

The concept and benefit of common (FAIR) vocabularies

Gwenaëlle Moncoiffé

NERC Vocabulary Server Management Team

National Oceanography Centre - British Oceanographic Data Centre

Presented virtually at European Glider Data
Management Workshop, 21 June 2022



This project has received funding the European Union's Horizon 2020 research and innovation programme under grant agreement No 951842.

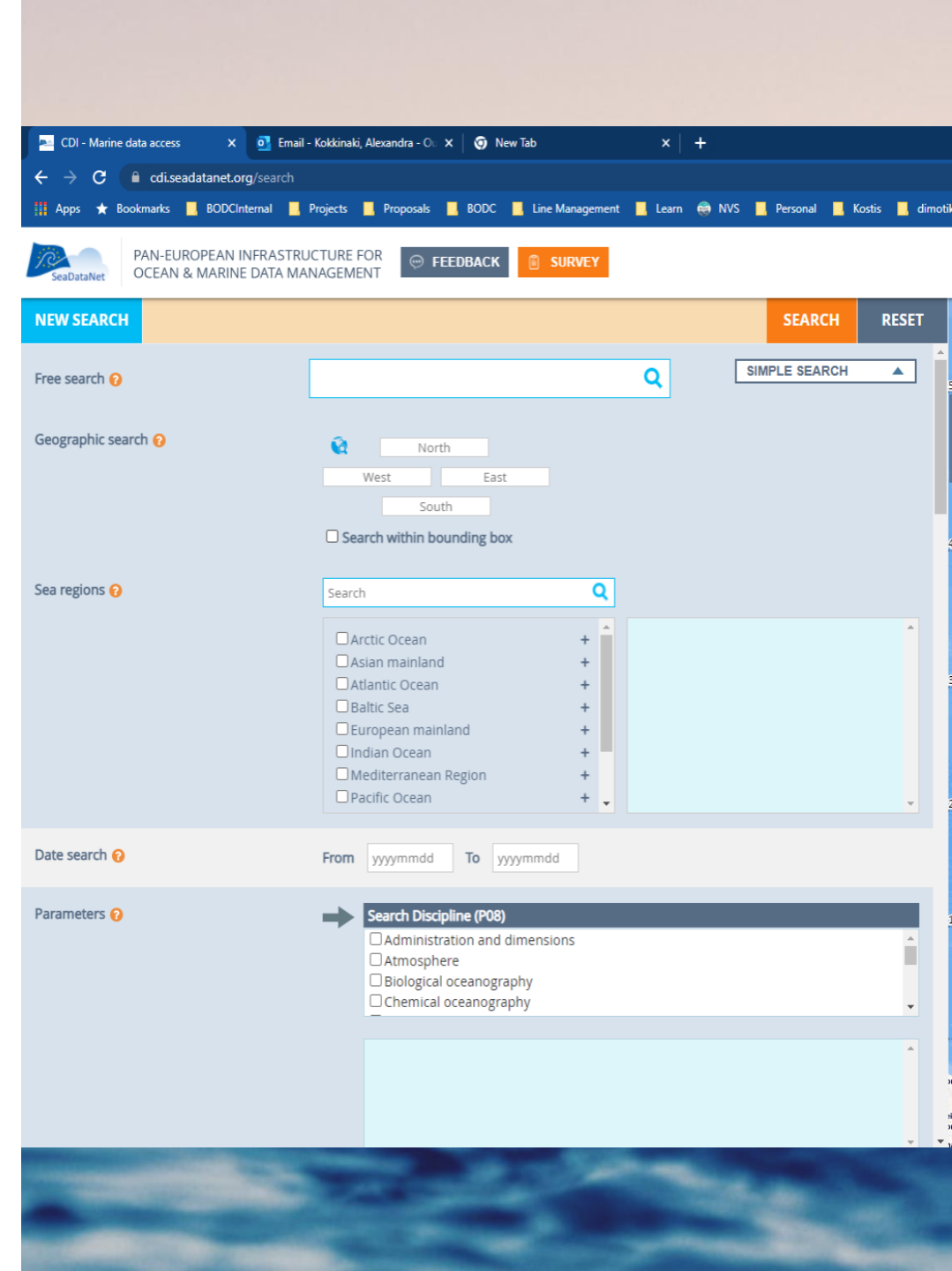
What are vocabularies for?

