



Ocean Data Interoperability Platform II

Deliverable 2.8 Minutes and actions of ODIP II workshop 4

Work package	WP2	ODIP workshops
Author (s)	Sissy Iona	HCMR
Author (s)	Dick Schaap	MARIS
Author (s)		
Author (s)		
Authorized by	Helen Glaves	NERC
Reviewer		
Doc Id	ODIP II_WP2_D2.8	
Dissemination Level	PUBLIC	
Issue	1.0	
Date	30 March 2018	



Abstract:

This deliverable documents the detailed minutes and actions from the 4th ODIP II workshop that took place at the Marine Institute, Galway, Ireland on 2 to 5 October 2018. It provides an account of the sessions, discussions and outcomes from the last of the four workshops planned for ODIP II.

Document History				
Version	Author(s)	Status	Date	Comments
0.1	Sissy Iona (HCMR)	DRAFT	8 May 2018	First draft
0.2	Dick Schaap (MARIS)	DRAFT	10 May 2018	Additional input to first draft
1.0	Helen Glaves	FINAL	20 May 2018	Final edits and sign-off



Contents

Executive Summary	6
1 Introduction	7
2 List of Participants	8
3 Workshop Agenda	10
4 Workshop proceedings	11
4.1 SESSION 1	11
4.1.1 Introduction	11
4.1.2 ODIP II: Overview of the project	11
4.1.3 ODIP II: Technical Objectives	11
4.2 SESSION 2 - ODIP 1+ Prototype Development Task (plenary)	11
4.2.1 ODIP 1 +: current status and future development activities	11
4.2.2 Progress of the semantic brokerage service	12
4.2.3 Rosetta Stone for vocabularies	12
4.2.4 NOAA One Stop/NCEI developments	12
4.2.5 Discussion	12
4.3 SESSION 3 - ODIP 2+ Prototype Development Task (plenary)	13
4.3.1 ODIP 2+: current status and planned development activities	13
4.3.2 R2R status	13
4.4 SESSION 4 – ODIP 4 Prototype Development Task (plenary)	13
4.4.1 Introduction	13
4.4.2 Progress with VRE analysis for SeaDataCloud	14
4.4.3 I - notebook – DIVA	14
4.5 SESSION 5 – ODIP 5 Prototype Development Task (plenary)	14
4.5.1 Progress towards a marine biology prototype	14
4.5.2 OBIS Environmental data scheme	15
4.5.3 MEOP Data schema	15
4.5.4 IMOS tracking data and current links with MEOP - OBIS	15
4.5.5 Discussion	15
4.6 SESSION 6: ODIP II: prototype impact assessment (WP4)	16
4.6.1 Prototype impact assessments	16
4.7 SESSION 7: Ocean Data Standards and Best Practices 1: Processes and Platforms	16
4.7.1 Introduction	16
4.7.2 Approach to Best Practices	16
4.7.3 ODIS	17



4.7.4	Brokering service	17
4.8	SESSION 8: Vocabularies.....	17
4.8.1	Plenary	17
4.8.2	Validating interlinks with the SUMMR Interlink Validation Tool.....	17
4.8.3	ODIP: Provenance of mappings progress towards solutions.....	18
4.8.4	European SeaDataCloud vocabularies	18
4.8.5	EARS Ontology for manual events - Eurofleets Automatic Reporting System ..	18
4.8.6	Australian Ocean Data Network and IMOS vocabularies	19
4.8.7	Canada Ocean Networks Canada vocabularies progress	19
4.8.8	USA Rolling Deck to Repository & EarthCube vocabulary progress.....	19
4.8.9	RDA Vocabularies Semantic Services Interest Group report.....	20
4.8.10	Discussion	20
4.9	SESSION 10: Linked Data Developments.....	20
4.9.1	Why Linked Data? Our experience in the OpenGovIntelligence project	20
4.9.2	Progressing Linked Data for SeaDataNet: Modelling the catalogues	21
4.9.3	Progressing Linked Data for SeaDataNet: Implementation	21
4.9.4	Implementing schema.org Datasets.....	21
4.9.5	Linked Data – an update from the USA.....	22
4.10	SESSION 11: Standard Data Formats	22
4.10.1	Introduction.....	22
4.10.2	Progress in SeaDataCloud	22
4.10.3	Experiences with using NetCDF 4	22
4.10.4	Advancing netCDF-CF.....	23
4.11	SESSION 12 Ocean Data Standards and Best Practices – Existing practices from prototypes and their evolution.....	23
4.11.1	Introduction.....	23
4.11.2	SWE	23
4.11.3	R2R and CSR.....	23
4.11.4	Recommendations on data standards from the Oceans of Tomorrow (OoT) projects	24
4.11.5	Best Practices, Data Formats Gaps & Directions	24
4.12	SESSION 13: ODIP 3+ Prototype Development Task: plenary.....	24
4.12.1	Introduction.....	24
4.12.2	SWE profiles.....	24
4.12.3	SWE Vocabularies.....	25
4.12.4	SWE editors.....	25



4.12.5	SOS services and clients	25
4.13	SESSION 14: Workshop wrap - up.....	26
4.13.1	Ocean Data Standards- Discussion and wrap up.....	26
4.13.2	BODC ODIP 3+ progress.....	26
4.13.3	Post-project sustainability of the Ocean Data Interoperability Platform	26
Appendix A	Workshop Programme	27
Appendix B	Terminology.....	37



Executive Summary

The 4th ODIP II workshop was hosted by the Marine Institute in Galway, Ireland from the 02 to 05 October 2017. This final ODIP II workshop was dedicated to progress reporting for the on-going prototype development tasks and recurring cross-cutting topics, and the finalizing of the outstanding project deliverables in preparation for the end of the project in March 2018.

A total of 48 attendees, from 31 organizations in 13 countries took part in the 4th ODIP II workshop, which included representatives from the regional and global marine data infrastructures participating in the project, as well as other selected experts in marine data management.

This deliverable documents the extended minutes and actions from the 4th ODIP II workshop. Full details of the agenda and logistical arrangements for this meeting are described in deliverable *D2.4 ODIP II workshop 4*. All materials relating to this workshop, including the presentations made during the sessions that are referenced in this document, are hosted on the IODE website at: https://www.iode.org/index.php?option=com_oe&task=viewEventRecord&eventID=2155 and also accessible via the ODIP II website at: http://www.odip.org/content/content.asp?menu=0480000_000000



1 Introduction

ODIP II - Extending the Ocean Data Interoperability Platform project, is the successor to the recent ODIP project. As for the previous project, the aim of ODIP II is to promote the development of a common global framework for marine data management by establishing interoperability between existing regional marine data systems in Europe, the USA and Australia, and with other relevant global data infrastructures such as GEOSS, IOC-IODE and POGO.

The workshops planned as part ODIP II are fundamental to the objectives, outcomes and success of the project. Each workshop aims to bring together representatives from selected regional marine data systems and related global data infrastructures together with other selected technical and domain experts, in an effort to promote and support the development of a common global framework for marine data management.

The 4th ODIP II workshop, the last in the series of four planned for the ODIP II project, was hosted by the Marine Institute, Galway, Ireland on the 02 to 05 March 2017. The agenda for this workshop was dedicated to providing an update on recent progress on the ODIP II development tasks, and to bringing the current project activities to a conclusion in preparation for the end of the project on 31 March 2018.

