Shared earth modeling (April 2013)

'Semantic' approach to oil and gas field modeling proposed in new publication from IFP Energies Nouvelles. 'Knowledge-driven' solution to inform next generation Energistics' Resqml protocol.

A new book by a team of researchers led by Michel Perrin and Jean-François Rainaud at IFP Energies Nouvelles, the French petroleum institute, breaks new ground in geological modeling, proposing a 'semantic' framework embracing model topology, stratigraphic relationships and grids.

Shared Earth Modeling* (SEM), a hefty 350 page work, is subtitled 'knowledge-driven solutions for building and managing subsurface 3D geological models.' For knowledge-driven, read the ubiquitous application of semantic-web ontology-based technology.

In his introduction, Total's Dominique Lefebvre compares the thrust of SEM to the actions of Aureliano Buendia in Gabriel García Márquez' 'One hundred years of solitude' who, in an effort to combat memory loss, labels every object in his village. Likewise, 'ontologies will be the labels of our geological models.' The image is particularly apt in the context of the 'graying workforce!'

SEM's early chapters provide an accessible introduction to earth modeling while introducing the 'knowledge framework' that describes stratigraphic relationships, faults and other objects. Various gridding techniques are discussed in the context of commercial tools and the current state of the art. Particular attention is given to Energistics' Resqml initiative for reservoir model data exchange. Here SEM distances itself from the current Resqml approach (and indeed most current XML-based exchange formats), contrasting their 'data driven' modeling with the 'knowledge-driven' approach of the SEM. This difference is illustrated in a chapter on seismic interpretation where faults and reflectors are viewed as ontology 'instances' and manipulated with the Stanford 'Protégé' editor. Several SEM authors were involved with Resqml although the current standard eschews the semantic vision.

The team is now working on a Resqml 2 proposal which will introduce simple ontological concepts and act as a bridge to a future semantic world. SEM winds up presenting several semantic applications that run atop of the framework including E-Wok (Oil ITJ Jan 09) and Petrolege and Stratalege developed at the University of Rio Grande do Sul, Brazil.

Ontology management is illustrated with a workflow editor running atop of an Ontodb database of data from Total's Alwyn field. A 'WebLab' platform is proposed to pull all the semantic strands together. All in all, SEM is a fascinating read. No oil and gas research department should be without a copy. More on SEM in this month's editorial.

* *Editions Technip 2013. ISBN 9782710810025.*

Click here to comment on this article

Reflections on Shared Earth Modeling and the semantic web (April 2013)

Neil McNaughton delves deeper into the subject of this month's lead. Can a book be news? It can, if it proposes a novel approach to intractable problems like interoperability and data management.

Our reports in this month's issue on the 2013 Microsoft global energy forum and the Houston PPDM data management symposium are at the same time interesting and disappointing. They are interesting because it is always good to hear how folks are using technology to actually achieve stuff. They are disappointing in that, despite IT's constant re-invention of itself, there is nothing much really that is *new*.

It was therefore interesting to receive a book for review that really does offer 'something new.' So much new stuff in fact that having devoted this month's lead to a brief presentation of Shared Earth Modeling* (SEM) I propose to continue in the same vein.

But first, can a book be news? You bet it can! If the concepts developed in SEM pan out, the whole industry will be revolutionized. Not just modeling applications and workflow, but many data management issues will be fixed too. But semantic technology, despite a decade of effort and backing from the great and good, especially Tim Berners-Lee (TBL), has singularly failed to set the world alight.

When I picked up SEM, I first thought, oh dear, here we go again. 20 or so authors—this is just going to be another collection of papers which some editor has thrown together. Wrong. SEM is a well written coherent whole which does a great job of introducing a tough subject and in tying a lot of disparate themes together. My only serious criticism of SEM is that it follows the inexplicable French publishing tradition of not having an index.

My next question for SEM (the concept rather than the book) is, will it work? The ideal means of introducing a new technology to a vertical such as oil and gas is just to steal it from another vertical. Thus was signal processing lifted from the telecom business and repurposed in geophysics back in the mid 20th century. Like wise the digital oilfield works because is uses horizontal technology from the process control business.

But while the intellectual property 'theft' paradigm has worked in the past, it is not the only game in town. The oil and gas industry has developed bags of its own IP –in specialist fields like seismic and deepwater. Why shouldn't the geo-modeling community lead the way in industrial application of semantics?

SEM is about more than geometry. A chapter on the use of ontologies for analyzing data in natural language is particularly interesting. If you think that this sort of stuff is pie in the sky, consider the case of IBM's Watson. The Jeopardy-winning machine runs on a stack of open source software that includes 'natural language processing, semantic analysis, information retrieval, automated reasoning and machine learning.'

Now Watson does not, as far as I can tell, use exactly the same 'semantic web' technology as SEM but it comes close. Watson has been reported as using various public 'sem web' data sources including Yago and dbPedia. The latter is a semantic interface to Wikipedia. Yago is another huge knowledge base that is generated automatically from Wikipedia, GeoNames and WordNet. Yago claims to hold 447 million facts about 9.8 million 'entities.'

The great thing about Watson though is that it is such a compelling use case. Answer a Jeopardy question, everyone can relate to that! Unfortunately, the same cannot be said for many of the component technologies of the semantic web. TBL's promise of a web of data and machine to 'machine interaction' were followed by a long and laborious period during which semantic standards have slowly evolved. Today, when you check out some of the technologies that underpin the SEM, they appear lacking in purpose.

But this is where SEM, the book, is so good—it confronts the obscure R&D stuff with the real world problem set of earth modeling and shows what might be a compelling application of the new technology.

What's next? Clearly there is a lot of mileage in SEM for the standards bodies. Energistics, PPDM for a start, but also pretty well all involved in the Standards Leadership Council. Someone needs to take a good look at the SEM approach and compare it with the only other major semantic standard in oil and gas, ISO 15926. And of course there is Resqml—pretty much the focus of SEM.

SEM lead author Jean-François Rainaud told us the following *'Resqml was partly inspired by the work done by the team behind SEM. But the Resqml group had to be realistic and stick with the technologies deployed in today's geomodeling application software. If we get too far ahead of ourselves, we would have a hard time getting take-up. We are currently working on Resqml V2 and will be introducing semantic relationships between geological objects and relationships. This will allow us to codify relationships such as 'Fault F2, on its hanging wall side interrupts horizon interpretation H1 on both sides.' We are also working to ensure that the Resqml V2 data model is structured so it can later incorporate more of the semantic concepts described in SEM. We have a paper in preparation** for the EAGE in London where we will be presenting Resqml V2.'*

Perhaps the proof of the pudding will come when the geometry component of say, ISO 15926, can be picked up in a semantic editor and reused alongside an ontology from the SEM. Even if such interoperability doesn't materialize, the idea of labeling objects with a tag that means something is probably better than using a Windows GUID as has been suggested elsewhere. Connecting said objects with domain-specific concepts would seem to make more sense than shoehorning everything into a database.

There's more on semantics in this issue—from OneGeology (page 4) and from PCA/Mimosa (page 5). Enjoy!

* *Editions Technip 2013. ISBN 9782710810025.*

** *SPE 164794-MS (authors from Total, Texas A&M, IFP, Paradigm, Schlumberger and Energistics).*

@neilmcn