Semantic Web for Earth and Environmental Terminology (SWEET) 2018

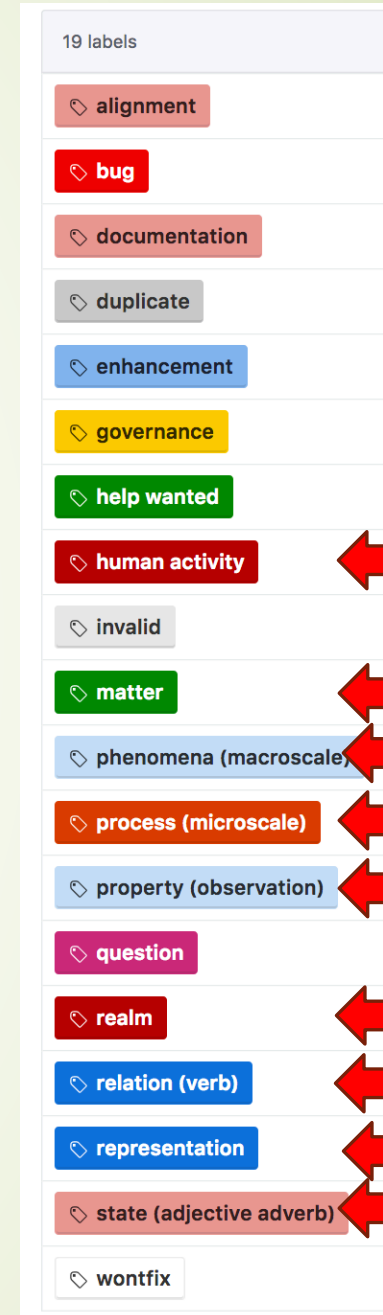Status, Future Development and Community Building

Agenda and Purpose

- Current status of SWEET e.g. What has the community been doing?

- Explore the abstract structure of SWEET [1] e.g. What are the primary information topics/categories/clusters and how are they currently extended?

- Highlight the development priorities moving forward.

- Explore ongoing alignment [2] activities between SWEET and other semantic or linked data resources.
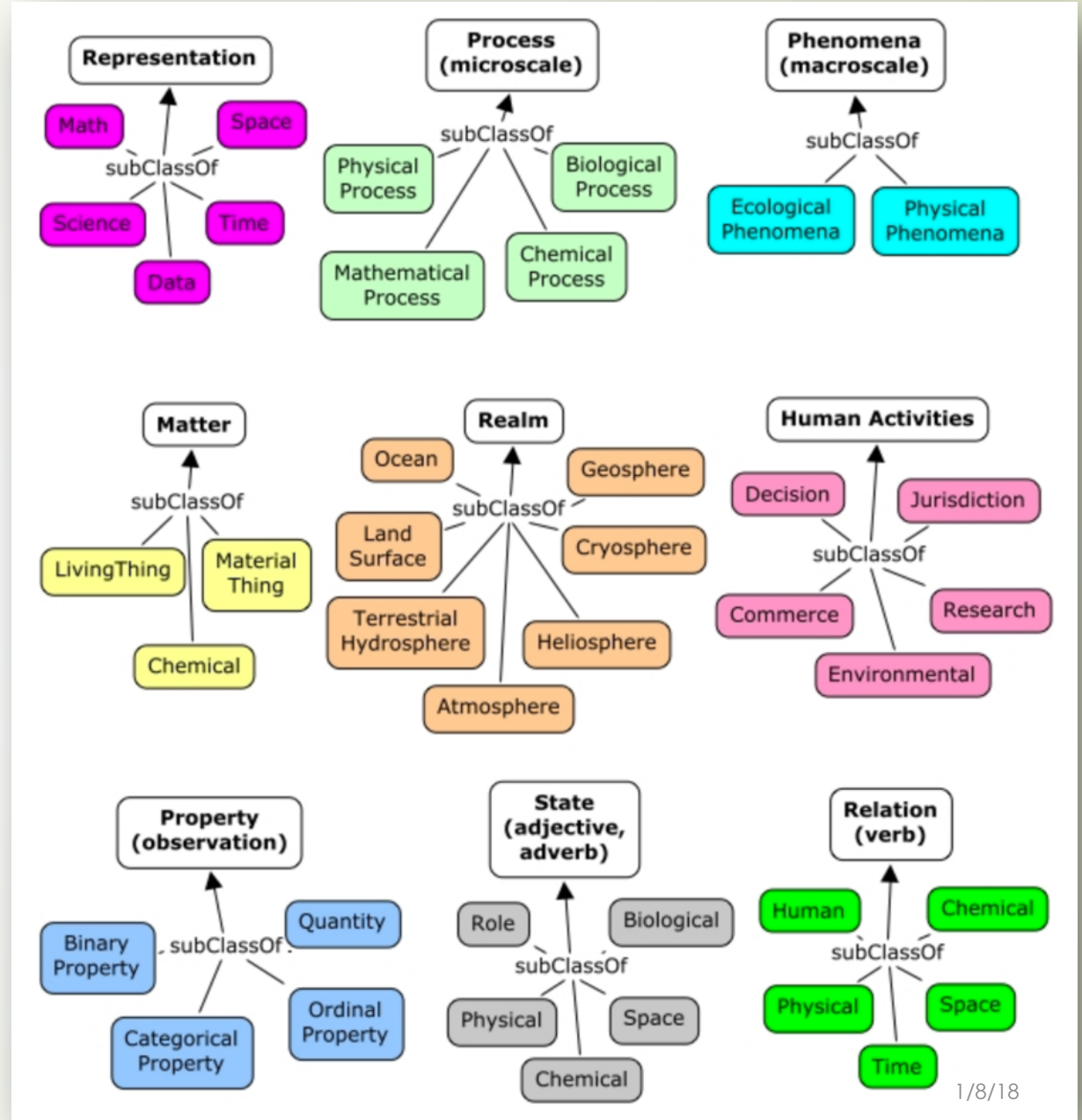
What is SWEET?