This deliverable documents the extended minutes and actions from the 4th ODIP II workshop. Full details of the agenda and logistical arrangements for this meeting are described in deliverable *D2.4 ODIP II workshop 4*. All materials relating to this workshop, including the presentations made during the sessions that are referenced in this document, are hosted on the IODE website at: https://www.iode.org/index.php?option=com_oe&task=viewEventRecord&eventID=2155 and also accessible via the ODIP II website at: http://www.odip.org/content/content.asp?menu=0480000_000000

2 List of Participants

Forty eight (48) attendees, from 31 organizations in 13 countries took part in the 4th ODIP II workshop. Of these, ten participants joined the meeting remotely via video conferencing. The participant list is shown below:

Adam Leadbetter	Marine Institute, Ireland
Adam Shepherd	BCO-DMO, WHOI, USA (remote participation)
Alessandro Oggioni	CNR-IREA, Italy
Alexandra Kokkinaki	BODC, United Kingdom
Angelo Lykiardopoulos	HCMR, Greece
Athanasia (Sissy) Iona	HCMR, Greece
Charles Troupin	GHER, Univ. of Liege, Belgium
Christian Autermann	52North, Germany
Christian Schafer-Neth	AWI, Germany
Cristian Munoz Mas	SOCIB, Spain
Chris Wood	BODC, UK (remote participation)
David Neufeld	NOAA, USA (remote participation)
Dick Schaap	MARIS, Netherlands
Elena Patrescano	OGS, Italy
Enrique Wulff	CSIC, Spain (remote participation)
Ethan Davis	UCAR, UNIDATA, USA
Fabrizio Papeschi	CNR-IIA, Italy
Francisco Souza Dias	VLIZ, Belgium
Friedrich Nast	BSH, Germany
Gwenaelle Moncoiffe	BODC, UK
Helen Graves	Coordinator NERC (BGS), UK
Jay Pearlman	IEEE, France
Jordi Sorribas	UTM-CSIC, Spain
Juan Luis Ruiz	UTM-CSIC, Spain
Juan Gabriel Fernandez	ICTS-SOCIB, Spain
Justin Buck	BODC, UK
Francoise Pearlman	IEEE, France
Karen Stocks	R2R-UCSD, USA (remote participation)
Leda Pecci	ENEA, Italy (remote participation)



Michèle Fichaut	IFREMER, France
Paolo Diviacco	OGS, Italy
Paolo Tagliolato	CNR-ISMAR, Italy
Peter Pissierssens	IOC/IODE, Belgium
Peter Thijsse	MARIS, Netherlands
Ramona Carr	Marine Institute, Ireland
Raymond Cramer	BODC, UK (remote participation)
Renata Dividino	Dalhousie University, Canada
Reyna Jenkyns	ONC, Canada (remote participation)
Rob Thomas	BODC, UK
Rob Van Ede	TNO, Netherlands
Sebastien Mancini	IMOS/AODN, Australia
Sergey Belov	RIHMI-WDC/NODC, Russian Federation
Simon Claus	VLIZ, Belgium
Simon Jirka	52North, Germany (remote participation)
Juan Luis Ruiz	UTM-CSIC, Spain
Susana Diez	UTM-CSIC, Spain
Thomas Loubrieu	IFREMER, France (remote participation)
Thomas Vandenberghe	BMDC, OD Nature, Belgium

3 Workshop Agenda

The agenda for 4th ODIP II workshop focused on progress reporting by the ODIP II prototype development tasks and associated cross-cutting themes. In addition, the workshop also addressed the activities necessary to finalize the remaining deliverables, preparation for the final reporting, and post-project sustainability of the Ocean Data Interoperability Platform including the associated wider ODIP community that has evolved throughout the two phases of the project.

The format of the 4th ODIP II workshop was largely the same as that adopted for previous meetings with a mix of plenary, discussion and breakout sessions (see Table 1 below).

The workshop agenda included a dedicated session for each of the existing prototype development tasks and the related cross-cutting topics (see Annex A for full workshop agenda).

Session	Title	Leader
1	Introduction	Helen Glaves
2	ODIP 1+ prototype development task	Peter Thijsse
3	ODIP 2+ prototype development task	Friedrich Nast
4	ODIP 4 prototype development task	Dick Schaap
5	ODIP 5 prototype development task	Francisco Souza Dias & Simon Claus
6	Prototype impact assessments	Michele Fichaut / Dick Schaap
7	Ocean Data Standards and Best Practices	Sissy Iona/Jay Pearlman
8	Vocabularies	Rob Thomas
9	Data and software publishing	Cyndy Chandler
10	Linked Data developments	Adam Leadbetter
11	Standard data formats	Justin Buck
12	Ocean Data Standards and Best Practices	Sissy Iona & Jay Pearlman
13	ODIP 3+ prototype development task	Simon Jirka
14	Wrap-up	Dick Schaap

4 Workshop proceedings

Detailed minutes were taken during the workshop that are documented below, along with any actions arising.

The presentations made during the workshop and referenced in this document are hosted on the IODE website at: https://www.iode.org/index.php?option=com_oe&task=viewEventAgenda&eventID=2155 and also accessible via the ODIP II website under the “Workshops” menu option at: http://www.odip.org/content/content.asp?menu=0480000_000000

4.1 SESSION 1

4.1.1 Introduction

Adam Leadbetter (MI), welcomed participants to the Marine Institute, Galway and opened the 4th ODIP II workshop. Helen Glaves (ODIP II Coordinator) thanked the hosts for the local logistical arrangements and provided some information on the format and objectives for the fourth and final ODIP II workshop. HG then invited participants to introduce themselves.

4.1.2 ODIP II: Overview of the project

Helen Glaves (NERC-BGS) gave an overview of the ODIP II project, explaining the concept, objectives and approach for developing interoperability across the different participating marine data systems, including the achievements to date. HG outlined the objectives for the workshop and stressed the need to wrap up the remaining actions in the prototype development tasks, cross cutting topics and impact assessment in preparation for the end of the project. The presentation can be accessed [here](#).

4.1.3 ODIP II: Technical Objectives

Dick Schaap (MARIS) introduced the overall approach for developing interoperability between the different data infrastructures to create a global framework for marine data management, and explained the technical challenges that were being met. DS gave an overview of the five prototype development task and the related deliverables produced to date. The presentation can be accessed [here](#).

4.2 SESSION 2 - ODIP 1+ Prototype Development Task (plenary)

4.2.1 ODIP 1+: current status and future development activities

Peter Thijssse (MARIS) explained the concept for the ODIP 1+ prototype development task that is focused on establishing semantic and syntactic interoperability between the SeaDataNet, IMOS and US NODC/NCEI data discovery and access services. The initial aim of this task was to implement a central broker service to harvest the three available metadata catalogues, and then visualize the result. PT explained that the initial aim has moved on and this task now aims to achieve a richer common format with a semantic mapping. Further details are included in the presentation available [here](#).

4.2.2 Progress of the semantic brokerage service

Fabrizio Papeschi (CNR-IIA) presented the progress that has been made on developing the pilot semantic brokerage service. He explained the broker system components, the main features, the architecture as well as the ODIP 1+ activities for establishing semantic discovery interoperability between the three regional catalogue services. He then demonstrated the prototype portal interface, several searching functionalities and how to constrain the discovery results. FP also showed an example of the semantic discovery from the CDI attribute vocabulary, and a second one for the instrument vocabulary. More details can be found in the presentation [here](#).

4.2.3 Rosetta Stone for vocabularies

Alexandra Kokkinaki (BODC) explained the prototype Rosetta Stone service, developed by BODC, to facilitate translations between terms in different vocabularies. The target vocabularies selected were: the European NERC Vocabulary Server (NVS); Australian AODN Vocabularies; and US NODC Vocabularies. In order to implement the translation service, BODC chose carefully selected terms from the following categories/vocabularies: platforms, instruments, organizations, and parameters. The terms were then mapped through the relation properties *owl:sameAs* and *skos:broader*. The Rosetta Stone service currently works for the following mappings: NODC to NVS and AODN to NVS. Further support for NODC to AODN or AODN to NODC mappings will require NODC and AODN to register these mappings in their SPARQL endpoint. Full details are available in presentation that can accessed [here](#).

