was illustrated by the active involvement in the Research Data Alliance (RDA), Marine Data Harmonization interest group by a number of the project participants HG also summarized the 1st ODIP workshop, which addressed six discussion topics and resulted in an extensive list of potential actions which have been carried forward to form the basis for the formulation of the three ODIP prototype interoperability development tasks which are currently being undertaken by selected partners.

4.3 ODIP prototype interoperability development tasks

4.3.1 Introduction

Dick Schaap (DS), ODIP Technical Co-ordinator, summarised the drivers for the ODIP activities. This includes comparing and analysing standards and best practices relevant to the management of marine data in the EU, USA and Australia; developing interoperability solutions between existing infrastructures; developing common standards, where possible, especially in new domains such as Sensor Web Enablement (SWE); establishing global discovery and exchange of marine and ocean data, between regional infrastructures including the GEOSS, Ocean Data Portal (ODP) and POGO services. DS also highlighted the technical approach used for the formulation of the ODIP prototype development tasks, which will be implemented by leveraging on the activities of current regional projects and initiatives such as SeaDataNet (EU), IMOS (Australia) and R2R (USA). ODIP will provide the communication and exchange platform where partners can meet, discuss and align their development activities. DS underlined that in practice this approach is relatively abstract and

challenging, especially considering the large geographical distances between those contributing to the individual tasks, the different time zones involved and limited number of opportunities for face to face meetings.

The three ODIP prototype projects that have been formulated and approved by the partners for development so far are:

- **ODIP 1:** Establishing interoperability between the SeaDataNet CDI, US NODC, and IMOS MCP data discovery and access services using a brokering solution, including moving towards interacting with the IODE-ODP and GEOSS portals
- **ODIP 2:** Establishing interoperability between cruise summary reporting systems in Europe, the USA and Australia including an evaluation of the potential use of the GeoNetwork software, and working towards delivery of cruise summary reports to the POGO portal
- **ODIP 3:** Establishing a prototype for a Sensor Observation Service (SOS) and formulating common O&M and SensorML profiles for selected sensors (SWE) installed on vessels and in real-time monitoring systems.

The theme of each ODIP prototype interoperability solution development task is quite different. ODIP 1 aims to deliver marine metadata and data from existing data infrastructures in Europe, the USA and Australia to the ODP and GEO portals through the implementation of a brokering solution.

ODIP 2 aims to contribute content from the cruise summary reporting (CSR) systems and cruise programmes from Europe, USA and Australia to the Partnership On Global Observations (POGO) portal. This task includes research on harvesting information from common vocabularies with an assessment of the potential use of GeoNetwork, innovations using a linked data approach and direct linking to on-board cruise event logging systems.

ODIP 3 is aiming to build up experience with sensor web enablement (SWE) which includes: 1) developing and demonstrating a number of common standards; 2) understanding the technology, and 3) developing guidance for the deployment of SWE

It should be noted that the ODIP 1 and 2 prototype development tasks will not deliver project specific portals which could potentially compete with ongoing initiatives, but will instead seek to align the development and activities of the regional marine data infrastructures with existing global portal initiatives such as IODE-ODP, GEOSS and POGO. The ODIP 3 prototype development task which is focused on sensor web enablement for marine monitoring systems has the potential to make a major contribution to the development of international standards in this area.

To ensure there is a balanced approach to the development of the prototype interoperability solutions, the three regions represented in the project (Europe, USA, and Australia) are all contributing to each of the three ODIP development tasks. In addition each region is also responsible for leading one of the development tasks:

- ODIP 1 by Europe through SeaDataNet (MARIS)
- ODIP 2 by the USA through R2R (LDEO)
- ODIP 3 by Australia through AODN (UTAS)

4.3.2 Session 1: ODIP Prototype Development Task 1

DS outlined the approach for ODIP 1 which will be based upon the regional data discovery and access services of SeaDataNet CDI service (Europe), US NODC service (USA), and MCP service (Australia). The aim is to establish a mechanism for the exchange of metadata between these three regional services and also delivery to the global GEO and IODE ODP portals. A decision has been taken to investigate the use of a brokering service concept.

Base services:

- **SeaDataNet Common Data Index (CDI) service:**

DS presented the key features of the SeaDataNet Common Data Index (CDI) service in Europe, which currently has more than 90 connected data centres from 34 European countries and includes more than 1.3 million data sets with global coverage for a wide range of marine disciplines. Further development and population of this service takes place through engagement in multiple European projects. The CDI is a marine metadata profile which uses the ISO 19115 content model. The associated XML encoding now follows the ISO 19139 standard and is fully INSPIRE compliant. Use is made of controlled vocabularies where ever possible. Tools are provided for preparing metadata and data entries in agreed formats e.g. ODI ASCII, NetCDF (CF) etc. Other services are also available for importing CDI entries and the discovery and access of data sets that are managed by the distributed network of data centres. Access to the data takes place by means of a shopping basket mechanism with a tracking and tracing system of request processing and downloading. Various interoperability services are available for the metadata, such as WMS, WFS, CS-W as well as service bindings in the CDI for supporting sensor web enablement (SWE) and visualisation tools. Further work is on going for OpenSearch and GeoNetwork metadata harvesting. The CDI shopping mechanism now also features a robot harvester for maintaining buffers of specific data sets with controlled access for individual user communities.

- **US National Oceanographic Data Centre (NODC) service:**

Ken Casey (KC) presented the key features of the US NODC service in the USA which includes in-situ, satellite and modelled data. All of these collections have standardised FGDC and ISO 19115-2 metadata. A distinction is made between granules (comparable to CDI) and collections (groups of granules) which are linked by identifiers. The NODC has more than more than 150 terabytes of data which represents 150,000 collections with a global coverage for a wide range of marine disciplines. These are made available via OAI-PMH. Both collections and granules are also accessible by a range of services: Geoportal REST API and web interface, CSW, and OpenSearch. The related data sets are accessible via HTTP, FTP, LAS, ArcGIS, OpenDAP, and TDS. Data are provided in range of formats, but CF-compliant netCDF-4 with templates for feature types is preferred. There is currently no user authentication/registration system, but it may be needed in the future for some specific services. US NODC maintains and makes use of a range of vocabularies, and is also experimenting with SKOS. Options for sharing the metadata are: harvesting directly from WAFs, retrieving multiple response formats (HTML, JSON, etc.) from Geoportal REST API and web interface, and using OAI-PMH for collections. US NODC has had good results with CS-W performance for > 1 million granules.

US NODC also maintains a directory of people. A number of similar systems were also discussed during the meeting including ORCID, IODE OceanExpert.etc.

- **AODN Marine Community Profile service:**

Sebastien Mancini (SM) presented the key features of the Australian Ocean Data Network (AODN) MCP service in Australia. AODN is a national network of organisations (government and science) which New Zealand has also recently joined. A dedicated web portal provides an overview and access to the data sets which have been collected and are currently managed by the AODN. The Marine Community Profile (MCP) metadata format is based upon the ISO 19115 and ISO 19139 standards. GeoNetwork is used for the discovery (CSW) and mapping (WMS) services, while WFS and OpenDAP (TDS) are used for downloading of the data. GeoNetwork is also used by the AODN portal for harvesting data in NetCDF (CF) format from the data providers. The AODN portal provides open access with no registration necessary for users. The discovery service is fully functional but currently lacks any form of ‘shopping’ mechanism for data access. AODN builds and maintains its own vocabularies, but also adopts others such as the Global Change Master Directory (GCMD) and those from the British Oceanographic Data Centre (BODC). Development of a faceted search is underway. At present the CSW works at a collection level which encompasses around 20,000 data collections. SM highlighted how the concept of a parameter is used in AODN metadata by adopting the Observations & Measurements (O&M) model.

Simon Cox (SC) suggested that use of a Universal Identifier (UID) between the metadata and associated data sets should be helpful for all discovery and access services.

Target services:

- **GEOSS portal**

Jay Pearlman (JP) presented the GEOSS portal which is one of the global data services that will potentially receive dynamic population from each of the three regional discovery and access services as part of the ODIP 1 prototype development task. JP provided background on the GEO initiative (2002 G8) which is charged with developing the Global Earth Observation System of Systems (GEOSS) for nine societal benefit areas. This initiative is based on the premise that “*Earth Observation Systems should be coordinated and shared internationally to answer Society’s need for informed decision making*”.

GEO currently has eleven ‘Communities of Practice’ including the ocean and coastal zone, as well as others which also have direct relevance for the marine domain. GEO is not a funding mechanism but its global cooperation inspires research programmes and can be seen as a framework to promote international collaboration. JP highlighted the recent Blue Planet initiative which provides significant opportunities in this area and which will underpin calls in funding programmes such as the European Commission’s Horizon 2020.

