Short Paper:

Deep Semantics in the Geosciences: semantic building blocks for a complete geoscience infrastructure

Brandon Whitehead,[1,2] Mark Gahegan[1]

[1]Centre for eResearch
[2] Institute of Earth Science and Engineering
The University of Auckland, Private Bag 92019, Auckland, New Zealand
{b.whitehead, m.gahegan}@auckland.ac.nz

Abstract. In the geosciences, the semantic models, or ontologies, available are typically narrowly focused structures fit for single purpose use. In this paper we discuss why this might be, with the conclusion that it is not sufficient to use semantics simply to provide categorical labels for instances—because of the interpretive and uncertain nature of geoscience, researchers need to understand how a conclusion has been reached in order to have any confidence in adopting it. Thus ontologies must address the epistemological questions of how (and possibly why) something is 'known'. We provide a longer justification for this argument, make a case for capturing and representing these *deep* semantics, provide examples in specific geoscience domains and briefly touch on a visualisation program called Alfred that we have developed to allow researchers to explore the different facets of ontology that can support them applying value judgements to the interpretation of geological entities.

Keywords: geoscience, deep semantics, ontology-based information retrieval

1 Introduction

From deep drilling programs and large-scale seismic surveys to satellite imagery and field excursions, geoscience observations have traditionally been expensive to capture. As such, many disciplines related to the geosciences have relied heavily on inferential methods, probability, and—most importantly—individual experience to help construct a continuous (or, more complete) description of what lies between two data values [1]. In recent years the technology behind environmental sensors and other data collection methods and systems have enabled a boom of sorts in the collection of raw, discrete and continuous geoscience data. As a consequence, the operational paradigm of many conventional geoscience domains, once considered data poor, now have more data than can be used efficiently, or even effectively. For example, according to Crompton [2], Chevron Energy Technology Corporation had over 6000 Terabytes of data, and derived products such as reports, and is rapidly expanding. This data deluge [3], while significant in its affect on capturing information related to complex earth science processes, has become a Pyrrhic victory for geoscientists from a computational perspective.

The digital or electronic facilitation of science, also known as eScience [4] or eResearch, coupled with the science of data [5] is fast becoming an indispensable aspect of the process of Earth science [6–8]. There are exemplar projects such as OneGelogy[1], which translates (interoperates) regional geologic maps in an effort to create a single map of the world at 1:1 million scale; as well as the Geosciences Network[2] (GEON) which houses a vast array of datasets, workflows, and tools for shared or online data manipulation and characterisation. Further, the National Science Foundation (in the U.S.A.) has funded EarthCube[3] which seeks to meld the perspectives of geoscientists and cyberscientists to create a framework for locating and interoperating disparate, heterogeneous information about the entire Earth as a comprehensive system. The major contributions that eScience can make is by providing ways to communicate the semantics, context, capabilities and provenance of the datasets, workflows, information and tools in order for researchers to have a firm understanding of the artefacts they are using, and how they are using them.

In this paper, we illustrate how multiple, multi-faceted semantic models are coordinated under the linked data paradigm to better reflect how geoscience researchers situate concepts with their own knowledge structures in an effort to contextualise observations, phenomena and processes. We look to expose which semantic, or ontological, commitments are needed to glean how science artefacts relate to researchers, methods and products (as data, or via theory) in order to transfer what is known about a place, and how it is known, as a useful analog for geoscience discovery. We use an interactive computational environment, known as *Alfred*, to view disparate ontologies that carry pieces of this 'knowledge soup' [9] as facets, and expose the relationships for discovery of new knowledge.

2 Geoscience Background

The geosciences are far from exact; the earth as a living laboratory provides plenty of challenges, not least to the task of representing and communicating semantics. While geoscientists are remarkable in their ability to utilise disparate knowledge in mathematics, physics, chemistry and biology to create meaning from observed phenomena, their theories are bound by the inherent problems associated with scale and place, cause and process, and system response [10]. The Earth's phenomena are complex, they often exhibit statistically unique signatures with several stable states while mechanical, chemical and biologic processes work in tandem, or asynchronously. Due to these often contradictory complications it has also been suggested that the Earth sciences exemplify a "case study to understand the nature and limits of human reasoning, scientific or otherwise" [11]. Adding to the complexity, "Geologists reason via all manner of maps, outcrop interpretation, stratigraphic

1 http://www.onegeology.org/

2 http://www.geongrid.org

3 http://earthcube.ning.com/

relationships, and hypothetical inferences as to causation" [12] and they do this simultaneously across geographic and temporal scales.

