Ocean Data Interoperability Platform

Deliverable D4.2: Final strategic analysis report

Workpackage	WP4	ODIP prototypes
Author (s)	Thomas Loubrieu	IFREMER
Author (s)		
Author (s)		
Author (s)		
Authorized by	Helen Glaves	NERC - BGS
Reviewers	Roy Lowry	NERC – BODC
	Simon Jirka	52^{o}NORTH
	Scott Bainbridge	AIMS
	Shane StClair	NOAA
Doc Id	ODIP_D4.2_Final_strategic_analysis	
Dissemination Level	PUBLIC	
Issue	1.0	
Date	14 October 2015	

Document History				
Version	**Author(s)**	**Status**	**Date**	**Comments**
0.1	Thomas Loubrieu	DRAFT	07 August 2015	First draft
0.2	Roy Lowry, Simon Jirka, Scott Bainbridge, Shane StClair	DRAFT	10 September 2015	Review, edit
1.0	Thomas Loubrieu	FINAL	14 October 2015	Update based on results presented at the ODIP II workshop in Paris, France
1.1	Helen Glaves	FINAL	22 October 2015	Final edits and sign-off

Contents

Executive summary **4**

1 Methodology **6**

1.1 Impact assessment sessions during 3rd ODIP workshop 6
1.2 Interim Strategic Analysis Report 6
1.3 Impact assessment sessions during 4th ODIP workshop 6

2 General benefits **6**

3 Impact assessments for individual prototypes **7**

3.1 ODIP Prototype 1 ..... **7**
3.1.1 Scope 7
3.1.2 General Benefits 7
3.1.3 Implications for the regional systems 8

3.2 ODIP Prototype 2 ..... **10**

3.2.1 Scope 10
3.2.2 General benefit 11
3.2.3 Implications for regional systems 12

3.3 ODIP Prototype 3 ..... **14**

3.3.1 Scope 14
3.3.2 General benefits 14
3.3.3 Implications at regional level 16

4 Cross-cutting topics **18**

5 Synthesis of implications and proposed action plans **19**

5.1 General review of identified impacts and implementation scenarios 20
5.2 Regional synthesis 23

6 Conclusion **24**

7 Terminology **25**

Executive summary

As outlined in the Description of Work for Work Package 4, “the ODIP workshops and the ODIP prototypes as well as the resulting feedback will provide the ODIP partners with an insight into possible common standards and/or interoperability solutions. However implementing these solutions in practice might have significant implications for the operational systems. Their adoption may require modifications throughout the existing system from the portal to the distributed data providers”.

This document describes the “Final strategic analysis” which has been conducted to identify the possible impacts and potential solutions for the adoption of the prototype interoperability solutions and the associated standards by the regional data infrastructures. This analysis is based upon the “Interim strategic analysis” documented in deliverable *D4.1 Interim strategic analysis report* and the status of the prototype development tasks as reported at the 4th ODIP workshop which took place in Liverpool, UK on 20th - 23rd April 2015.

Appraisals of the three cross-cutting topics which have been addressed during the ODIP workshops are also included in this document. These topics, which are data citation and publication, vocabularies and data ingestion, were identified by the ODIP partners as being of fundamental importance for all of the ODIP development tasks and also of direct relevance for the wider marine community.

Finally this report includes a proposed action plan to address some of the issues that have been identified by the impact analysis. It is also suggested that this plan could potentially be used as part of the road map for those agencies responsible for the development and operation of the regional marine e-infrastructures.

1 Introduction

As outlined in the Description of Work for Work Package 4, “the ODIP workshops and the ODIP prototypes as well as the resulting feedback will provide the ODIP partners with an insight into possible common standards and/or interoperability solutions. However implementing these solutions in practice might have significant implications for the operational systems. Their adoption may require modifications throughout the existing system from the portal to the distributed data providers”.

This document describes the “Final strategic analysis” which has been conducted to identify the possible impacts and their potential solutions. This analysis is based upon the “Interim strategic analysis” documented in deliverable *D4.1 Interim strategic analysis report* and the status of the prototype development tasks as presented at the 4th ODIP workshop which took place in Liverpool, UK on 20th - 23rd April 2015.

This analysis shows the improvements which are expected from the interoperability solutions proposed by each of the prototype development tasks:

- **ODIP 1**: Establishing interoperability between SeaDataNet CDI, US NODC, and IMOS MCP data discovery and access services including making use of brokerage services for interacting with the IODE-ODP and GEOSS portals
- **ODIP 2:** Establishing interoperability between cruise summary reporting (CSR) systems in Europe, the USA and Australia including making use of GeoNetWork for interacting with the POGO portal
- **ODIP 3:** Establishing a Sensor Observation Service (SOS) prototype for selected sensors (OGC Sensor Web Enablement) installed on vessels and in real-time monitoring systems

The potential improvements provided by the implementation of the prototype interoperability solutions are considered in the context of the ODIP objective to “remove barriers hindering the effective sharing of data across scientific domains and international boundaries”. However, implementation of these solutions may also have implications for the existing regional or global systems or even require new replacement systems to be developed. These implications are carefully detailed in the analysis and the specific components or affected infrastructures are identified (e.g. SeaDataNet, Eurofleets, JERICO, POGO, R2R, AODN, GEOSS, Ocean Data Portal, etc.) so that the implementation can be analysed in more detail and scheduled to minimise the impact on the systems and its users.

In addition to the prototype interoperability solutions a number of cross-cutting topics (vocabularies, data publication and citation, data ingestion) that are relevant for all of the regional data infrastructures have also been addressed during the workshops. The current state of the art in these areas and the impact of implementing the associated standards and best practice is reviewed in this deliverable.

The final section of this deliverable outlines possible strategic implementation scenarios which are based on a synthesis of the identified impacts for the regional data infrastructures. These scenarios also include a draft task plan with suggested contributions from existing or planned projects and frameworks.

2 Methodology

2.1 Impact assessment sessions during 3rd ODIP workshop

During the 3rd ODIP workshop held on 5th – 8th August in Townsville, Australia, brainstorming sessions were organised to get input from partners on the perceived impacts for the regional data infrastructures of implementing the prototype interoperability solutions and/or the agreed standards.

For each prototype solution, the partners were asked to identify the general benefits for users, the regional systems and data centres. They were also asked to identify potential implications, constraints or costs related to the implementation of the prototype functionality in the operational systems.

Some general benefits have been identified for each prototype that are documented in section 3 below. The details of the specific benefits and impacts are detailed in the following sections and are documented as reported during the workshop sessions.

2.2 Interim Strategic Analysis Report

As a result of the impact assessment sessions conducted during the 3rd ODIP workshop a detailed interim strategic analysis report (deliverable *D4.1 Interim strategic analysis report*) was produced which was reviewed by the project partners. Strategic implementation scenarios have been formulated based on a synthesis of the inventory of impacts that have been compiled during the impact assessment sessions that took place during the ODIP workshops.

