

Community Ontology Repository Working Sessions

(Also see: [Semantic Technology Committee business notes](#))

Community Ontology Repository Working Sessions	1
(Also see: Semantic Technology Committee business notes)	1
2019 Community Ontology Repository Working Session	1
Attendees	1
Agenda (proposed)	2
Notes	2
Pre-agenda discussion	2
Show how COR can be used to achieve use cases	3
Needs from COR	3
Post-meeting	5
2018 Community Ontology Repository Working Session	6
Attendees	6
Some Links	6
Discussion	6

2019-10-01

Attendees

Annie Burgess, John Graybeal, Carlos Rueda, Lewis McGibbney

Apologies

Agenda

1. Add Ruth's use cases to
<https://esipfed.github.io/stc/UseCases/STCUseCasesAndRequirements.html#UseCases>
[Ruth,Lewis]
2. Add Blake Regalias use case to
<https://esipfed.github.io/stc/UseCases/STCUseCasesAndRequirements.html#UseCases>

3. COR: Terms of Use and enforcement of same (if people put non-ESIPy vocabs on COR, what is our process?) [John]
4. COR: Latest status from Annie re CloudFlare [Annie]
5. COR: How we might approach the need for good metadata (MOD, Widoco, other...) [John]
6. COR: How we might approach better presentation of ontologies (Widoco?) [Lewis]
7. A pre-meeting workshop on FAIR Semantics and FAIR Repositories, at which I'll give a 10-minute ontology repository roundup [John]
8. using supervisor (<http://supervisord.org>) on COR

Minutes

- On #4
 - we are having issues communicating with David Basendyne on generic COR issues. A better way for us to communicate with him would be sending specific bug reports.
 - A valid solution for us working this one is simply for us to take COR off of Cloudflare. There should not be any issue with moving COR back to Network Solutions (which is what the DNS provider was previously before moved to Cloudflare).
- On #5 & #6
 - Action item on John to provide more information on MOD for us to evaluate moving forward.
- On #7
 - Information only
- On #8
 - Check action item for Lewis

Action Items

- ~~Lewis will add use cases for both Ruth and Blake to <https://esipfed.github.io/stc/UseCases/STCUseCasesAndRequirements.html#UseCases>~~
- John to send group ToU examples for review - <https://github.com/mmisw/orr-portal/issues/123>
- ~~Annie to send information to David Basendyne regarding reverting COR to use Network Solutions DNS resolution.~~
- John to provide more information on MOD
- Lewis to create a wiki document regarding use of health checking and service provisioning solution.
- ~~Lewis to reach out to Sky Bristol, USGS~~

2019 Community Ontology Repository Working Session

Friday, July 19 • 10:00am - 11:30am

Attendees

- Kai Blumberg, Developer on Environment Ontology
- Annie Burgess, ESIP
- John Graybeal, Stanford Univ.
- Lewis McGibbney, JPL
- Beth Huffer, Lingua Logica,
- Mark Schildhauer, UC Santa Barbara
- Simon Cox, CSIRO
- Sky Brystol
- Abdullah Alowairdhi
- (remote - late) Brandon Whitehead, popokotea
- (remote) Carlos Rueda, MBARI

Agenda (proposed)

- Show how COR can be used to achieve use cases
 - What is our idealized use case?
- Solicit inputs for needs
- Show details of running and operating COR

Notes

Pre-agenda discussion

The meeting started with conversation about COR's status and objectives. Beth asks can we show that our use cases are being met? Example: there are use cases for tagging, how will these be met? Do we know what people are using the system.

Resources like ENVO are used. One good outcome is to have something parallel to OBO, within ESIP or earth sciences, so that the earth science community can use it.

Yes people are using the system, esp. for SWEET

What is objective for COR? It would be good to talk about that. Are we saying OBO is already doing what COR is doing? No, it's biological, it would be nice to have a repository in earth science that is analogous to OBO. If COR functions like OBO in providing a set of ontologies. Should we try to work together in some way, or is that not necessary? *PLB (post-meeting)*: This

is needed. Otherwise people will start spinning up redundant efforts and not talk to each other. Federated search across portals needed.

Who are the people who are using COR? We don't really know, any more than NASA knows. We don't need to know about 'who' so much as we need to provide easy access.

Discussion of whether COR is SWEET. It is not, SWEET is just a major user.