Semantic annotation of datasets

- Standardise
- Harmonise
- Catalog
- Validate
- Search
- Automation

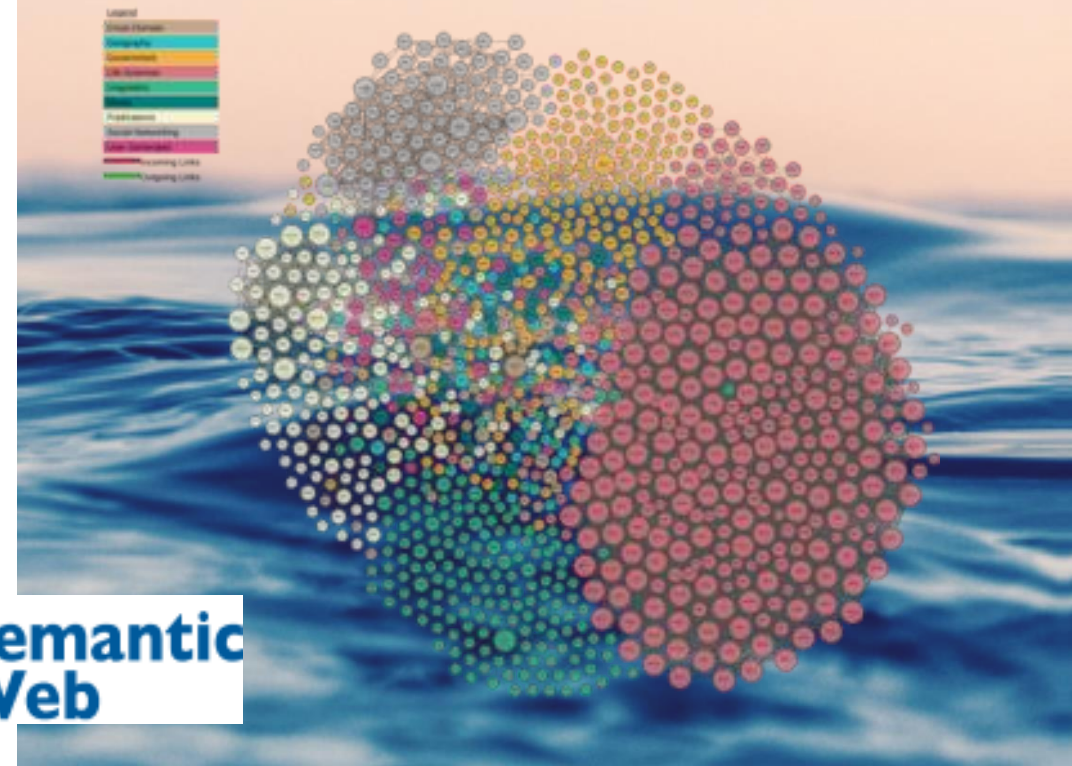
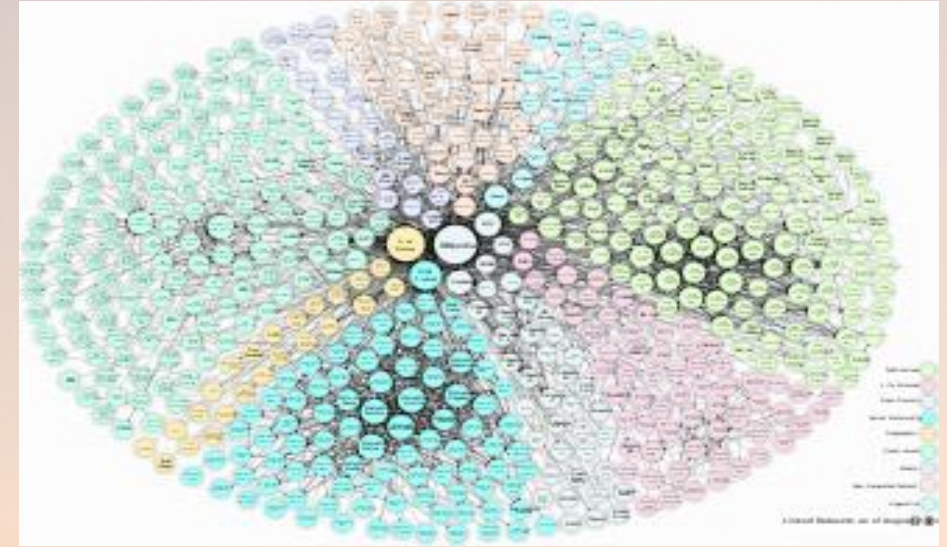
Semantic alignment and mapping

