CF meeting
9-11 June 2020

Discussion topic notes

Standard names: general discussion on any aspect (Alison Pamment)
37 attendees

- The latest standard name table:
  http://cfconventions.org/Data/cf-standard-names/72/build/cf-standard-name-table.html
- The CEDA standard name editor: http://cfeditor.ceda.ac.uk/proposals/1
- A useful reference for background reading is Jonathan Gregory's presentation at the 2018 CF meeting: http://www.met.rdg.ac.uk/~jonathan/talks/CF180620.pdf

These topics are also being suggested for discussion

- Publish CF Standard Names as an Ontology (Alison Pamment)
  https://github.com/cf-convention/discuss/issues/51
- International effort to create a data standard for chemical oceanographic data (?)
  - biological as well?
  - (M. Lankhorst, Scripps): Should there be thousands of standard names like "concentration of XXX"? Or is there a smarter way, such as "concentration of substance", and the actual substance is a coordinate (or something else)? Applies to chemical substances as well as biological species.
  - (N Galbraith, WHOI) Roy Lowry has introduced a mechanism for adding biological species without needing to add a standard name for each, using the species name as a coordinate variable (I think I got that right!) added in V 1.8
    Chem variables had already been defined in standard names, so this has not
  been implemented for those.
  - (F. Santiago-Mandujano, U. Hawaii). We have a list of about 40 biogeochemical variables that don't have a CF name. Do I have to request a variable name individually or can I submit them as a batch?
  - (N Galbraith WHOI) F. Santiago-Mandujano, I can answer that! There's a github template, and you can request an unlimited number of names at once. Just need a definition and canonical units
  - (F. Pacheco, U.Hawaii). Template link
    https://github.com/cf-convention/discuss/blob/master/.github/ISSUE_TEMPLATE/standard-name-template.md?
- DOIs for standard name table?

- (Trevor Smith, Ouranos): I'm curious about the possibility of integrating more climate indices into the standard names table, more wondering how far we can broaden the scope (e.g. agroclimatology : specialized metrics).
  - Ethan Davis: I think some discussion on this may be happening in the "Cell Methods" breakout (Google Doc)
  - (Trevor Smith): I'm not surprised, I know it's something that bleeds into questions of climatology (also, my colleagues are in that session).

NOTES
Kevin:
Mathew Biddle: some communities worry about CF having mission creep
- CF started in Climate and Forecast but has grown beyond that.
- How do we get various domains involved in CF?
- Carolina Berys - does CF have a presence at scientific meetings? (other than when presenters organically mention it in their presentations)
  - Alison Pamment - no I don't think we actively go out and promote CF but maybe we should think of doing this as AGU/EGU meetings for example (
  - Ethan Davis - We have held sessions on CF at AGU and EGU, including one for the upcoming AGU, and other meetings. (Carolina - ok great! I haven't attended AGU and EGU but I do attend the Ocean Sciences meeting, maybe more niche venues would be good as well :) Ethan: Agree, a few of us had a poster at the last Ocean Sciences and also staffed the Data Help Desk. But plenty of other domain meetings that could use some CF attention. (Carolina - yep. I know it is cumbersome but a session might be helpful. I was at the data help desk too, sometimes I feel like it can be preaching to the choir. But right on with the efforts, I'm not trying to take away from what has been done)
  - Mathew Biddle - ESIP?? (Carolina +1)
    - Ethan: There will be a CF session at this summers ESIP mtg +1 (link)

Ethan Davis: It would be good to have CF materials available for posters, presentations, training

Mathew Biddle:
- Example dataset at BCO-DMO w/ bacterial abundance/DOC/etc note nerc_identifier attribute: https://erddap.bco-dmo.org/erddap/info/bcodmo_dataset_471701/index.html

variable	FCM_T0_0		float	
attribute	FCM_T0_0	_FillValue	float	NaN
attribute	FCM_T0_0	actual_range	float	171000.0, 257000.0
attribute	FCM_T0_0	bcodmo_name	String	bact_abundance
attribute	FCM_T0_0	description	String	Bacteria as determined by flow cytometry at T0 (0 hrs)
attribute	FCM_T0_0	long_name	String	FCM T0 0
attribute	FCM_T0_0	nerc_identifier	String	https://vocab.nerc.ac.uk/collection/P02/current/BNTX
attribute	FCM_T0_0	units	String	cells/ml

-

- This is our current ad-hoc approach, not CF compliant...