4.2.4 NOAA One Stop/NCEI developments

David Neufeld (NOAA), presented recent developments in the OneStop project. The project aims to improve NOAA's data discovery and access framework. It is addressing data format and metadata best practices, improving collection-level metadata management and granule level metadata systems to accommodate the wide variety and vast scale of NOAA's data holdings. OneStop provides a seamless search experience between collections and granules files, and ensures more data are available through modern web services. DN demonstrated the functionalities of the OneStop data discovery interface using several exemplars. He also mentioned that, although first released in December 2016, this system is not yet fully operational due to the vast historical data resources that need to be incorporated into the system.

4.2.5 Discussion

The group discussed several issues related to the ODIP 1+ prototype development including how the vocabulary mappings will be done and what parameters to map. The question of how other partners, other than those already discussed (NCEI, AODN, CDI), can connect to the broker was also raised. It was agreed that there is a need for documentation on how to implement this service. PT concluded the session by outlining the work plan for this task over the coming months until the end of the project. Details of this work plan can be viewed [here](#).

4.3 SESSION 3 - ODIP 2+ Prototype Development Task (plenary)

4.3.1 ODIP 2+: current status and planned development activities

Friedrich Nast (BSH) introduced the session and presented the current status of ODIP 2+ including harmonization with the ICES system and planned future activities. There are now more than 3000 cruise summary reports in the CSR database, with 750 being new additions from the US via the R2R project. (A further 13 new cruises having been added since the previous workshop.) The other recently added CSRs are updated records from the ICES database which are still in V0 format (free text, no controlled vocabularies). FN identified suggested that a possible goal would be for R2R to adopt the same harvesting method as that used for SeaDataNet before the end of ODIP II.

Australia has submitted new CSRs to the database since the previous workshop (cruises from the Aurora Australis 2014 – 2017). There are also a total of 2100 new CSRs added by European partners. However, there are only four SeaDataNet partners currently connected to the system for regular harvesting of CSRs.

FN continued by explaining the current harmonization activities between BSH and ICES for regular exchange of CSR records that includes daily updates. He then concluded with an overview of future planned activities including providing IODE with a research vessel cruise catalogue system, introducing persistent identifiers for CSRs (DOI) and scientists (ORCID), and the automatic generation of CSR from shipboard systems. Full details are available in the presentation that can be accessed [here](#).

4.3.2 R2R status

Friedrich Nast (on behalf of Karen Stocks) gave the progress report for the ODIP 2+ activities undertaken by the US R2R partners. Thirteen CSRs from the RV Falkor were added to the database and harvested by POGO. It was observed that this could be an alternative mechanism for cruise reporting i.e using an automatic system like that used on RV Falkor to generate CSRs. The plan for the future is for CSR production on RV Falkor to continue (funded by Schmidt Ocean Institute that operates RV Falkor). Some further details were also given about the work being done on Cruise ID and Persistent Identifiers by R2R. Further information can be found in the presentation available [here](#). KS also gave a more detailed presentation on this topic during the Best Practices session (see Section 4.11.3 below).

4.4 SESSION 4 – ODIP 4 Prototype Development Task (plenary)

4.4.1 Introduction

Dick Schaap (MARIS) introduced the concept of the 'digital playground' being formulated in the ODIP 4 prototype development task. The initial approach for this task includes reviewing and analyzing relevant existing projects and initiatives, looking for common development opportunities, as well as specific potential solutions for creating VREs suitable for users in the marine domain. DS emphasized that this prototype is not focusing on content but on getting a better insight into the suitable technologies and the best ways to implement them.

4.4.2 Progress with VRE analysis for SeaDataCloud

Peter Thijsse (MARIS) presented an analysis of possible options for a suitable virtual research environment (VRE) architecture. The SeaDataCloud VRE needs to facilitate both collaborative and independent research. It must provide functionality for combining the SeaDataNet data resources with those from other external sources, and deliver large capacity and high performance computing for big data processing and state-of-the-art web visualisation services. The VRE must also respect the privacy of users and the different data policies currently in place. It should allow users to configure virtual work spaces for individuals or groups to work on specific projects, or allow producers to decide whether their outcomes will be shared in the public domain or remain private.

The H2020 SeaDataCloud project will use the EUDAT's infrastructure for deploying a marine VRE. (Architectures from other projects such as those developed by EVER-EST, BlueBridge NecTAR/MarVL were also analyzed to determine if they might be suitable solutions for SeaDataCloud.) Also the conclusions of this analysis were presented and how these could be combined with the EUDAT's infrastructure. He also presented the VRE technical options. Click [here](#) for the details on the presentation.

4.4.3 I - notebook – DIVA

Charles Troupin (ULg) introduced the concept of I-notebooks for documenting workflows. I-notebooks provide an interactive computational environment for combining code, text, results, and figures. CT noted that I-notebooks are not a virtual research environment, but are defined as interactive computational environments which can be a component of other VREs. I-notebooks are not a new concept have been in existence for more than 15 years, but their use has evolved. I-notebooks are very useful for documenting workflows, for example, they provide the capability to generate a climatology that can be replayed and amended. Examples of existing I-notebooks include Jupyter, R-Markdown and Apache Zeppelin. Details of further examples of I-notebooks, including a comparative analysis of the different tools, is included in the associated presentation accessible [here](#). CT concluded his presentation with an example of DIVA that includes automatic data download. This example is accessible [here](#).

4.5 SESSION 5 – ODIP 5 Prototype Development Task (plenary)

4.5.1 Progress towards a marine biology prototype

Simon Claus (VLIZ) introduced the session dedicated to the marine biology prototype development task. SC continued by presenting a proposal for a marine biology orientated prototype focusing on animal tracking data. Bio-telemetry is a powerful technology for remotely collecting data and information that provides unprecedented insights into animal movement behavior, social interactions, and the physical environment. Using organisms as sensor platforms provides very valuable information on the physical state of the oceans, but also on the occurrence, behavior and migration of the tagged animals.

Exchanging data between biological and physical data systems will add value. However, a key issue is that different communities use different data systems and data standards. In this context possible next actions include: 1) identification of relevant datasets within MEOP,

OTN and OBIS databases that contain occurrence and physical data; 2) cross matching data schemas between systems to identify commonalities and differences; 3) identify a pilot dataset to integrate occurrence/physical data into relevant data systems and propose mechanisms for interlinking them; 4) identify suitable science use case; and 5) propose a plan for automation. More details are available [here](#).

4.5.2 OBIS Environmental data scheme

Ward Appeltans (IOC/IODE) presented the progress of the OBIS pilot project OBIS-ENV-DATA. This project was initiated in 2015 following a recommendation from the 23rd session of IODE (IODE-XXIII.4). It was subsequently renamed to "OBIS-Event-Data" during the 24th session of IODE (recommendation IODE-XXIV.3). The main objective of OBIS-ENV-DATA is to expand OBIS to include environmental data. WA then explained the database schema which is based on the Darwin Core standard. All measurement types, values and units are standardized to the BODC NERC controlled vocabularies (NVS2.0). WA also showed a test data set, a Satellite Relay Tagging Program - Near real-time CTD profile data, and described how it is stored in the database. Further details of this presentation are available [here](#).

4.5.3 MEOP Data schema

Justin Buck (BODC) explained how BODC supports the near-real-time supply of SMRU Seal Tag data, via the Met Office, to the GTS for operational use.

JB explained that both the Marine Mammals Exploring the Oceans Pole to Pole (MEOP) consortium and this seal tag data are of international importance. MEOP datasets have a DOI, but this is limited in terms of discoverability and accessibility. The aim for MEOP is to make data that is not readily accessible, discoverable and deliverable through the BODC portal and the EMODnet data infrastructure. JB then explained what is required to progress the aims of MEOP. He concluded by outlining the work that has been done with the SMRU to facilitate automated data submission. Further details of this use case can be found in the presentation available [here](#).

