SWEET governance and roadmapping working session

Host: Brandon Whitehead
20 July 2021

ESIP SUMMER MEETING

Microsoft
esri
Element 84

NASA MODIS Image of
Lake Erie, May 24, 2021

Session Format

- Arrival and introductions
- Context and session goals
- SWEET survey results
- Guest speaker!!
- Pause for questions
- Breakout groups for discussion topic 1
- Reassemble and report back to group
- Breakout groups for discussion topic 2
- Reassemble and report back to group
- Capture 3 takeaways from session
- Concluding remarks

Context

- SWEET (Semantic Web for Earth and Environmental Terminology) was initially conceived and developed internally at NASA JPL during the early part of this century
- Released to the public domain (CC0 1.0) in 2017
- Subsequently migrated to a publicly available Github repository
- Members of the ESIP STC serve as custodians, curators and developers of the resource
- The utility and future direction of the SWEET ontology is being evaluated

Goals for this session

- Primarily a brainstorming and requirements gathering session
- What goals should be on SWEET's roadmap?
    - What would enhance the utility of SWEET?
- What does a framework for SWEET community (self) governance entail?
    - What are the qualities of successful self-governed communities and/or projects?

...

- Document discussion points
- Synthesize discussion points and distribute to wider community of practice
- Actions to follow

A Synthesis of SWEET Survey Responses

Goals of survey

- Better understand how many members are in the Earth, Space, and Environmental science semantics community;
- Better understand what percentage of the community use SWEET;
- Better understand how SWEET is used within the community;
- Better understand which aspects of semantic resources the broader Earth, Space, and Environmental science community value in an effort to inform its development.

Please note

For the purposes of the survey, a semantic resource is considered to be synonymous with semantic artifact, semantic model and Knowledge Organization System (KOS), etc.; these include:

- Controlled vocabularies
- Data models
- Thesauri
- Taxonomies
- Classifications
- Property graphs
- Ontologies
- Or, any similar structure used to describe or define data or organize concepts and terminology relevant to the domain of discourse

Summary

- 18 questions total
- 7 questions on semantic resource value, assessment and (re)use
- 6 SWEET specific questions
- 5 questions with optional personal information

35 responses

Raw results collated here:
https://docs.google.com/document/d/19-9HKQATzbmG_T157OcySZyI96FQji_Bhnle8WHfZ2w

Part 1. Views on Semantic Resources

Q1: Of the semantic resources you use, which of the following do they include? (1 of 2)

Response (35)	Count	Percentage	Quartile
Persistent identifiers	28	80.0	1
Hierarchical relationships	26	74.3	1
Natural language definitions	26	74.3	1
Natural language labels	25	71.4	1
Dereferenceable URIs	24	68.6	2
Provenance	24	68.6	2
References to external vocabularies	24	68.6	2
Established governance model and/or principles	21	60.0	2

Q1: Of the semantic resources you use, which of the following do they include? (2 of 2)

Logical axioms	20	57.1	3
Specific relationships	20	57.1	3
Synonyms, antonyms, etc.	18	51.4	3
Stable releases	17	48.6	3
Multi-language support	16	45.7	4
Maintenance schedule	12	34.3	4
Other (please specify) [free text response]	5	14.3	4
Not sure	1	2.9	4
None	0	-	-
All of the above (values included in individual counts)	9	25.7	-

Q2: Which of the following do you consider most important when assessing semantic resources? (1 of 2)

Response (35)	Count	Percentage	Quartile
Natural language labels	23	65.7	1
Dereferenceable URIs	22	62.9	1
Hierarchical relationships	22	62.9	1
Persistent identifiers	22	62.9	1
Natural language definitions	21	60.0	2
Provenance	20	57.1	2
References to external vocabularies	19	54.3	2
Specific relationships	18	51.4	2

Q2: Which of the following do you consider most important when assessing semantic resources? (2 of 2)

Established governance model and/or principles	16	44.6	3
Logical axioms	15	42.9	3
Stable releases	15	42.9	3
Synonyms, antonyms, etc.	15	42.9	3
Multi-language support	13	37.1	4
Maintenance schedule	12	34.3	4
Other (please specify) [free text response]	3	8.6	4
Not sure	1	2.9	4
None	0	-	-
All of the above (values included in individual counts)	7	20.0	-

If you have re-used any semantic resources "as is" (e.g. without modification) please list them here.

- OWL / RDF / RDFS (6)
- GCMD (4)
- DCAT (4)
- NVS vocabularies (4)
- SWEET (4)
- SOSA/SSN (3)
- ENVO (2)

- OBO ontologies (2)
- DublinCore (2)
- PROV (2)
- GeoNames (2)
- DBPedia (2)

(22 responses)

Please list the semantic resources which you consider to be most important or of highest value.

