Grant agreement no: 654310

Ocean Data Interoperability Platform

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

Workpackage	**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	2 November 2016

Document History				
Version	**Author(s)**	**Status**	**Date**	**Comments**
1.0	Dick M.A. Schaap (MARIS)	DRAFT	2 November 2016	First draft using contributions from several partners

Contents

Executive Summary

In the framework of the ODIP II project prototypes are worked out in order to evaluate and test selected potential standards and interoperability solutions for establishing and demonstrating improved interoperability between the regional infrastructures and towards global infrastructures. As a first strand of prototyping for the ODIP II project it was discussed and agreed between partners to expand the three prototype projects that had been successfully initiated and implemented in the predecessor ODIP project. These concerned:

- **ODIP 1**: Establishing interoperability between SeaDataNet CDI, 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 systems in Europe, US 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 at research vessels and in real-time monitoring systems.

The specifications for the expansion have been formulated in the period between the 1st ODIP II Workshop and the 2nd ODIP II Workshop and documented in the ODIP II Deliverable D3.1. In summary, the targets for the expansions concern:

- **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; and analysing 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. Lead by MARIS with contributions by European, USA and Australian partners;
- **ODIP 2+:** Further population of the Cruise Summary Reports database for USA cruises from R2R, initiating contributions for Australian cruises, and including ICES legacy CSRs; analysing the upgrading of the CSR metadata format, Schema and related tools and services following earlier R2R suggestions; and publishing CSRs also in RDF by means of a SPARQL endpoint. Lead by BSH with contributions by European, USA and Australian partners;
- **ODIP 3+:** Establishing OGC Sensor Web Enablement (SWE) standards for facilitating the interoperable sharing of oceanographic observation data and metadata; analysing 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. Lead by 52°North with contributions by European, USA and Australian partners.

This ODIP II Deliverable D3.3 gives a detailed report on the activities that have been undertaken and the progress that has been achieved up till the end of Month 18 for the three ODIP II expansion projects and the cross-cutting topic on vocabularies. In several aspects good progress has been achieved, while for other aspects a start has been made. The development activities for these three prototypes and cross-cutting topics are continued and the final results will be reported in the Deliverable D3.4 that will be released near the end of

the ODIP II project. D3.4 will then also report on the progress and results achieved for the second strand of ODIP II prototypes for which specifications are documented in Deliverable D3.2.

1. Introduction

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

This is done in practice by organising four international workshops over the three years lifetime of the project to present, compare and discuss approaches and standards applied. The workshops involving relevant domain experts provide insights into commonalities and differences and contribute to identify opportunities for the development of common standards and interoperability solutions. As a follow-up ODIP prototypes projects are formulated and worked out in order to evaluate and test selected potential standards and interoperability solutions for establishing and demonstrating improved interoperability between the regional infrastructures and towards global infrastructures. A complication has arisen for ODIP II in comparison to ODIP that only EU partners have achieved extra funding, while contributions from USA and Australian partners must be brought up from their own institute funds. This gives even more emphasis on the approach that actual ODIP II prototype developments should be done largely by leveraging on the activities of current regional projects and initiatives of the ODIP II partners. Therefore ODIP II prototype projects must be formulated taking these constraints into account. This also implicates that additional developments outside of ongoing projects should be done largely by the European ODIP II partners. Fortunately the European partners will start from 1st November 2016 with the EU supported SeaDataCloud project as successor to the SeaDataNet II project and this will give extra synergy options.

As a first strand of prototyping for the ODIP II project it was discussed and agreed between partners during the first 2 ODIP II Workshops to expand the three prototype projects that had been successfully initiated and implemented in the predecessor ODIP project. The specifications for the expansions have been documented in the ODIP II Deliverable D3.1.

This ODIP II Deliverable D3.3 will give a detailed report on the activities that have been undertaken and the progress that has been achieved up till the end of Month 18 for the three ODIP II expansion projects as well as for the cross-cutting topics. It should be noted that this Deliverable will give a momentary report as the developments are ongoing and will be continued till the end of the ODIP II project. The full results will then be reported in Deliverable D3.4 that will also cover the results of the second strand of ODIP II prototypes as specified in Deliverable D3.2.

2. Progress of the three prototype expansions

The specifications for the expansion have been formulated in the period between the 1st ODIP II Workshop and the 2nd ODIP II Workshop and documented in the ODIP II Deliverable D3.1. In summary, the targets for the expansions concern:

- **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; and analysing 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. Lead by MARIS with contributions by European, USA and Australian partners;
- **ODIP 2+:** Further population of the Cruise Summary Reports database for USA cruises from R2R, initiating contributions for Australian cruises, and including ICES legacy CSRs; analysing the upgrading of the CSR metadata format, Schema and related tools and services following earlier R2R suggestions; and publishing CSRs also in RDF by means of a SPARQL endpoint. Lead by BSH with contributions by European, USA and Australian partners;
- **ODIP 3+:** Establishing OGC Sensor Web Enablement (SWE) standards for facilitating the interoperable sharing of oceanographic observation data and metadata; analysing 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. Lead by 52°North with contributions by European, USA and Australian partners.

The activities that have been undertaken and the progress that has been achieved up till the end of Month 18 for the three ODIP II expansion projects are described in the following paragraphs.

2.1 ODIP Prototype 1+

The earlier ODIP Prototype 1 is the basis for the expansions and it concerns interoperability of regional discovery services (SeaDataNet, NCEI and AODN) towards the global GEOSS and ODP portals. In the present situation there is **metadata** brokerage at **collections level** and entries are included in GEOSS and ODP with return links to granule level at regional portals and their data access options.

The ODIP Prototype 1+ is led by MARIS with expected contributions from partners from the three regions. Activities have been undertaken and are planned for each of the targets of the **ODIP Prototype 1+** project:

1) establishing semantic interoperability between the three regional discovery and access services at metadata level:

The idea is to develop a translation service that will interact with the GEO-DAB brokerage service. The translation service will capture knowledge of which terms are identical to, similar or specialisations of other terms used across the regional systems (anyway

SeaDataNet, NCEI and AODN) and publish in a machine to machine system. A start has been made with identifying vocabularies used in the three regional systems.

SeaDataNet makes use of:

- SeaDataNet Common Vocabularies as hosted by NERC-BODC 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;
- SeaDataNet EDMO codes for organisations
- SeaDataNet EDMERP codes for projects
- SeaDataNet CSR codes for Cruise Summary Reports

The data centres in the SeaDataNet infrastructure might use other vocabularies in their own local systems, but as SeaDataNet partner each connected data centre has to map its local vocabularies to those used in SeaDataNet for achieving a harmonised population of the SeaDataNet metadata directories and data resources.

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

- NASA's Global Change Master Directory, 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, also use is made of other vocabularies:

- 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 data centres, recently bundled as National Centres for Environmental Information (NCEI), oversee and manage an enormous data collection, which also goes back a long time. This implicates 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 their metadata.

The Australian Integrated Marine Observing System (IMOS) makes use of the SeaDataNet Common Vocabularies, where-ever possible. The participation in ODIP has been very favourable for this decision and IMOS has become also a regular contributor for new terms to these vocabularies. Furthermore IMOS has become the Australian national node for entering and maintaining Australian organisations (so far > 250) in the SeaDataNet EDMO directory. While developing the overarching Australian Ocean Data Network (AODN) it appeared that 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. But, these terminologies are often integral to how their internal (or public-facing) systems operate. Therefore IMOS has begun with developing mappings between institutional terminologies and the AODN common vocabulary as the primary means of creating standardized vocabulary usage within metadata that is required to underpin the AODN data delivery infrastructure.

Taking this situation with the three regional systems into consideration it seems that it should be feasible to determine mappings between vocabularies as being used in comparable metadata fields by SeaDataNet and AODN. However undertaking a comparable exercise with NCEI will be a considerable challenge, also taking into account that NCEI is using a very rich metadata format approach with many additional attributes in comparison with the SeaDataNet CDI and AODN MCP metadata formats. The latter have been set up from their start as common core implementations of the ISO19115 standard with attributes supported by controlled terms, where-ever posssible. In that sense one can say that both SeaDataNet CDI and AODN MCP formats are more lean and contain less information than the NCEI metadata format which has been developed and populated by bringing together legacy metadata systems.

In stead of trying to map the various vocabularies to their full extent another approach will be implemented. The plan is now to make use of the comparisons between the metadata formats of the three regional systems as earlier done by CNR as part of the configuration of the GEO-DAB Brokerage Service. Thereby the metadata formats of each of the three systems have been analysed and mapped to the Brokerage Reference Schema to configure the brokerage. The GEO-DAB service thus provides CS-W and OAI-PMH services for each of the three regional systems giving converted XML records while sustaining the original terms for specific attributes. The services can be found at :

OAI-PMH service for SeaDataNet collections:

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

CS-W service Version 2.0.2 Service – HTTP POST method for SeaDataNet collections:

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

with SEADATANET as parameter for retrieving the SeaDataNet output.

OAI-PMH service for NCEI collections:

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

CS-W service Version 2.0.2 Service – HTTP POST method for NCEI collections:

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

with ‘apiso:parentIdentifier’ = NODC

OAI-PMH service for AODN collections:

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

CS-W service Version 2.0.2 Service – HTTP POST method for AODN collections:

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

with ‘apiso:parentIdentifier’ = AODNCSWCORE

MARIS has recently harvested the XML output of each of the three regional services using OAI-PMH which resulted in:

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

The following steps are planned:

- The harvested XML files will be loaded into a database, configured to the brokerage reference schema;
- Per attribute shortlists will be made of the range of terms as originating from each of the three regional systems;
- These shortlists will be the further subject of mappings in contact with the managers of the regional systems, also determining in detail which source vocabularies have been used;
- If successful, these mappings will provide the required basis for the planned vocabulary translation service;
- Options will be considered for the publication of the mappings in human and machine readable formats. An "ODIP Rosetta Stone" vocabulary service might be provided and maintained. One or more vocabulary services host all mappings that are required for catalogue harmonisation;
- Analyse how and arrange that the ODIP Rosetta Stone vocabulary service will interact with the GEO-DAB brokerage service in order to expose by CSW and AOI-PMH XML metadata from the regional systems with harmonised vocabularies.

Previously DAB has already developed a functionality for semantic discovery, utilizing textual terms expansion and navigation. This is documented in the publication "Methodologies for augmented discovery of geospatial resources", available at:

https://books.google.it/books?hl=en&lr=&id=-rOeBQAAQBAJ&oi=fnd&pg=PA305&ots=Jt7S8575W8&sig=7vYboPsrYQDHLBVpRpenIkNptqg#v=onepage&q&f=false

This experience will be taken into account when working out the last bulleted action.

2) establishing horizontal interoperability by developing an ODIP interface using WMS – WFS and possibly OpenSearch:

The metadata entries for the three regional systems are included in the global GEOSS and ODP portals. The aim of the expansion is to develop and publish also a Discovery and Access service at the ODIP portal, using WMS-WFS and possibly OpenSearch protocol.

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

As a first step MARIS has been in contact with AODN and NCEI to identify possible existing WMS-WFS and OpenSearch services at the regional systems:

AODN: The metadata describing all the dataset collections available on the AODN portal (https://portal.aodn.org.au/) are accessible through an instance of Geonetwork (https://catalogue-portal.aodn.org.au/geonetwork/srv/eng/main.home). Each collection XML contains Lat-Lon box details which can be used for feeding a map layer at collection level. Moreover, in the online resources section of each collection XML WMS and WFS services are available such as 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 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 the layer used in step 3 of the AODN portal when the user can select multiple download options.

Analysing the XML as provided by the GEO-DAB brokerage service for AODN it appears that the WMS service is included; however the tag with the indication of the specific layer is missing. CNR will undertake action to correct this. Also CNR will update the harvesting as progress is made with the AODN portal and catalogue which now includes also entries from other Australian organisations next to only IMOS.

NCEI: The Oceans section of NCEI (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. Combined, these technologies enable NODC to provide access to its data holdings and products through some of the commonly-used standardized Web services. The Geo portal (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 Lat-Lon box details which can be used for feeding a map layer at collection level. There are not WMS and WFS services mentioned 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 brokerage harvests from the same Geo portal. Analysing the XML as provided by the GEO-DAB brokerage service for NCEI it appears that the link to the image of the detailed map are included.

It can be concluded that all three services are thus providing Lat-Lon box information concerning the data collections and the links in the collection metadata for displaying a detailed map at granule level of the individual observation locations that are part of a collection, in WMS for SeaDataNet and AODN, while as static image for NCEI. There are no overarching WMS / WFS or OpenSearch services from AODN and NCEI at catalogue level.

Following this analysis the next steps are planned:

- CNR will upgrade the brokerage for AODN concerning used catalogue and including the WMS layer tag in the XML output; MARIS will use the database with harvested XML entries (see ad 1 above) to extract the Lat-Lon boxes for all collections to feed an overarching WMS – WFS service;
- MARIS will build a common WMS – WFS user interface on the ODIP portal, giving search options towards the connected regional systems at metadata level

- The WFS will retrieve the collection metadata, including a detailed map of the individual observation locations within a collection, using the identified WMS or image links;
- Possibly the “ODIP Rosetta Stone” vocabulary service (see ad 1) will be integrated for harmonisation of used vocabulary terms
- Analyse and arrange a dynamic maintenance of the user interfaces.

Note: As part of the earlier EuroGEOSS project CNR has developed advanced REST/JSON APIs to interface with the GEO-DAB service. See: http://api.eurogeoss-broker.eu/dab/api-rest-docs/. It will be analysed by MARIS if these APIs might be used for (part of) the required functionality.

3) setting up 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:

For this subject so far limited activity has taken place; however this will alter in the second period, also under influence of the new SeaDataCloud and EMODnet projects that will give extra focus to INSPIRE compliance and applying the INSPIRE implementing rules at data level.

A start has been made with the ODIP II activities focused on the NetCDF (CF) format. NetCDF is already an OGC standard thanks to the work of ODIP partners CNR (Stefano Nativi) and UNIDATA (Ben Domenico). NetCDF (CF) is widely used in the oceanographic and meteorologic communities. Unfortunately there are various flavours of NetCDF (CF) in use by different observing programmes and research groups. Therefore it is a challenge for ODIP II to develop and recommend common netCDF (CF) profiles for the marine community. Thereby it should be noted that NetCDF is to be considered as an exchange format and not as a storage format.

A first discussion took place during the 1st ODIP II Workshop. There are several CF attributes lists which cover different topics and the groups behind these lists are not working together. Some of these groups extend the CF terms but not in a compliant manner. In CF there is quite a number of degrees of freedom. The guidelines are loose, the set of conversions are branched, two different versions deal with gridded data, 1.5 and 1.7, while the 1.6 version deals with point data. Some people label data as CF compliant but it is not, even if they pass the CF checker. The official checkers are for gridded data only 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 checker can be modified to accept multiple plug-ins to satisfy different needs.

It is realised that there can not be one global implementation of CF conventions in great details because different communities have and will keep having different needs. For that purpose BODC (Roy Lowry) proposes a layer structure: the CF conventions at the bottom, then a layer with the specific community conventions for all data types. In fact, this was already done in the SeaDataNet NetCDF (CF) format (no gridded data) where a part of the profile included an attribute with the SeaDataNet parameter codes in addition to the standard name. Other layers on top could be added for specific types of data, e.g. for bathymetry. BODC proposed that this approach could be adopted for developing CF compliant NetCDF profiles for specific gridded datasets.

Another aspect of the NetCDF challenge is to gather and formulate best practice. For instance, usually people include one instrument in one NetCDF files, but communities such as OceanSITES include multiple instruments which makes management more difficult.

Furthermore CNR (Stefano Nativi) has participated on behalf of ODIP II in the NetCDF Summit Workshop, 24-26 May 2016, Boulder, USA, which was organised by UNIDATA and funded by NSF (see http://www.unidata.ucar.edu/events/2016CFWorkshop/#home).

CNR participated in particular to advance CF - NetCDF (also in the perspective of the SeaDataNet NetCDF (CF) and the INSPIRE data models) and contributed with a presentation touching upon inter alia:

- Metadata conventions for netCDF -spanning across datasets aggregation and granularity;
- Injecting vocab URIs and Linked Data elements into netCDF.

The presentation can be downloaded from:

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

This presentation was also triggered by Jonathan Hodge (CSIRO) on the relationship between NetCDF header metadata and standalone "document" metadata. There is often a situation where we have software such as THREDDS which can generate ISO metadata from the headers as well as a standalone ISO metadata record in something like Geonetwork. It would be useful to have a recommended approach regarding the best way to marry these two together and avoid the duplication that we often end up with. CSIRO has started using the "metadata_link" attribute in the NetCDF global header with a link to a metadata record. In this example, it would be useful if tools such as THREDDS find that and make use of it in the generation of its metadata.

