

Managing and sharing common vocabularies through NVS services

...OPTION FOR MANAGING ARGO VOCABS?

Presented by Robin McCandliss - using slides from Gwen Moncoiffe (Leader of NVS service at BODC)



**National
Oceanography Centre**
NATURAL ENVIRONMENT RESEARCH COUNCIL



**British Oceanographic
Data Centre**
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE
ENVIRONMENT

Why use vocabularies?

Divided by a common language



What is the NERC Vocabulary Server (NVS)?

Essentially a server that provides persistent URIs for:

- vocabulary term (often referred to as “concept”)
- vocabulary lists (collections of related concepts)
- vocabulary thesaurus (aggregation of concepts from one or more collections)



What is the NERC Vocabulary Server (NVS)?

By extension, it also refers to a set of web services and applications designed to manage and share vocabularies served by the NERC Vocab Server;

They can be shared as a **trusted resource** because:

- Each vocabulary is referenced as a **unique resource identifier (URI)**;
e.g. <http://vocab.nerc.ac.uk/collection/L22/current/TOOL0668/>
- The identifier provides access to **machine and human readable** information;
- The information is structured and managed according to **W3C standards**;
- The NVS supports **mapping to internally and externally hosted vocabularies**;
- Once published the **URIs are permanent**;
- They can be **deprecated but never removed or redefined**;

Structure based on SKOS standard

Simple Knowledge Organisation System (SKOS)

A concept has:

- Identifier e.g. “31”
- Preferred Label e.g. “research vessel”
- Alternative Label -
- Definition e.g. “A platform of any size operating on the surface of the water column in unpredictable locations that is specifically equipped, manned and operated for scientific, usually oceanographic, research.”

<http://vocab.nerc.ac.uk/collection/L06/current/31/>

An example from the SeaVox platform categories vocabulary (L06)



An example from the SeaVox Device Catalogue (L22)

↑ -- Sea-Bird SBE 41 CTD --

URI <http://vocab.nerc.ac.uk/collection/L22/current/TOOL0668/>

Identifier () SDN:L22::TOOL0668

Preferred label (en) **Sea-Bird SBE 41 CTD**

Alternative label (en) SBE 41 CTD

Definition (en) A self-contained unit comprising the MicroCAT temperature, conductivity and pressure sensors and a pump that is designed specifically for deployment on profiling floats, particularly Argo. The unit is designed to provide stable salinity data accurate to 0.005 PSU for periods in excess of three years without any form of maintenance. Temperature is within 0.002 C (stability 0.0002 C/year) and pressure within 2 dbar (stability 0.8 dbar/year). During float ascent spot samples are taken and transmitted to the float controller. More information is given in http://www.seabird.com/products/spec_sheets/41data.htm.

Version Info () 1

Deprecated() false

Broader <http://vocab.nerc.ac.uk/collection/L05/current/134/>

Is broadly related with “Temperature sensor”

Broader <http://vocab.nerc.ac.uk/collection/L05/current/130/>

Is broadly related with “CTD”

Broader <http://vocab.nerc.ac.uk/collection/L05/current/350/>

Is broadly related with “Salinity sensor”

Date () 2014-03-11 15:55:41.0

Management tools

Externally governed vocabularies are increasingly managed via the NVS VocabEditor – Allow external editors to submit and edit new terms and mappings to vocabularies they are authorised to access (access control)

https://www.bodc.ac.uk/resources/vocabularies/vocabulary_editor/

VocabEditor Client (version 1.0)

List options

You have been authorised as an editor on the list(s) presented below. Please proceed with one of the following options:

- **Mappings – Bulk Upload:** Click Mappings – bulk upload button to upload mappings
- **Edit - single List:** Select a list to edit
 1. Select the list you require
 2. Click Edit single list
 3. Select one of the options

Please note – for security reasons, if you fail to interact for a period of more than 30 minutes your session will be closed. Any updates submitted prior to closing the session will be queued in the pending updates holding area and the changes will occur overnight during the scheduled vocabulary list update.

Options

Key	List ID	Short name	Definition	Version	Modified
<input type="radio"/>	M09	MEDIN EV method types	Terms used to classify techniques used to assess the socio-economic value of an ecosystem.	1	2013/11/15:02:00:10
<input type="radio"/>	M10	MEDIN EV methods	Terms used to describe techniques used to assess the socio-economic value of an ecosystem.	1	2013/11/15:02:00:10
<input type="radio"/>			Terms used to classify the conditions and processes through which natural ecosystems sustain		

Content and technical governance

- **Collections and governance**
- 231 vocabulary collections accessible through the NVS
- 70 owned and governed by BODC
- 48 are managed by BODC on behalf of SeaDataCloud, EMODnet, SeaVox, SWE content governance authorities
- Remainder (113) are owned by 25 different governing authorities
- **Technical governance and support**
- assured by BODC
- On-going work improving access to governance and version information, improving user interfaces (part of the European SeaDataCloud project)

Collections library

A01	A02	A03	A04	A05	B02	B03	B04	B05	B06	B07	B09	B11	B12	B20
B21	B22	B39	B75	B76	C00	C10	C16	C17	C18	C19	C30	C31	C32	C33
C34	C35	C36	C37	C38	C39	C40	C41	C43	C45	C46	C47	C48	C60	C61
C62	C64	C67	C71	C72	C75	C77	C86	C87	C88	C89	C96	C97	C98	
E01	E02	EVO	F02	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11
G12	G13	G14	G15	G17	G18	G20	G21	G22	G23	G25	G26	G28	G29	GBX
GGS	GS1	GS2	GS3	GS4	GS5	GS6	GS8	GS9	GSA	GSB	GSC	GXM	H01	H02
H03	H04	H05	HA2	I01	I02	I03	I10	I11	I12	I13	I14	I15	L02	L03
L04	L05	L06	L07	L08	L10	L11	L12	L13	L14	L15	L18	L19	L20	L21
L22	L23	L24	L26	L27	L30	L31	L33	L34	L35	M01	M03	M04	M05	M06
M09	M10	M11	M12	M13	M14	M15	M16	M17	M18	N01	N02	N03	N04	N05
P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15
P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P35
P36	P37	P38	P64	Q01	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10
S11	S12	S13	S14	S15	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30
V12	V22	V23	W01	W02	W03	W04	W05	W06	W07	W08	W09	W10		



Steps to hosting vocabularies on the NVS

Agree a governance model e.g.

- an ARGO vocabulary governance committee?
- one or multiple authorised vocabulary managers including a member of the BODC Vocab Management team

Identify collections and make them SKOS compliant by providing:

- Title e.g. Climate and Forecast Standard Names
- Short name e.g. CF Standard Names
- Definition e.g. Terms used for definitive but not necessarily comprehensive descriptions of measured phenomena in the CF conventions (a content standard for data stored in NetCDF)

Ensure concepts are unambiguous by providing:

- A preferred label
- An optional alternative label
- A definition description of the term

See for example: <http://vocab.nerc.ac.uk/collection/W03/current/>



Steps to hosting vocabularies on the NVS

- BODC Vocabs Team provides infrastructure and expert advice
- Community develops the vocabularies (terms, definitions etc.)
- Community to identify Argo vocabs for management
- Community to ensure Argo vocabs are clean, well-defined and SKOS compliant
- Could pilot taking an example of a good, well-defined Argo vocab into NVS?
- Get Argo vocabs into good shape beforehand → reduces effort to get them into NVS
- Note: If significant BODC effort required to get Argo vocabs into NVS, then this would need funding
- Will require time and effort over the longer term

