Mini-series Wrap-up

Earth Science Ontology Dialog" mini-series was designed to explore the current status and application of multi-level ontologies towards developing a semantically enabled cyberinfrastructure for the Earth Science Community. In addition, one key mission of the mini-series was to bring together members of both communities (Earth Science and ontology/semantics) into a meaningful dialog.

This mini-series of events are co-organized or supported by members of the EarthCube community, Ontolog community, SOCoP community and IAOA community

Organizers:

DaliaVaranka GaryBergCross KrishnaSinha KrzysztofJanowicz LeoObrst
Mark Schildhauer MikeDean NaicongLi NancyWiegand PascalHitzler PeterYim

Objective of the Earth Science Ontolog mini-series

- to help start collaborative investigations and evolve a path to implement methods for community supported development of both data and service ontologies, as well as to demonstrate application of ontologies during the five stages of data related activities identified below:
  - 1) enabling providers to SHARE data
  - 2) enabling ease of access to both centralized and distributed data
  - 3) enabling discovery of data of interest to the end user
  - 4) enabling integration and fusion of data
  - 5) enabling modeling capabilities

Virtual Panel Sessions for this Mini-Series

- Session-1: **Value Proposition of Ontology and Semantic Technology for the Earth Science Community,** co-chairs: Krishna Sinha & Leo Obrst
- Session-2: **Ontology Development and Application across Earth Science Systems Lifecycle,** co-chairs: Gary BergCross & Naicong Li
- Session-3: **Heterogeneity-preserving Data Interoperability: Methods and Challenges,** co-chairs: Krzysztof Janowicz & Pascal Hitzler
- Session-4: **Ontologies for Earth Sciences,** co-chairs: Dalia Varanka & Mark Schildhauer
- Session-5: **Tutorials,** co-chairs: Nancy Wiegand & Mike Dean

with facilitation and infrastructure support provided by
Peter Yim, Co-convener of **ONTOLOG**

Speakers at the EarthScienceOntolog **Mini-series**

1. Krishna Sinha, Virginia Tech
2. Krzysztof Janowicz, University of California, Santa Barbara
3. Dalia Varanka, U.S. Geological Survey
4. Pascal Hitzler, Kno.e.sis Center , Wright State University
5. Mike Dean, Raytheon BBN Technologies Corp.
6. Rick Hooper, Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI)
7. Peter Fox , Rensselaer Polytechnic Institute (RPI)
8. M. Santoro, Institute of Atmospheric Pollution Research - Italian National Research Council (CNR), Rome, Italy
9. Naicong Li, The Redlands Institute, University of Redlands
10. Simon Scheider , Münster Semantic Interoperability Lab (MUSIL)
11. Prateek Jain, IBM TJ Watson Research Center
12. Benjamin Adams, National Center for Ecological Analysis and Synthesis (NCEAS), University of California, Santa Barbara
13. Giancarlo Guizzardi, Federal University of Espírito Santo (UFES), Brazil
14. Luis Bermudez, Open Geospatial Consortium (OGC)
15. Boyan Brodaric, Geological Survey of Canada
16. Norman Morrison on behalf of the EnvO Consortium
17. Thomas Huang, Jet Propulsion Laboratory, California Institute of Technology
18. Mark Schildauer, NCEAS, University of California, Santa Barbara
19. Nancy Wiegand, University of Wisconsin
20. Leo Obrst, MITRE Corporation
21. Dave Kolas, Raytheon BBN Technologies Corp.

Highlighted Semantic themes from the series (adapted from abstracts)

- Semantic technologies and ontologies are key building blocks for next-generation scientific infrastructures and workflow systems, and support scientists in publishing, discovering, and integrating data.
- Implementation of valuable legacy data using semantic technologies provides the opportunity to reexamine the way the data were historically modeled and interpreted compared to the way they can be used to meet updated objectives.
- Aligning Earth Science data management and use with the state of the art in Semantic Web technologies.
- Semantic Web technologies (including the synergistic combination of ontologies and linked open data) appear to be widely applicable to large scale earth science data management and applications.

- A centerpiece of this effort (CUAHSI) is a metadata catalog that is drawn from the more than 90 published services and indexes tens of millions of time series. This discovery services is mediated by a list of search terms, a controlled vocabulary of "leaf concepts" that are arranged into a taxonomy from general to specific terms.
- In 2004 languages and tools were mature enough to develop workable applications. Along the way, we refined a development methodology which has enhanced our collaborative semantic web efforts and broadened the developer base
- Enhancing geospatial resource discovery capabilities can be achieved by augmenting the searchable descriptions of resources. Examples of additional descriptions (that is, something not searchable with typical geospatial discovery services) are: semantic information and user-generated annotations. This is the task assigned to a middleware component called "Semantic Broker".

