Community Development of the SWEET semantic system for Earth and Environmental Data – A Call for Interest

Robert J. Rovetto

ESIP Semantic Technology Committee member

rrovetto@terpalum.umd.edu

https://purl.org/space-ontology

Brandon Whitehead ESIP Semantic Technology Committee chair



Introduction

Organizer: Robert Rovetto

- Member ESIP clusters & partner orgs.
 - Semantic Tech Committee, Etc.
- Conceptual modeling, Formal ontology, Terminology. Ethics, methodology.
- Focus on space ontology... https://purl.org/space-ontology
- Cert. commercial mariner.
 Water rescue training & focus.
- Aspiring PhD student, actively applying rrovetto@terpalum.umd.edu

Meetings: https://tinyurl.com/y6anucqp Services: https://tinyurl.com/yas7trzy

Co-organizer: Brandon Whitehead

- Member ESIP clusters & partner orgs.
 - Semantic Tech Committee Chair
- Environmental Data Scientist, Manaaki
 Whenua -- Landcare Research

Agenda GoogleDoc: https://tinyurl.com/4wfhcadz

Questions for attendees: https://tinyurl.com/SWEET-Questions

Acknowledgement of indigenous history & cultures.

Acknowledgements

- o Input by B.Whitehead.
- o Input by L.McGibbney (former SemTech chair)
- o Made by R.Rovetto. Min. person hours: 70 hours work. No funding to declare.

Images with shadows, beginning slide 8, from PPTs (see references)

Session Roadmap

Session Roadmap

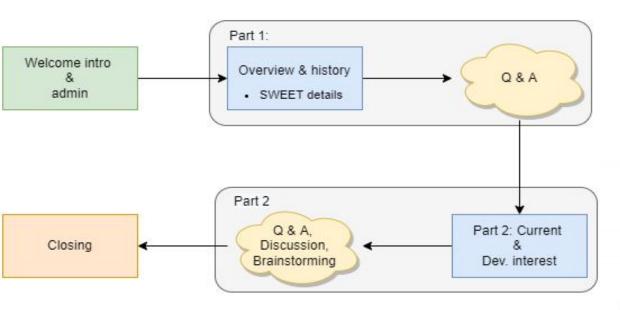


Diagram by Robert Rovetto. Contact: rrovetto@terpalum.umd.edu Hire for diagraming & conceptual modeling: https://tinyurl.com/4bu74bmc



Questions for participants

- What are your research topics/expertise?
- What types of data do you work with?
- Do you have modeling challenges (and what are they)?
- What would you like to do with your data that you presently cannot?
 - How might SWEET help with that?

Can answer at:

https://tinyurl.com/SWEET-Questions or agenda page

In a nutshell

What is SWEET ☐ a semantic technology. a knowledge organization system Why should I care? potential to help with your data needs, r&d, innovation,... What problems is it solving?

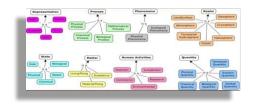
content search, retrieval; syn, homyn support, ...innovate (Why important to develop) \Box for the same, for Earth sci, for innovation exploration

Computational Ontology(s)

- · Various definitions. sometimes called semantic model, conceptual model, knowledge graph, ...
 - Living catalog https://tinyurl.com/yrm78zpn
- Various dev. methodologies → automated, manual, ...
- Can be developed to various degrees of complexity, size, and abstraction
 - sometimes used as other KOSs: as a taxonomy, a controlled vocab, as a database, ...
- Pros & Cons as compared with other KOS
- one of many KOS composed of categories (metadata), capable of being formally defined (for computation & semantics) & structured (e.g., like a taxonomy)
 - A vocabulary + computable rules that constrain the meaning and use of the vocabulary \rightarrow for formation of expressions that can be computed in machine inference, data queries, ...
- Act as a semantic layer, providing an interpretation of the meaning of data elements
- As a terminological system, its terms can be mapped to others
- Can be *linked* to other ontologies

Ontology Consulting: https://tinyurl.com/34u9w6wx

Caution: Should ask... does a given ontology provide the desired/right meaning (or semantics) for your data? Owners & users should examine the (in)formal meaning expressed in any ontology (and what it assumes, imports) to determine suitability