In order to discern the categories and components of the Earth as a system, the geoscientist requires a trained eye, what anthropologists call "professional vision" [13], which often necessitates years of experience and mentoring. This contextualised view of the world uses a long view of time, and becomes adept at distinguishing infrequent catastrophic events from those more frequent via the feedback loops between processes and components [13]. However, these feedback loops are often not well understood due to the fragmented nature of geoscience observation and data. This has required the geoscience community of practice to develop the means by which their observations are understood. Most notably, instead of constructing a specific research question and testing it, geoscientists often use the method of 'multiple working hypotheses' [14] and work toward reducing what is not known, instead of working towards some axiomatic truth. Indeed, the ability to abstract earth processes to a rational metaphoric justification could be considered an art form.

As such, geology is often referred to as an interpretive science [15]; where empirical evidence is not possible, a story often emerges. Interpreting meaning in the geosciences revolves heavily around the inherent allusion in hypothesis, methods, models, motivations, and often more importantly, experience. Understanding the knowledge any researcher creates requires understanding that person's research methods and the rationale behind their decision processes, which requires the ability for knowledge components to change roles as one tries to demystify the scale in context and perceptions from which they are constrained. Often, what is determined to be a result is steeped in probability as a function of a desired resource. To date, the research and research tools used throughout geoscience domains are largely situational; capturing tightly coupled observations and computations which become disjointed when the view, filter, or purpose is altered, even slightly, to that which is more representative of an earth system science.

3 Semantic Modelling in the Geosciences

As the previous section suggests, the semantic nature of geoscientific ideas, concepts, models, and knowledge is steeped in experiential subjectivity and often characterised by what can or cannot be directly observed, directly or indirectly inferred, and, in many cases, the goals of the research. As the Semantic Web [16] has gained traction and support, a subset of Earth science researchers have been intrigued by the possibility of standards, formal structure, and, ultimately, ontologies in geoscience domains, mainly because, as Sinha et al., have stated, "From a scientific perspective, making knowledge explicit and computable should sharpen scientific arguments and reveal gaps and weaknesses in logic, as well as to serve as a computable reflection of the state of current shared understanding" [17].

As evidenced by the dearth of semantic models, or ontologies, in the earth sciences [18], the often-conflicting ideals and knowledge schemas are proving to be significant hurdles for ontological engineers. Most of the semantic models in Earth science communities would be considered weak [19], lightweight (sometimes referred to as 'informal') [20, 21] or implicit [22]. These include taxonomies, or controlled vocabularies—like the American Geophysical Union's (AGU) index of terms,[4] glossaries [23], thesauri [24], or a typical data base schema. Conversely, semantic models created with the aspiration of eventuating to strong, heavyweight or formal ontologies are limited. In cases where published formal domain ontologies do exist [25], they are often not openly available within the community.

One openly available ontology of note is the upper-level ontology SWEET: Semantic Web for Earth and Environmental Terminology [26]. This formal ontology was created to tag the huge repositories of satellite imagery created and housed by NASA. As a result, the concepts used in SWEET are very high level and the granularity of the ontology is, in most cases, not detailed enough to differentiate between the thousands of resources an active research geoscientist might find useful,

There is a middle ground between the two ends of the semantic spectrum, in the form of Mark-up Languages, which is quite promising. In the geosciences, two exemplar Mark-up Languages do exist; the Geography Mark-up Language (GML) [27], and Geoscience Mark-up Language (GeoSciML) [28]. The two languages were created to serve mainly as a translation schema for data sharing and interoperability, and do provide a level of formalisation and weakly typed relationship structure.

4 A Case for *Deep* Semantics

In this research we use the relative lack of formalised structures in the geosciences as an opportunity to start from scratch and take a slightly different approach to ontological engineering in the domain. We try get away from the highly restrictive, monolithic, overarching structures and focus on a more complete picture of the relational patterns of geoscientific artefacts. To summarise Gahegan et al., [29]: we are looking to expose the 'web of multi-faceted interactions' between observations, theory, data, motivations, methods, tools, places and people. To focus the modelling effort, we asked the following questions: How is something known? Which entities support a research artefact? Who has been publishing about a topic, concept, place, research method or data product? What was the inference path a geoscientist traveled from a piece of evidence to an interpretation?