The GEOSS portal makes use of the GEO Common Infrastructure which provides the Discovery and Access Broker (DAB) and resource registration for Earth observation data, information and services for the nine societal areas. At present GEOSS has 20 major data infrastructures connected, including SeaDataNet. The portal gives an overview and potential access for around 7 million data collections and 65 million data granules. Work is ongoing to provide access to additional collections and granules, and further systems development is also being undertaken by CNR (Italy).

Simon Cox (SC) indicated that GEOSS is adopting the Observations & Measurements model which uses vocabularies for observations and also supports connections to other vocabularies.

- **IODE Ocean Data Portal (ODP)**

Tobias Spears (TS) and Sergey Belov (SB) presented the IODE Ocean Data Portal (ODP) which is another global service that will receive dynamic population from each of the three regional data discovery and services as a result of the activities in the ODIP 1 prototype development task.

The International Oceanographic Data and Information Exchange (IODE) of UNESCO-IOC was established in 1961 and strives to enhance marine research, exploitation, and development through exchange of data and information between member states.

The Ocean Data Portal initiative started in 2007 to facilitate discovery, evaluation and access to marine data, and delivers services and facilities for both data users and providers. The portal makes use of GeoNetwork and dedicated ODP services for the purposes of data discovery and supports CSW 2, OAI-PMH, SRU, OGC WxS, WebDav, THREDDS (GeoNetwork based), HTTP REST, SOAP, JSON and RSS/ATOM protocols.

The ODP vocabulary is based on the SeaDataNet vocabularies with some extensions for climatology and other fields. Data access comprises both pull and push services with data pulling facilitated by HTTP REST and SOAP (online mode) and data pushing using FTP and email (online, delayed). The metadata profile is based on the ISO19115/19139 group of standards. The ODP also has native support for single sign-on (SSO) authorization and authentication but this is not enabled yet. For the purposes of ODIP 1 the three regional services could potentially be considered as IODE-ODP Associated Data Units (ADU) which includes metadata exchange with data access via the metadata (URI-based via portal or directly). The metadata exchange could use CSW.

SB provided background on the technical development of the ODP infrastructure which is undertaken through an in-kind contribution by the Russian Federation. It is largely based upon the Unified State System of Information on the World Ocean – ESIMO (Russia) which currently connects 35 national organizations including the Russian Academy of Science.

ODIP prototype 1: approach

ODIP Prototype 1 use will be made of the brokerage service concept (at metadata level) to develop interoperability between the regional data infrastructures.

- **Brokerage service concept**

Stefano Nativi (SN) presented the brokerage service concept. There are many multi-disciplinary and multi-organisational data and information infrastructures which are heterogeneous in their domains, interoperability and organisation. The aim is to realise a community of communities using the systems of systems approach. Despite the large-scale heterogeneity of existing systems this can be achieved by building on the autonomous infrastructures and using two main interoperability implementations: 1) standardization by means of federation and 2) intermediation by means of brokerage.

The brokerage approach has many benefits, such as lowering barriers for acceptance by existing repositories, accelerating interconnection between services and removing the need to impose common specifications. The brokering concept was developed further in the EU FP7 EURO-GEOSS project and then adopted for GEOSS with the results highlighted by Jay Pearlman in his earlier presentation on the GEO initiative. The concept is being further developed further in the USA by the National Science Foundation (NSF) funded Advanced Cooperative Arctic Data and Information Service (ACADIS) project (<http://nsidc.org/acadis/>) the NSF's Brokering Building Blocks for Earth Cube BCube project (<http://workspace.earthcube.org/bcube>) and the EU FP7 funded GEO-WOW project (<http://www.geowow.eu/>).

Simon Cox (SC) reiterated that for this type of approach it is very important to make use of Universal Identifiers (UID) between the metadata and associated data sets. He explained that GEOSS originally tried to register data providers and their services, but that this proved to be very complex and therefore unworkable. The brokerage concept solved this by being more an engineering approach. SC also emphasised the need to analyse the rules for granularity: collections versus granules.

Shawn Smith agreed that the issue of data granularity is a very interesting topic and proposed that a number of use cases should be defined and analysed which should include existing practices within the three regional services.

- **Potential activities and work plan for ODIP 1**

Dick Schaap (DS), the leader of ODIP 1, presented a proposed work plan for the further development and implementation of the prototype interoperability solution. The following activities will be required:

- Discussion and choice of common metadata model and supporting vocabularies as output for the brokerage service
- Mapping regional metadata models and supporting vocabularies onto a common model for each of the three regions to feed the brokerage converters
- Setting up web services (CSW or REST / SOAP) at each of the regional portals to support metadata exchange with the brokerage services
- Analysing the uptake of output from the brokerage service by the ODP and GEOSS portals
- Analysing the granularity of exchange data
- Analyse how the regional portals can receive and integrate the brokerage output from other regions
- Analyse how to overcome possible complications of authentication, authorization and access services

This draft work plan was generally well received by partners but, following discussion, it was concluded that some of these actions require further refinement. It was agreed that DS will expand the work plan further to include time lines, actors, expected results and planned interactions with regional and global initiatives. This will also provide input for ODIP deliverable *D3.2 Results and conclusions from prototype analysis*.

4.3.3 Session 2: ODIP prototype development task 2

Bob Arko (BA) introduced ODIP 2 which aims to establish interoperability between the regional cruise summary reporting systems in Europe, the USA and Australia, possibly using a GeoNetwork solution, with the objective of delivering cruise summary reports from these services to the POGO portal. This will enhance data discovery by filling the existing gaps in the global portals, for example POGO currently has very few CSRs from the US system. It will also re-establish cruise identifiers as reference points for distributed data sets and provide hooks into the semantic web. ODIP 2 also aims to stop schema creep by adopting a common schema and vocabularies, and assess whether this works in practice.

- **SeaDataNet CSR, POGO and Eurofleets**

Sissy Iona (SI) presented the Cruise Summary Reporting (CSR) and Cruise Programmes (CP) which are maintained in Europe. The CSR gives a low level inventory of cruises with reference to observations and also provides a tool for cruise tracking.

CSR is part of the SeaDataNet metadata infrastructure which currently contains in excess of 53.000 entries from more than 2000 research vessels going back as far as 1873. It also includes historic European CSRs held in the ICES database from 1960 onwards as well as cruise summary reports for a number of US research vessels while working in European waters.

The CSR has recently been upgraded to the ISO 19139 standard schema and makes use of the SeaDataNet controlled vocabularies. It is based upon the ISO 19115 and 19115-2 (for measurements and samples part) content models. The format and interface also supports linking to track charts (image / WMS) and station lists. The CSR web map service (WMS), which is currently being piloted, will display cruise summary reports that have been supplied with their navigation tracks in GML. Work to harvest cruise summary reports in XML format from the data centres using GeoNetwork is also on-going. The XML cruise summary reports entries are also provided as web service and are available in the SeaDataNet CDI service. This facilitates retrieval of data sets from specific cruises. The SeaDataNet cruise summary report entries can be managed by an online content management system or by using the MIKADO XML editing software. The central CSR directory and services are managed by the SeaDataNet partner Bundesamt für Seeschifffahrt und Hydrographie (BSH) from Germany. The cruise programme (CP) entries can also be gathered by online CMS or by CSV file exchange or by MIKADO XML. Central management is done by BODC while MARIS takes care of the portal and services.

CSR is also used by the POGO research vessel portal. The European Centre for Information and Marine Science and Technology (EurOcean) maintains a global directory of research vessels which are linked by the ICES ship codes to the CSR. The POGO portal which is managed by some of the SeaDataNet partners (BODC, MARIS and BSH) is focused on ocean-going research vessels (length >60 metres). It provides directories for research vessels, cruise summary reports (CSR) and cruise programmes (CP). The details of planned cruises are captured by BODC from global research vessel operators using a simple CP format.

As part of the EU FP7 EUROFLEETS project work is ongoing to establish a streamlined flow of cruise data information directly from the shipboard systems into the EVIOR portal and to get direct input from European research vessel operators. The EVIOR portal gives access to the directory of European research vessels, cruise summary reports and cruise programmes for all European research vessels. In addition it features a prototype for a Dynamic Vessel Tracking and Events System using the onboard systems.