A comparable example can be found at SeaDataNet. SeaDataNet delivers downloading users ODV data files with limited metadata, accompanied by a CSV file with the full CDI metadata of all the related ODV data files. It happened that users tossed away the CDI metadata, only used the ODV data files and then complained about the scarcity of metadata. This is being tackled now by SeaDataNet by the ODV software that users can freely download and use for various analyses, checks and visualisations. An extra functionality has been added that the CDI metadata is read and written into the ODV files during ODV import. This creates so-called metadata enriched ODV files which stand-alone provide users with the required information.

Further steps are planned such as:

- Compile an inventory of the CF profiles that are being used within the ODIP community;
- Compile an inventory of the CF checkers, who are using which, what can be plugged in;

These inputs will provide a basis for the further activities.

2.2 ODIP Prototype 2+

The earlier ODIP Prototype 2 is the basis for the expansions and it concerns Cruise Summary Reporting from Europe (SeaDataNet), USA (R2R) and Australia (MNF) towards the global POGO portal. In the start situation there is already CSR V3.0 reporting from Europe and USA while Australia has undertaken preparatory mappings.

The ODIP Prototype 2+ is led by BSH with expected contributions from partners from the three regions. Activities have been undertaken and/or are plannned for each of the targets of the **ODIP Prototype 2+** project:

1) **further population of the CSR directory by ODIP II regions;**
2) **analysing the upgrading of the CSR metadata format, Schema and related tools and services following earlier R2R suggestions;**
3) **publishing CSRs also in RDF by means of a SPARQL endpoint.**

In the reporting period activities for ODIP Prototype 2+ have concentrated on the first target. Activities for the other two targets are planned in the second period when optimal synergy can be reached with the new SeaDataCloud project. The progress of activities for the first target is as follows.

<u>CSR Statistics</u>

The following table gives the total CSR submission since 1. April 2015 from all cruises (including non-POGO ships).

Partner country	Total submissions (including updates)	Increase for POGO
Europe	1706	451
USA	748	704
Australia	1	1

Table 2.2.1: CSR submission since start ODIP II project

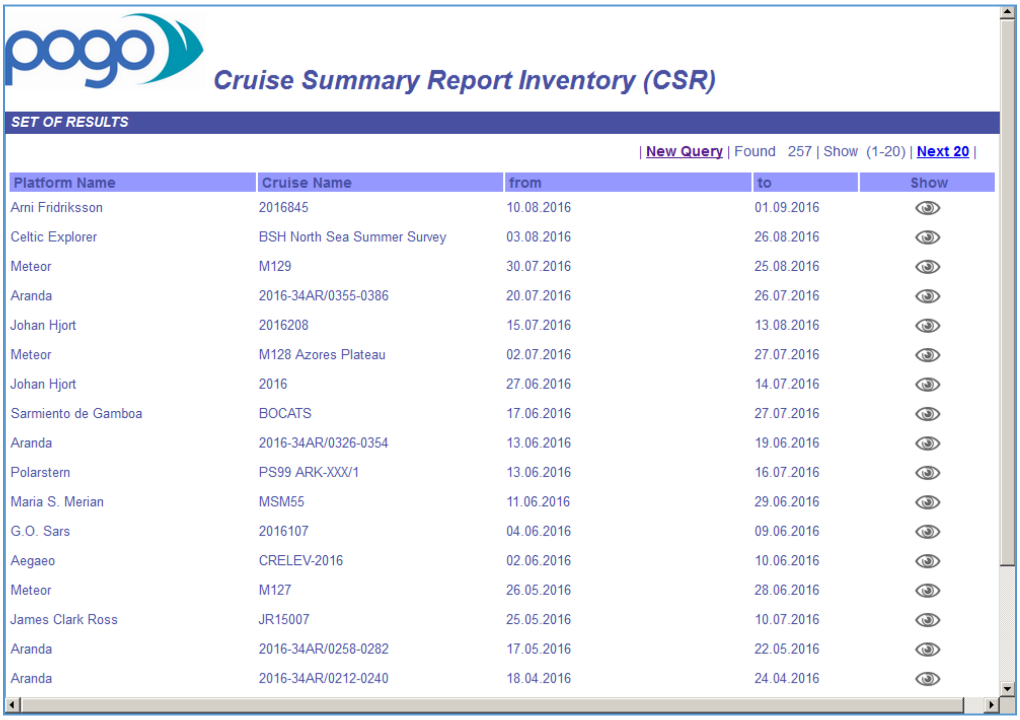
Unfortunately the Australian partners had reorganisation issues and could not provide much input. CSIRO hopes to be able to provide CSRs of the new ship “RV Investigator” beginning from next year.

The next table gives the CSR entries that qualify for the POGO portal as these concern ocean-going vessels. Since April 2015 257 CSR entries have been published on the POGO CSR website.

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 2.2.2: CSR submission to POGO portal since start ODIP II project


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	
G.O. Sars	2016107	04.06.2016	09.06.2016	
Aegaeo	CRELEV-2016	02.06.2016	10.06.2016	
Meteor	M127	26.05.2016	28.06.2016	
James Clark Ross	JR15007	25.05.2016	10.07.2016	
Aranda	2016-34AR/0258-0282	17.05.2016	22.05.2016	
Aranda	2016-34AR/0212-0240	18.04.2016	24.04.2016	

Figure 2.2.1: CSR published at the POGO portal

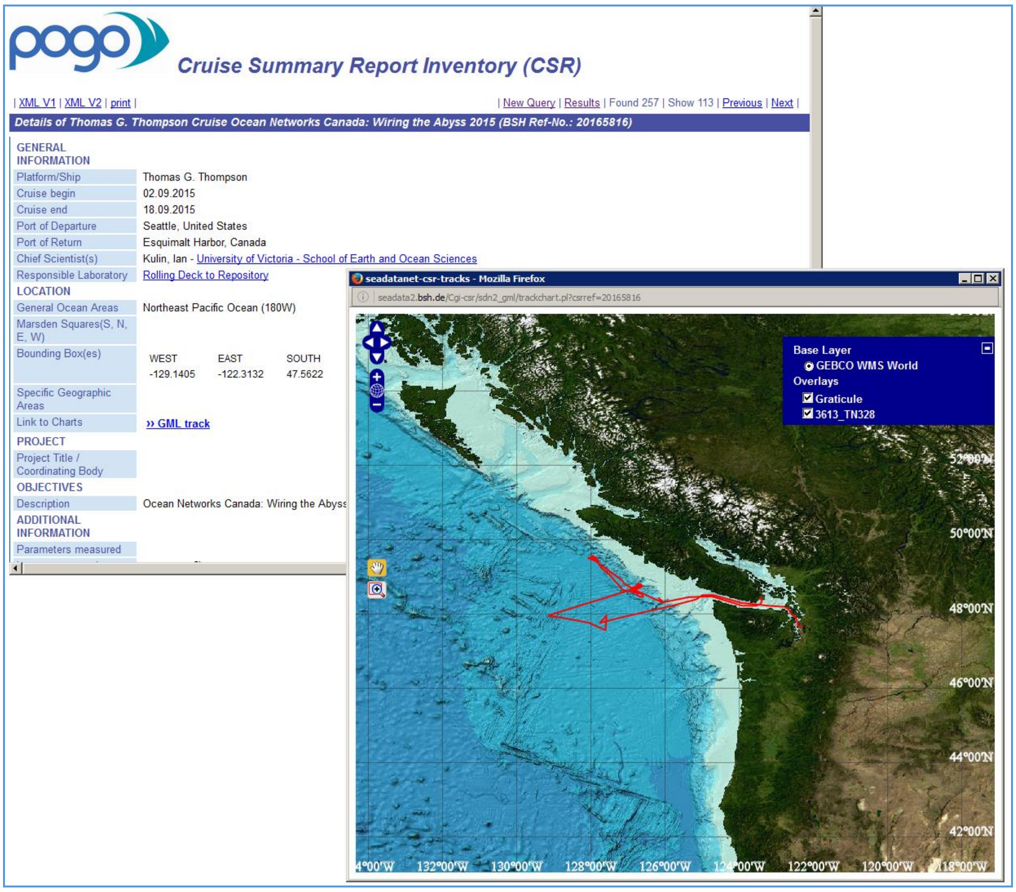


Figure 2.2.2: Example of dynamic cruise chart via WMS at the POGO portal

The CSR Schema which supports tracks in GML enables WMS services offering detailed information on the navigational tracks. An example from an R2R cruise is shown above.

Harvesting progress

Since April 2015 the harvesting has been operating effectively and steadily retrieving CSR records from connected partners on a weekly basis. The contents are automatically validated for mandatory fields, vocabularies and some basic consistency checks on e.g. coordinates (bounding boxes and moorings) and reference dates etc. A harvesting portal has been implemented for further visual checks by BSH and by the connected partners, which is illustrated in the image below. Up till now only 4 centres are connected for regular harvesting. Further centres have started to build up their GeoNetwork server for CS-W harvesting. However there are still some issues with firewalls. It is expected that the new SeaDataCloud projecf will stimulate several NODCs to upgrade their way of submitting new CSR entries. At present it is done mostly by sending XML entries by e-mail.

SeaDataNet

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
20087127	8030070	Pourquoi pas?	ESSRESON08-1	18.07.2008	28.07.2008	IFREMER / TSI/AS- ACOUSTIC SEISMIC	OK

Figure 2.2.3: Illustration of the CSR harvesting portal

GeoNetwork

All submitted CSRs (via online, email, ftp, harvesting) are also accessible from the BSH GeoNetwork portal http://seadata.bsh.de/geonetwork-sdn/ under resource type 'series'. The records can be downloaded as XML-, RDF- and PDF-format.

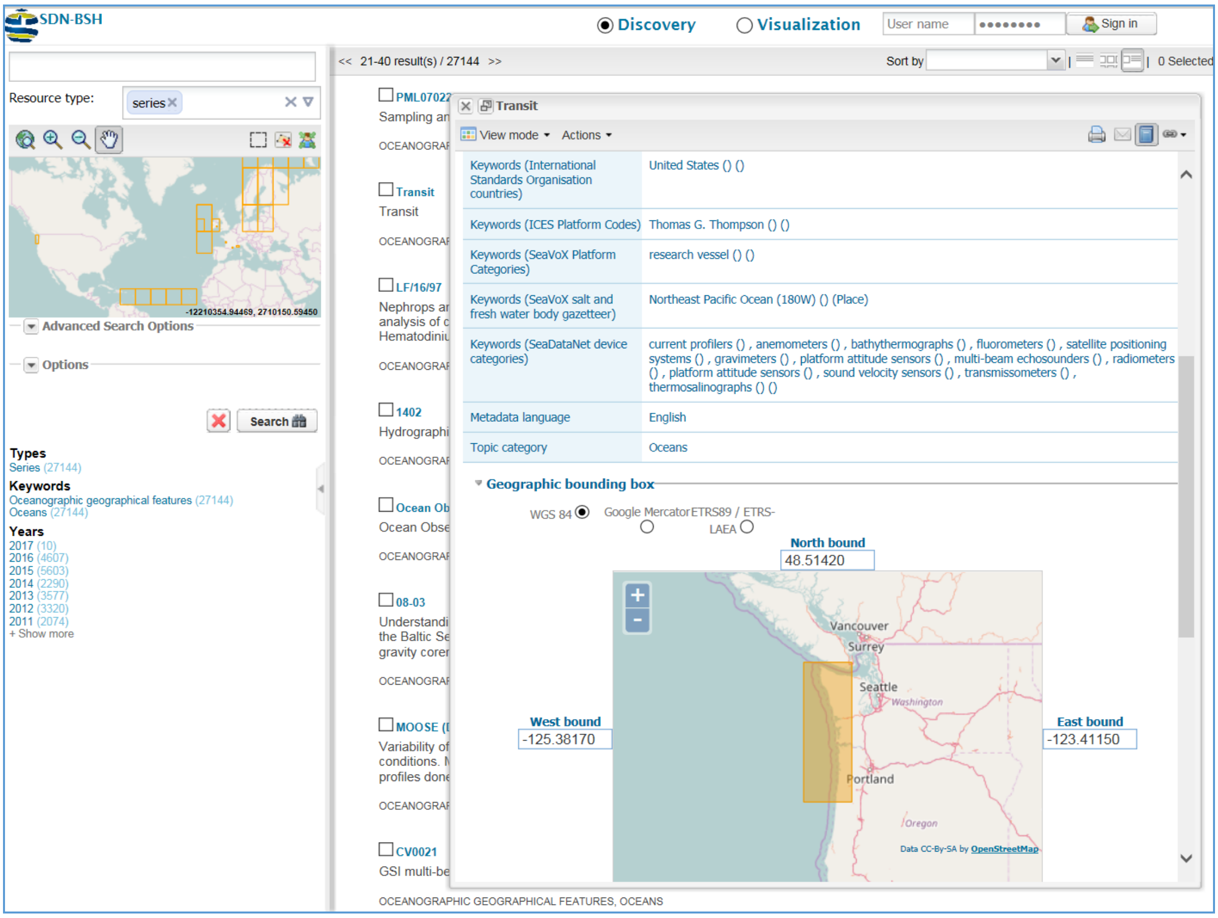


Figure 2.2.4: BSH GeoNetwork portal for CSRs

CSR harmonisation with ICES

A large number of the CSRs in the central inventory originated from the ICES ROSCOPs from 2004 and before. The contents of these legacy records could not be updated to the present standards by the SDN partners due to missing metadata. Among these records are a large number of cruises from Australia, Canada and the USA containing valuable information on earlier cruises.

In the last years ICES has started to partially adopt the SeaDataNet standards and vocabularies. BSH created a special CSR web service such that ICES can also have access to the SeaDataNet CSRs for the cruise metadata to their datasets. Since July 2015 the ICES cruise metadata are daily synchronized with the SeaDataNet central CSR inventory based on the “amendment date” of the record.

In the meantime ICES also has a CSR web service so that BSH can harvest the updated and upgraded ICES legacy as well as new CSR records. The harmonisation from ICES to the CSR central inventory has recently been implemented. All the Australian, Canadian and US CSRs (via the ICES web service) have been updated which BSH has in the CSR central inventory from ICES. Also most of the other legacy CSR records could be matched to the ICES CSR reference numbers so that in the future it should be no problem to update these records for the POGO and SeaDataNet portals. All these records now have ICES as collating centre and the ICES reference number as local CSR ID in the CSR database. All other records which could not be matched have been set the "deprecated" flag so that these do not appear on the CSR homepage anymore. Presently it is not yet feasible to synchronise the ICES CSRs automatically via web service because ICES still uses the old CSR format (**NOT** ISO19115-2) and many mandatory fields are still missing. The aim is to have an automatic process in the future for the harmonisation of CSRs with those from ICES. Fisheries and Oceans Canada (Mathieu OUELLET) has worked closely with ICES to update all the Canadian CSRs. He also indicated interest in joining the CSR harvesting process, perhaps in 2017.

Vocabularies

The primary vocabulary requirement for Prototype 2 is the extension of the content of the SeaDataNet C17 platform vocabulary served by NVS to cover all vessels of interest to the US R2R project plus all cruises in the CSR database at BSH, including V0 records. Existing coverage of Australian research vessels is considered adequate. This work is being done in such a manner that it makes positive steps towards full synchronisation between C17 and the platform code system at ICES.

Historically, there were three platform metadata systems. One was at ICES, issuing platform identifiers known as ICES codes. Two more were operated by different groups in the US NODC, issuing identifiers known as NODC codes and WOD codes. In 2004 it was agreed to rationalise this through full harmonisation of the ICES and NODC codes with a mapping to the WOD codes. This has been a work in progress ever since. Significant progress has 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 to modify the vocabularies then carry the changes through into legacy data. These issues were not confined to obscure vessels – at one time the primary UK research vessel RRS Discovery was affected. 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.

SeaDataNet took the pragmatic decision in 2005 to create C17 as a subset of ICES ship codes that had been carefully checked, sufficiently researched to be sure that any issues such as multiple hulls 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 a SeaDataNet partner preparing a Cruise Summary Report identified a vessel as missing from C17. BODC worked with the partner 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 wasn't in the ICES system then it was added.

The work for ODIP Prototype 2 is a case of extending the work described above to cover the R2R ship list and also a list of ships provided by BSH. At the start of ODIP II there were nearly 3000 ship codes for American vessels in the ICES system for which the only identification metadata attribute was the name. Of these over 100 were identified that were research vessels in the broadest sense of the word, which included vessels associated with academic institutions, NOAA ships and naval hydrographic and oceanographic survey vessels. These were prioritised in a project covering all US vessels and have now been added to C17 together with over 1000 American warships. A list of 40 vessels missing from the BSH system has been obtained and a quarter of these have 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 has been significantly assisted through the development of tooling to automate the comparison of metadata held in ICES and BODC systems. This comprises a simple schema in the BODC Oracle system comprising two tables of identical structure. One is populated from C17 by 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 may easily be compared and contrasted using SQL. 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 Prototype 3+

The ODIP Prototype 3+ continues the earlier activities as undertaken for the ODIP Prototype 3 and concerns further elaborating of Sensor Web Enablement (SWE) standards and services in concertation with many on-going regional projects and OGC. It will build upon the earlier results and explore further enhancements and new technological approaches. Thereby also great attention will be given to establishing synergy and tuning between the ongoing regional projects.

