Ocean Data Interoperability Platform

Deliverable D3.3: Progress and Results from Prototype Developments 1

Work package	WP3	ODIP Prototypes
Author (s)	Dick M.A. Schaap	MARIS
Author (s)	Rob Thomas, Roy Lowry, Alexandra Kokkinaki	BODC
Author (s)	Adam Leadbetter	MI
Author (s)	Simon Jirka	52°North
Author (s)	Stefano Nativi, Alessandro Oggioni, Paolo Tagliolato	CNR
Author (s)	Anne Che Bohnenstengel	BSH
Authorized by	Helen Glaves	NERC
Reviewer		
Doc Id	ODIP II_D3.3	
Dissemination Level	PUBLIC	
Issue	1.0	
Date	3 December 2016	

Abstract:

The Ocean Data Interoperability Platform project is developing a series of prototype solutions that establish interoperability between existing regional and relevant global marine data infrastructures. These interoperability solutions aim to demonstrate effective sharing of data between these data systems and support development of a common global framework for marine data management.

Initial formulation and development of these interoperability solutions formed part of the activities in the previous ODIP project. This deliverable describes the continued development and expansion of the three ODIP prototype interoperability solutions as part of the current ODIP II project.

Document History				
Version	**Author(s)**	**Status**	**Date**	**Comments**
0.1	Dick M.A. Schaap (MARIS)	DRAFT	2 November 2016	First draft using contributions from several partners
0.2	Helen Glaves	DRAFT	14 November 2016	Editing/proof reading
1.0	Helen Glaves	FINAL	30 November 2016	Final edits and sign-off

Contents

Executive Summary

Prototype development tasks have been formulated within the framework of the ODIP and ODIP II projects to evaluate and test selected potential standards and solutions for establishing and demonstrating improved interoperability between regional data infrastructures and also with global e-infrastructures.

During the first ODIP II workshop, which took place in September 2015, the project partners agreed to expand the three prototype development tasks that had been successfully initiated and implemented in the previous ODIP project. These prototype development tasks are:

- **ODIP 1**: Establishing interoperability between SeaDataNet CDI, NCEI (formerly US NODC) and IMOS MCP data discovery and access services using a brokerage service, and moving towards interacting with the global IODE-ODP and GEOSS portals;

- **ODIP 2**: Establishing a common deployment and interoperability between cruise summary reporting (CSR) systems in Europe, the USA and Australia by making use of GeoNetWork, and moving towards interaction with the global POGO portal;

- **ODIP 3**: Establishing a prototype for a Sensor Observation Service (SOS) and formulating common O&M and SensorML profiles for selected sensors (SWE), installed on research vessels and in real-time monitoring systems.

The specifications for the expansion of these prototype development tasks have been formulated in the 7-month period between the first and second ODIP II workshops and documented in the ODIP II deliverable *D3.1 Definition of ODIP Prototypes 1*. In summary, the aims of the prototype expansion projects are as follows:

- **ODIP 1+**: Analysing options for establishing semantic interoperability between the three regional discovery and access services at the metadata level as well as establishing horizontal interoperability by developing an ODIP interface using WMS – WFS and possibly OpenSearch capabilities. This task will also address the analysis necessary for setting-up the transformation services for converting the SeaDataNet ODV data format to both the Observations & Measurements (O&M) data model following the INSPIRE guidelines, and to an extended SeaDataNet NetCDF (CF) data format. This task will be led by MARIS with contributions by European, US and Australian partners;
- **ODIP 2+:** Further population of the Cruise Summary Reports database for US cruises from R2R, initiating contributions for Australian cruises, and incorporating ICES legacy CSRs. The task will also undertake an analysis for upgrading the CSR metadata format, schema and related tools and services following suggestions previously received from the R2R project partners; and also publishing CSRs in RDF by means of a SPARQL endpoint. This task will be led by BSH with contributions from European, US and Australian partners;
- **ODIP 3+:** Establishing OGC Sensor Web Enablement (SWE) standards for facilitating the interoperable sharing of oceanographic observation data and metadata. This task will include an analysis of the handling of large volumes of data within SWE-based infrastructures, use of lightweight technologies such as JSON and REST as complementary technologies,and use of RDF-based approaches for supporting the discovery of marine sensors and data sets; and synchronising efforts for metadata/SensorML editors. Task will be led by 52°North with contributions from European, US and Australian partners.

This deliverable *D3.3 Progress and Results from Prototype Developments 1* gives a detailed report of the activities that have been undertaken and the progress that has been achieved up to the end of Month 18 for the three ODIP II prototype expansion tasks, as well as the cross-cutting topic on vocabularies. Significant progress has already been achieved in several areas but some of the activities have only recently commenced. The development activities for the three prototype interoperability solutions and cross-cutting topics are continuing and the final results will be reported in the deliverable *D3.4 Progress and results from prototype developments* that will be released in M32 of the ODIP II project. Deliverable D3.4 will also report on the progress and results achieved for the new ODIP II prototype development tasks, specifications for which are documented in deliverable *D3.2 Definition of ODIP prototypes 2*.

1 Introduction

The "Extending the Ocean Data Interoperability Platform" (ODIP II) project is promoting the development of a common global framework for marine data management by establishing interoperability between existing regional e-infrastructures in Europe, the USA and Australia and with global infrastructures such as GEOSS, IOC-IODE and POGO.

To establish interoperability between the different data systems four workshops are being organised during the three years of the project that bring together relevant domain experts and other technical specialists to provide insights into commonalities and differences between the individual marine data systems. They will also discuss and compare their respective approaches, best practices and standards, and identify opportunities for the development of common standards and interoperability solutions.

The initial set of ODIP II prototyping activities, which are extensions of those interoperability solutions developed in the previous ODIP project, were discussed and agreed between partners during the 1st and 2nd ODIP II workshops. The specifications for the expansion of the prototype development tasks have been documented in the ODIP II deliverable *D3.1 Definition of ODIP Prototypes 1*.

This deliverable *D3.3 Progress and Results from Prototype Developments 1* provides a detailed report on the activities that have been undertaken and the progress that has been made up to the end of Month 18 (September 2016) for the three ODIP II prototype expansion projects as well as for the cross-cutting topic addressing vocabularies. It should be noted that this deliverable gives a report of current progress for on-going development activities that will be continued until the end of the ODIP II project. The final results will be reported in deliverable *D3.4 Progress and results from prototype developments* that will also cover the results of the new prototype development tasks that being initiated as part of the ODIP II project which are specified in deliverable *D3.2 Definition of ODIP prototypes 2*.

One issue that has arisen for ODIP II is that, unlike the previous phase of the project, only the European partners have received funding for the continuation of the project. Any contributions made to the project by partners from the US and Australia must be supported by funding from their own institutes. This places even more emphasis on the need for an approach to the prototype development tasks that leverages the activities of existing regional projects and initiatives being undertaken by the ODIP II partners. This situation also dictates that the new ODIP II prototype development tasks must be formulated taking these constraints into account and that any additional development activities outside the scope of the ongoing regional projects should be done largely by the European ODIP II partners.

2 Progress of the three prototype expansions

The specifications for the expansion of the existing ODIP prototype development tasks have been formulated in the 7-month period between the 1st and 2nd ODIP II workshops and documented in deliverable *D3.1 Definition of ODIP Prototypes 1*.

The targets for the prototype development task expansions can be summarized as follows:

- **ODIP 1+**: Analysing options for establishing semantic interoperability between the three regional discovery and access services at metadata level as well as establishing horizontal interoperability by developing an ODIP interface using WMS – WFS and possibly OpenSearch capabilities. This task also aims to analyse the set-up of transformation services for converting the SeaDataNet ODV format to the Observations & Measurements (O&M) data model following INSPIRE guidelines, and to an extended SeaDataNet NetCDF (CF) data format. This task is led by MARIS with contributions from European, USA and Australian partners

- **ODIP 2+:** Enhanced population of the Cruise Summary Reports (CSR) database by adding further cruises from the R2R service in the USA, initiating contributions for Australian cruises, and including ICES legacy CSRs. This task will also analyse the upgrading of the CSR metadata format, schema and related tools and services in line with suggestions previously made by the R2R partners. This activity will also support publishing CSRs in RDF by means of a SPARQL endpoint. This task is led by BSH with contributions from partners from Europe, the USA and Australia

- **ODIP 3+:** Establishing OGC Sensor Web Enablement (SWE) standards to facilitate the interoperable sharing of oceanographic observation data and metadata. This task will analyse the handling of large volumes of data within SWE-based infrastructures, use of lightweight technologies such as JSON and REST as complementary technologies, and use of RDF-based approaches for supporting the discovery of marine sensors and data sets. It will also encourage synchronisation of existing efforts to develop metadata/SensorML editors. This task will be led by 52°North with contributions from European, US and Australian partners.

The activities that have been undertaken and the progress that has been achieved by the three expanded ODIP prototype development tasks up to the end of Month 18 are described below.

2.1 ODIP 1+ Prototype Development Task

ODIP1+ is an extension of the ODIP 1 prototype development task initiated during the previous project. This task addresses interoperability of the selected regional marine discovery services (SeaDataNet, NCEI and AODN) with the global GEOSS and ODP portals. At present there is metadata brokerage at the collections level and entries are included in GEOSS and ODP with return links to data at the granule level in the regional portals where it can be accessed and downloaded in most cases.

The ODIP 1+ prototype development task is led by MARIS with planned contributions from partners in the three participating regions.

Activities have been undertaken with further ones planned to address each of the objectives of the ODIP 1+ prototype development task:

2.1.1 Establishing semantic interoperability between the three regional discovery and access services at the metadata level

The concept for this task is to develop a translation service that will interact with the GEO-DAB brokerage service. This service will capture knowledge about which terms are identical, similar or specialisations of other terms used across the participating regional systems (SeaDataNet, NCEI and AODN) and publish this information in a machine-to-machine system. This activity has started by identifying vocabularies used in the three regional systems which can be broken down as follows:

SeaDataNet makes use of:

- SeaDataNet Common Vocabularies for various attributes such as discovery parameters, platforms, sea regions, access restriction policies, coordinate reference systems, data transport formats, parameter usage vocabulary, units, instrument types, instruments, Climate and Forecast (CF) standard names, ISO Country codes, and ICES platform codes. Hosted by NERC-BODC http://seadatanet.maris2.nl/v_bodc_vocab_v2/welcome.asp
- SeaDataNet EDMO codes for organisations http://seadatanet.maris2.nl/v_edmo/welcome.asp
- SeaDataNet EDMERP codes for projects http://www.seadatanet.org/Metadata/EDMERP-Projects
- SeaDataNet CSR codes for Cruise Summary Reports http://www.seadatanet.org/Metadata/CSR-Cruises

The component data centres in the SeaDataNet infrastructure may also use other vocabularies in their own local systems, but each connected data centre has to map its local vocabularies to those used in SeaDataNet for the purposes of harmonised population of the metadata directories and data resources.

NOAA's National Centres for Environmental Information (NCEI) make use of:

- NASA's Global Change Master Directory (GCMD), for earth science, data centres, locations, instrument/sensors, platforms/sources and projects
- NODC Vocabularies for people, projects, institutions, ICES platform codes, sea names, data types, observations, instruments, ISO country codes, and Climate and Forecast (CF) standard names

Next to these primary vocabularies, use is also made of other vocabularies including:

- Ocean Exploration and Research (OER) Discovery Keywords
- Getty Thesaurus of Geographic Names
- Library of Congress Subject Headings
- SeaDataNet Common Vocabularies
- Geographic Names Information System

- GEBCO Gazetteer of Undersea Feature Names

NOAA and its associated data centres, which have recently been reorganised to become the National Centres for Environmental Information (NCEI), oversee and manage an enormous data collection that extends back a long time. This implies that various code lists have been used over time and that harmonising these to one common set of vocabularies will have major consequences for the marking-up of the large data collections and the associated metadata.

The Australian Integrated Marine Observing System (IMOS) makes use of the SeaDataNet Common Vocabularies, wherever possible. The decision to adopt these vocabularies has largely been influenced by the IMOS participation in ODIP. IMOS has also become a regular contributor of new terms to these vocabularies. Furthermore, IMOS has become the Australian national node for entering and maintaining Australian organisations in the SeaDataNet EDMO directory with more than 250 relevant entries to date.

IMOS is also a component of the overarching Australian Ocean Data Network (AODN), and during its development it became apparent that many of the larger Australian institutions (e.g. CSIRO Oceans and Atmosphere, AIMS, Australian Antarctic Division, GA, Bureau of Meteorology) already use some form of in-house vocabularies. Many of these vocabularies are not formalized, published or well governed but the terminology is often integral to how their internal (or public-facing) systems operate. IMOS has therefore begun to develop mappings between institutional level terminologies and the AODN common vocabulary as the primary means of establishing standardized vocabulary usage for metadata that is required to underpin the AODN data delivery infrastructure.

Considering the current status of the three regional systems it should be feasible to determine the mappings between the various vocabularies being used for comparable metadata fields by SeaDataNet and AODN. However, undertaking a similar exercise with NCEI will be a considerable challenge due to their very rich metadata format which has many additional attributes when compared to the SeaDataNet CDI and AODN MCP metadata formats.

The SeaDataNet CDI and AODN MCP metadata formats have been set up as common core implementations of the ISO19115 standard with attributes supported by controlled vocabularies wherever possible. However, both the SeaDataNet CDI and AODN MCP formats can be considered to be more limited and contain less information than the NCEI metadata format which has been developed and populated in part by bringing together legacy metadata systems.

Instead of trying to map the various vocabularies to their full extent another approach will be implemented. Use will be made of the comparisons of the different metadata formats used by the three regional systems that was previously done by CNR. Each of the metadata formats from the three regional data systems have been analysed and mapped to the Brokerage Reference Schema as part of the configuration of the GEO-DAB Brokerage Service. As a result, the GEO-DAB service provides CSW and OAI-PMH services for each of the three regional systems that delivers converted XML records while maintaining the original terms for specific attributes. Examples of these services are listed below.

SeaDataNet

The relevant OAI-PMH service for SeaDataNet collections is available at:

http://seadatanet.essi-lab.eu/gi-cat/services/oaipmh?verb=ListIdentifiers&metadataPrefix=ISO19139&set=SEADATANET

A CS-W service Version 2.0.2 Service – HTTP POST method for SeaDataNet collections which uses SEADATANET as the parameter for retrieving the SeaDataNet output can be found at:

http://seadatanet.essi-lab.eu/gi-cat/services/cswiso

.

NCEI Services

The relevant OAI-PMH service for NCEI collections is available at:

http://seadatanet.essi-lab.eu/gi-cat/services/oaipmh?verb=ListIdentifiers&metadataPrefix=ISO19139&set=NODC

A CS-W service Version 2.0.2 Service – HTTP POST method for NCEI collections that uses 'apiso:parentIdentifier' = NODC can be viewed at:

http://seadatanet.essi-lab.eu/gi-cat/services/cswiso

AODN Services

The relevant OAI-PMH service for AODN collections is available at:

http://seadatanet.essi-lab.eu/gi-cat/services/oaipmh?verb=ListIdentifiers&metadataPrefix=ISO19139&set=AODNCSWCORE

A CS-W service Version 2.0.2 Service – HTTP POST method for AODN collections that uses 'apiso:parentIdentifier' = AODNCSWCORE can be viewed at:

http://seadatanet.essi-lab.eu/gi-cat/services/cswiso

A recent exercise (October 2016) to harvest the XML output from each of the three regional services using the OAI-PMH services resulted in:

- 28670 XML entries from NCEI
- 496 XML entries from SeaDataNet
- 127 XML entries from AODN

Next steps:

- The harvested XML files will be loaded into a database that is configured to the brokerage reference schema;
- Shortlists will be made of the range of terms for each attribute that have originated from each of the three regional systems;
- These shortlists will then be subject to further mapping in consultation with the managers of the regional systems which will include details of the source vocabularies that have been used;
- If successful, these mappings will provide the basis required for the vocabulary translation service which is planned in ODIP II;
- Options will be considered for the publication of the mappings in both human and machine readable formats. An “ODIP Rosetta Stone” vocabulary service might be provided and maintained. One or more vocabulary services will host all of the mappings that are required for catalogue harmonization;
- An analysis of how the ODIP Rosetta Stone vocabulary service will interact with the GEO-DAB brokerage service will be undertaken. The best solution will be implemented in order to expose XML metadata from the regional systems with harmonized vocabularies as CSW and OAI-PMH services. Functionality for semantic discovery, utilizing expansion of textual terms and navigation has previously been developed for the GEO-DAB broker which will be taken into account as part of this activity. Further information about this additional functionality is documented in the publication Methodologies for Augmented Discovery of Geospatial Resources (M. Santoro et al; 2012)[1]

2.1.2 Establishing horizontal interoperability between the regional marine data systems

The metadata entries from the three regional marine data systems are currently exposed in the global GEOSS and ODP portals. The ODIP 1+ expansion task aims to also develop and implement a discovery and access service on the ODIP portal using WMS-WFS and possibly the OpenSearch protocol.

SeaDataNet currently has operational WMS-WFS and OpenSearch services, both at collections level and at the granules level (i.e. locations (points, tracks, polygons) of individual observations are mapped with the related metadata also made available).

The first step in this activity is to identify similar WMS-WFS and OpenSearch services that are being delivered by the regional systems:

AODN: The metadata describing the dataset collections available on the AODN portal (https://portal.aodn.org.au/) are accessible through a GeoNetwork instance (https://catalogue-portal.aodn.org.au/geonetwork/srv/eng/main.home). Each collection XML contains spatial information in the form of a latitude-longitude box which can be used for populating a map layer at the collection level. Moreover, in the online resources section of each collection XML

[1] Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications (2012) DOI: 10.4018/978-1-4666-0945-7.ch009

there are WMS and WFS services available e.g. general service (http://geoserver-123.aodn.org.au/geoserver/wms) and a tag indicating the specific layer (for example**: imos:soop_co2_trajectory_map**)

An overview of all AODN collection layers can also be retrieved from a GeoServer interface at: http://geoserver-123.aodn.org.au/geoserver/web/

Most of the layers follow a filename pattern.

- "_map" in the layer name means that this layer is used as a WMS (to represent the data of a particular dataset collection on step 2 of the AODN portal)
- "_data" in the layer name means that this layer is used as a WFS service. This layer contains all the measurements for the corresponding dataset and this is used in step 3 of the AODN portal when the user can select multiple download options.

Analysis of the XML provided by the GEO-DAB brokerage service for AODN appears to indicate that the WMS service is included; however, the tag indicating the specific layer is missing. Partner CNR will undertake corrective actions to remedy this issue. CNR will also update the harvesting process as progress is made with the AODN portal and catalogue which now also includes entries from other Australian organisations.

NCEI: The oceans section of NCEI (formerly US-NODC) has implemented numerous interoperable data technologies to enhance the discovery, understanding, and use of the vast quantities of oceanographic data in the NODC archives. When combined these technologies enable NODC to provide access to its data holdings and products through some of the commonly-used standardized web services. The Geoportal (http://data.nodc.noaa.gov/geoportal/) is used as an integrating technology, bringing together various data access, visualization, discovery services, and metadata into a user-focused framework. It gives a user interface on top of the catalogue of collections. Each NODC collection XML contains spatial information as a latitude-longitude box which can be used for populating a map layer at the collection level. WMS and WFS services are not identified in the XML; however, in many cases there is a link to an image included which gives a detailed map of the observations at granule level. The GEO-DAB broker also harvests information about the NCEI collections from the Geoportal. Analysing the XML as provided by the GEO-DAB brokerage service for NCEI indicates that the link to the image of the detailed map is included.

It can therefore be concluded from this evaluation of the individual regional data systems that all three services are providing spatial information in the form of a latitude/longitude box for the data collections. Links are provided in the collection level metadata for displaying a detailed map of the locations of the individual observations (granules) that make up the collection that are delivered as a WMS in the case of SeaDataNet and AODN, and as a static image for NCEI. It should also be noted that there are no overarching WMS, WFS or OpenSearch services available from AODN and NCEI at the catalogue level.

Following this analysis of the regional systems the next steps are:

- Partner CNR will upgrade the brokerage to incorporate the catalogue being used by AODN including the WMS layer tags in the XML output. Partner MARIS will use the harvested XML entries in the database to extract the spatial information (latitude-longitude boxes) for all of the collections in order to populate an overarching WMS – WFS service;

- MARIS will build a common WMS – WFS user interface on the ODIP website that provides search options for the connected regional systems at the metadata level
- The WFS will retrieve the collection metadata, including a detailed map of the locations of individual observations within a collection using the WMS or image links
- The “ODIP Rosetta Stone” vocabulary service (described in section 2.1.1 above) will be integrated into this service to support harmonisation of vocabulary terms
- Analyse and implement dynamic maintenance of the user interfaces.

Note: As part of the previous EuroGEOSS project CNR developed advanced REST/JSON APIs to interface with the GEO-DAB service. See: http://api.eurogeoss-broker.eu/dab/api-rest-docs/. These APIs will be analysed to determine if they can be used for any part of the required functionality.

2.1.3 Transformation services for converting the SeaDataNet ODV format to the Observations & Measurements (O&M) data model following INSPIRE guidelines and to an extended SeaDataNet NetCDF (CF) data format

There has been limited activity to address this aspect of the ODIP 1+ task to date but more attention will be paid to this topic during the second part of the project, due in part to the influence of the new SeaDataCloud and EMODnet projects that will have an increased focus to INSPIRE compliance and applying the INSPIRE implementing rules at the data level.

Initial ODIP II activities have focused on the netCDF (CF) data format. The netCDF format is already an OGC standard as a result of the work done by CNR and UNIDATA that are ODIP II partners. The netCDF (CF) data format is widely used in the oceanographic and meteorological communities. Unfortunately, there are various netCDF (CF) profiles in use by different observing programmes and research groups. It is therefore a challenge for ODIP II to develop and recommend common netCDF (CF) profiles for the marine community. For this reason, netCDF should be considered as an exchange format and not as a storage format.

An initial discussion regarding the netCDF format took place during the 1st ODIP II workshop (September 2015). It was noted that there are several CF attribute lists which cover different topics and the groups behind these lists are not working together. Some of these groups are extending the CF terms but not in a compliant manner.

The CF conventions are also relatively flexible with guidelines that are also quite loose. The CF conversions are also multifaceted with two different versions for dealing with gridded data (1.5 and 1.7), while version 1.6 is specifically for use with point data. Another issue is that some data is identified as being CF compliant and even passes the CF checker when it is not. This occurs because the official checkers are only suitable for gridded data and not for point data. IMOS has adopted the US-IOOS NetCDF CF checker and extended it by adding an IMOS plug-in. The IOOS netCDF checker can be modified to accept multiple plug-ins to satisfy different needs.

It has been recognized that there cannot be a single global implementation of the CF conventions because different communities have, and will continue to have, different requirements for its use. In an attempt to address this situation partner BODC has proposed a layered structure that is based on the CF conventions with a layer for the specific community conventions for all data types above it. Certain aspects of this solution have already been implemented in SeaDataNet netCDF (CF) format for point data where part of the profile includes an attribute for the SeaDataNet parameter codes in addition to the standard name. Other layers could also be added on top for specific types of data, e.g. bathymetry. BODC proposed that this approach could be adopted for developing CF compliant NetCDF profiles for specific gridded datasets.

Another aspect of the NetCDF activities is to gather and formulate best practice. For instance, typically one instrument is in included in a single netCDF file, but communities such as OceanSITES include multiple instruments which makes data management more difficult.

To support this activity partner CNR participated, on behalf of the ODIP II project, in the NetCDF Summit Workshop which was organised by UNIDATA and funded by the NSF that took place on 24-26 May 2016 in Boulder, Colorado, USA. (More details of this workshop can be found on the UNIDATA website: http://www.unidata.ucar.edu/events/2016CFWorkshop/#home)

The objective of CNR participating in this workshop was to advance netCDF – CF (also from the perspective of the SeaDataNet NetCDF (CF) and the INSPIRE data models) and contribution of a presentation addressing:

- Metadata conventions for netCDF spanning dataset aggregation and granularity;
- Injecting vocabulary URIs and Linked Data elements into netCDF.