During the discussion which followed, Roy Lowry (RL) explained that every data centre includes its Local_CSR_ID as a unique identifier in the cruise summary report (CSR) XML. The central master (BSH) then uses the combination of Local_CSR_ID and EDMO_code for the CSR author to manage unique CSR entries at the central portal and also adds a Central_CSR_ID for its services. The Central portal provides web services with Local_CSR_IDs for internal use and with Central_CSR_IDs for external use. These IDs are persistent and used for recognising updates and new CSR additions. BSH also provides the European cruise summary reports to ICES. However ICES maintains its own CSR ID numbering system and various countries outside Europe also report to the ICES CSR database on a limited basis.

It was also highlighted that there is a need to be aware of potential duplicate cruise summary reports when the CSR collections from SeaDataNet, R2R and Australia are brought together. This is due to those research vessels which have sailed through international waters and were on a global scale charter/barter arrangement. A potential solution is for each region to only focus on their own vessels.

The use of UUIDs as a good way to manage unique metadata records in a distributed environment was also discussed. However, this will not solve the issue of governance for the preferred origin of information or consolidation of duplicates.

- **Use of GeoNetwork for harvesting cruise summary reports (CSR)**

Thomas Loubrieu (TL) presented how SeaDataNet is going to make use of GeoNetwork for harvesting the cruise summary report XMLs from the national cruise summary report providers for delivery to the central CSR centre. Cruise summary reports are generated by the providers as ISO 19139 XML files using MIKADO and then prepared for the GeoNetwork exchange. IFREMER has developed a function in GeoNetwork to load the CSR XMLs into the local internal GeoNetwork database (including vocabulary relationships). The harvesting is managed by interaction between the internal GeoNetwork databases at local and central level. The ISO 19139 XML file can then be retrieved from the central internal database. This approach is now being tested for both the CSR and also for the CDI as part of the SeaDataNet project. GeoNetwork provides a CSW service and a discovery user interface on top of its metadata. This has been deployed by IFREMER to demonstrate it works for both the CDI and CSR.

- **MANIDA: a German marine data portal project**

Ana Macario (AM) submitted a presentation about the national MANIDA project in Germany which looks at improving the flow of cruise data and metadata from 13 German research vessels. Validated cruise metadata (including track lines) are already available for two research vessels. Weather station and thermosalinograph sensor data are also being made available in 10 minute intervals. There is currently a plan to implement SensorML and SOS services. The archived data are managed in the PANGAEA (OAI-PMH), BSH DOD, and AWI bathymetric databases, while any related publications will be made available via OAI-PMH in EPiC, OceanRep, and HZG publications. The near-real time data will be made available in the German coastal research monitoring system COSYNA via SOS and WFS. (It was also noted that the RV Polarstern and RV Heincke research vessels could potentially participate in the EUROFLEETS EVIOR prototype for a dynamic vessels tracking and events system.)

- **Cruise metadata in Australia (Marine National Facility)**

Sebastien Mancini (SM) described how metadata for Australian research vessel cruises is stored in the CSIRO Marine and Atmospheric Research metadata system, MarLIN. It records details of surveys including cruise plans, tracks and voyage data. The database also holds metadata on datasets from Marine National Facility (MNF) as well as other sources. A new Australian research vessel will be launched in the near future and there is an opportunity for development of the cruise summary reporting system. Other research vessels are currently operating separate systems. CSIRO has been identified as best contact for ODIP 2 as they are responsible for the MarLIN database.

- **R2R - Cruise reports in USA**

Bob Arko (BA) presented progress on the Rolling Deck to Repository (R2R) project. The objective of R2R is to provide uniform stewardship of routinely-collected environmental sensor data from the US academic research fleet.

As of November 2013 the R2R database has records for: 3266 cruise IDs, 31 research vessels, 24 device classes and 168 device make and models. It also includes the IDs for in excess of 3000 people from more than 400 organisations. There is also information from over 200 portals and 25 programmes.

R2R publishes cruise records in multiple formats including: ISO 19115-2 metadata records; OGC Web Services; W3C “Linked Open Data” with SPARQL endpoints; DOI-DataCite metadata records which are all using controlled vocabularies and best practices where available. For ISO 19115-2 it uses the vanilla schema, as provided by the NOAA/R2R Template, and it is adopting the SeaDataNet CSR 3.0 Schema. The NOAA/R2R records are published in a GeoNetwork (GN) 2.8 based web portal.

Future activities planned for R2R are: 1) to modify GeoNetwork to ingest 19115-2; 2) resolve authentication problems in GeoNetwork 2.10 and 3) to ingest SeaDataNet CSR 3.0 records. R2R already made progress with mapping around 100 of its cruises onto the SeaDataNet CSR schema and as a result also mapped to a number of the SeaDataNet vocabularies. However R2R has retained its own codes for person, organization, program, device make and model and port by using *gmx:Anchor* elements. *There are also further plans for mapping of the R2R codes to the SeaDataNet device make and model, port, and data category codes, and to map cruise tracks to the SeaVox water bodies, GEBCO undersea feature names, and Marsden squares. This will be followed by work to resolve the remaining codes for person, organization and program.*

- **Discussion**

The discussion was moderated by Bob Arko as leader of ODIP 2. The expected outcome of ODIP 2 is that the global POGO portal will use the SeaDataNet CSR 3.0 Schema. This will require the regional systems in the USA and Australia to either adopt or achieve interoperability with this schema and its vocabularies. GeoNetwork is seen as a good option for exchanging the CSR XML entries between the regions and for their delivery to the POGO portal.

Roy Lowry suggested that each region should focus on building its own portal which can then develop interoperability with the other regional systems and also be harvested by BODC, MARIS and BSH for delivery to the POGO portal. There would then be no need for the Australian regional system to adopt the SeaDataNet CSR and vocabularies but can

instead maintain its own systems. Their activities for the ODIP project could then be focussed on developing interoperability with the SeaDataNet CSR model and vocabularies.

BA introduced a number of issues to consider:

- global convention for file IDs
- multiple CSRs for the same cruise
- Use of <gmx:Anchor> instead of <sdn:codeList>
- where to host non-SDN schemas
- can we accommodate non-CSR schemas
- how to get GeoNetwork to support 19115-2
- possible new elements: field role (e.g. Chief Scientist) for a person; funding awards; vessel (Platform) operator and scheduler; EventLogs
- create *ActivityCode* for <gmi:significantEvent>
- how to solve the most difficult vocabularies: person; organization and program

RL pointed out that the European Directory of Marine Organisations (EDMO) is used by SeaDataNet. EDMO contains mostly European organisations, but already has quite a number of international entries because these are needed for marking up CSR and EDMERP (project) entries where there is international involvement. The EDMO user interface is filtered for European countries but the EDMO web service lists all entries. DS will inform ODIP partners of details for the EDMO web service and will set up an account for R2R to manage USA entries in the online EDMO content management system.

The European Directory of Marine Environmental Research Projects (EDMERP) which allows entry and discovery of European research programme / project details could potentially be used more widely for programmes. EDMERP, which is already used for the SeaDataNet CSR profile, has a user interface and an ISO web service available. Entries can be made online via the content management system or by using the MIKADO software tool.

For the purposes of CSR governance there is a need to put some basic rules in place whereby it is possible to distinguish between European, US and Australian research vessels.

It was agreed that BA, as ODIP 2 task leader, will draft and finalise the ODIP 2 work plan on the basis of the discussions during the workshop. It will include activities, time lines, actors, expected results, and planned interactions with other regional and global initiatives. This will also be used as input for the ODIP Deliverable D3.2. *Results and conclusions from prototype analyses*. Further development of the work plan will be done between partners and key stakeholders by e-mail.

4.3.4 Session 3: ODIP Prototype project 3

Sebastien Mancini (SM) introduced prototype development task ODIP 3 which will aim to establish a prototype for a Sensor Observation Service (SOS) and formulate common O&M and SensorML profiles, including the use of common vocabularies where possible, for selected sensors (SWE) installed on vessels and in real-time monitoring systems. There is an increasing need for and uptake of marine data from a variety of platforms and instruments. The sensor web enablement (SWE) architecture comprises two major components: 1) the underlying conceptual models and encodings and 2) the model used for the specification of the services. SWE comprises SensorML profiles, O&M profiles and SOS services. SOS supports three key operations: core, transactional and enhanced profiles. The

core profile comprises: *GetCapabilities* which provides a description of the service interface and available data; *DescribeSensor* provides sensor metadata in SensorML; and *GetObservation* returns measured values and associated metadata in O&M format. Note: an observation is defined as “*action whose result is an estimate of the value of some property of the feature-of-interest, obtained using a specified procedure*”.

