

## Semantic Technologies Committee Business and Planning Meeting

July 22, 2016

### Ontology Portals

- need to select one online repository
- Does this group feel that an evaluation of different features is important?
  - We need to publicize
  - We need to separate use cases, functionality
  - Continued evaluation
    - When we talk about reevaluation and continuing it – is it formalized?
- Wiki page was set up for ideas but never fleshed out
- If we should go forward with one portal vs two portals. Don't have enough information on what is necessary in order for a positive evaluation
- This is a good test case for what this would look like – ESIP leadership is behind it
  - Who would oversee it is still to be worked out
- Are we talking about multiple servers for the different portals? – how much does this cost?
  - This is an issue and constraint
- Line P:
  - Part of what we are asked to do since June is to deploy something on the ESIP TestBed
  - We just moved over to the ESIP Cloud Structure to see if we could provide this as a test bed
- Thinks ESIP evaluation is immature. The portal has been available for two years and there are no users.
  - The only evaluator has been Lewis. Only external user, as well
- Ontologies are incredibly useful, and somehow we need to disseminate that viewpoint and enforce that it is important.
  - If people were creating a small ontology or reusing others. They would need a portal
  - In the meantime, ESIP has solved a problem by providing a stable home for these services to live
- 2 years and not much action. How much energy are we going to put into this?
  - Concerned whether one or even two are necessary
- If you look at the wiki document, it needs to be populated: only way to do it in a meaningful way is from users
- Can people utilize - people will have confidence.
- The document that we have currently is too verbose and that is a barrier.
  - Do we have these collated?
  - We only need about 5 to go forward. Evaluation is going to take a bit of time. If people are happy putting their resources in more than one place. It shouldn't be too much of a problem to put it in more than one place.

- We want to know the usage of that API as one way to see how much it is being used
- Are there people who are trying to use the system and getting frustrated? This could be an indicator.
  - Really based in uncertainty. Even though ESIP was interested, we didn't have a clear message that this was the place within the community.
  - 2 years is a relatively short amount of time for infrastructure. Sometimes it takes longer to get into people's heads.
- We may need to have a more formal evaluation process. People who have expressed interest in using semantics. We provide the resource, ask them to try it, and then see how they feel about the resource.
  - We have a hard time identifying what success looks like. If we have something more specific (Usability Study).
  - People may not know how to use the resource, and we can provide the resource, and then ask them how that works and if they would use it again. Provide a clear pathway.
  - There are so many usability issues that don't get addressed. We can come up with hypotheticals. Want real user studies.
- Important component: governance. Who can upload stuff? Who can put up new ontologies? What is accepted/not accepted?
  - You can identify an authority for that ontology. This would be registered. People could use it, but change management would be in the hands of the owner.
  - That partially answers the question, but a lot of specifics to be addressed.
- Anyone who gets a login and API key can upload. Minimal governance.
  - Technical requirements to be met, but then could fairly accessible to users
- The Bioportal system is completely open. Chose a UN and PW, when you upload the ontology, it just has to be in OWL, RDF, or CSS. No approval process, no vetting. Everything is wide open.
- Set of criteria for evaluation? We don't have a concrete plan for evaluation.
  - Maybe spend some time looking at evaluation.
- We have some experience in the Ontology Design Portal. One way they wanted to do the eval was to have a set of people to give comments and reviews. In the beginning it worked, but towards the end it fizzed out. They didn't follow through with the review. We have to ask them to give their time, which is very hard. We have now asked them to set aside dedicated time.
- We are talking about elements best addressed in a roadmap. Governance, etc. need to be spelled out, which does not exist at this point in time. Having looked at the medical/bioinformatics community, they don't have a way to deal with semantics.
- Identifying the scope is the first step. If you don't have an end-goal, then you won't be able to identify the level that you want. You are in prototype efforts now but want to grow. That is where the work plan is most important.
- **Action Item:** work plan/road map
  - Gather up the use cases and start with that. Document them between now and the next meeting. Get some scenarios for how to use the portal. Once we start

documenting them, then we could have a portal usage session at the Winter Meeting. Walk them through the scenarios and how they are of most help. Reach out to some people to understand their needs. Then at the next meeting we will be at a position for a formal evaluation.