SWEET Basics

- Semantic Web for Earth and Environmental Terminology (SWEET)
 - a suite of modular ontologies implemented in OWL
 - originally in the DARPA Markup Language (DAML) [7]
 - thousands of terms across relevant topics
- Developed at NASA JPL by Dr. Rob Raskin
 - used GCMD keywords. Decomposed into facets
- ESIP is current the steward of SWEET
 - approx. Mar 2017 SWEET files migrated to ESIP, by L. McGibbney (JPL)
 - https://github.com/ESIPFed/sweet
 - https://wiki.esipfed.org/Semantic_Technologies
 - living work. versions 3.x



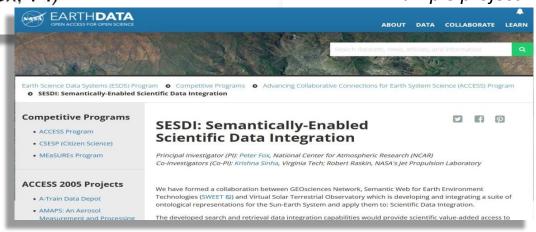
NASA Support for SWEET

- AIST (2002-05)
 - SWEET development
- AIST/ACCESS (2006-09)
 - SESDI (Semantically-Enabled Science Data Integration) (Peter Fox, PI)

Example project

SWEET Users

- ESML- Earth Science Markup Language
- ESIP Earth Science Information Partner Federation
- GEON- Geosciences Network
- GENESIS- Global Environmental & Earth Science Information System
- IRI- International Research Institute (Columbia)
- LEAD- Linked Environments for Atmospheric Discovery
- MMI- Marine Metadata Initiative
- · NOFSIS
- PEaCE- Pacific Ecoinformatics and Computational Ecology
- SESDI- Semantically Enabled Science Data Integration
- VSTO- Virtual Solar-Terrestrial Observatory



Some SWEET Benefits & Applications (1)

"Special emphasis on improving search for NASA Earth science data resources

- Atmospheric science, oceanography, geology, etc.

Provide a common semantic framework for describing Earth science information and knowledge

SWEET provides semantics tags to interpret data

SWEET supports model interoperability

- Earth Science terms
- Compatibility of model parameterizations, modules

SWEET supports grid interoperability

- Earth Science terms
- Grid concepts

Contributions of sweet - Improved data discovery without exact keyword matches" [1]

Some SWEET Benefits & Applications (2)

Circa 2011 and later

Used in NextGen Network Enabled Weather (NNEW) Ontology, within the US NextGen Air Transportation System project

2012

"we plan to replace part of the data-attributes ontology by importing the relevant Semantic Web for Earth and Environmental Terminology (SWEET) ontologies (Raskin and Pan 2005). SWEET provides **support for scientific and numerical concepts**, such as scientific units, scientific relations, provenance, and data representation. We believe that ontologies should carry as few ontological commitments as possible" [2, emph added]

Side note to be aware of: Relates to both technical & ethical aspects of ontology/semantics. Contact: rrovetto@terpalum.umd.edu to discuss.

Some SWEET Benefits & Applications (3)



- Consults knowledge base to find alternative meanings
 - Clustered by: synonyms, parent, children
- Enables discovery of resources without [9] exact keyword match

2014:

"This investigation suggested that the SWEET ontologies do sufficiently represent the EES domain, and are broad in their topic coverage." [5, emph added]

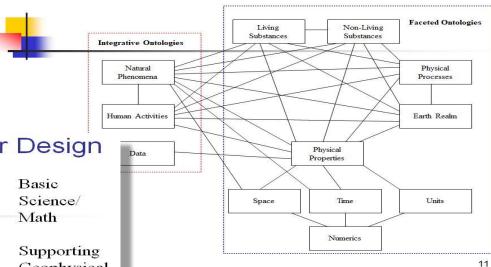
2021:

- "We retrieve phenomena and domain information using SWEET" [3, an ESIP presentation]
- "To create knowledge representations of science carried out in these publications, we use existing ontologies such as GCMD and SWEET." [4, an ESIP presentation]