- Interoperability
- Data integration



What do we mean by “vocabularies”?

- Units of knowledge
- Terms used to represent “things”
- Organised collections of related terms
- E.g. controlled lists, thesauri
- Terms can be linked with relationships
 - *e.g. Broader, Narrower, Same as, Related*
 - *And also more expressive relationships using ontologies*
- Foundation on which to build more complex reasoning



Why “FAIR” vocabularies?

- FAIR vocabularies enable the “I” of FAIR
- Findable, Accessible, Interoperable and Re-Usable
- Applies to data AND vocabularies
- Data cannot be interoperable without adoption of common vocabularies and/or alignments between vocabularies
- They need to be served in formats that can be read by both humans and machines

Temp(degC)

?

sw_temperature

http://vocab.nerc.ac.uk/collection/A05/current/EV_SEATEMP/

http://vocab.nerc.ac.uk/standard_name/sea_water_temperature/

<http://vocab.nerc.ac.uk/collection/P01/current/TEMPPR01/>

Human- and Machine accessibility

Concept

Temperature

URI http://vocab.nerc.ac.uk/collection/A05/current/EV_SEATEMP/

Within Vocab [AtlantOS Essential Variables](#)

Alternative Labels [Sea temperature](#)

Definition The temperature of seawater expressed against the International Temperature Scale 1990 (ITS-90) standard.

Date 2016-06-23T15:58:53

Identifier SDN:A05::EV_SEATEMP

Note accepted

Has Current Version 1

version 1

Same As

[OD1:OD1SEATEMP](#)

Sea temperature

Narrower

[P07:CFSN0381](#)

sea_surface_temperature

[P01:TEMPST01](#)

Temperature of the water body by CTD or STD

[P07:CFSN0335](#)

sea_water_temperature

[P01:TEMPTC03](#)

Temperature of the water body by in-situ thermistor (third sensor)

[P01:TEMPPR01](#)

Temperature of the water body

Alternate Formats

Other formats for this page:

[RDF/XML](#)

[Turtle](#)

[JSON-LD](#)

Alternate Profiles

Other views of this page:

[Alternate Profiles](#) ?

[PUV html view](#) ?

[I-ADOPT html view](#) ?

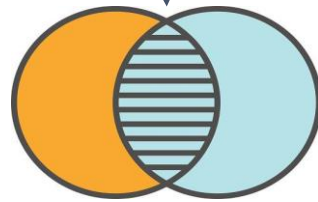
Make data FAIR with FAIR vocabularies



Annotate (encode) data with standardised
FAIR terminologies



Improve
Findability



Enable
Interoperability



Optimise
Re-usability

Different words - same things

British & American English

americanenglish.state.gov

British

torch

/tɔːtʃ/



American

flashlight

/'flæʃ.laɪt/



British

mobile phone

/'məʊ.baɪl fəʊn/



American

cell phone

/'sɛl 'foʊn/



British

rubbish

/'rʌb.ɪʃ/



American

garbage (trash)

/'gɑːb.ɪdʒ/



British

bin

/bɪn/



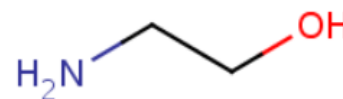
American

trash can

/'træʃ 'kæn/



Images ©Shutterstock.com and may not be reproduced separately without permission of ©Shutterstock.



