

Vocabularies: the Semantic Glue for Interoperability

POLAR DATA FORUM II
LISE QUESNEL
ROY LOWRY



**National
Oceanography Centre**
NATURAL ENVIRONMENT RESEARCH COUNCIL



noc.ac.uk

Sharing Data

Why are Vocabularies important for sharing data?

What is a Vocabulary?

Introducing NVS2.0 and how it is used

Exchange Standards allow Interoperability

Data in **standard** formats (e.g. CF-compliant NetCDF)

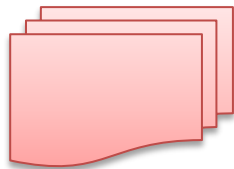
Metadata with **standard** content (e.g. ISO19115)

Content structured for **machines** (e.g. ISO19139 XML)

Structure with standard **web resource** elements (e.g. RDF)



Focus on Content



= ?



Plain text Metadata

Why are Vocabularies Important?

To capture **expertise** in agreed, well-defined **descriptions**

To enable **population** of a given field in a **metadata model**

Why are Vocabularies Important?

To capture **expertise** in agreed, well-defined **descriptions**

To enable **population** of a given field in a **metadata model**

With this Peer-reviewed expertise i.e.



Semantic 'glue' of Interoperability!

Why are Vocabularies Important?

To capture **expertise** in agreed, well-defined **descriptions**

To enable **population** of a given field in a **metadata model**

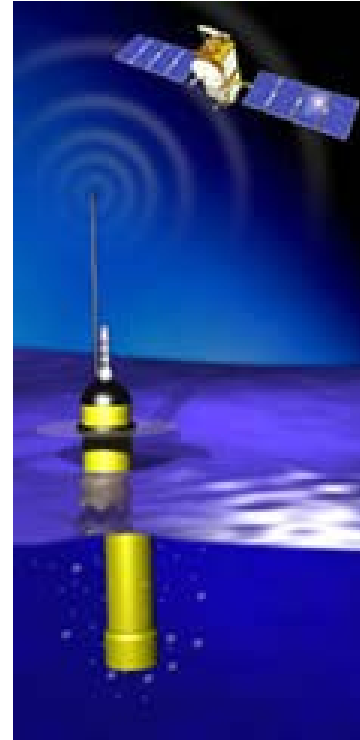
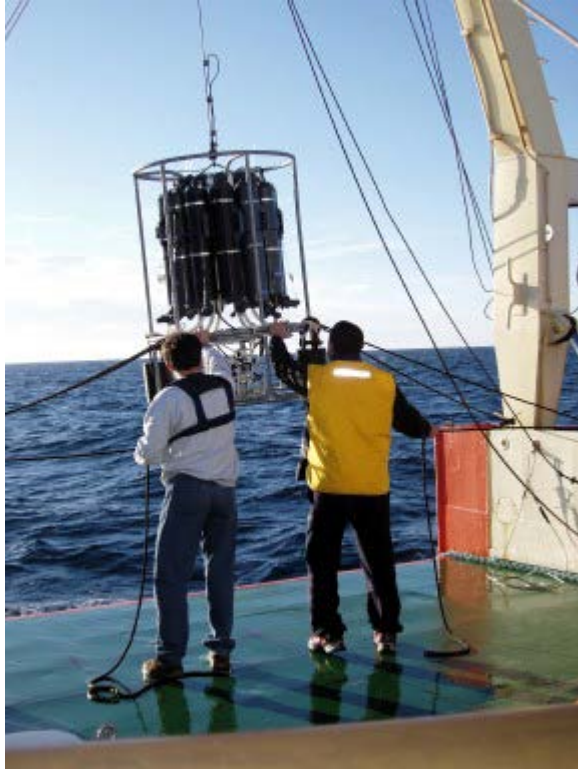
With this Peer-reviewed expertise i.e.



Semantic ‘glue’ of Interoperability!

- ✓ *Data centres **automate** Discovery & Sharing*
- ✓ *Cross-domain experts **map** one vocabulary to another*
- ✓ *Embedded web structures bring **new knowledge***

What is a Vocabulary?