The ODIP 3+ Prototype is led by 52°North with expected contributions from partners from the three regions. Activities have been undertaken and/or are plannned for each of the targets of the **ODIP Prototype 3+** project:

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. Several activities have taken place.

SWE Technologies:

The NeXOS project (EU FP7, with participation of 52°North and IFREMER) develops a full Sensor Web architecture and implementation for marine sensors. This comprises on the one hand 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. This has resulted in several developments that were contributed to ODIP Prototype 3+:

- A Sensor Web Viewer for stationary in-situ measurements
- Extension of the Sensor Web Viewer to display data collected by mobile sensors/sensor platforms (see **Error! Reference source not found.**)
- Several SOS deployments at NeXOS partners which can be used for interoperability testing and demonstration
- SOS server supporting optimised data encodings for publishing sensor data (EXI encoding of the SOS Result Handling operations)

Furthermore, the NeXOS activities were also closely linked with the Marine Sensor Web Profile and SensorML metadata editor developments.

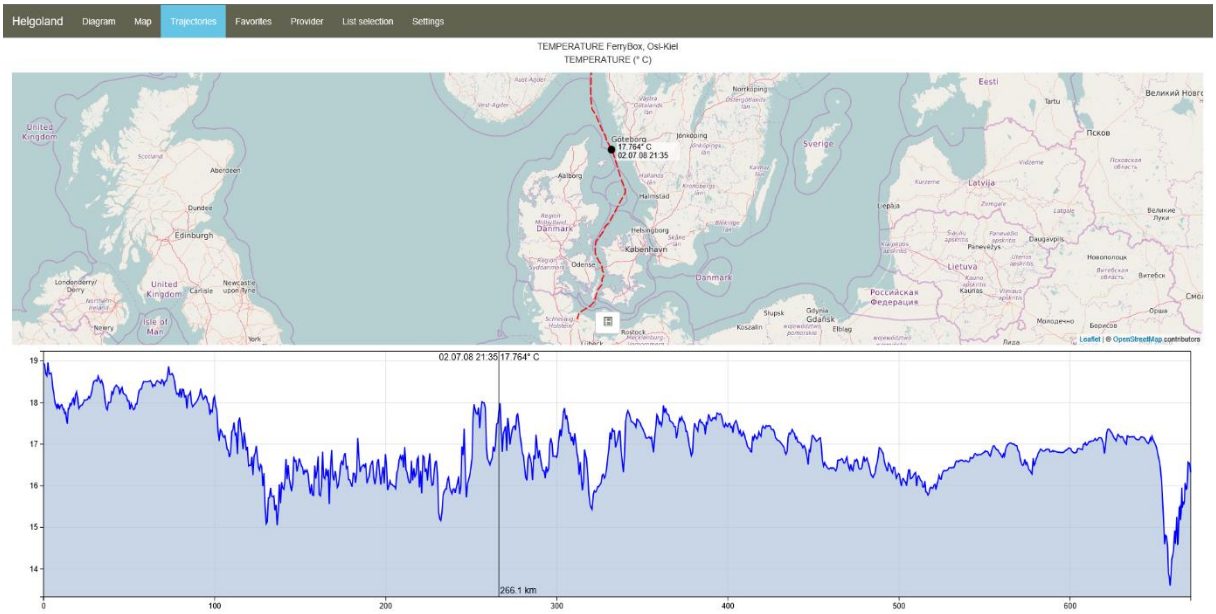


Figure 2.3.1: NeXOS Sensor Web Viewer for Mobile Sensor Platforms

Within the FixO^3 project (EU FP7, with participation of 52°North, IFREMER and further partners), the NeXOS developments were enhanced and adjusted to visualise observation data from different fixed ocean observatories. Currently the SOS deployments are in progress so that more SOS servers are likely available for demonstration purposes of ODIP Prototype 3+ in the next reporting period.

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

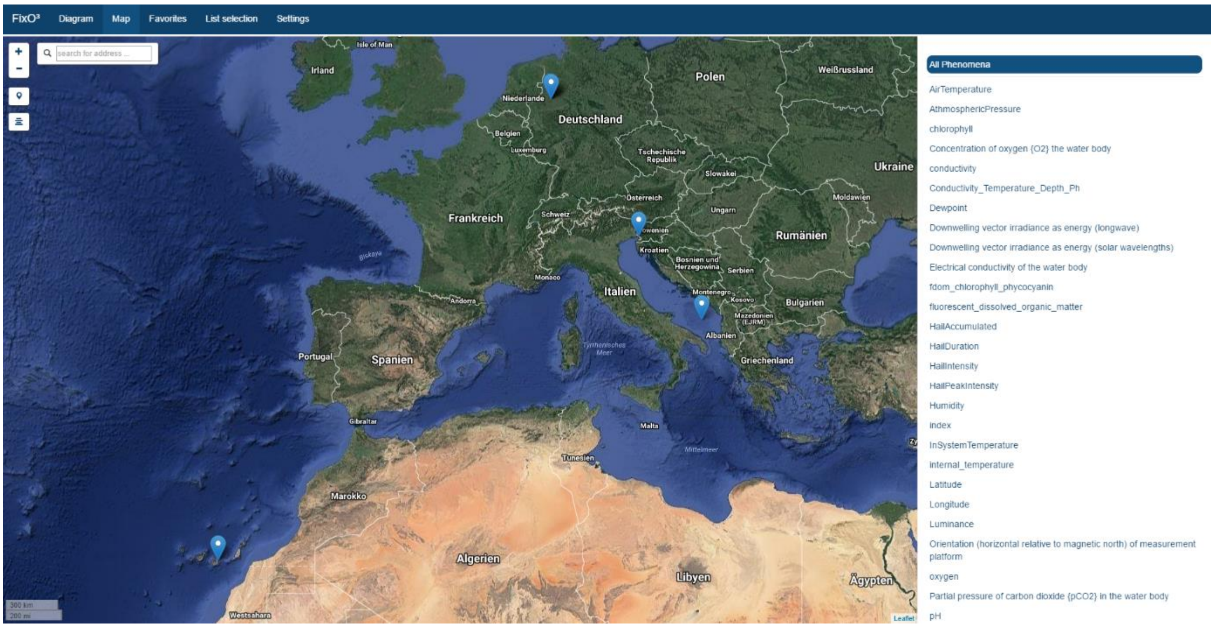


Figure 2.3.2: Sensor Web Viewer (FixO3 and NeXOS)

In this context, also the Sensor Nanny activities of IFREMER need to be mentioned. While the SensorML editor part of this software suite is explained below in a separate section, it also offers a comprehensive set of tools for managing, publishing, discovering and visualising marine observation data and sensors (see **Error! Reference source not**

found.). 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.

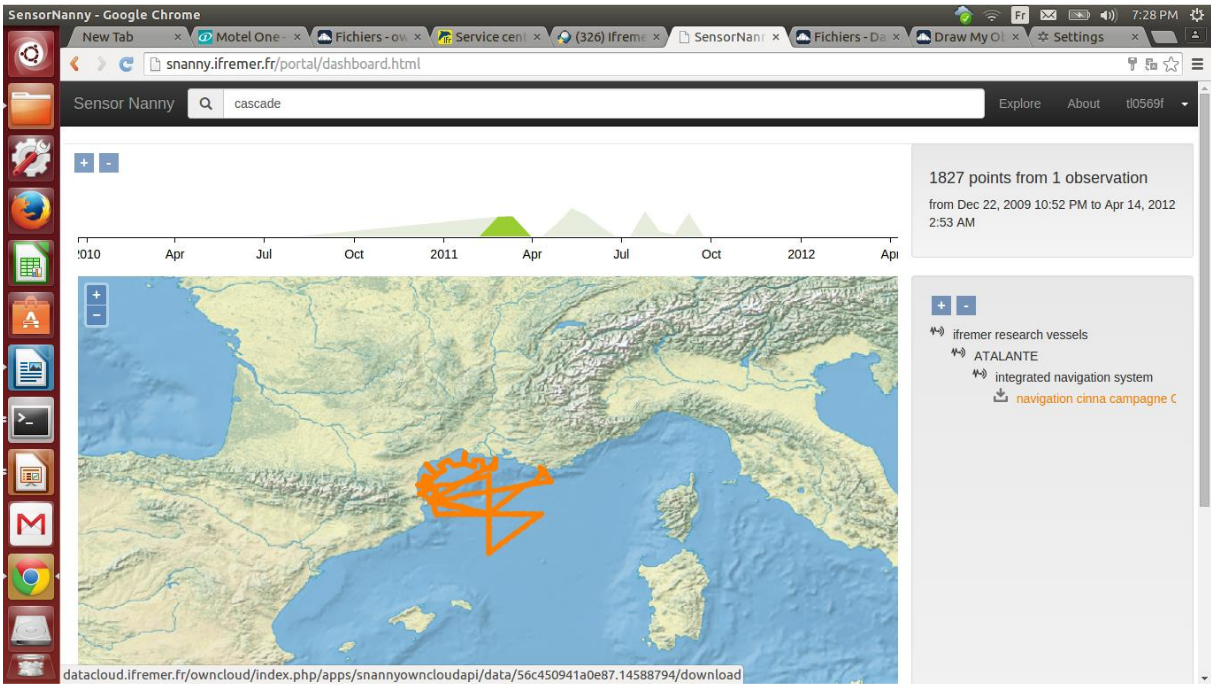


Figure 2.3.3: Sensor Nanny Tool for Data Discovery

Another tool to be mentioned is the SOSInserter developed by TNO as part of ODIP. This is an extension of the FME (Feature Manipulation Engine) of 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 receives the capability to publish observation data sets via the transactional SOS operations on instances of the 52°North SOS. The availability of this tool will lower the barrier for owners of existing observation data to make their data available in the Sensor Web.

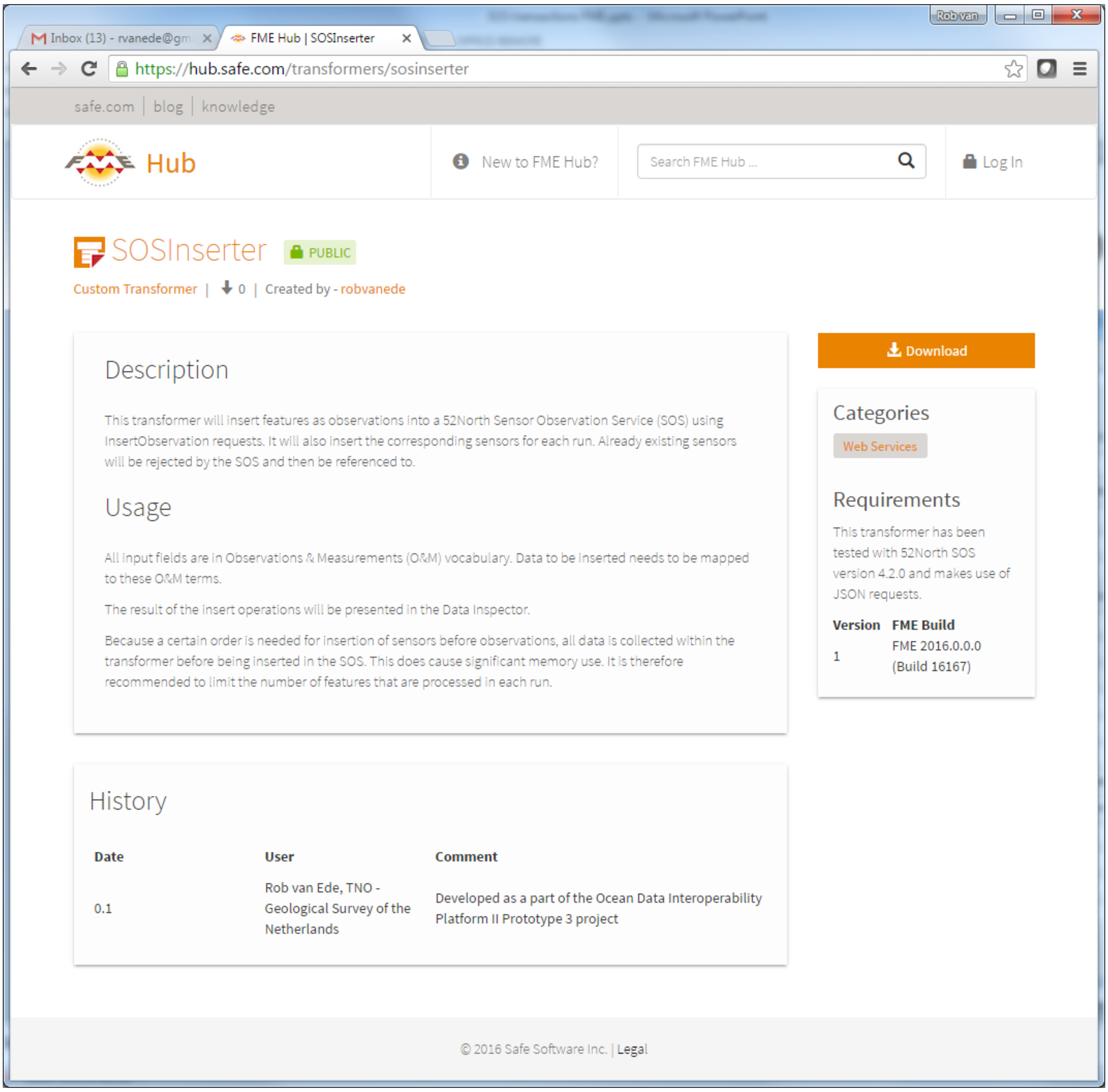


Figure 2.3.4: SOS Inserter Developed by TNO available on FME Hub

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 how the SWE specifications are applied. The OGC SWE standards were intentionally designed in a domain independent manner so that they can be applied in as many scenarios as possible. This idea goes hand in hand with a high level of flexibility within the specifications so that the same goals could be achieved in different ways.

To avoid interoperability issues, a common approach is needed how SWE specifications shall be applied for marine data and sensors. Furthermore, this should be complemented by a use of vocabularies (see below) to ensure not only common syntax but also common semantics.

Consequently ODIP II has joined this activity with partners from several projects and initiatives: AODN, BRIDGES, envri+, EUROFLEETS/EUROFLEETS2, $FixO^3$, FRAM, IOOS, Jerico/Jerico-Next, NeXOS, RITMARE, SeaDataNet, SenseOcean, X-DOMES, and the new SeaDataCloud.

During the predecessor ODIP Prototype 3 project a strong focus of the work has been on collecting usage examples of SWE specification in different projects and organisations. These findings are documented in a Wiki hosted by 52°North (see **Error! Reference source not found.**). This includes not only examples of O&M and SensorML but also endpoints of available SOS instances.

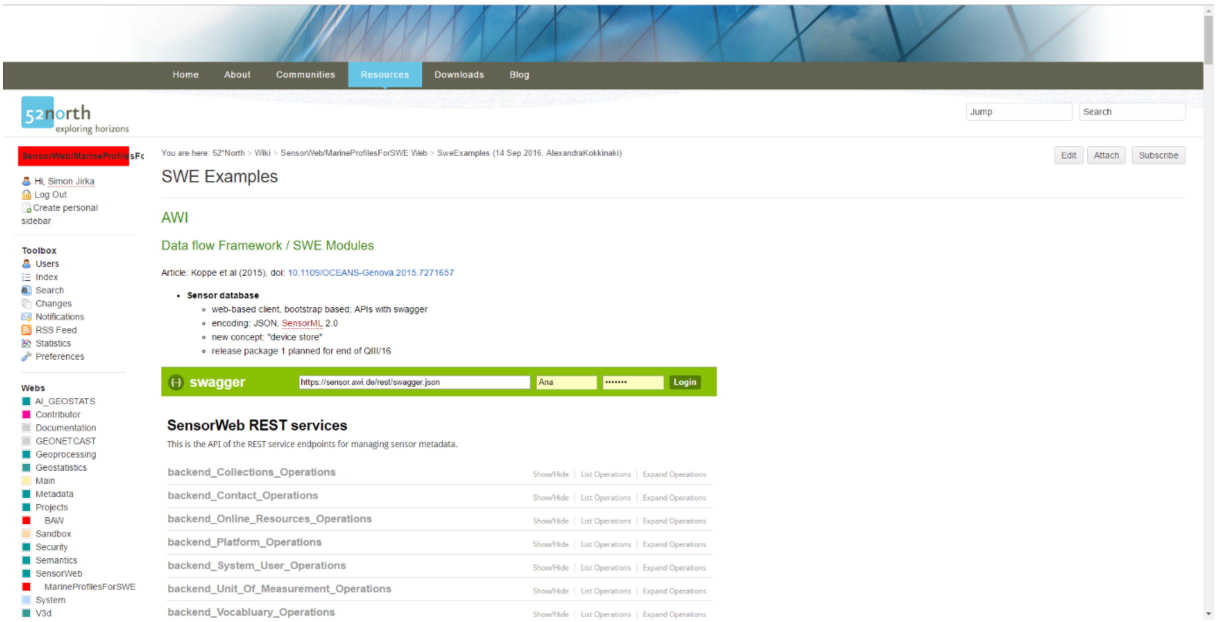


Figure 2.3.5: Marine SWE Profiles Wiki

In addition, a first model 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 actual 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. **Error! Reference source not found.** shows an exemplary overview of the possible elements that describe an instance of an instrument.