2.3 Impact assessment sessions during 4th ODIP workshop

During the 4th ODIP workshop, which was held in Liverpool, UK on the 20 – 23 April 2015, the impacts of the individual prototype interoperability solutions were reviewed based on the progress that had been made since the previous workshop. In addition, the strategic implementation scenarios proposed in the interim strategic analysis report (D4.1) were also reviewed and subsequently completed.

This document is the '*Final Strategic Analysis Report'* which is an update of the deliverable *D4.1 Interim Strategic Analysis Report* and includes further input from the impact assessment sessions organized during the 4th ODIP workshop.

3 General benefits

A number of general benefits have been identified which apply to all of the ODIP prototype interoperability solution:

- Development of common standards and best practices utilising the larger pool of developers and resources available across the ODIP community significantly saves on time and resources by reducing the unnecessary duplication of effort within the individual data systems.
- Meeting of the community of practice in ODIP workshops helps to share knowledge and best practices on tools, systems and standards (DOI, ISO19139 etc.). This also potentially allows wider access to emerging technologies (big data, web technologies, SWE etc.) that are otherwise unavailable to smaller or isolated groups.

4 Impact assessments for individual prototypes

4.1 ODIP Prototype 1

4.1.1 Scope

ODIP prototype development task 1 (ODIP1) aims to “Establish interoperability between SeaDataNet CDI, US NODC, and IMOS Marine Community Profile data discovery and access services including making use of brokerage services for interacting with the IODE-ODP and GEOSS portals ” (see deliverable *D3.1 Definition of prototypes 1*).

The GEO-DAB brokering service ingests OAI-PMH metadata services at the regional centre level. The brokering service then translates the metadata records into the Generic Brokerage Reference Schema which can then be harvested by the IODE-ODP and GEOSS global systems thus facilitating delivery of metadata from the regional data systems in the global services.

As a result of the activities in the ODIP 1 prototype development task the number of observations which have been harvested from European, US and Australian marine data centres and exposed via the IODE-Ocean Data Portal and GEOSS portal has increased significantly.

4.1.2 General Benefits

4.1.2.1 For users:

As a result of implementing the ODIP 1 prototype interoperability solution, marine observations from regional data systems will be exposed in global portals:

- IODE Ocean Data Portal: http://www.oceandataportal.net/portal/portal/odp-theme/data/relatedprojects
- GEOSS portal: http://www.geoportal.org/web/guest/geo_search_overview

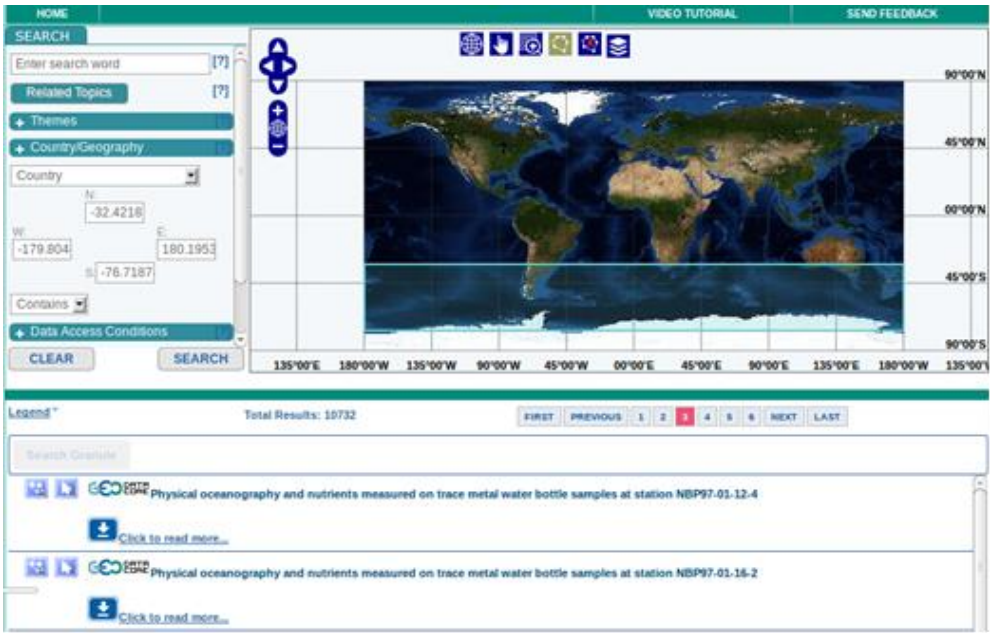

Figure 1 Marine data for the Antarctic in the GEOSS portal

These global portals therefore become a single point of access for a range of data available in the regional data systems. They also enable multi-parameter aggregations of datasets over a wider geographical area and with this enhanced “data fusion” facility also make a significant contribution to dealing with larger scale complex problems.

- **Gateways to other communities**

The federation of access to datasets exposed in the regional systems via the global portals using standardized back-end services (CSW/ISO19139, OAI-PMH, OpenSearch) will potentially promote the use of the data outside the scientific community particularly if new delivery methods, for example via apps for mobile devices, are made available.

4.1.2.2 For regional systems

The following benefits for the regional data systems from the ODIP 1 prototype development task have been identified:

- **A tool for optimizing resources**

The global portals federating data from the regional systems will allow the localised e-infrastructures to identify other data in their own waters they had not been aware of previously. ODIP 1 is generally increasing data sharing and re-use of observations, and as a result is reducing the use of resources required for re-collection of data.

- **A marketing tool to expose better the datasets**

The visibility of data available via the regional systems will be increased as a result of prototype development task ODIP1. This increased exposure will in turn raise the profile of the data providers as individuals and/or organizations and may also lead to subsequent funding opportunities.

4.1.3 Implications for the regional systems

4.1.3.1 General implications

- **Deprecation management**

To enable proper sharing of the metadata records for datasets, the definition and implementation of common rules will be required. It is generally agreed that once published (and potentially used) datasets cannot be deleted from the registries. Obsolete datasets should be deprecated or superseded. This rule needs to be detailed and implemented in each regional system.

- **Granularity**

The granularity of datasets available in the data systems is also an issue which needs to be addressed. If this problem is not resolved, the global portals will be populated with heterogeneous records describing anything from single observations to collections of data e.g. records will range from a single CTD profile to the whole collection of ARGO observations. There is no obvious definition for the granularity of the datasets, as, depending on the science area, the volume or “value” of single observation is different (e.g. one ocean physics water column profile differs significantly in extent and size to that of a single seismic profile). The global portals can resolve this granularity issue by carefully targeting records they are harvesting from the regional systems. The catalogue being harvested is also

required to properly manage the granularity of their datasets (i.e. observations, collections etc.). This issue was discussed extensively during the 2nd ODIP workshop and documented in the deliverable *D2.4 Minutes and actions of ODIP workshop 2*.

- **Data Access (view, download)**

The interoperability of datasets published in the global portals could be improved so that the global systems can provide seamless access to the datasets through the regional centres. Data access is currently provided by a redirection from the regional data services or data provider to the location where the data can be downloaded by the user. Ideally the user should get seamless access to the datasets via the global portal. However, this issue is not only technical but also strategic, for example, the data providers and regional systems require recognition in the form of a visible presence (e.g. logo) associated with the resources that they are providing.