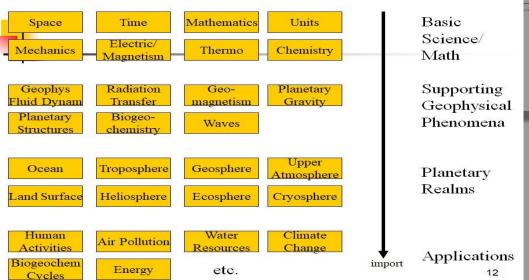
SWEET Versions

v. 1 to 3.5 (as of 2021)

SWEET 1.0 Ontologies

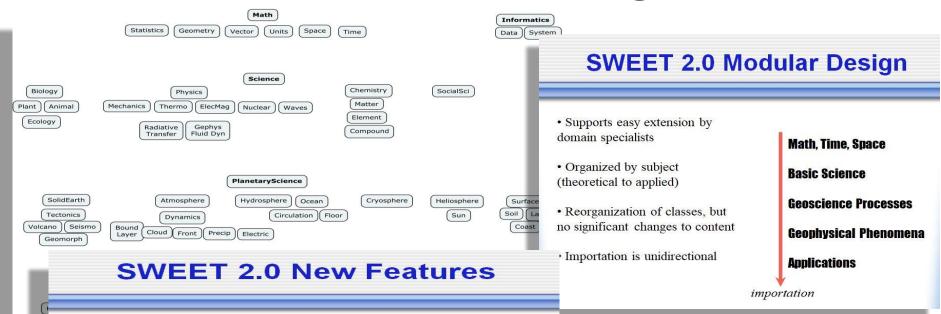


SWEET 2.0 Ontologies: Modular Design

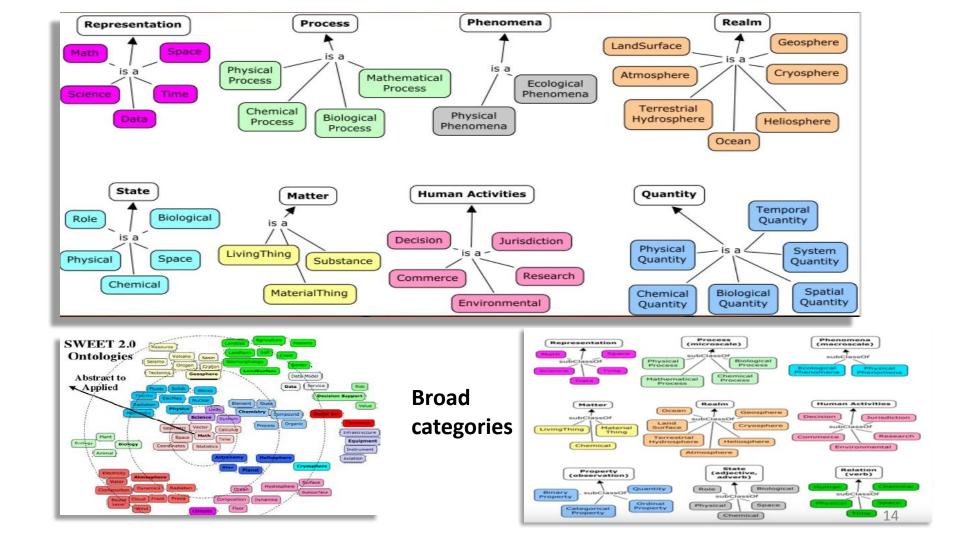


Potential for SME to contribute to specific topics

SWEET 2.0 ontologies



- Organized by subject
 - Makes it easy for domain specialists to add new modules
- Smaller, modular ontologies
- 23 ontologies -> 80 ontologies



Selected Broad category/facet Selected Category terms Faceted Science Ontology Earth realms: atmosphere, Ocean LandSurface SWEET Ontologies and Their Interrelationships Physical properties: temperature, composition, area, ... **Substances:** CO2, water, lava, hydrogen, ... Non-Living Living Living substances Humans, fish, ... Substances Substances Integrative Ontologies Diffusion, absorption, ... **Processes** Phenomena **Integrative Science Ontologies** Phenomena Thunderstorm, Deforestation, Physical pro Human Activities magnetism, convection, ...) **Human Activities** Fisheries, Industrial Processing Data Properties **Numerical Ontologies** Intervals, numeric relations Space Time **Functions** Statistical concepts Numerics Spatial concepts/entities: Relations (above, etc.); 0-, 1-, 2-, 3-D; coor Temporal concepts/entities: Instants, durations; Relations (after, etc.) "Each box represents a separate ontology, and a connecting line indicates where major properties are used to define concepts Spectral band UV. . . . across ontology spaces" Data Ontology

Format, data model, ...