As these questions might suggest, a *deep* semantic support structure should provide a conceptual richness that permeates the depth of a specialised set of concepts and provide a mechanism for defining how an artefact came to be represented. A *deep* semantic structure should provide enough specificity in the concepts and relations that

[4] http://www.agu.org/pubs/authors/manuscript_tools/journals/index_terms/

the terms can be used to differentiate complex but real situations via the written materials describing these situations, not simply how it was labelled, or tagged, by the creator, librarian or data curator. Exposing this story behind the data, or, more formally, the epistemological connections, *deep* semantics works towards constraining the conceptual uncertainty of the procedural knowledge by explicitly representing and exposing the semantics of research artefacts as the scientist has orientated them for his/her evidence, and thus, the resultant interpretation. This effectively frees the research scientist to focus on the declarative knowledge supporting the probabilities in the numerical components.

The ability to locate resources has become increasingly important as data storage continues to increase. What carries a heavier weight is the ability to locate a data product at the time it is most useful, by being able to distinguish a resource's *when* and *where* relatively rapidly. With a *deep* semantic view, we are able to begin pursuing the *why* and *how* of the conceptual structures that support geoscientific knowledge and discovery. In the information sciences this is often referred to as precision and recall. Deep semantics adds epistemological underpinnings and a level of context to precision and recall while adding facets to the constraint and delineation mechanisms.

5 Ontology Inception and Use

Building the relationship structures, as described in this section, of the disparate parts of geoscience research artefacts creates a contextualised, and in this case visual, representation of the network of ontological components that support a concept. We treat every ontological component as linked data supported by domain specific terminological ontologies [30]. We use the full version of the Web Ontology Language (OWL Full) to promote emergent constraints via relationships when possible. OWL Full was chosen due to its compatibility with other modeling approaches, most notably Resource Description Framework (RDF), as well as reducing the restrictions on class definitions. The latter is necessary in the Earth sciences as it is quite common to find a concept, or identifier, that is a class name as well as an instance. In addition, given the nature of geoscience knowledge, it should not be logically impossible to arrive at a conclusion that is not yet known to the system through the ontological framework. We felt this fits more in line with the process of Earth science, which relies quite heavily on reducing what is not known rather than enforced, top-down, logical constraints depicting what is known axiomatically. It is this connected interworking of heterogeneous semantic models ranging from weak to strong, lightweight to heavyweight, and informal to formal which join together as linked data, that we have come to refer to as *deep* semantics. The remainder of this section describes how each individual ontology was constructed.

5.1 Basin and Reservoir ontology

We endeavoured to create a framework for formal geoscience knowledge as it applies to sedimentary basins and reservoirs in the energy and petroleum industry under the aegis of recognised industry Subject Matter Experts (SMEs). The SMEs participated in knowledge acquisition exercises [31] orchestrated to discuss fundamental concepts and their meanings as interpreted and explained through their formal and experiential mastery. As concepts emerged, they were explicitly described, often through diagrams and examples, to the satisfaction of the other participants. Prior to each workshop, a set of concepts had been extracted from a survey of applicable literature in the domain to serve as exemplars for the types of concepts found in research artefacts that differentiate and describe specific geoscience situations and models. These concepts were periodically re-introduced to the SMEs to ensure structure that was being created had semantically tenable end-points. This process allowed interrelationships between fundamental and domain-level concepts to be exposed and characterised. As the exercise progressed, clusters of concept and relationship types became apparent. The open nature of the knowledge acquisition exercise allowed the participants to navigate through the conceptual neighbourhood that they had created. As such, there are areas of the concept space that are defined more rigorously than others.

The workshops culminated two ontological frameworks: a Basin ontology [32, 33] and a Reservoir ontology. The Basin ontology focuses on concepts corresponding to basin characterisation. The core concepts are related to properties and other classes through select earth processes (e.g., the Basin class is related to the Strata class via a tectonic processes, such as subsidence). The Reservoir ontology was created quite similarly as the Basin ontology, with the exception being that the contributing SMEs were well versed in petroleum reservoir characterisation and modeling instead of basin characterisation.