What are the measured phenomena?

unique term = concept

BCMT Metal concentrations in biota

PT0X Concentration of phycotoxins in biota

BCPB Concentration of polychlorobiphenyls (PCBs) in biota

unique term = concept

aggregation = vocabulary

BCMT Metal concentrations in biota

PT0X Concentration of phycotoxins in biota

BCPB Concentration of polychlorobiphenyls (PCBs) in biota

Parameter Discovery (P02) has 449 concepts
Parameter Usage (P01) has 35025!

unique term = concept

aggregation = vocabulary

- BCMT Metal concentrations in biota
- PT0X Concentration of phycotoxins in biota
- BCPB Concentration of polychlorobiphenyls (PCBs) in biota

90 Manual biota samplers

383 Metal analysers

concept

vocabulary

Relationships = Thesaurus / Ontology

concept

vocabulary

BCMT Metal concentrations in biota

PTOX Concentration of phycotoxins in biota

BCPB Concentration of polychlorobiphenyls (PCBs) in biota

90 Manual biota samplers

383 Metal analysers

vocabulary

Concepts, Vocabularies, Thesauri

<http://vocab.nerc.ac.uk/collection/P02/current/BCMT/>

<http://vocab.nerc.ac.uk/collection/P02/current/PTOX/>

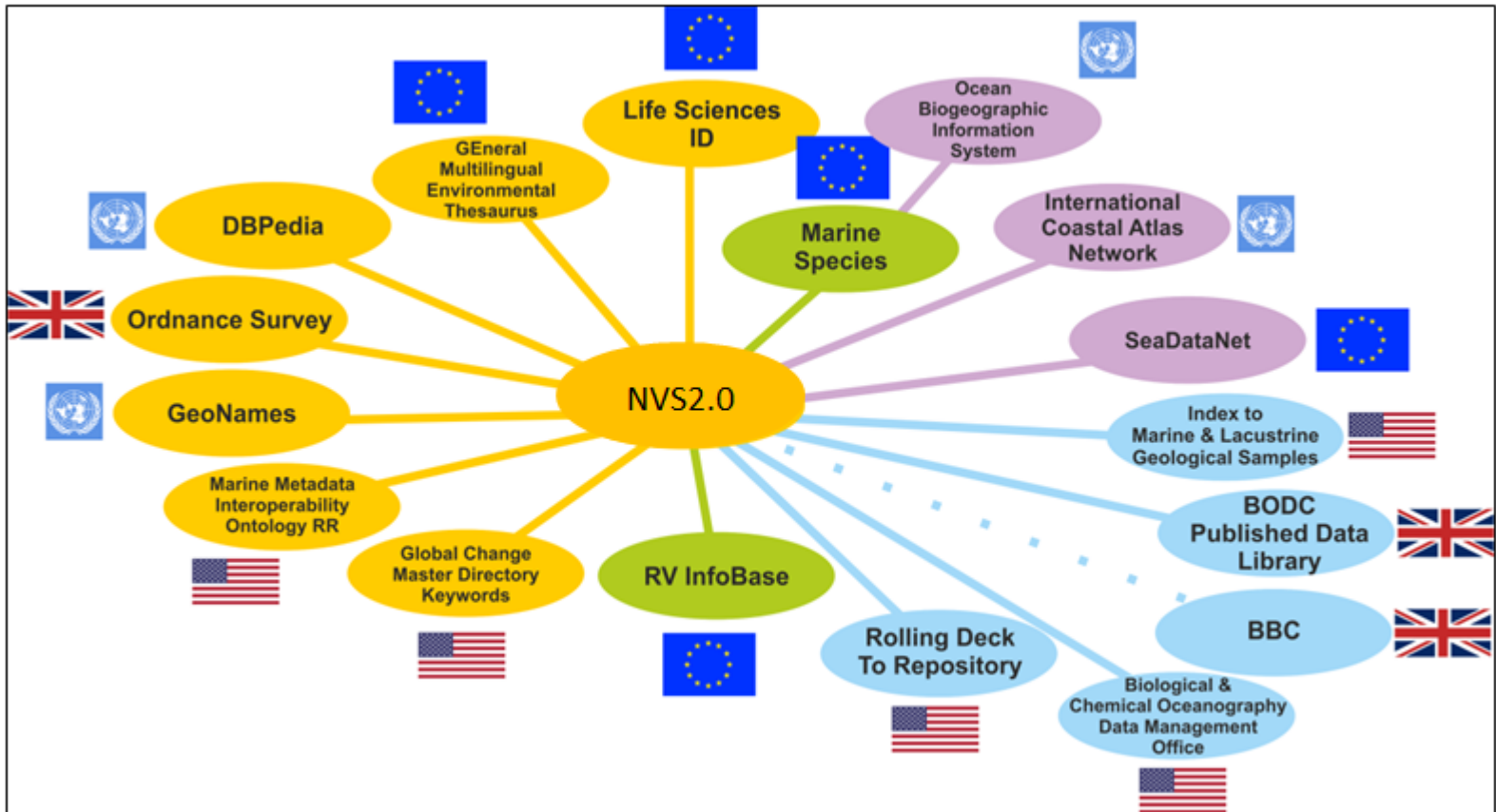
<http://vocab.nerc.ac.uk/collection/P02/current/BCPB/>