4.5.4 IMOS tracking data and current links with MEOP - OBIS

Sebastien Mancini (AODN) presented recent IMOS developments on animal tracking data and current links with MEOP-OBIS. IMOS undertakes animal tracking using both acoustic telemetry and satellite tagging (bio-logging). The bio-logging component is focused on the Southern Ocean and southern Australia. The Antarctic and the surrounding Southern Ocean are one of the most important, yet least observed of marine habitats. The merging of oceanography and marine mammal ecology advances our understanding of the world's oceans and its top predators, and allows us to predict how these species will be affected by future climate changes. In the presentation [here](#) you can see how IMOS CTD data are collected by seals and how they linked with the MEOP-OBIS system, a related [video](#) for a month of data as well as a [video](#) for all deployments are also available.

4.5.5 Discussion

The discussion that followed focused on the availability of biological data and how the data management plans for these activities could improve the flow of data into ODIP II related activities.

4.6 SESSION 6: ODIP II: prototype impact assessment (WP4)

4.6.1 Prototype impact assessments

Thomas Loubrieu (Ifremer) summarized the objectives of WP4 which is conducting the impact assessments for the ODIP prototype interoperability solutions. He explained the methodology used and the results of the initial analysis, including the potential benefits identified for each prototype. TL remarked that implementing or adopting the ODIP II prototype solutions might have significant implications for the existing operational systems including the need for modifications throughout from the portals to the distributed data providers.

TL commented that the impact analysis could also potentially provide new insights which are relevant for wider implementation of the ODIP II prototype solutions in other existing systems. Two strategic analysis reports, which will provide details of the implications and benefits of the ODIP II interoperability solutions for the existing systems, are foreseen by the end of the project. The initial analysis, which will form the basis for the deliverable D4.1, was discussed and reviewed during the workshop. It can be viewed [here](#). The final impact analysis will be submitted at the end of the project as deliverable D4.2.

4.7 SESSION 7: Ocean Data Standards and Best Practices 1: Processes and Platforms

4.7.1 Introduction

Jay Pearlman (IEEE) introduced the session and presented the candidate best practices that have already been identified. JP also explained the difference between a best practice and a standard. A standard is something that is considered by an authority or by general consent as a basis for comparison (dictionary.com) or something set up and established by authority as a rule for measure of quantity, value or quality (Merriam-Webster). A best practice is a procedure or set of procedures that is preferred or considered standard within an organization, industry, etc. (dictionary.com); or a method or technique that has been generally accepted as superior to any alternatives because it produces results that are superior to those achieved by other means or because it has become a standard way of doing things. (Wikipedia). JP concluded by explaining the activities planned for the future. Further details can be accessed [here](#).

4.7.2 Approach to Best Practices

Cristian Munoz (SOCIB) presented the work done over the last 6 months by the Best Practices (BP) working group. A best practice is a procedure that has been shown by experience to produce optimal results. A community best practice is formed when a methodology is repeatedly shown to produce superior results relative to other methodologies with the same objective. For example, the European FixO3 project created documentation for fixed-point ocean observatories that was compiled and reviewed by the project experts. However, such documents often do not have a proper structure that allows them to be returned through a search. A proposal has been put forward for the creation of an indexed repository to expose semantically indexed documents on the web. This proposed solution will be discussed at a joint workshop on “*Evolving and Sustaining Ocean Best Practices*”

during November 2017 that includes participation by the AtlantOS and ODIP II projects, JCOMMOCG and IODE. More details of this proposal can be found [here](#).

4.7.3 ODIS

Peter Plssiessens (IOC/IODE) explained that the Ocean Data and Information System (ODIS) has been established in response to an audit of IODE by the Intergovernmental Oceanographic Commission (IOC) during April 2016. A draft concept paper was submitted to the 24th IODE session for approval. The proposed approach stated that the implementation should be undertaken with partners within and outside the IODE, and should build upon the success of established systems. The end goal is to establish ODIS such that it is inclusive and implemented to support stakeholders at all levels. PP observed that many of the ODIP II activities appear to be comparable to those necessary for the implementation of ODIS. He suggested that ODIS could build upon the legacy of ODIP. More detailed background information and a description of ODIP are available [here](#).

4.7.4 Brokering service

Stefano Nativi (CNR-IIA) presented the brokering service and the associated best practices and standards that could potentially be derived from it. The ODIP Broker is a complex system with many components for the discovery, evaluation, access and transformation of data sets.

Lessons learned so far are that there are different perspectives to be considered when formalizing best practices and design specifications. SN then described the analysis of four different points of view and their respective requirements: 1) the business view is based on use cases and their requirements; 2) the information view is based on conceptual (abstract) models and schema models (implementations); 3) the engineering view needs a system of systems architecture applying the brokering pattern and 4) the functional process which includes the main functionalities. Full details are available in the presentation which is accessible [here](#).

4.8 SESSION 8: Vocabularies

4.8.1 Plenary

Rob Thomas (MI) introduced the vocabularies cross-cutting theme and the progress that has been made on the activities that were defined during the previous workshop. The complete list of these activities including recent progress is available [here](#).

4.8.2 Validating interlinks with the SUMMR Interlink Validation Tool

Alan Meehan (ADAPT Centre, Trinity College Dublin) gave an invited keynote presentation entitled “Validating Interlinks with the SUMMR Interlink Validation Tool”. The tool is used for maintenance and reuse of vocabulary mappings and RDF interlinks and provides SPARQL templates for maintenance and reuse use cases. AM explained the motivation for using this type of tool and the choice of templates that were selected. He also explained the types of mapping that the tool can perform and gave an example of a case study. AM described how the SUMMR tool is applied to DBpedia interlink management process. SUMMR is an open

source available for download and use. Future work includes examining automatic detection and repair of invalid mappings between datasets. The presentation is accessible [here](#).

4.8.3 ODIP: Provenance of mappings progress towards solutions

Alexandra Kokkinaki (BODC) explained that provenance information ensures that there is confidence in the mappings carried out for NVS2.0. Details of who carried out the mappings and their reliability will be stored alongside the mappings.

AK mentioned that capturing provenance information about the mapping does create some practical problems. For example, storing this provenance information will require changes to all of the mappings, and one triple will need to be represented by at least four triples which leads to a lot of unnecessary load for NVS2.0. However, this problem could be addressed by using Named Graphs which are a set of triples labelled with an URI (quad). This then turns the RDF triple model into a quad model by extending a triple to include an additional item of information.

Named Graphs provides the ability to group triples around a URI which underpins a number of features including: tracking provenance of RDF data, replication of RDF graphs, managing RDF datasets, versioning, access control etc. More technical details can be found in the presentation available [here](#).

AK also reminded everyone that NVS 1.0 is now obsolete and will shortly be retired. Users should point their applications and services to NVS 2.0.

4.8.4 European SeaDataCloud vocabularies

Gwen Moncoiffe (BODC) reported on the common vocabulary developments in SeaDataCloud and other related European projects. The NERC Vocabulary Server (NVS) can be accessed either machine-to-machine or using several different browser interfaces (NVS Search, NVS Editor and NVS Vocabulary Builder). NVS 2.0 content is administered through the Vocabulary Management Group.

Six SeaDataCloud deliverables will document the various aspects of the on-going or future vocabulary development activities that include: 1) Improving the transparency of the vocabulary governance model, 2) developing new vocabularies, 3) undertaking a review of P02 and P03 vocabularies, 4) exposing versioning and history of concepts through NVS, 5) documenting provenance of mappings, and 6) further developing the Vocabulary Builder tool. GM also presented some illustrations of these activities which can be viewed [here](#).

4.8.5 EARS Ontology for manual events - Eurofleets Automatic Reporting System

Thomas Vandenburghe (RBINS-BMDC) presented an overview of the Eurofleets Automatic Reporting System (EARS), a software package for managing research cruises. He explained that an oceanographic event is any circumstance, fact or situation that happens on board ship that should be logged e.g. sampling, acquisition and deployment.

TV then demonstrated how events are created in EARS. The EventDefinition templates are stored in the EARS ontology, and when the EventDefinition is created it is timestamped. The EARS front-end application can manage cruises and programs, browse terms and their definition, manage terms and create events from terms.