20
synonyms!

Synonyms

1-amino-2-hydroxyethane

2-amino-1-ethanol

2-Amino-ethanol

2-aminoethan-1-ol

2-aminoethyl alcohol

2-Hydroxyethylamine

Aethanolamin 

Aminoethanol

β-aminoethanol

β-aminoethyl alcohol

β-ethanolamine

β-hydroxyethylamine

colamine

ETA

Ethanolamine

glycinol

Hea

MEA

MONOETHANOLAMINE

monoethanolamine

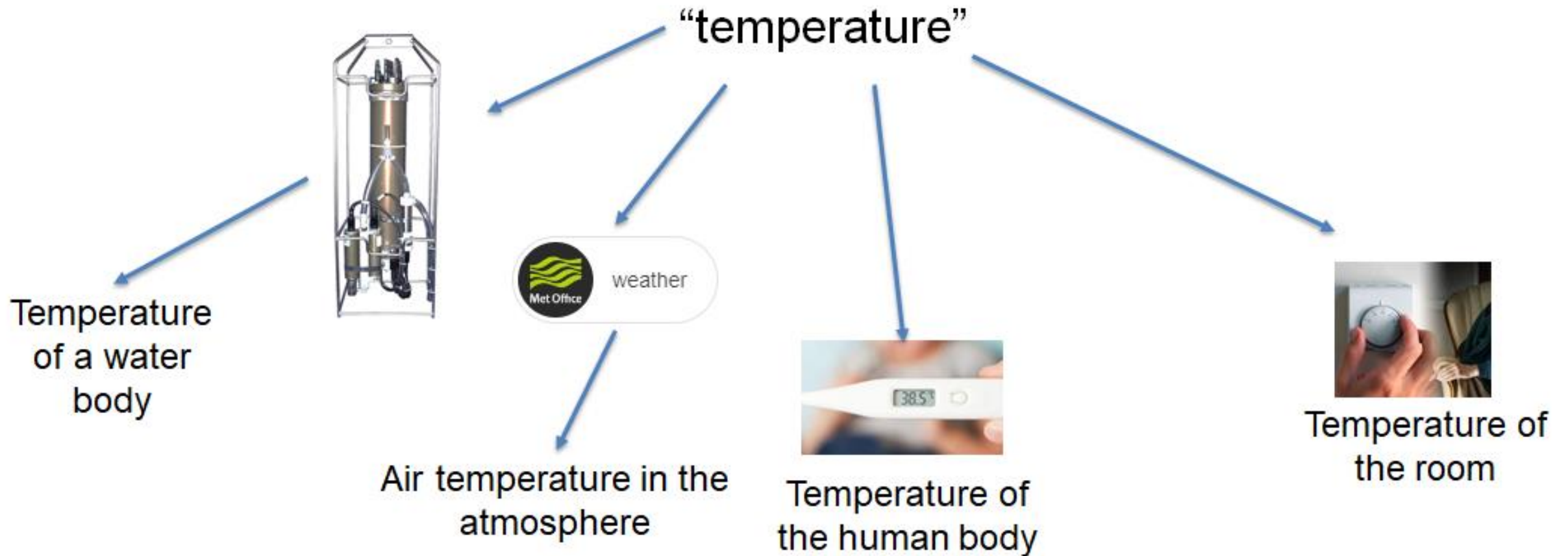
Same word - different meanings



State	1140ml (40 fl oz)	570ml (20 fl oz)	450ml (15 fl oz)	285ml (10 fl oz)
NSW/ACT	Jug	Pint	Schooner	Middy
Victoria	Jug	Pint	Schooner	Pot
Queensland	Jug	Pint	Schooner	Pot
Western Aust.	Jug	Pint	Schooner	Middy
Southern Aust.	Jug	Imperial Pint	Pint	Schooner
Tasmania	Jug	Pint	Schooner	Ten
N. Territory	Jug	Pint	Schooner	Handle