The presentation is available at:

http://www.odip.org/media/odip/org/documents/nativi_presentation-3_boulder_24-26_may-pub.pdf

This presentation was also prompted by CSIRO that are addressing the relationship between NetCDF header metadata and standalone "document" metadata. A situation is frequently encountered where software such as THREDDS can generate ISO metadata from the headers as well as a standalone ISO metadata record using GeoNetwork or similar. It would therefore be useful to have a recommended approach regarding the best way to integrate the two to avoid duplication. CSIRO has started using the "metadata_link" attribute in the NetCDF global header with a link to a metadata record. In this situation it would be useful if tools such as THREDDS utilize this metadata_link for generation of the metadata.

The SeaDataNet data system is a comparable example, providing users downloading ODV data files with limited metadata and an accompanying CSV file with full CDI metadata of all the related ODV data files. One issue that has arisen is that users discarded the CDI metadata and only used the ODV data files and then complained about the lack of metadata.

This is now being tackled by SeaDataNet through the ODV software that users can freely download and use for various data analyses, checks and visualisations. Extra functionality has been added to read and write CDI metadata into the ODV files during import. This creates so-called metadata enriched ODV files which provide users with the required stand-alone information.

This development activity will be continued throughout the remainder of the ODIP II project with the following development steps planned:

- Compile an inventory of netCDF CF profiles that are being used within the ODIP community
- Compile an inventory of netCDF CF checkers including details of which ones are being used and by whom, plus capturing information about compatible plug-ins

The outcomes of these activities will then be used as input for development of further ODIP II prototype interoperability solutions.

2.2 ODIP Prototype 2+

This activity is an expansion of the ODIP 2 prototype development task that was initiated during the previous project. It is focused on the integration of regional cruise summary reporting systems in Europe (SeaDataNet), the USA (R2R) and Australia (MNF) with the global Partnership for Observation of the Global Oceans (POGO) portal. As a result of the activities in the previous ODIP project the SeaDataNet CSR V3.0 schema has now been adopted for cruise summary reporting in Europe and the USA, while in Australia the preparatory mappings necessary for its adoption have been carried out by partners.

The ODIP 2+ prototype development task is led by European partner BSH with contributions expected from all three participating regions. A number of activities have been undertaken or are planned to address the following objectives for ODIP 2+:

1) further population of the CSR directory
2) carrying out the analysis necessary for upgrading the CSR metadata format, schema and related tools and services
3) publication of CSRs in RDF through a SPARQL endpoint.

To date the activities of the ODIP 2+ prototype development task have concentrated on the further population of the CSR directory. Activities to address objectives 2) and 3) are planned for M18 – M36 in an effort to maximize the opportunities to align these activities with the new H2020 SeaDataCloud project which begins in November 2016 (http://cordis.europa.eu/project/rcn/207433_en.html).

2.2.1 Population of cruise directory

As a result of the activities in ODIP 2+ there has been a significant rise in the number of cruise summary reports that have been added to the central cruise catalogue managed by BSH and also exposed in the POGO portal since 1 April 2015 as shown in Table 1 below.

Partner region	Total CSR submissions (including updates)	CSRs added to POGO portal
Europe	1706	451
USA	748	704
Australia	1	1

Table 1 Submission of additional CSRs since M1 of the ODIP II project (April 2015)

Unfortunately, due to issues arising from the reorganisation of the Australian agencies contributing to the ODIP II project, input of CSRs from this region has been minimal. However, CSIRO will endeavour to provide CSRs for the new ship, RV Investigator, from 2017.

Table 2 below provides a breakdown of the CSR entries that qualify for inclusion in the POGO portal as ocean-going vessels (greater than 60m long). Since April 2015 a total of 257 additional CSR entries have been published on the POGO CSR website http://www.pogo-oceancruises.org/

Ship	Responsible country for cruise	Number of CSRs
Aegaeo	Italy	1
Aranda	Finland	1
Aranda	Sweden	14
Arni Fridriksson	Norway	1
Atlantis	United States	14
Beautemps-Beaupre	France	12
Belgica	Belgium	3
Celtic Explorer	Germany	3
Celtic Explorer	Ireland	7
Celtic Voyager	Ireland	6
D. Carlos I	Portugal	1
Falkor	United States	8
G.O. Sars	Norway	20
James Clark Ross	United Kingdom	5
James Cook	United Kingdom	7
Johan Hjort	Norway	14
Kilo Moana	United States	13
L'Atalante	France	16
Le Suroit	France	1

Marcus G. Langseth	United States	5
Maria S. Merian	Germany	12
Meteor	Germany	15
OGS Explora	Italy	4
Pelagia	United Kingdom	1
Polarstern	Germany	3
Poseidon	Germany	11
Pourquoi pas?	France	13
Roger Revelle	United States	8
Sarmiento de Gamboa	Spain	1
Thalassa	France	9
Thomas G. Thompson	United States	14
Vizconde de Eza	Spain	4
Walther Herwig III	Germany	7
Zirfaea	Netherlands	3
TOTAL		**257**

Table 2 Breakdown of the CSRs added to the POGO portal since the beginning of ODIP II (April 2015)

Figure 1 below shows an example listing of CSRs delivered by a query submitted via the POGO portal.

Cruise Summary Report Inventory (CSR)

SET OF RESULTS

| New Query | Found 257 | Show (1-20) | Next 20 |

Platform Name	Cruise Name	from	to	Show
Arni Fridriksson	2016845	10.08.2016	01.09.2016	
Celtic Explorer	BSH North Sea Summer Survey	03.08.2016	26.08.2016	
Meteor	M129	30.07.2016	25.08.2016	
Aranda	2016-34AR/0355-0386	20.07.2016	26.07.2016	
Johan Hjort	2016208	15.07.2016	13.08.2016	
Meteor	M128 Azores Plateau	02.07.2016	27.07.2016	
Johan Hjort	2016	27.06.2016	14.07.2016	
Sarmiento de Gamboa	BOCATS	17.06.2016	27.07.2016	
Aranda	2016-34AR/0326-0354	13.06.2016	19.06.2016	
Polarstern	PS99 ARK-XXX/1	13.06.2016	16.07.2016	
Maria S. Merian	MSM55	11.06.2016	29.06.2016	

Figure 1 Example listing of CSRs published in the POGO portal

Figure 2 shows a cruise summary report from the R2R project displayed in the POGO graphical user interface and illustrates how the CSR schema, which supports tracks in GML, allows detailed navigational tracks to be provided via a WMS.

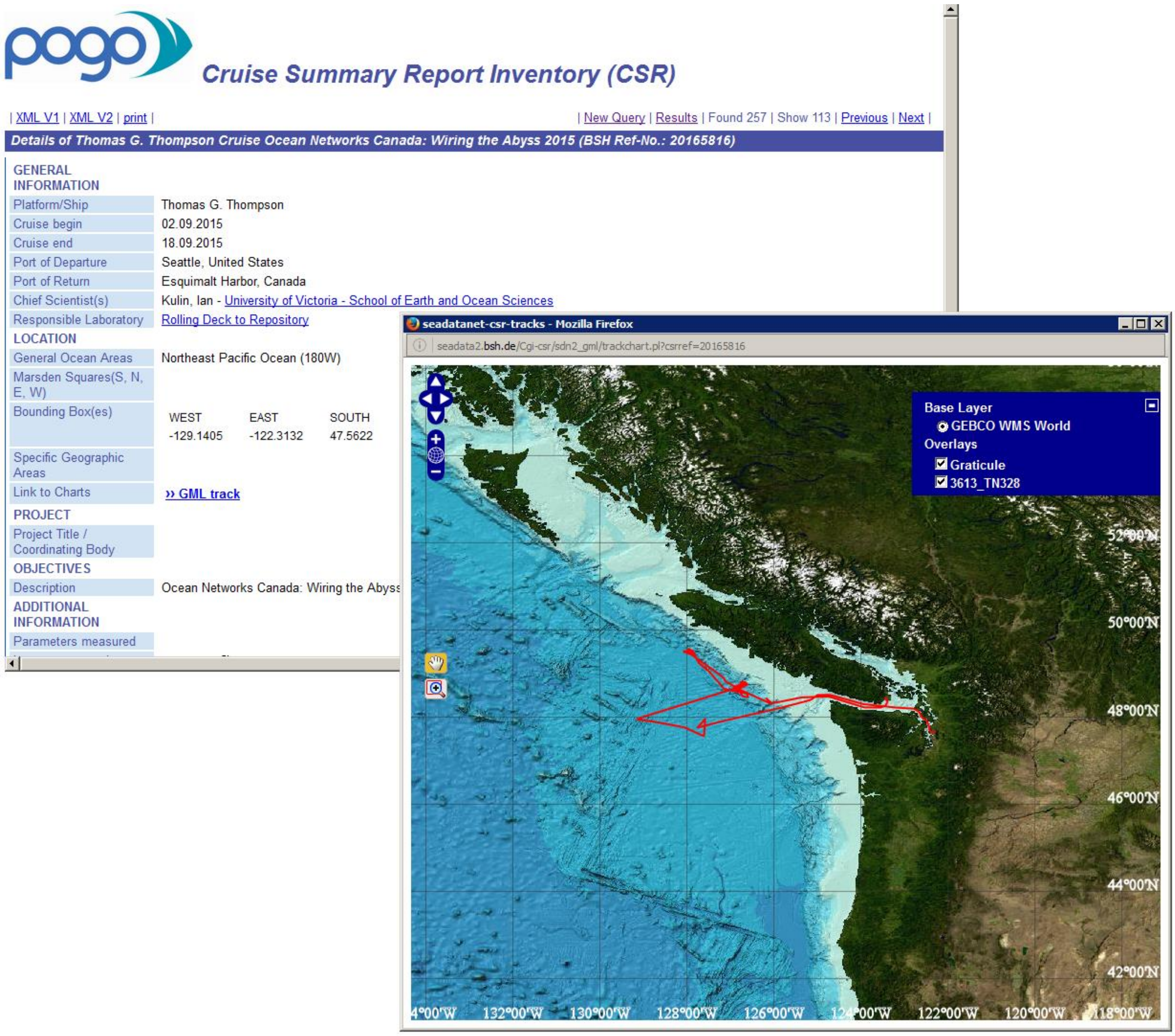


Figure 2 Example of dynamic WMS cruise chart displayed in the POGO portal

CSR Harvesting: progress

Harvesting of cruise summary reports from those partners connected to the CSR catalogue has continued on a weekly basis since April 2015. The contents are automatically validated for mandatory fields and vocabulary usage along with some basic consistency checks e.g. on coordinates (bounding boxes and moorings), reference dates etc. A harvesting portal has also been implemented for conducting further visual checks on the CSRs both by BSH and the relevant partners, which is illustrated in Figure 3 below. There are currently only four centres connected for regular harvesting of CSRs but other partners have begun to set-up GeoNetwork servers to support CS-W harvesting of their CSRs. However, there are still some issues with local system security which are causing problems for some partners. The new SeaDataCloud project is also expected to encourage several NODCs to upgrade their

mechanism for submitting new CSR entries which is currently done by sending XML records by e-mail.
All CSRs submitted (via online, email, ftp, harvesting) are also accessible from the BSH GeoNetwork portal http://seadata.bsh.de/geonetwork-sdn/ under resource type ‘series’ (see Figure 4). The records can be downloaded in XML, RDF and PDF format.

Cruise Summary Report Harvesting - IFREMER

16 records harvested on 2016-10-11

BSH Ref view	Local cruise ID	Platform	Cruise	Begin	End	Laboratory	status
20136610	7480160	Cotes De La Manche	PECH4	23.11.2007	30.11.2007	Universite D'Angers / Laboratoire Des Bio-Indicateurs Actuels Et Fossiles (Biaf)	OK
20126642	11100080	Alis	EXBODI	28.08.2011	29.09.2011	Museum National D'Histoire Naturelle / Departement Systematique & Evolution	OK
20106226	10020070	Le Suroit	ESSINF	10.07.2010	15.07.2010	IFREMER / NSE-DEPARTEMENT NAVIRES ET SYSTEMES EMBARQUES	OK
20106221	10020020	Le Suroit	BOBGEO2	18.07.2010	25.07.2010	IFREMER / GM-MARINE GEOSCIENCES	OK
20096021	9030060	Pourquoi pas?	BOBGEO	14.10.2009	28.10.2009	IFREMER / GM-MARINE GEOSCIENCES	OK
20096018	9030020	Pourquoi pas?	ESSNAUT 2009	09.05.2009	19.05.2009	IFREMER / GENAVIR LA SEYNE SUR MER	OK
20096003	9010030	L'Atalante	VALIDOP	26.06.2009	14.07.2009	IFREMER / CENTRE DE BRETAGNE	OK
20087328	8480100	Cotes De La Manche	PECH6	07.07.2008	13.07.2008	Universite D'Angers / Laboratoire Des Bio-Indicateurs Actuels Et Fossiles (Biaf)	OK
20087320	8480010	Cotes De La Manche	PECH5	01.03.2008	08.03.2008	Universite D'Angers / Laboratoire Des Bio-Indicateurs Actuels Et Fossiles (Biaf)	OK
20087133	8030140	Pourquoi pas?	ESSRESON08-2	08.12.2008	13.12.2008	IFREMER / TSI/AS- ACOUSTIC SEISMIC	OK

Figure 3 CSR harvesting portal

CSR harmonisation with ICES

A large number of the CSRs in the central inventory originated from the ICES ROSCOPs system that includes information for cruises up to and including those from 2004. However, the contents of these legacy records could not be updated to the present standards for CSRs due to missing metadata elements. Among these records are a large number of cruises from Australia, Canada and the USA which contain valuable legacy information.

In the recent years ICES has started to partially adopt the SeaDataNet standards and vocabularies. To support this approach BSH has created a special CSR web service that allows ICES to access SeaDataNet CSR cruise metadata for their datasets where available. Since July 2015 the ICES cruise metadata have been synchronized with the SeaDataNet central CSR inventory on a daily basis using the “amendment date” of the individual records.

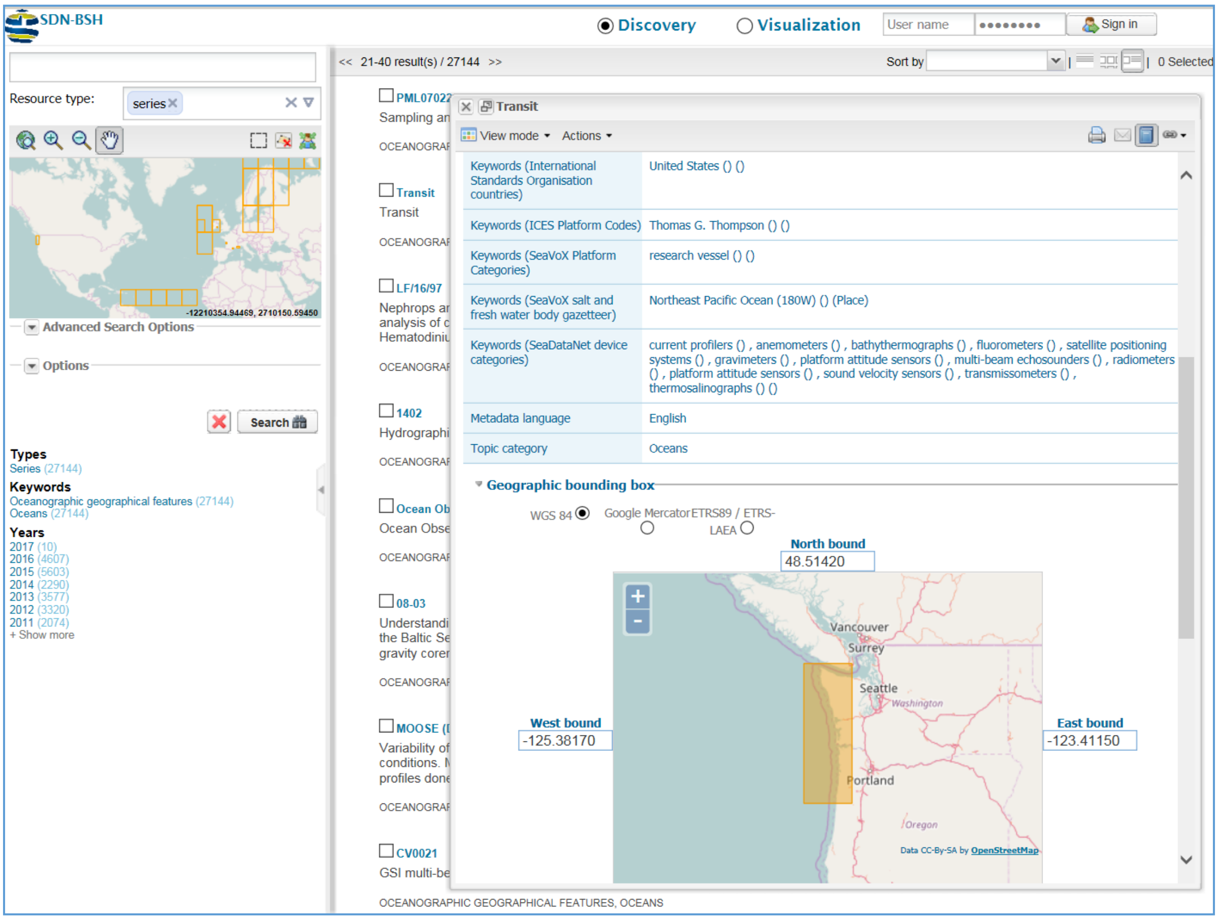


Figure 4 BSH GeoNetworks CSR portal

In addition, ICES also has a CSR web service that allows BSH to harvest both the updated and upgraded ICES legacy and new CSR records.

The harmonisation between ICES and the CSR central inventory has recently been implemented. All the Australian, Canadian and US CSRs (delivered via the ICES web service) have been updated and includes in the central CSR inventory.

In addition, most of the other legacy CSR records could be matched to the ICES CSR reference numbers so that in the future it should be possible to update these records for inclusion in the POGO and SeaDataNet portals. All of these records now have ICES as collating centre and the ICES reference number as local CSR ID in the CSR database. Those records ICES which could not be matched have been set the "deprecated" flag so that these no longer on the CSR homepage.

It is not currently feasible to automatically synchronise the ICES CSRs via the web service because ICES still uses the old CSR format which lacks many of the mandatory metadata fields defined in the ISO19115-2 standard used in the updated SeaDataNet CSR schema. The aim for the future is to have an automatic process for the harmonisation of CSRs with those from ICES. This activity will also be supported by Fisheries and Oceans Canada (FOC) that has been working closely with ICES to update all Canadian CSRs. FOC has also indicated interest in joining the CSR harvesting process in the future.

Vocabularies

The primary requirement for the ODIP 2+ prototype development task is the extension of the content of the SeaDataNet C17 platform vocabulary to cover all vessels of interest to the US R2R project plus all cruises in the CSR database (including V0 records) operated by BSH. Existing coverage of Australian research vessels is considered to be adequate for current requirements. This work also supports the move towards full synchronisation between the C17 vocabulary and the ICES platform code system.

Historically, there were three metadata systems for platforms. One was operated by ICES that issued platform identifiers known as 'ICES codes'. The other two were run by different groups in the US NODC that issued identifiers known as NODC codes and WOD codes.

In 2004 it was agreed to rationalise the various systems through full harmonisation of the ICES and NODC codes including a mapping to the WOD codes. This has been a work in progress ever since with significant progress having been made, but there is still work to be done. The work is straightforward for most platforms, but becomes much more difficult in cases where several platforms have carried the same name at different times. In such cases it is not uncommon to find equivalent codes in the different systems referring to different platforms of the same name.

Resolving problems like this requires significant amounts of work to research the histories of the platforms involved and then careful co-ordination between the organisations involved is required to modify the vocabularies and then propagate the changes through into the legacy data. It should be noted that these issues were also not confined to obscure vessels, for example, an issue with the primary UK research vessel RRS Discovery has arisen due to this vessel name have been used several times for different ships. The issues were further complicated by weak content governance in the past allowing codes to be assigned to entities that were not platforms, such as fixed stations including ferry routes, and a significant number of typographical errors.

In 2005 SeaDataNet took the pragmatic decision to create the C17[2] vocabulary as a subset of ICES ship codes that had been carefully checked and sufficiently researched to be sure that any issues such as multiple platforms of the same name had been resolved, and had sufficient metadata attributes to ensure unique identification was possible . Over the past decade, this subset has been growing steadily on an 'as required' basis. Typically, if during the preparation of a Cruise Summary Report a vessel was identified as missing from C17 then BODC, who manages the vocabulary, worked with the individual concerned to research the platform, add metadata into the ICES system (automatically migrated to C17) and accept the code into C17 by changing its status from 'deprecated' to 'accepted'. If the vessel was not already in the ICES system it was then added.

The ODIP 2+ prototype development task is extending the work described above to include the R2R ship list and also a list of ships provided by partner BSH. At the start of the ODIP II project there were nearly 3000 ship codes for US vessels in the ICES system for which the only identification metadata attribute was the name. Of these over 100 were identified as research vessels in the broadest sense, which included vessels associated with academic institutions, NOAA ships and naval hydrographic and oceanographic survey vessels. These ships were prioritised in a project covering all US vessels and have now been added to the C17 vocabulary together with more than 1000 US warships.

[2] http://vocab.nerc.ac.uk/collection/C17/current/

A list of 40 vessels currently missing from the BSH system has been generated with a quarter of these having already been resolved. Work is continuing on the other 30, but these are the most difficult cases with either known issues or very little available information and so progress will be slow.

The work to align the platform vocabularies has been significantly assisted by the development of tools to automate the comparison of metadata held in the ICES and BODC systems. This approach uses a simple schema in the BODC Oracle system comprising two tables with identical structures. One of these tables is populated from the C17 vocabulary via a call to a PL/SQL procedure. The other is populated by a Java application that issues service calls to the ICES system. Once populated the two tables can easily be compared using SQL queries to identify any discrepancies. To date this system has been used to identify ICES codes missing from C17, and to generate reports of ICES codes that BODC recommend for deprecation in the ICES system, such as stations and ferry routes.

2.3 ODIP 3+ Prototype Development Task

The ODIP 3+ prototype development task, which is continuing the activities undertaken by ODIP 3 in the previous project, is addressing the further elaboration of Sensor Web Enablement (SWE) standards and services in concertation with many on-going regional projects as well as the international Open Geospatial Consortium (OGC) standards body. It will build upon the earlier results and explore further enhancements and new technological approaches with an emphasis on establishing synergy and alignment between ongoing regional projects addressing the topic of SWE.

The ODIP 3+ Prototype is led by 52°North with contributions expected from partners in the three participating regions. Activities have been undertaken or are planned for each of the targets of the ODIP 3+ prototype development task:

2.3.1 Evaluation of Sensor Web Technologies for Marine Applications

The aim of this sub-task is to evaluate different Sensor Web technologies, including the OGC Sensor Web Enablement (SWE) standards for facilitating the interoperable sharing of oceanographic observation data and metadata.

