

# 2020 - 2021 SWWG Meetings

August 5, 2021 (Yes, it's the first week, this month.)

Topic:

We will continue learning to semantically annotate a model, using the ARIES system. At the August meeting, we will develop talking points and key questions in preparation for a productive discussion with subject matter experts. We hope to recruit an experienced knowledge engineer to guide this session.

July 8, 2021

Topic:

On our June call, we agreed that a useful path forward as a group would be to work on making one of the simpler models in the USGS Model Catalog semantically interoperable in ARIES. Leslie guided us toward a simple liquefaction model used by USGS that is described in the attached paper.

So, as optional homework for our next meeting on July 8, those interested could:

1. Read the attached article
2. Think about how you'd semantically represent the following concepts in your own work. I've attached a small slide deck that describes the semantic annotation process in ARIES; the flowchart I showed in the last meeting is on slide 19. I'm particularly interested in the first four, as we've never connected to ontologies or controlled vocabularies related to earthquakes in ARIES in the past. If anyone knows of good earthquake/hazard semantic resources those would be very helpful to bring to the table. Here's the list of the concepts:
  - a. Probability of soil liquefaction
  - b. Soil liquefaction susceptibility
  - c. Peak ground velocity (cm/s)
  - d. Shear wave velocity over the first 30 m (m/s)
  - e. Distance to coast
  - f. Distance to rivers
  - g. Distance to water bodies
  - h. Water table depth (m)

On the call, we can discuss potential benefits of semantic annotation and a process for annotating the needed concepts and making the data and models interoperable within ARIES.

Cheers,  
Ken

Links to attachments: [article slides](#)

Links shared during the meeting:

Does not have all the terms that we are looking for, but [https://esipfed.github.io/stc/sweet\\_lode/phenGeolSeismicity.html](https://esipfed.github.io/stc/sweet_lode/phenGeolSeismicity.html)

also, not as formal, but a trusted source, [https://www.iris.edu/hq/inclass/fact-sheet/vocabulary\\_for\\_earthquakerelated\\_topics](https://www.iris.edu/hq/inclass/fact-sheet/vocabulary_for_earthquakerelated_topics)

and the ESIP ontology portal probably has more: <http://cor.esipfed.org/>

Not sure but possibly their (USGS) code repo would have some useful links to data (in addition to talking to subject matter experts) <https://github.com/usgs/groundfailure>

A high-level technical description of k.LAB for technical partners <https://docs.integratedmodelling.org/technote/>

About use of ARIES <https://www.biorxiv.org/content/10.1101/2021.02.23.432363v1.abstract>

June 10, 2021

Topic:

## After the CDI Workshop: What next?

It was good to see many of you at our Semantic Web 101 session at the CDI Workshop.

This week, let's talk about next steps.

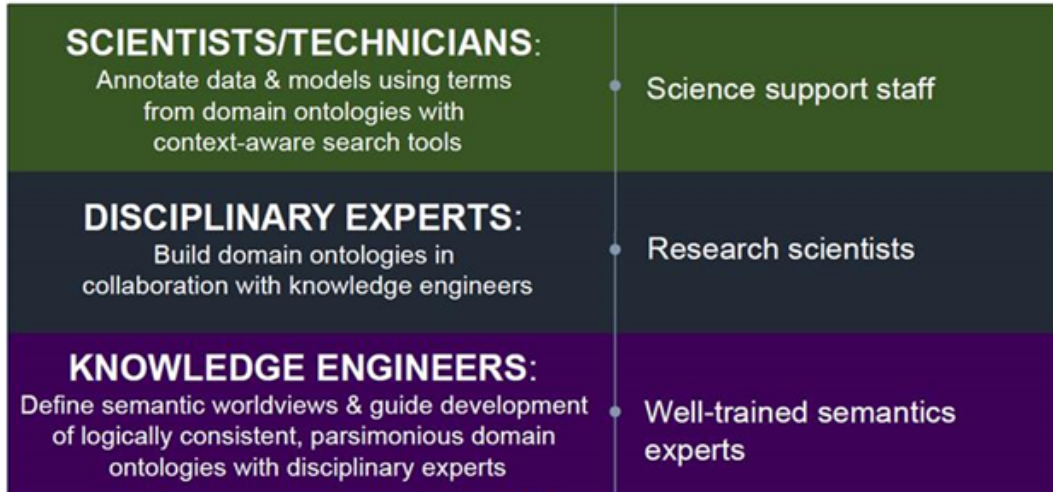
Sponsor training on semantic methods and resources? Spin off a project or two, perhaps for a CDI proposal? Investigate participation in ESIP collaboration areas?

What do you think would be fun and productive?

## Notes:

The Semantic Web 101 session at the workshop was great. Thank you to all who contributed.