<http://vocab.nerc.ac.uk/collection/L05/current/90/>

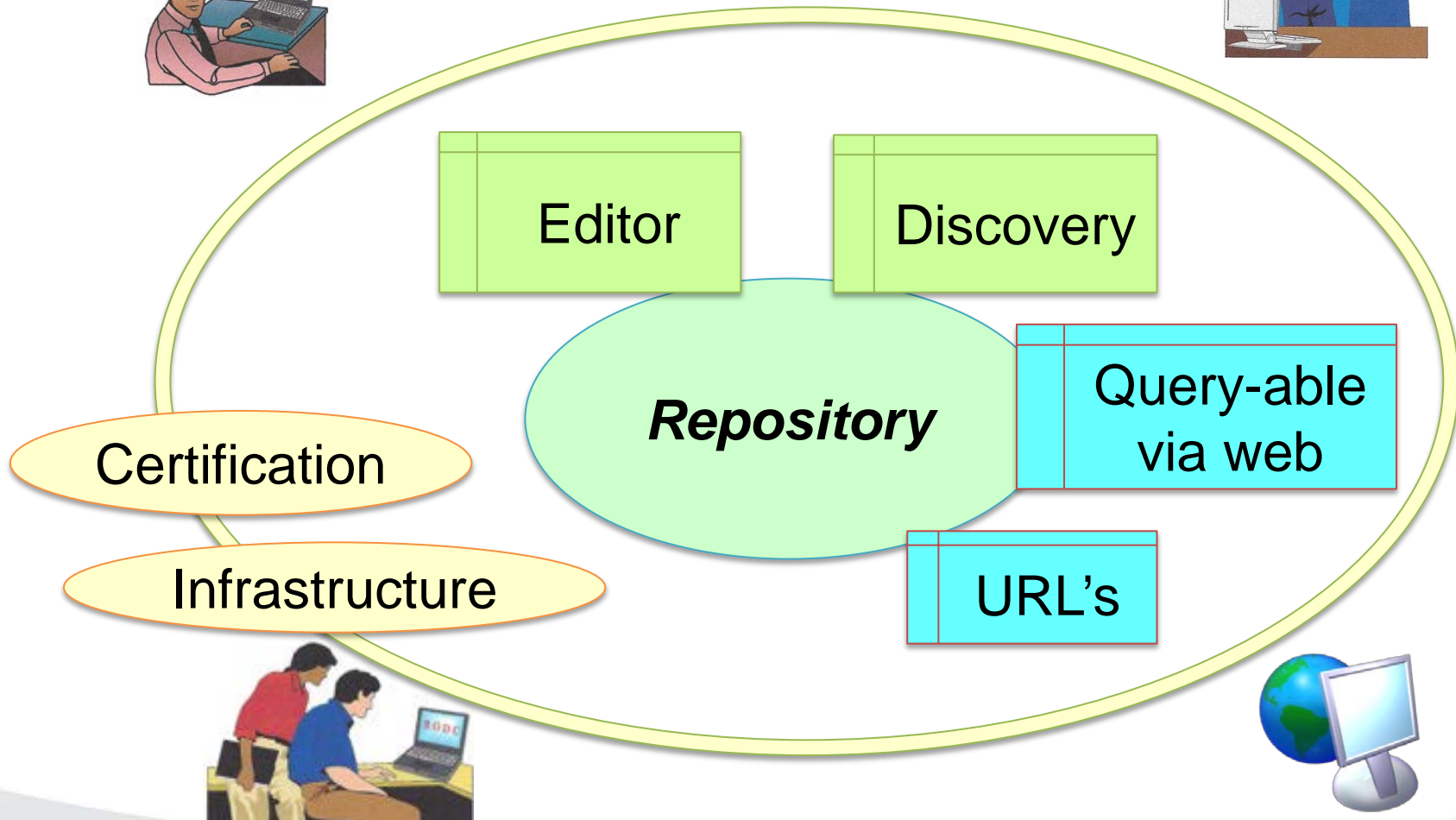
<http://vocab.nerc.ac.uk/collection/L05/current/383/>