Precise compound vocabularies sometimes needed to resolve ambiguity and make context more explicit and machine-readable

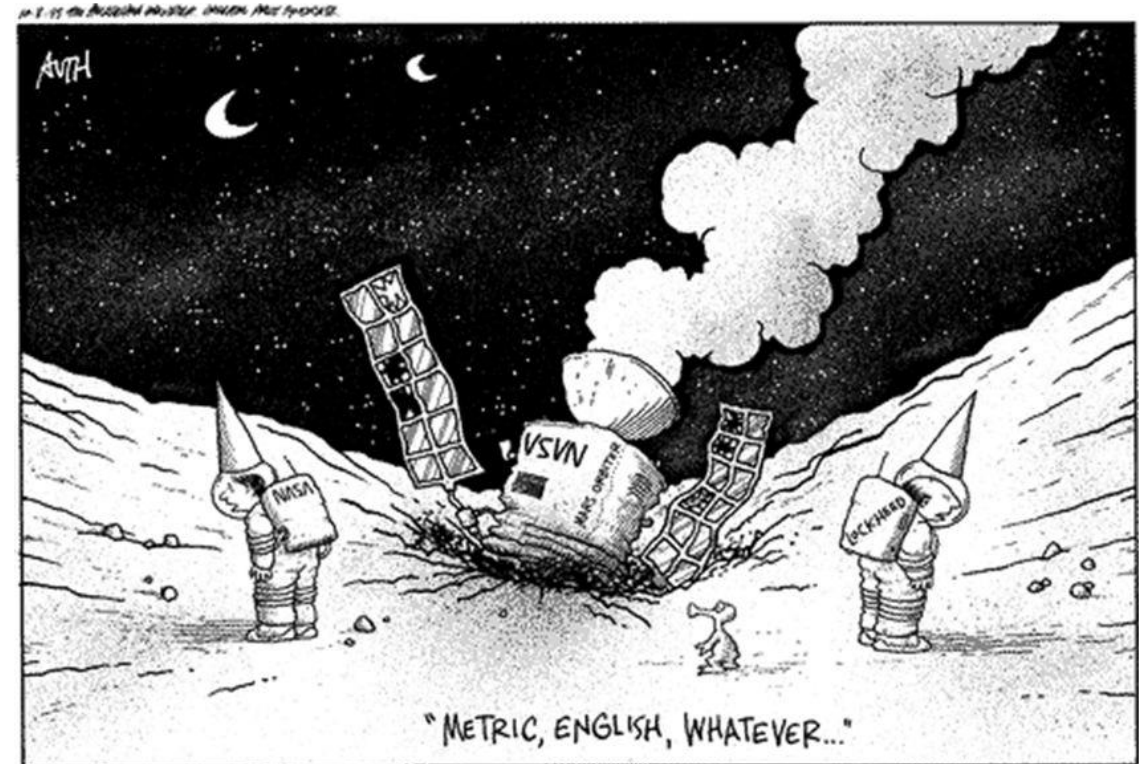


Standardisation is hard!

By capturing units in
machine- and human
readable formats

engineers might have
realised something was
not quite right

and a validation software
could have reported the
error.



Remember the Mars Climate Orbiter incident from 1999?

It could have been saved by
controlled vocabularies!!



Thank you

For more information :

✉ contact@groom-h2020.eu

✉ Twitter : @GROOM2RI

✉ www.groom-h2020.eu