SWE Technologies

The NeXOS project[3] (EU FP7, with participation by 52°North and IFREMER) is developing a full Sensor Web architecture and implementation for marine sensors. This comprises the sensor/platforms which are equipped with a firmware that supports the plug-and-play integration of sensors based on SensorML encoded metadata, as well as the interoperable publication of observation data. These activities have resulted in several developments that have also contributed to the ODIP 3+ prototype development task including:

- A Sensor Web Viewer for stationary in-situ measurements

[3] http://www.nexosproject.eu/

- Extension of the Sensor Web Viewer to display data collected by mobile sensors/sensor platforms (see Figure 5)
- Several SOS deployments by NeXOS partners which can be used for interoperability testing and demonstration purposes
- SOS server supporting optimised data encodings for publishing sensor data (EXI encoding of the SOS Result Handling operations)

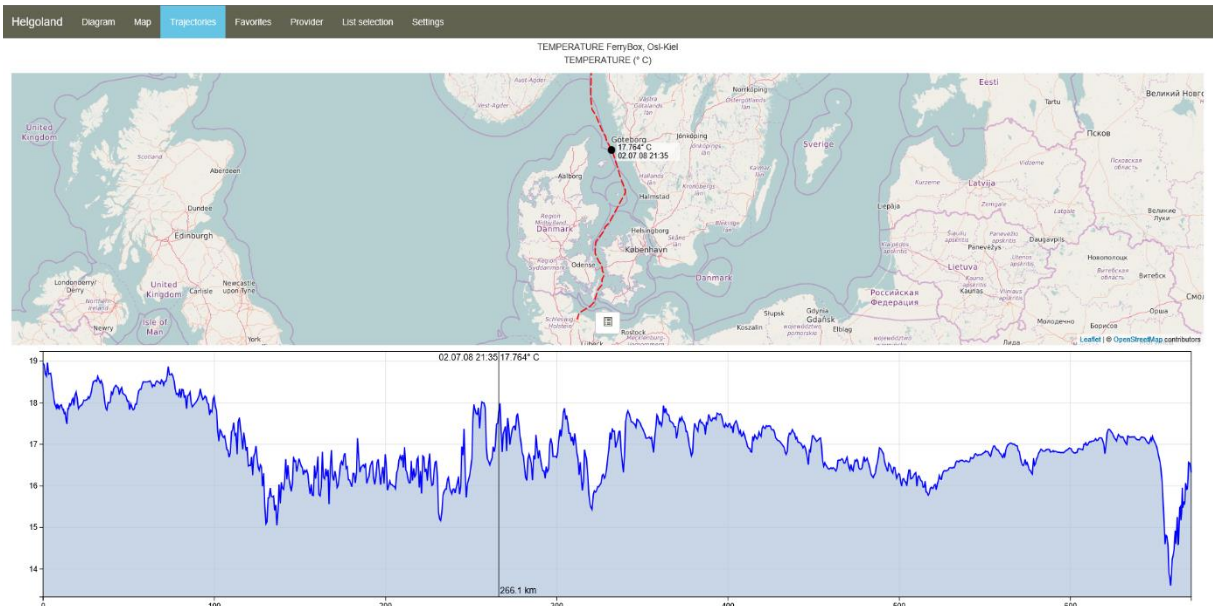


Figure 5 NeXOS Sensor Web Viewer for mobile sensor platforms

The developments of the NeXOS project were also further enhanced and adjusted to visualise observation data from different fixed ocean observatories within the FixO^3 project[4] - an EU FP7 project which included participation by 52°North, IFREMER and other ODIP II partners (see Figure 6 below).

The SOS deployments are currently in progress so further SOS servers are likely to be available for demonstration of the ODIP 3+ prototype solution during the next reporting period.

Furthermore, both the NeXOS and FixO^3 activities were closely linked with the Marine Sensor Web Profile and SensorML metadata editor developments.

[4] http://www.fixo3.eu/

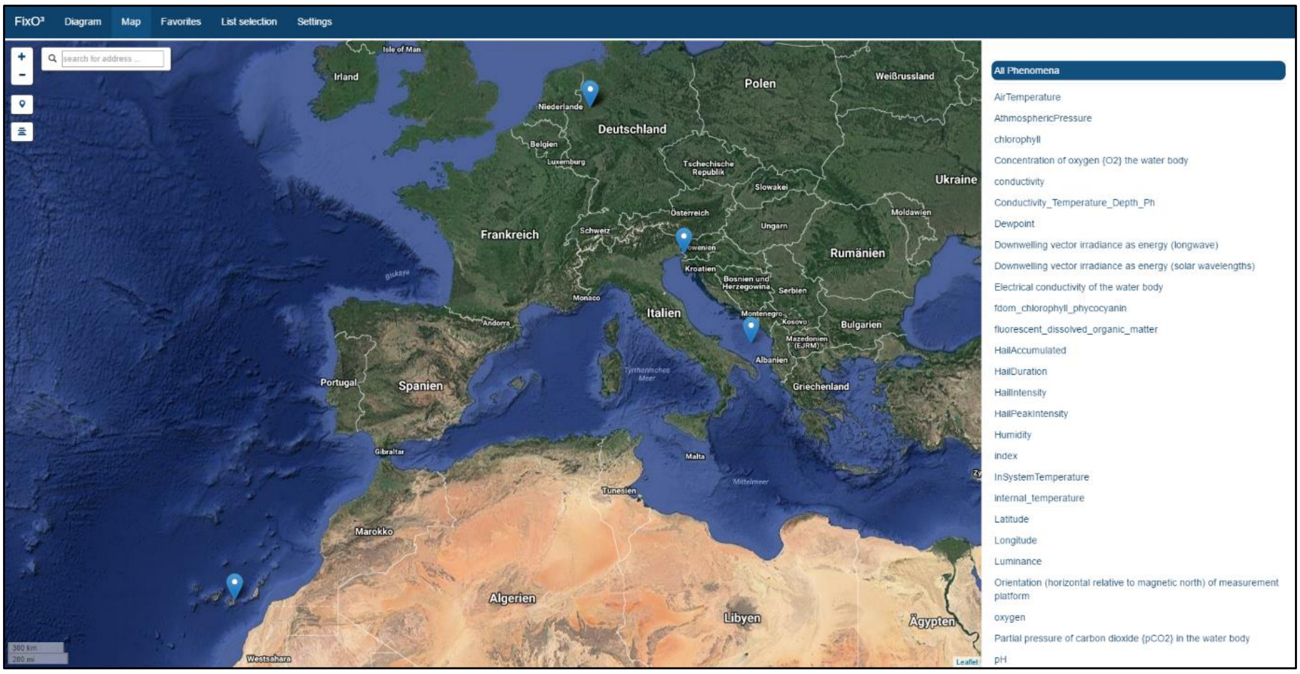


Figure 6 Sensor Web Viewer (FixO³ and NeXOS)

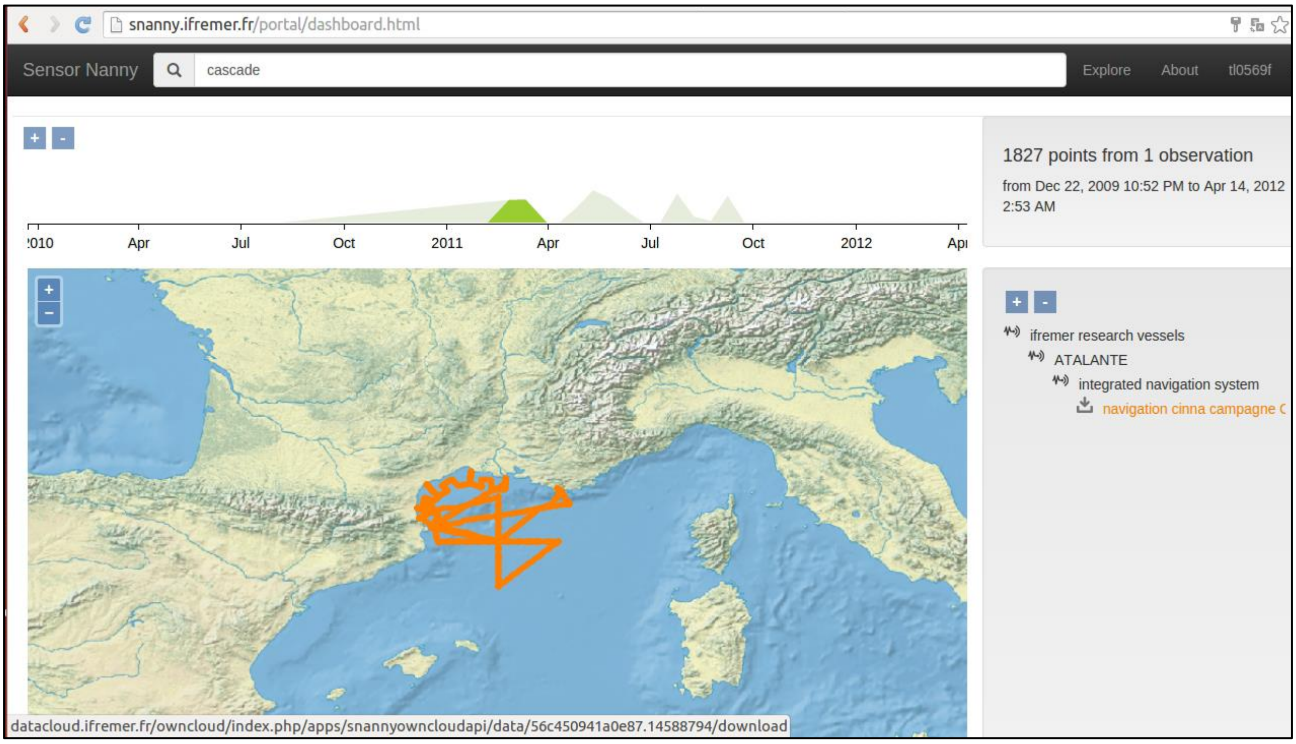


Figure 7 Sensor Nanny tool for data discovery

Sensor Nanny

The Sensor Nanny[5] software developed by IFREMER also needs to be mentioned in the context of sensor web enablement (SWE) technologies. It offers a comprehensive set of tools for managing, publishing, discovering and visualising marine observation data and sensors (see Figure 7 above). It uses technologies such as OwnCloud (e.g. for sharing data via the cloud) and ElasticSearch for enabling discovery within the published data. A central feature is an easy mechanism for "dropping" observation data sets into a cloud infrastructure so that other researchers can access it. In addition, the Sensor Nanny software suite also includes a SensorML editor which is described in more detail below.

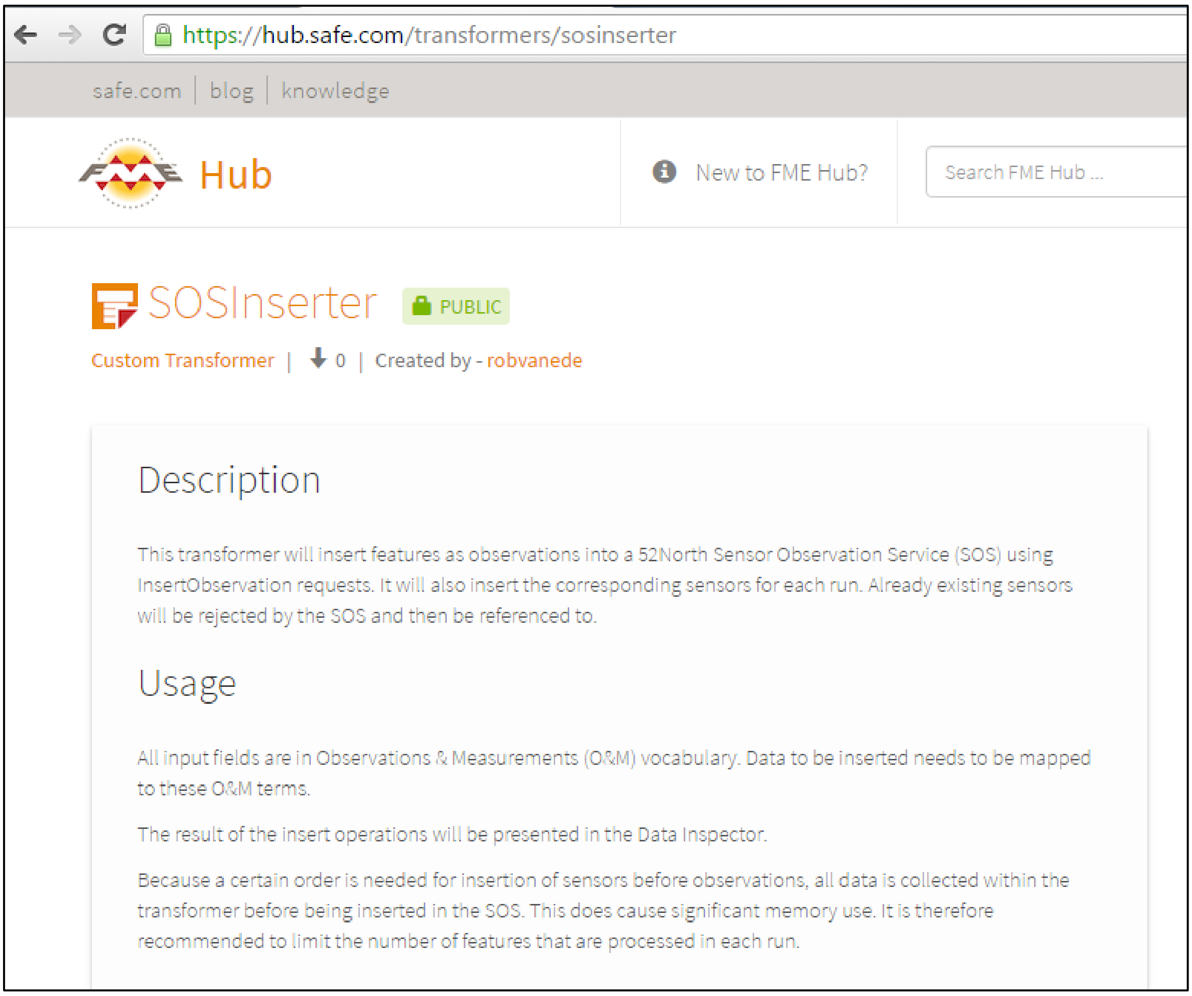


Figure 8 SOS inserter available on FME Hub

Another tool to be noted is the SOSInserter which has been developed by TNO as part of the ODIP activities. It is an extension of the FME (Feature Manipulation Engine) from Safe Software, a widely used tool for converting data sets between an extremely large range of source and target data formats. Based on an intuitive user interface, the FME allows users to describe the transformations they want to perform. With the SOSInserter Plug-In, the FME has

[5] http://snanny.ifremer.fr/dashboard.html

the capability to publish observation data sets via the transactional SOS operations on instances of the 52°North SOS. The availability of this tool, which is available for download via the FME Hub (see (see Figure 8), will lower the barriers for owners of existing observation data to make their data available via the Sensor Web.

Marine Sensor Web Profiles

The broad range of activities using SWE standards for marine observation data and systems leads to a risk of incompatible approaches for how the SWE specifications are applied. The OGC SWE standards were intentionally designed to be domain independent so that they can be applied in as many scenarios as possible. This idea goes hand in hand with a high degree of flexibility within the specifications so that the same goals could be achieved in different ways.

To avoid interoperability issues, a common approach to the application of SWE specifications for marine data and sensors is needed. Furthermore, this should be complemented by the use of vocabularies to ensure not only common syntax but also common semantics.

Consequently, ODIP II has created a joint activity with partners from several other relevant projects and initiatives such as: AODN, BRIDGES, ENVRIplus, EUROFLEETS/ EUROFLEETS2, FixO[3], FRAM, IOOS, Jerico/Jerico-Next, NeXOS, RITMARE, SeaDataNet/ SeaDataCloud, SenseOcean and X-DOMES.

The previous ODIP 3 prototype development task had a strong focus on collecting usage examples of SWE specifications in different projects and organisations. These exemplars, which include not only examples of O&M and SensorML but also endpoints of available SOS instances, have been collated on a wiki hosted by partner 52°North as shown in Figure 9.

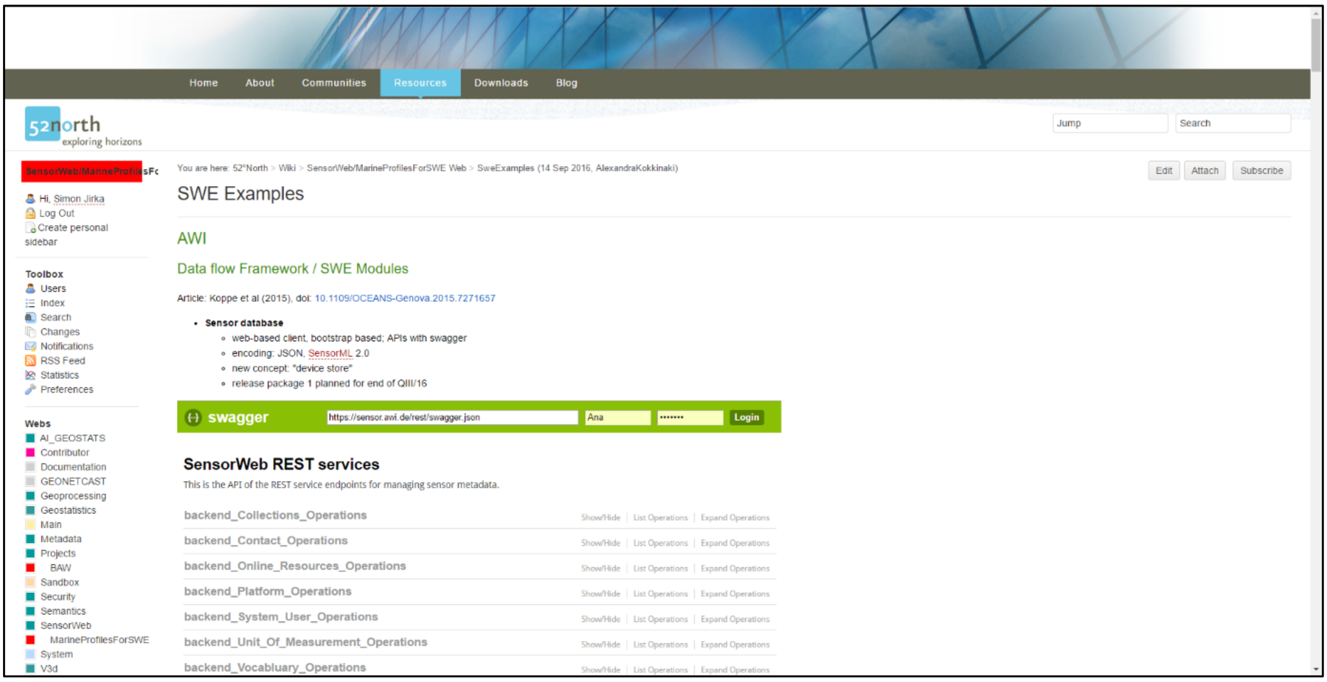


Figure 9 Marine SWE profiles wiki

In addition, an initial model for how a SensorML profile for marine sensors could be structured was developed. This has resulted in a hierarchical approach that distinguishes between:

- Sensor platforms (e.g. vessels, gliders, buoys)
- Instruments (attached to sensor platforms)
- Detectors (as component of an instrument)

However, depending on the individual organisation, only selected layers of the three levels above might be used. Furthermore, the proposed profile distinguished between the descriptions of types and instances. This means that a manufacturer could provide a description of a sensor type (e.g. through the ESONET/FixO3 Yellow Pages) while a sensor operator would only have to provide the specific information of the sensor instance of this type that is deployed by the operator. Figure 10 shows an exemplar overview of the possible elements that describe an instance of an instrument.

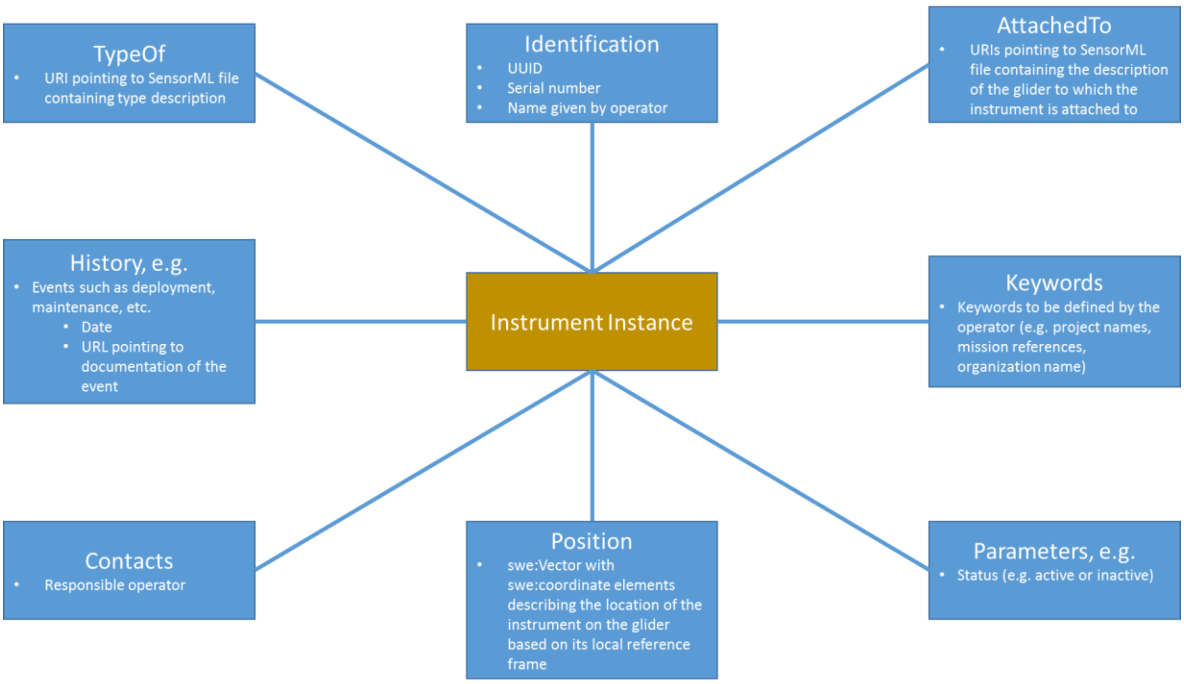


Figure 10 Exemplar for the development of a marine SensorML profile (Description of an Instrument Instance)

Analysing the handling of large volumes of data within SWE-based infrastructures

Another challenge that is relevant to the field of sensor web in the marine domain is the handling of large volumes of observation data.

During the first phase of ODIP II the main focus of this topic was on identifying open issues and questions that included:

- How large heterogeneous spatio-temporal datasets can be organized, managed, and provided to Sensor Web applications
    - What are the typical request scenarios of observation data for search, download, visualization and processing?
    - Are current Sensor Web standards capable of and suitable for handling massive observation data sets?
- How views on big data sets and derived information products can be made accessible in the Sensor Web?
    - Which service standards are appropriate? SOS would be the most obvious choice, but are WPS, WCS/WCPS also suitable options?
    - Which conceptual models and encodings, e.g. O&M or NetCDF, are appropriate?
- How can big observation data sets be processed efficiently?
    - How does the underlying storage structure influence performance?
    - How does the WPS handle situations in which transferring datasets is hard to achieve?
    - Can the WPS be used as a Rich-Data-Interface for big observation databases?
    - How can predefined, parameterized or even interactive analyses be realized?
    - How could a query language that enables on-demand analysis of time series data look like?
    - How could a combined analysis of multiple datasets of different origins be accomplished with such high volumes of data?

During the second period of ODIP II these questions will be investigated by prototypical implementations in a cooperation between partners 52°North and AWI.

2.3.2 Use of lightweight complementary technologies such as JSON and REST

In addition to the XML-based SWE standards, there are further technologies such as JSON and REST which allow more lightweight implementations for handling observation data.

The potential advantages of these technologies such as reduced data volumes as well as easier data processing on client platforms, which may potentially be resource constrained, are being investigated as part of the ODIP 3+ prototype development task. Two activities are currently being undertaken:

1) The work of CSIRO in particular shows how marine observation data infrastructures could benefit from the use of these lightweight technologies.

   In order to investigate the question of how this could also be mapped to the OGC Sensor Web Enablement standards, a mapping of the concepts behind the Marine SWE profile specification to JSON and REST will be developed by 52°North. However, as the Marine SWE profile is still in development, this activity will be conducted during the next reporting period of the ODIP II project (M18 – M36).

2) Establishing of a link between ESRI's ArcGIS for Server GeoEvent Extension. In this context, 52°North and ESRI are currently cooperating to investigate how data from Sensor Web services (i.e. SOS instances) could be consumed by the GeoEvent Extension in order to detect relevant events (e.g. critical wind speeds measured by NOAA buoys) within sensor data streams. This development is currently ongoing with initial results expected by the end of 2016.

2.3.3 Use of RDF-based approaches to support discovery of marine sensors and data sets

Several partners, including BODC which took a lead role, have worked on the creation of an overview of terms that are needed for inclusion in SWE documents in conjunction with the process to define the Marine Sensor Web Profiles. The resulting list of terms has been documented in the Marine SWE Profiles Wiki (see Figure 11 below).