TV described the underlying EARS architecture. TV concluded his presentation with list of the terms and vocabularies that have been included in the EARS cruise summary reports for compatibility with CSRs. Further details are available in the presentation that can be accessed [here](#).

4.8.6 Australian Ocean Data Network and IMOS vocabularies

Sebastien Mancini (AODN) presented the Australian National Environmental Information Infrastructure (NEII) Vocabulary Service for creating, curating and publishing environmental vocabularies to make them fully accessible. The NEII vocabulary service uses the Australian National Data Service's (ANDS) Research Vocabularies Australia (RVA) platform as its backbone. The ANDS platform provides a nationally consistent approach and leverages existing investment. The NEII Linked Data API provides additional flexibility and system redundancy.

A major challenge for NEII has been communication and engagement with the various vocabulary owners (NEII, GA, DoEE, IMOS). However, the current challenge is the system level integration of vocabulary content within and across organisations.

SM continued with an overview of the Australian Vocabulary Special Interest Group that provides a forum for discussion and activities relating to the use and creation of controlled vocabularies in research, data, information and collection management. Further details on NEII and the AVSIG can be found in the presentation available [here](#).

4.8.7 Canada Ocean Networks Canada vocabularies progress

Reyna Jenkyns (ONC) presented the progress on implementing vocabularies and related applications at Ocean Networks Canada (ONC). RN outlined the Oceans 2.0 API, which is a collection of web services for accessing the ONC data. The API services are split into those for discovery, delivery and standards (ERDDAP & OGC SOS).

RJ continued with an overview of other relevant activities include the use of vocabularies for Dive Logging (ROV Dive Annotation Standards), platform descriptions and marine protected area management (MPA) tools. Of particular note for the ODIP II community is that the NERC 2.0 vocabulary server has been adopted for all of these activities and applications RJ concluded her [presentation](#) with some details of the current challenges being addressed and the future directions for ONC.

4.8.8 USA Rolling Deck to Repository & EarthCube vocabulary progress

Adam Shepherd (WHOI) gave an overview of the R2R practices and use of DOI and other persistent identifiers. The motivations for R2R using persistent identifiers can be summarised using the 3 Rs: reproducibility (what journals want), reuse (what funders want), and recognition (what scientists want).

AS continued with some details of current R2R practices that could also be adopted more widely as best practices that include: a) the creation of DOIs for cruises and R2R datasets, b) integration of external PIDs into R2R Linked Data Graphs, c) creating reciprocal links with partners, and d) leveraging existing vocabularies e.g. NVS 2.0.

AS explained that there are still some challenges with adopting persistent identifiers including: lack of community consensus on all identifier types e.g. those for instruments, projects, organizations, significant amount of backfill mapping is still required, and identifier

systems are not perfect, for example, a person can have more than one ORCID. The presentation can be viewed [here](#).

4.8.9 RDA Vocabularies Semantic Services Interest Group report

Adam Shepherd (WHOI) provided an update on the recent activities of the RDA Vocabularies Semantic Services Interest Group. VSSIG was previously known as the Vocabulary Services (VSIG) Interest Group. AS reported that along with the name change there has been a change in the co-chairs with John Graybeal (who was previously working in data and semantics for the Earth Science, but who is now involved in the biomedical field at Stanford University), and Yann Le Franc (CEO of e-Science Data Factory, Paris and EUDAT partner) now leading this interest group.

During the 10th RDA Plenary in Montreal, the restructured VSSIG defined a number of activities to be undertaken which relate primarily to vocabulary governance and ontologies Canada. The full list of VSSIG activities can be found in the presentation [here](#). ODIP II partners with an interest in this topic can get involved via the Slack discussion group at: <https://vocabulary-services.slack.com/> and also join the VSSIG e-mail list <http://bit.ly/rda-vssig>

4.8.10 Discussion

During the discussion that followed participants reviewed the NVS wish list that was originally defined during the first ODIP II workshop held in Paris during September 2015, and subsequently updated and expanded during the second workshop in Boulder, Colorado during May 2016. The full list, which includes vocabulary tooling, content, mapping, best practices etc., can be reviewed [here](#).

4.9 SESSION 10: Linked Data Developments

4.9.1 Why Linked Data? Our experience in the OpenGovIntelligence project

Adegboyega Ojo (Insight Centre for Data Analytics) gave an overview of the European OpenGovIntelligence project that aims to improve public services using multidimensional statistical data.

OpenGovIntelligence is a three years project, starting in February 2016, which has 12 participants from seven countries. It aims to provide data-driven public services, rather than the raw data, in an effort to fulfil the needs of society. As a result society should be involved in service co-production to ensure that these public services address their needs.

Statistical Data is one of the five thematic dataset categories in highest demand from re-users across the EU (EC, 2014), which are often organized as data cubes where each cell contains a measure described based on a number of dimensions. Linked Open Statistical Data (also known as linked data cubes) have the potential to address data interoperability and facilitate data integration on the Web, and to perform analytics on top of multiple (previously isolated) datasets in order to co-create added value services.

AO then presented an overview of six on-going pilot projects/use cases that covered: 1) managing government vehicles in Greece; 2) business planning in Lithuania; 3) environmental planning in Belgium; 4) addressing unemployment in Trafford (UK); 5)

exposing real estate data in Estonia; and 6) use of ocean data for search and rescue services in Ireland. AO concluded his presentation by giving some further detail about Pilot 6 (marine data for search and rescue applications) which is being developed in partnership with the Marine Institute. More details can be found [here](#).

4.9.2 Progressing Linked Data for SeaDataNet: Modelling the catalogues

Adam Leadbetter (MI) reported on the Linked Data Developments in the SeaDataCloud project. He explained that SeaDataNet is using linked data as a method for publishing structured data so that it can be interlinked and therefore become more useful through semantic queries.

Existing Linked Data patterns are re-used for the SeaDataNet catalogues (EDMO, EDMED, EDMERP, CDI, CSR, EDIOS), which will provide better understanding outside of SDN, better interoperability with other organisations and better INSPIRE compliance. AL gave an example of using the linked data pattern for cruise summary reports (CSR) that can also be adopted for ODIP II.

He then outlined some of the challenges that have been encountered with this approach and described some detailed examples of how they are being resolved e.g. interoperability of EDMO with other external vocabularies. AL concluded the presentation with an overview of the next steps that include: developing RDF implementation of catalogues, and creating the required Schema.org mappings. More examples of the challenges encountered in implementing linked data for SeaDataNet and other details are available [here](#).

4.9.3 Progressing Linked Data for SeaDataNet: Implementation

Alexandra Kokkinaki (BODC) reported on the progress that has been made in publishing the SeaDataNet catalogues EDMED and EDIOS as linked data (via SPARQL endpoints) using the linked data analysis carried out by the Marine Institute. She explained the approach, and gave several examples of the URLs for the two catalogues, SPARQL queries and the results of the queries for the two catalogues.

AK provided a list of the technologies being used for this approach which are: Java application, Jena RDF API, PL/SQL functions, Jena Fuseki, and Cron job overnight. She concluded by describing the issues with dates and contact references that have been encountered and the ongoing efforts to resolve them. Full details are available in the presentation [here](#).

4.9.4 Implementing schema.org Datasets

Adam Leadbetter (MI) explained that schema.org is compatible with all search engines (Google, etc). The schema facilitates the discovery of public datasets. Examples were shown of its implementation in ERDDAP data server, Drupal content management software, and THREDDS data server.

AL reported on the next steps for implementing Schema.org that include: confirming the proposed implementation with Google; proposing the code merger for ERDDAP; extending Linked NetCDF for schema.org; and building an EarthCube registry of related scientific resources. Full details of this schema.org implementation are available [here](#).

