[PS]Improving E&P Data Interoperability Through the Development of a Reusable Earth Science Ontology for Basin Characterization*

Melanie A. Everett[1], Scott Hills[1], Mark Gahegan[2], Brandon Whitehead[3], and Boyan Brodaric[4]

Search and Discovery Article #40809 (2011)
Posted September 30, 2011

*Adapted from poster presentation at AAPG Annual Convention and Exhibition, Houston, Texas, USA, April 10-13, 2011

[1]Chevron Energy Technology Company, Houston, TX (MEverett@chevron.com)
[2]Centre for eResearch, University of Auckland, Auckland, New Zealand
[3]School of Environment, University of Auckland, Auckland, New Zealand
[4]Geological Survey of Canada, Ottawa, ON, Canada

Abstract

Data and information volumes relevant to the upstream energy industry continue to expand in a growing number of diverse, distributed repositories, which makes discovery and integration of information exponentially more difficult for knowledge workers. We believe that improving this situation requires adoption of a knowledge management infrastructure reliant on semantic technologies and industry-wide semantic standards. Central to these standards is the concept of a reusable ontology, i.e., a codified organization of concepts and their relationships that can be shared across domains, accessed by various types of applications, serve as an anchor for more detailed ontologies, and may itself anchor into higher-level abstractions for reasoning across a broader domain.

To demonstrate this idea, we have developed a reusable earth science domain ontology, the BASIN ontology, focused on concepts associated with basin characterization. This ontology could serve as an anchor point for future domain-specific ontologies intended for use along the entire upstream value chain from hydrocarbon exploration to production. To further promote semantic interoperability, the BASIN ontology will itself be anchored into upper-level earth science ontologies in the public-domain, such as the Semantic Web for Earth and Environmental Terminology (SWEET) developed by NASA.

The central concept of the BASIN ontology is the basin class, which is related to other classes through select earth processes (e.g., the basin class is related to the strata class via tectonic processes, such as subsidence). This high-level abstraction of primary basin concepts allows for application of the ontology to a wide variety of E&P earth science projects. The ontology was developed by

resolving semantic ambiguities and differences in the terminology used by practitioners within the domain through knowledge elicitation from key E&P subject matter experts. This disambiguation resulted in identification of central concepts represented as classes, their definitions, and the establishment of critical connections between classes.

Current applications benefiting from the BASIN ontology include both corporate knowledge management and enterprise semantic search. However, the development of the BASIN ontology was driven fundamentally by a desire to increase semantic interoperability throughout the geoscience community and encourage development of semantic standards having value to the upstream energy industry.

Improving E&P Data Interoperability Through the Development of a Reusable Earth Science Ontology for Basin Characterization

Melanie A. Everett[1], Scott J. Hills[1], Mark Gahegan[2], Brandon Whitehead[2], Boyan Brodaric[3]

[1]Chevron Energy Technology Company

[2]Centre for eResearch Univ. of Auckland

[3]Geological Survey of Canada

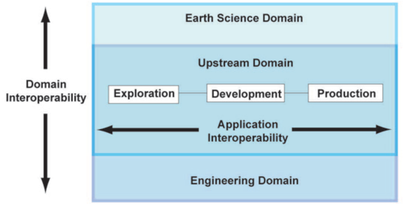
Leveraging Ontologies in the Upstream Energy Industry

Data and information volumes relevant to the upstream energy industry continue to expand and diversify at exponential rates.

Improved management of these resources is contingent upon adoption of semantic technologies and standards, which includes the use of **ontologies** - codified frameworks that describe a common set of concepts or terms and their relationships to each other.

Here, we present the **BASIN ontology**, which is the first formal descriptive framework of many of the features and processes associated with sedimentary basin and petroleum system formation.

The BASIN ontology is a **reusable ontology**, in that it can be shared across domains and applications. Core concepts in the BASIN ontology not only overlap the earth sciences and engineering domains, but are also applicable across the upstream value chain, from hydrocarbon exploration to production (shown above). This overlap emphasizes the reusability of the BASIN ontology.



The interoperability of the BASIN ontology allows it to be used in a variety of information management applications, including corporate knowledge management, enterprise search, and database schema mapping.