Identifier	Preferred Label	Definition
	Delivery	Instrument or platform is delivered to the owner
	Deployment	Placement of an instrument or platform to where it will make measurements
	Calibration	~~Calibration is a method of improving instrument performance by removing uncertainty or bias in the sensor outputs~~ UPDATED 'An event in the life of a sensor or instrument in which the coefficients used to convert its outputs into engineering units, or improve its accuracy, are determined and applied.'
	Recovery	Return of an instrument or platform from where it has made measurements
	LoanStart	Instrument or platform is supplied to a third-party for a limited period of time
	LoanEnd	Instrument or platform is returned by third party
	RepairStart	Instrument or platform is taken out of service for correction of a malfunction
	RepairEnd	Instrument or platform is returned to service following correction of a malfunction
	Failure (instead PartialFailure, TotalFailure?)	Unexpected malfunction of instrument or platform.
	Maintenance	Instrument or platform is under maintainance
	Loss	Instrument or platform is lost
	Comissioning	
	Decommissioning	Instrument or platform is withdrawn from service
	Repair	Instrument or platform is under repair
	Upgrade	Raise (instrument or platform) to a higher standard, in particular improve (instrument or platform) by adding or replacing components.
	Calibration	
	PartialFailure	
	TotalFailure	

More SensorML terms under different sections

Comment (Robert Huber)on the below Capabilities terms:

Imho it should be avoided to define quality constraints of sensor output within the capabilities section, this needs to be placed at the output section for each output separately. -- RobertHuber - 2016-06-20

SECTION	ATTRIBUTE	CODVAL	CODNAM	CODALT	DEFINITION
Capabilities	Accuracy	CAPB	Accuracy		The maximum difference that exists between an actual value and an indicated value output from a sensor. The actual value is usually measured by a reference standard.
Capabilities	Dambing Ratio	CAPB	Damping ratio		Dimensionless measure describing how oscillations in a system decay after a disturbance. Many systems exhibit oscillatory behavior when they are disturbed from their position of static equilibrium
Capabilities	Frequency	CAPB	Frequency		The number of cycles or events of a waveform.

Figure 11 Extract from the Overview of Collected Terms for Marine Sensor Web Profiles

This activity brings significant added value to the OGC SWE standards which are mainly focused on syntactic interoperability. The use of vocabularies as part of the marine SWE profiles will make it possible to ensure that different organisations refer to the same concepts, parameters, values, etc. through the use of the same vocabulary terms.

Furthermore, this approach makes it possible for more powerful semantic web technologies to be applied in the future. The work being done on vocabularies by BODC is described more fully below.

BODC SWE Vocabularies

The flexibility of SensorML can result in many different variations of sensor description that leads to a reduction in interoperability and discoverability via the web. To resolve this issue, it is important to bring together potential user communities, identify lists of required terms, define them and then use controlled vocabularies to publish them according to selected standards.

To achieve this objective, the SWE Marine Profiles group was created by partners from several projects and initiatives who joined forces to develop common marine profiles of OGC SWE standards that can be used in multiple projects and organizations. The SWE Marine Profiles mailing list is essentially a discussion list, where members are allowed to post their own items which are then distributed to the entire mailing list. For the purposes of vocabulary building, this list was given the responsibility of acting as the SensorML vocabulary content governance body, which is important in order to stay up-to-date and aligned with ongoing developments in this field.

The publication of SensorML implementations by different projects revealed the lack of published vocabularies for terms and properties definitions, and the need for common vocabularies to refer to the same terms coherently in the marine domain. There are essentially two sections in SensorML that would benefit from the use of vocabularies: the “term definition” and the “term value”.

For “term values”, SWE Marine Profiles members agreed to use existing concepts available in NERC Vocabulary Server version 2.0 (NVS2.0) http://vocab.nerc.ac.uk/). The following vocabularies were identified as adequate for providing the term values:

- Observable property: NVS2.0 Collections: P01 that lists terms used to describe individual measured phenomena, and P07 that lists the Climate and Forecast standard names, have been nominated for use as SensorML observable properties
- Instrument Type: NVS2.0 Collection L05: lists device categories. It is used for the classification of instruments and procedures.
- Platform Type: NVS2.0 Collection L06: list of platform categories that are to be used for classifying platforms
- Roles: NVS2.0 Collections G04, C86 list roles and are used to populate the SensorML role property
- Feature of Interest: NVS2.0 Collection C19 is the Salt and Fresh Water Body Gazetteer that can be used to create a rich list of features of interest
- Manufacturer: NVS2.0 Collections L35, C75: can both be used to populate the manufacturer property, since they refer to organisations and manufacturers respectively.

SensorML consists of sections which include several terms. In NVS2.0, each section is modelled as a new vocabulary with a unique URI that lists a set of domain relevant terms.

Following the NVS2.0 URL pattern, SensorML vocabularies are all grouped under the 'W0X' notation as shown in Table 2.3.1, although there is no semantic relevance between the vocabulary's subject and the notation. Each vocabulary is self-documented and refers to the SWE Marine Profiles group as its creator and owner. BODC is the manager and moderator and NERC is the publisher.

URI	Title
http://vocab.nerc.ac.uk/collection/W03/current/	SensorML History Event Types
http://vocab.nerc.ac.uk/collection/W04/current/	SensorML Capability Section Terms
http://vocab.nerc.ac.uk/collection/W05/current/	SensorML Characteristic Section Terms
http://vocab.nerc.ac.uk/collection/W06/current/	SensorML Classification Section Terms
http://vocab.nerc.ac.uk/collection/W07/current/	SensorML Identification Section Terms
http://vocab.nerc.ac.uk/collection/W08/current/	SensorML Contact Section Terms

Table 3 List of URIs and descriptions for the published SensorML collections

As part of the SenseOCEAN[6] project BODC is working on the standardisation of sensor metadata enabling 'plug and play' sensor integration. The approach combines standards, controlled vocabularies and persistent URIs (Uniform Resource Identifiers) to publish sensor descriptions, their data and the dataset metadata as 5 star Linked Data and OGC SWE (SensorML, Observations & Measurements) implementations (see Figure 12 and Figure 13). Sensors can therefore become more discoverable, accessible and useable via the web and their data can be combined with other sensor or Linked Data datasets to create knowledge.

As part of the SenseOCEAN project sensor metadata and data are encoded using OGC standards with the sensors and instruments transmitting unique resolvable web links to the persistent OGC SensorML records that are published by BODC. The URIs, based on content negotiation, resolve to either Linked Data or Open Geospatial Consortium Sensor Web Enablement (OGC SWE) descriptions. Sensor data, which is either observations or sensor metadata, are also delivered through a SPARQL endpoint and a Sensor Observation Service (SOS) server allowing them to be more discoverable and accessible via the web.

To support this service, the sensor owners must register their sensors on the web in order to receive a unique URI for each one. As shown in Figure 14 and Figure 15, users can add specific information for their sensor instances along with event information such as deployment, calibration installation etc.

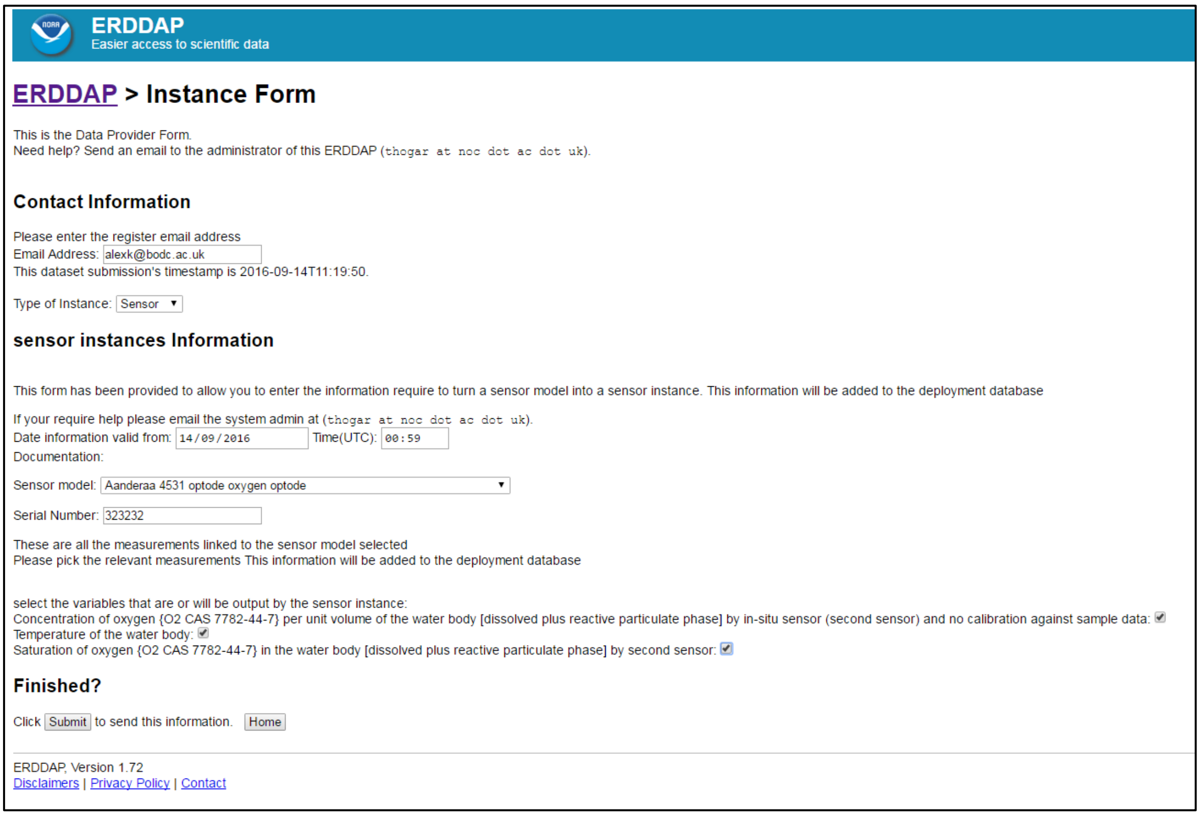
[6] http://www.senseocean.eu/

```
- <sml:PhysicalSystem xsi:schemaLocation="http://www.opengis.net/swes/2.0 http://schemas.opengis.net/swes/2.0/swesDescribeSensor.xsd http://www.opengis.net/sensorml/2.0 http://schemas.opengis.net/sensorML/2.0/sensorML.xsd
  http://www.isotc211.org/2005/gmd http://schemas.opengis.net/iso/19139/20070417/gmd/gmd.xsd http://www.isotc211.org/2005/gco http://schemas.opengis.net/iso/19139/20070417/gco/gco.xsd http://www.opengis.net/gml/3.2 http://schemas.opengis.net
  /gml/3.2.1/gml.xsd" gml:id="TOOL0969_102">
    <gml:description>TOOL0969_102</gml:description>
  - <gml:identifier codeSpace="uniqueID">
      http://linkedsystems.uk/system/instance/TOOL0969_102/
    </gml:identifier>
  - <sml:keywords>
    - <sml:KeywordList>
        <sml:keyword>O2Sat_2</sml:keyword>
        <sml:keyword>Temp</sml:keyword>
        <sml:keyword>WC_dissO2_uncalib_2</sml:keyword>
      </sml:KeywordList>
    </sml:keywords>
  - <sml:validTime>
    - <gml:TimePeriod gml:id="validityPeriod">
        <gml:beginPosition>2016-05-25</gml:beginPosition>
        <gml:endPosition>Indeterminate</gml:endPosition>
      </gml:TimePeriod>
    </sml:validTime>
  - <sml:documentation>
    - <sml:DocumentList>
      - <sml:document xlink:arcrole="http://sensorml.com/ont/swe/property/OperationsManual">
        - <gmd:CI_OnlineResource>
          - <gmd:linkage>
            - <gmd:URL>
                http://www.aanderaa.com/media/pdfs/oxygen-optode-4531.pdf
              </gmd:URL>
            </gmd:linkage>
          - <gmd:name>
            - <gco:CharacterString>
                Aanderaa 4531 optode oxygen optode OperationsManual
              </gco:CharacterString>
            </gmd:name>
          - <gmd:description>
            - <gco:CharacterString>
                Aanderaa 4531 optode oxygen optode OperationsManual
              </gco:CharacterString>
            </gmd:description>
          </gmd:CI_OnlineResource>
        </sml:document>
      - <sml:document>
```


Figure 12 Sensor instance defined by URI in SensorML


```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:arpfo="http://vocab.ox.ac.uk/projectfunding#" xmlns:dcat="http://www.w3.org/ns/dcat#"
 xmlns:dcterms="http://purl.org/dc/terms/" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:geo="http://www.opengis.net/ont/geosparql#"
 xmlns:obs="http://def.seegrid.csiro.au/isotc211/iso19156/2011/observation#" xmlns:obs-prop="http://environment.data.gov.au/def/op#" xmlns:odo="http://ocean-data.or
 xmlns:schema="http://schema.org#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:prov="http://www.w3.org/ns/prov#" xmlns:qudt="http://qudt.org/schema/qudt#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:sf="http://www.opengis.net/ont/sf#" xmlns:skos="http://www.w3.org/2004/02/skos/core#"
 xmlns:time="http://www.w3.org/2006/time#" xmlns:vcard="http://www.w3.org/2006/vcard/ns#" xmlns:ssn="http://purl.oclc.org/NET/ssnx/ssn#"
 xmlns:gr="http://purl.org/goodrelations/v1#" xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:WM30="http://purl.oclc.org/NET/ssnx/meteo/WM30"
 xmlns:lso="http://linkeddev.bodc.ac.uk/ontology/" xmlns:DUL="http://www.loa-cnr.it/ontologies/DUL.owl#">
 ▼<ssn:SensingDevice rdf:about="http://linkedsystems.uk/system/prototype/TOOL0969/">
    <rdf:type rdf:resource="http://purl.org/goodrelations/v1#ProductOrServiceModel"/>
    <rdf:type rdf:resource="http://http://schema.org/ProductModel"/>
    <rdf:type rdf:resource="http://purl.org/goodrelations/v1#Entity"/>
    <rdf:type rdf:resource="http://www.w3.org/ns/prov#Entity"/>
    <rdf:type rdf:resource="http://www.w3.org/ns/prov#Agent"/>
    <dcterms:created>2016-05-10</dcterms:created>
    <prov:qualifiedAttibution rdf:resource="http://linkedsystems.uk/system/attribution/TOOL0969/1"/>
    <prov:qualifiedAttibution rdf:resource="http://linkedsystems.uk/system/attribution/TOOL0969/2"/>
    <rdf:type rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/351/"/>
    <skos:notation rdf:resource="http://vocab.nerc.ac.uk/collection/L22/current/TOOL0969/"/>
    <gr:hasManufacturer rdf:resource="http://vocab.nerc.ac.uk/collection/L35/current/MAN0007/"/>
    <lso:hasPhysicalCharacteristic rdf:resource="http://linkedsystems.uk/system/physicalcharacteristic/TOOL0969/current/"/>
    <lso:hasPhysicalCharacteristic rdf:resource="http://linkedsystems.uk/system/physicalcharacteristic/TOOL0969/current/"/>
    <skos:altLabel>Aanderaa_4531_oxygen_optode</skos:altLabel>
    <ssn:implements/>
    <dcterms:subject>WC_dissO2_uncalib_2</dcterms:subject>
    <ssn:observes rdf:resource="http://vocab.nerc.ac.uk/collection/P01/current/DOXYUZ02/"/>
    <dcterms:subject>O2Sat_2</dcterms:subject>
    <ssn:observes rdf:resource="http://vocab.nerc.ac.uk/collection/P01/current/OXYSZZ02/"/>
    <dcterms:subject>Temp</dcterms:subject>
    <ssn:observes rdf:resource="http://vocab.nerc.ac.uk/collection/P01/current/TEMPPR01/"/>
    <DUL:includesObject rdf:resource="http://www.aanderaa.com/media/pdfs/oxygen-optode-4531.pdf"/>
    <skos:prefLabel>Aanderaa 4531 oxygen optode</skos:prefLabel>
    <dcterms:description>Aanderaa 4531 oxygen optode</dcterms:description>
    <dcterms:title>Aanderaa 4531 oxygen optode</dcterms:title>
    <ssn:hasMeasurementCapability rdf:resource="http://linkedsystems.uk/system/measurementcapability/TOOL0969/TEMPPR01/current/"/>
    <ssn:hasMeasurementCapability rdf:resource="http://linkedsystems.uk/system/measurementcapability/TOOL0969/DOXYOP01/current/"/>
  </ssn:SensingDevice>
  <gr:BusinessEntity rdf:about="http://vocab.nerc.ac.uk/collection/L35/current/MAN0007/"/>
</rdf:RDF>
```


Figure 13 Sensor model defined by URI in RDF


ERDDAP
Easier access to scientific data

ERDDAP > Instance Form

This is the Data Provider Form.
Need help? Send an email to the administrator of this ERDDAP (thogar at noc dot ac dot uk).

Contact Information

Please enter the register email address
Email Address: alexk@bodc.ac.uk
This dataset submission's timestamp is 2016-09-14T11:19:50.

Type of Instance: Sensor

sensor instances Information

This form has been provided to allow you to enter the information require to turn a sensor model into a sensor instance. This information will be added to the deployment database

If your require help please email the system admin at (thogar at noc dot ac dot uk).
Date information valid from: 14/09/2016 Time(UTC): 00:59
Documentation:

Sensor model: Aanderaa 4531 optode oxygen optode

Serial Number: 323232

These are all the measurements linked to the sensor model selected
Please pick the relevant measurements This information will be added to the deployment database

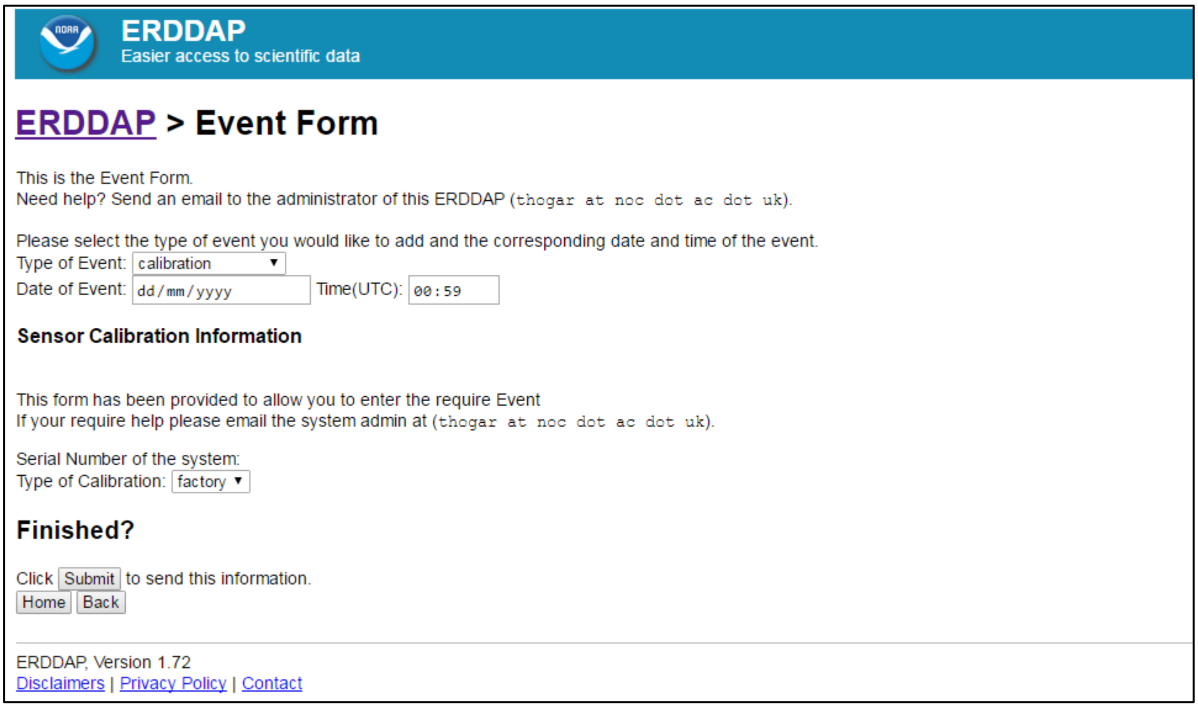
select the variables that are or will be output by the sensor instance:
Concentration of oxygen {O2 CAS 7782-44-7} per unit volume of the water body [dissolved plus reactive particulate phase] by in-situ sensor (second sensor) and no calibration against sample data: ☑
Temperature of the water body: ☑
Saturation of oxygen {O2 CAS 7782-44-7} in the water body [dissolved plus reactive particulate phase] by second sensor: ☑

Finished?

Click Submit to send this information. Home

ERDDAP, Version 1.72
Disclaimers | Privacy Policy | Contact

Figure 14 Registration form for an individual sensor (a "sensor instance")


ERDDAP
Easier access to scientific data

ERDDAP > Event Form

This is the Event Form.
Need help? Send an email to the administrator of this ERDDAP (thogar at noc dot ac dot uk).

Please select the type of event you would like to add and the corresponding date and time of the event.
Type of Event: calibration
Date of Event: dd/mm/yyyy Time(UTC): 00:59

Sensor Calibration Information

This form has been provided to allow you to enter the require Event
If your require help please email the system admin at (thogar at noc dot ac dot uk).

Serial Number of the system:
Type of Calibration: factory

Finished?

Click Submit to send this information.
Home Back

ERDDAP, Version 1.72
Disclaimers | Privacy Policy | Contact

Figure 15 Sensor event form

Synchronising efforts for metadata/SensorML Editors

While the development of the marine Sensor Web profiles addresses the content and structure of data and metadata files, the creation of the corresponding content presents a different challenge. The data measured by sensors can be inserted into OGC Observations and Measurements documents in an automated manner but the creation of the corresponding metadata is usually a much more labour intensive task. It will therefore be necessary to provide tools which make this process as easy as possible in order to encourage sensor manufacturers to provide the required comprehensive metadata. As a result, there are currently several projects that are addressing this topic but it should be noted that each of these initiatives has a different primary focus. However, they are still likely to gain significant benefit from further cooperation and alignment. An important common topic will be the use of the vocabularies as they are currently being enhanced in the context of the marine Sensor Web profile development activities described above. Using this approach the metadata files generated by the SensorML editors would rely on the use of common terms so that not only is syntactic interoperability achieved but also some degree of semantic interoperability.

The SensorML editors included in the ODIP 3+ prototype development tasks are outlined below:

Sensor Nanny

As described above, Sensor Nanny is a comprehensive framework of different components for dealing with marine observation data and metadata that is publicly available via GitHub (https://github.com/ifremer). One of the components of Sensor Nanny is a SensorML editor for describing sensor systems and their components (see Figure 16)

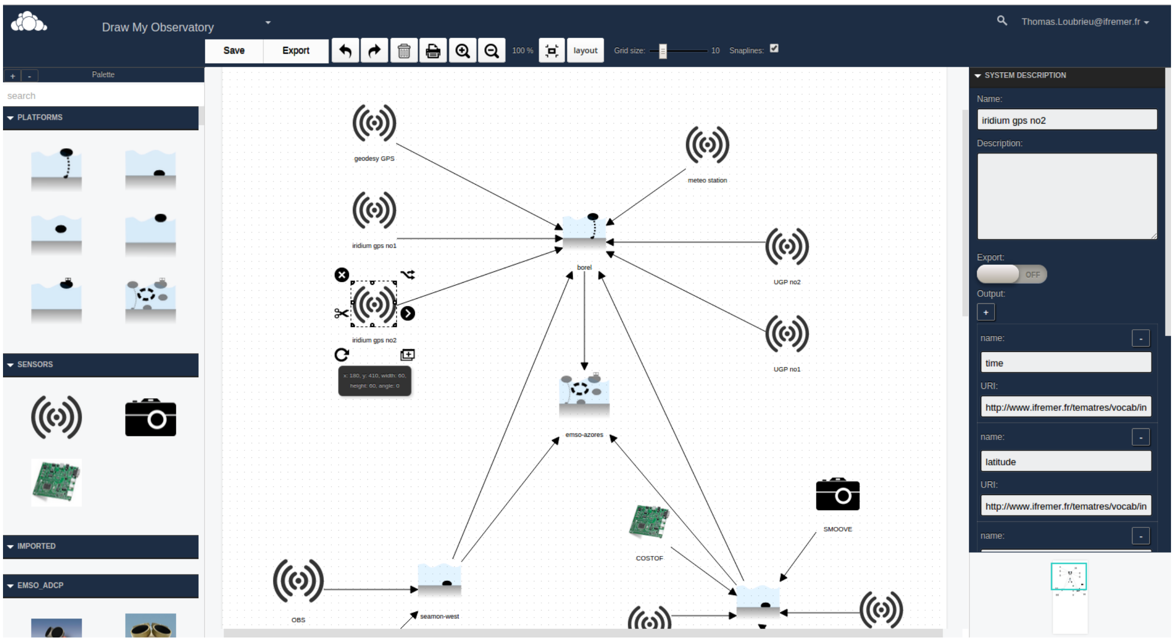


Figure 16 Sensor Nanny SensorML editor

Core features of Sensor Nanny comprise:

- Cloud support
- An intuitive graphical editor for describing/drawing the relationships between components and sensor systems
- A close link to the data publication mechanisms of Sensor Nanny
- Use of pre-defined system models from the $FixO^3$ Yellow Pages

RITMARE EDI-NG SensorML editor

EDI-NG is an HTML-based editor for sensor metadata, developed by CNR in cooperation with CSIC (see Figure 17 below). It is also openly available via GitHub: https://github.com/SP7-Ritmare/EDI-NG

Besides the HTML based client component, it also includes a separate server component that is used as the backend for the editor (e.g. for handling the XML creation). A key feature of EDI-NG is its high level of configurability. This means that it can be configured to different application scenarios using XML templates. These templates ensure a high level of flexibility so that users can define not only the required metadata fields but also add further constraints on permitted values, etc.