An additional problem is also the management of authentication for restricted datasets in an environment where different user directories and management policies are implemented. The federation of identity (openID, Shibboleth) will help but requires further development at the regional system level.

- **From syntactic to semantic interoperability (common vocabularies)**

The ODP and GEOSS portals are currently only aggregating regional systems at a syntactic level. This essentially means that the datasets are identified by specific attributes (e.g. keywords) by the aggregating systems where these have been defined with a semantic role at regional system level (e.g. type of acquisition device). The ability to manage the semantics at the IODE and GEOSS portal level will require 1) additional information on the format of metadata records at the regional system level 2) a more detailed mapping at brokering component level.

4.1.3.2 Impacts at the European level

In order to maintain the delivery of data from the European SeaDataNet marine data service to the global portals, the CDI server operated by MARIS must continue to provide CSW, OAI-PMH and OpenSearch interfaces for the collections of observations (‘super-CDIs’). This service needs to be maintained and continuously qualified in order to populate the GEOSS and Ocean Data Portal.

The GEO-DAB brokering service is a key component of prototype ODIP1. It is hosted and operated at ESSI-LAB (http://www.essi-lab.eu) in Italy. The ESSI team is also responsible for mapping the different regional system contributions (as back-end services) to the homogeneous services used in the GEOSS portal.

For the purposes of maintaining an operational architecture, the GEO-DAB brokering component must be sustainable in terms of both funding and the facilities necessary for its hosting and operations. The ESSI-LAB team is currently focused on research and development, two options for the long-term sustainability of the broker service should therefore be considered:

- **option a**: update the scope and skills of ESSI-LAB allowing them to continue to host the GEO-DAB service
- **option b**: migration of the GEO-DAB brokering service to a suitable host organisation already responsible for maintaining operational systems. ESSI-LAB could then still be responsible for the configuration of the GEO-DAB system and mapping of the information flows.

4.1.3.3 Impacts at USA level

At the US NODC, OpenSearch and CSW along with a REST API service have been enabled and are publicly available at http://data.nodc.noaa.gov/geoportal through the use of the open source Geoportal Server. In addition, an OAI-PMH service has been established, with a base URL of http://data.nodc.noaa.gov/cgi-bin/oai-pmh. As an example, http://data.nodc.noaa.gov/cgi-bin/oai-pmh?verb=ListIdentifiers&metadataPrefix=oai_dc will list all of the identifiers associated with more than 20,000 collections in the NODC archive (the resulting list takes several minutes to display in a web browser). This new service can also now be ingested by the GEO-DAB brokering service to allow the ODP and GEOSS global systems to harvest and expose metadata for datasets in the US-NODC regional system.

Based on experiences gained through ODIP interactions, the US NODC has begun testing a SKOS service for exposing their vocabularies to the wider community via an interoperable service. However, it is currently only available internally at NODC but this service will significantly enhance semantic interoperability of the US-NODC dataset descriptions once it becomes fully operational and available to external users.

4.1.3.4 Impacts at the Australian level

The Australian Ocean Data Network (AODN) provides standard interfaces for dataset discovery using the CSW, OpenSearch and OAI-PMH standards via a GeoNetwork application (see Table 1 below). The OAI-PMH interface can be used by GEO-DAB broker. In order to achieve semantic interoperability it will be necessary to fully map the Marine Community Profile (MCP) vocabularies on to those used by SeaDataNet in Europe.

OAI-PMH	https://catalogue-123.aodn.org.au/geonetwork/srv/eng/oaipmh?verb=Identify
CSW	https://catalogue-123.aodn.org.au/geonetwork/srv/eng/csw?REQUEST=GetCapabilities&service=CSW
Opensearch	https://catalogue-123.aodn.org.au/geonetwork/srv/eng/portal.opensearch

Table 1: standard interfaces provided by AODN for dataset discovery

4.2 ODIP Prototype 2

4.2.1 Scope

The ODIP 2 prototype development task aims to establish interoperability between cruise summary reporting (CSR) systems in Europe, the USA and Australia (see deliverable *D3.1 Definition of prototypes 1* for details). This includes the use of GeoNetwork at the regional system level for interacting with the Partnership for Observation of the Global Oceans (POGO) portal available at: http://seadata.bsh.de/csr/retrieve/pogo_index.html

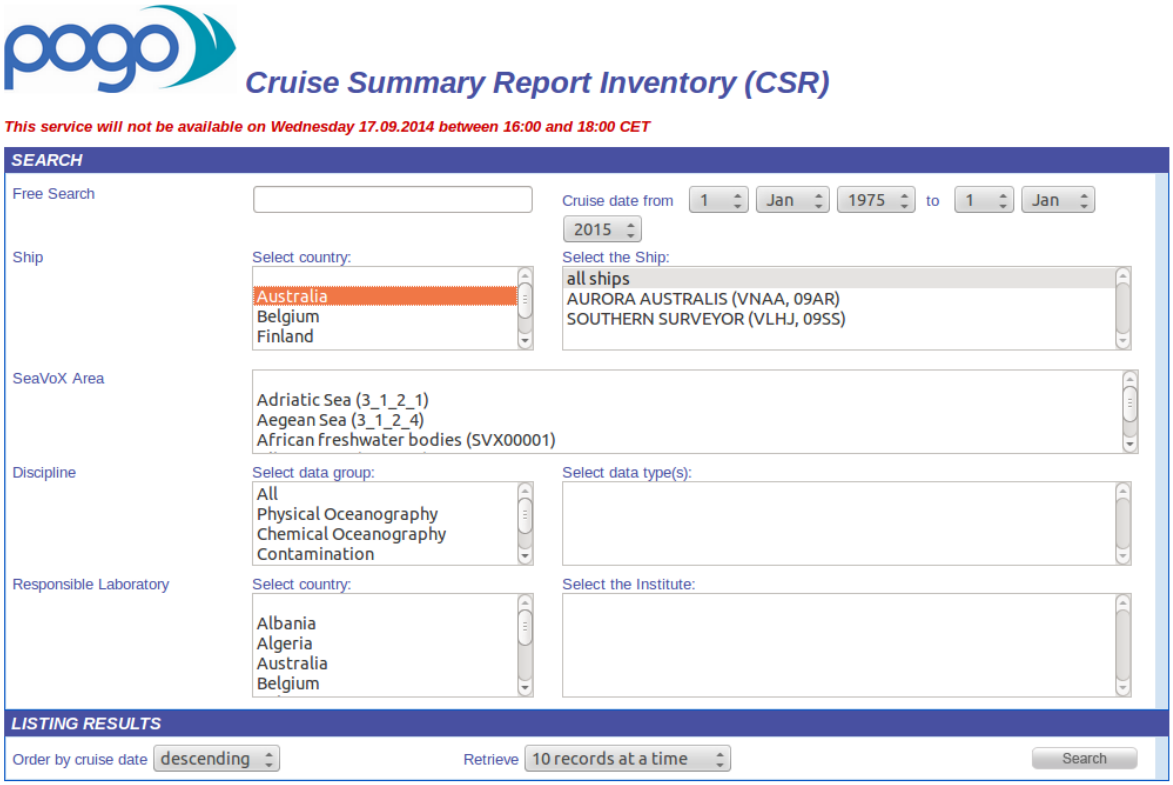

Figure 2 POGO portal operated by BSH in Germany

4.2.2 General benefit

4.2.2.1 For users

As a result of the development work done for the ODIP 2 prototype development task, details of all scientific cruises operated by organisations in the USA, Australian and Europe and exposed in the regional systems, will be also available in the POGO portal.