http://vocab.nerc.ac.uk/scheme/EMODNET_PEST/current/

NVS2.0



NVS2.0 at BODC



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

Simple search for vocabularies

Advanced search for vocabularies

Simple search within a vocabulary

Advanced search within a vocabulary

Click search button with empty search string to get list of all available concepts

Search text Collection ID

Collections library

A01	A02	A03	A04	B02	B03	B04	B05	B06	B07	B09	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	EVO	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	GG5	GS3	GS4	GS5	GS6	GS8	GS9	GSA	GSB	GSC	GXM	I01
I02	I03	I10	I11	I12	I13	I14	I15	L02	L03	L05	L06	L07	L08	L10
L11	L12	L13	L14	L15	L18	L19	L20	L21	L22	L23	L24	L26	L27	L30
L31	L33	M01	M03	M04	M05	M06	M09	M10	M11	M12	M13	M14	M15	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	S01	S02	S03	S04	S06	S07	S08
S09	S10	S11	S12	S13	S14	S15	S21	S22	S23	S25	S26	S27	V12	V22
V23	W01	W02												

Vocabulary search

help

Parameter discovery (P02)

Search in P02 SeaDataNet Parameter Discovery Vocabulary

Search text

Vocabulary

[advanced options](#)

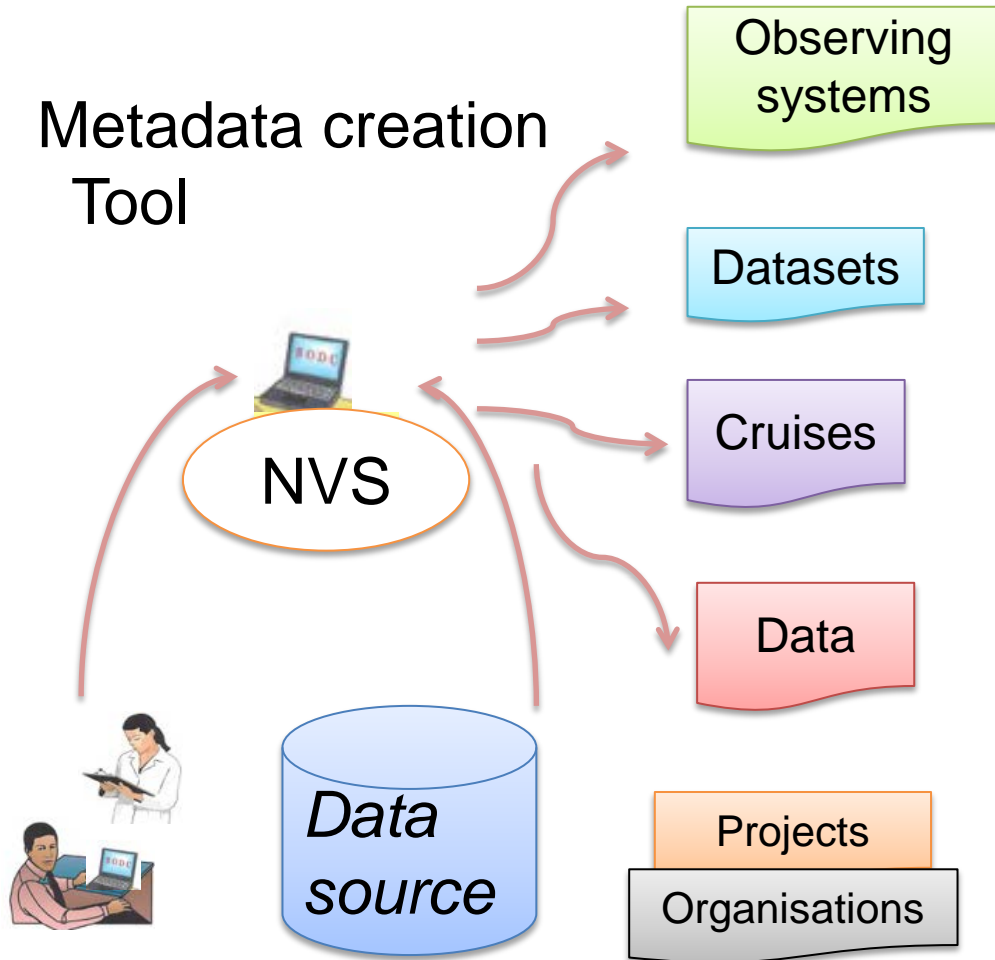
Found 102 records | Show (1 - 10) |

[download results](#) | [sta](#)

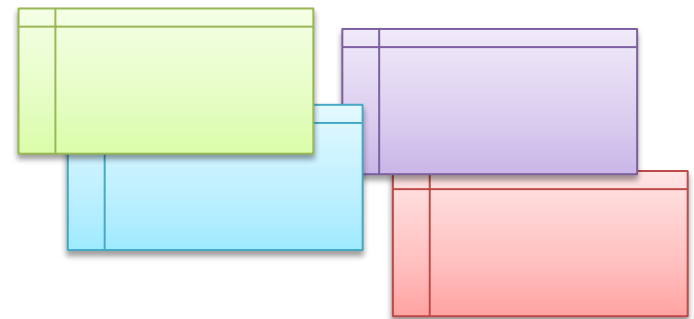
▲ Identifier ▼	▲ PrefLabel ▼	▲ Definition ▼	▲ D
TSED	Concentration of suspended particulate material in the water column	Parameters quantifying SPM concentration in the water column by weight, by volume or in terms of turbidity	200-17
CPWC	Chlorophyll pigment concentration in the water column	Includes all variants of chlorophyll expressed in terms of per unit volume or unit area of the water column. Does not include data expressed per unit mass of SPM.	200-30
PCHW	Concentration of polycyclic aromatic hydrocarbons (PAHs) in the water column	PAH concentrations in all phases of the water column. This group includes concentrations per unit volume of the water column in the particulate phase, but not concentrations per unit weight of SPM.	200-25