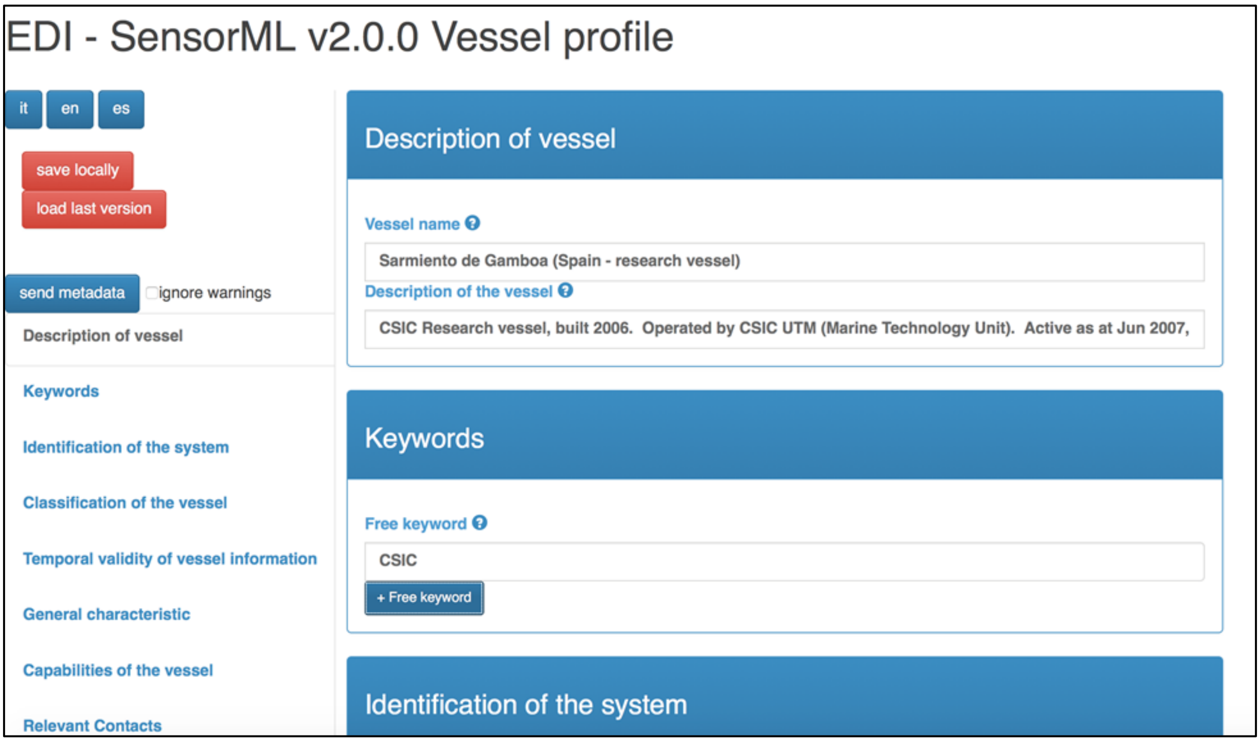


Figure 17 EDI SensorML editor

smle

The third SensorML editor contributing to the ODIP 3+ prototype development task is smle which is being developed by 52°North as part of the NeXOS and FixO3 projects (see Figure 18) . This open source client is also available via GitHub [https://github.com/52°North /smle](https://github.com/52°North/smle)

When compared to the other two examples of SensorML editors described above, smle has a stronger focus on integration with other Sensor Web Enablement components. For example, smle is able to use any SOS server for storing the SensorML files created, as long as it supports the transactional operations of the OGC SOS specification.

Core features of smle are:

- Supports the use of SOS servers for metadata storage
- Domain-independent (could also be used beyond marine sensors)
- Web-based implementation
- Supports the use of sensor type information from the ESONET/FixO^3 Yellow Pages

Current on-going activities are investigating how smle could be enhanced by a profile definition language that allows the configuration of the users client interfaces as well as the validation of content entered by users.

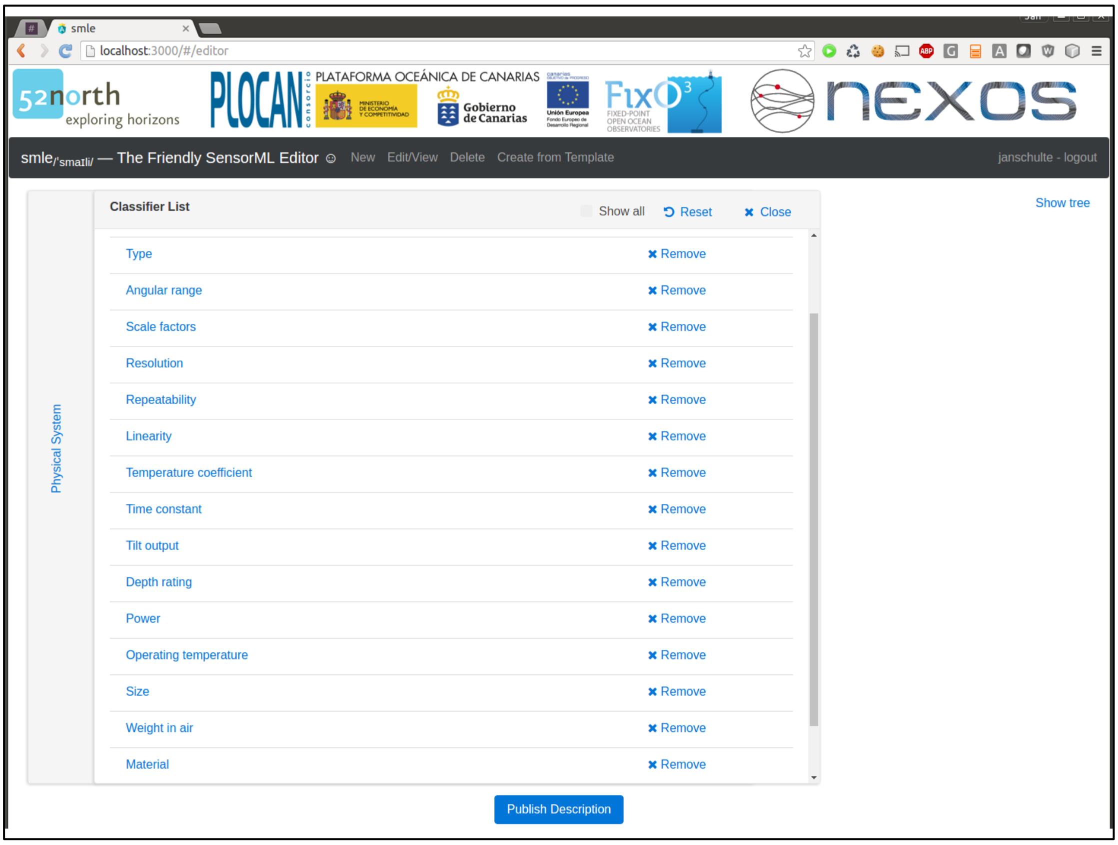


Figure 18 smle SensorML editor

A comparative table of the functionalities of the different XML editors that has been prepared by CNR and is included in this document as Appendix C.

2.4 Linking the Sensor Web and Global Infrastructures

Linking Sensor Web services to global infrastructures such as GEOSS is another key element of both Work Package 3 and the development of the ODIP II prototype interoperability solutions. The current approach is similar to that used by the FP7 GEOWOW[7] project for the hydrology domain. In general terms, it will rely on the marine Sensor Web profile (see above) to provide a common entry point to the marine Sensor Web services. This would allow GEOSS Common Infrastructure components such as the GEOSS Discovery and Access Broker (GEO-DAB) to harvest metadata from marine SWE services and to make these resources discoverable and accessible through the GEOSS portal. However, as the marine Sensor Web profile is still in development, this activity will be expanded further once it is completed.

2.5 Promoting Sensor Web Technology

The advances made by the ODIP 3+ prototype development task have been presented at several workshops and conferences in order to promote the adoption of the resulting advances in Sensor Web technology. The most significant of these events in the context of the ODIP 3+ prototype development activities are outlined below:

1) SWE Workshop at Oceanology International 2016

A half-day workshop at the Oceanology International 2016 Conference (London, UK; March 2016) was organised by the EU-funded Eurofleets2 project[8] in cooperation with a range of other European and US projects, including ODIP II. During this workshop several of the outcomes of the ODIP 3+ prototype were demonstrated including the vocabulary development work done by BODC. This was complemented with an introduction to the sensor web by 52°North and the Universitat Politècnica de Catalunya.

The workshop, which was attended by more than 60 people, aimed to initiate and promote dialogue between researchers and industry. It also demonstrated the need for continued dialogue and cooperation between the various projects and initiatives in the manner that is key to the approach being taken by ODIP II.

More information about this workshop including the agenda, presentations and conclusions, is available at: http://eurofleets.maris2.nl/swe-workshop/

2) EGU General Assembly 2016

The European Geosciences Union (EGU) General Assembly in Vienna, Austria during April 2016 also presented an opportunity for a dialogue between those working on the ODIP 3+ prototype development task and other experts and researchers that were also attending the conference. Presentations by the ODIP II partners at EGU included the vocabulary developments being undertaken by BODC, the work on Sensor Nanny by IFREMER,

[7] http://www.geowow.eu/

[8] http://www.eurofleets.eu/np4/home.html

development of marine Sensor Web profiles by 52°North, and the challenges of big data sets in the Sensor Web, also by 52°North.

`

3) 10th GEO European Projects Workshop (GEPW-10)

The Sensor Web activities in ODIP 3+ as well as other related ongoing projects were also presented by 52°North at the 10th GEO European Projects Workshop 2016 in Berlin, Germany during May 2016. The main focus of this presentation was to raise awareness of the Marine Sensor Web Profile development activities that are an important element of the ODIP 3+ prototype development task. The discussions during the workshop provided an opportunity to identify further potential contributors to the specification process for the Marine Sensor Web Profile. Further information on the GEPW-10 workshop including all presentations are available at: https://ec.europa.eu/easme/en/geo-european-projects- workshop-2016

4) 52°North Geospatial Sensor Webs Conference 2016

The 52°North Geospatial Sensor Webs Conference 2016 took place during August 2016 in Münster, Germany. It is a continuation of the previous series of Sensor Web workshops organised by 52°North which brought together around 60 participants from research, industry, and public administration.

A significant proportion of the conference programme addressed the topic of marine applications for Sensor Web technologies. A number of the presentations were therefore of direct relevance to the ODIP 3+ prototype development task such as the work done on vocabularies by BODC and Sensor Web tools by 52°North (see Figure 19). The resulting discussions also provided the opportunity to engage more widely with those involved in this field of research.

More detailed information about this event is available at: http://52°North.org/about/other-activities/geospatial-sensor-webs-conference

Figure 19 ODIP II presentation during the 52ºNorth Sensor Web Conference

3 Cross-cutting topics

It was agreed during the 1st ODIP II workshop that the cross-cutting topics identified during the previous project (i.e. vocabularies, data publishing/citation and persistent identifiers) would continue to be addressed as part of ODIP II**.** Below is a report on the progress and results of the work done on the vocabularies topic as part of ODIP II.

3.1 Vocabularies

A number of activities addressing different aspects of vocabulary development are currently on-going within the regions participating in ODIP II.

3.1.1 European vocabulary development activities

NVS search tool

The British Oceanographic Data Centre (BODC) has developed a search interface to make interrogation of the NERC Vocabulary Service (NVS) easier using a standard web browser. This tool enables a user to carry out simple or advanced searches either across or within vocabularies. The interface can be found at: https://www.bodc.ac.uk/data/codesandformats/vocabularysearch/ and is also illustrated in Figure 20 shown below.

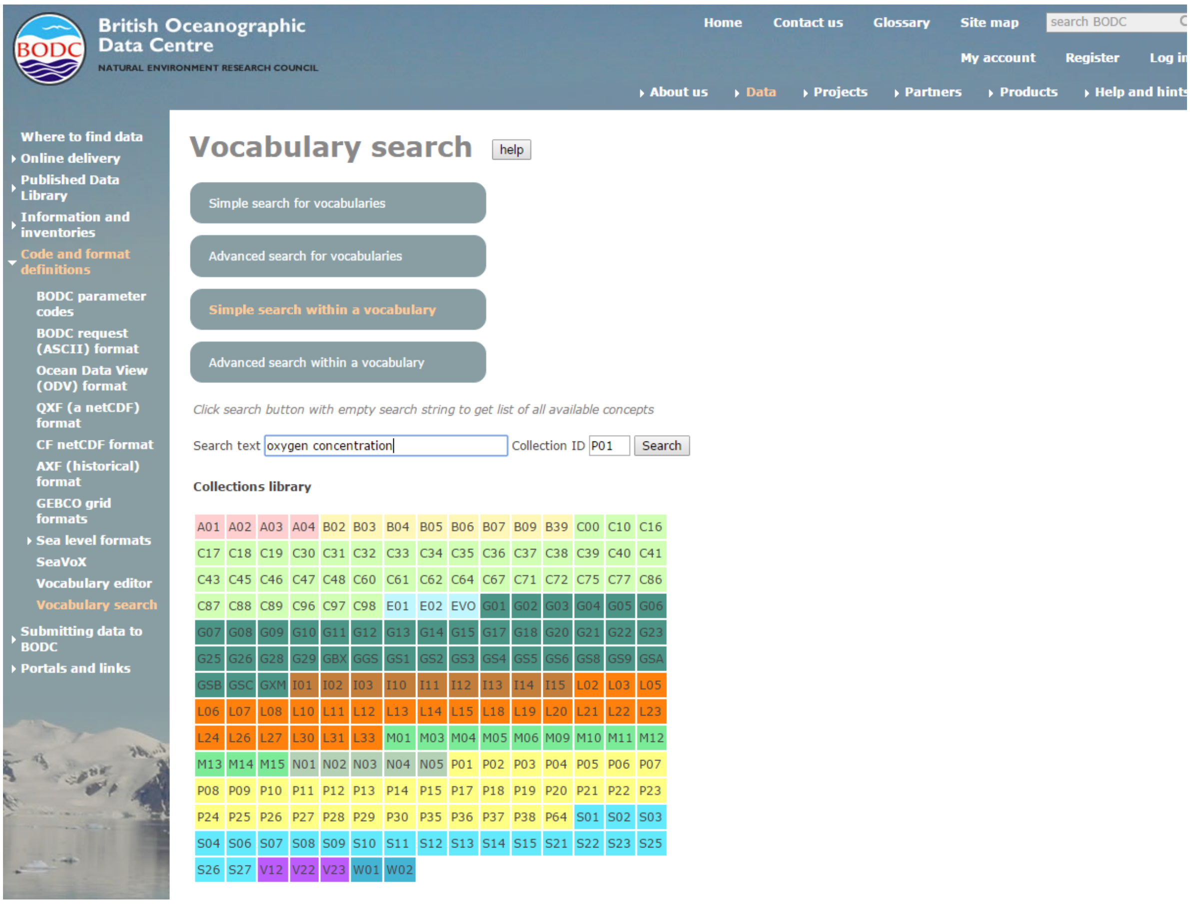


Figure 20 NVS Vocabulary Search tool showing "periodic" table of available content

Content negotiation

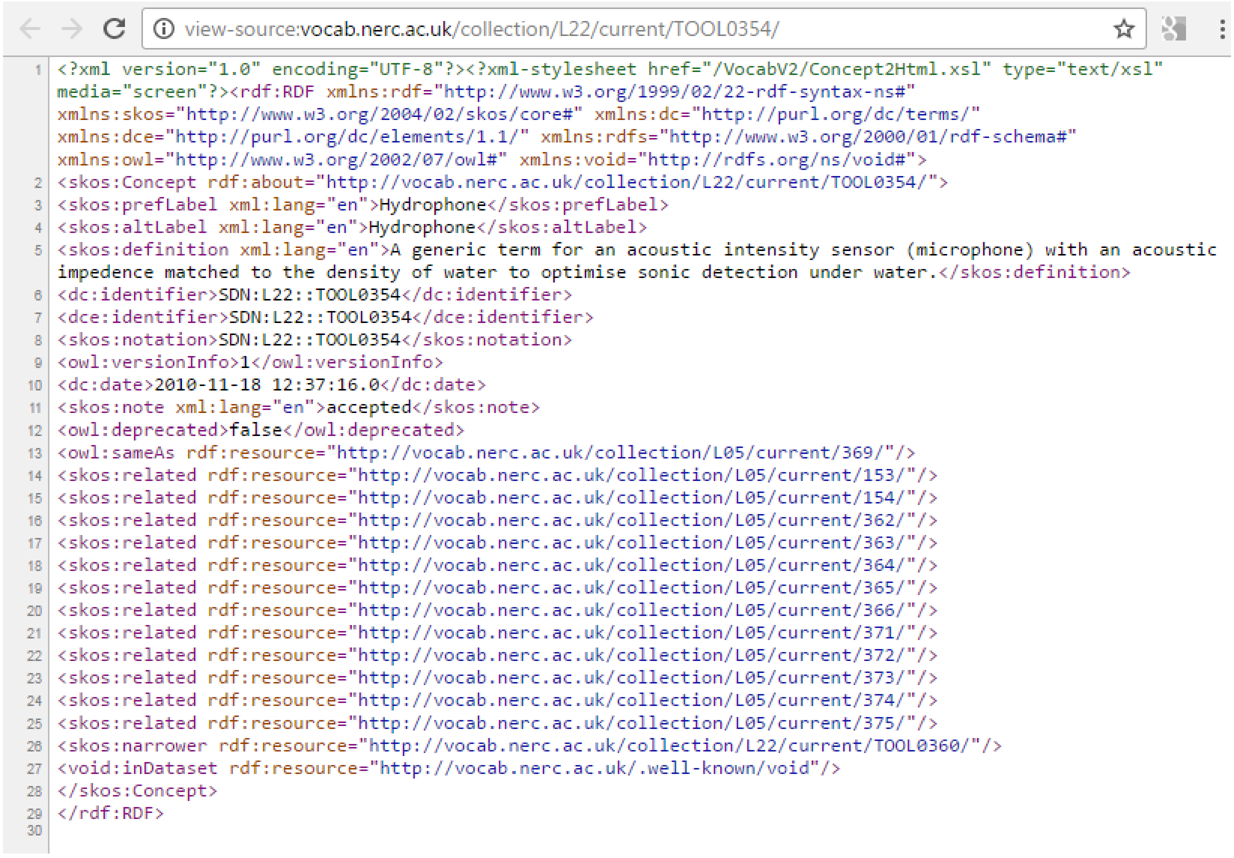
The results of queries submitted to the NVS 2.0 http://vocab.nerc.ac.uk/ are now delivered using content negotiation. The results of searches via a web browser are returned as html (Figure 21) while machine-to-machine interrogation of the NVS returns results as RDF-XML (Figure 22).

vocab.nerc.ac.uk/collection/L22/current/TOOL0354/

⇑ -- Hydrophone --

URI	http://vocab.nerc.ac.uk/collection/L22/current/TOOL0354/
Identifier ()	SDN:L22::TOOL0354
Preferred label (en)	**Hydrophone**
Alternative label (en)	Hydrophone
Definition (en)	A generic term for an acoustic intensity sensor (microphone) with an acoustic impedence matched to the density of water to optimise sonic detection under water.
Version Info ()	1
Deprecated()	false
Same as	http://vocab.nerc.ac.uk/collection/L05/current/369/
Narrower	http://vocab.nerc.ac.uk/collection/L22/current/TOOL0360/
Related	http://vocab.nerc.ac.uk/collection/L05/current/153/
Related	http://vocab.nerc.ac.uk/collection/L05/current/154/
Related	http://vocab.nerc.ac.uk/collection/L05/current/362/
Related	http://vocab.nerc.ac.uk/collection/L05/current/363/
Related	http://vocab.nerc.ac.uk/collection/L05/current/364/
Related	http://vocab.nerc.ac.uk/collection/L05/current/365/
Related	http://vocab.nerc.ac.uk/collection/L05/current/366/
Related	http://vocab.nerc.ac.uk/collection/L05/current/371/
Related	http://vocab.nerc.ac.uk/collection/L05/current/372/
Related	http://vocab.nerc.ac.uk/collection/L05/current/373/
Related	http://vocab.nerc.ac.uk/collection/L05/current/374/
Related	http://vocab.nerc.ac.uk/collection/L05/current/375/
Date ()	2010-11-18 12:37:16.0

Figure 21 NVS 2.0 query results delivered as html (web page)


view-source:vocab.nerc.ac.uk/collection/L22/current/TOOL0354/


```
<?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet href="/VocabV2/Concept2Html.xsl" type="text/xsl"
media="screen"?><rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:dc="http://purl.org/dc/terms/"
xmlns:dce="http://purl.org/dc/elements/1.1/" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:void="http://rdfs.org/ns/void#">
<skos:Concept rdf:about="http://vocab.nerc.ac.uk/collection/L22/current/TOOL0354/">
<skos:prefLabel xml:lang="en">Hydrophone</skos:prefLabel>
<skos:altLabel xml:lang="en">Hydrophone</skos:altLabel>
<skos:definition xml:lang="en">A generic term for an acoustic intensity sensor (microphone) with an acoustic
impedence matched to the density of water to optimise sonic detection under water.</skos:definition>
<dc:identifier>SDN:L22::TOOL0354</dc:identifier>
<dce:identifier>SDN:L22::TOOL0354</dce:identifier>
<skos:notation>SDN:L22::TOOL0354</skos:notation>
<owl:versionInfo>1</owl:versionInfo>
<dc:date>2010-11-18 12:37:16.0</dc:date>
<skos:note xml:lang="en">accepted</skos:note>
<owl:deprecated>false</owl:deprecated>
<owl:sameAs rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/369/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/153/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/154/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/362/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/363/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/364/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/365/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/366/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/371/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/372/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/373/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/374/"/>
<skos:related rdf:resource="http://vocab.nerc.ac.uk/collection/L05/current/375/"/>
<skos:narrower rdf:resource="http://vocab.nerc.ac.uk/collection/L22/current/TOOL0360/"/>
<void:inDataset rdf:resource="http://vocab.nerc.ac.uk/.well-known/void"/>
</skos:Concept>
</rdf:RDF>
```


Figure 22 NVS 2.0 query results delivered as RDF XML

Direct vocabulary search URLs

A user can go directly to the search page for a known vocabulary by adding its code to the search tool URL, for example, adding L22 to the NVS URL https://www.bodc.ac.uk/data/codes_and_formats/vocabularysearch/L22/ links to the SeaVox Device Catalogue. This provides users with vocabulary search links tied to specific fields in metadata documentation.

NVS editor tool

A web-based tool for editing vocabularies and associated mappings has been deployed by BODC (https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_editor/). This tool allows a user assigned with the content governance role for a particular vocabulary to make edits (insert, modify, deprecate) to the terms within the vocabulary (see Figure 23). There is also functionality for the vocabulary to be edited using a bulk upload option where a large number of edits need to be made. The user guide for the NVS vocabulary editor is provided in Appendix B of this document.