The Basin and Reservoir ontologies have been created to interoperate with each other to coordinate the delineation of scale dependent ambiguities in research artefacts. To further promote semantic interoperability, both of these ontologies have natural contact points for semantic correlation with upper-level earth science ontologies in the public-domain, such as SWEET, as well as with other domain-specific ontologies from hydrocarbon exploration and production, to hydrologic and paleoclimate modeling, should they become available.

5.2 Agent and Resource ontology

Two of the more important facets of this research are the actual research artefacts and the researchers, or creators, of those artefacts. Fortunately, librarians have already spent a significant amount of time developing a standard for metadata that captures the types of information that we wanted to capture from resources. We used the Agent profile from the Dublin Core Metadata Initiative[5] (DCMI) to describe authors, contributors, software, companies, and research groups. There are other types of agents, of course, and DCMI is set up to handle these distinctions, but for our

[5] http://dublincore.org/

purposes a subset was all that was required. The Resource profile from DCMI is used to describe any artefact produced by research. This can include publications, abstracts, presentations, and data products. Again, the schema for the DCMI framework allows for a plethora of types, but a subset was all that was required here.

5.3 Task ontology

The Task ontology was constructed to provide a framework for actions that are completed during research. These include observations, methods and processes like data collection, data manipulation, statistical methods, etc. Items in the Task ontology link to Resources as outputs and inputs, and to Agents as creators, contributors, reviewers, etc. The concepts in this specification are often chained together to create large structures, and are helpful in delimiting clusters of information. The Task ontology was created by, first, describing a small set of exemplar concepts that related strongly to key components of the semantic models described in the previous two sections. Once the initial concepts were introduced, the structure was extended by defining known superclasses and subclasses, and then supplanting those core concepts with text mining utilising basic natural language processing principles.

5.4 Oilfield Glossary and World Oil and Gas Atlas ontology

The Schlumberger Oilfield Glossary[6] is a fantastic on-line resource covering an expansive number of topics. The Oilfield Glossary ontology was constructed from harvesting the information hosted on this web site. Due to research limitations, it was more beneficial to create a local copy of this data and convert that to a series of triples than to develop a script to query the site interactively. During the data conversion, all partial relationships within the structure, and the links to the corresponding web page were preserved.

The World Oil and Gas ontology was created by manually entering information provided in the summaries, graphs, and tabular data, as depicted in the World Oil and Gas Atlas [34], into an electronic format. Once in an electronic format, a script was generated to alter the format, along with a little manual editing, to OWL.

6 Early Results: Powder River Basin Use Case

This use case illustrates how research artefacts associated with the Powder River Basin, located in the central part of the U.S.A., can be visualised and navigated via a knowledge computation platform referred to here as *Alfred*. This platform allows for navigating and manipulating disparate multifaceted structures in one graph space. The user loads an ontology into the system, which is then represented as a facet. Any facet can be docked to any length of the graph border. Through docking a facet (up to four), *Alfred* provides a space for the user to follow their interests in their linked data exploration.

---

6 http://www.glossary.oilfield.slb.com/

We proceed from the perspective of a research scientist with an interest in the Powder River Basin. The user enters "Powder River" into *Alfred's* the search field. The resulting graph, shown in Figure 1, shows two concepts matching the search term within a neighbourhood of related concepts. One of the Powder River concepts (highlighted with a yellow outer ring) is symbolised using a gold circle coupled with a black 'basin' object from a scalable vector graphic (SVG). The edge pointing to the yellow circle labeled Basin, denotes it is an instance of the *Basin* class (yellow disc) from the Basin ontology. The other symbol labeled Powder River is a grey triangle which signifies it as a member of the World Oil and Gas ontology (in the structure, these two concepts are in fact connected via an *owl:sameAs* relation, but this type of relation has been suppressed in the current view for readability). Three conference abstracts, symbolised by a red circle, with an SVG in the shape of a book, relate via a *references* edge to the Powder River concepts, as well as a few concepts symbolised by black diamonds, which are delineating concepts from the Oilfield Glossary ontology.