Until very recently only European cruises were available in a single portal. Moving towards delivery of the CSRs available in the regional systems through an international system is therefore not only a benefit at a global level but also represents an enhancement for organisations in Australia and the USA that do not currently have single portals showing their national scientific cruise catalogues. The development activities in ODIP2 have already resulted in an increase in the number of US cruises registered in the POGO portal and this number is expected to rise significantly over the coming months.

In addition, using the same standards (e.g. OGC/CSW and ISO19139) for cruise descriptions in the back-end services of the regional systems also makes them interoperable. As a consequence, the aggregated information will potentially also be available to users for other applications either within or outside the scientific community, and using different technologies (e.g. web, mobile devices etc.)

The POGO portal gives users a synthesized view of who is observing the ocean, where and what is being observed, and how it is observed. The cruise summary reports from the regional systems aggregated in the POGO portal will provide homogeneous information on:

1 organizations responsible for the cruises

2 location of the observations (bounding box and track of the vessel)

3 disciplines studied on-board (geology, ocean physics, fisheries, etc.)

4 devices used for observations (CTD, multi-beam, etc.)

This up to date information is especially valuable for conducting cross-calibration of equipment, for example, for checking ARGO delayed mode processing using ARGO profiles versus CTD profiles.

The information available in the POGO portal on past cruises is also highly valuable for the purposes of organizing and planning new observation programs and cruises, especially in those marine basins with international boundaries (e.g. Antarctic, Arctic, North Atlantic North etc.).

4.2.2.2 For regional systems

The following benefits of the ODIP2 prototype have been identified for the regional systems:

- Improved knowledge of available observations/data

At the data management level, the aggregation of cruise descriptions assists in following-up datasets with the scientists to ensure that the datasets are deposited with a repository and ensure the long-term preservation of the observations. This compilation of datasets also helps where access to some information is restricted through moratorium.

- Improved quality of Cruise Summary Reports (CSR)

The use of controlled vocabularies and constraints on metadata profile descriptions improves the quality of the information.

Generation of standardised metadata records has been made easier through the use of the MIKADO tool which has been developed by IFREMER and co-funded by the EU SeaDataNet project. The MIKADO tool can be used in batch mode or through a manual form interface to improve accuracy and efficiency of metadata generation.

- Shared open-source framework for metadata management

Use of the GeoNetwork open-source software facilitates dissemination of the metadata records at the regional level. As a result of the ODIP 2 prototype development task, the GeoNetwork application for the management of cruise summary reports has been updated. Work has been carried out by IFREMER (Europe) and CSIRO/ UTAS (Australia) to modify GeoNetwork to allow it to use the CSR ISO19139 profile and reference vocabularies in metadata records, and also enable the harvesting of it .

4.2.3 Implications for regional systems

4.2.3.1 General implications

To provide seamless access to datasets acquired during cruises via the cruise summary report requires the use of the linked data concept. To support this approach, under-way observations and details of other experiments deployed during the cruise need to be

recorded in the CSR. The current status of providing the necessary links is different for each regional system and will be described in the following sections.

One modification to existing systems that would potentially provide greater scope for recording and discovery of cruise information on a global scale would be for the POGO portal to include ships with a length of less than 60 metres which are currently excluded from the database.

4.2.3.2 Impacts at the European level

The implications of the ODIP 2 prototype interoperability solution at the European level are:

- European reference directories and vocabularies have been updated to include to entries for non-European resources. For example, details of 173 US organizations and 44 other non-European organizations have been added to the European Directory of Marine Organisations (EDMO). In addition the NVS L22 vocabulary for the model of observation devices used by European marine data services has also increased by 30% in the period from December 2013 to April 2015. These directories will also continue to expand as more CSR records are created.
- Bundesamt für Seeschifffahrt und Hydrographie (BSH), in Germany hosts and operates the POGO portal but is not a partner in the current ODIP project. In order to facilitate the upgrade of the POGO portal and fulfil the requirements for the ODIP 2 prototype interoperability solution, including an overhaul of the current interface, BSH will be included as a project partner in the follow-on H2020 ODIP II project.
- A number of corrections have been identified that should be applied to the CSR ISO19139 schema currently defined and used in SeaDataNet:
    - GML3.2.1 should be used in ISO19115-2 instead of GML3.2.0. GML3.2.0 is an un-official version of GML and is therefore not supported by OGC.
    - there are currently only a limited number of contact details included in the CSR records and where these do exist the roles are not adequately managed. Roles should be added to identify the full science party using personal persistent identifiers (e.g. ORCID, ResearcherID, etc.)
    - to improve management and use of controlled vocabularies, “gmx:Anchor” tags and “xlink” attributes should be utilised as much as possible. The methods currently used for referencing vocabularies are very heterogeneous.
    - the ISO19155-1 metadata standard is currently in use but the is likely to be an upgraded to ISO19115-2 in the future which will assist with the addition of digital object identifiers (DOIs) to CSR records.

These upgrades are not equally important but all of them will necessitate an upgrade of the supporting software (MIKADO, GeoNetwork etc.). In addition a migration of existing CSR records managed at BSH would be required.

4.2.3.3 Impacts at Australian level

The Marine National (MNF) operated by CSIRO is the only node connected to the ODIP 2 prototype solution in Australia. As a result only cruise summary reports for the vessels RV Southern Surveyor and the RV Investigator will be made available for delivery. The Australian regional node should be extended to cruises operated by other institutions e.g. AIMS etc. in the future. In addition, internal mapping to the EDMO directory and the other relevant vocabularies needs to be carried out.

4.2.3.4 Impacts at US level

Due to some mandatory information not being included in current US cruise summary reports (e.g. contact information for full science party including roles), the R2R facility provides the CSR output to POGO as a supplementary activity. Although the same standards are used, the duplicate flow of information incurs additional operational and maintenance costs.

Organizations providing cruise summary reports in the USA will also need to work on mappings to EDMO and the reference vocabularies.

4.3 ODIP Prototype 3

4.3.1 Scope

The ODIP 3 prototype development task aims to establish a Sensor Observation Service (SOS) for selected web-enabled sensors (SWE) installed on vessels and in real-time monitoring systems (see deliverable *D3.1 Definition of prototypes 1* for further details).