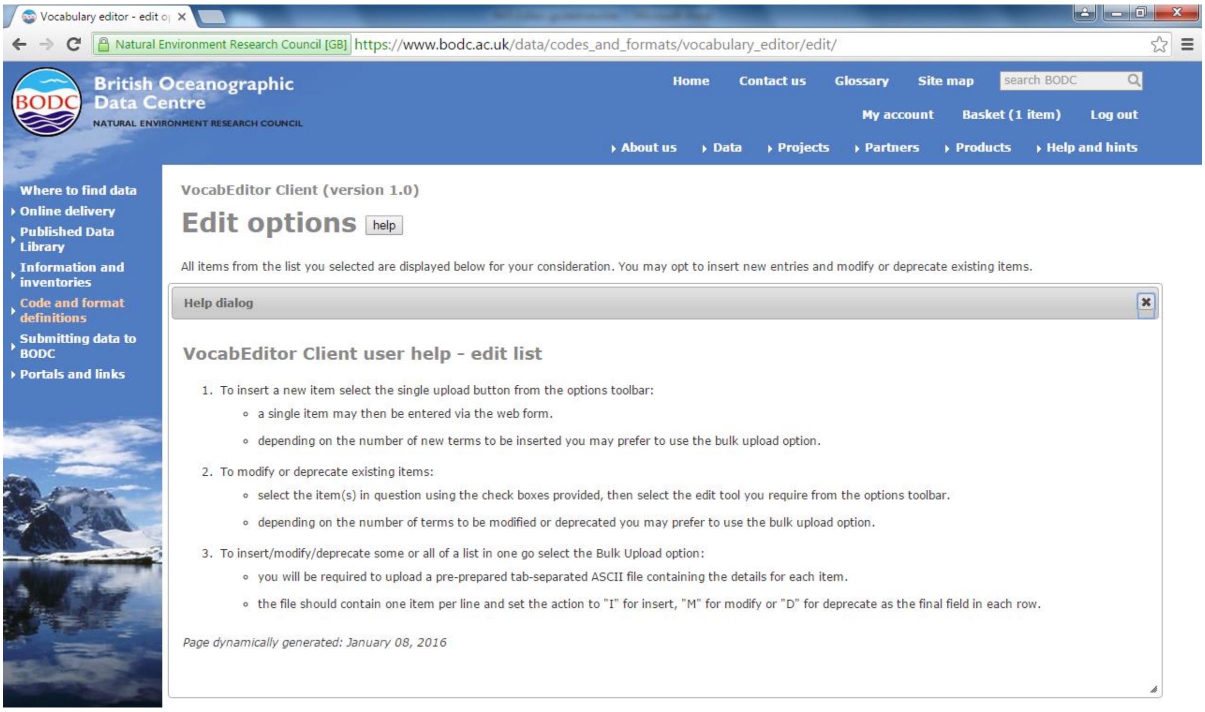


Figure 23 NVS vocabulary editor

NVS Vocabulary builder tool

Building on the exposure of the three semantic models that are used to generate the Parameter Usage Vocabulary (P01) and the semantic models to build biological entities and matrices, BODC has developed and deployed a prototype vocabulary builder tool. This tool is designed to facilitate searching and submission of new terms to three vocabularies: chemical terms from the P01 vocabulary, biological entity terms held in S25 and matrix terms held in S26. This tool can be found at: https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_builder/

P01 - Chemical substance terms

The vocabulary builder allows a user to construct the preferred label for a term from the underlying semantic model as shown in Table 4. As the terms are selected from each component vocabulary they appear in the preferred label box and the number of terms that are relevant to that selection is displayed. The user can choose to see the results at any stage in the construction of the term.

Semantic model component	Component vocabulary	Chemical	Biological	General
Measurement property	S06	Concentration	Abundance	Temperature
Measurement statistical qualifier	S07	-	-	standard deviation

Chemical substance	S27	lead {Pb CAS 7439-92-1}		
Biological entity	S25		Bacteria	
Measurement-matrix relationship	S02	per unit dry weight of	per unit mass of	of the
Matrix	S26	biota	sediment	water body
Biological entity as a matrix	S25	Mytilus galloprovincialis {IT IS: 79456: WoRMS 140481}		

Table 4 Breakdown of the P01 semantic models used to build a vocabulary term

Once a term has been constructed, and if it does not already exist, a create button is displayed that allows a user to submit their code for consideration (see Figure 24). The term is then checked for suitability and if appropriate is published to the NERC Vocabulary Server.

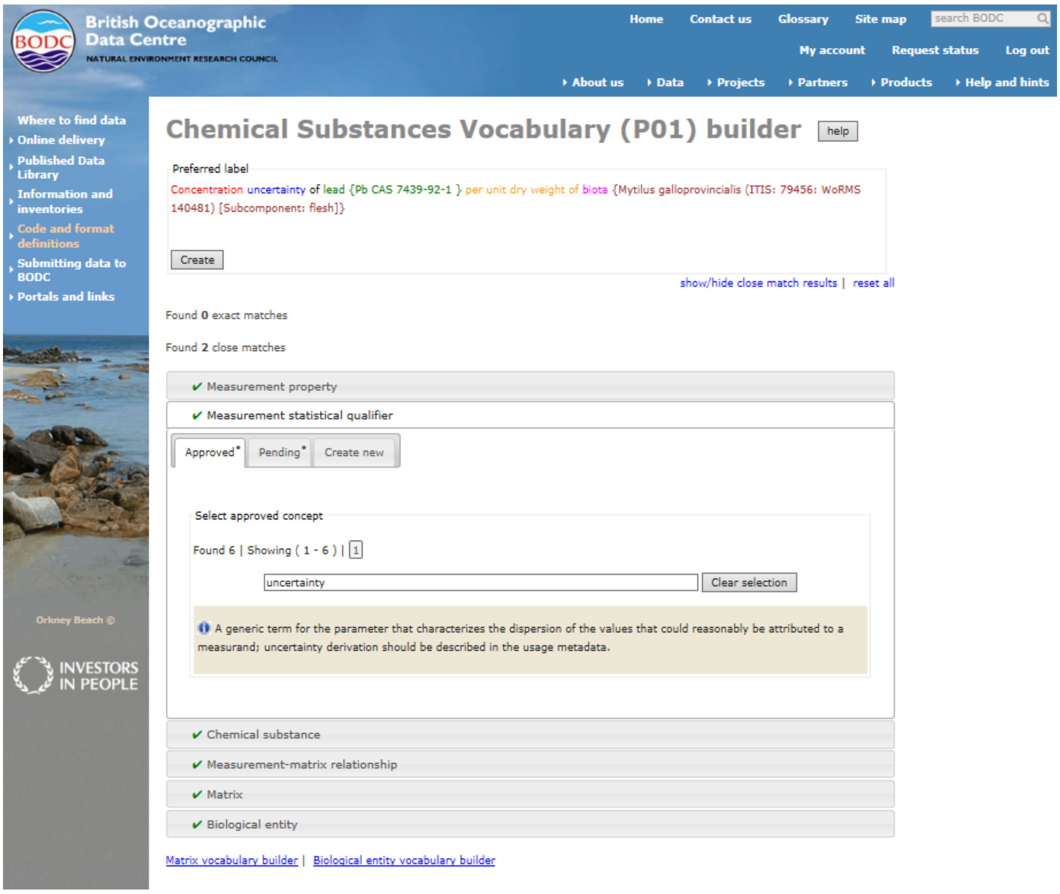


Figure 24 NVS vocabulary builder for chemical substances within the P01 vocabulary

Ongoing and future vocabularies work

BODC has planned the following future activities for the further development of the NVS:

- Shut down of NVS1.0 to allow enriched predicate sets for mappings. This will then facilitate the implementation of unit conversions through use of richer predicates.
- Overlay SKOS with OWL for enhanced semantic reasoning.
- Provide multilingual support for key vocabularies.
- Provide provenance of mappings.
- Expand the vocabulary builder to include P01 terms constructed from the biological and general purpose models.

3.1.2 Australian vocabulary developments

IMOS is now using the Australian National Data Service (ANDS) Tool Suite in production mode. This toolset comprises the Pool Party Semantic Suite for editing, SISSVoc for Linked Data publishing and the ANDS RVA Portal for vocabulary submission, access and searching (https://vocabs.ands.org.au/#!/?p=1).

The Australian Ocean Data Network (AODN) vocabularies (https://vocabs.ands.org.au/#!/?p=1&subjects=Marine) re-use the BODC vocabularies wherever possible. Through participation in ODIP, IMOS has been able to expand the number of Australian organisations registered in the European Directory of Marine Organisations (EDMO), (http://seadatanet.maris2.nl/v_edmo/browse_step.asp?step=00054), as well as adding terms to the existing BODC vocabularies.

Since the first ODIP II workshop IMOS focus has shifted from assisting ANDS in developing the tools necessary to underpin vocabulary use and management to expanding the scope and content of AODN vocabularies and their application. This progress report highlights a number of issues that AODN has encountered along the way and in some cases discusses how there are addressed.

Vocabulary Versioning – Implications for Vocabulary Usage

The AODN vocabularies carry a version number, which applies to a vocabulary scheme. At any point in time the current version of a vocabulary is available from its publication point using ‘current’ as part of the service endpoint request for a specific vocabulary (e.g. https://vocabs.ands.org.au/repository/api/lda/aodn/aodn-discovery-parameter-vocabulary/current/concept for retrieving the current AODN discovery parameter vocabulary),, Alternatively a request can be issued that replaces ‘current’ with the actual version number (e.g., version1-1). Individual vocabulary terms are retrieved by appending the term URI to the service endpoint that carries either ‘current’ or the ‘vocabulary version number’ (e.g. https://vocabs.ands.org.au/repository/api/lda/aodn/aodn-discovery-parameter-vocabulary/version-1-1/resource?uri=http://vocab.aodn.org.au/def/discovery_parameter/entity/401).

Version information is also written into the scheme description of all downloaded vocabulary files, using an ‘owl:versionInfo’ property.

Several of the published (versioned) AODN vocabularies are used to populate extended ISO 19115 elements (e.g. the mcp:dataParameter element) in AODN (Marine Community Profile) dataset metadata records. In the case of MCP metadata, these controlled AODN vocabularies perform a similar function to that of the SeaDataNet Common Vocabularies in the SeaDataNet Common Data Index (CDI). They are integral for standardised indexing and searching within the IMOS Data Portal that is largely driven by dataset-level metadata.

Apart from listing the vocabulary term label and its URL, the mcp:dataParameter element in MCP 2.0 permits the recording of the version of the vocabulary from which a term has been drawn. So, if the metadata content is properly completed, a named term can always be linked back to its parent (versioned) scheme.

These various versioning modalities were built into the AODN vocabularies and metadata because it was recognized that vocabularies are dynamic, particularly during the early days of their formation, and it is important for data providers to be able to track changes occurring as vocabularies evolve. Using out-of-date vocabularies could mean that providers are unaware of new terms available with which to mark-up metadata, or terms that they have used may no longer be valid.

Since the AODN vocabularies are relatively immature when compared to the BODC vocabularies, it is anticipated that their structure and detail will not be that stable in the medium term, even if updates are released on a managed basis. Although AODN borrows heavily from the more stable BODC vocabularies, some of the vocabularies are unique and growing in form and detail (e.g., Organisations and Geographic Extents vocabularies). This situation requires that AODN has a robust migration path/recommended process in place for updating vocabularies used in metadata.

Once AODN had identified the requirements for managing the use of versioned vocabularies in metadata submitted for publishing AODN datasets via the IMOS Portal, activities then moved on to developing a flow-chart for this process. It was at this point that a number of related problematic issues became evident. For example, when a vocabulary was updated and a new version published which may contain additional and/or deprecated terms AODN were confronted with the following decisions in relation to the metadata:

- Should the portal infrastructure only accept new metadata records that carry terms from current vocabularies, or accept metadata marked up with any version up to the current version?
- If any previously valid set of terms are concurrently supported, how would the hierarchy-based portal navigation facets be portrayed if the cumulative versions of a vocabulary resulted in a complex, non-user friendly structure in the classification hierarchy of the vocabulary (associated with trying to portray all versions at once)?
- Should existing metadata, already indexed in the portal Infrastructure be treated differently to metadata yet to be ingested (e.g. should new metadata be validated for compliance using only current vocabularies, whilst accepting already indexed metadata as was submitted)?

- If already indexed metadata should be updated to conform to the new vocabulary version and provider expectation was that this be done in an automated fashion within the portal infrastructure, how would AODN reconcile the fact that updating already ingested metadata creates a difference between the copy of Portal-indexed metadata and that held by providers?
- If it is acceptable to centrally alter (already indexed) provided metadata records what provenance information should be associated with these metadata records and how should this information be managed given that GeoNetwork has limited native provenance management support?
- How would a transition period, between issuing a new version of a vocabulary and expecting compliance in its usage by providers be handled from a communications and a technical perspective?
- If it is possible to change a vocabulary by deprecating terms with no term replacement, or alternatively by using multiple replacements, how do you automate the process of updating the metadata without any form of human intervention?
- Instead of rejecting records because they do not conform, or changing them in the centrally indexed store to make them conform, could AODN alternatively index (or re-index) non-conforming records by doing an on-the-fly mapping of terms which does not alter the original record, only its indexing. In such a scenario the record would still contain out-dated vocabulary terms if inspected, but it would be indexed according to the new vocabulary terms.

There is little best practice guidance available surrounding these issues. However, communications with BODC provided information on some current and proposed future practices to be adopted by SeaDataNet in relation to aspects of these problems. For example, the BODC vocabulary governance rules prohibit term deprecation without a single term replacement and if it is found that a term needs further refinement it is kept as the parent of introduced, more detailed child terms. This makes automating updates easier with respect to the issues relating to the deprecation of terms as identified above.

SeaDataNet also intends to centrally update metadata to reflect changes in vocabulary versions, but it is still unclear how the provenance information regarding such changes will be handled. It is assumed that this central updating leads to inconsistencies between the records indexed and exposed in the portal and those held locally by the data provider. A discussion will be needed with the data providers to gauge their reaction to such a situation.

In summary, AODN is leaning towards the approach where automated mapping is used as records are ingested (or re-indexed) when an updated vocabulary is loaded into GeoNetwork. Although providers will receive a general notification when vocabularies have been updated, they will also receive a notification if their records are using out-dated vocabularies. Providers will be notified of records that require automated mapping and they will be encouraged to update their own metadata at source for re-submission. However, all mapped records will still be re-indexed so if a provider fails to make the record conformant it has been done for them in a virtual sense through mapping.

Vocabulary Mapping

Many of the larger Australian institutions (e.g. CSIRO Oceans and Atmosphere, AIMS, Australian Antarctic Division, GA, Bureau of Meterology) already use some form of in-house vocabularies. Many of these vocabularies are not formalized, published or well governed. However, these terminologies are often integral to how their internal (or public-facing) systems operate and they are reluctant to replace them with an alternative vocabulary. This is particularly the case when the proposed alternative vocabulary does not contain equivalent, or similar terms.

Given that it is unlikely that institutions will abandon their institutional vocabularies in the short term, mappings between institutional terminologies and the AODN common vocabulary will be the primary means of creating standardized vocabulary usage within metadata that is required to underpin the AODN data delivery infrastructure.

IMOS has therefore begun working with partners to help them formalize their institutional vocabularies with a view to encouraging the inclusion of a mapping between the local terminologies and the common AODN vocabularies. This exercise has revealed some structural issues surrounding vocabulary classification in typing L05 (Device) categories. For example. The example of the different terms for photographic equipment in the L05 vocabulary, which are all at the same level in the vocabulary, can be used to characterise these issues:

> **Camera**: “All types of photographic equipment that may be deployed in aircraft or satellites including stills, video, film and digital systems”.
>
> **Underwater Camera**: “All types of photographic equipment that may be deployed underwater including stills, video, film and digital systems.”
>
> **Laboratory Camera**: ”All types of photographic equipment that are hand-held or part of laboratory apparatus including stills, video, film and digital systems”.

The classification in L05 illustrated above delineates cameras on the basis of where they are operating, rather than on the intrinsic properties of the object itself, resulting in a splitting of the term into three alternatives. Using the classification structure applied here, it could be argued that ‘Camera’ – as defined, should be labelled ‘Airborne Camera’ or ‘High Altitude Camera’, to indicate that it is not a generic term for all classes of Camera.

A further complication is that CSIRO has a vocabulary that splits cameras based on a different principle, that of photographic method and then place of operation. Their classification appears as follows:

> **Cameras**
>
> **Cameras**| Cine Cameras
>
> **Cameras**| Still Cameras| Surface Still Cameras
>
> **Cameras** | Still Cameras| Underwater Still Cameras

Cameras | Video Cameras| Surface Video Cameras

Cameras |Video Cameras| Underwater Video Cameras

In trying to map between the two classifications, the task is made more difficult due to the AODN Instrument Vocabulary having re-used a BODC vocabulary that lacks a broad definition for the root term ‘Camera’ that could have had an all-encompassing definition. If ‘camera’ were broadly defined it would always be possible to broadly match the various CSIRO camera types back to the broader term ‘Camera’. As it stands, only ‘Underwater Still Cameras’ and ‘Underwater Video Cameras’ can be effectively ‘closeMatched’ or (perhaps ‘exactMatched’) to the BODC term ‘Underwater Camera’. It is not immediately apparent how to map ‘Surface Still Cameras’ and ‘Surface Video Cameras’. They might possibly be matched to the term ‘Laboratory Camera’ but this would be confusing if the cameras used were actually placed in the field, which is likely in coastal biology.

This example highlights that it will always be useful in vocabularies, such as those issued by BODC, to include a broad definition for root category terms based on the intrinsic properties of the object concerned, in order to facilitate mapping. This does not negate a further finessing of the root term into additional classes of a thing based on more subjective, or fine-grained criteria. However, the root term should be sufficiently generic so that it can be easily mapped onto in the event that criteria for subdividing terms narrower than it differ between various vocabularies.

As most of the interactions between AODN and BODC have been in relation to adding missing terms to vocabularies there has been little discussion regarding the possibility of recasting terms in L05 for other user applications where such an edit may have impacts on internal (BODC) vocabulary relationships. This is an issue for discussion that can potentially be taken forward during future workshops.

Appendix A: Figures and Tables

Appendix B: Terminology

Term	Definition
AODN	Australian Ocean Data Network
API	Application Programming Interface (API): a set of routine definitions, protocols, and tools for building software and applications
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 (CSW): OGC standard for exposing a catalogue of geospatial records in XML on the Internet
DataCite	Global non-profit organisation that provides persistent identifiers (DOIs) for research data to support improved citation https://www.datacite.org/
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 intiative 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/
GEO	Group on Earth Observations: a voluntary partnership of governments and organizations supporting a coordinated approach to Earth observation and information for policy making

GEO-DAB	Brokering framework developed and implemented by GEO for interconnecting heterogeneous and autonomous data systems http://www.geodab.net/
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: international initiative linking together existing and planned observing systems around the world http://www.earthobservations.org/geoss.php
GitHub	Distributed revision control and source code web-based Git repository hosting service.
GML	Geography Markup Language (GML): XML grammar defined by the OGC to express geographical features
ICES	International Council for the Exploration of the Sea http://www.ices.dk/
IMOS	Integrated Marine Observing System: Australian monitoring system; providing open access to marine research data http://imos.org.au/
INSPIRE	EU Directive (May 2007), establishing an infrastructure for spatial information in Europe to support Community environmental policies, and policies or activities which may have an impact on the environment.
IOC	Intergovernmental Oceanographic Commission of UNESCO (IOC/UNESCO).
IODE	International Oceanographic Data and Information Exchange" (IODE) of the "Intergovernmental Oceanographic Commission" (IOC) of UNESCO
IOOS	US Integrated Ocean Observing System https://ioos.noaa.gov/
ISO	International Organization for Standardization http://www.iso.org

jOAI	Java-based OAI software that supports the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), version 2.0 http://www.dlese.org/oai/
JSON	JavaScript Object Notation: an open-standard format that uses human-readable text to transmit data objects consisting of attribute–value pairs. It is the most common data format used for asynchronous browser/server communication.
MarineID	Registration and authentication services for selected marine data services including SeaDataNet and EMODnet
MCP	Marine Community Profile: ISO19115 profile developed by Australian Ocean Data Centre Joint Facility (AODCJF) for marine data
MIKADO	Java-based software tool, for creating XML metadata records for the SeaDataNet directories EDMED, CSR, EDMERP, CDI and EDIOS.
MNF	Marine National Facility is owned and operated by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) http://mnf.csiro.au/
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.
NCEI	NOAA's National Centers for Environmental Information https://www.ncei.noaa.gov/
O&M	Observations and Measurements: OGC standard defining XML schemas for observations, and for features involved in sampling when making observations
ODP	Ocean Data Portal: data discovery and access service, part of the IODE network http://www.oceandataportal.net/portal/
ODV	Ocean Data View: a software package for the interactive exploration, analysis and visualization of oceanographic and other geo-referenced profile, time-series, trajectory or sequence data

OGC	Open Geospatial Consortium: international voluntary consensus standards organization http://www.opengeospatial.org/
OIA-PMH	Open Archives Initiative Protocol for Metadata Harvesting https://www.openarchives.org/pmh/
OpenDAP "	Open-source Project for a Network Data Access Protocol: a data transport architecture and protocol widely used by earth scientists https://www.opendap.org/
OpenSearch	Collection of technologies that allow publishing of search results in a format suitable for syndication and aggregation http://www.opensearch.org/Home
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.
RDF	Resource Description Framework (RDF): family of W3C specifications for conceptual description or modeling of information that is implemented in web resources https://www.w3.org/RDF/
REST	REpresentational State Transfer (REST): an architectural style, and an approach to communications often used in the development of web services
SensorML	OGC standard providing models and an XML encoding for describing sensors and process lineage
SOS	Sensor Observation Service: a web service to query real-time sensor data and sensor data time series. Part of the Sensor Web

SPARQL "	SPARQL Protocol and RDF Query Language: a semantic query language for databases, able to retrieve and manipulate data stored in Resource Description Framework (RDF) format http://www.w3.org/TR/rdf-sparql-query/
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
US-NODC	US National Oceanographic Data Centre (now the NOAA National Centres for Environmental Information) https://www.nodc.noaa.gov/
W3C	World Wide Web Consortium: main international standards organization for the World Wide Web http://www.w3.org/
WCS	Web Coverage Service Interface Standard: OGC standard defining Web-based retrieval of coverages i.e. digital geospatial information representing space/time-varying phenomena http://www.opengeospatial.org/standards/wcs
WFS	Web Feature Service: standards allowing requests for geographical features across the web using platform-independent calls
WMS	Web Map Service: standard protocol for serving geo-referenced map images over the Internet
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/

Appendix C: Using the NVS Editor to manage vocabularies and mappings

The NERC Vocabulary Server (NVS) editor can be located from the BODC homepage as shown or by using the following link:

https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_editor/.