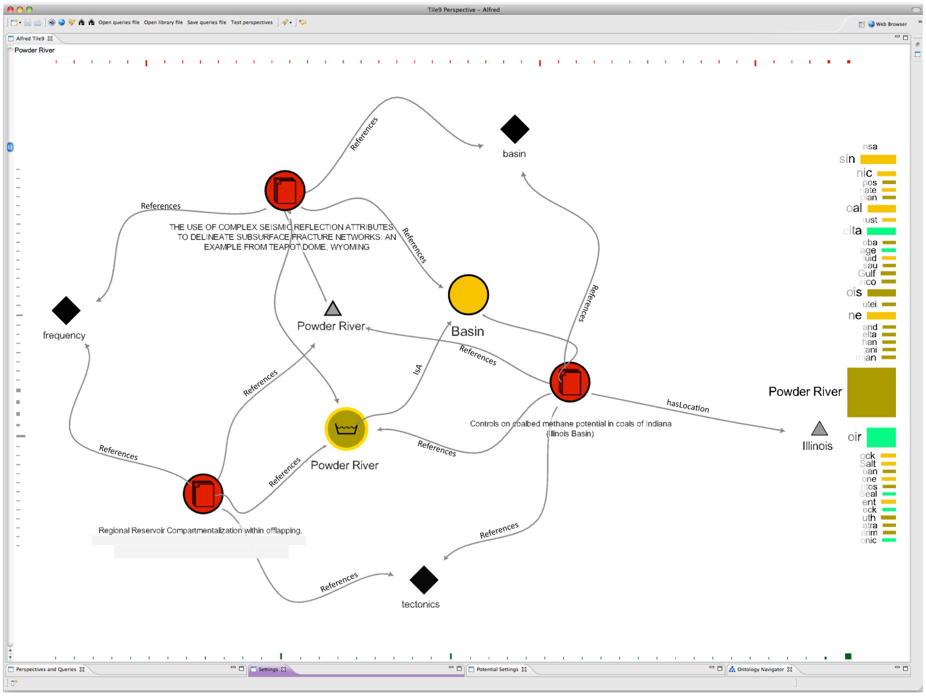

Fig. 1. Local graph neighbourhood of the concept representing the Powder River Basin. The current view shows three published artefacts, how the concept Powder River is linked in the hierarchy (it is an instance of Basin) as well as a few terms from the Oilfield Glossary.

At this stage, the user has a few options. The user can select something from the graph, adjust the filter settings to increase the type and/or level of information displayed in the graph, or start a new search. To continue with the example in the use case, we assume the users interest has been piqued by one of the research artefacts sharing a relationship with the Powder River Basin concept. If the user were interested in seismic reflection data, they might select the artefact purporting to deal with complex seismic reflection attributes (red disc with book SVG, lower left) via a double click.

Upon this click action, the graph re-centres itself using the user selected node as the central concept, as depicted in Figure 2. This selection reveals a deeper structure associated with that particular artefact. In this view, the creator of the artefact, symbolised by a dark green circle surrounding a SVG of a person, has emerged along with several concepts found in the Oilfield Glossary. The relationship to the Powder River and Basin concepts have persisted to the new concept layout (lower middle of Fig. 2). Several concepts from the Task ontology (blue circles) have emerged, potentially signifying relationships to the data (seismic data) as well as concepts related to analysis mechanisms (phase coherence) associated with this particular research artefact. This view also provides the user with contextually similar research artefacts by displaying the research outputs that share a relation with other concepts in the known ontologies.