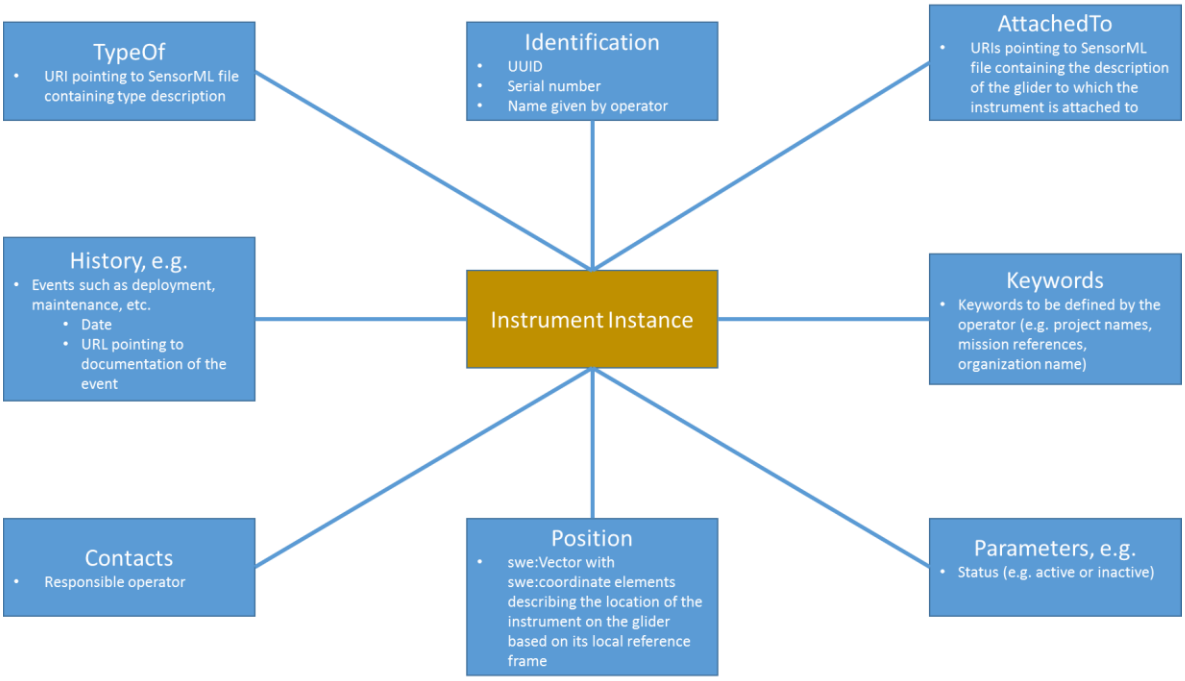


Figure 2.3.1: Excerpt fom the Ideas for the Marine SensorML Profile Development (here: Description of an Instrument Instance)

2) 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 regarding this topic was on identifying open issues and questions:

- How can big heterogeneous spatio-temporal datasets be organized, managed, and provided to Sensor Web applications?
  - What are 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 can views on big data sets and derived information products be made accessible in the Sensor Web?
  - Which service standards are appropriate? SOS would be a natural choice, but what about WPS, WCS/WCPS?
  - 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 cooperation between 52°North and AWI.

3) Use of lightweight technologies such as JSON and REST as complementary technologies:

In addition to the XML-based SWE standards, there are further technologies which allow more lightweight implementations for handling observation data. Especially JSON and REST are becoming more and more important when building lightweight applications.

Thus, as part of Prototype 3+ potential advantages such as reduced data volumes as well as easier data processing on (potentially resource constrained) client platforms are investigated. Currently there are two activities to report:

Especially the work of CSIRO has shown how marine observation data infrastructures could benefit from lightweight technology. In order to investigate the question how this could also be mapped to the OGC Sensor Web Enablement standards, it is planned to develop a mapping of the concepts behind the Marine SWE profile specification to JSON and REST (52°North). However, as the Marine SWE profile is still in development, this activity is planned to be conducted during the next reporting period.

The second activity concerns the 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. First results are expected for the end of 2016.

4) Use of RDF-based approaches – vocabularies - for supporting the discovery of marine sensors and data sets:

In close conjunction with the Marine Sensor Web Profiles definition, several partners, especially BODC which took a leading role, have worked on the creation of an overview of terms that are needed within SWE documents. The resulting list of terms has been documented in the Marine SWE Profiles Wiki (see earlier **Error! Reference source not found.** and Figure 2.3.7 below).

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

Furthermore, this approach opens the door for applying in the future more powerful semantic Web technologies.

The work on vocabularies by BODC is further specified in the following paragraph.

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.
Capabilities	Latency	CAPB	Latency		The time taken for a device or system after receiving an input signal to change its output signal from its previous state to a

Figure 2.3.7: Excerpt from the Overview of Collected Terms for the Marine Sensor Web Profiles

BODC SWE Vocabularies:

The flexibility of SensorML can result in many different variations of sensor descriptions which reduce interoperability and discoverability via the web. To resolve this, 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 standards.

The SWE Marine Profiles group, was created as a solution 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 broadcast to all of the other mailing list members. For the purposes of vocabulary building, the list was given the responsibility to act as the SensorML vocabulary content governance, which is important in order to stay up-to-date and in sync with ongoing developments.

The publication of SensorML implementations by different projects, revealed the lack of published vocabularies for term and property 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 by the use of vocabularies: the **term definition** and the **term value**.

For "term values", SWE Marine Profiles members agreed to use existing concepts in NVS2.0. The following collections were identified to adequately serve term values:

- Observable property: NVS2.0 Collections P01, P07
- Instrument Type: NVS2.0 Collection L05
- Platform Type: NVS2.0 Collection L06
- Roles: NVS2.0 Collections G04, C86
- Feature of Interest: NVS2.0 Collection C19
- Manufacturer: NVS2.0 Collections L35, C75

NVS2 Collection P01, which lists terms used to describe individual measured phenomena and P07 which is a list of the Climate and Forecast standard names, have been nominated to serve SensorML observable properties.

The L05 collection, which lists device categories, is used for the classification of instruments and procedures. L06, in the same respect, provides a list of platform categories, to be used for classifying platforms. G04 and C86 list roles and are used to populate the SensorML role property. C19, which is the Salt and Fresh Water Body Gazetteer, can be used to create a rich list of features of interest. L35 and C75 can be both 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, holding a unique URI, listing a set of domain relevant terms. Following 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 2.3.1: Table listing the URI and the description of the published SensorML collections

As part of the SenseOCEAN 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 datasets metadata as 5 star Linked Data and OGC SWE (SensorML, Observations & Measurements) implementations. Thus sensors become more discoverable, accessible and useable via the web and their data can be combined with other sensor or Linked Data datasets and form knowledge.

In the SenseOcean project, sensor metadata and data are encoded in OGC standards where sensors and instruments transmit unique resolvable web linkages to persistent OGC SensorML records published in BODC. The URIs, based on content negotiation, resolve to either Linked Data or Open Geospatial Consortium Sensor Web Enablement (OGC SWE) descriptions. Sensor data, being either observations or sensor metadata, are also delivered through a SPARQL endpoint and a Sensor Observation Service (SOS) server thus being more discoverable and accessible through the web.

Sensor owners need to register their sensors on the web, where they receive a unique URI for each registered sensor. As shown in figures 2.3.8 and 2.3.9 users can add specific information for their sensor Instances along with event information such as deployment, calibration installation etc.

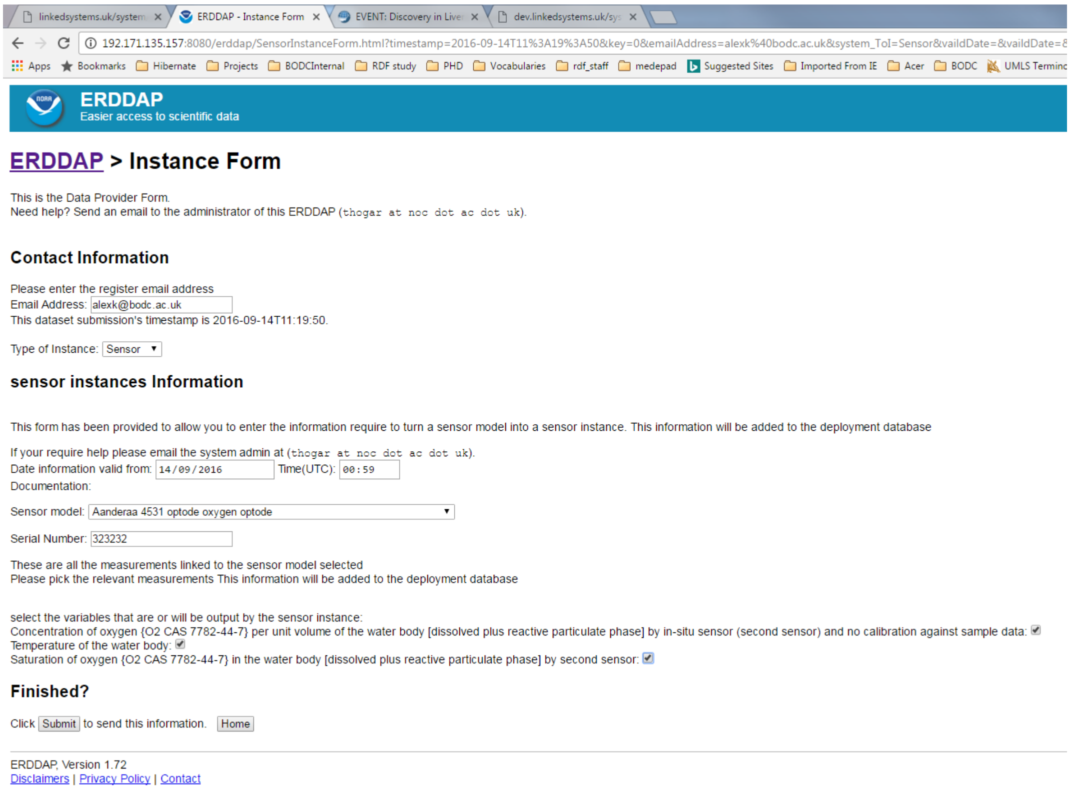


Figure 2.3.8: Instance Form

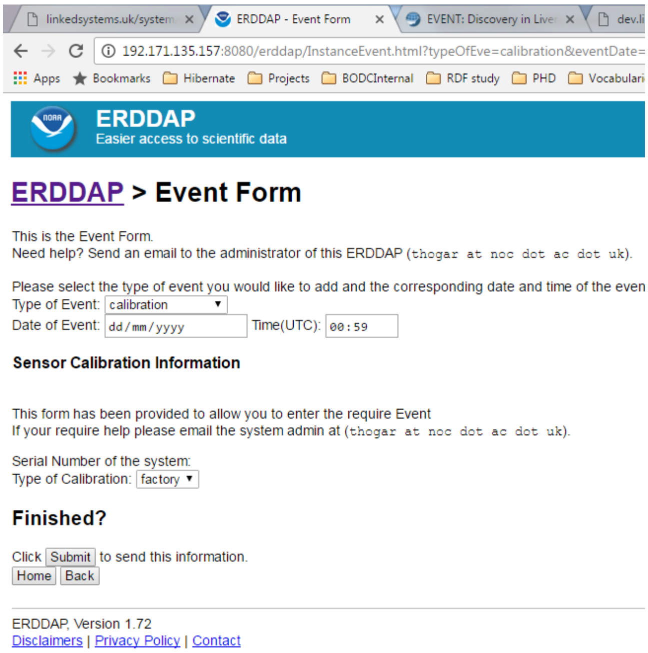
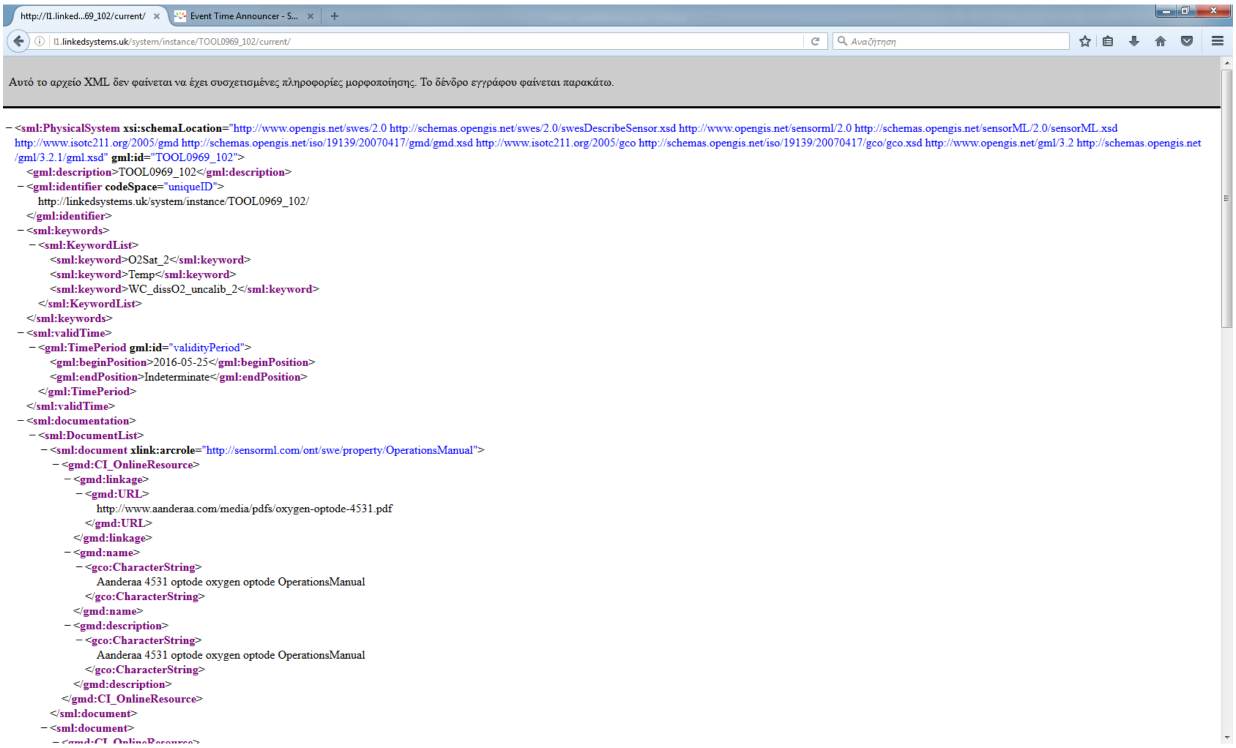