4.9.5 Linked Data – an update from the USA

Adam Shepherd (WHOI) gave an update on the use of persistent identifiers for Linked Data. The current practices in U.S. ocean sciences include: DOIs for journal papers/articles, DOIs for datasets, DOIs for documents, FundRef codes for awards, IGSNs for samples, and ORCIDs for people. AS presented several examples of how BCO-DMO uses persistent identifiers (PIs) to improve the discovery of related research resources. He also provided an overview of the US GeoLink project. AS continued by explaining the issues encountered with using PIs as properties. He concluded his presentation by giving an overview of the BCO-DMO use of persistent identifier patterns including the SPARQL implementation. Full details are described in the presentation [here](#).

4.10 SESSION 11: Standard Data Formats

4.10.1 Introduction

Justin Buck (BODC) introduced the session on data formats by giving some context for the presentations to follow. He outlined the ongoing efforts to align the SeaDataNet NetCDF format with O&M. Use of NetCDF4 is becoming more widespread with feature beginning to be incorporated into in-situ earth observation formats. Two relevant workshops took place since the last ODIP II meeting: 1) AtlantOS Transatlantic Ocean Data Harmonization Workshop (June 2017), and 2) EarthCube NetCDF-CF workshop (September 2017). The presentation can be viewed [here](#).

4.10.2 Progress in SeaDataCloud

Ray Cramer (BODC) reported on the progress made to review and specify how the SeaDataNet NetCDF/ODV formats can be used as the basis for an INSPIRE compliant data format, following O&M. The main tasks are: a) review feasibility of transforming SeaDataNet formats into INSPIRE O&M data standards (following analysis of INSPIRE data implementation rules); b) review feasibility of merging CDI metadata into SeaDataNet ODV and NetCDF files to enable delivery of metadata-enriched data sets as part of the CDI service; c) review implications of migrating from NetCDF V3.6 to V4.0 (time-permitting); and d) formulate a SeaDataNet NetCDF (CF) format for gridded data, including CDI metadata. Details of each task are described in the presentation which can be found [here](#).

4.10.3 Experiences with using NetCDF 4

Charles Troupin (ULg) began his presentation by noting that netCDF is the acronym for 'network Common Data Form' and not format as often used. He also highlighted that according to the Unidata Best Practices the correct usage is netCDF, netcdf, or NetCDF.

CT explained that NetCDF is a suite of software libraries and self-describing, machine-independent data formats. It has been an OGC standard since 2011, with the Climate and Forecast (CF) extensions added in 2013. The improvements for netCDF version 4, released in 2008, were explained. NetCDF is also now being used in several programming languages and the results of the comparison between several visualization and analysis tools were shown. CT concluded the presentation by recommending the wider implementation of netCDF-4. Full details can be found in the presentation that can be accessed [here](#).

4.10.4 Advancing netCDF-CF

Ethan Davis (UCAR / Unidata) provided an overview of the EarthCube “Advancing netCDF-CF” project, which is funded by the US National Science Foundation. This community driven project is developing netCDF-CF extensions, and developing prototype implementations in existing software packages.

NetCDF-CF has a number of supported scientific data types which include: gridded data; time series at a station, soundings, aircraft tracks; unstructured grids (e.g., triangular mesh); and CF-Radial (radial data for radar and LIDAR). Further data types are under development, including support for: time series for polylines or polygons (based on OGC Simple Geometries) and satellite swath data. Future netCDF developments include: CF data model, implementation of linked data for CF, support for data streaming, other on-going CF versioning activities. Full details can be found [here](#).

4.11 SESSION 12 Ocean Data Standards and Best Practices – Existing practices from prototypes and their evolution

4.11.1 Introduction

Jay Pearlman (IEEE) explained that the aim of the session is to discuss and decide which of the ODIP II outputs could be endorsed as recognised community best practices. Those that are selected can be recommended for global dissemination and adoption through the IODE ODSBP initiative, which aims to achieve broad agreement and commitment to the adoption of a number of standards and best practices ocean data management. Further details are included in the presentation available [here](#).

4.11.2 SWE

Christian Autermann (52°North) gave an update on recent developments of marine profiles for OGC Sensor Web Enablement Standards. Many oceanology projects and organizations are starting to use OGC SWE standards, and a common approach for how to apply them is needed. There are specific requirements that should be met such as: covering the main functionalities offered by sensor web technologies; ensuring compatibility between SWE applications in different organizations and projects; defining a set of minimum requirements which must be fulfilled for compliance with SWE profiles; use of vocabularies; allowing customization. CA concluded by giving an update on the current status of the relevant activities, and also provided some information on the next steps that are planned. Full details are included in the presentation available [here](#).

4.11.3 R2R and CSR

Karen Stocks (R2R-UCSD) and Friedrich Nast (BSH) presented current practices for cruise reporting within the R2R and CSR initiatives that could potentially be used as candidate best practices. These include: minting DOIs for cruises and R2R datasets, integration of external PIDs into the R2R linked data graph, creation of reciprocal links with partners, documenting the procedures from cruise planning to data archiving. See section 1.3.2 for further details of the relevant R2R activities. Details of the CSR activities are included in the presentation accessible [here](#).

4.11.4 Recommendations on data standards from the Oceans of Tomorrow (OoT) projects

Justin Buck (BODC) presented the recommendations on data standards from the Oceans of Tomorrow (OoT) projects that include: CommonSense, NeXOS, SCHeMA, and SenseOCEAN. These recommendations, derived from the joint deliverables produced across the OoT projects, are for the implementation of a number of OGC and W3C standards. Those from OGC include: Observations & Measurements (O&M), PUCK Protocol Standard, Sensor Model Language (SensorML), Sensor Observation Service (SOS), Sensor Planning Service (SPS), Sensor Alert Service (SAS), Smart Sensor Interface for Sensors and Instruments (SEISI), and Efficient XML Interchange (EXI). The recommended W3C standards are: Semantic Sensor Network (SSN), Resource Description Framework (RDF), and SPARQL Protocol and RDF Query Language (SPARQL).

Complementary activities for the individual OoT projects are addressing: marine SWE profiles, linked data, access control and cyber security, and PIDs for instruments / platforms / deployments. Planned next steps include formalising this work to converge on agreed recommended standards that can be promoted and adopted more widely. Full details of the OoT activities and outcomes are outlined in the presentation available [here](#).

4.11.5 Best Practices, Data Formats Gaps & Directions

Adam Leadbetter (MI) began his presentation by explaining that the best practices and standards that are implemented for the use of a dataset depend on the audience (user), the intended application, and the approach being taken.

There are several resources on the web from organisations such as W3C and Data One that document best practices related to data usage, data formats publications and on how to effectively work with data through all stages of its lifecycle. Some examples were shown using different standards depending on their directions. Full details can be found in the presentation available [here](#).

4.12 SESSION 13: ODIP 3+ Prototype Development Task: plenary

4.12.1 Introduction

Simon Jirka (52°North) began the session with an overview of the agenda and the topics that will be addressed. The introductory slides can be viewed [here](#).

4.12.2 SWE profiles

Simon Jirka (52°North) explained that various projects are contributing to the development of SWE in the marine domain (NeXOS, ODIP, SeaDataCloud, and others), and there are several tools available for developing SWE profiles.

The likely outcome of these activities are OGC Best Practice documents containing profiles for SOS, SensorML and O&M. SJ highlighted that the JERICO-NEXT project has also developed new SensorML templates. Currently the documentation is often created as part of project deliverables but coherent documentation for profile developments is needed. SJ proposed setting up a Github repository for joint editing of profile documentation. A repository is already available but it needs to be moved to a neutral place (e.g. dedicated

GitHub org for the working group). Further details of the SWE profile development activities is available [here](#).

4.12.3 SWE Vocabularies

Alexandra Kokkinaki (BODC) explained that the marine domain has started implementing the Open Geospatial Consortium's (OGC) Sensor Web Enablement (SWE) standards with the aim of making all types of sensors, transducers and sensor data repositories discoverable, accessible and useable via the Web. These standards include: Sensor Model Language (SensorML), Observations and Measurements (O&M), Sensor Planning Service (SPS), Transducer Mark-up Language (TML), Sensor Alert Service (SAS), Sensor Observation Service (SOS) and Web Notification Service (WNS). She explained their characteristics, and how the interoperability and discoverability via the web can be reduced. A detailed description of the work undertaken by BODC to implement the SWE standards is available [here](#).