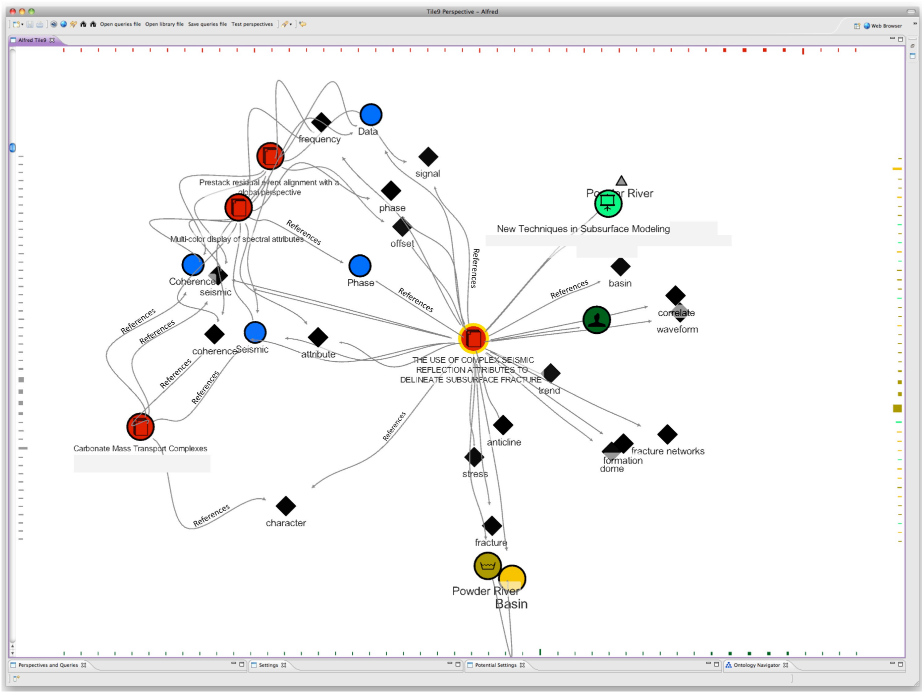

Fig. 2. Local graph neighborhood of a research artefact related to the Powder River Basin. The current view shows the artefacts creator, as well as other concepts from the semantic structures known to the system.

In the example depicted in Figure 2, three research artefacts appear to share several relationships (mostly a *reference* edge) with concepts found in the Oilfield Glossary, as well as the Tasks ontology. This clustering suggests there are other research products that have utilised the same, or similar, methods and data that were used in the research artefact of interest. This is worth mentioning here as the related data, tasks, and concepts allow the user to explore and glean the concepts and structures that support a research artefact. The ability to navigate, what has become, the epistemological lineage of a research artefact cultivates a formal representation of the symbiotic components of geoscience research products and geoscience knowledge.

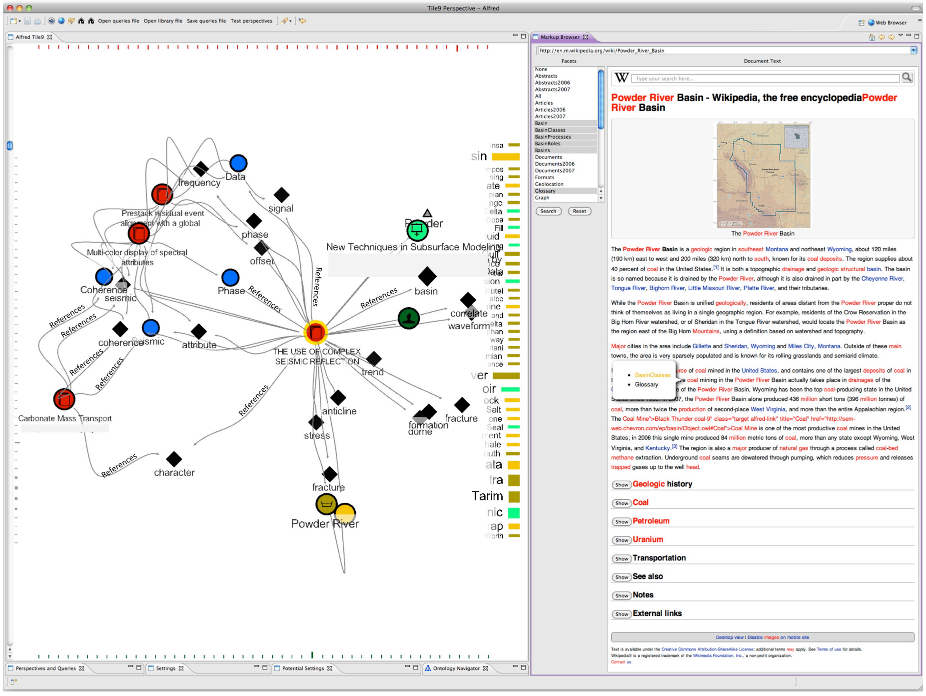

Fig. 3. A local graph neighbourhood shown with a web page that is marked up with red text using the concepts from the ontologies loaded into the system. The user can go back and forth between graph space and the web page in order to better refine the context and scope of research artefacts and web enabled content.