Scale factors, offset, ...

Source, ...

15

Physical

Processes

Earth Realm

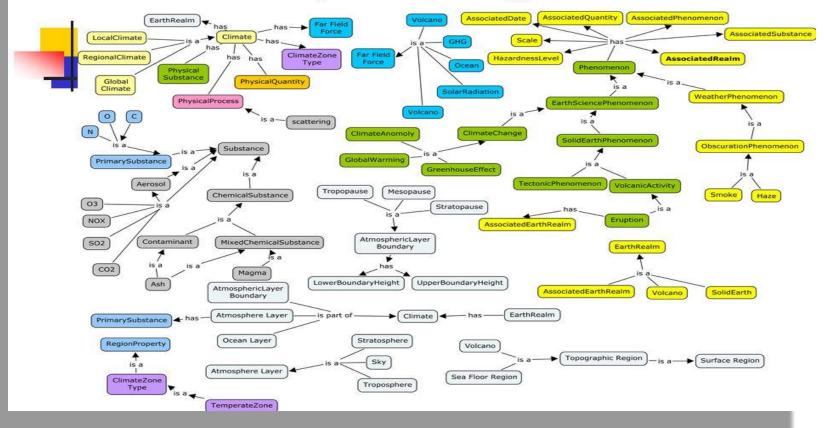
Units

Dataset characteristics:

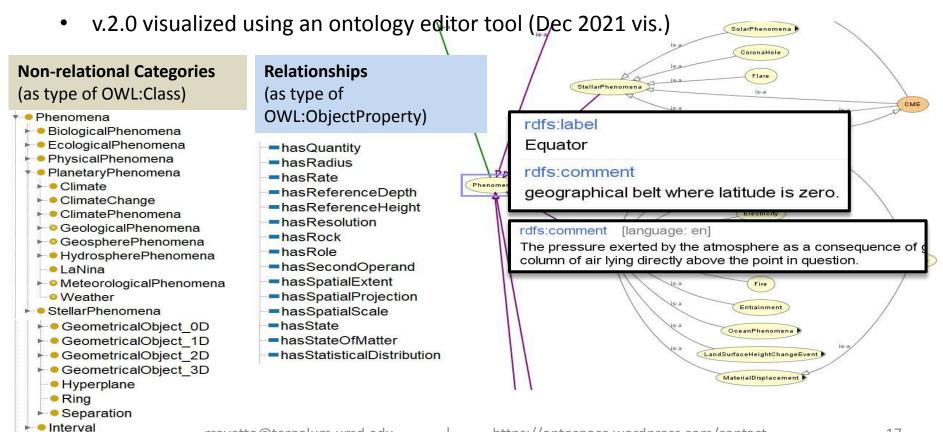
Provenance:

Parameters:

Atmosphere Ontology



Some SWEET concepts



New in SWEET 2.1

- Version 2.1 provides a large module set for STATE
 - Mostly adjectives and adverbs
 - Examples: Role, color/band, shape, size, equilibrium type, activity level, connectedness, impact, substance form
 - Qualitative analog of "Quantity"
 - Enables greater separation of adjective from object to improve modularity
 - Mostly represented by OWL individuals
 - "types" of categorical, ordinal, and cardinal

SWEET 2.1 Statistics

- Classes: 4400
- Individuals (mostly States): 2200
- Relations: 600

Example Development & Modeling Activity

CF vs SWEET Representation

CF (traditional single-attribute parameter name): tendency_of_mole_concentration_of_dissolved_inorganic_phosphorus_in_sea_water_due_to_biological_processes

SWEET (multi-attribute parameter name):

- Quantity= mole concentration
- Transformation= tendency
- State= dissolved, inorganic
- Substance= phosphorous
- Medium= sea_water
- Process= biological processes

[10]