The key objectives of the ODIP 3 prototype development task are:

- Tabulate and explore the existing work done on fixed and mobile platforms to find areas of homogeneity
- Develop minimum level templates for SOS elements (SensorML and O&M templates)
- Look at ‘bronze’, ‘silver’ and ‘gold’ level templates to raise the bar on minimum requirements
- Support and develop these templates via vocabularies, metadata
- Agree on a ‘reference implementation’ (anticipated released of 52 North SOS 3.6)
- Creation of demonstration clients

- **What’s happening in Australia**

SM presented the situation in Australia where a virtual machine (VM) has been set up with latest version of 52 North by the Australian Institute of Marine Science (AIMS). This makes setting-up new instances of the 52 North SOS on the same VM or new virtual machines much easier. The current SOS is providing real-time data from the Davies reef Sensor network (every 10 minutes) via the Integrated Marine Observing System (IMOS). The SOS is open so anyone can post queries to this service (code available on request). It currently uses minimal SensorML and O&M templates as a starting point for future development. There is also minimal security at the service level to allow potential users to access these systems. It is open for anyone to test/try and can be quickly re-built, cloned or used as a software test bed.

SM indicated that there other groups working on sensor web enablement in Australia. This includes the Bureau of Meteorology (Andrew Woolf and Geoff Williams) who are working on an implementation of a 52 North SOS as a component of the National Environmental Information Infrastructure. The objective is to deliver historical data from large datasets using SOS. However a problem has been encountered because the 52 North SOS version 3.5 could not handle the required volumes of data. As a result a set of improvements have been developed that will be included in next release of 52 North SOS version 3.6. It was also noted that CSIRO is also involved with the development of sensor web enablement and is working on SOS with a number of group.

- **What’s happening in the USA**

Derrick Snowden (DS2) presented the SOS activities in US IOOS. DS2 gave some background about the US IOOS Program which has 6 or 7 years of experience with SWE in the context of 11 portals used for exposing ocean information at regional scales. IOOS aims to use SOS to provide a homogeneous service level on top of the heterogeneous regional systems. The SWE family of standards can be applied to time series, profiles and trajectory data obtained from buoys, tide gauges, ADCPs, gliders, and other common oceanographic

data collection platforms. Several templates have been developed (V1.0) based on the 1.0 version of the relevant SWE standards and work is ongoing for migration to version 2.0.

These templates, which US IOOS has published in the public domain, have been implemented in two types of servers: 1) 52°North (SOS) 4.0 and 2) the ncSOS plugin for the THREDDS data server. A number of additions have been made to the 52°North server and use is made of an SOS-injector between the data sources and the 52°North server. The SOS Observation Injector is a Java toolkit that can be used to enter sensor observations into an IOOS customized 52°North SOS. This server supports *GetCapabilities*, *DescribeSensor* (*sensor and station*), *GetObservation* (*O&M, NetCDF, CSV, KML, JSON*).

For the ncSOS server, a THREDDS plugin is used to convert between NetCDF CF version 1.6 and O&M/SensorML. Functionality includes a number of clients and graphics utilities. There are also a number of SOS testing tools such as a NetCDF compliance tester.

DS2 sees a number of potential collaboration activities between US IOOS and ODIP 3:

- Open Source project via Github
- SOS 1.0 migration to SOS 2.0 (and SWE/SensorML/O&M)
- Combine various CF netCDF manuals
- Templates for shipboard data (trajectory/trajectoryProfile)
- Load testing and sharing of experience with high volume installations
- Brokering and extending ingest routines (SOS-injector)
- Cloud deployments and canned AMI's

It is agreed that there are a large number of potential opportunities for US IOOS to contribute to ODIP which will be explored further after the workshop.

- **What's happening in Europe**

Jordi Sorribas (JS) presented the SWE activities in Europe which are on-going as part of the SeaDataNet 2 and EuroFleets 2 projects. The activities are combined where possible and aimed at streamlining the flow of data from research vessels and fixed monitoring stations. The focus in the EuroFleets 2 project is on development of software and tools for underway Ship Summary Reporting, onboard data acquisition and remote e-access, while SeaDataNet is focusing on defining SensorML and O&M profiles for a range of instruments, sensors, and observation types. The objectives are to increase the level of information provided from the instrumentation, establish a common method for accessing real-time data and instrument descriptions, and enhance existing SeaDataNet metadata descriptors for both the CDI and EDIOS.

The first steps towards sensor web enablement (SWE) were made for seismic data sets as part of the FP7 funded Geo-Seas project. Specific SensorML and O&M profiles were defined as extensions to the existing CDI format to describe the seismic instrumentation. These extended CDI metadata profiles are supported by the SeaDataNet vocabularies and can be generated with the MIKADO XML editor. In addition the O&M object is also used to access the seismic visualisation tools which have been integrated into the operational SeaDataNet and Geo-Seas CDI services and provide access to several thousand CDI entries for seismic data sets. This development activity has continued in the SeaDataNet 2 and EuroFleets 2 projects.

A roadmap for SWE has been established which includes three distinct phases: 1) specification of requirements; 2) development of the descriptions for services and 3) implementation of services. Current activities have now reached the implementation phase with an on-board system for research vessels (EARS) now being refined that will support the registration of events and data, as well as the automatic generation of daily ship summary reports for the EVIOR portal. The complete vessel system and its instrumentation and calibration procedures have been described using SensorML and O&M. A similar approach has been taken for fixed platforms with the SensorML and O&M profiles now available. To date this approach only covers time series, profiles and trajectory data and further work is needed for the inclusion of gridded data and instruments such as radar and swath. Observation & Measurements profiles have been defined for specific observation data types and links have been integrated with the SeaDataNet CDI and EDIOS metadata services. A number of issues have been identified while coding such as vocabularies that are not yet complete, use of URNs or URLs and the need for a common approach to identifiers.

Future activities in the SeaDataNet and EuroFleets projects will focus on a number of important aspects including: 1) the review and refinement of the SensorML and O&M profiles, 2) how to adopt MIKADO for generating the new profiles, 3) setting up services to test SOS implementations and real-time data access.

The next tasks for the SeaDataNet and Eurofleets partners will be:

- Compile and inventory of instrument SensorML records & O&M structures
- Compile an inventory of vocabularies and registry services
- Formalisation of test cases for vessels and fixed platforms
- Install SOS services accessing real-time underway vessel data
- Establish a working group to propose templates for SensorML and O&M profiles and data models
- Development of a (smart) client (or clients) and tools to enhance functionality

The goal is to integrate a SOS prototype with existing data services e.g. data centres, vessels, operator's sites etc. Comparable work on SWE is also taking place in the new EU FP7 NEXOS project which will have synergy with the SeaDataNet - EuroFleets activities and also the ODIP 3 project.

- **Sensor Observing Systems (SOS) and 52 North**

Simon Jirka (SJ) presented the developments by 52 North for SOS servers. The current version is SOS 3.x with release of version 3.6 planned for mid-December 2013 which will have a full SOS 1.0 implementation with elements from SOS 2.0. This will also include the contributions of Geoff Williams to give a performance boost, support streaming XML encoding, and provide a more efficient capabilities cache.

Development of SOS version 4.x is also underway which will have a SOS 2.0 reference implementation (KVP), full SOS 2.0 and core elements of SOS 1.0. It will also include a lightweight SOS profile and SOS 2.0 Hydrology profile including WaterML 2.0. Delivery of SOS version 4.0 which includes extensions for US IOOS SOS is expected in January 2014.

52°North is also developing a 52°North Sensor Web Client and a Mobile Sensor Web Client. Other development work includes implementation of SensorML 2.0, and client development for a REST-API and Javascript API.

- **O&M and SensorML update**

Simon Cox (SC) presented the minimum requirements for SOS and also introduced a star system to describe the functionalities:

- * GetCapabilities
- * GetObservation (feature, property, procedure, time, *space*)
- ** GetObservation - by ID
- ** GetResult
- ** GetFeatureOfInterest
- ** DescribeSensor
- *** InsertSensor/DeleteSensor
- *** InsertObservation/InsertResult
- **** GetDataAvailability

The OGC SOS for observation data access provides the `getObservation`, `getResult`, `describeSensor` and `getFeatureOfInterest` operations. SC outlined the O&M vocabulary which comprises: Observation, Procedure, Observed property, Result, Feature of interest, Phenomenon time, Result time, and Valid time. Using this neutral terminology supports cross-domain data discovery and fusion.

SC also reported that SensorML version 2.0 is making good progress and several examples of templates are available. However SC is becoming less convinced that everything has to be solved with XML and that there are other potential lightweight alternative such as *FL.