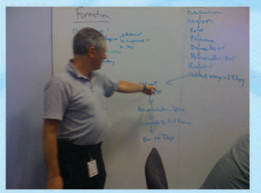
Development of the BASIN Ontology

The BASIN descriptive framework was developed through a community knowledge exercise[1] orchestrated to discuss fundamental concepts and their meanings, as interpreted by leaders in basin characterization.

Concepts were first surfaced through examples of specific basin studies shared by the experts, at which point they were explicitly described through diagrams to the satisfaction of the other participants. This process allowed interrelationships between fundamental and domain-level concepts to be exposed and disambiguated.

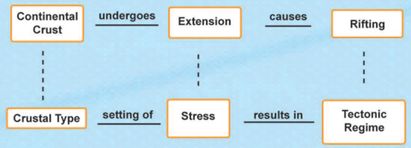
As the exercise progressed, key concepts and relationship types became apparent and were captured in a concept map-representation of the BASIN ontology.

The resulting descriptive framework shows "basin" as a central concept surrounded by connectors, which depict relationships of the basin to various processes and physical properties. Some areas of the ontology space are more rigorously fleshed out than others, reflecting the organic nature of the knowledge acquisition exercise.



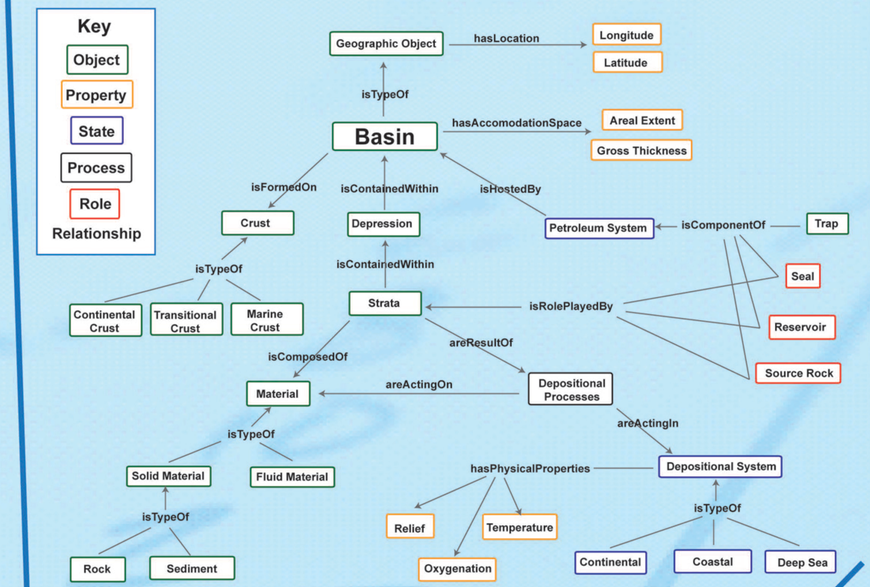
Subject matter experts Jeff Johnson (Chevron ETC), Barry Katz (Chevron Fellow, shown far left), and Marty Perlmutter (Chevron ETC, shown left) participated in the knowledge extraction exercise that resulted in the BASIN ontology.