- There is an interest from the community to have a centralized resource. Our counter message: we are in the evaluation process. We are not there yet and need more time. Not even in a position to agree that it needs to be a centralized resource. Not there yet.
  - Does it need to be a centralized place?
- As a committee, we need a common message. We need a common end-goal.
- Next telecon: 3rd week of August.
- Regarding coming up with concrete use cases as requirements: can we talk about why the semantics committee got created in the first place? Seems like there must have been some people who wanted it.
  - Use case that seems common: Earth Scientist wants to gather data. Working on seawater salinity. What should he name the table in his database? Should he pick his own or use somewhat common ones in his domain? To find those, he should go to an ontology for seawater and find properties for standards. Still believe there should be a portal to find those. There may be properties not recorded before. Then you can augment the ontology. This makes your own data more valuable since it can more easily be linked and reused in the future.
- There is a community that believes this is important, or the committee wouldn't have been created. It isn't hard to understand why both these two portals are important and are helpful to the community. How do we get everyone in a good place where needs are met and are actually using the systems that are there.
- Idea: We aren't sure how the ontologies are used. We want them used by earth scientists. Good for the committee to connect with other groups and data managers and will have use cases for us.
- The spectrum of skill level and needs is much broader than initially thought. We need to have a more concerted effort to describe what the solution is.
- The ontology portal is not why we became a committee. One component of a wide variety of things. It sometimes sounds like our sole reason for being is to host an ontology portal.
- The earth science community has interest in semantics. Unlike health care, we don't have a dedicated focus in the same way. There isn't a place to go to understand the best practices and their steps forward. We need a community consensus on a roadmap. What has been done, what needs to be done, where are we ultimately headed. ESIP is a good place to bring that together. Need to put it out as a formal, published document.
- Why we need a road map: Struggled with coming up with a unified approach to approaching semantics. Lots of different people doing research. What is lacking: haven't come across the state of the art for semantics in earth science today.
  - We need a monograph/position paper. Open source type paper. Something people can have right away. What does the paper contain? Put together an outline as to what the paper should look like.

- International Association for Ontology and its Applications: very applied.
  - Need to come to consensus between North America and Europe. So as to capture state of the art as it is in 2016. Need an articulation.
- All the projects are deep data - If we could use successful projects that have been developed instead of new case studies. Follow those methods in order to use them for the roadmap. Using successful projects and their methodologies as a starting point, and then having discussions.
- People using PROV and Provenance of datasets
  - Working Group for PROV - how to use this successfully?
- Semantics Technologies Committee is strong. The roadmap is not just for an ontology portal, but larger than that. There is a lot of ground to cover there.
- Transdisciplinary - using semantics. Leveraging concepts and philosophy. Applied is different from philosophy. Part of the roadmap should confront this.

Ontologies this committee could publish to increase data interoperability - save for next telecon

## **SWEET**

- The goal for sustainability for this is to have it contributed to by people. Want to have the development to occur under ESIP. Right now there is no real contribution model. ESIP Fed GitHub account could be the host. Those who want to contribute could do so via the github account. Need the process to be finished first.
- Part of the recommendation for posting spatial data on the web: reference URIs
  - Provides opportunity of a practical opportunity for application and working out governance models
  - This is a knowledge resource being developed and this is the community for that
  - ESIP good on taking over the resource.
    - There may be more than just getting people to take it over. There should also be a review process. Suspect that it will be work. Maybe someone needs to be responsible for it. Someone needs to review the changes. A lot could go wrong. It will be work.
      - Big difference between governance to serve and then governance to review. Requires a dedicated leader to the task to the latter. If we are committed to this 'management' role, we need to identify that leader.
- Consistent question: review process.
- Members of the W3C could use them as an example of practice. These resources are very well documented. There will be a number of people at AGU that will be presenting on this. Would like to see other come forward to work on code management. The more the merrier. Having it as an agenda item consistently would be helpful.
-

### **Common Concepts**

- Mainly will talk about this at August meeting
- Idea of creating 'common core' a la Dublin Core
- Can promote it as an ESIP resource, people could do a horizontal integration.
-