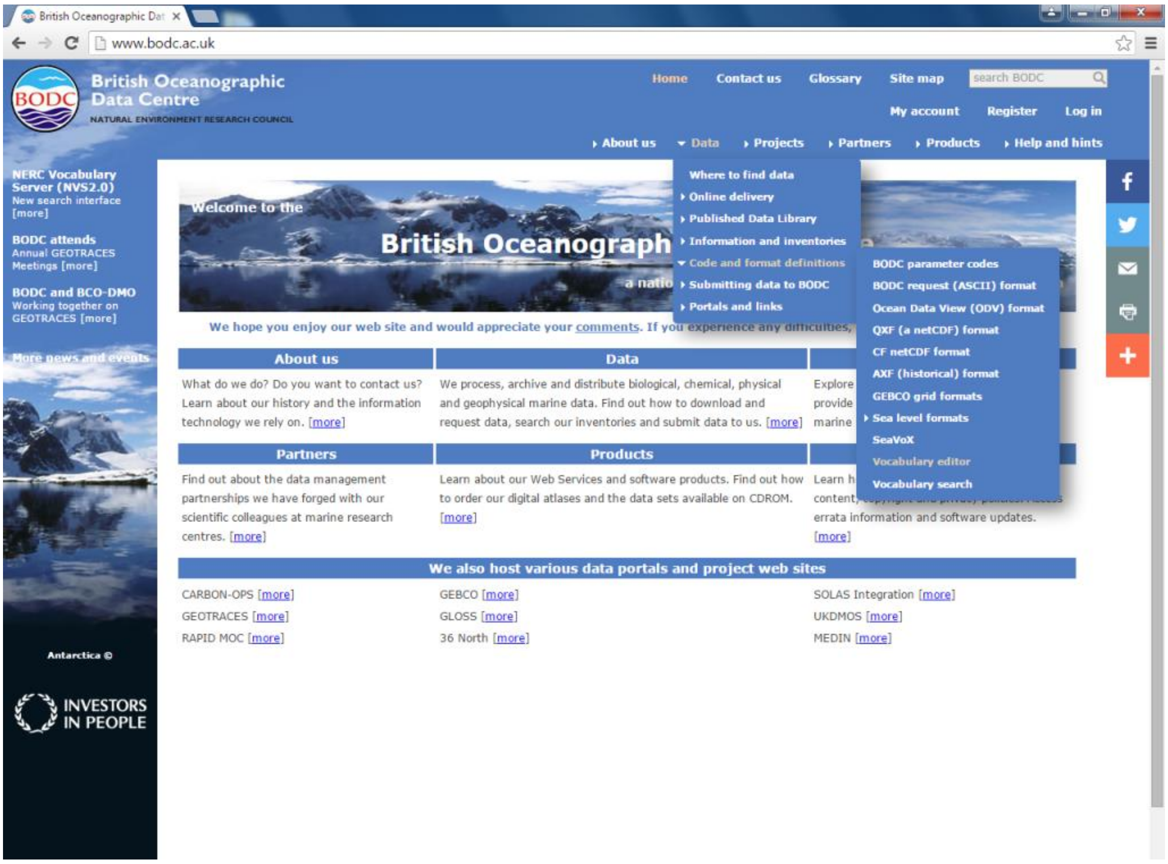


Figure B.1: Accessing the NVS Editor from the BODC homepage.

The editor is intended to be used for updating existing vocabularies. If a new vocabulary needs to be set up then details of the vocabulary (Name, Description and Governance) should be emailed to BODC (enquiries@bodc.ac.uk). Users should search the NVS to determine if appropriate vocabularies already exist and can seek advice from the BODC Vocabularies Management Group. It may be that an existing vocabulary can be extended. Once the “container” for the newly requested vocabulary is in place, the vocabulary can be populated through the NVS Editor.

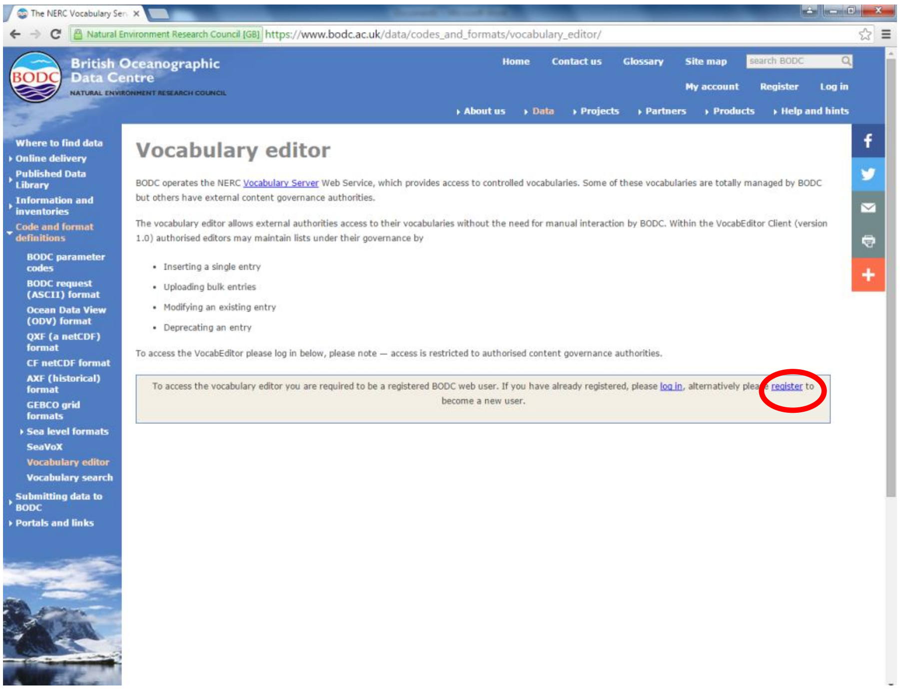


Figure B.2: The editor requires a user to be registered before use.

Registration is through a simple web form. The details of which vocabularies a user wishes to have update permissions on should be supplied in the free text field. BODC will confirm with the appropriate governance group that the user should be granted these permissions.

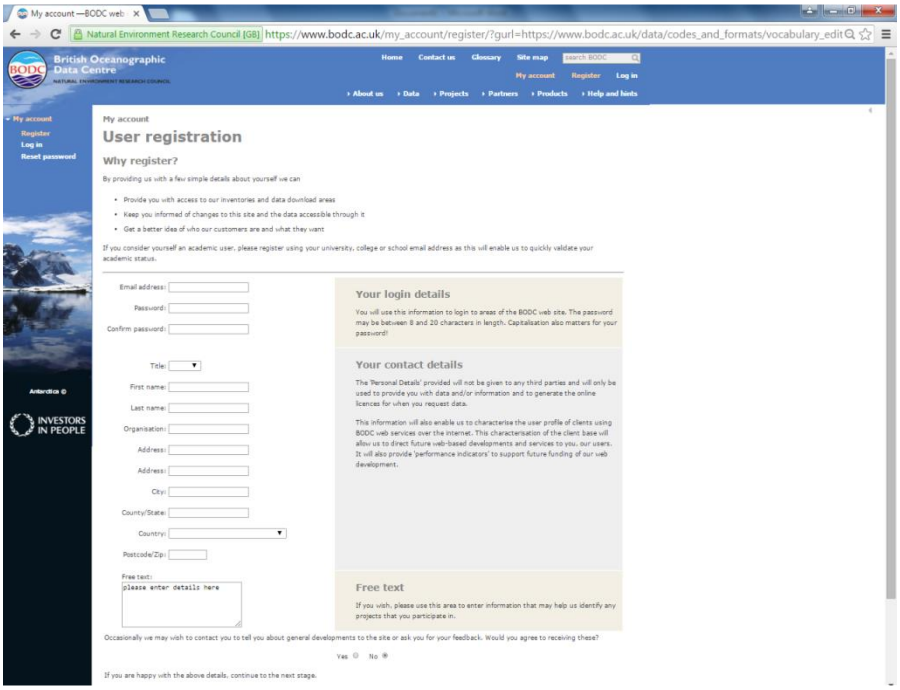


Figure B.3: BODC website user registration page.

If a user is already registered and needs new permissions for a vocabulary they should email enquiries@bodc.ac.uk.

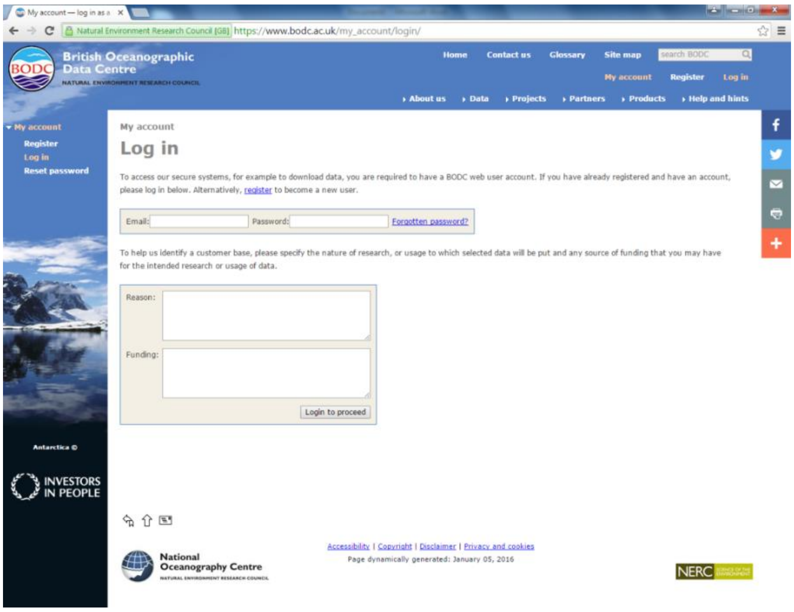


Figure B.4: BODC website login page.

After logging in the user will be presented with a list of the vocabularies that they have permission to edit. Clicking on the hyperlinked List ID will take you to the vocabulary as displayed on the NVS2.

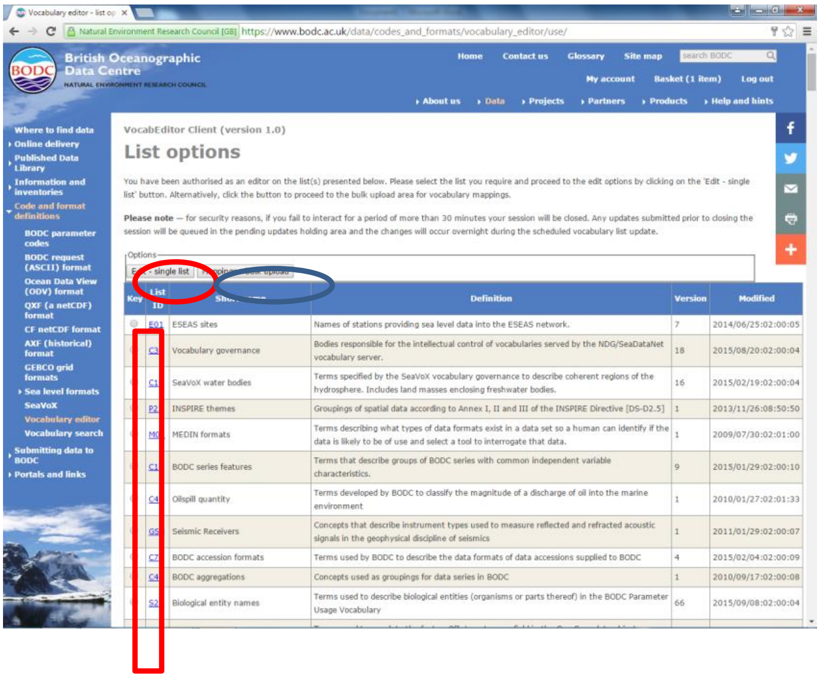


Figure B.5: NVS Editor page showing vocabularies available for editing by the user.

Select the appropriate vocabulary using the radio button next to the vocabulary List ID and choose the “Edit – single list” option highlighted in red.

There is an option to add mappings for terms by using the “Mappings – bulk upload” option highlighted in blue.

Editing a vocabulary list

There are online help details available when using the editor.

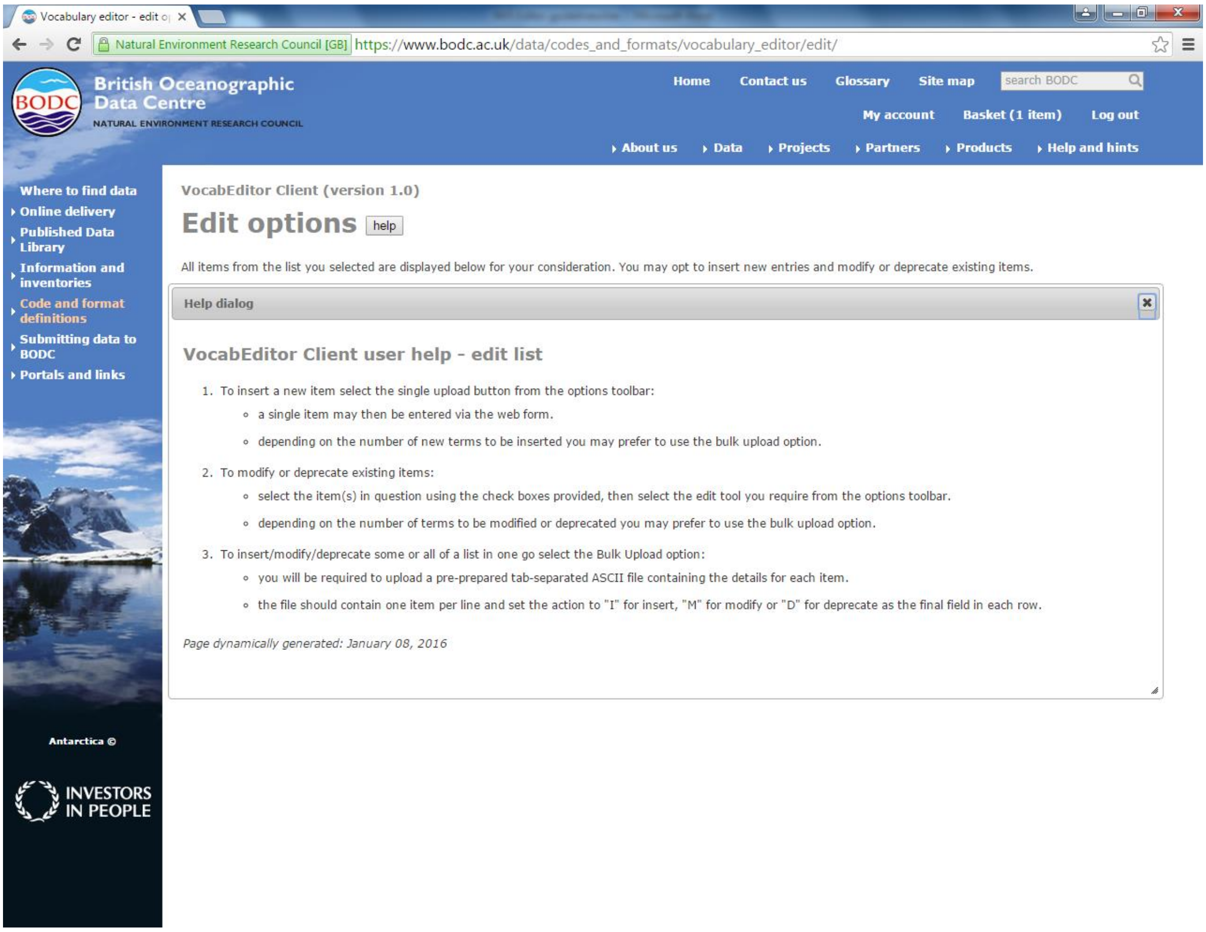


Figure B.6: NVS Vocab Editor user help for making edits to a vocabulary.

Having selected the vocabulary to edit, the user will be presented with the terms already within the list and four options for editing (Single insert/Bulk update/Modify/Deprecate).

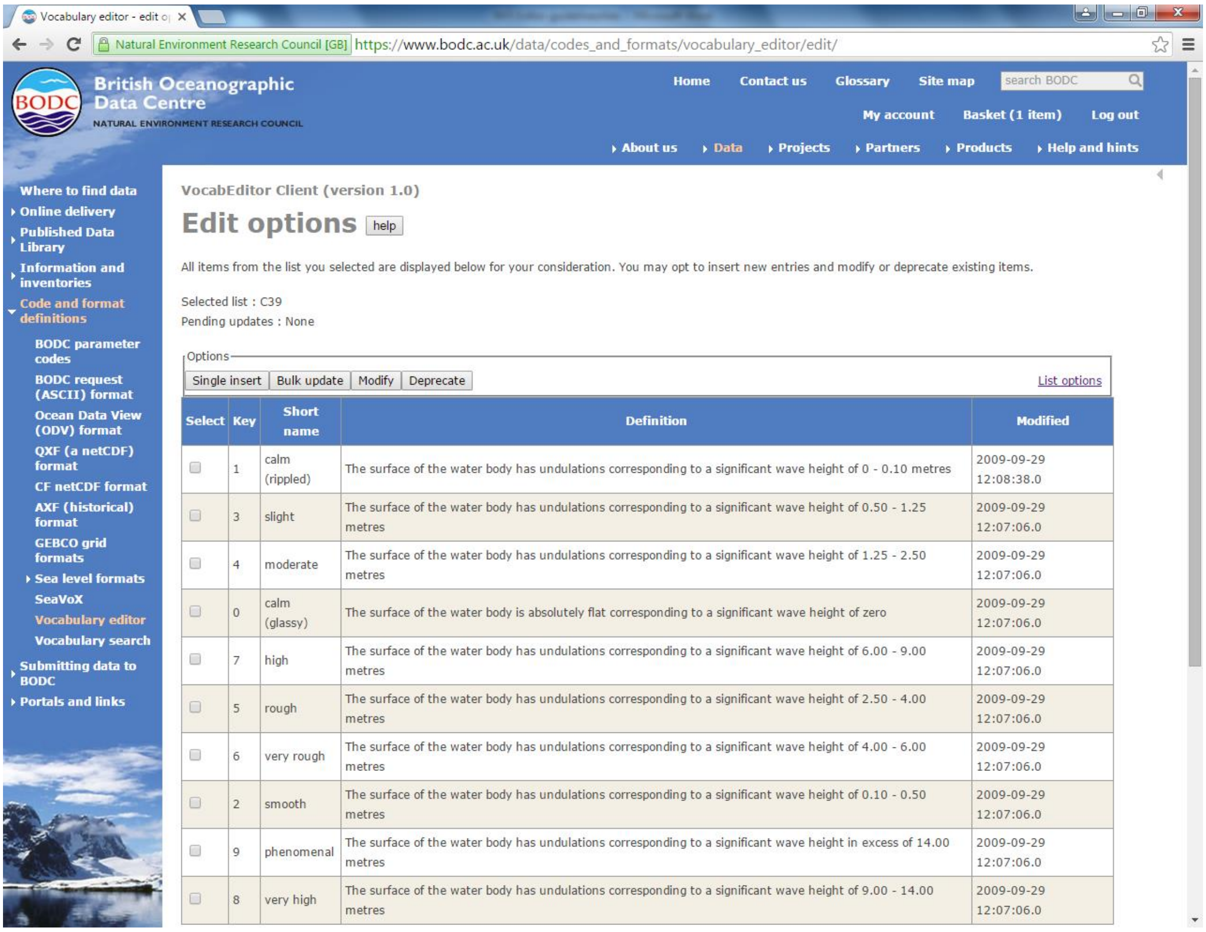


Figure B.7: Example of how a selected vocabulary's terms are listed for editing.

Single insert

The user manually types text into each of the required fields

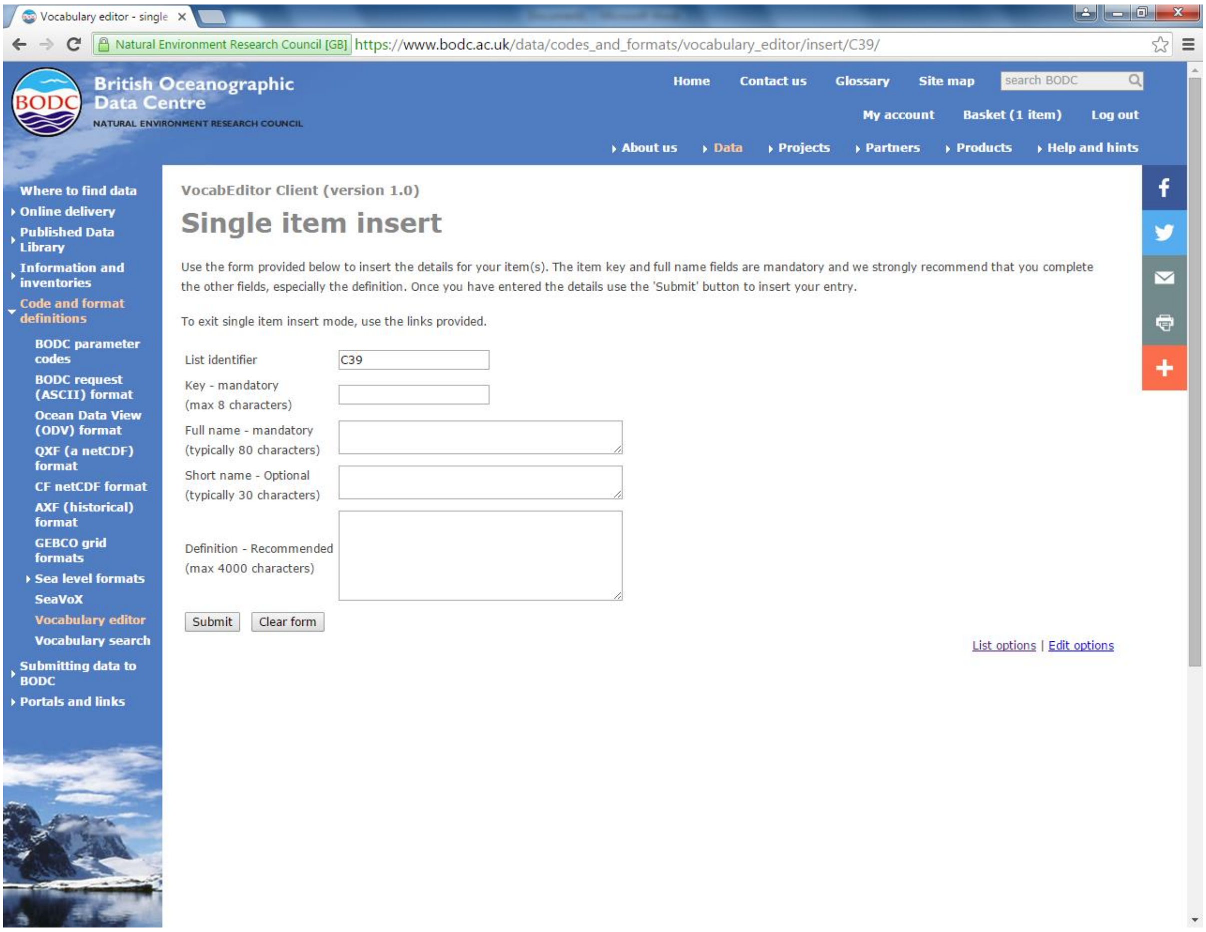


Figure B.8: How to insert a single item.

<u>Bulk update</u>

The upload file must be TAB delimited with five elements per row as detailed in the next figure.

Figure A.8: Web form for inserting a single item into a vocabulary.

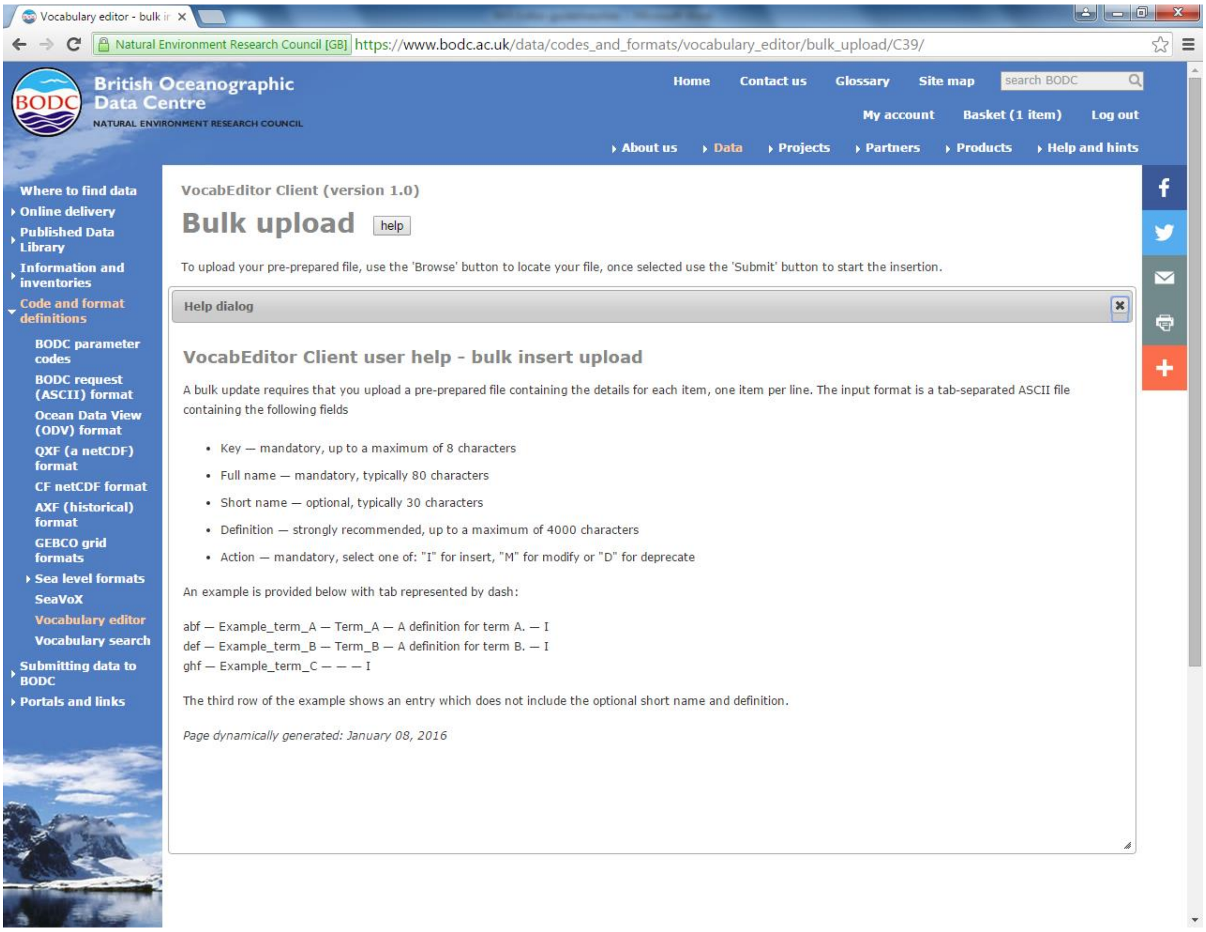


Figure B.9: NVS Vocab Editor user help detailing how to set up a file for the bulk update of a vocabulary.