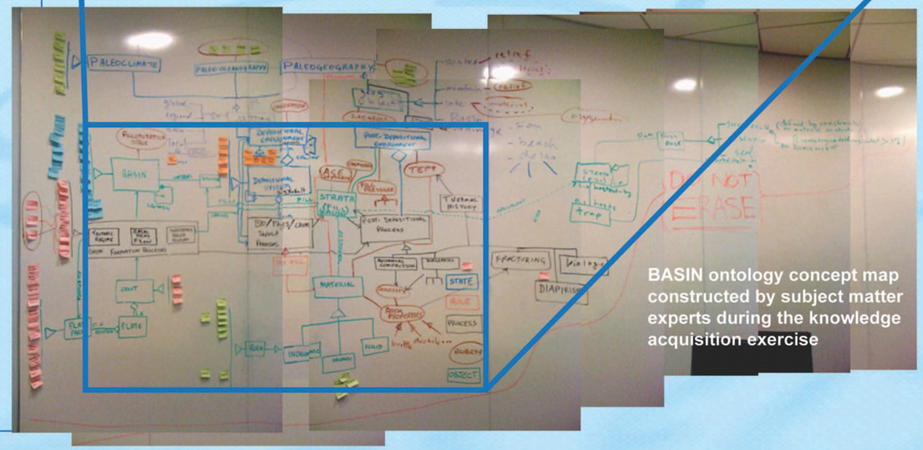
Basin Example:
Central Sumatra Basin



Abstracted Concepts and Relationships

Excerpt from the BASIN Ontology Concept Map



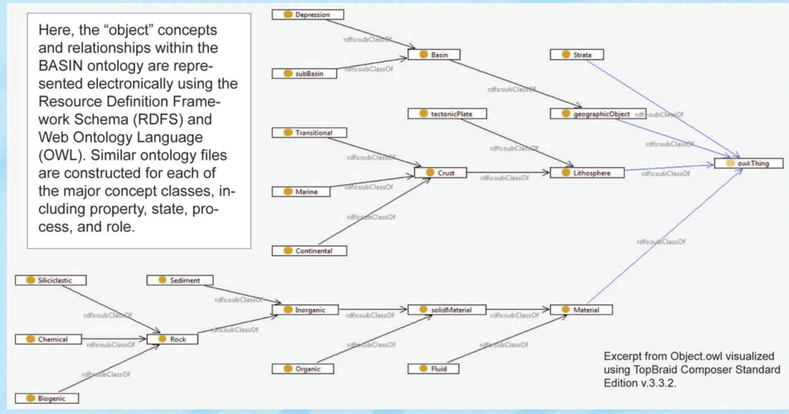


BASIN ontology concept map constructed by subject matter experts during the knowledge acquisition exercise

From Concept Map to Semantic Web Language

Modeling the BASIN ontology in semantic web languages allows its use in digital applications, such as enterprise search. This encoding increases interoperability of the BASIN ontology by anchoring it into similarly modeled structures like NASA's mid-level Semantic Web for for Earth and Environmental Technology[2] (SWEET ontologies). The cross-domain framework that the BASIN ontology creates also has the potential to connect and integrate domain-specific data standards, such as Energistics' RESQML[3], an XML exchange format for earth model data.

Here, the "object" concepts and relationships within the BASIN ontology are represented electronically using the Resource Definition Framework Schema (RDFS) and Web Ontology Language (OWL). Similar ontology files are constructed for each of the major concept classes, including property, state, process, and role.



Excerpt from Object.owl visualized using TopBraid Composer Standard Edition v.3.3.2.

Applications of the BASIN Ontology and Forward Plan

Applications expected to benefit from the BASIN ontology include corporate knowledge management, enterprise semantic search, and database schema mapping.

The exercise used to develop ontologies has the potential to surface complex and intuitive knowledge held by subject matter experts. The ontology provides a framework within which to capture and disseminate this knowledge. As such, ontologies and the exercises used in their development can serve as training tools for new staff.

For applications such as enterprise semantic search, descriptive frameworks such as the BASIN ontology provide a means of disambiguating concepts that carry multiple meanings, even within the same domain. For example, the term "migration" to a petroleum geochemist may hold a different meaning to a seismologist.

Development of the BASIN ontology was also driven by a desire to increase data interoperability throughout geoscience and upstream communities. Applications leveraging the reusability of the BASIN ontology include database schema mapping or "translation".

We are discussing with Energistics[4] the opportunity to promote within the upstream energy community collaborative development and adoption of standard vocabularies and semantic structures, such as the BASIN ontology. This effort would represent a step towards mitigating the impact of the continued growth of information resources.

REFERENCES

[1]Ribes, B., Bowker, G. 2009. Between meaning and machine: Learning to represent the knowledge of communities. *Information and Organization* 19(4):199-217.

[2]*Semantic Web for Earth and Environmental Technology (SWEET).* Last updated 17 Dec. 2010. NASA. Accessed 17 Mar. 2011. <http://sweet.jpl.nasa.gov>

[3]*RESQML Standards.* Last updated 2011. Energistics. Accessed 17 Mar. 2011. <http://www.energistics.org/resqml-standards>

[4]*energistics: The Energy Standards Resource Centre.* Last updated 2011. Energistics. Accessed 21 Mar 2011. <http://www.energistics.org>

ACKNOWLEDGEMENTS

We would like to thank the following people for their contributions to this work: Jeff Johnson (Chevron ETC), Barry Katz (Chevron Fellow), Marty Perlmutter (Chevron ETC), Doug Gregory (Chevron ETC), Frank Chum (Chevron ETC), and Chien-Fen Chiu (Chevron ITC).