- **Discussion led by Sebastien Mancini (Australia)**

Sebastien Mancini (SM), ODIP 3 development task leader (together with Roger Proctor and Scott Bainbridge), led the discussion on a possible set up of the ODIP 3 prototype development task and the drafting of the associated work plan.

SM advised that as a minimum a template using SensorML 2.0 and O&M 2.0 should be implemented by ODIP 3 with the assistance of Simon Cox. The O&M version 2.0 is available and this will make it easier to describe collections next to observations. However, there is also concern about which vocabularies to use as US IOOS currently uses CF standard names (accessed from the Marine Metadata Interoperability project's Ontology Register and Repository). Further discussion on this topic will take place in session 4 of the workshop (Vocabularies). SensorML 2.0 will be used for time series and trajectories.

The plan for ODIP 3 is to set up three test environments in Europe, the USA and Australia for a feedback loop from theory through to practice. The 52°North SOS server is the most fully featured server implementing both the SOS Core Profile and the SOS Transactional Profile, while ncSOS which is only implementing only the Core Profile, can be considered and integrated into the ODIP 3 project as a good solution in certain situations. There is also the possibility that the European partners will introduce an alternative. The concept is that the client interface established at AIMS in Australia will then be implemented on top of the three regional SOS services as a common interface.

It was agreed that SM together with RP and SB as the task leaders will draft and finalise the ODIP 3 work plan following the discussions during the workshop. It will include activities,

time lines, actors, expected results, and planned interactions with regional and global initiatives.

4.4 Additional discussion topics

Using the original list of potential workshop topics compiled by partners during the initial ODIP kick-off meeting and following discussions at the project side meeting held at the IMDIS conference in September 2013 a short list of additional discussion topics for the 2nd workshop was identified. All ODIP partners were then asked to indicate their priorities using an on-line survey and as result an additional two discussion topics were included in the meeting programme.

4.4.1 Session 4: Vocabularies

The vocabularies session was led by Roy Lowry (EU), John Graybeal (USA) and Simon Cox (Australia).

- **What's happening in Europe**

Roy Lowry (RL) presented an update on the vocabulary developments in Europe and specifically the NERC Vocabulary Server (NVS 2.0) and its further developments for the SeaDataNet and EMODNet Chemistry projects. The NVS 2.0 is stable and since early 2013 it has received 860.000 calls from 1268 IP addresses with a balance between the use of the SOAP and REST protocols. The SeaDataNet initiative has also recently undertaken a successful migration from the NVS 1.0 to NVS 2.0. All SeaDataNet repositories, formats, tools (MIKADO, NEMO, ODV) and services now use NVS 2.0 which has a better version management and truly unique URNs for concepts. The parameter description vocabulary used by SeaDataNet (P01) includes approximately 28.000 terms and more have also been added for handling biological terms in SeaDataNet. A new vocabulary for aggregated parameters defined as mappings to P01 terms (P35) was initiated and populated with one test term. It will be populated further in the coming months with the priority being given for parameters in use in EMODNet Chemistry. The P35 vocabulary supports a fully automated aggregation and validation process for data sets marked up using P01 vocabulary gathered from multiple sources. Progress has also been made in exposing the semantic model which underlies the P01 terms. The semantic model (O&M concept) was agreed and both JSON and RDF representations have been developed. Further work on developing the user interfaces to make it easier for data centres to query and identify relevant P01 terms is planned in the coming months. Users will also be able to propose P01 extensions using a form which can be processed more easily as part of governance. The P01 vocabulary has already been extended by mapping 1779 ICES contaminants for the biota chemical/matrix combinations used in UK data. This has been done as part of the EMODNet Chemistry project which is attempting to populate the SeaDataNet portal with a complete set of marine chemistry data.

Development of the P01 semantic model was undertaken in consultation with CSIRO (Australia) who also provided feedback on the RDF representation. In addition CSIRO demonstrated the use of a SISSVoc facade over the NVS content which underpins the interoperability of the NVS 2.0. Interactions with partners from the USA provided information on the US NODC SKOS, and the mappings between the NVS 2.0 and MMI-ORR for SeaVox platforms and device categories have been created and loaded on both services.

- **What's happening in the Australia**

Simon Cox (SC) explained that there are many other vocabularies being served outside the ODIP community which can have a complimentary use for the project. However there is a need to be careful about different definitions and meanings. Examples of available vocabularies include OFKN which has a focus on Linked Open Vocabularies, QUDT which offers a conceptual model for units, quantities, and dimensions via URIs and ChEBI which provides URIs for chemical substances. SC used these examples to show how a comparable semantic model to that previously presented by RL for P01 can be built. This has been converted to SKOS and is now included in the SSSVoc Search service that CSIRO operates. It has URIs to authoritative vocabs such as ChEBI and DUTQ. Underneath SSSVoc are an RDF demonstrator and a SPARQL endpoint. SC has also included around 60 of the NVS 2.0 vocabularies in SSSVoc which are accessed on-the-fly and not by local buffering.

SC concluded that vocabularies should be standardized, harmonized and published, and that where ever possible the ODIP community should extend / re-use existing vocabularies. SC also suggested that the P01 semantic model should be exposed by RDF via a SPARQL endpoint. RL, SC and Adam Leadbetter (AL) agree on this and will work together on the relevant services required for the ODIP prototypes.

- **What's happening in USA**

John Graybeal (JG) presented progress on the Marine Metadata Interoperability (MMI) project. JG is now working at Marinexplore which is aggregating large numbers of public data sets using a metadata rich strategy. JG also highlighted the large number of vocabulary developments going on in the USA.

The Rolling Deck 2 Repository (R2R) project is focusing on the vocabulary mappings required to publish R2R Cruise Summary Reports in the SeaDataNet 3.0 ISO schema. The BCO-DMO initiative is looking at the concept of event logging for research vessels together with European colleagues (EuroFleets). It is also undertaking several mapping exercises and developing faceted search on top of SKOS. As part of the Shipboard Automated Meteorological and Oceanographic Systems (SAMOS) activities Florida State University is mapping the SAMOS QC flags on to those used by SeaDataNet. Scripps Institution of Oceanography (SIO) is making progress with mapping the R2R organisation and port vocabularies to those used in SeaDataNet (EDMO and C38 respectively). However, there are still a lot of missing mappings which should be populated in the coming months as part of ODIP Prototype 2.

US NODC is exploring using SKOS in their vocabularies and will also begin consolidating/merging the controlled vocabularies used across the three NOAA National Data Centers (NODC, NGDC, and NCDC). MMI-ORR has successfully been migrated to TAMU-CC for operations and several new developments have taken place, such as VINE which is a vocabulary integration tool used for making mappings to external ontologies.

JG also highlighted the shift from closed to open practices, such as using linked open data principles. The idea of creating a human search interface for terms is also desirable. It is also advisable to create APIs to search for and harvest metadata. JG agrees with SC that there is scope for a common wiki for semantic projects.

The issue of unique identifiers for people also needs to be addressed. Although this is not currently within the scope of SeaDataNet it is being addressed by initiatives such as ORCID and the IODE's Ocean Expert database.

4.4.2 Session 4: Vocabularies (continued)

At the end of Day 3 of the meeting, following Session 5, there was an opportunity to return the discussion regarding the use of vocabularies in the ODIP community.

- **Discussions led by Roy Lowry (EU), Simon Cox (Australia) and John Graybeal (USA)**

Vocabularies are instrumental in all of the ODIP prototype development tasks. These vocabularies should be used and linked by means of concept URIs which return a document for each concept. There is a need to check which vocabularies, concepts and mappings are already available in each of the ODIP prototype tasks and what additional ones might also be required. This information should be submitted to the ODIP vocabularies group which consists of Roy Lowry, Simon Cox and John Graybeal. It is expected that ODIP 3 will have extra requirements for vocabularies to support the SensorML and O&M profiles, such as instruments, their types and manufacturers, calibration, etc.

It was also reported that, as part of the SeaDataNet project, a subcontract is being negotiated with JCOMM/OPS together with ICES and BODC for a global Register of Platforms (buoys, gauges, etc.) in order to bridge between the operational oceanography and data archiving communities. This development is highly relevant for ODIP.

- **Checking the status of actions from the 1st ODIP Workshop**

This session, which was led by Roy Lowry (RL), aimed to assess the amount of progress that has been made since the 1st ODIP workshop and identify specific tasks for the first half of 2014 including any additional work on the vocabularies required by the ODIP pilot projects.