To achieve this objective selected ODIP partners are evaluating and implementing various solutions for providing sensor web enablement (SWE) services for accessing descriptions of the systems at sea and observations e.g. 52^{o} North, ncSOS, sensorCloud, Oceanotron, etc.

These solutions and their interoperability have been discussed during the ODIP workshops with two categories of data dissemination architecture having been considered:

- **atomic data services** which provide datasets at small granularity and with limited capacity for discovery. These systems rely heavily on a "central" catalogue which supports the discovery services: ncSOS and THREDDS Data Server based in-situ observation dissemination as used by IOOS (USA), IMOS (Australia), and IMEDEA (Spain). This approach is also used by SeaDataNet in Europe with the Common Data Index (CDI) central catalogue and Download Manager data access service.
- **collection data services** which provide datasets in a ‘collection’ of observations and provide discovery functions (for example through OGC Filter encoding) directly on these datasets. Other services such as Oceanotron, sensorCloud, or to some extent general GIS servers, providing web feature services (WFS) operate in a similar manner.

The atomic data services are generally oriented to heterogeneous datasets because each one is potentially very important in terms of the information it provides (e.g. HF radar, large moorings, geo-seismic profiles, etc.). The collection data services are more suitable for aggregations of homogeneous observations (i.e. of the same shape or feature type), that are usually small (e.g. simple time series, profiles). This approach is particularly efficient when the collection is an aggregation of observations for a target user group in response to their specific requirements (e.g. quality level).

4.3.2 General benefits

4.3.2.1 Users

The implementation of the ODIP 3 prototype by the marine data infrastructures will allow users to access a greater volume of observations from sensors with a reduced latency.

For users of the data, sensor web enablement will be a step towards open data, with observations potentially made available in near real-time and even remotely from platforms at sea, as demonstrated by the European SensOcean (http://www.senseocean.eu/) project and the British Oceanographic Data Centre (BODC).

Users will also benefit from ad hoc web portals which can be quickly developed on top of the combined standard data flows, especially when accessible with RESTful API in JSON format. This approach has been demonstrated at CSIRO (see Logan river flood portal at http://loganvis.meteor.com/) and also by the 52°North Sensor Web Client REST API, which aggregates multiple SOS servers as a REST proxy

4.3.2.2 Regional centre benefits

The users of the ODIP3 prototype solution are both observation providers and dataset consumers. The data providers benefit from secure and efficient repositories for their observations. By considering their data providers as users, data centres are able to obtain more datasets.

4.3.2.3 The benefits for observation providers

For observation providers the standardisation of sensors and observation descriptions combined with the advent of tools to edit these descriptions has made the submission of observations to data centres easier. Once the SWE editing and ingestion tools are fully operational (see developments done at CNR: RITMARE and on-going at IFREMER: Sensor Nanny), any justification used by providers for not submitting their datasets to the repositories will be invalidated.

Early submission of datasets to the repository significantly enhances the quality of the data especially if this process if automated with simultaneous control of the data formats and referencing of controlled vocabularies. In addition, the management of duplicate data, which has always been an issue in distributed data management architectures, may be partly solved by tagging platforms, sensors and observations with universally unique identifiers (UUID) at an early stage, before the data becomes distributed.

In the future, the system used to make SWE records available by observatory operators may function in a similar way to social networks where they can publish observations and be informed about the progress of their peers. Moreover the SWE repositories will also facilitate the publication and submission of data to infrastructure catalogues such as GBIF, SeaDataNet etc. This publication can be automated from the description of sensors and observations in the SWE repositories.

4.3.2.4 The benefits for data centres

The ODIP3 prototype development task relies on state-of-the-art software which still requires further development and/or maturation. In this respect, the regional data infrastructures benefit from the shared development activity with a reduced requirement for resources, as well as independent validation of the outcomes by the ODIP partners.

The standardized flow of information for observations and sensor descriptions helps to aggregate datasets from heterogeneous platform types and also potentially reduces the cost and delays associated with the development of web user interfaces especially where RESTful API and JSON encoding is already available.

At the European level, the availability of SWE compliant services is mandatory to comply with the INSPIRE directive for the dissemination of observation data.

The standardization of observations and sensor descriptions is reducing the latency of the data centres to collect them. This effect is also very important for invoking a prompt reaction when a problem arises with a platform out at sea.

The information available in the sensorML records can also be used for provenance information which in turn supports the certification of the laboratories, for example, by tracking the calibration of equipment.

The ODIP 3 prototype development task has been identified as a key component for observation ingestion systems.

4.3.3 Implications at regional level

4.3.3.1 General implications

- **Implementing SWE systems at sea**

As seen above, the implementation of SWE compliant systems at sea, for example on-board research vessels, is extremely useful for the efficient delivery of observational data to the data centres. However, the cost of upgrading on-board vessel systems to make them SWE compliant as well as the necessary bandwidth required to transmit the information in real-time to the shore-based repositories also needs to be taken into consideration.

This process also requires the involvement of the sensor manufacturers and some on-going initiatives, e.g. SensOcean, Eurofleets in Europe, have tried to develop this type of engagement with the makers of these systems. However, science is a not a critical customer for the manufacturers. The oil and gas industry and the navy are the main customers for these systems and therefore have the greatest influence with the equipment manufacturers

In addition emerging standards such as the OGC SensorThings API should be considered in this context. However, the REST- and JSON-bindings for this standard would be more useful for oceanography if they were published as SOS and O&M enhancements.

- **Standardization of modern web SWE APIs and formats**

A number of initiatives have been responsible for the development of a diverse range of RESTful API and JSON encodings (e.g. 52^{o}North, CSIRO, CSIC, IFREMER, RITMARE) due to it having become more commonly used in web development. However, none of these solutions are currently aligned to an OGC standard and it is therefore necessary to merge these initiatives into a single standard. Initial discussions, indicated that the RESTful API proposed by 52^{o}North is a good candidate standard, while some standardization of sensorML and O&M implementation in JSON is also currently being considered by OGC (using JSON-LD for linked data and geoJSON for geospatial information).

- **Scalability and performance of SWE systems**

The regional data centres manage millions of observations and it is generally acknowledged that a scalable SOS server is required. In addition, the flexibility of the back-office repository is also a key factor. Unlike RDBMS, the current no-SQL databases (e.g. mongoDB, couchBase), provide the data repository with the flexibility necessary to compliment the versatility of SWE standards.

Under a project run in Australia, version 3.6 of the 52°North SOS was modified to deliver faster responses to large datasets and queries. Some of this code has been incorporated into version 4 to deal with issues of performance under heavy loads but as version 4.x uses a different architecture it is not certain how the performance gains developed under version 3.6 will be realized under the new version 4.

However, 52°North SOS version 4 provides full support for SOS v2 with partial support for SOS v1 and so all development should now be done using v4.x of the 52°North product. This will ensure that all systems developed will be fully compliant with SOS v2 and this in turn means that systems should now be developed around SOS v2 and not the older SOS v1.