Gwen mentioned the [I-ADOPT working group](#) An RDA Working group working on cross-domain interoperability of the terminologies used for the description of observable properties.

Kevin mentioned FAIR maturity model:
https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

References to external vocabularies discussed\\

Gwen: FAIR-ness criteria for semantic resources/terminology/artefacts is also being worked on at present. See e.g. the work being done under FAIRsFAIR [essential criteria(WIP)](#). It is also a task of an RDA Interest Group called [VSSIG](#) (Vocabulary and Semantic Services Interest Group)

Should define requirements for external vocabularies to be included in CF

CF Standard Names table available as Ontology
"Official" URI for CF standard names -
- Need cfconventions.org redirections to NVS CF standard names URIs
- Also, host TTL file on cfconventios.org

ACTION: Alison will organize a Zoom, etc. to discuss redirection of URI
- Alison, Fran, Ethan, Gwen, Alex

ACTION: Ethan will organize zoom call to discuss CF Standard Names build process and clean up
- Include: Alison, Fran, Ethan, Andrew Barna…

Links copied from chat:
From Gwen Moncoiffe (BODC) to Everyone: 07:16 PM
This is the link:
https://www.rd-alliance.org/groups/interoperable-descriptions-observable-property-terminology-wg-i-adopt-wg
From Kevin O'Brien (UW/JISAO) to Everyone: 07:22 PM
FAIR data maturity model: https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

Summary:

1. Q: How do we get people from other scientific domains involved in CF (given that it has expanded beyond the original Climate-Forecast community)?

   A: We should try to have a CF presence at more scientific conferences e.g. AGU, EGU, ocean conferences. It would be good to have some pre-prepared CF materials available for posters, presentations, training that we can all use as needed.

2. Q: Should we decide to include large sets of new standard names, e.g. to serve the ocean chemistry community, in the standard name table or should we point to external vocabularies as we are now doing for biological taxa?

   A: If a 'suitable' external vocabulary already exists then we should use it. By 'suitable' we mean something that passes FAIRness tests. See e.g. the work being done under FAIRsFAIR essential criteria(WIP) and RDA Interest Group VSSIG (Vocabulary and Semantic Services Interest Group). Every term in external vocabularies should have an individual URI that resolves to something!
   If there is no suitable external vocabulary then the CF community should do the work of agreeing the new standard names so that data producers can get on and write their data.

3. Q: We discussed the issues currently being raised in https://github.com/cf-convention/discuss/issues/51 ("Publish CF Standard Names as an Ontology").

   A: We think it is possible to have URIs for each standard name in CF "namespace", i.e. to include "CF" somewhere in the URI. This could be based on the terms served in NVS and would probably need redirection at DNS level. We will arrange a Zoom meeting to discuss the technical details and take this forward in the near future. Having a Turtle file on the CF website for use with ontology tools is a good idea, and would be even more useful if we can set up the redirection.

4. Q: Is it possible to view diffs between standard name versions in a simple way?

   A: At the moment there isn't an easy way to do this on the CF website or in GitHub because of the way the files are organised in the repository. A Python tool has been developed (by Harry Singh?) for finding diffs between XML versions of the standard name table and this could be advertised on the CF website. Also we could tidy up the way the standard names files are stored in the website repo so that it would be easier to see diffs using the features of GitHub.

- EUMETSAT has a tool for viewing vocabularies that is currently restricted to WMO formats. It's here: https://vocabulary-manager.eumetsat.int/. You can use it to view e.g. the differences between WMO BUFR & GRIB table versions, which are in the end pretty similar to CF Standard Names. - user driven tool - could request CF to be added