The following actions were identified during the 1st ODIP workshop and documented in deliverable *D2.2: Minutes of the 1st ODIP Workshop*. Progress on these individual actions was reported during this session.

Action 1.1 Implementation of SPARQL technology and mappings between vocabularies (SKOS)

A SPARQL endpoint has been established in Australia (SISSVoc) and further work is ongoing for federating from different vocabulary services such as NVS 2.0, ChEBI, etc. SISSVoc can also potentially serve as a pilot portal. It was agreed that SISSVoc should be made more visible on the ODIP website. This will be actioned by Simon Cox and Dick Schaap. There is also a need to prioritize the vocabularies and mappings which are relevant for the ODIP prototype development tasks.

Action1-2: Establishing thesaurus-based semantic aggregation of data marked-up using the BODC/SeaDataNet parameter usage vocabulary (P01)

This action focussed on the development of semantically aggregated controlled vocabularies and associated governance model. This activity is being progressed in Europe through the development of the P35 vocabulary (see above) and by SISSVoc in Australia. A further action could be to identify further applications for this kind of vocabulary in the ODIP prototype development tasks.

Action1-3: Formally document vocabulary governance within the NERC Vocabulary Server.

There is an on-going discussion in the NERC Information Strategy Group which is establishing a mission-critical register of services. NVS 2.0 is recognised as a candidate service and additional resources for development are also being put in place. Further work is also needed to address the NERC Vocabulary Server (NVS) governance with reference to the ISO 19135 standard.

Action1-4: Harmonisation of the conceptual models and controlled vocabularies used for event logging on research vessels with particular reference to Eurofleets and R2R

Karien de Cauwer (KdC) and Cyndy Chandler (CC) provided a report on this action which forms part of the cooperation between the R2R and EuroFleets projects under ODIP 2. A major challenge is that the existing conceptual models have significant differences and work is currently ongoing to overcome these issues. The two projects are also compiling a list of the terms used by the individual communities to create a single combined list.

Action1-5: Develop a unified approach to the utilisation of controlled vocabularies under the NERC Vocabulary Server governance using GeoNetwork.

Action1-6: Develop a unified approach to the utilisation of controlled vocabularies under the NERC Vocabulary Server governance in other metadata standards such as O&M and SensorML

Actions 1.5 and 1.6 concern the development of a unified approach to utilising the NVS in GeoNetwork and also O&M and SensorML. This will be part of ODIP prototype development tasks 1 and 2.

- **Possible use of gmx:Anchor**

Thomas Loubrieu (TL) spoke about how to handle external references in metadata (e.g. ISO 19139). In the current SeaDataNet 3.0 ISO Schema for cruise summary reporting, use is made of different codings: some external references are stored with codeList while others are stored as text in a Character String tag, or directly as tags or attributes values. TL outlined how references to these external directories or thesaurus could be done in a more homogeneous way by always using the **gmx:Anchor** tag which allows storage of a remote URL (as the external reference) and a label.

The implications for introducing this solution into cruise summary reporting at this point were discussed and it was agreed that it would have significant implications for the service chain and tools. It was therefore agreed not to adopt it now, but the potential impacts of its implementation should be described in deliverable *D4.1. Interim strategic analysis report*

For the purposes of metadata validation (against vocabulary references) the reference can be checked by applying a schematron rule on a SKOS list or alternatively the URL in xlink:href status can be verified. If it resolves on the web this means that the reference exists, if not this means that the reference does not exist. This second option can also be easily incorporated in the schematron. TL finishes by explaining a feasible procedure for editing references in the metadata through GeoNetwork (GNW).

4.4.3 Session 5: Data publication and citation

This session was led by Cyndy Chandler (CC) together with Justin Buck (JB) and Lesley Wyborn (LW).

- **ANDS Research Data Citation**

Lesley Wyborn (LW) outlined the procedures for research data citation at the Australian National Data Service. ANDS started investing in the establishment of research data collections in cooperation with research institutes and universities during 2009. The aim is that Australian researchers can easily publish, discover, access and use/re-use research data.

Data citation is important because it eases reuse and validation of data, it makes it possible to track the impact and reach of data, it recognizes and rewards data producers, it increases academic and institution profile, and it connects all research outputs. Data citations were previously quite short and imprecise, but now tend to be more structured and precise, and also often include a **DOI** (Digital Object Identifier) to facilitate retrieving the data set. For example CSIRO publishes such an attribution statement (Data Citation) as part of its data access portal.

LW encourages that the motto should be "Access + Share = Cite". This is also adopted in the USA where we see increasing access to the results of federally funded scientific research. The National Science Foundation (NSF) now allows for citable data (i.e. with a DOI) to be listed as an outcome of research in a similar way to journal articles. The data set itself is considered as a product which needs to be citable and accessible.

There is a need to build awareness among researchers and create a culture of data citation. ANDS is running a community building campaign which includes videos and flyers. In addition, ANDS is providing a Data Citation Toolkit which includes general information and functions for minting DOIs for use in the ANDS data citing service. It can also be used for implementing data citation across a wide range of institutions and researchers.

There are a range of DOIs for scientific articles and increasingly DOIs are also being adopted for data sets e.g. Datacite.org already has more than 2 million DOIs for data sets. DOIs can also be used for linking a researcher to individual datasets. By registering personal details in a catalogue such as ORCID and linking to DOIs for software, publications and data sets an individual can build their research profile. Publishers are also increasingly adopting this approach and encouraging authors to include DOIs for their data sets in papers submitted for publication.

The Scientific Data Initiative which will raise awareness of data citation and urge researchers to publish data sets will be launched in the USA during spring 2014. This initiative will also stimulate the establishment of approved repositories. Work is now underway to develop reliable data citation trackers across the different media to count the number of individual data citations.

- **What's happening in Europe**

Justin Buck (JB) highlighted the push from publishers and scientists for data citation: Publishers want to link journal articles to the data, while scientists want credit for data set creation and usage. The adoption of DOIs is a good approach and in particular the use of DataCite DOIs. The British Oceanographic Data Centre (BODC) can issue DOIs for datasets in collaboration with DataCite as part of a wider NERC and global approach to dataset publication.

For the purposes of data citation, datasets MUST be static, fulfil strict (meta) data requirements and the data sets must become freely available when a DOI is issued. These DOIs can be found at the BODC Published Data Library (PDL) webpages and are also available in the SeaDataNet EDMED directory of data sets. However, the scope of EDMED is wider and also includes restricted data resources and those without DOIs.

The Ocean Data Publication Cookbook has been produced jointly by UK, USA and IODE It gives the criteria and best practice for publication and citation of data, and is freely available from the IODE portal.

There are a number of unresolved issues relating to data citation including how to attach DOIs to open time series and also whether persistent identifiers are the same regardless of the status of the data, versioning, and granularity. This can be solved by distinguishing between two separate timelines: event / measurement / OBSERVATION time; ingestion / update / STATE time. DataCite provides a dynamic data policy to deal with these kinds of data sets; however there are still some caveats.

BODC and IFREMER are currently analysing data citation for the Argo programme of ocean floaters. The Argo programme produces more than 200 publications annually. The problem is how to cite Argo data at a given point in time. To solve this issue for the real-time data stream, IFREMER has minted a Digital Object Identifier (DOI) for the Global Data Assembly Centre (Argo GDAC) as a whole. These are sufficient for Argo if long term reproducibility of the data is not required by the user. However, IFREMER is also minting individual DOIs for monthly granularity snapshots at the GDAC level to enable reproducibility.

- **Data Citation in USA**

Cyndy Chandler (CC) explained that data publication involves domain scientists, data managers and library scientists. It provides the opportunity to strengthen the bonds between professionals working in those communities. The challenge is to develop a system that supports proper citation of intellectual work that also encourages increased sharing of research data. The SCOR/IODE/MBLWHOI Library Project (USA) is assigning persistent identifiers to data referred to in traditional journal articles which are stored in institutional libraries, and where the data held by data centers are packaged and served in formats that can be cited. The project has developed a number of use cases to identify best practices for tracking data provenance and clearly attributing credit to data creators/providers so that researchers will make their data accessible. CC emphasized the incentive that this gives to researchers to publish their data sets. Data citation metrics are also starting to be collected and will potentially be adopted by funding agencies as part of performance measurement. Libraries have used DOIs for a number of years and are now becoming the de facto standard for data sets.

There are also a number of relevant related activities including: continued support and interest from IODE/SCOR; the Research Data Alliance (RDA) Marine Data Harmonization Interest Group (led by Helen Graves), the RDA Data Citation Working Group and the CODATA Data Citation Standards and Practices Task Group.