4.3.3.2 Implications at the European level

As already discussed, in Europe the use of SWE standards for sensor and observation descriptions requires new or updated vocabularies. At a technical level, these vocabulary services need to comply with standards such as SKOS and SPARQL as used by the RITMARE web forms application. (However, it should be noted that the SeaDataNet vocabulary services already comply with those standards).

4.3.3.3 Implications at the Australian level

It has been noted that for Australian SWE applications there is a requirement for the quality control information to be stored together with the observations. In the OGC Observations and Measurements (O&M) conceptual model a “result Quality” object is designed to host this information. However, further work is required to determine how it could be applied for marine observations.

The importance of sharing and exposing a larger set of examples of valid sensorML and O&M to help regional data centres to design their own SWE templates has been stressed. It would be of particular interest to push some new marine examples to the SensorML repository at: http://www.sensorml.com/sensorML-2.0/examples/

In addition all SOS developments should now be made to the version 2 standard using version4.x of the 52°North SOS product but with the caveat that full performance testing of this version has yet to be completed b would be useful.

4.3.3.4 Implications at the US level

US ocean observing centres (especially IOOS and affiliated regional associations) serve sensor data via OGC SOS. In order to facilitate and enrich SOS interoperability between regional data assembly centres, IOOS has funded and coordinated the creation of a domain specific SOS application profile including standards and response templates which dictate metadata requirements, vocabularies, XML formats, unit systems, and other aspects of SOS functionality. The application profile dictates that SOS implementations use CF-1.6 discrete sampling geometries [1], standard observable property vocabularies (CF standard names [2] and IOOS parameters [3]), and the concept of hierarchical observing assets (because networks contain many platforms and stations contain many sensing instruments). The current version of the IOOS application profile is Milestone 1, which describes SOS 1.0 responses. These guidelines and templates are available at the IOOS SOS Guidelines GitHub site [4].

IOOS also sponsors the development of two SOS implementations capable of serving the custom application profile: i52n-sos [5] and ncSOS [6]. The i52n-sos uses 52°North SOS [7] as its core (currently the 4.x version) and adds custom features to conform to the IOOS SOS application profile. By leveraging the 52°North SOS core, i52n-sos includes many advanced features including SOS 2.0 core and transactional operations, advanced bindings (XML, JSON, REST), and support for multiple database platforms. In addition, i52n-sos adds custom encodings for IOOS profile XML and NetCDF responses, test data generation, and other advanced features. The i52n-sos developer contributes bug fixes and generally applicable features to the core 52°North SOS codebase (for example, the i52n-sos NetCDF encoder was recently adapted and merged into the main 52°North SOS codebase).The i52n-sos stores sensor data in a database (usually Postgres/PostGIS) and is well suited for serving active sensor streams.

In addition, the European NeXOS project provided some funding to integrate some of the work on NetCDF output and the hierarchical procedures into the 52°North SOS v4

The ncSOS integrates with UNIDATA's THREDDS Data Server [8] as a plug-in and serves sensor data stored in CF-1.6 discrete sampling geometry NetCDF files. The ncSOS implements the core SOS 1.0 operations and serves IOOS Milestone 1.0 compliant responses. Data served by ncSOS is stored as NetCDF files rather than in a database. The ncSOS is well suited for serving large time series aggregations of sensor data. However, due to an early adoption, most of the developments and resources are based on SOS v1 and an upgrade to SOS v2 would be valuable for global interoperability with other regional components.

[1] http://cfconventions.org/Data/cf-conventions/cf-conventions-1.6/build/cf-conventions.html#discrete-sampling-geometries

[2] http://cfconventions.org/standard-names.html

[3] http://mmisw.org/ont/ioos/parameter

[4] http://ioos.github.io/sos-guidelines

[5] http://ioos.github.io/i52n-sos

[6] https://github.com/asascience-open/ncSOS

[7] http://52north.org/communities/sensorweb/sos/

[8] http://www.unidata.ucar.edu/software/thredds/current/tds/

5 Cross-cutting topics

As well as the prototype development tasks the ODIP project has also addressed a small number of cross-cutting topics which are relevant for all of the prototype development tasks. These topics are:

- Vocabularies
- Data publication and citation
- Data ingestion

These three topics have been discussed during workshops and the significant benefits and impacts of these activities for partners and the prototype development tasks have been noted. These topics will also be taken forward for further consideration in ODIP II.

Vocabularies are vital for managing reference terms in any domain and this point has been stressed in the impact assessment analysis for each prototype in earlier sections of this report. Some challenges for the management of vocabularies have been identified by a number of authorities in the field, for example, implementation of virtual collaborative spaces to manage vocabularies or use of shared governance for vocabulary lists depending on the expertise needed (e.g. cheBI for chemical entities of biological interest at http://www.ebi.ac.uk/chebi/).

Unique identifiers for people (i.e. non-ambiguous identification of researchers, observatory operators etc.) has also been recognised as a key issue for all partners. Although other systems (e.g Google Scholar, Research Gate, Ocean Expert etc.) or standards (e.g. FOAF http://xmlns.com/foaf/spec/) for identifying researchers have been considered by the ODIP partners, the ORCID system is currently a preferred option for many organisations and individual researchers. Further work will be done to develop a white paper which makes an assessment of the potential options for assigning unique identifiers for people in the marine domain.

Data publication and citation has also been discussed during dedicated sessions at the ODIP workshops. Digital Object Identifiers (DOI) are now used to enable dataset citation and some internationally agreed guidelines have been written on the subject (http://www.iode.org/mg64). However, the approach for allocating persistent identifiers to continuously updated datasets (e.g. ARGO floats), is not as well advanced. A mechanism for the citation of dynamic data has been developed by the RDA Data Citation working group (https://rd-alliance.org/group/data-citation-wg.html) and it has been proposed that ODIP partners adopt this solution for selected applications e.g. implementation by IFREMER for ARGO floats. Further work will also be done by selected members of the ODIP consortium in partnership with the RDA Data Citation working group to develop some specific implementation use cases in order to test and evaluate the solutions that are currently being proposed.

The topic of **data ingestion systems** has also been addressed during ODIP workshops and partners were asked to provide their own experiences in this area. As a result several relevant examples have been highlighted including Send2NODC in the USA and RITMARE in Europe. A number of challenges associated with data ingestion systems, for example, contacting new data providers such as those companies doing environmental observations for renewable energy installations or oil industry applications, have been highlighted. A number of potential implications of these solutions have also been identified e.g. security of the systems used for the delivery of data.

6 Synthesis of implications and proposed action plans

A synthesis of the implications of adopting the ODIP prototype interoperability solutions and/or the common standards has given rise to a number of possible actions to address these challenges. These actions, which are described below, were also discussed by the project partners during the 4th ODIP workshop.

6.1 General review of identified impacts and implementation scenarios

Rules on metadata management (deprecation, data access service encodings, vocabularies, granularity), and moving towards a global marine metadata profile.