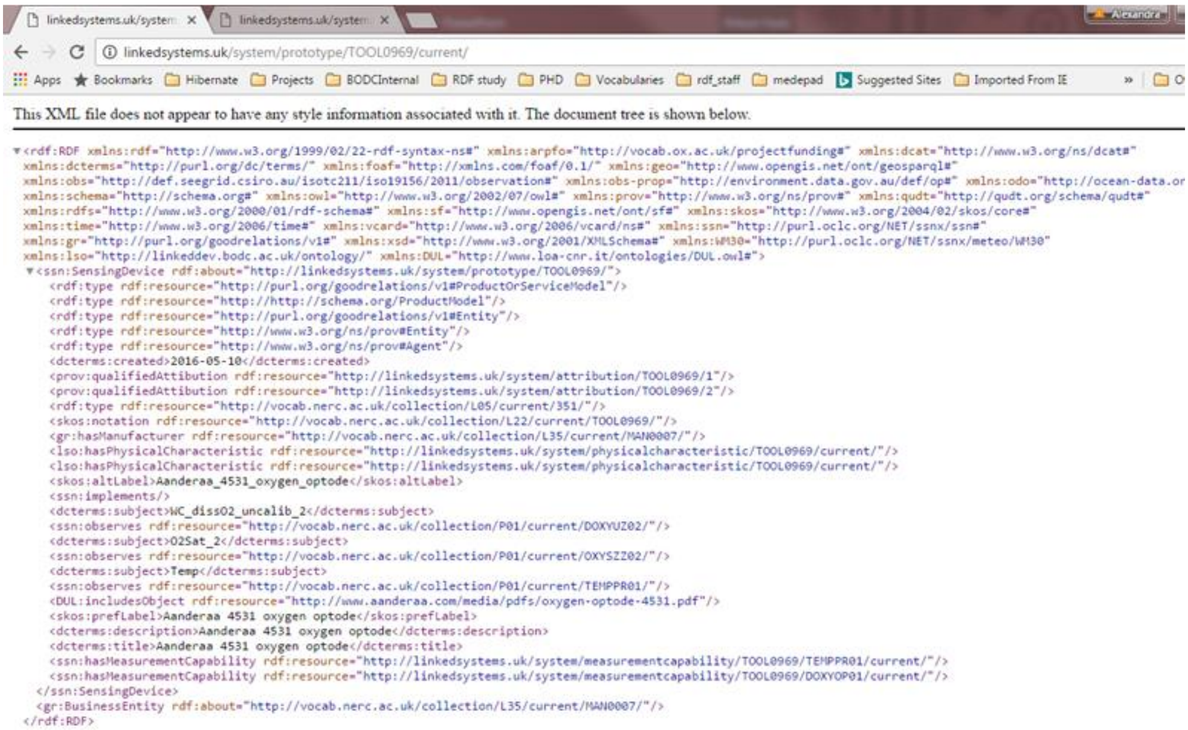
Figure 3.3.9: Event form



Αυτό το αρχείο XML δεν φαίνεται να έχει συσχετισμένες πληροφορίες μορφοποίησης. Το δένδρο εγγράφου φαίνεται παρακάτω.


```
- <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 2.3.10: Sensor Instance defined by URI in SML.



This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
▼<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 2.3.11: Sensor model defined by URI in RDF.

5) 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 is a different challenge. While the data measured by sensors can be inserted into OGC Observations and Measurements documents in an automated manner, the creation of the corresponding metadata is usually a much more work intensive task. Thus, if sensor manufacturers and operators shall be encouraged to provide comprehensive metadata, it is necessary to provide tools which make this process as easy as possible. Consequently there are currently several projects which address this topic. Subsequently, three developments that are part of Prototype 3+ are introduced. It shall be noted that each of these projects have different main focuses but in the future, they would strongly benefit from further cooperation and alignment. An important common topic wll be the use of the vocabularies as they are currently enhanced (see above) in the context of the marine Sensor Web profile developments. This way, the metadata files generated by the presented SensorML editors would rely on common terms so that not only syntactic interoperability is achieved but to a certain extend also semantic interoperability.

Sensor Nanny

Sensor Nanny developed by IFREMER is a comprehensive framework of different components for dealing with marine observation data and metadata. It is openly available via GitHub (https://github.com/ifremer). One of the components of Sensor Nanny is an editor for describing sensor systems and their components (see **Error! Reference source not found.**).

Core features of Sensor Nanny comprise

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

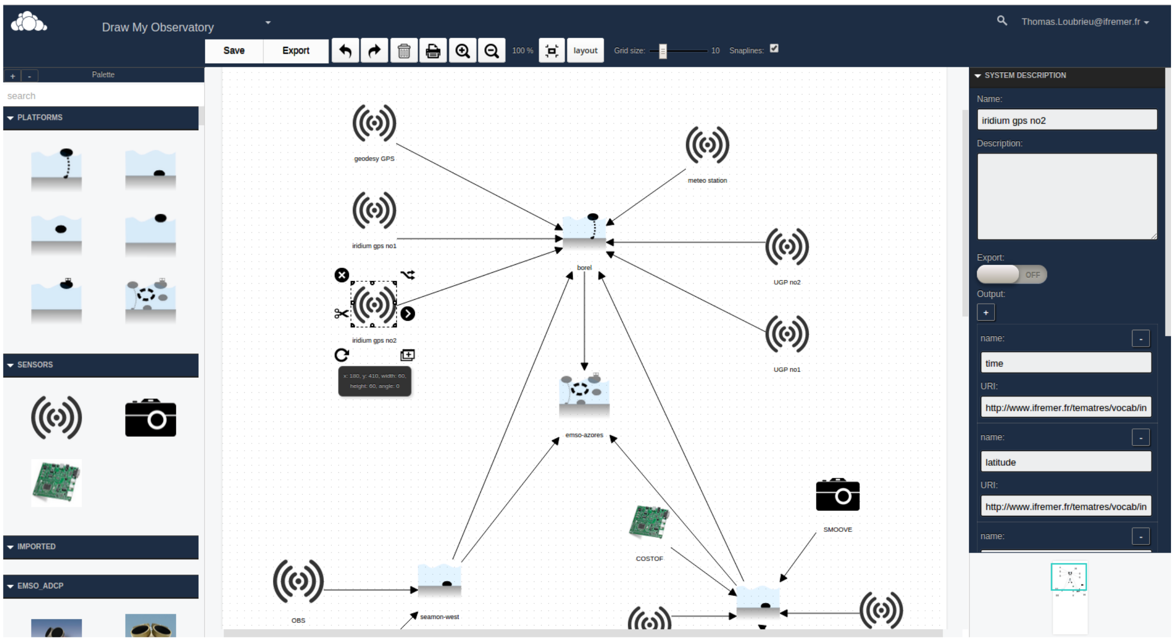


Figure 2.3.12: Sensor Nanny SensorML Editor

RITMARE EDI-NG

EDI-NG is an HTML-based editor for sensor metadata, developed by CNR in cooperation with CSIC. Like Sensor Nanny, it is as well openly available via GitHub: https://github.com/SP7-Ritmare/EDI-NG

Besides the HTML based client component, it also include a separate server component that is used as backend for the editor (e.g. for handling the XML creation). For EDI-NG a core feature is its high level of configurability. This means that this SensorML editor 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 make further constraints on allowed values, etc.

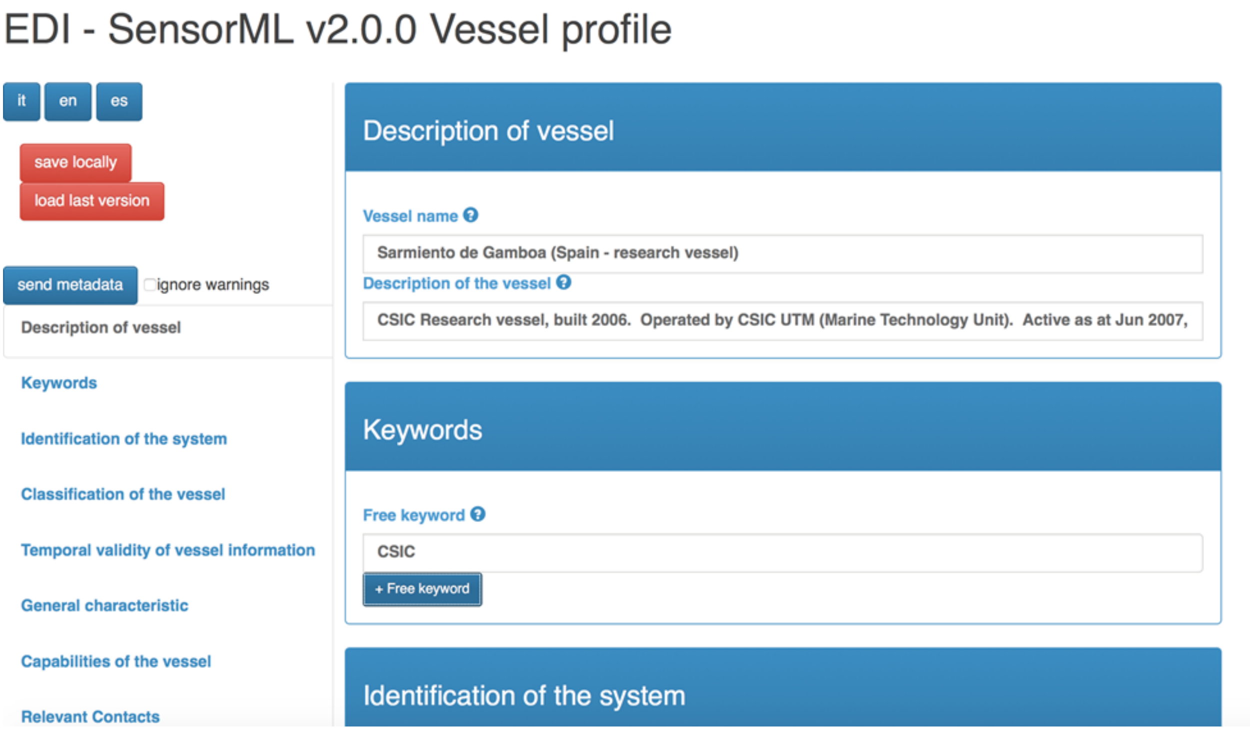


Figure 2.3.13: EDI SensorML Editor

smle

The third SensorML editor contributed to Prototype 3+ is smle, developed by 52°North. This open source client, which is developed as part of the NeXOS and FixO3 projects, is available via GitHub, too: https://github.com/52°North /smle

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

Core features of smle comprise:

- Support the use of SOS servers for metadata storage
- Domain-independent (could also be used beyond marine sensors)
- Web-based implementation
- Support the use of sensor type information from the ESONET/FixO3 Yellow Pages

Currently, there are ongoing activities to investigate, how smle could be enhanced by a profile definition language that allows the configuration of the client users interfaces as well as the validation of content entered by users.

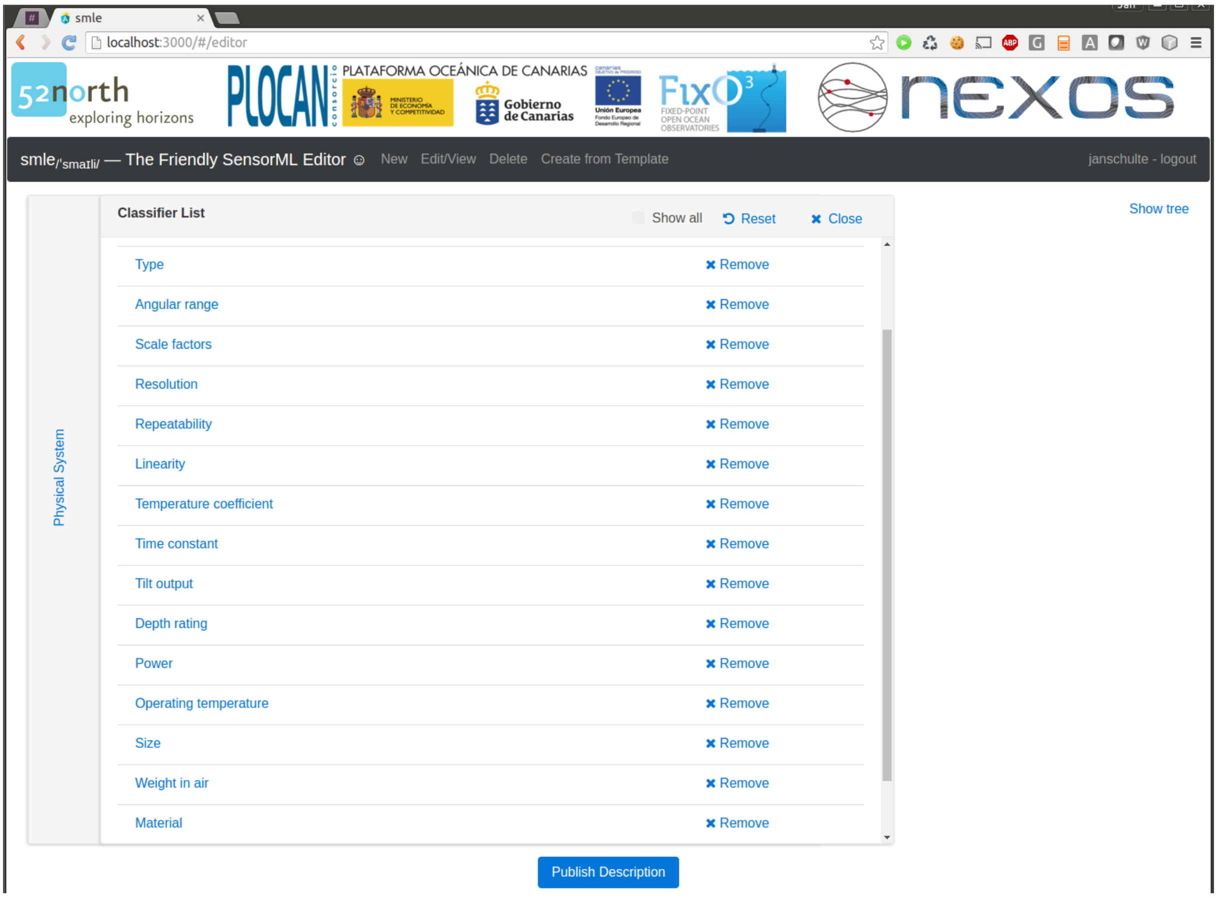


Figure 2.3.14: smle SensorML Editor

Appendix C gives a comparative table between XML editors and has been prepared by CNR.

6) Linking between Sensor Web and Global Infrastructures such as GEOSS

The linking between Sensor Web services and global infrastructures such as GEOSS is another part of this work package. Currently, it is expected to follow a similar approach like in the hydrology domain (GEOWOW project). The general approach will be to rely on the marine Sensor Web profile (see above) to offer a common entry point into marine Sensor Web services. This would allow GEOSS Common Infrastructure components such as the GEOSS Discovery and Access Broker to harvest metadata from marine SWE services and to make these resources discoverable and accessible through GEOSS.

As the marine Sensor Web profile is still in development, this activity will be intensified after the completion of the profile.

7) Promoting Sensor Web Technology

The developments performed within Prototype 3+ were presented at several workshops and conferences in order to promote the application of Sensor Web technology. Especially the following four events should be mentioned in the context of these prototype development activities:

SWE Workshop at Oceanology International 2016