4.12.4 SWE editors

Christian Autermann (52°North) began with an overview of SWE metadata editors and the SeaDataCloud SWE Ingestion Service.

He continued with a description of the Sensor Nanny web application that provides services for observatories to manage their observations on the "cloud". The application sits alongside the reference services (vocabularies, organization directory) and standard profiles (OGC/Sensor Web Enablement) provided by SeaDataNet.

Sensor Nanny provides an on-line editor to graphically describe observatory (acquisition and processing systems). It currently supports auto-completion from SKOS configurations (but not from BODC yet, but coming soon) and can be applied to outputs, identifier terms, classifier terms, contact roles. SeaDataCloud SWE Ingestion Service is used as an online service to describe observatories (or networks of observatories) and as an ingestion service. It is based on the OGC Sensor Observation Service (SOS) 2.0-Standard. Wherever possible it relies on enhancing existing (open source) software components. The specification of the SWE Ingestion Service is nearly complete and it will be published as a SeaDataCloud deliverable. Some further details are also available [here](#).

4.12.5 SOS services and clients

Simon Jirka (52North) presented some further activities in the ODIP 3+ prototype development task related to SOS services.

The AWI NearRealTime database contains 307 billion observations starting from June 2009. The data, which is about 15 minutes old, are mobile and stationary measurements taken from research vessels, stations, and buoys. This installation is based on a 52°North SOS server with a backend that is based on data from database and metadata from sensor.awi.de. This service supports the following operations: GetCapabilities, DescribeSensor, GetObservation, GetFeatureOfInterest, GetDataAvailability. However, it currently only serves a subset of the available data as hourly mean values that are not publicly available as yet. Further information on this ODIP 3+ development activity are included in the presentation available [here](#). In addition, the recent developments on the Sensor Web Viewer "Helgoland" can be found [here](#).

4.13 SESSION 14: Workshop wrap - up

4.13.1 Ocean Data Standards- Discussion and wrap up

Jay Pearlman (IEEE) summarized the discussions on the best practices and the next steps which include: defining BP processes and implement initial operation capability; introduce peer review for best practices; expand uptake of the BP repository (JCOMM, EMSO, etc.); address transition from best practices to standards; define best practice and standards outputs from the ODIP II prototype solutions; submission of the proposed best practices to IODE/ODSBP project for wider adoption by the ocean community; and identify experts outside the ODIP II community to conduct the review of the proposals.

The potential candidate best practices for the IODE ODSBP are: brokering prototype solution; documenting of cruise reporting (note that CSR metadata model and xml encoding are already submitted to IODE); vocabularies; SWE standards and vocabularies used in SWE; Data formats (NetCDF4). Details of this activity are included in the presentation available [here](#).

4.13.2 BODC ODIP 3+ progress

Justin Buck (BODC) presented an overview of the various EU projects and research infrastructures that have SWE related requirements including SenseOCEAN, EMSO (EMSODev/Fix03), EMODNet data ingestion, and ENVRIplus.

JB explained that a project is needed to address the common requirements for sensor web enablement identified by this group of projects/infrastructures. The initial suite of requirements includes creation and publication of platform and sensor metadata and production of O&M for historic ANIMATE (PAP site) on an operational SOS server to meet the goals of EMSO ERIC and the EMODNet data ingestion SWE prototype. The next phase of work, planned for December 2016 /January 2017 aims to include linked data elements and possibly ERDDAP. Metadata elements have already been incorporated into the Oceanids C2 project that is building a UK marine glider/AUV DAC. More details of this work are included in the presentation available [here](#).

4.13.3 Post-project sustainability of the Ocean Data Interoperability Platform

Dick Schaap (MARIS) wrapped up the workshop with some initial comments on the post-project sustainability of ODIP II. He suggested there may be an opportunity for a further ODIP project that would include wider international cooperation. However, further lobbying and exploration of this opportunity with the European Commission is still required to clarify what may be possible going forward.

Dick Schaap closed the 4th ODIP II workshop and thanked the hosts at the Marine Institute for the excellent local organization of the logistics. He also acknowledged the partners for their participation and contributions to make the final ODIP workshop a resounding success.

Appendix A Workshop Programme

AGENDA

DAY 1: Monday, 02 October 2017

Session 1

- | | | |
|-------------|-------------------------------------|---|
| 8:45 – 9:00 | Registration | |
| 9:00 – 9:15 | Welcome | <i>Adam Leadbetter (Marine Institute)</i> |
| 9:15 – 9:30 | Workshop introduction and logistics | <i>Helen Glaves (ODIP II project coordinator)</i> |
| 9:30 – 9:45 | Introductions | <i>(Name, Country, institution, main responsibility, expectations for this workshop: 30 seconds max.)</i> |

ODIP II Overview

- | | | |
|---------------|----------------------------------|--|
| 9:45 – 10:00 | ODIP II: overview of the project | <i>Helen Glaves (ODIP II Coordinator)</i> |
| 10:00 – 10:10 | ODIP II: technical objectives | <i>Dick Schaap (ODIP II Technical Coordinator)</i> |

Session 2

ODIP 1+ Prototype Development Task: plenary

10:10 – 11:00 ODIP 1+: current status and planned development activities
Led by Peter Thijsse (MARIS)

- *Aims and approach – Peter Thijsse (MARIS)*
- *Progress of semantic brokerage service – Enrico Boldrini (CNR)*
- *Rosetta Stone for vocabs – Alexandra Kokkinaki*
- *IMOS/AODN catalogue - * Sebastien Mancini - IMOS (UTAS)*
- *NOAA One Stop/NCEI developments (remote contribution) - David Neufeld (NOAA)*

11:00 – 11:15 Coffee break

ODIP 1+ Prototype Development Task: plenary

11:15 – 12:15 Discussion *Led by Peter Thijsse*

Session 3

ODIP 2+ Prototype Development Task: plenary

12:15 – 12:45 ODIP 2+: current status and planned development activities
Led by Friedrich Nast

- *CSR status – Friedrich Nast (BSH)*
- *R2R status – Cyndy Chandler on behalf of Karen Stocks (UCSD)*
- *Adopting CSR reporting and discovery on the IODE website – Dick Schaap (MARIS)*

12:30 – 13:00 Discussion *Led by Friedrich Nast*

13:00 – 14:00 Lunch

Session 4

ODIP 4 Prototype Development Task: plenary

14:00 – 15:00

ODIP 4 (Creating a 'digital playground')

Led by Dick Schaap

- *Introduction – Dick Schaap (MARIS)*
- *Progress with VRE analysis for SeaDataCloud – Thomas Loubrieu (IFREMER)*

15:00 – 15:15 Break

ODIP 4 Prototype Development Task: plenary (continued)

15:15 – 16:00

ODIP 4 (Creating a 'digital playground')

Led by Dick Schaap

- *I - notebook – DIVA – Charles Troupin / Alexander Barth (ULg)*
- *VRE Architecture options – Peter Thijsse (MARIS)*
- *Virtual Geophysics Laboratory (VGL) - Carsten Friedrich (CSIRO)*

16:00 – 17:00

Discussion

Led by Dick Schaap



DAY 2: Tuesday, 03 October 2017

9:00 – 9:10 Introduction/Announcements

Helen Graves /Adam Leadbetter

Session 5

ODIP 5 Prototype Development Task: plenary

9:10 – 10:00 Progress towards a marine biology prototype

*Led by Francisco Souza Dias
Simon Claus*

- *Proposal for biological prototype focusing on marine mammal tracking data - Simon Claus*
- *OBIS Environmental data scheme - Ward Appeltans*
- *MEOP Data scheme - Justin Buck*
- *IMOS tracking data and current links with MEOP-OBIS - Sebastien Mancini*