At this final stage of the use case, we illustrate how the conceptual neighbourhood of the graph can synchronise with other web enable components, in this case browser content. As portrayed in Figure 3, a user can open a web page and search for known semantic components on that page. When the search has completed, all known concepts are now displayed with red text within the browser. Further, by hovering over any syntactic component on the corresponding web page (in this instance, it is the mobile version of the Wikipedia page for the Powder River Basin) the user is presented with a pop-up dialogue populated by the referring ontology. If a term is situated in multiple ontologies, each one will be listed in the pop up, with the text providing a live link back to the graph space. In this way, a user could bring their conceptual neighbourhood with them as they peruse web content and use the highlighted text references to help filter for relevance. This has proved to be particularly helpful with the increasing number of peer reviewed publications available in a web friendly format.

The Powder River Basin use case illustrates how *deep* semantics can benefit geoscientists by providing a mechanism to visualise and explore the components that comprise a knowledge construct. When a geologist purports to know how to characterise a particular basin, other geologists and engineers naturally want to know what data and analysis methods were used to support that interpretation. How was the stratigraphy interpreted? What was the timing of the tectonic events? What is the burial history? *Deep* semantics allows other geoscientists to explore these supporting entities and the decisions that were made along the path to any particular explication.

7 Concluding Remarks

Geoscience ontologies are typically quite lightweight, or implicit, and are engineered for one specific purpose. As such, the semantic structures in the geosciences fail to capture the complexities and intricacies inherent in the domain knowledge. Ontologies like SWEET are a great start at a general upper level structure for geoscience domains, but other than providing a label for instances, these structures are far removed from capturing the level of detail necessary to empower domain scientists, or knowledge engineers, with useful components for day-to-day meaningful research activities. In this paper we illustrate how a *deep* semantic structure serves to differentiate research products by capturing epistemological commitments of geoscience research artefacts using ontologies throughout the spectrum of formalisation. This deep semantic structure provides the conceptual backbone for geoscientific search, discovery and enquiry.

Acknowledgments. The authors would like to thank Dr. Will Smart and Sina Masoud-Ansari, at the University of Auckland's Centre for eResearch, for their contributions to the Alfred framework used to illustrate the use case in this paper. The authors would also like to acknowledge the generous support of the New Zealand International Doctoral Research Scholarship (NZIDRS), which helped make this research possible.

8 References

1. Brodaric, B., Gahegan, M.: Experiments to Examine the Situated Nature of Geoscientific Concepts. Spatial Cognition and Computation. 7, 61–95 (2007).
2. Crompton, J.: Putting the FOCUS on Data. W3C Workshop on Semantic Web in Oil & Gas Industry. , Houston, USA (2008).
3. Bell, G., Hey, T., Szalay, A.: Beyond the Data Deluge. Science. 323, 1297–1298 (2009).
4. Hey, T., Trefethen, A.E.: Cyberinfrastructure for e-Science. Science. 308, 817–821 (2005).
5. Hey, T., Tansley, S., Tolle, K. eds: The Fourth Paradigm: Data-intensive scientific discovery. Microsoft Research, Redmond, Washington (2009).
6. Sinha, A.K. ed: Geoinformatics: Data to Knowledge. Geological Society of America (2006).