At the Oceanology International 2016 Conference (London - UK, March 2016) a half-day workshop was organised as initiative of the EU Eurofleets2 project in cooperation with a range of other European and US projects, including ODIP II. During this workshop several developments resulting from the ODIP prototype 3+ were demonstrated such as, the Vocabulary developments performed by BODC. This was complemented by a Sensor Web introduction by 52°North and the Universitat Politècnica de Catalunya.

The workshop which had ca. 60 attendees allowed and supported the dialog between research and industry. Furthermore, it showed the need for dialogue and cooperation between different projects, in the way as it is a core element of the ODIP II project.

More information about this workshop is available here: http://eurofleets.maris2.nl/swe-workshop/

EGU General Assembly 2016

Also at the EGU 2016 (Vienna – Austria, April 2016), the opportunity of having a dialogue between the Prototype 3+ participants as well as external researchers was used. Presentations comprised for example the Vocabulary developments (BODC), the work on Sensor Nanny (IFREMER), a poster on the challenges of big data sets in the Sensor Web (52°North) as well as talk on the development of marine Sensor Web profiles (52°North).

GEPW-10

At the 10th GEO European Projects Workshop 2016 in Berlin - Germany, 52°North used the opportunity to present the Sensor Web activities of Prototype 3+ as well as other ongoing, related projects. Special focus of this activity was to raise awareness for the Marine Sensor Web Profile development activities that are an important element of Prototype 3+. The discussions at the workshop helped to identify further potential contributors to these exchange and specification process.

52°North Geospatial Sensor Webs Conference 2016

The 52°North Geospatial Sensor Webs Conference 2016 took place in August 2016 in Münster - Germany. As continuation of the previous series of Sensor Web workshops organised by 52°North this conference brought together ca. 60 participants from research, industry, public administration, and application. During this conference, the topic of applying Sensor Web in marine applications took a significant share so that there were not only presentations of relevant activities (including Prototype 3+ developments such as the Vocabulary activities of BODC and the Sensor Web tools of 52°North) but also direct discussions between involved researchers.

More detailed information about this event is available here:

http://52°North .org/about/other-activities/geospatial-sensor-webs-conference

Figure 2.3.15: 52°North Sensor Web Conference (Key note by Dick Schaap – ODIP II Technical Coordinator)

3. Cross-cutting topics – progress of Vocabularies

At the 1st ODIP II Workshop it was decided to continue also with the cross-cutting topics, such as vocabularies, data publishing and persistent identifiers. In the following progress and results are reported for the vocabularies.

3.1 European developments by BODC

NVS search tool:

BODC has developed a search interface to make interrogation of the NERC Vocabulary Service (NVS) easier through a desktop 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/codes_and_formats/vocabulary_search/

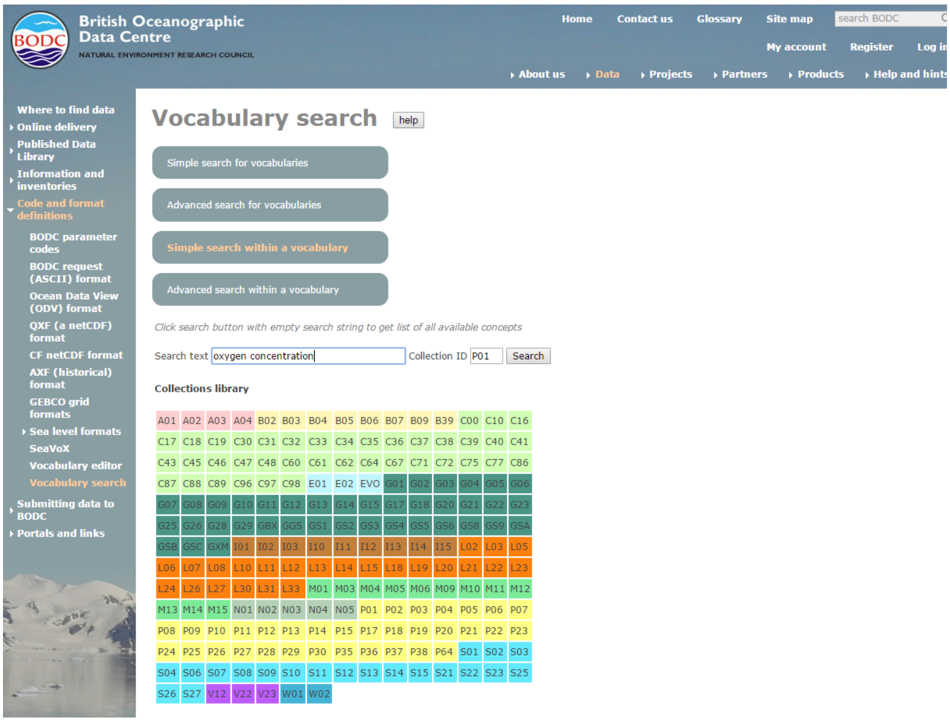


Figure 3.1.1: NVS Vocabulary Search showing “periodic table” of NVS vocabulary content.

Content negotiation:

The NVS pages are now returned using content negotiation, so when a web browser is used to call the NVS the results are presented as html and in the case of a machine-to-machine interaction RDF-XML is returned.

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 3.1.2: NVS vocabulary term presented as human readable webpage in a computer browser through html.

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 3.1.3: NVS vocabulary term presented as RDF XML for machine-to-machine interactions.

Direct vocabulary search URLs

A user can go to a search page for a known vocabulary by adding the vocabulary code to the search tool URL. For example:

https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/L22/

in the case of the SeaVox Device Catalogue. This facilitates providing users with vocabulary search links tied to a particular metadata field in metadata documentation.

NVS editor tool

This tool can be found at:

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

A web based tool for editing vocabularies and associated mappings has been deployed. 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. 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 is provided in Appendix B of this Deliverable D3.3.

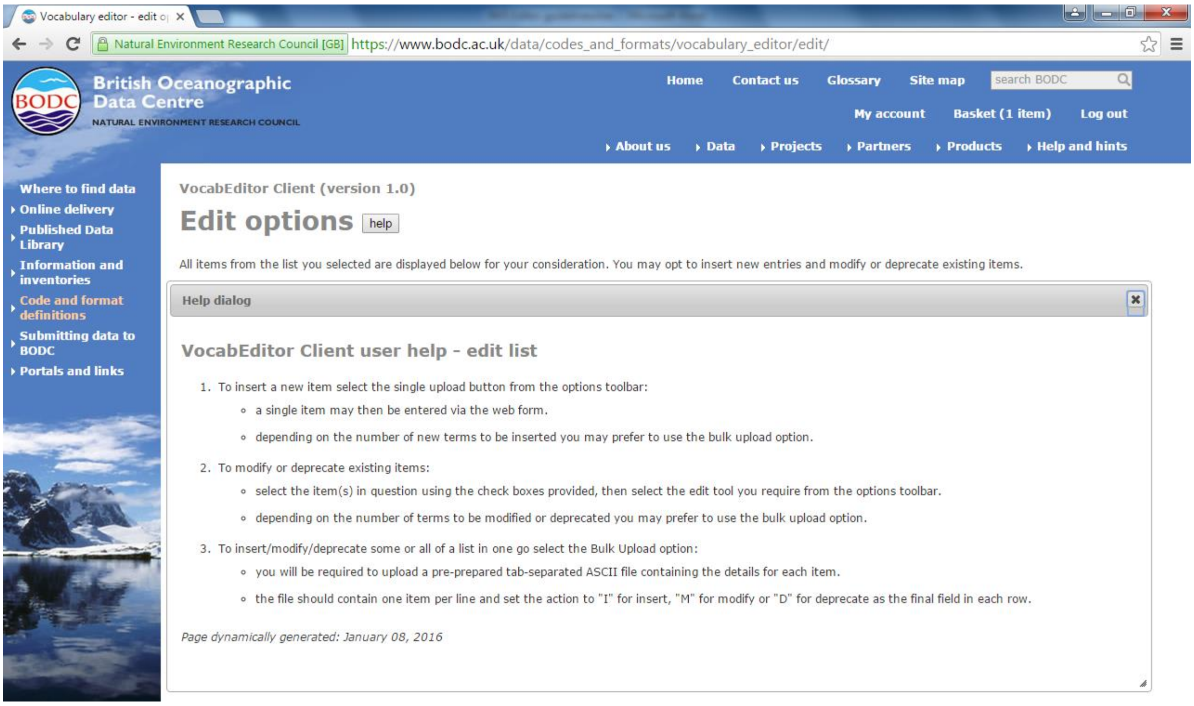


Figure 3.1.4: NVS editor.

NVS Vocabulary builder tool

This tool can be found at:

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

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 built and deployed a prototype vocabulary builder tool to facilitate searching and the 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.