Ontology Schematic Phenomena EathRealm hasAssociatedEarthRealm WheatherPhenomena Atmosphere Ocean hasAssociatedEarthRealr Wind Hydrometeor AssociatedSubstance Substance [10] Fog Precipitation Cloud hasAssociatedSubstance OxygenObject HydrogenObject Chemistry Math Physics import WaterObject Space Property SWEET EarthRealm Time Process, Phenomena hasAssociatedSubstance Substance LiquidWater Ice Data import Stratospheric Specialized Biogeochemistry Chemistry domains

Some SWEET findings

Table 2 - SWEET Usability Measures

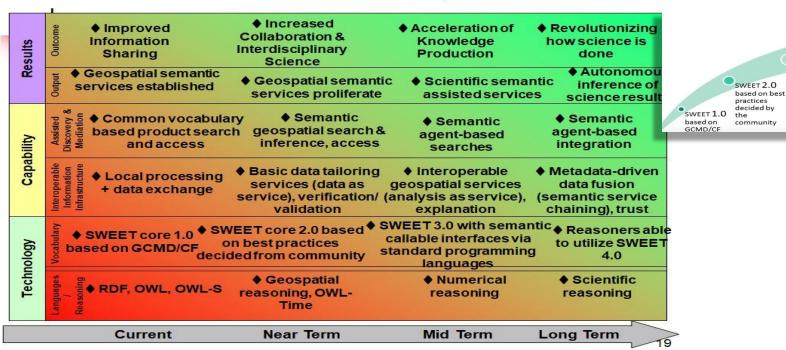
Metric	Evaluation output
User types using SWEET	Knowledge Engineer/ Project Manager Application User/ Ontology Developer
For what type of use was it conceived in first place?	Improving search for NASA Earth science data resources
How easy is it to extend SWEET?	Modular Design / Organized by Subject / Simple steps for extension
Cost of using the SWEET ontology	Free/ Open Source
SWEET Popularity	NASA/ The ESIP Federation/Taught Widely in Semantic Web Courses in Universities
Time to open SWEET	Less than a minute on a normal commodity machine [8]



SWEET Future Community Plans

- Gain further support from Earth system science community
 - Workshop at Summer '08 Meeting of eSIP Federation
- Submit SWEET as community standard to NASA Earth Science Standards and Processes Working Group

Semantic Web Roadmap



SWEET

Semantic

able to

utilize

Reasoners

4.0

SWEET 3.0

interfaces via

programming

semantic

callable

standard

languages

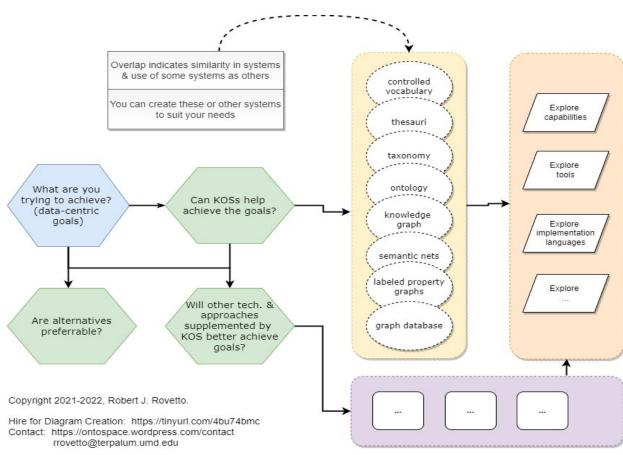
Background considerations: do you need a KOS, semantic system?

Knowledge organization systems, semantic resources, ...

Related systems & approaches

- Formal concept analysis
- NLP, Machine learning
- Model-based systems engineering
- ..

Can semantic approaches, in general, help?



Some considerations for your data goals/needs

- Various types of semantic systems, KOS, & tech. for your data
 - KOS: from thesauri to taxonomies to knowledge graphs. Varying functionalities, limitations, ...
 - Some (or none!) may be most efficient for your needs.
 - Ontologies can be made to varying degrees of complexity.
 - Can "automatically generate the tags during the indexing process. Automatic tag creation involves natural language processing to ascertain the meaning of a term based on its context. In some cases, terms have multiple meanings, and tools such as Latent Semantic Analysis" [7, emphasis added]
- Consider also non-semantic/semantic-web systems and approaches
- Ability & precedent vision to develop SWEET to higher degrees of complexity
 - a matter of will, resources, etc. to do it

End of Part 1 (overview & history) – Questions - Discussion

What do you find interesting thus far?

https://tinyurl.com/SWEET-Questions

- Consider the original questions
 - What are your research topics/expertise?
 - What types of data do you work with?
 - What would you like to do with your data that you presently cannot?
 - How might SWEET help with that?
 - Are you interested in contributing to SWEET?
 (e.g., SME, content, use-case development, applications, misc., etc.)
- Questions, ideas?