BCAH	Concentration of polycyclic aromatic hydrocarbons (PAHs) in biota	PAH concentrations in either whole organisms or their constituent parts	200-30
BCPB	Concentration of polychlorobiphenyls (PCBs) in biota	PCB concentrations in either whole organisms or their constituent parts	200-30
BCMT	Metal concentrations in biota	Metal concentrations in either whole organisms or their constituent parts.	200-12

NVS2.0 & EU SeaDataNet-2

Metadata creation
Tool



Metadata Discovery



Map from <http://www.seadatanet.org/Overview/Partners>

NVS2.0 & EU SeaDataNet-2

Identification Where When What How Who Where to find the data Cruise/Station Documentation Quality Others

Parameters

- RBHY - Molecular biology parameters
- MPMN - Moored instrument depth
- NTUP - New production in water bodies
- NTRA - Nitrate concentration parameters in the water column
- AMOX - Nitrification rate in the water column**
- NTRI - Nitrite concentration parameters in the water column
- NTSD - Nitrogen concentrations in sediment
- NTSP - Nitrogen concentrations in suspended particulate matter
- NTPW - Nutrient concentrations in sediment pore waters
- SAMO - Nutrient fluxes between the bed and the water column
- R410 - Ocean colour and earth-leaving visible waveband spectra
- OPBS - Optical backscatter
- CSDE - Organic sulphur compound dynamics in the water column
- OMBI - Organometallic species concentration parameters in water bodies
- OMPW - Organometallic species concentration parameters in sediment pore waters
- OMWC - Organometallic species concentration parameters in water column
- DMST - Organosulphur species concentration parameters in water bodies
- SICO - Other fluxes between the bed and the water column
- DCMX - Other halocarbon concentrations in water bodies
- HLDE - Other halocarbon dynamics in water bodies
- OMET - Other meteorological measurements