Generally speaking, although metadata standards for datasets are largely based on the ISO19139 standard and sensor web enablement (SWE) is used for observations and systems, there are subtle differences between the profiles that are used in Europe, the USA and Australia. This kind of divergence is not unusual across the different communities and the individual standards have the flexibility to allow it. However, the global marine community should make a particular effort to promote the use of the mature ISO19139 standard which is used for both datasets and cruises summary reports (CSR). In this context, guidelines or profile specifications for using the ISO19139 standard should be provided in the following areas:

- Obsolescence management: how the metadata of a deprecated or superseded object should be encoded.
- Data access service encoding: users increasingly expect service oriented discovery metadata enabling data visualization, download or transformation. These services are often encoded in the metadata but they need to be more homogeneous so that any front-end portal can make use of this information to embed datasets in a geo-viewer system or a 'shopping cart' download service.
- Use of common vocabularies: although unification of vocabulary references has been agreed this still needs to be implemented in the regional metadata profiles.
- Granularity of metadata: this is still an issue which needs a set of rules to be defined and also requires user-friendly portals with discovery capabilities for potentially heterogeneous datasets that include anything from a single CTD profile to a complete climatology model for the entire planet.

These issues will be discussed further as part of the ODIPII project. The marine community will also take advantage of the discussions in the relevant Research Data Alliance interest and working groups e.g. Marine Data Harmonisation and Vocabulary Services interest groups, and the Data Description Registry Interoperability work group.

The IODE Ocean Data Standards process (http://www.oceandatastandards.org/) has been recognised as the most suitable mechanism for worldwide publication and dissemination of a unified marine metadata profile.

Federations of identity for restricted data access

To develop a global infrastructure for marine data discovery and access which includes the Ocean Data Portal and GEOSS portal as proposed in the ODIP 1 prototype development task, requires the issue of access (visualization or download) to restricted datasets, or at least the traceability of the usage of datasets, to be addressed. This will require a universal mechanism for authenticating users.

As well as the web industry standards and other identity providers (Google, Facebook, Yahoo, etc.), regional research communities are also organizing the provision of an infrastructure for user authentication such as the eduGAIN service provided by GEANT in Europe (http://services.geant.net/edugain/Pages/Home.aspx).

The regional marine communities are also building user directories which include service desks and in some cases customer relationship management facilities e.g. Marine-ID (http://www.marine-id.org/) in Europe. The SeaDataNet project is currently working on this topic and it will also be addressed in ODIP II.

Adoption of Sensor Web Enablement (SWE) standards in the marine community

The flexibility and fitness for purpose of the sensor web enablement (SWE) standards such as the Observation & Measurements and SensorML conceptual models for the marine community has been proven during the ODIP project. However there are some technological issues which are delaying its wider adoption in the marine community.

By undertaking some strategic actions, the end-to-end flow of data from the observing systems to the end-user portals can leverage the usage of these standards and, by doing so, streamline and improve the quality the data management processes (see section 4.3.2 above).

Sensor web enablement (SWE) at sea

The first step to the wider adoption of sensor web enablement (SWE) is to raise awareness of the relevant standards among the hardware manufacturers, software providers and those responsible for integrating them into the systems on board vessels or automated platforms. There is already collaboration for the purposes of standardisation within these communities (e.g. IEEE); and data managers should also seek to connect with these communities so that standards are also shared from the hardware to the datasets.

Two existing European projects are currently carrying out development activities related to promoting the wider adoption of SWE. The Eurofleets project is developing software for research vessels while NeXOS is creating integrated multi-functional sensor systems, and also contributing to the new generation of sensor web enablement (SWE) standards with the Open Geospatial Consortium (OGC). These two projects provide potential opportunities for collaboration to address these issues within the European marine community but there is also a need to identify similar opportunities to connect with these communities in the USA and Australia.

A final consideration is that the large commercial customers, such as the oil and gas companies or the military, have a greater influence on the equipment manufacturers than the marine research community. ODIP II should therefore seek collaborative opportunities with relevant industrial and military partners in order to strengthen a dialogue with the manufacturers of the sensors etc.

Sensor web enablement (SWE) on the modern web

Sensor Web Enablement implementations have been defined with modern technologies (XML). Currently efficient and reactive web development projects are often based on lighter technologies (RESTful/JSON interfaces, no-SQL repositories etc.) which ease development and improve the cost-effectiveness for the creation of web applications. These technologies also support better scalability for high performance applications.

Although some implementation APIs are provided (SensorCloud, 52°North Sensor Web Client REST-API), there is no agreed standard on these modern web interface definitions. The prototype cloud observation systems e.g. sensorCloud, sensorNanny are also promising technologies but these have not yet been widely adopted by marine data managers. Further

development and maturation of these technologies will be done in ODIP II and other new related projects.

In addition, OGC should consider working groups for the implementation of other standards beyond the XML cornerstone (e.g. similar to the OGC SensorThings API but with a broader focus beyond the pure internet of things applications; a good approach would be to work on REST/JSON enhancements of the SOS and O&M standards).

Vocabularies and reference directory management

Vocabularies and reference directories such as the European Directory for Marine Organization (EDMO) are fundamental components of a global e-infrastructure for marine data management. The mapping of regional data management system onto these reference services requires both expertise and significant effort. These mappings have been started by partners in the ODIP project and this activity will continue in ODIPII.

The ODIP project has shown that vocabulary management can be distributed depending on the field of expertise (e.g. ChEBI for chemical entities of biological interest, WoRMS for world register of marine species). However to provide a cross-cutting marine thesauri (e.g. devices, marine observation parameters) whilst acting as proxy for providing a proper management process (governance) and interfaces on thesauri managed elsewhere, necessitates central components such as those provided by BODC for vocabularies (NVS) and MARIS for marine organisations (EDMO).

The maintenance of the NVS and EMDO services is supported by SeaDataNet and carried out by BODC and MARIS respectively. However, in order to establish sustainable state-of-the-art global services, further investment of resources will be required. In particular, this will allow improvement of the online collaborative features of these systems and open up the management of thesauri to dedicated teams of experts throughout the participating regions. This will ensure a proper governance framework for the vocabularies is maintained whilst reducing the work required for the internal maintenance of the service. ODIPII will contribute to the definition and governance rules for such a service.

Cruise summary reports

Cruise summary report integration is highly heterogeneous depending on the regional system (see 6.2). However some common actions are required to improve the availability of scientific cruise descriptions in a global e-infrastructure.

The POGO cruise database has been defined as the federating component for cruise summary reports (CSR). It is operated by BSH as part of the of the consolidated SeaDataNet activities for cruises metadata management at the European level. However, one significant issue with the POGO database is that it does not currently manage cruise information for non-ocean-going (length <60m) vessels. This issue will be addressed by ODIPII in consultation with the POGO consortium. BSH is also a partner in the ODIP II project and will upgrade the POGO portal as part of their contribution to the project.

Finally, although the SeaDataNet ISO19139 CSR metadata standard has been accepted for use by all of the participating regional systems, some upgrades have been recommended and the implementation of the interface in regional systems still requires some significant work.