Part 2: SWEET Today & SWEET Development

- Updates
- Contributing page on GitHub
- Feedback from 2020 survey by B.Whitehead w/input by R.Rovetto
- Q&A Discussion
 - Neutral SMEs for each topic covered by SWEET
 - Propose, Create, Verify content, e.g. . . .

 - Verify definitions
 SME quality control (e.g., verify accuracy)

Some SWEET updates at

- ESIP is current steward of SWEET
 - https://github.com/ESIPFed/sweet
 - https://github.com/ESIPFed/sweet/releases v. 2.3 □ 3.4
 - Semantic Technology Committee
 - https://wiki.esipfed.org/Semantic Technologies
 - https://github.com/ESIPFed/sweet/wiki/Publications
 - Literature: https://tinyurl.com/Biblio-SWEET
 - SWEET is a living work

Example updates via SemTech:

- 2018: "URI transition and governance, transition from OWL to Turtle serialization, linked data dispatch via the ESIP Community Ontology Repository [5], and ongoing alignment activities" [6]
- Can use GitHub for development ideas & suggestions

SWEET online - the ontologies

- Contributing description page:
 - https://github.com/ESIPFed/sweet/blob/master/CONTRIBUTING.md

Development

The development process for SWEET follows the Review-then-Commit software development process. For a Subsection below

How to work with us on Github, using git command line:

Browse on ESIP repository: http://cor.esipfed.org/



http://sweetontology.net/phenEcology http://sweetontology.net/phenElecMag http://sweetontology.net/phenEnergy http://sweetontology.net/phenEnvirImpact http://sweetontology.net/phenFluidDynamics http://sweetontology.net/phenFluidInstability http://sweetontology.net/phenFluidTransport http://sweetontology.net/phenGeol http://sweetontology.net/phenGeolFault http://sweetontology.net/phenGeolGeomorphology http://sweetontology.net/phenGeolSeismicity http://sweetontology.net/phenGeolTectonic http://sweetontology.net/phenGeolVolcano http://sweetontology.net/phenHelio http://sweetontology.net/phenHydro http://sweetontologv.net/phenMixing http://sweetontology.net/phenOcean http://sweetontology.net/phenOceanCoastal http://sweetontology.net/phenOceanDynamics http://sweetontology.net/phenPlanetClimate

List of SWEET ontology files (subject-specific)

Some SWEET feedback (2020) (1)

"If I want to make additions or suggestions where do I go to do that [understand SWEET dev.]? Is it possible to merge my domain ontology with SWEET?"

- https://github.com/ESIPFed/sweet/
- (Idea) Exploratory dev. track: solicit persons to download versions, merge their domain ontology, in order to create innovating beneficial use-cases for Earth Sci data.
 - Example: 2005 with hydrogeology [11]
- Identify approaches, challenges, desiderata, etc.

Some feedback (2)

"Reach out to a wide swath of earth sciences and run them through an exercise of identifying their terminology and explaining how they relate to other terms in plain language. Getting input needs to be accessible and the approach structured."

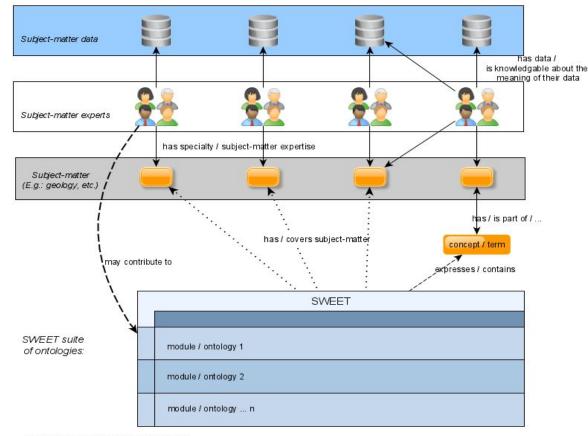
• In organizing this session, perhaps first-steps, brainstorming ideas, ...