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 {ITIS: 79456: WoRMS 140481}		

Table 3.1.1: Breakdown of the P01 semantic models used to build a vocabulary term with vocabularies used for each pat of the model.

P01 - Chemical substance terms

The vocabulary builder allows a user to construct the preferred label for a term from the underlying semantic model. 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.

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

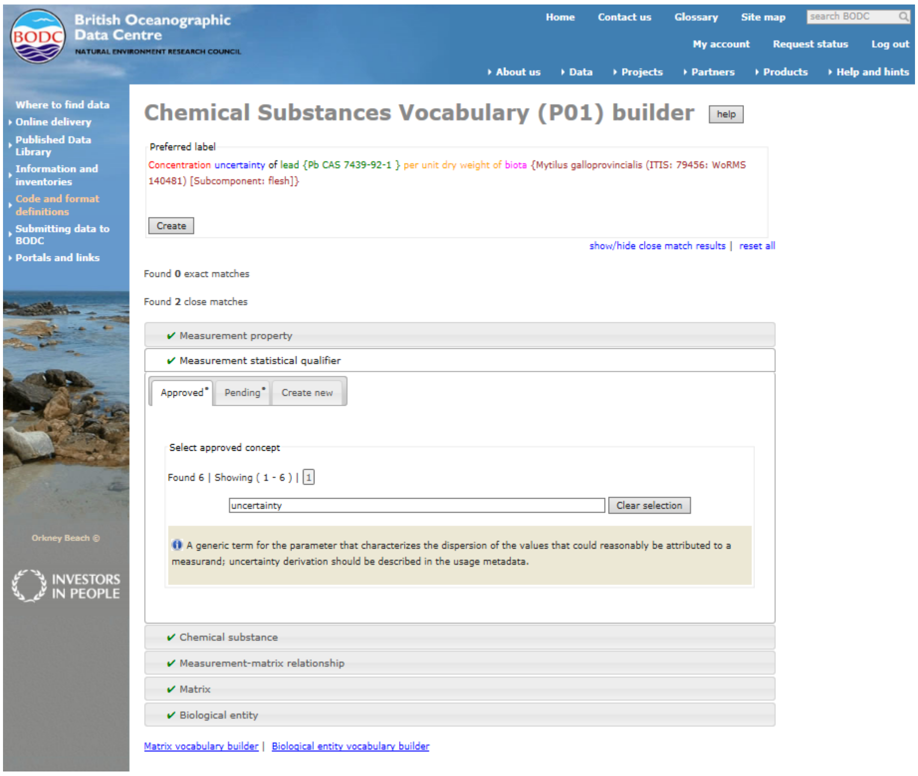


Figure 3.1.5: NVS vocabulary builder for chemical substances within the P01 vocabulary.

Ongoing/future work for vocabularies

BODC has planned the following activities:

a. Shut down of NVS1.0 to allow enriched predicate set for mappings. This will then facilitate implement unit conversions through use of richer predicates.
b. Overlaying SKOS with OWL for enhanced semantic reasoning.
c. Providing multilingual support for key vocabularies.
d. Providing provenance of mappings.
e. Expand the Vocabulary builder to include P01 terms built from the biological and general purpose models.

3.2 Australian developments by IMOS

IMOS is now using the Australian National Data Service (ANDS) Tool Suite in production mode. This toolset is comprised of 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 can be found at:
https://vocabs.ands.org.au/#!/?p=1&subjects=Marine
and re-use BODC vocabularies where-ever possible. Through participation in ODIP, IMOS has been able to expand the number of Australian organisations registered in EDMO (http://seadatanet.maris2.nl/v_edmo/browse_step.asp?step=00054), as well as add terms to 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), or by issuing a request 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. The use of these controlled AODN vocabularies, in MCP metadata, performs a similar function to that of the SeadataNet Common Vocabularies in the SeadataNet Common Data Index. They are integral for standardised indexing and searching within the IMOS Data Portal, which 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 used 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 in to 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 non-current vocabularies could mean that providers are not aware 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 immature (as 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 in place a robust migration path/recommended process for updating used vocabularies in metadata.

Thinking that AODN had covered the hooks necessary for managing the use of versioned vocabularies in metadata submitted for publishing AODN datasets via the IMOS Portal, AODN moved on to flow-chart this process. It was at this point a number of related problematic issues became evident. 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:

a) 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?

b) 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)?

c) Should existing metadata, already indexed in the Portal Infrastructure be treated differently to metadata yet to be ingested (i.e., should new metadata be validated for compliance by insisting on the use of only current vocabularies, whilst accepting already indexed metadata as was submitted)?

d) If already indexed metadata should be updated to conform to using 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?

e) 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?

f) 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?

g) 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?

h) Instead of rejecting records because they don't 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 (in particular Roy Lowry) shed light 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 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 mentioned at (g).

SeaDataNet also intends to centrally update provided metadata to reflect changes in vocabulary versions, but we are not aware of how the provenance information regarding such changes will be handled. It is assumed that this central updating cause a difference between provided records that are indexed and those held by the provider locally. How do providers feel about this situation?

In summary, AODN is leaning towards the approach mentioned at (h) where an 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 an individualized 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 be re-indexed anyway, 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. But, these terminologies are often integral to how their internal (or public-facing) systems operate and they are loathe to replace them with an alternative. 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.

So, IMOS has 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. To characterise these issues refer to the following terms in L05, which are all at the same level in the vocabulary.