Name * Nitrification rate in the water column

Code P02 * AMOX

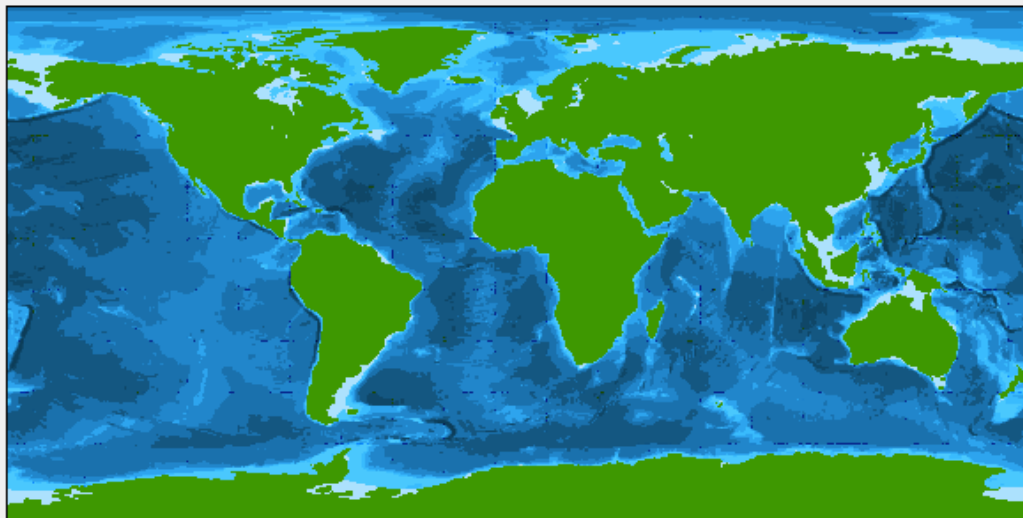
NVS2.0 & EU SeaDataNet-2

Sign In **SEADATANET COMMON DATA INDEX (CDI) V3**

Tools ?

Navigation and utility tools including zoom in/out, pan, and map controls. Includes buttons for Enlarge, Help, Position, and Index.

Datasets 0
Basket Reset



Layer control ?

Expand Add layer

- CDI entry Points ?
- CDI entry Tracks ?
- CDI entry Areas ?

- Grid Lines ?
- Regional sea ?
- Regional sea labels ?
- Main sea ?
- Main sea labels ?
- Bathymetry ?
- Blue Marble ?
- World ?

Lat/long ?

Upper-left ? Lower-right ?

Search

Search Clear ?

Free search

Disciplines - Parameter groups

- All
- Administration and dimensions
 - > Administration and dimensions
- Atmosphere
 - > Atmospheric chemistry
 - > Meteorology

Discovery parameters

- Suspended particulate material grain size parameters
- Temperature of geological units
- Temperature of the water column
- Temperature variation in the water column
- Terrestrial detritus in the water column suspended particulate material**
- Terrestrial mapping

Cruise/Station name

Instrument type

All

NVS2.0 & Ocean Data Interoperability Platform (ODIP)

Data management organisations from EU, US & Australia

- ✓ Fostering best practices & common standards
- ✓ Creating prototypes to enable transfer of technologies

NVS2.0

underpins linkage of research cruise programmes

included in new ontology to describe sensor metadata

Future Directions & Challenges for BODC

Future Directions

Enable Experts to semi-automate mappings
Increase our Linked Data footprint

Challenges facing all Organisations

Funding & Continuity Planning
Sustainability of Content & Technical Governance



*NVS supported via NERC funding
from its beginnings*