The content of the vocabulary bulk upload file should be laid out as tab delimited text as follows; containing Key, Full name, Short name, Definition and Action in that sequence.

abf	Example_term_A	Term_A	A definition for term A.	D
def	Example_term_B	Term_B	A definition for term B which is being modified.	M
ghf	Example_term_C	Term_C	A definition for term C.	I

On loading, each row in the file will be reported with a status code from the following list.

VocabEditor Client user help - List terms upload- status codes

Status code	Meaning
200	Success - item inserted - currently queued in the holding area.
400	Specified insert is already member of the list or is currently queued in the holding area.
401	User not authorised - user session has elapsed please login again.
403	User not authorised - insufficient permissions for list specified.
404	Incorrect format. The list identifier not a valid list.
500	Oracle error - please try again (please contact enquiries@bodc.ac.uk if problem persists).

The rows that are unsuccessful should be corrected accordingly and reloaded or the user contact BODC to resolve any issues over permissions.

Modify

Terms to be modified are selected, then all fields except the Key can be modified. The action will be applied by clicking the “Modify” button.

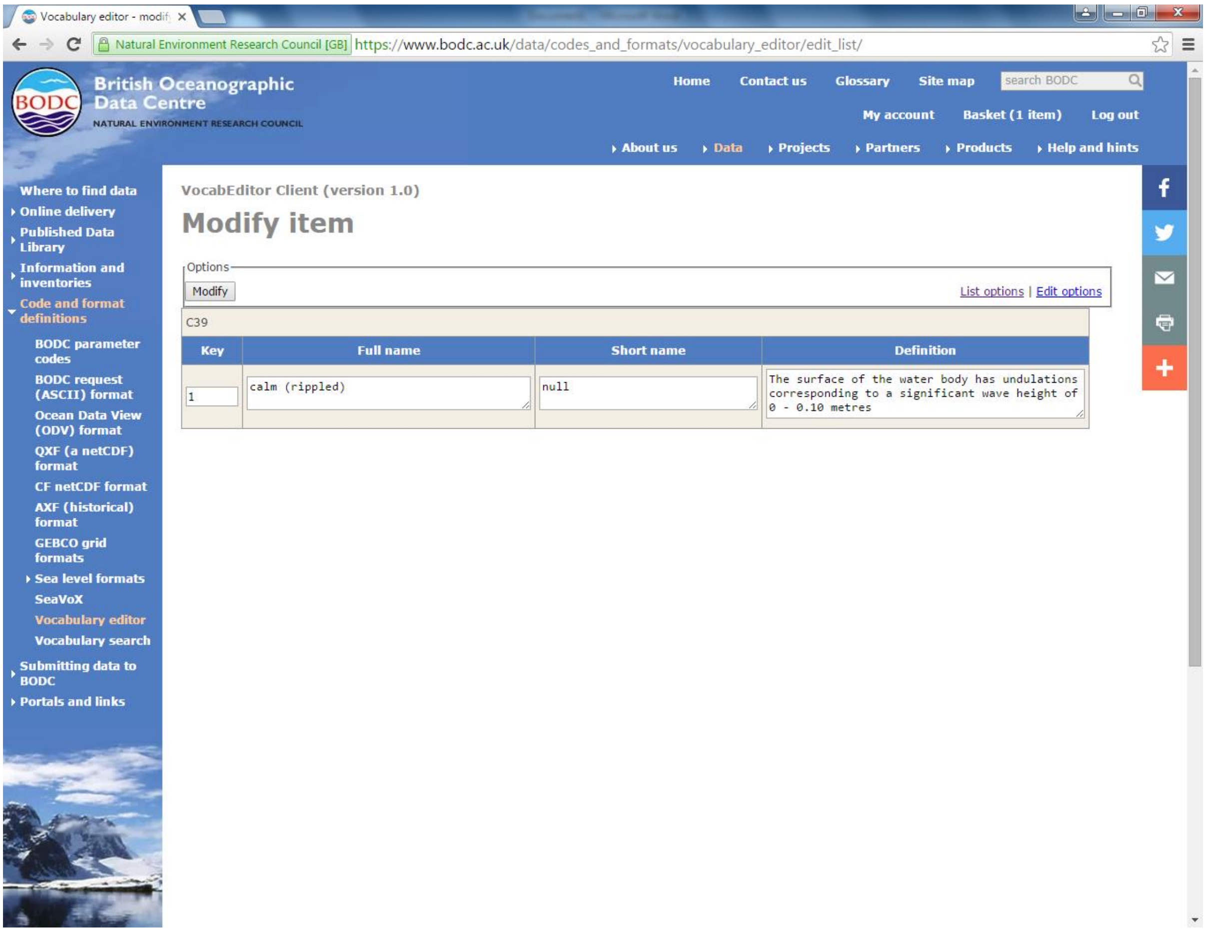


Figure B.10: NVS Vocab Editor web form for modifying an existing item.

The key field cannot be changed. The full name, short name and definition should only be changed to clarify details of the concept NOT to refer to a different concept. If a term is to be replaced then the deprecate function should be used and a new concept added to the list.

Deprecate

This requires terms to be selected prior to clicking the “Deprecate” button. The user is then given a second chance to confirm the term(s) to be deprecated.

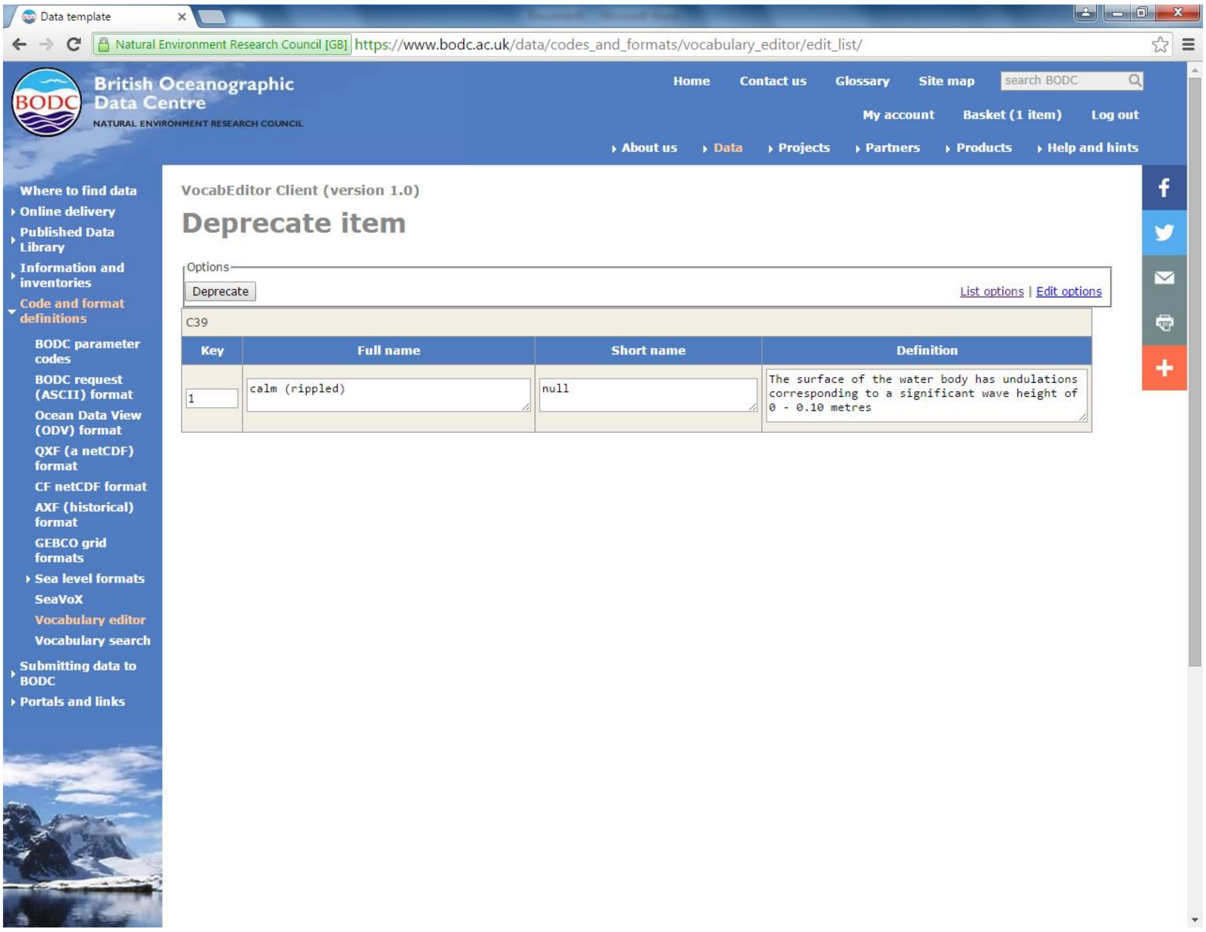


Figure B.11: NVS Vocab Editor web form for deprecating an existing item.

Mappings

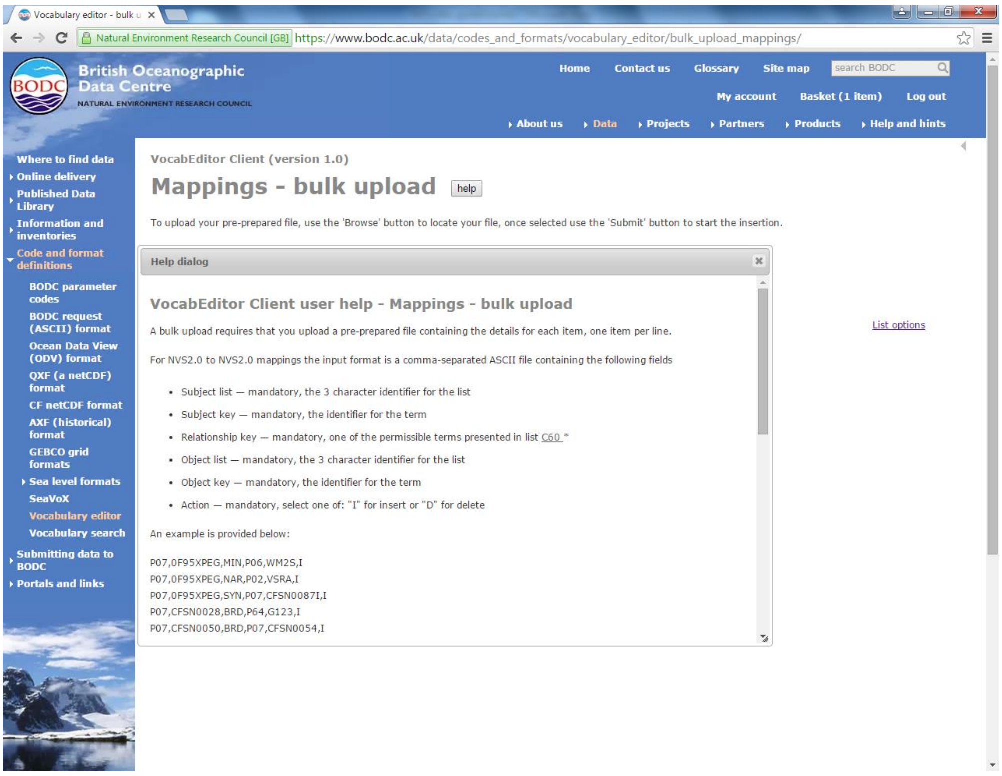


Figure B.12: An example layout for inserting mappings between two NVS vocabulary lists.

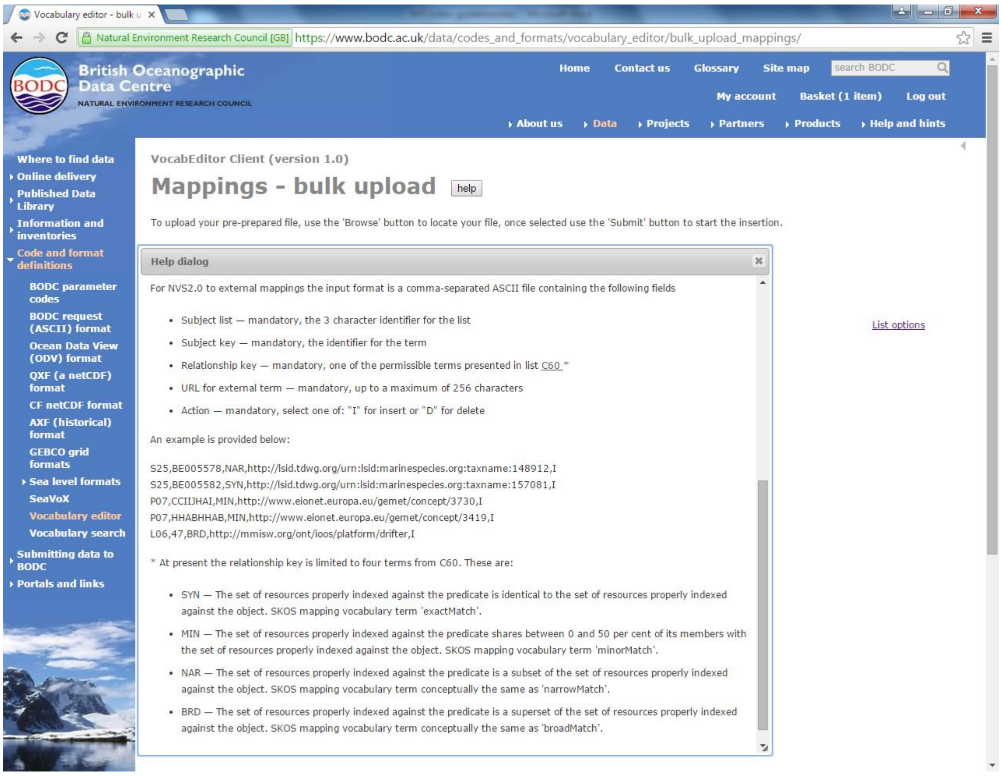


Figure B.13: The following screen print provides the example layout for inserting mappings between an NVS vocabulary list and a non-NVS served concept URL.

The content of the mappings bulk upload file should be laid out as comma separate value text as follows; containing Subject List, Subject Key, Relationship Key, URL for external term and Action in that sequence.

S25,BE005578,NAR,http://lsid.tdwg.org/urn:lsid:marinespecies.org:taxname:148912,I
S25,BE005582,SYN,http://lsid.tdwg.org/urn:lsid:marinespecies.org:taxname:157081,I
P07,CCIIJHAI,MIN,http://www.eionet.europa.eu/gemet/concept/3730,I
P07,HHABHHAB,MIN,http://www.eionet.europa.eu/gemet/concept/3419,I
L06,47,BRD,http://mmisw.org/ont/ioos/platform/drifter,I

After loading the mappings each row will be returned with a status code from the following list.

VocabEditor Client user help - Mappings - status codes

Status code	Meaning
200	Mapping successfully uploaded.
400	Mapping specified already pending.
401	User has insufficient permissions to generate a mapping for this list.
409	An empty or null mapping, i.e. a blank line in your input file.
410	Incorrect format. Expected mapping is: Subject list, Subject key, Relationship key, Object list, Object key.
412	Incorrect format. Invalid option as a mapping term cannot be mapped to itself.
414	Incorrect format. List C60 (the relationship option) is not a permitted value for the Subject and/or Object list.
415	Incorrect format. One or more of the list or key terms entered were not found.
416	Incorrect format. Expected mapping is Subject list, Subject key, Relationship key and URL.
417	Invaild URL. A mapping to a NVS2.0 URL (i.e. vocab.nerc.ac.uk) is not permitted. Please use the NVS2.0 to NVS2.0 option for such mappings.
418	Invalid URL. An invalid response was returned from the external URL. Please check that it exists.
500	Oracle error. Please try again (please contact enquiries@bodc.ac.uk if the problem persists).

The rows that are unsuccessful should be corrected accordingly and reloaded.

Appendix D – Comparative table of XML editors

The following table compares several characteristics of available SensorML open source editors and can be intended as an extension of the OGC web pages of SensorML related softwares (http://www.ogcnetwork.net/SWE_Software).
This tentative comparison should serve as a starting point for the ODIP discussion and activity about synchronizing efforts for metadata/SensorML editors.

	GET-IT EDI[1]	SensorML process edi- tor (recently named SensorML Library)[2]	ensorML Schema Browser[3]	SensorML Profile Library[4]	Pines SensorML Editor[5]	SensorML Editor[6]	OpenSensorHub SensorML editor[7]	SensorNanny-drawmyob-servatories[8]	ISTSOS[9]	52North - smle[10]
Reported in OGC webpage[11]	y (within comment)	Y	Y	Y	Y	Y	N	N	N	N
Software type	editor	editor and software library	webpages with graphical representation of a selection of SensorML	file repository (see descr.)	editor	editor	editor	editor	editor (embedded in SOS)	editor
Development group	GET-IT, CNR IREA, CNR-ISMAR	UAH - SensiaSoft, Botts Innovative Research Inc.	UAH - ensiaSoft, Botts Innovative Research inc.	UAH - ensiaSoft, Botts Innovative Research inc.	Xuesong Liu (Civil and Environmental Engineering, Carnegie Mellon University)	Botts innovative Research	SensiaSoft	IFREMER	SUPSI	52North initiative

1 http://www.get-it.it/ (website), https://github.com/SP7-Ritmare (software)
2 http://code.google.com/p/sensorml-data-processing/ (software - deprecated), https://github.com/sensiasoft/lib-sensorml (software - current)
3 http://www.botts-inc.net/SensorMLv1.0_Browser.html (website)
4 http://code.google.com/p/sensorml-profile-library/ (software)
5 http://lxspine.googlepages.com/pine\%27ssensormleditor/ (website and software)
6 https://code.google.com/archive/p/sensorml-tools/downloads/ (software)
7 https://github.com/opensensorhub/sensorml-editor (software)
8 https://github.com/ifremer/snanny-drawmyobservatory (software)
9 http://istsos.org/index.html (website), https://sourceforge.net/projects/istsos/ (software)
10 http://52north.github.io/smle (website), http://blog.52north.org/2016/05/24/usability-improvements-for-smle/ (software)
11 http://www.ogcnetwork.net/SWE_Software

	GET-IT EDI[1]	SensorML process edi- tor (recently named SensorML Library)[2]	ensorML Schema Browser[3]	SensorML Profile Library[4]	Pines SensorML Editor[5]	SensorML Editor[6]	OpenSensorHub SensorML editor[7]	SensorNanny-drawmyob-servatories[8]	ISTSOS[9]	52North - smle[10]
Description	template-driven metadata authoring tool that can be easily customized to any XML-based metadata format and to a specific workgroup, institute, or project.	Library for the execution of processes represented in SensorML. It is a process chain execution engine (not an editor of SensorML).	Webpages pointing to views of SensorML schema (similar to XML representation utilities like in generic XML editors). Currently no software seems to be available.	Repository for executable SensorML process model instances, as well as RelaxNG profiles of the core SensorML schema (not an editor).	Program to explore and modify SensorML models	This product is used to create and edit SensorML instances. It may be downloaded as a standalone eclipse application.	A web based viewer/editor to create your SensorML document. This SensorML viewer/editor is used by OSH but can also be used as a standalone web editor.This editor allows to view any SensorML documents (V2.0) and edit the current content. The project has been designed using GWT.□Currently under development. (March,2016)[12] The sw allows the EDITION of existing SensorML document instances.	Graphical composition of predefined SensorML of specific Sensors. App for OwnCloud.	SOS server with SensorML editor embedded in the management interface of the SOS	SensorML editor which enables web-based editing of SensorML descriptions
Last updated	2016	2016	unavailable	2009	2011 (?)	2009	2016	2016	2016	2016
Status	stable	stable	-	-	stable	stable	under dev.	beta	stable	under dev.
Licence	GPL v.3	Mozilla Public License, version 2.0	undefined	Mozilla Public License 1.1	undefined	Mozilla Public License 1.1	Mozilla Public License 2	GNU AFFERO GENERAL PUBLIC LICENSE	GPL v.2	Apache License 2.0
SensorML 1	Y	unclear[13]	Y	Y	Y	Y	N	N	Y	unclear
SensorML 2	Y	Y	N	N	N	N	Y	Y	N	Y
Exensibility to other MD schemas	Y	-	-	-	N	N	N	N	N	N
Sensor model support	Y (prototypical sensors to be edited with the same UI; Partial support for derivation constraints, under development)	Y	-	-	N	?	Y (under dev.)	Y (predefined models only; extendable with other models through custom JSON documents for new models information)	Y (prototypical sensor instances; prototypes can be chosen from sensors already in the system; constrained to the profile supported by the software)	Y

12 See also the video by M. Botts https://www.youtube.com/watch?v=PDdWYLbFmKY

13 Documentation does not provide such information. Apparently the source code has only references to SensorML 2.0.0

	GET-IT EDI[1]	SensorML process edi- tor (recently named SensorML Library)[2]	ensorML Schema Browser[3]	SensorML Profile Library[4]	Pines SensorML Editor[5]	SensorML Editor[6]	OpenSensorHub SensorML editor[7]	SensorNanny-drawmyob-servatories[8]	ISTSOS[9]	52North - smle[10]
Profiles support	Y (by EDI template Language)	-	-	-	Y (validation only by RelaxNG and schematron)	?	Y (under dev.)	?	?	?
UI type	web form	-	-	GUI-visualization	Java desktop application (with GUI)	?	web form	GUI, web app for owncloud	web form	web form
Standalone module	Y	-	-	-	Y	Y	Y	Y	N	?
Integrated in other systems	Y (e.g. GET-IT)	Y	-	-	N	?	Y (OpenSensorHub)	Y (SensorNanny)	Y (part of the management interface of IstSOS software)	?
Type of support SWE Common Data Model ``definition'' attributes (anyURI) [14]	runtime SPARQL queries	-	-	-	manual(?)	?	undocumented (values encoded within configuration?)	local	local db, dynamic (the user can add new values, URIs manually, which will be available for new insertions)	?
Programming language	JAVA, Javascript	JAVA	-	-	JAVA	JAVA	JAVA	php, javascript	python, javascript	Javascript (AngularJS framework), TypeScript
Persistence	JBO, SOS	-	-	-	?	?	XML	JSON	SOS db	? (declared SOS but unclear at this state of development)

SensorML editors: tentative enumeration and comparison of available software characteristics.
Legend: ``-'' = non applicable; ``?'' = not enough information within documentation or other sources (code, examples, executable if any).

[14] E.g. manual insertion of URIs, values stored in local db (static or dynamic lists), remote runtime queries, ...