- SWEET is a highly modular ontology suite with 6924 concepts (Classes, Object Property, Data Property and Individuals) in 225 separate ontologies* covering Earth system science. SWEET is a mid-level ontology and consists of nine top-level concepts that can be used as a foundation for domain-specific ontologies that extend these top-level SWEET components.

- SWEET's own domain-specific ontologies, which extend the mid-level ontologies, can provide users interested in further developing a particular domain with a solid set of concepts to get started.

- SWEET ontologies are written in W3C Turtle; the Terse RDF Triple Language and are publicly available under the Apache License v2.0.

* Numbers are accudate as of January 2018

19 labels

- alignment
- bug
- documentation
- duplicate
- enhancement
- governance
- help wanted
- human activity
- invalid
- matter
- phenomena (macroscale)
- process (microscale)
- property (observation)
- question
- realm
- relation (verb)
- representation
- state (adjective adverb)
- wontfix

High-level Information Categories

Current status of SWEET

- As of mid 2017 version 3.x of SWEET was recently released under new community development, management and governance by ESIP's Semantic Technologies Committee (STC) [3].

- Development follows a review-then-commit (RTC) policy for source code contributions. The development guidelines can be seen at [4]. These are automatically shown to any contributor upon either creation of a new Github issue or pull request.

- Advances in SWEET since its transition from NASA JPL over to the open source, community-driven management and governance structure now overseen by the ESIP STC include (i) URI transition and governance, transition from OWL to Turtle serialization, linked data dispatch via the ESIP Community Ontology Repository [5], and ongoing alignment activities with existing semantic technology resources such as the OBO Foundry collection, W3C SOSA/SSN, W3C PROV-O, etc.,

URI transition and governance

- The canonical details for this topic can be found at [6]

- **Essentially, it boils down to the URI transition from http://sweet.jpl.nasa.gov/… to http://sweetontology.net/...**

- The '#' character has also been substituted for '/' e.g. http://sweetontology.net/reprDataProduct#Dataset to http://sweetontology.net/reprDataProduct/Dataset

- There is now no file suffix e.g. *.owl, content negotiation is managed by the hosting platform (COR) in an adhoc manner, available serializations include OWL, Turtle, RDF/XML, N-Triples, etc. An example would be http://sweetontology.net/stateEnergyFlux?format=rdf

Transition from OWL to Turtle serialization

- SWEET ontologies are written in W3C Turtle [7]; the Terse RDF Triple Language

- We have tools available to ensure that the content of each file is 'prettified' e.g. alphabetically ordered, blank nodes use [...] syntax for clearer syntax as supported by the OWL-API.