A next step is to address the issue of identifiers for people. The Open Researcher and Contributor ID (ORCID) initiative is building a registry of unique researcher identifiers which will provide persistent identifiers for named people and facilitate links to other resources/outputs created by the researcher.

- **ESRI developments**

Dawn Wright (DW) provided details of a small exemplar created by ESRI which demonstrates the use of DOIs for geographical maps and base material that have been included in a peer reviewed monograph for a conference. This can be found at <http://www.esri.com/oceansolutions>.

- **Discussion led by Cyndy Chandler (USA), Justin Buck (EU) and Lesley Wyborn (Australia)**

The idea of IFREMER minting DOIs for each data product that will be produced and published as part of the SeaDataNet and EMODNet Chemistry projects was put forward. These data products will be described using the ISO 19139 metadata schema and SeaDataNet vocabularies, and included in the SEXTANT catalogue (based on GeoNetwork). The process for adding DOIs is easy and also ensures that older versions of data products continue to be available. The minting of DOIs for larger data sets described in the SeaDataNet EDMED Directory in the same way as is already done for Argo data will be considered. It was also highlighted that any DOI that is allocated must point to a suitable landing page.

Sissy Iona (SI) emphasised the importance of having a robust method for registering and assigning persistent identifiers for people. There are a number of systems already in existence e.g. ORCID, IODE OceanExpert but there is a need to identify other potential mechanisms. It was agreed that there is currently a lack of sufficient information about these systems within the ODIP community. To address this issue an ODIP Prototype project 4 will be formulated which will prepare a position paper on registries for people with the objective of linking researchers to data publications, cruise summary reports, etc. This task will be led by Helen Glaves (HG).

It was proposed to facilitate the sharing of information on data publishing and citation across the project community by maintaining a wiki at the ODIP website (or similar), and also by participating in relevant activities such as those led by EUDAT, RDA, Belmont Forum, etc. This will potentially be explored by Cindy Chandler (CC).

4.5 Session 6: Feedback on sessions

Each of the ODIP prototype development task leaders gave a summary of additional ideas for their development tasks which have been formulated as a result of the discussions during the workshop. This included an outline of a draft work plan for each of the three ODIP prototype development tasks.

It was agreed that over the next few weeks each of the prototype development task leaders will expand their draft concepts to produce a fully developed work plan for each of the ODIP prototype development tasks. Each one will provide details of planned activities, expected results and the project partners or stakeholders who will contribute to the task. These work plans will be used as the basis for the development of the ODIP prototype interoperability solutions and also provide input for deliverable *D3.2 Results and conclusions from prototype analysis* which is due for delivery in January 2014

Ben Domenico (BD) recommended keeping an eye on relevant future developments, and to not forget cloud computing and the concept of big data. Both of these concepts are not currently addressed by the existing ODIP prototypes but could be included as new discussion topics for the 3rd ODIP workshop.

It was emphasised that each of the ODIP prototype development tasks should share information within their groups and also with other members of the ODIP community. It was agreed that it would be more efficient to have online tools available for both a document repository and software developments. Ideas put forward were to use COLLA (maintained by OGS) which was outlined in the ODIP description of work or alternatively to use GitHub or something similar. Paolo Diviacco stated that he is a strong believer in collaborative tools such as COLLA to support working groups. However use of this type of tool does require some setting up in line with the requirements for each of the ODIP prototype development tasks and also necessitates moderation by the task leader. It is agreed to evaluate the use of this type of tool.

Simon Cox (SC) and Roy Lowry (RL) summarized the discussions on vocabularies and agree to explore the possibilities for an additional ODIP prototype development task exploring the options for a vocabularies wiki and a SPARQL endpoint for federated vocabularies.

4.6 Session 7: Workshop wrap-up

4.6.1 Further potential prototype projects

Helen Graves (HG) summarized the proposed additional ODIP prototype development tasks which are: 1) drafting of a position paper on unique identifiers for people and 2) development of a wiki for vocabularies which can also potentially be used as a forum to support the ODIP prototype development tasks.

As part of ODIP work package 4 (WP4) IFREMER is required to prepare an impact assessment for the ODIP partners and also the wider community, of adopting the prototype interoperability solutions. This document will include an evaluation of migrating to gmx:Anchor and the implications of introducing DOIs for EMODNet data products as discussed during the workshop.

Sytze van Heteren (SvH) proposed looking towards supporting a multilingual approach as part of the ODIP prototypes. Sebastien Mancini suggested that quality control should be considered for inclusion as a discussion topic for the 3rd ODIP workshop.

4.6.2 Plans for next 8 months

• Deliverables

Helen Graves (HG), the ODIP project co-ordinator, reminded the partners about the deliverables that are due for completion over the next few months. They are:

- D1.6: Minutes of Steering Committee 3 (Helen Graves)
- D1.7: Six monthly progress report: year 2 (Helen Graves)
- D2.3: Short report on the organisation and programme of the 2nd ODIP Workshop (Sissy Iona)
- D2.4: Minutes and actions from the 2nd ODIP Workshop (Dick Schaap and Sissy Iona)
- D2.5: 3rd ODIP Workshop (Sissy Iona and Scott Bainbridge)

- D3.2: Work plan for development and implementation of the three ODIP Prototype projects (Dick Schaap with contributions from Bob Arko for ODIP 2 and Sebastien Mancini - Roger Proctor for ODIP 3)
- D3.3: Definition of possible extra ODIP Prototype projects (Helen Glaves)
- D4.1: Interim strategic analysis report with impact assessment (Thomas Loubrieu)
- D5.4: Promotional leaflets and posters (Helen Glaves)

- **Work package 4**

WP4 which is led by IFREMER starts immediately after the workshop. It will focus on assessing the potential impact of wider adoption of selected common standards and the integration of the ODIP interoperability solutions into the national and regional systems.

The original work plan for WP4 included exploring options for interoperability solutions and/or common standards, differentiating between possible shorter term 'easy wins' and longer term efforts. However, in practice these activities have been integrated into the ODIP prototype development tasks which form part of WP3. These are now much more evolved and defined as developing full scale services and standards, where originally prototypes were foreseen as components with an analysis of integration in WP4.

- **Project promotion and dissemination**

HG outlined a number of forthcoming dissemination activities for the project including:

- **American Geophysical Union (AGU) Fall Meeting**, San Francisco, USA (December 2013): ODIP will make two poster presentations
- **Research Data Alliance (RDA) plenary meeting**, Dublin, Ireland (March 2014)
- **European Geosciences Union (EGU)**, Vienna, Austria (May 2014)

ODIP will be represented in the Marine Data Harmonisation Interest Group which includes a number of people from the ODIP partner organisations and is co-chaired by Helen Glaves (ODIP coordinator)

ODIP will hold two splinter meetings 1) public information session for the purposes of community building and 2) partners meeting for the purposes of monitoring progress on the ODIP prototype development tasks. In addition ODIP will hold a joint workshop with the COOPEUS and BCube projects. The topic to be addressed and the exact format of the workshop will be decided in consultation with the COOPEUS and BCube project leaders.

Helen Glaves reminded partners that they should notify the leader of WP5 about any promotion and dissemination activities that they undertake.

Ben Domenico (BD) indicated that he reports on ODIP at UNIDATA meetings. He also sees opportunities for a formal ODIP presentation at those meetings once the prototypes have made further progress. Stefano Nativi (SN) highlighted the GEO X meeting that will take place in Geneva, Switzerland during January 2014. The data brokerage service will be presented at that meeting.

- **3rd ODIP workshop**

Sissy Iona (SI) informed the meeting that the 3rd ODIP workshop will take place at the Australian Institute of Marine Sciences (AIMS) in Townsville, Australia. It is planned to take place during the second week of August 2014. Partners will receive further information from SI and Scott Bainbridge, the local host, in due course. Details about the arrangements for the workshop will also be posted on the ODIP website. Drafting of the agenda will begin in the near future and a first draft will be presented for discussion during the partners meeting at EGU.

- **ODIP students effort**

As part of the USA participation in ODIP a number of students are actively engaged in a number of activities. The students were in attendance at the workshop and HG encouraged the ODIP partners to interact with the students as much as possible.