We discussed the different roles of USGS staff in producing data and models that are semantically ready. Ken contributed this model:



We would like to provide learning opportunities so that USGS staff can use semantics in their own work, recognize semantic capabilities that are increasingly present in the software we use every day, and recognize the need for knowledge engineering expertise.

We decided to do a small project together, to semantically annotate a very simple model. Leslie will help us identify the model, from her work with the USGS Model Catalog. Ken will provide the ARIES flowchart to guide our work. On July 8 we will look at the model and the flowchart and decide how to go forward together. This might be the start of a larger model to be proposed later.

## March 11, 2021

Topic:

### **Our session at the 2021 CDI Workshop**

We hope that session participants will come to understand

what semantic web is and why it would be useful for our work in USGS. We hope they also gain an idea what would be involved in actually using semantic web. We hope the working group will be re-invigorated as a result of working together on the session and attracting new members.

We're thinking of a series of presentations with time for questions and discussion.

- Ken offered to speak about the ARIES project.
- We would like an update from Scott Peckham, who spoke at the 2019 CDI Workshop about interdisciplinary work.
- Beth Huffer was suggested as a speaker.
- Fran offered to find a speaker who could demonstrate how semantic web is used for FAIR data.
- We need a foundational presentation about the basics of semantic web.

## February 11, 2021

Topic:

**USGS needs a knowledge graph capability for using [schema.org](https://schema.org) vocabularies. How can we do that?**

## November 12, 2020

Topic:

**But what is a model? Semantics in the USGS Model Catalog.**

Presented by the USGS Model Catalog team

A new catalog for USGS Scientific Models is in early stages of development. How can this tool use semantics to make sure its users find what they want? We'll lead a presentation and discussion on our plan for what terms are used, how these terms are applied, and how to keep moving forward. The team is very interested in gathering ideas for how to improve the utility of the catalog.

## Outline

1. What is the model catalog (<https://data.usgs.gov/modelcatalog/>)
2. What the role of semantic web in the catalog... (filters, tags, browse topics)
3. Our plans for gathering and using information from other systems for the catalog
4. What are some other possibilities? (think big)

## October 8, 2020

Topic: A Geospatial Knowledge Graph for National Topographic Data, presented by Dalia Varanka.

This month we will have a Microsoft Teams meeting. Please use the link on the calendar invitation.

Here are [Dalia's slides](#).

## October 8, 2020

Topic: A Geospatial Knowledge Graph for National Topographic Data, presented by Dalia Varanka.

This month we will have a Microsoft Teams meeting. Please use the link on the calendar invitation.

Here are [Dalia's slides](#).

## June 11, 2020

Topic: Discuss a [Forbes article](#), "The Semantic Zoo – Smart Data Hubs, Knowledge Graphs and Data Catalogs." This is a basic introduction to some relatively new Semantic Web technologies. Let's look at it together, understand what the new options are, and consider whether our work could benefit from using some of them. Here's a [PDF](#) of the article.

This month we will have a Microsoft Teams meeting.

## May 14, 2020

Topic: Presentation by Brian Wee about an ESIP project to use concept maps.

This month we will use the ESIP GoToMeeting account.

Abstract:

The Federation of Earth Science Information Partners (ESIP) Agriculture and Climate Cluster (ACC: <https://tinyurl.com/esipagclimate>) recently initiated an experiment to use concept maps for documenting science-informed, data-driven workflows for climate-related adaptation, mitigation, and response planning (<https://tinyurl.com/cy2020-ag-climate>). These types of planning activities are inherently transdisciplinary in nature, involving individuals from multiple organizations, scientists from different disciplines, and policies promulgated at various jurisdictional levels. Concept maps are ideal for capturing transdisciplinary processes that unfold across a complex suite of interdependent actors. We intend for these concept maps to be repurposable for self-paced education, transdisciplinary engagement, machine-assisted "context-aware" knowledge discovery, and machine-assisted planning.

The latter two objectives are facilitated by machine-readable concept maps. A number of ACC members were involved in an "ESIP Lab" project to explore these ideas (<https://tinyurl.com/esiplabreportv4>). Ultimately, we envision repurposable agroecosystem best management practices, traceable to data and models, for empowering "a cadre of transdisciplinary resilience engineers to adopt and continuously develop climate adaptation solutions" (<https://tinyurl.com/resiliencegenome>).

The 2020-05-14 presentation is publicly available at: <https://tinyurl.com/20200514-d2d-conceptmap>

Recording: <https://www.youtube.com/watch?v=uJ0k33H-2BY>

## April 9, 2020

Topic: Part of enabling USGS products to be FAIR (Findable, Accessible, Interoperable, and Reusable) will be providing vocabularies that are also FAIR that can be used in metadata and data dictionaries. That gets to the heart of semantic web. Join us this month to discuss "[Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web](#)" by Daniel Garijo and Mar'a Poveda-Villal'on. We hope to learn together about semantic web and maybe develop some suggestions for USGS progress toward enabling FAIR.