- More information can be found at [8]

Linked data dispatch via the ESIP Community Ontology Repository (COR)

- Canonical documentation can be located at [9]

- Full service GUI is available at http://cor.esipfed.org

- Core REST API documentation can be located at http://cor.esipfed.org/ontapi

- The COR software is being used on several other projects e.g. Marine Metadata Interoperability (MMI) project https://mmisw.org/.

- The COR software is maintained at the MMI Github https://github.com/mmisw and has an active team of developers.

- Both follow-on sessions below will cover MUCH more information on this

  - GeoSem Symposium – 1:30-2:15pm & 3:30-4-15pm (Carlos Rueda, Felimon Gayanilo)

  - COR Developer Workshop - **Thursday**, January 11 • 3:30pm - 5:00pm http://sched.co/D6oC (Carlos Rueda, Beth Huffer, Lewis McGibbney)

Development Priorities (General)

- ISSUE-20 Annotation of SWEET terms [10]

  - Very few SWEET terms have human-readable annotations.

  - The definitions of SWEET terms are implied by the axiomatization, which is often incomplete.

  - Every SWEET term should have at least

    - rdfs:label (and/or skos:prefLabel or dct:title) for the name

    - rdfs:comment (and/or skos:definition or dct:description) for a text definition

  - We hope the YAMZ effort will help to move this issue forward. YAMZ will be covered at 12:45-1:30pm & 2:45-3:30pm today!

  - Also see ISSUE-33 [11] for more information.

Development Priorities (Alignments)

- Coordination with Open Bio Ontologies (OBO) [12]; preliminary results can be found at [13]. This has resulted in alignments across 15 OBO resources… we need to solicit input from domain experts now to verify alignments before they are formalized.

- Formal alignment between SOSA, SSN and SWEET [14]. Formalizes 7 concepts including ***Observation***, ***Result***, ***Sample***, ***Variable***, ***Property***, ***Sampling*** and ***Statistical Sample***. N.B. alignment mapping graphs are currently limited to one file and are maintained separately from the source and target(s) being mapped. Depending on mutability of both SWEET and potential alignment targets, this does increase maintenance… discussion on managing this is required.

- A generic alignment tool named the ***SWEET Alignment Manager*** (SAM) has been created and now resides at [15]. SAM provides a pipeline and results for aligning SWEET with various ontologies. This is semi-automated but the goal is to learn from a seed of initial curated equivalence axioms using the Agreement Maker Light (AML) (16). AML is a lightweight ontology matching system specialized on the biomedical domain but applicable to any ontologies. It can be used to generate alignments automatically, as a platform for reviewing alignments, or as an alignment repair system (both automatically and interactively).

References

1. https://github.com/esipfed/sweet
2. https://github.com/ESIPFed/sweet/tree/master/alignments
3. http://wiki.esipfed.org/index.php/Semantic_Technologies
4. https://github.com/ESIPFed/sweet/blob/master/CONTRIBUTING.md
5. https://cor.esipfed.org
6. https://github.com/ESIPFed/sweet/wiki/SWEET-IRI-Patterns-for-Ontologies-and-Their-Terms
7. http://www.w3.org/TR/turtle/
8. https://github.com/ESIPFed/sweet-tools
9. https://github.com/ESIPFed/sweet/wiki/sweetontology.net
10. https://github.com/ESIPFed/sweet/issues/20
11. https://github.com/ESIPFed/sweet/issues/33
12. http://www.obofoundry.org/
13. https://github.com/cmungall/sweet-obo-alignment
14. https://github.com/ESIPFed/sweet/blob/master/alignments/sweet-ssn-mapping.ttl
15. https://github.com/ESIPFed/sweet-tools/tree/master/sweet-alignment-manager
16. https://github.com/AgreementMakerLight/AML-Jar