7. Sinha, A.K., Arctur, D., Jackson, I., Gundersen, L. eds: Societal Challenges and Geoinformatics. Geological Society of America, Boulder, CO (2011).
8. Keller, G.R., Baru, C. eds: Geoinformatics: Cyberinfrastructure for the Solid Earth Sciences. Cambridge University Press, Cambridge, UK (2011).
9. Sowa, J.F.: Crystallizing Theories out of Knowledge Soup. In: Ras, Z.W. and Zemankova, M. (eds.) Intelligent Systems: State of the Art and Future Directions. pp. 456–487. Ellis Horwood, New York (1990).
10. Schumm, S.A.: To Interpret the Earth: ten ways to be wrong. Cambridge University Press, Cambridge, UK (1991).
11. Raab, T., Frodeman, R.: What is it like to be a geologist? A phenomenology of geology and its epistemological implications. Philosophy and Geography. 5, 69–81 (2002).
12. Baker, V.R.: Geosemiosis. GSA Bulletin. 111, 633–645 (1999).
13. Kastens, K., Manduca, C.A., Cervato, C., Frodeman, R., Goodwin, C., Liben, L.S., Mogk, D.W., Spangler, T.C., Stillings, N.A., Titus, S.: How Geoscientists Think and Learn. Eos. Trans. AGU. 90, 265–266 (2009).
14. Chamberlin, T.C.: The Method of Multiple Working Hypotheses. Science. 148, 754–759 (1899).
15. Frodeman, R.: Geological reasoning: Geology as an interpretive and historical science. GSA Bulletin. 107, 960–968 (1995).
16. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American. 284, 34–43 (2001).
17. Sinha, A.K., Ludäscher, B., Brodaric, B., Baru, C., Seber, D., Snoke, A., Barnes, C.: GEON: Developing the Cyberinfrastructure for the Earth Sciences. (2003).
18. Babaie, H.A.: Ontological relations and spatial reasoning in earth science ontologies. In: Sinha, A.K., Arctur, D., Jackson, I., and Gundersen, L. (eds.) Societal Challenges and Geoinformatics. pp. 13–27. Geological Society of America, Boulder, CO (2011).
19. Obrst, L.: Ontologies for semantically interoperable systems. Proceedings of the twelfth international conference on Information and knowledge management. pp. 366–369. ACM, New Orleans, LA (2003).
20. Sowa, J.F.: Future Directions for Semantic Systems. In: Tolk, A. and Jain, L.C. (eds.) Intelligence-Based Systems Engineering. pp. 23–47. Springer-Verlag Berlin (2011).
21. Giunchiglia, F., Zaihrayeu, I.: Lightweight Ontologies. In: Liu, L. and Özsu, M.T. (eds.) Encyclopedia of Database Systems. pp. 1613–1619. Springer, US (2009).
22. Sheth, A., Ramakrishnan, C., Thomas, C.: Semantics for the Semantic Web: The Implicit, the Formal and the Powerful. International Journal on Semantic Web & Information Systems. 1, 1–18 (2005).
23. Neuendorf, K.K.E., Mehl, Jr., J.P., Jackson, J.A.: Glossary of Geology. American Geosciences Institute (2011).
24. Deliiska, B.: Thesaurus and domain ontology of geoinformatics. Transactions in GIS. 11, 637–651 (2007).
25. Zhong, J., Aydina, A., McGuinness, D.L.: Ontology of fractures. Journal of Structural Geology. 31, 251–259 (2009).
26. Raskin, R.G., Pan, M.J.: Knowledge Representation in the Semantic Web for Earth and Environmental Terminology (SWEET). Computers & Geosciences. 31, 1119–1125 (2005).
27. Lake, R., Burggraf, D.S., Trninić, M., Rae, L.: Geography Mark-Up Language (GML). John Wiley & Sons, Ltd, West Sussex, England (2004).
28. Sen, M., Duffy, T.: GeoSciML: Development of a generic Geoscience Markup Language. Computers & Geosciences. 31, 1095–1103 (2005).
29. Gahegan, M., Luo, J., Weaver, S.D., Pike, W., Banchuen, T.: Connecting GEON: Making sense of the myriad resources, researchers and concepts that comprise a geoscience cyberinfrastructure. Computers & Geosciences. 35, 836–854 (2009).
30. Sowa, J.F.: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Course Technology Cengage Learning, Boston, MA (2000).
31. Ribes, D., Bowker, G.: Between meaning and machine: Learning to represent the knowledge of communities. Information and Organization. 19, 199–217 (2009).
32. Whitehead, B., Gahegan, M., Everett, M., Hills, S., Brodaric, B.: Fostering the exchange of geoscience resources for knowledge exploration and discovery. Proceedings of eResearch Australasia Conference. p. 2p. , Gold Coast, Australia (2010).
33. Everett, M., Hills, S., Gahegan, M., Whitehead, B., Brodaric, B.: Improving Semantic Interoperability through the Development of a Reusable E&P Earth Science Ontology. American Association of Petroleum Geologists (AAPG) Annual Convention and Exhibition. , Houston, USA (2011).
34. Guoyo, L.: World Atlas of Oil and Gas Basins. WIley-Blackwell, Hoboken, NJ, USA (2010).