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 chosen above in L05 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.

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 than it perhaps otherwise needed to be because the BODC vocabulary (re-used in the AODN Instrument Vocabulary) lacked 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'. Perhaps they should be matched to '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 doesn't negate a further finessing of the root term into additional classes of thing based on more subjective, or fine-grained criteria. But the root term should be sufficiently generic that it can be easily mapped to, should the splicing and dicing criteria for terms narrower than it differ between various vocabularies, which is a frequent occurrence.

Since most of AODNs dealings with BODC have been in relation to adding missing terms, AODN is unsure how amenable BODC is to recasting terms in L05 where such an edit may have impacts on internal (BODC) vocabulary relationships?

Appendix A: 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 B - 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/.

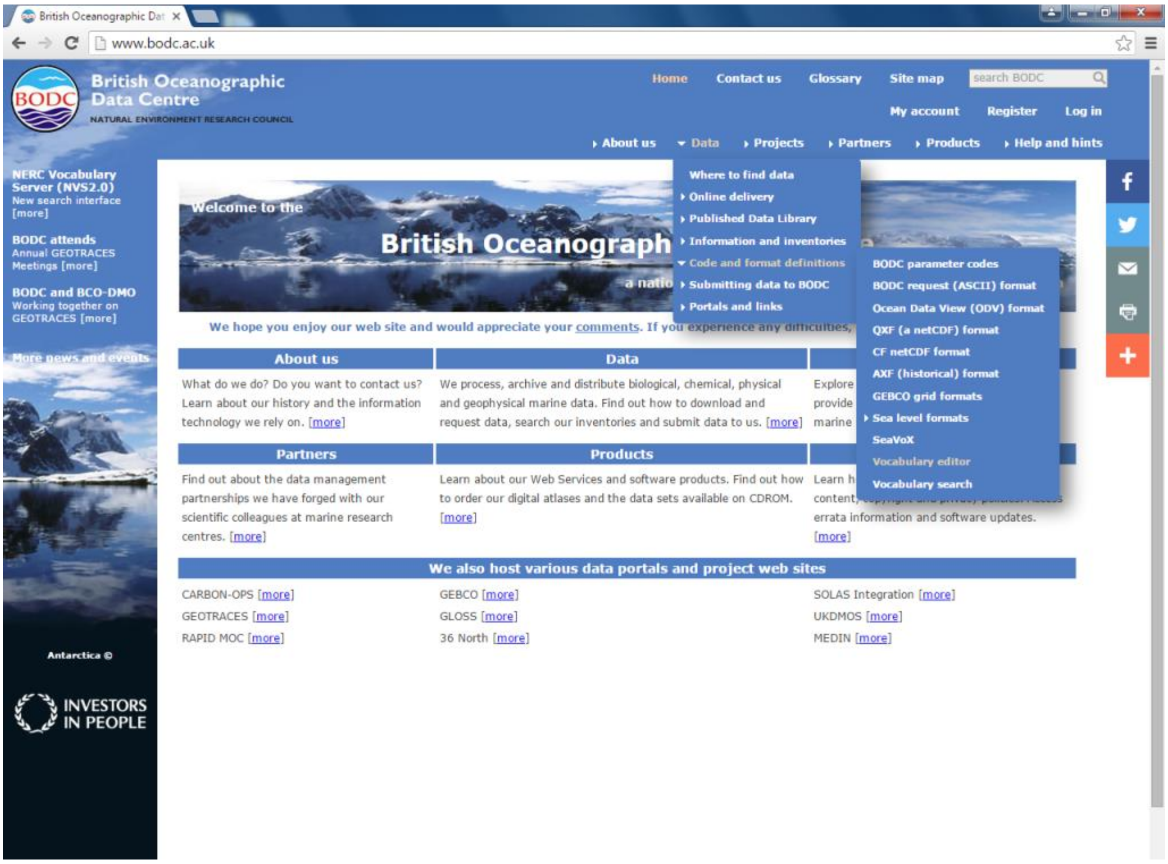


Figure B.1: Accessing the NVS Editor from the BODC homepage.

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.

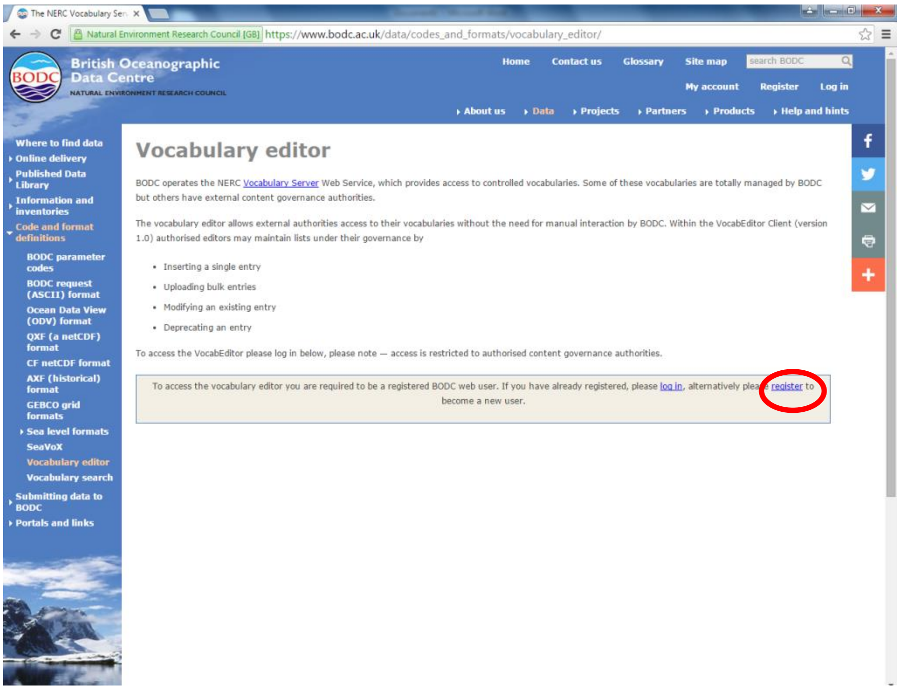


Figure B.2: The editor requires a user to be registered before use.

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.

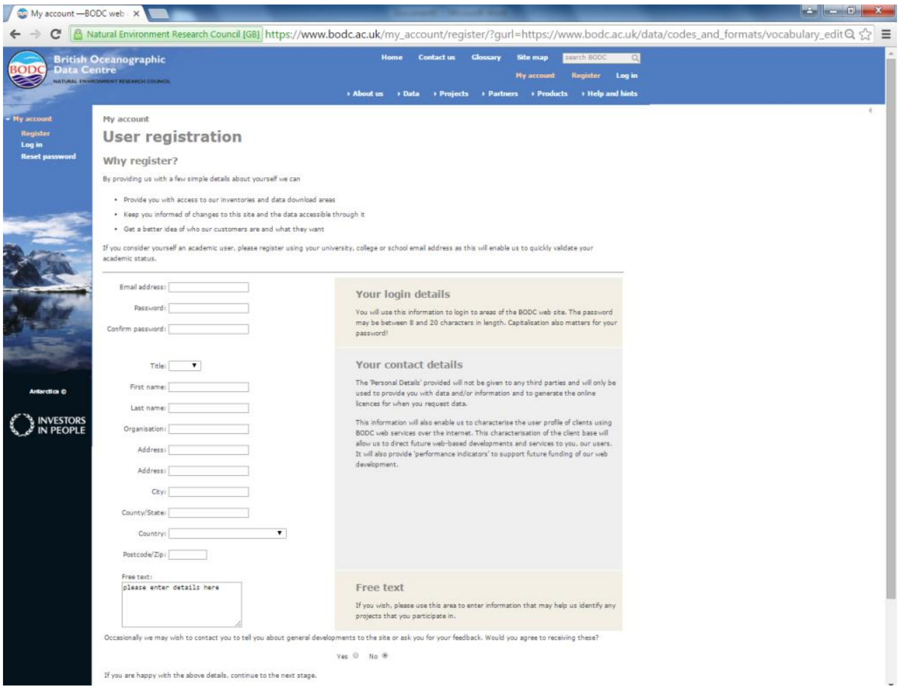


Figure B.3: BODC website user registration page.

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

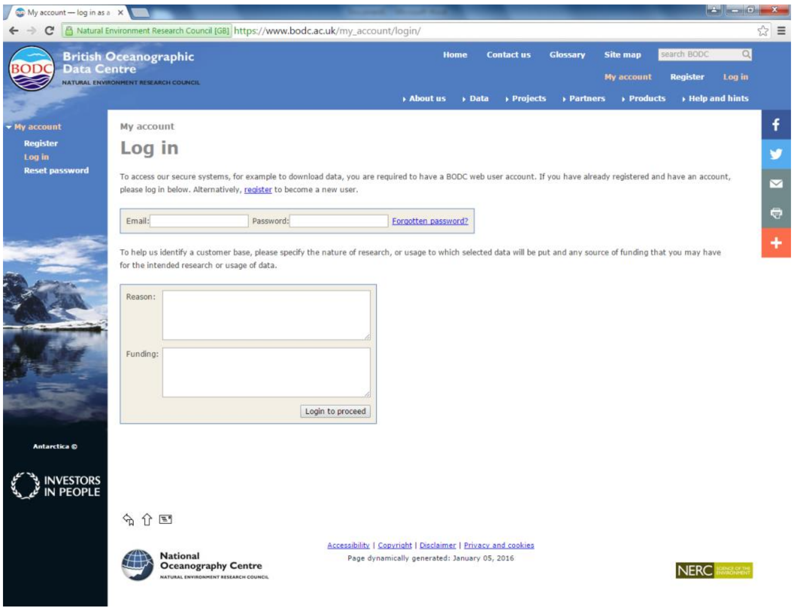


Figure B.4: BODC website login page.

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.

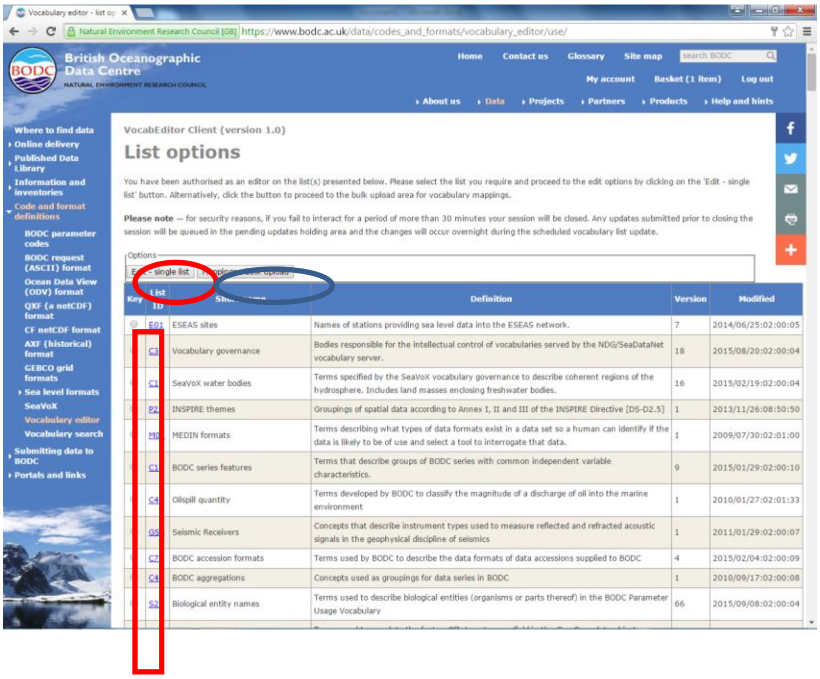


Figure B.5: NVS Editor page showing vocabularies available for editing by the user.

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.

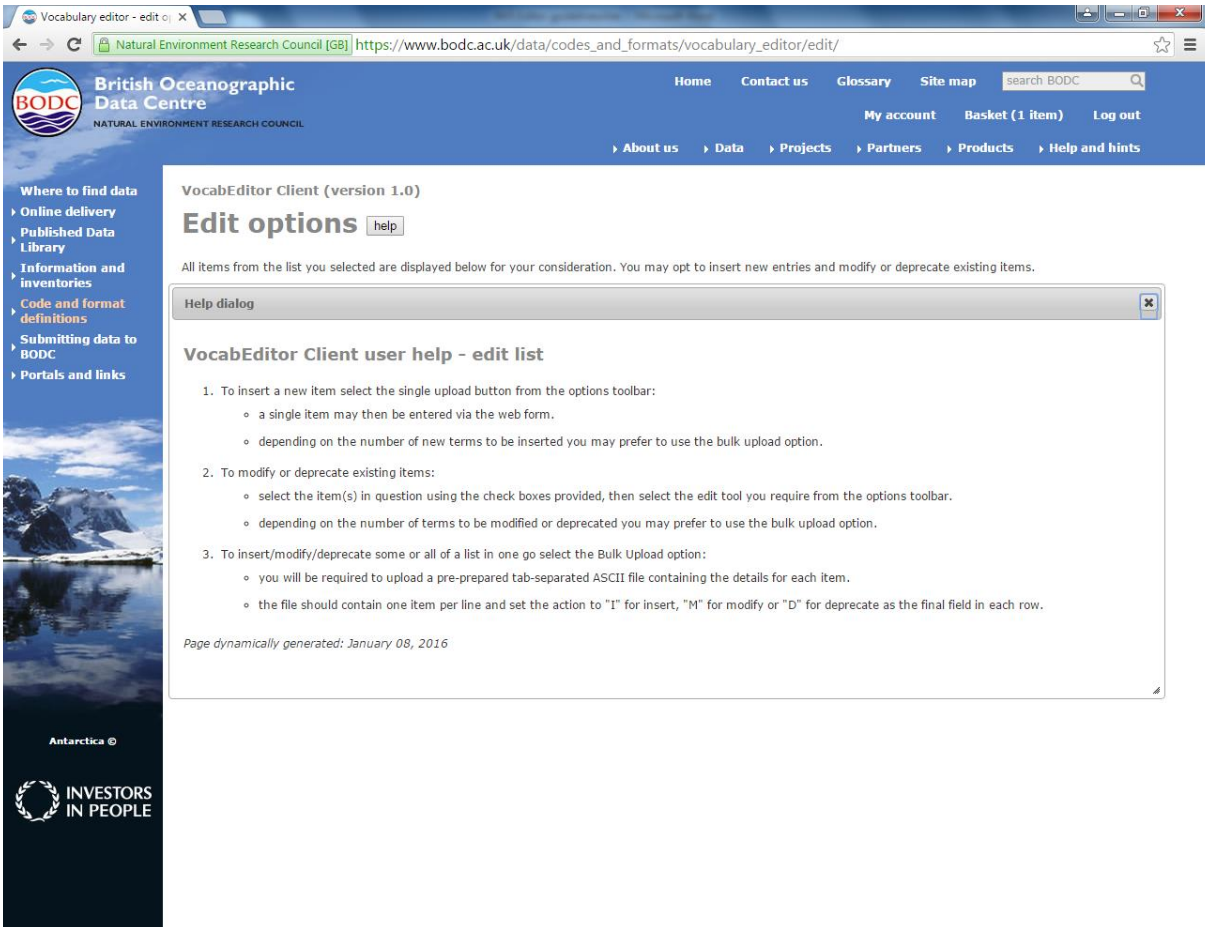
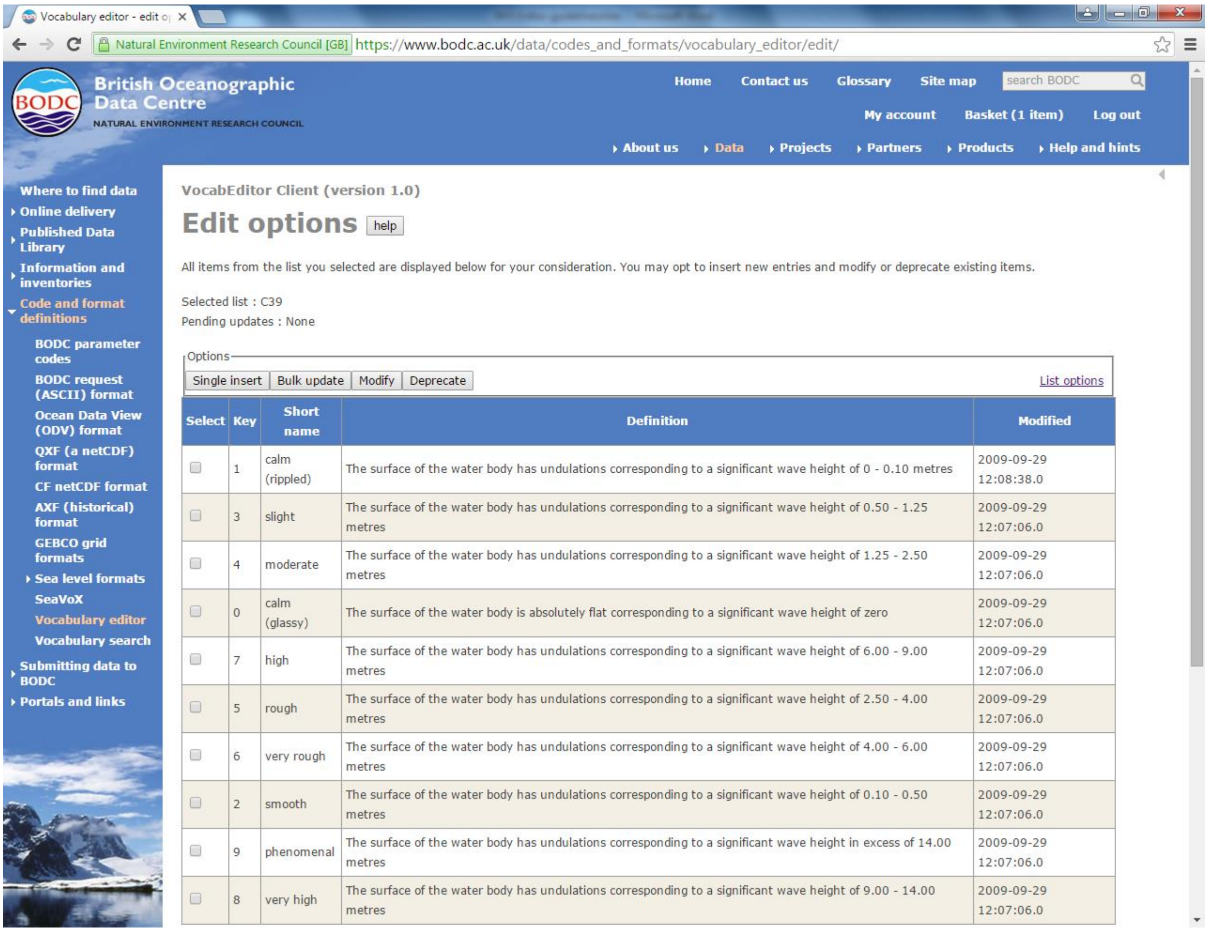


Figure B.6: NVS Vocab Editor user help for making edits to a vocabulary.

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).



Select	Key	Short name	Definition	Modified
	1	calm (rippled)	The surface of the water body has undulations corresponding to a significant wave height of 0 - 0.10 metres	2009-09-29 12:08:38.0
	3	slight	The surface of the water body has undulations corresponding to a significant wave height of 0.50 - 1.25 metres	2009-09-29 12:07:06.0
	4	moderate	The surface of the water body has undulations corresponding to a significant wave height of 1.25 - 2.50 metres	2009-09-29 12:07:06.0
	0	calm (glassy)	The surface of the water body is absolutely flat corresponding to a significant wave height of zero	2009-09-29 12:07:06.0
	7	high	The surface of the water body has undulations corresponding to a significant wave height of 6.00 - 9.00 metres	2009-09-29 12:07:06.0
	5	rough	The surface of the water body has undulations corresponding to a significant wave height of 2.50 - 4.00 metres	2009-09-29 12:07:06.0
	6	very rough	The surface of the water body has undulations corresponding to a significant wave height of 4.00 - 6.00 metres	2009-09-29 12:07:06.0
	2	smooth	The surface of the water body has undulations corresponding to a significant wave height of 0.10 - 0.50 metres	2009-09-29 12:07:06.0
	9	phenomenal	The surface of the water body has undulations corresponding to a significant wave height in excess of 14.00 metres	2009-09-29 12:07:06.0
	8	very high	The surface of the water body has undulations corresponding to a significant wave height of 9.00 - 14.00 metres	2009-09-29 12:07:06.0

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

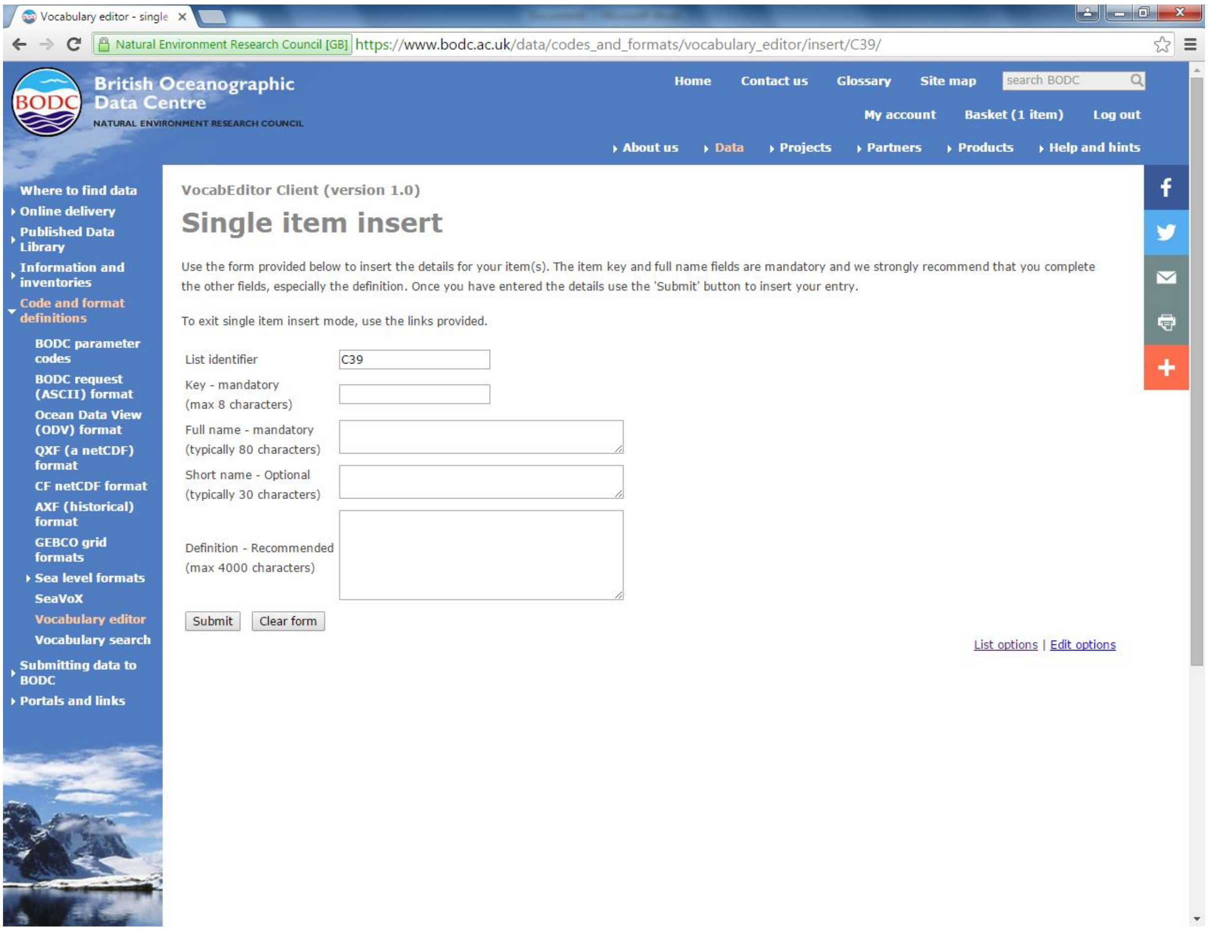


Figure B.8: How to insert a single item.

Bulk update

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.

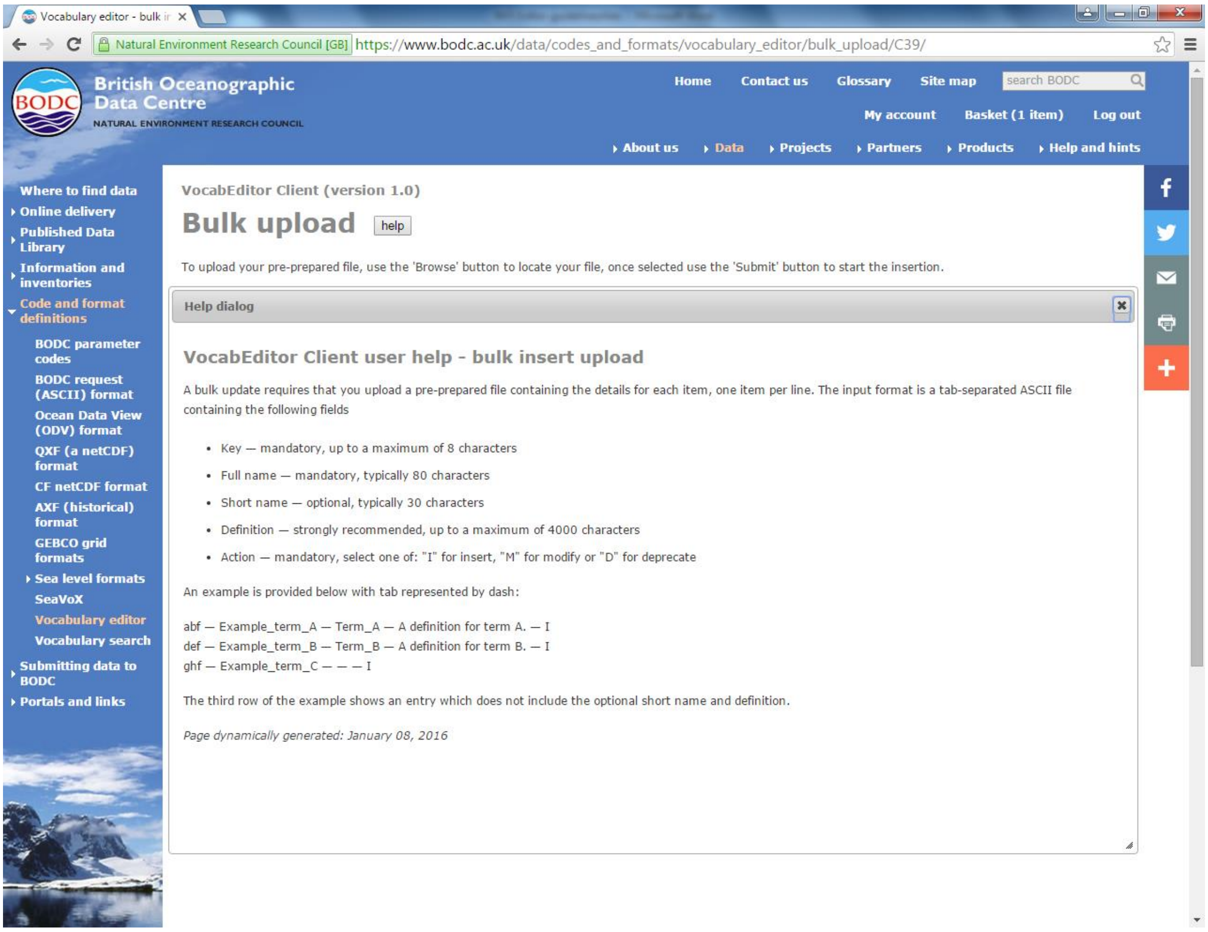


Figure B.9: NVS Vocab Editor user help detailing how to set up a file for the bulk update of a vocabulary.

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.

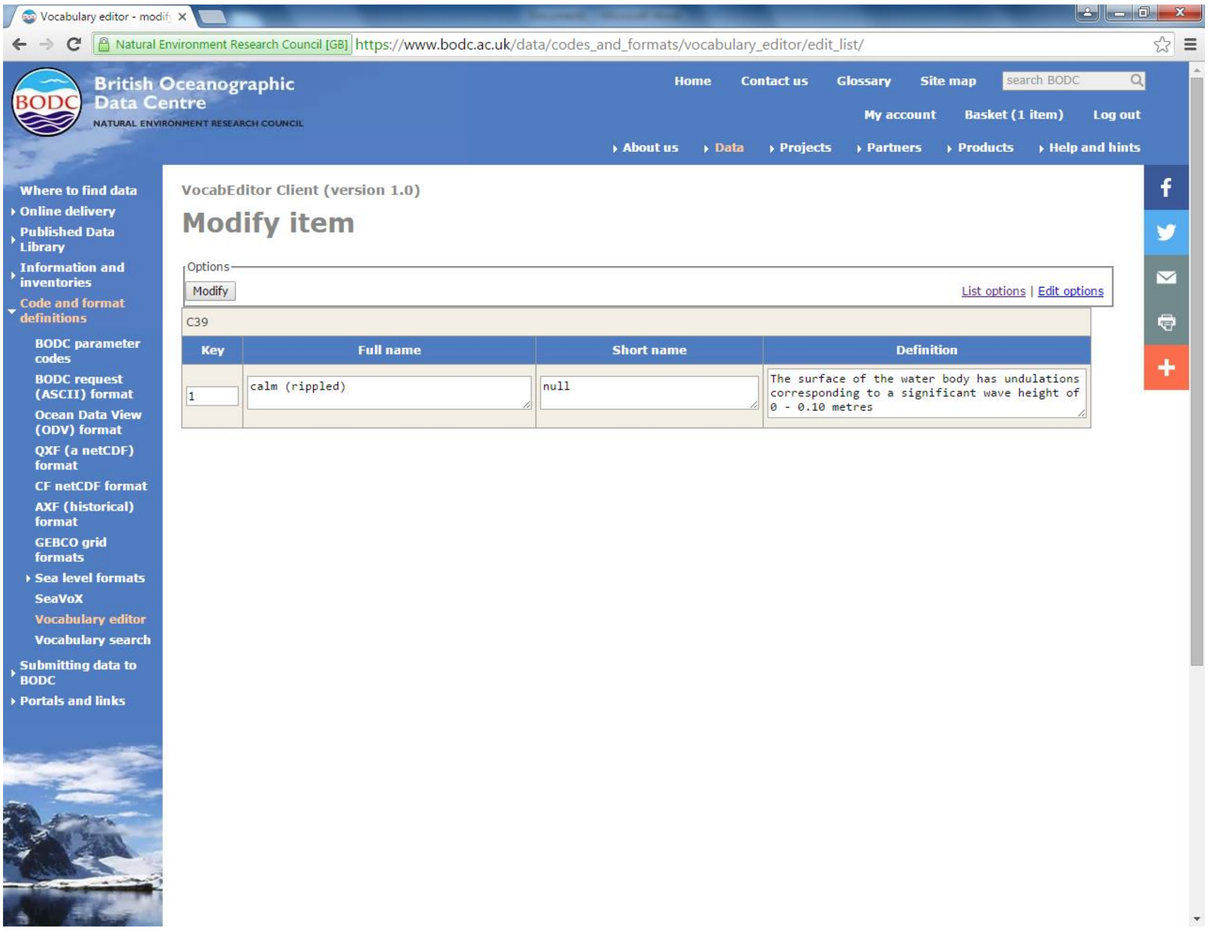


Figure B.10: NVS Vocab Editor web form for modifying an existing item.

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.

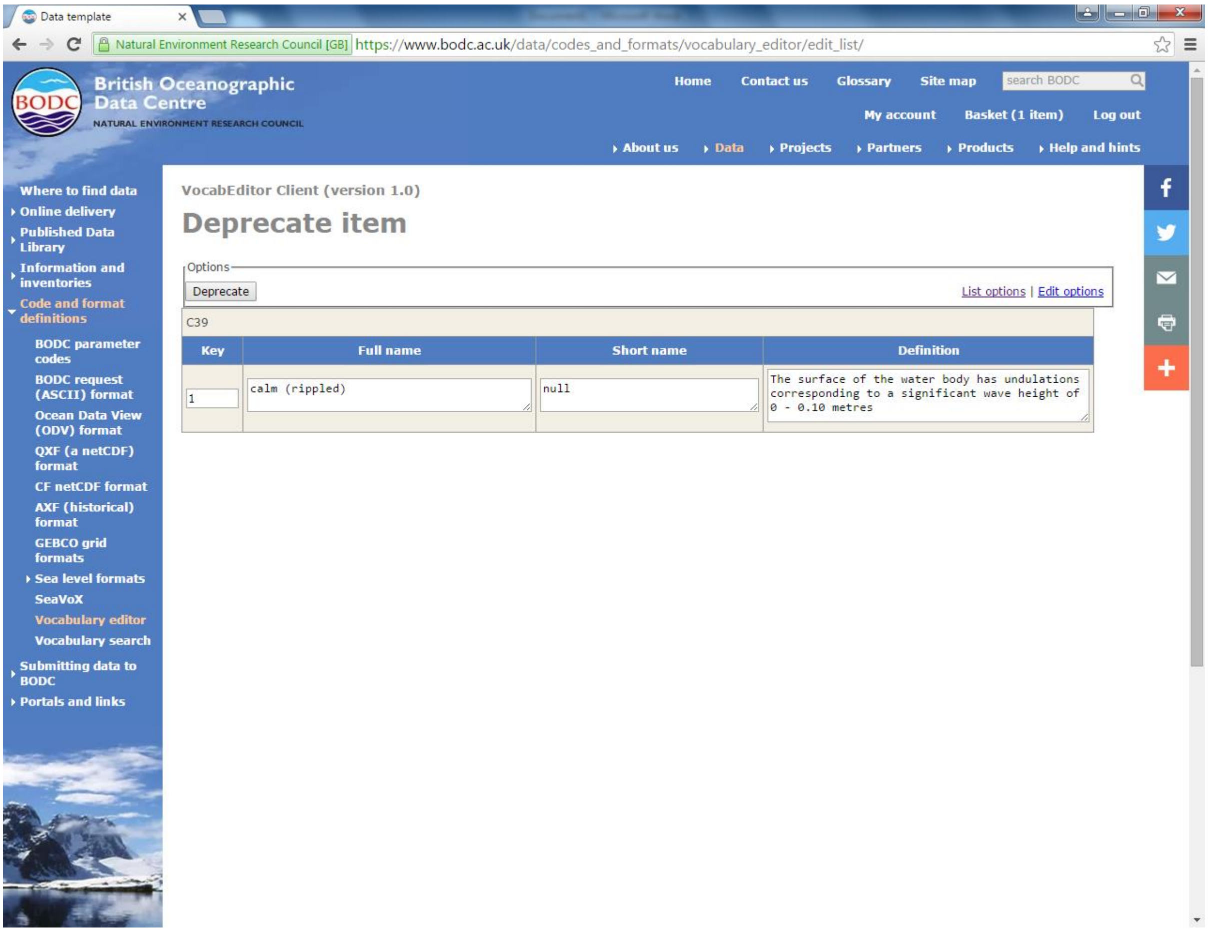


Figure B.11: NVS Vocab Editor web form for deprecating an existing item.

Mappings

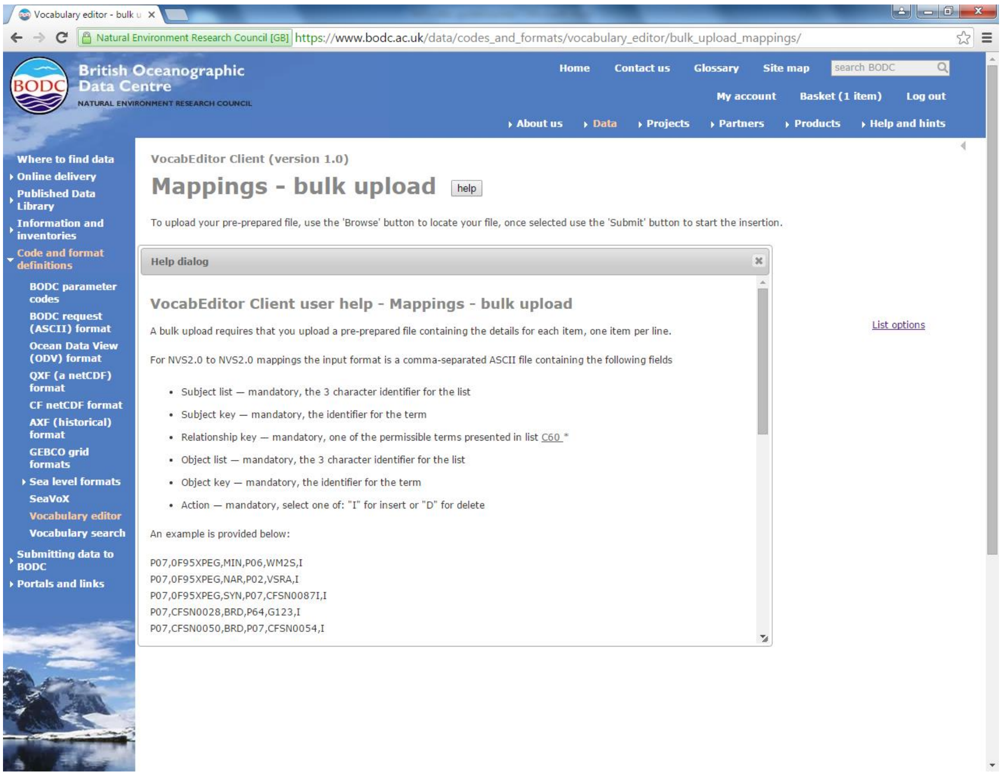


Figure B.12: An example layout for inserting mappings between two NVS vocabulary lists.

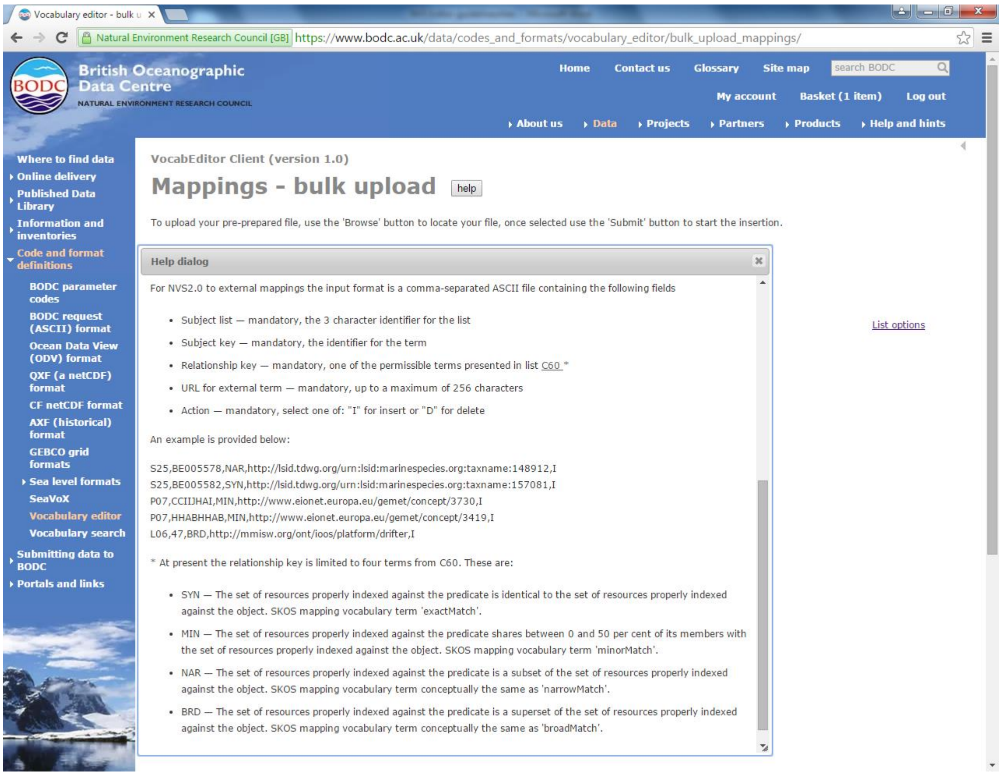


Figure B.13: The following screen print provides the example layout for inserting mappings between an NVS vocabulary list and a non-NVS served concept URL.

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 C – Comparative table of XML editors as prepared by CNR

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)	?
Programmin g 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, ...