During the meeting the students were given the opportunity to introduce themselves, their research interests and also report on their impressions of the workshop. The ODIP students are:

- Linyun Fu. Institution: Rensselaer Polytechnic Institute. Degree Program: PhD. Mentor: Bob Arko.
Linyun will work on a project to deploy a Linked Data API (LDA) for R2R; the LDA will create a "Web Services"-style interface to the R2R Catalog
- Yanning "Yu" Chen. Institution: Rensselaer Polytechnic Institute Degree Program: PhD. Mentor: Cyndy Chandler.
Yu will develop a vocabulary mapping tool to facilitate term mapping between R2R devices and the SDN devices served via NVS 2.
- Renata Ferreira. Institution: University of California San Diego. Degree Program: BSc. Mentor: Karen Stocks.
Renata will create mappings between R2R controlled vocabularies and SeaDataNet and other community vocabularies.
- Nkem Dockery. Institution: FSU. Degree program: MSc. Mentor: Shawn Smith.
Nkem and Jocelyn will develop linked data approaches to link SAMOS data to R2R cruise information and also a cross-walk between SAMOS and EU vocabularies.
- Jocelyn Mandalou. Institution: FSU. Degree program: BSc. Mentor: Shawn Smith.
Unfortunately Jocelyn was not able to attend the ODIP Workshop. She will work together with Nkem on a linked data approach (see above)

The students reported that the workshop had given them an opportunity to see how their studies (which relate to computer sciences) are connected to the real world (natural sciences) and the management of technical challenges faced by the oceanography community. They expressed their clear interest in interacting with the ODIP community and contributing to the project development tasks through their own related activities



- **Closing remarks**

Helen Glaves (HG) closed the workshop by thanking the local organisers at Scripps Institution of Oceanography, in particular Rita Bauer and Karen Stocks, for their excellent logistical arrangements. HG also thanked all participants for their contributions and making the 2nd workshop such a success.

HG also mentioned that all of the presentations and associated videos from the workshop are available via the ODIP website (www.odip.org) under the “Workshops” menu option. The presentations and videos are hosted by IODE at:

http://www.iode.org/index.php?option=com_oe&task=viewEventAgenda&eventID=1420

There is also a short link on the ODIP website on the 2nd Workshop page at: http://www.odip.org/content/content.asp?menu=0350000_000000

5 Proposed follow-up and actions

As a result of the 2nd workshop a number of additional actions have been formulated. These supplement the planned project deliverables and the on-going actions from the 1st ODIP workshop. These actions are:

Action 2WS-1: ODIP prototype development task leaders will finalise their work plans following the discussion and suggestions made during the 2nd ODIP workshop. Each work plan will include activities, time lines, actors, expected results and interactions with regional and global initiatives. Finalization of these work plans will be done by e-mail communication with the designated contact person from each partner organisation or other identified stakeholder.

The final work plans will be forwarded to the ODIP Technical Coordinator for inclusion in ODIP deliverable *D3.2 Results and conclusions from prototype analysis*.

Special attention will be given in the work plans to a number of the issues discussed during this 2nd ODIP workshop. These are:

- Universal Identifiers (UID) between metadata and associated data sets
- Granularity: looking at the issues around collections and granules through the analysis of selected use cases.
- Analysis of the requirement for unique CSR IDs for the purposes of registration and governance in cruise summary reporting (CSR) systems
- Potential duplicates when the cruise summary reports collections from SeaDataNet, R2R and Australia are brought together. This is a result of research vessels sailing through international waters.
- Review what is reported to the ICES CSR database from outside Europe and how to deal with those records
- Explore how to connect the research vessels RV Polarstern and RV Heincke into the EUROFLEETS EVIOR prototype for a dynamic vessels tracking and events system via the MANIDA services
- Use of EDMO for organisations and possible use EDMERP for projects in the R2R services
- Check which vocabularies, concepts and mappings are already available in each of the ODIP prototype development tasks and what additional ones may be required. Submit these to the ODIP vocabularies group
- Identify further applications in the ODIP prototype development tasks for parameter aggregation vocabularies such as P35.

Action 2WS-2: The ODIP Coordinator will prepare a position paper on unique identifiers for people (ORCID, OceanExpert, etc.). This will focus on providing a link between researchers and data publication/citation, cruise summary reports, etc. This may lead to an ODIP prototype development task 4. This paper will also form part of deliverable *D3.3 Definition of prototypes 2*

Action 2WS-3: Simon Cox will make SISSVoc and associated supporting information more visible on the ODIP website.

Action 2WS-4: IFREMER will mint digital object identifiers (DOIs) for each data product that is produced and published as part of the SeaDataNet and EMODNet Chemistry projects. This will ensure that that older versions of data products remain available. The minting of DOIs for larger data sets described in the SeaDataNet EDMED directory will also be considered (this is already being done by EuroArgo).

Action 2WS-5: Cyndy Chandler (WHOI) will explore options for sharing information on data publishing and citation, e.g. by maintaining a wiki on the ODIP website. This will include relevant activities in EUDAT, RDA, etc. (This activity can potentially be documented in *D3.3. Definition of prototypes 3*)

Action 2WS-6: The ODIP Co-ordinator will prepare a draft agenda for the 3rd ODIP workshop for discussion at the European Geosciences Union conference (EGU) in Vienna, Austria which takes place on 27 April to 02 May 2014. The topics of cloud computing and concepts for big data will be considered for inclusion as topics for discussion, next to progress on the now 4 ODIP Prototype projects and actions.

Action 2WS-7: Paolo Diviacco to explore and implement online tools for sharing information and software developments around the ODIP prototype development tasks. Ideas so far are to use COLLA, GitHub, wiki or something similar.

Action 2WS-8: Partners to inform the Co-ordinator on promotion and dissemination activities that they undertake.

Action 2WS-9: Sissy Iona will progress the preparations for the 3rd ODIP Workshop which will be held in Townsville, Australia during August 2014. Partners will be kept informed via e-mail and the ODIP website.

Appendix A Terminology

Term	Definition
CDI	Common Data Index metadata schema and catalogue developed by the SeaDataNet project
COOPEUS	EU-NSF funded project promoting open access and sharing of data and information produced by environmental research infrastructures
CSR	Cruise Summary Reports is a directory of research cruises.
EARS	EARS, the Eurofleets Automatic Reporting System, is being developed to record a wide range of actions to account for any possible event during a cruise, including any malfunction occurring or any observation done
EDIOS	European Directory of the initial Ocean-observing Systems: EuroGOOS initiative which provides an overview of the ocean measuring and monitoring systems operated by European countries
EVIOR	European Virtual Infrastructure in Ocean Research (EVIOR): portal developed by the EUROFLEETS 2 project to provide up-to-date information about research vessels (RV), cruise programmes, completed cruises and special equipment
GEOMAR	GEOMAR Helmholtz Centre for Ocean Research Kiel: German institute for marine sciences
GeoNetwork	An open source catalogue application for managing spatially referenced resources. It provides a metadata editing tool and search functions as well as providing embedded interactive web map viewer
GCMD	The Global Change Master Directory

IMOS	Integrated Marine Observing System: Australian monitoring system; providing open access to marine research data
OceanRep	An open access digital collection containing the research output of GEOMAR staff and students
ODP	Ocean Data Portal: data discovery and access service, part of the IODE network
IOC	Intergovernmental Oceanographic Commission of UNESCO (IOC/UNESCO).
IODE	International Oceanographic Data and Information Exchange (part of IOC)
ODV	Ocean Data View (ODV) data-analysis and visualisation software tool.
O&M	Observations and Measurements: OGC standard defining XML schemas for observations, and for features involved in sampling when making observations
OGC	Open Geospatial Consortium: an international industry consortium to develop community adopted standards to “geo-enable” the Web
POGO	Partnership for Observation of Global Oceans: forum for leaders of major oceanographic institutions around the world to promote global oceanography:
QUDT	Quantities, Units, Dimensions and Data Types: NASA-sponsored initiative to formalize Quantities, Units of Measure, Dimensions and Types using ontologies expressed in RDF/OWL
RDA	Research Data Alliance: an international initiative promoting open data sharing across technologies, disciplines, and countries to address the grand challenges of society.
SensorML	OGC standard providing models and an XML encoding for describing sensors and process lineage



SDN	SeaDataNet: EU-funded pan-European e-infrastructure for the management and delivery of marine and oceanographic data
SKOS	Simple Knowledge Organization System
SOS	Sensor Observation Service: a web service to query real-time sensor data and sensor data time series. Part of the Sensor Web
SPARQL	a query language for databases, able to retrieve and manipulate data stored in a Resource Description Framework (RDF) format
SWE	Sensor Web Enablement: OGC standards enabling developers to make all types of sensors, transducers and sensor data repositories discoverable, accessible and useable via the web
R2R	Rolling Deck to Repository: a US project responsible for the cataloguing and delivery of data acquired by the US research fleet.
WebEx	On-line web conferencing and collaboration tool