We haven't done a lot of work on COR, this is not strictly a bad thing—the system has worked well without undue attention. If we look at the requirements we can see it satisfies almost all those requirements. Does it satisfy the use cases about needs? We need to look at those and see if they increase the requirements on COR. If there's something COR can do to make him more enthusiastic about dereferencing ontology terms, it would be some handy utilities.

If we can convince our sponsors that this has value for them, it would be a good thing. Doesn't think anyone even knows about it. Can take these use cases back to NASA, it would help introduce it.

Need to invite key folks like Beth and Mark to next meeting and get your inputs on use cases, just in a conversation. (Beth and Mark concerns about COR providing better UX for *science researchers* dereferencing a COR-housed term. Is primarily ontologists using COR now)

COR is not a finished product, still needs development. It isn't like we have a group of developers all working on COR. If we had end user engagement we might get more developers.

Show how COR can be used to achieve use cases

Toolkit developed on how to use software with COR (PyPI tool kit). Carlos and Lewis put together demo showing how data can be pulled from COR with minimal Python code, using an existing Python library in GitHub repository.

Lewis demonstrated a message pub/sub application using COR to demonstrate how events and other messages can be published to the repository and read by other user, prompting questions and discussion. Is it that different from GitHub messaging? No, not that different. Is this an expected/accepted use of COR? John has some concerns about what will happen if it becomes heavily used, in terms of the impact on users of COR and their impressions of the COR system, but as a demo illustrates advantages of LinkedData as a data publication/exchange technology. In discussion, Lewis agreed it is not intended as a production system, that for production purposes you might well choose other purpose-built infrastructures. But what this shows is that you can easily build on an infrastructure like COR.

Needs from COR

Moving to a discussion of COR's goals and needs to go forward.

1. Is the vision of COR to service an end user research scientist client?
 - a. If you go to the EMBL-OLS it is a very user friendly and useful representation of a term that also allows exploring related terms. Similarly with ONTOBEE. In COR we don't see any of that.
 - b. The scientist/researchers need some accessible representation of an ontological term to better understand the value of semantics. Is the vision of COR to engage with the research scientists in this way (e.g. that those scientists dereference their terms through COR)? Some of us think that's one purpose, yet.
2. Carlos suggests we look at <https://mmisw.org/cfsn> as simple more human-useful UI of a term
 - a. This illustrates how the back end serves as a foundation for building services. Is an example of client UI that exposes a particular and useful service, powered by the same backend that can power a variety of other clients.
 - b. It could be re-purposed for integration into COR as well.
3. Is COR going to be an ontology repository for named-graphs (triple stores).
 - a. Yes, it is, John says definitively.
4. We don't want to close off investigatory uses like the one Lewis presented.
 - a. This is an operational tool but also an exploratory tool supporting research use.
 - b. If something becomes promising, maybe it needs to spin off into a separate production system so as not to interfere with the production purpose of COR.
5. COR is an enabler. We should focus on the API to enable tools and APIs to use it.
 - a. E.g. API-enabled tagging, cross-referencing, etc.
 - b. Leverage what is already happening
6. You're not going to attract more users by (only) working on the back end
 - a. There's a critical need for at least some more researcher-useful representation of the term when a SWEET URI (or any URI) is dereferenced
7. Both backend and frontend are super important indeed.
8. There needs to be a more concerted effort related to the requirements and use cases, we need those.
 - a. Lewis reviewed the requirements in the original evaluation document, put together by members of the user community.
 - i. All these requirements were met (except one one only partially so far).
 - ii. What we really need is use cases that real people need.
 - b. In fact a number of use cases already have been written, and are in the same document. COR already supports many of these as well. In response to a concern that the use cases need to be connected to the document, Lewis reviewed the connections already in the document.
 - c. When we say "what is the purpose of COR", let's not forget that it can and should have many purposes. Part of the value of a semantic solution like COR is that it can address many different use cases.
9. Figuring out pain points is still needed. What are the real problems that need to be solved?