- Automatic discovery of relevant and interoperable resources for the decision problem at hand needs formalization of decision problems and the approach adopted for problem solving, as well as semantic annotation of existing resources.
- The ontology and semantic web community has undergone several strategic shifts in terms of paradigms, from holistic ontology standardization, over ontology alignment with top-level ontologies, to a pluralist translation and similarity paradigm. Ontology engineering is currently changing from top-down engineering to bottom construction, e.g., in the form of design patterns. Interestingly, all of these paradigms were motivated by the goal of semantic interoperability. However, the role of semantic heterogeneity differs considerably among them: The more recent paradigms face the challenge of heterogeneity preservation rather than heterogeneity resolution.

- Identifying relationships between entities is a natural behavior of human mind. Humans have been doing it since our existence on this planet. The same human behavior has been replicated in various fields such as databases and semantic web. ……..BLOOMS for the identification of relationships between ontology schemas. The system computes alignments with the help of noisy community-generated data available on the Web. Currently, BLOOMS uses Wikipedia and the Wikipedia category hierarchy for this purpose
- As ontologies become more mainstream in the Earth sciences and related research areas there is interest in measuring the semantic similarity of entities and entity classes described in those ontologies. Application areas for semantic similarity measurement range from semantic interoperability and ontology matching to understanding uncertainty in geographic feature types and land-use classes to semantics-enabled geographic information retrieval.
- Formal representations of concepts within OGC standards and the ontology behind the new GeoSPARQL.

- The quality of an Ontology as a Reference Model of Consensus depends not only on its suitability to capture relevant notions describing that domain, but also on how truthful it is to the underlying domain in reality. Looking at the Ontology as a logical theory, it should be axiomatized in a way that it will have as possible models exactly the intended ones, i.e., exactly those models which represent state of affairs deemed acceptable by the underlying conceptualization. Under this view, one fundamental methodological issue is: how can we help ontology engineers identify the exact formalization that will produce such a result?
- Pressing scientific and societal issues, coupled with the increasing availability of water data, require an integrated approach to surface and subsurface water representations. However, emerging representations are largely disconnected at present, most notably within geospatial data standards, thus impeding joint use of data for key aspects such as estimating water balances.
- The Environment Ontology (EnvO) is a community-developed ontology which provides a controlled, structured approach to support the annotation of any organism or biological sample with environment descriptors. EnvO contains terms for biomes, environmental features, and environmental material which enable the description of environments from coarse to fine levels of granularity

- Using knowledge base technology to drive design decisions and as an instrument for validation not only cuts cost in product development, but it also enables much more thorough validation of the design prior to fabrication.

- An ecosystem of Earth and environmental science vocabularies, markup languages, and ontologies has arisen in recent years. These representations of the meanings of Earth science terms vary greatly in thematic scope, intended uses, foundational design principles, and axiomatic structures and inferencing capabilities. Common issues include the need for definitions of terms that come from authoritative sources, better use of richer, formalized representations for describing relations among concepts, and a need for clearer governance mechanisms that enable systematic approaches to evaluating, populating, updating and using vocabularies and ontologies.

Highlighted themes from Mini-series that will enable sharing-access-discovery-interoperability across distributed and heterogeneous resources

Infrastructure building blocks	Legacy data	Semantic Web	linked open data
Controlled vocabulary	Semantic Broker	Semantic annotation	Heterogeneity preservation
Ontology schemas	Category hierarchy	Semantic similarity	Formalization of underlying concepts
Concepts within OGC standards	Geospatial data standards	Community-developed ontology	Cost benefit of knowledge base technology
Meanings of Earth science terms vary	Semantic Collaboratory	Beyond vocabulary to data ontology	Process and Service ontology

Summary

The mini-series identified

1. The value of semantics in sharing, accessing, discovering and interoperability: glue that binds EarthCube together
2. The need to utilize ontologies for knowledge management : science driver for EarthCube
3. The need to collaborate : leading to successful implementation of EarthCube

Thank You

Presenters, Participants and Organizers

Mini-Series archived material available at

http://ontolog.cim3.net/cgibin/wiki.pl?EarthScienceOntolog