

USGS integrated modeling on the semantic web

Hi folks –

We're looking to submit a CDI proposal that would provide proof-of-concept for a USGS integrated predictive modeling infrastructure on the semantic web. This is a big lift, but we've already built much of the infrastructure (see references below). We believe it's possible to build a relatively simple proof of concept that could later be moved to production scale within a larger effort. The approach has several key advantages – the semantic underpinnings to connect diverse scientific disciplines, ability to connect models of different paradigms (including data-driven approaches), and to better navigate scale differences.

We're looking for partners interested in building that proof of concept. Key requirements are 1) that we address a multidisciplinary challenge (which will immediately challenge us to build ontologies for semantic annotation of data and models that cross disciplines – a challenge that existing domain ontologies struggle with, but that we're equipped for), and 2) that the initial models be relatively simple (to be tractable within the limited scope of a CDI proposal).

If you're interested in partnering on or have good ideas on an integrative science case study, or have any questions about our proposed approach, please let me know!

Proposed steps in CDI proposal:

1. Generate common semantics for a small, multidisciplinary integrative modeling problem, via a 1-week semantics workshop.
2. Serve & semantically annotate key data resources on the web (i.e., Semantic USGS Science Data Catalog, e.g., data already hosted on sciencebase.gov/geoserver). This sets the stage for wider use of semantic interoperability in ScienceBase, and a framework for other data repositories to follow to allow their products to work together.
3. Code/wrap key models and contextualization information in the k.LAB semantic modeling environment, enabling their use within a linked data & modeling environment.
4. Provide a few model runs that illustrate the results.

Later steps toward production-level (not part of this CDI proposal):

1. Model calibration/testing & validation of predictive power
2. Production version of above to address additional integrative science challenges (building new cross-disciplinary ontologies and models)
3. Setup semantic modeling node to enable full control of confidential/not-yet-released data
4. Web explorer for key products – Decision Support System-style "front-end" for nontechnical users
5. Semantic annotation navigator: Design a metadata editor plugin that enables semantic annotation of new products in ScienceBase by technical support staff (i.e., GS-9 level).

References

<http://www.integratedmodelling.org>

Villa et al. 2017. *F1000Research* 6:686.

Villa et al. 2014. *PLOS ONE* 9(3):e91001.