- ENVO (4)
- NERC vocabs (3)
- QUDT (3)
- SSN/SOSA (3)
- SWEET (3)
- RDFS (2)
- OWL (2)
- PROV (2)

(35 responses)

# What are your use-cases for developing or utilizing semantic resources?

- Discovery (12)
- Data integration (5)
- Annotation of datasets (4)
- Interoperability support (3)
- Metadata (2)
- Automated reasoning (2)

- Access
- Clarify measurement semantics
- Data publication
- Linking research artifacts
- Persistent Identifiers
- Precise data content definitions
- Resource provenance
- Spatially enabled Env Data
- Topographic mapping

25 responses

Part 2. SWEET specific questions

Q8. Do you currently use SWEET?

35 responses

Legend:
- Yes (blue)
- No (red)

No: 77.1%
Yes: 22.9%
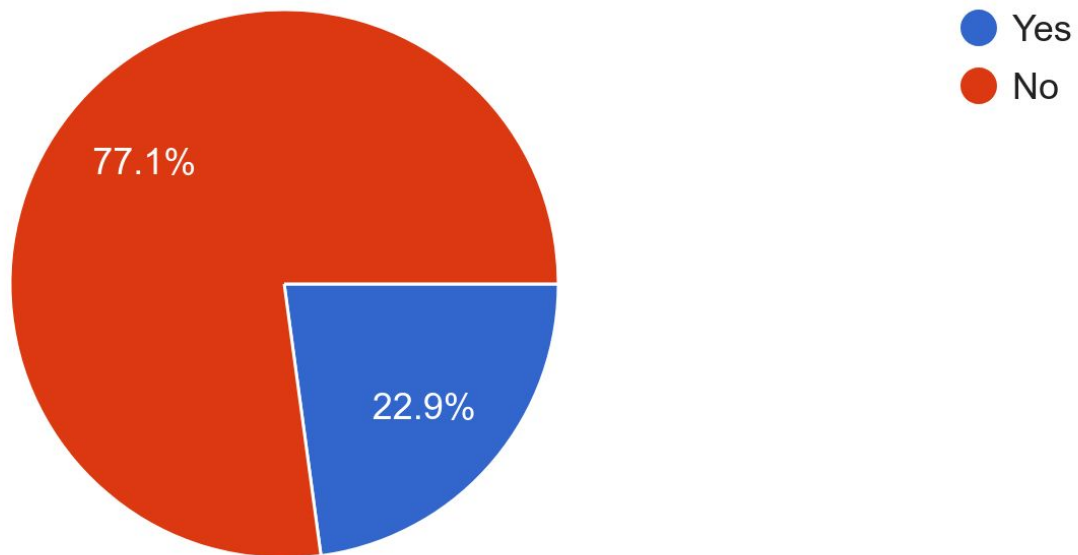
Q12: In what way do you think SWEET could be improved?

- Improving hierarchy and definitions
- SWEET needs well documented: use cases, success stories, users (banner page?), principles and guidelines
- Rigorous analysis and axiomatization of concepts and relations

- Got to get demonstrable ontological skill in there! What are some cool queries against SWEET that can be demoed? If I map data collection X to SWEET's classes, what does this give me? Can I perhaps cross-map to other data collections?

Q13: Do you have any comments about the future direction of SWEET that you would like to share with the community?

- Just make sure it aligns with other initiatives both within and outside of the Earth Science
- Need to increase visibility across federal agencies and extend the usage. Need specific maintenance team for different subjects in SWEET.
- Success stories would be helpful
- The roadmap is not explicit.

Part 3. Optional information

- Identified primarily as academic and/or research
- Identified primarily as well established in current role

Questions?

-

And now to our guest speaker...

Welcome, Bruce!

- Open Scientist Handbook: https://openscientist.pubpub.org/
- The New Media Institute: https://tnms.org
- https://cybersocialstructure.org/

Topic 1 Breakout

- 10 minutes
- -
- What goals would you like to see on the SWEET Roadmap?
    - Why is the goal important?
    - What does accomplishing the goal allow?

Topic 2 Breakout

- 15 minutes
- -
- What does a framework for SWEET community (self) governance entail?
  - What are the qualities of successful self-governed communities and/or projects?

Further points of interest

- STC calls on last Tuesday of every month @4pm EST
  - ESIP events calendar – https://www.esipfed.org/get-involved/telecon-calendar

- Initiating SWEET working sessions following this Summer Meeting

- SeREEn Federation (Semantic Resources for Earth and Environment)
  - Project in its infancy – repo here: https://github.com/ESIPFed/SeREEn
  - Currently part of ESIP STC
  - Calls are second Tuesday of every month @4pm EST