Some feedback (3)

"Need to increase visibility across federal agencies and extend the usage. Need specific maintenance team for **different subjects** in SWEET."

- (Idea) Neutral SME potentially help with their subject covered by SWEET
 - Note: neutral = focus on their subject/discipline/science. No preference/advocating for particular product, ontology, way of semantically modeling, characterization, etc.

Idea for community participation to develop & utilize beneficial applications of SWEET



Example Tasks:

Task: SME provides a set of subject-matter concepts or vocabulary.

Task: SME proposes definition.

If a computable def. is desired, the proposed def. can be translated into the selected computable formalism

Task: ...

Diagram by Robert J. Rovetto, 2021-2022.
Contact: rrovetto@terpalum.umd.edu | https://ontospace.wordpress.com/contact

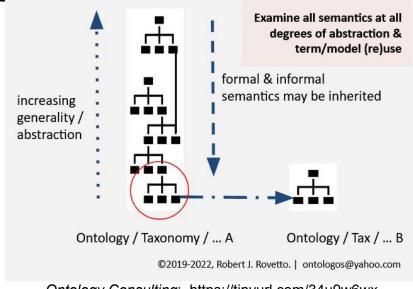
Some feedback (4)

"I'd like to make use of e.g. to align our vocabularies to, and I've always liked the facet based approach to the structure which I recommend to others when asked about constructing a controlled vocabulary."

- Identify vocabularies
- Re: alignment/harmonization/mappings ...
 - different approaches by distinct activities/groups. automated, manual,...
 - There are types & degrees of alignment/harmonization/mappings.
 - Not all groups will perform thorough harmonizations of terms or ontologies. So caution is in order (next slide).
- Explore facets...
 - See library science...faceted approaches are also used in other KOSs, e.g., taxonomies

A genote on mappings/harmonization/alignment

- Distinctions among each activity. Here taken as synonymous.
- Care is in order when doing mappings, harmonization, etc. ...
- Same terms, but different intended meaning/use?
- Ambiguous or imprecise descriptions of the harmonized terms or ontologies
- Should explicitly state degree of analysis
 & mapping/harmonization/...
- Examples
 - 'related to' as very generic
 - more specific may be needed
- Examine all degrees of abstraction of the given mapped terms/ontologies



Ontology Consulting: https://tinyurl.com/34u9w6wx

Some feedback (5)

"list datasets that are mapped to SWEET [or use SWEET on specific datasets]"

 (Idea) Create a list or catalog of datasets. identify datasets, apply SWEET, discover desiderata, ...

Candidate Development Tasks & Activities

What are some ways to do a hackathon to dev. SWEET?

General interest

Identify or propose use-cases

Identify or create applications of SWEET → apply SWEET (e.g., some in references, SWEET literature, ESIP presentations)

Explore organizational aspects: e.g., ideas for (sub)scoping of topic areas covered by SWEET, organizing SWEET individual ontologies (modules), ...

Serve as SME

Quality control of descriptions/definitions of terms (e.g., for your subject expertise), both natural language, and computable

Propose additional domain terms/vocabulary/jargon, metadata. Propose additional SWEET ontologies/modules

Develop logical axioms/restrictions/constraints, computable definitions, etc.

Documentation products: SWEET manual, ...

...?...

rrovetto@terpalum.umd.edu

Questions / Discussion

- Q&A / discussion / breakout rooms
 - What do you find interesting? Areas for innovation? Unexplored applications to other data science techniques?
 - How might SWEET help with your data goals/needs?
 - Are you interested in contributing to SWEET?
 (e.g., SME, content, use-case development, applications, misc., etc.)
- Takeaway points
- Time-permitting: Demo:
 - Open SWEET online: Github → Pylode rendering
 - Opening SWEET in...COR? ontology-editor on desktop? online visualization?
 - Audience chose which SWEET module to open
 - ...other ideas...