10:10 – 10:45 Discussion

*Led by Francisco Souza Dias &
Simon Claus*

10:45 – 11:00 Break

ODIP II: prototype impact assessment (WP4)

11:00 – 11:45 Prototype impact assessments

Dick Schaap/Michele Fichaut / Thomas Loubrieu

11:45 – 12:30 Discussion

Led by Dick Schaap

Session 6

Ocean Data Standards and Best Practices (1): Processes and Platforms

12:30 – 13:00 Plenary *Led by Sissy Iona & Jay Pearlman*

- *Intro – Jay Pearlman and Sissy Iona*
- *Approaches for best practice repository – Cristian Munoz*
- *ODIS – Peter Pissierssens*

13:00 – 13:30 Discussion *Led by Jay Pearlman and Sissy Iona*

13:30 – 14:30 Lunch

Session 7

Vocabularies

14:30 – 15:30 Plenary *Led by Rob Thomas (MI)*

- ***Guest talk:*** *Validating interlinks with the SUMMR Interlink Validation Tool - Alan Meehan (ADAPT Research Centre, Trinity College Dublin)*
- *ODIP: Provenance of mappings progress towards solutions - Alexandra Kokkinaki (BODC)*
- *Europe SeaDataCloud vocabularies progress - Gwen Moncoiffe (BODC)*
- *Australia Australian Ocean Data Network & IMOS vocabulary progress - Sebastien Mancini (UTAS)*
- *USA Rolling Deck to Repository & EarthCube vocabulary progress - Cyndy Chandler (BCO-DMO/IODE)*
- *Canada Ocean Networks Canada vocabularies progress Reyna Jenkyns (ONC)*
- *RDA: Vocabularies Specials Interest Group report Adam Shepherd (WHOI)*

15:30 – 16:00 Break



Vocabularies (continued)

16:00 – 17:00

Discussion

*Led by Rob Thomas (MI)
& Gwen Moncoiffe (BODC)*

- *What can and should be achieved in the final 6 months of the project?*
- *What developments should be built on for future proposals?*

DAY 3: Wednesday, 04 October 2017

9:00 – 9:10 Introduction/Announcements

Helen Glaves /Adam Leadbetter

Vocabularies (continued)

09:10 – 09:45

Discussion

*Led by Rob Thomas (MI)
& Gwen Moncoiffe (BODC)*

- *What can and should be achieved in the final 6 months of the project?*
- *What developments should be built on for future proposals?*

Session 8

Linked Data Developments

9:45 – 11:00

Plenary

Led by Adam Leadbetter

- *Why Linked Data? Our experience in the OpenGovIntelligence project - Prof Adeboyega Ojo, Insight Centre for Data Analytics, NUI Galway)*
- *Progressing Linked Data for SeaDataNet: Modelling the catalogues - Adam Leadbetter (MI)*
- *Progressing Linked Data for SeaDataNet: Implementation (Alexandra Kokkinaki (BODC)*
- *Implementing schema.org Datasets - Adam Leadbetter (MI)*
- *Linked Data – an update from the USA Adam Shepherd (WHOI)*

11:00 – 11:15

Break

Linked Data Developments

11:15 – 12:30

Discussion

Led by Dick Schaap

12:30 – 13:30 Lunch

Standard Data Formats

13:30 – 14:30 Plenary *Led by Justin Buck*

- *Progress in SeaDataCloud – Lesley Rickards (BODC)*
- *Experiences with using NetCDF 4 – Charles Troupin (ULg)*

14:30 – 15:15 Discussion *Led by Justin Buck*

15:15 – 15:30

Break

Ocean Data Standards and Best Practices (2) Existing practices from prototypes and their evolution

15:30 – 16:45 Plenary *Led by Sissy Iona & Jay Pearlman*

- *Intro – Jay Pearlman and Sissy Iona*
- *SWE – Christian Autermann*
- *R2R and CSR – Cyndy Chandler and Friedrich Nast*
- *Brokering service – Stefano Nativi*
- *Evolving from practices to standards (incl ODSBP) – Jay Pearlman & Sissy Iona*

16:45 – 17:15 Discussion *Led by Sissy Iona & Jay Pearlman*



Thursday, 05 October 2017

9:00 – 9:10

Announcements

Adam Leadbetter

Session 11

ODIP 3+ Prototype Development Task: plenary

9:10 – 10:00

ODIP 3+: current status and planned development activities

Led by Simon Jirka

- *SWE profiles – Simon Jirka (52 North) + Thomas Loubrieu (IFREMER)*
- *SWE Vocabularies – Alexandra Kokkinaki (BODC)*
- *SWE editors - TBC*
- *SOS services and clients – Simon Jirka (52 North)*

10:45 – 11:00

Break

Workshop wrap-up

11:00 – 12:00

ODIP prototype development projects

Feedback from each group on activities during the workshop and next steps (max. 10 minutes each):

ODIP 1+ - Peter Thijsse

ODIP 2+ - Friedrich Nast

ODIP 3+ - Simon Jirka

ODIP 4 - Dick Schaap

ODIP 5 - Francisco Souza Dias / Simon Claus



12:00 – 12:30 Plans / options for the continuation and post-project sustainability of the Ocean Data Interoperability Platform

Led by Dick Schaap

12:30 – 13:00 Closing remarks

*Helen Glaves (remotely) &
Dick Schaap*

13:00 – 14:00

Lunch

Appendix B Terminology

Term	Definition
AODN	Australian Ocean Data Network
API	Application Programming Interface
CDI	Common Data Index metadata schema and catalogue developed by the SeaDataNet project
CF	Climate and Forecast conventions: metadata conventions for the description of Earth sciences data, intended to promote the processing and sharing of data files http://cfconventions.org/
CSR	Cruise Summary Reports is a directory of research cruises.
CSW	Catalog Service for the Web
DIVA	Data Interpolating Variational Analysis tool for spatial interpolation
DOI	Digital Object Identifier (DOI): a unique persistent identifier for objects which takes the form of a unique alphanumeric string assigned by a registration agency
EDMO	European Directory of Marine Organisations
EMODnet	EU-funded initiative to develop and implement a web portal delivering marine data, data products and metadata from diverse sources within Europe in a uniform way. http://www.emodnet.eu/
ERDDAP	NOAA Environmental Research Division's Data Access Program
EUDAT	European Data Infrastructure, http://www.eudat.eu/
GIS	Geographic Information System
GitHub	A distributed revision control and source code management (SCM) system (GIT) repository web-based hosting service which offers all of the distributed revision control and source code management (SCM) functionality of Git as well as adding its own features
GTS	Global Telecommunication System



ICES	International Council for the Exploration of the Sea
IMOS	Integrated Marine Observing System: Australian monitoring system; providing open access to marine research data
INSPIRE	Infrastructure for Spatial Information in the European Community
IOC	Intergovernmental Oceanographic Commission of UNESCO (IOC/UNESCO).
IODE	International Oceanographic Data and Information Exchange (part of IOC)
ISO	International Organization for Standardization http://www.iso.org
NOAA	National Oceanic and Atmospheric Administration
NetCDF	Network Common Data Form (NetCDF): a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data.
NVS2	NERC Vocabulary Server Version2
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
ORCID	Open Researcher and Contributor ID: a non-proprietary alphanumeric code to uniquely identify scientific and other academic authors and contributors http://orcid.org/
POGO	The Partnership for Observation of the Global Oceans: a forum created by the major oceanographic institutions around the world to promote global oceanography. http://www.ocean-partners.org/
R2R	Rolling Deck to Repository: a US project responsible for the cataloguing and delivery of data acquired by the US research fleet.



RDA	The Research Data Alliance (RDA) builds the social and technical bridges that enable open sharing of data.
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: a W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary
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
URI	Uniform Resource Identifier, a string of characters used to identify a resource
VRE	Virtual Research Environment
XML	Extensible Markup Language: a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable http://www.w3.org/XML/