- a. Lewis suggests that we need to find the users who can describe for us what they are looking for. Invite them one or two at a time to COR meetings over the coming months, to give us additional direction.
 - b. A key concern is that we are not funded to actually meet these needs. There are no front-end developers waiting around for instruction, so let's not get hopes up about implementing needed improvements.
 - c. Back-and-forth ensued about what comes first, making it more user-friendly, advertising it more, or getting more contributions?
 - i. They all depend on each other.
 - ii. See also comment about 'content' (after meeting closed)
 - d. We need our code to go to a common place where the community can consult and add our own knowledge
 - e. We need this to be a community asset
10. We need a crystal clear answer about 'why are semantics important?'
- a. Any utilities that users can provide that use this asset will help demonstrate that.
 - b. There is a useful 2015 publication in this regard:
<https://doi.org/10.1609/aimag.v36i1.2560> ("Why the Data Train Needs Semantic Rails")
11. Sky Brystol reports his experience as data system developer: COR serves as a critical repository to store his semantic assets where the community can access them
- a. He thinks the access points for developers (e.g., the backend) are powerful
 - b. Several comments to the effect COR needs to continue appealing to data managers as a priority
 - i. Thinking is that the Earth scientists are individual users, but not so many of them need to use COR right now to resolve terms
 - ii. *(post-meeting)* Suggestion for a tool: a widget like BloPortal's that could be inserted into code or web sites, lets end users find/select a term to be used in their metadata (see John G for specs here :->)
12. This seemed to close the circle in the discussion on the different needs for the system.
(See also follow-on comments below.)

Post-meeting

Quick hits after the meeting officially closed (but many people still there):

- 12. There is something else that makes a repository much more important: content.
 - 1. Toward that end, Nonong suggests bringing some data over from MMI ORR (<https://mmisw.org/ont>), where it is not as readily discovered.
 - 2. Several agree this would be valuable, and is timely.
 - 3. Will have to consider how to best do this.
- 13. Note that COR was in fact built as a bridge for simple users to submit vocabularies easily and do mappings easily.

- a. Not necessarily as a high-end feature-packed tool, but basic capabilities to capture content
- b. This could be leveraged to build awareness and content for the system
- c. Some of these 'easy' things apparently need to be advertised more

2018 Community Ontology Repository Working Session

Friday 9:30-11A

Attendees

(Huffer, Burgess, Graybeal, Aulenbach, Portelli, Richard, O'Brien, Schildhauer, Lafia, Narock, (a couple other folks); Remote: Rueda)

Some Links

- <http://cor.esipfed.org> — Entry page for COR
- <http://cor.esipfed.org/ont> — Ontology browse home page for COR
- <https://esipfed.github.io/cor/> — Overview page for COR
- <https://github.com/ESIPFed/cor> — Github pages for COR **deployment**
- <https://github.com/ESIPFed/cor/issues> — Github pages for COR **deployment** issues
- <https://github.com/mmisw> — Github pages for MMI ORR **software**
- <https://github.com/mmisw/orr-portal/issues> — GitHub pages for MMI ORR **software** issues

Discussion

If you turn these into issues, you can sign off the ticket at <https://github.com/ESIPFed/cor/issues/48>

- Should be able to expand columns on main view
- Search for 'water temperature' has three blank nodes at the beginning (is disconcerting for the average user)
- Lindsay Portelli expresses some (highly preliminary) interest in supporting system (possibly a developer-user)
- <http://www.ontobee.org/> style has nice term resolution and presentation useful for resolving GUID to get an idea of term meaning; compare with COR interface. EBI OLS has nice features to track term edits (mentioned via comment).
- COR as a place to submit ontologies vs. COR as a place to search ontologies
 - John notes a good tool should be able to do both.
- Getting back to the SWEET/ENVO alignment is important and useful
- Semantic Tech should do some curation on ontos, to prevent sloppy redundant knowledge modeling
- EarthCube is going down the road of resource curation
- Would be nice if there is a bulk process to add to COR
 - API development

- COR as a place for discovery of existing (curated) resources
- The “Owner”, if organizational, should be resolved to an authority that vouches for the good quality of the ontology. (SWEET’s lack of definitions, for example, unqualifying it from full endorsement.)
- “ESIP expertise that creates consistency and leads to continuity”
- Are there libraries that can be used with COR?
 - Yes, see <https://github.com/ESIPFed/corpy> (in progress); let Carlos know of others/interest
- GUI for building an ontology (Sara Lafia)
 - Interactive tool to explore ontology development
- Try to agree upon some kinds of standards for ontology that aren’t Earth-science related (Lindsay Portelli)
 - SWEET has some of them, but we should leverage work from other communities.
 - common core (from OBO) for people, time very helpful, but are there others?
- Love to see filters (mature ontologies, complete) (OWL or SKOS)
 - HOW do we represent that across ontologies, with what metrics or concepts is it established?
 - There are ways of doing this—BioPortal has some useful metrics, AgroPortal has some new categories. Certainly COR and BioPortal would both benefit from more user-friendly expressions of ontology quality