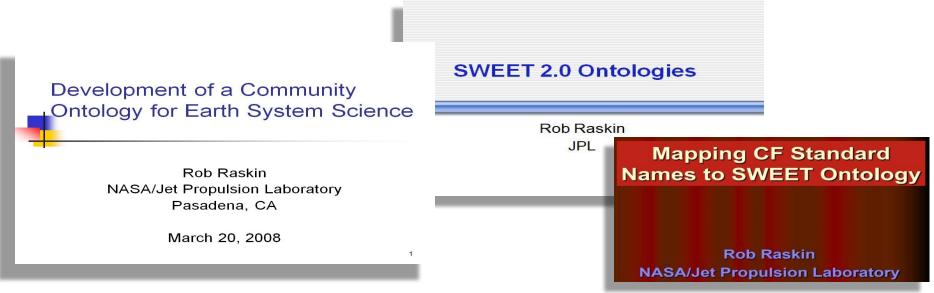
Thank you

- For you attention
- For your participation
- To ESIP staff, and fellow cluster members www.esipfed.org
- Pamela "Charley" Hayley, www.wayforagers.org

A session presentation for January 2022 ESIP Winter meeting

References (1)

Source material: Slides from past presentations by late Dr. Rob Raskin & others are raised images with shadows. Other material is quoted from cited references. Other content is by Rob Rovetto as indicated.



References (2)

Session Agenda Document: https://docs.google.com/document/d/1vRUIE3QqF68AF9JunAFj-zIY_uXoRrdroxLyZ3w9DAQ/ Session Meeting page:

- https://sched.co/gkoe
- https://2022esipjanuarymeeting.sched.com/event/gkoe/community-development-of-the-sweet-semantic-system-for-earth-and-environment-data-a-call-for-interest

SWEET Bibliography created by Robert Rovetto: https://tinyurl.com/Biblio-SWEET

- [1] "NASA and The Semantic Web", PowerPoint Presentation (PPT), Naveen Ashish, NASA Ames Research Center. https://ti.arc.nasa.gov/m/pub-archive/849h/0849%20(Ashish).pdf
- [2] "An Ontology-Driven Framework and Web Portal for Spatial Decision Support" Li, Raskin, Goodchild, Janowicz, https://onlinelibrary.wilev.com/doi/abs/10.1111/j.1467-9671.2012.01325.x
- [3] "Improving Earth Science dataset search with publications content via Knowledge Graph linkage Improving Earth Science dataset search with publications content via Knowledge Graph linkage", 2021 https://esip.figshare.com/articles/poster/Improving Earth Science dataset search with publications content via Knowledge Graph linkage/15054306
- [4] "Linking Science Data and Research to Enable Data and Knowledge Discovery" https://esip.figshare.com/articles/poster/Linking Science Data and Research to Enable Data and Knowledge Discovery/13611227
- [5] "SWEET ontology coverage for earth system sciences", DiGiuseppe, Pouchard, Noy, 2014.
- [6] "Semantic Web for Earth and Environmental Terminology (SWEET) 2018 ", PPT, Lewis McGibbney (previous ESIP SemTech Chair), 2018 https://esip.figshare.com/articles/presentation/McGibbneyL sweet 20180110 pdf/5782122
- [7] "Enabling semantic interoperability for earth science data", R.Raskin, M.Pan, C.Mattmann at 4th NASA Earth Science Technology Conference, https://www.academia.edu/912221/Enabling semantic interoperability for earth science data?auto=download
- [8] "A Survey on SWEET Ontologies and their Applications", Sumita Barahmand, Mohsen Taheryian, Saud Al-Ashri, Bhaskar Upadhya
- [9] "Development of a Community Ontology for Earth System Science", PPT, R.Raskin, 2008. https://slidetodoc.com/development-of-a-community-ontology-for-earth-system/
- [10] "SWEET 2.0 Ontologies", PPT, R.Raskin https://wiki.earthdata.nasa.gov/download/attachments/50792120/SWEET_SPG.ppt?version=1&modificationDate=1435254939370&api=v2
- [11] "Developing a Modular Hydrogeology Ontology Extending the Sweet Ontologies," Ajay Tripathi, https://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1002&context=geosciences_theses
- The SWEET 2.0 graph diagram with concentric circles is found in "Evolution in data and product management for serving operational oceanography, a GODAE feedback", and was cited as being from NASA SWEET (therefore, may be originally be from NASA JPL).
- Green arrow image from Video (see 'Other SWEET Links' slide)