Discussion:

We began with the observation that accessible and interoperable vocabularies are essential to successful implementation of FAIR. This is recognized by one of the original FAIR principles.

I2: (Meta)data use vocabularies that follow the FAIR principles.

Our discussion focused on sections 2 and 3 of the paper, URIs and Documentation.

An additional metadata property, for describing thesauri that will be used to assist in finding data and metadata, is the structure of the thesaurus as a whole, for example, hierarchical.

Permanent URIs that resolve to useful information about vocabularies and terms are essential for interoperability of data and metadata, as well as linking terms in different vocabularies. To extend the usefulness of USGS vocabularies for interoperability we will need regularity, stability, and accessibility. This is not a problem within the scope of USGS "ontology engineers" (the target audience of the paper we're discussing) but needs the attention of those who set policy for the USGS network. **We will call this problem to the attention of the team that we expect to be set up to coordinate enabling FAIR in USGS.**

Future discussions and action on the topic of FAIR vocabularies would benefit from focusing on use cases. Possible cases to consider: the use of term lists in SharePoint and Drupal; semantic web techniques to improve discovery of items in long lists like the open innovation opportunities relevant to homeschooling and COVID-19.

## March 12, 2020

Topic: A practical example of semantic technology in action: assessing the status of biodiversity in the world's oceans

Sky Bristol (USGS Core Science Systems, <https://orcid.org/0000-0003-1682-4031>) presented a big use case, in which semantic standards are needed to enable multiple large datasets of ocean biological and ecological observations need to be integrated in order to answer important questions about the effects of human activities on ocean ecosystems, as well as the sustainability of human uses of ocean resources. Several groups are working on ontologies that provide standard terms for the biota, ecosystem components, and relationships. A next step will be normalizing all the data with ontologies. This is an opportunity to assist with real world semantic web work. [Slides](#).

## February 13, 2020

Our meeting will be cancelled because of a conflict with the Director's Town Hall about the budget. The following discussion will be rescheduled, perhaps for March 12.

Topic: discussion with Sky Bristol. Since his work with SemantEco, Sky has been working on:

The problem of semantic harmonization across biological observation datasets flowing into OBIS, now leading to an ontology to enable integration toward the [Essential Ocean Variables for Biodiversity](#);

Decision ontologies that explicitly identify key decisions by regulatory and conservation management groups and then knowledge graph the scientific inputs to those, which ties to the Global Change Information System model for linking findings to evidence;

The management team for EarthMAP, working to figure out what all that really means and what we are going to do about it in the near term and over the next 30 years. Sky's question for us: "What all are you working on that we can tie into the endeavor?"

## January 9, 2020

Topic: Look together at the use of semantic technology in "SemantEco: a semantically powered modular architecture for integrating distributed environmental and ecological data"

<https://doi.org/10.1016/j.future.2013.09.017>, by Evan W. Patton, Patrice Seyed, Ping Wang, Linyun Fu, F. Joshua Dein, R. Sky Bristol, and Deborah L. McGuinness. (See also <https://pubs.er.usgs.gov/publication/70073700> and [this link](#) – thank you, Dalia, for getting us this better copy of the article.)

Notes:

The paper seems interesting, but we have lots of questions. We are inviting Sky Bristol to join us at a future meeting to address these questions:

- Is there anything in SemantEco that is transferable to EarthMap?
- Did anything come of the hope to create a reusable ecosystem of semantic tools for integrating data and regulations?
- Has anybody implemented SemantEco in a live, online system that we can see?

The data system seems useful for applications beyond ecosystems, for example mining industries.

We also need to learn more about the basics of semantic web. We exchanged some citations of resources, and Dalia volunteered to lead a future meeting with a short summary of semantic techniquet that her program uses. We are also starting a channel at the CDI Microsoft Team site to exchange information, questions, answers.

Here are some of the resources that were suggested:

On graph databases:

- <https://neo4j.com/blog/why-database-query-language-matters/>
- <https://towardsdatascience.com/an-introduction-to-graph-databases-cd81a0d5aa12>

For ontologies and reasoning:

- Semantic Web for the Working Ontologist (<https://www.elsevier.com/books/semantic-web-for-the-working-ontologist/alleman/978-0-12-385965-5>)
- Demystifying OWL for the Enterprise ([https://www.morganclaypoolpublishers.com/catalog\\_Orig/product\\_info.php?products\\_id=1252](https://www.morganclaypoolpublishers.com/catalog_Orig/product_info.php?products_id=1252))