6.2 Regional synthesis

This section provides a synthesis of the implications of adopting the individual solutions described above (i.e. vocabularies, metadata, SWE etc.) for each of the regional data infrastructures at the end of the ODIP project in September 2015. The proposed frameworks for addressing these issues are shown in red in the table below.

Topic	Region		
	Europe	USA	Australia
Reference services (e.g. vocabularies , EDMO)	Maintain and upgrade agreed core services i.e. NVS, EDMO SeaDataNet, ODIPII	Mapping of existing regional vocabularies on to NVS and EDMO ODIPII	Mapping of existing regional vocabularies on to NVS and EDMO ODIPII
Federation of identity	Explore possible solutions (e.g. eduGain, marine-id) SeaDataNet, ODIPII	Explore possible solutions ODIPII	Explore possible solutions ODIPII
Standard metadata profiles	Establishment of a common working group; solution endorsement and dissemination RDA Marine Data Harmonization Interest Group and Data Description Registry Interoperability Working Group , ODIPII, IODE Ocean Data Standards project		
Operational brokering service	Maintain and make operational GEO-DAB RDA Brokering Working Group, ODIPII		
Cruise summary reports	POGO portal and CSR upgrades ODIPII	Maintain CSR interface ODIPII	Federate additional institutions delivering cruise summary reports ODIPII
Sensor web enablement (SWE)	Adoption of tools; technology deployment SeaDataNet, Eurofleets, NexOS, ODIPII, equipment manufacturers, oil and gas industry, navy	Adoption of tools; technology deployment ODIPII, equipment manufacturers, oil and gas industry, navy	Adoption of tools; technology deployment ODIPII, equipment manufacturers, oil and gas industry, navy

7 Conclusion

The outcomes of the final impact analysis documented in this report indicate that the ODIP objective to develop a collaborative platform for marine e-infrastructures which contributes to "the removal of barriers hindering the effective sharing of data across scientific domains and international boundaries" has been successful in a number of areas

As discussed in the general benefits section above, the combination of development activities (prototypes) and discussions (sessions in workshops) has proved to be effective in bringing together a number of IT, data management and domain experts to create common interoperability solutions and identify the agreed standards necessary to underpin the establishment of a common global framework for marine data management. The positive contributions and relevant feedback provided by this pool of experts, who regularly met during the ODIP workshops, demonstrates the efficacy of the selected partners and the relevance of the subjects addressed.

The performance indicators identified for the prototype development tasks indicate that a number of the stated objectives for the ODIP project have been achieved. These achievements are of two kinds: 1) IT and technical developments promoted by ODIP e.g. GEO-DAB brokering service for prototype development task ODIP1, GeoNetwork deployment and usage for Cruise Summary Reporting in the regional data centres as part of prototype ODIP 2, and the various experiments in prototyping or testing SWE implementations for prototype 3; and 2) progress in data management techniques, leading to the exposure of datasets and cruise information available in the regional data infrastructures via international data systems such as GEOSS, ODP and POGO. Currently the majority of these contributions are coming from the European SeaDataNet e-infrastructure, but additional data from the Australian and US services will also be available in the near future. A significant amount of progress has also been made in completing the relevant reference directories (e.g. EDMO for organizations) and vocabularies (e.g. instrument types), and mapping them to the regional systems to allow federation of the metadata records.

The impact of the **cross-cutting topics** on vocabularies, data publication/citation and data ingestion has proved to be highly positive due to the very significant contributions that have been made by the various experts involved in the project, and the outcomes described in section 5 above.

Finally, as well as the concrete outcomes of ODIP, another tangible achievement of the project has been to provide the participating marine data management systems in Europe, Australia and the USA with an accurate status report and a detailed work plan to efficiently improve the sharing and re-usability of their marine data holdings across scientific domains and international boundaries. This work plan can be used as a guide for future development activities and especially the ODIPII project which started in April 2015. This plan, detailed in section 6, could also potentially be used as part of the road map for those agencies responsible for the development and operation of the regional marine e-infrastructures.

Appendix A Terminology

Term	Definition
AODN	Australian Ocean Data Network
ARGO	Global array of more than 3,000 free-drifting profiling floats that measures the temperature and salinity of the upper 2000 m of the ocean
CDI	Common Data Index metadata schema and catalogue developed by the SeaDataNet project
CSIRO	Commonwealth Scientific and Industrial Research Organisation: the federal government agency for scientific research in Australia
CSR	Cruise Summary Reports is a directory of research cruises.
CSW	Catalogue Service for the Web (CSW): standard for exposing a catalogue of geospatial records in XML
CTD	Instrument used to determine the conductivity, temperature and depth of the ocean
EDMO	European Directory of Marine Organisations
EUROFLEETS	FP7 project creating a pan-European distributed research fleet infrastructure that provides coordinated access to European marine research vessels and equipment
GBIF	Global Biodiversity Information Facility: international open data infrastructure for biodiversity data and information
GEO-DAB	GEO Discovery and Access Broker (DAB): a middleware component that interconnects the heterogeneous and distributed data systems contributing to GEOSS
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
GEOSS	Global Earth Observation System of Systems: built by the Group on Earth Observations (GEO) to connect the producers of environmental data and decision-support tools with end users
IODE	International Oceanographic Data and Information Exchange (part of IOC)
IMOS	Integrated Marine Observing System: Australian monitoring system; providing open access to marine research data
ISO 19139	Standard providing XML implementation schema for ISO 19115 specifying the metadata record

	format. May be used to describe, validate, and exchange geospatial metadata prepared in XML
JERICO	FP7-funded project developing a European coastal marine observatory network
ODP	Ocean Data Portal: data discovery and access service, part of the IODE network
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	System for providing researchers with a unique persistent identifier
MCP	Marine Community Profile: Australian marine metadata standard
MIKADO	Tool for generating metadata in XML format created by SeaDataNet
NODC	National Oceanographic Data Centre, USA
NVS	NERC vocabulary server
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH):protocol to harvest the metadata descriptions of the records in an archive
OpenSearch	Collection of technologies that allow publishing of search results in a format suitable for syndication and aggregation.
OPenID	Open standard and decentralized protocol developed by the non-profit OpenID Foundation that allows users to be authenticated by certain co-operating sites (known as Relying Parties or RP) using a third party service
POGO	Partnership for Observation of the Global Oceans
R2R	Rolling Deck to Repository: a US project responsible for the cataloguing and delivery of data acquired by the US research fleet.
ResearcherID	System to provide authors with unique identifiers
SeaDataNet	SeaDataNet: EU-funded pan-European e-infrastructure for the management and delivery of marine and oceanographic data
SensorML	OGC standard providing models and an XML encoding for describing sensors and process lineage

Shibboleth	A single sign-on (log-in) system for computer networks
SKOS	Simple Knowledge Organization System:W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary.
SPARQL	SPARQL Protocol and RDF Query Language (recursive acronym)
SOOS	Southern Ocean Observing System
SOS	Sensor Observation Service: a web service to query real-time sensor data and sensor data time series. Part of the Sensor Web
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
UTAS	University of Tasmania, Australia