W3C Workshop on Semantic Web in Energy Industries; Part I: Oil & Gas — A Position Paper

Clark & Parsia, LLC (http://clarkparsia.com/)

30 September 2008

1 Introduction

Semantic Web technology ("SW tech"), based primarily on W3C and IETF standards, has to date been applied most often in fields like health care life sciences, enterprise IT, financial services, and defense. The movement from R&D into core production applications has begun in earnest and, despite the worsening economic picture these days, we can expect to see increasing numbers of deployed, stable systems built on SW tech that become central to real operations. In short, industries that adopted SW tech early — HCLS, defense, and financial — have already begun to reap the business benefits of their investment in SW tech, in some cases dramatically.

What we have not seen yet, however, is an equivalent investment in SW tech from the energy sector, either from core players in that sector, or even from green, alternative, or fringe players. In fact, this reluctance to invest so far is mirrored in most of the surrounding "heavy industry" fields like construction, architecture, and engineering disciplines (process, civil, chemical, petroleum, etc.).

And yet all of these fields, despite their reliance on real-world technologies beyond IT, are knowledge disciplines where one may reasonably expect to see SW tech paying off in concrete benefits. And, further, most of these fields have been undergoing serious renaissance in the degree to which they are investing in cutting-edge IT initiatives. In this regard, consider an article from *The Economist*, 5 June 2008, "From blueprint to database,"[1] which describes the rise in Building Information Modeling (BIM) systems in the architectural, design, and building maintenance fields, as well as the new Model Based Systems Engineering[2] initiatives in both academia and industry. Add to these ISO 15926, POSC Caesar, Product Modeling Ontology (PMO, from the SWOP, an EU-funded R&D project) — suddenly

[1] See `http://snipurl.com/3yoce`

[2] See `http://www.mbse-org.org/`

the picture looks quite different.[3]

Clearly something is happening in the heavy industries, including O&G. But what is it and what should we do about it?

1.1 Answering Two Questions

To answer that question, and thus to orient ourselves to the larger trends which the W3C's Oil & Gas Workshop is responding to, we have to ask and answer two questions:

1. What can you do with SW tech?
2. what kinds of IT problems do O&G companies have?

1.2 SW Tech Capabilities

The most general characterization of SW tech is that it's a family of related technologies which are abstract and declarative — rather than concrete and procedural or imperative — that are consistent with and finely-tuned for use in Web applications, both public and private, and Web-like information systems for the purpose of machine legibility, information exchange, and information reuse.

In other words, SW tech is a set of standards, plus a lot of software infrastructure, for doing knowledge representation on, for, and with the Web generally conceived.

This cluster of tools and standards has been used among early adopters for the following kinds of application: information integration and decision support and analysis.

Information Integration A major thrust of SW tech application and ROI is in information integration efforts, which constitute a significant portion of all IT spending globally. For heterogeneous enterprises with legacy and new systems, as well as diverse trading and business partners, and complex data flows, SW tech provides a set of declarative and abstract technologies with which to manage the complexity of integration efforts.

OWL ontologies, in particular, are growing in popularity as a means of building an abstract, declarative, and high level description of a problem or data domain, which is then used as a kind of "abstract data contract" or "semantic integration hub" in order to facilitate integration. In short, systems that need to be integrated target their efforts at the abstract data model layer, as expressed in OWL, such

[3]The point isn't that these initiatives are using SW tech; but that they are the sort of advanced, nex-gen IT efforts that either could benefit from SW tech or are sufficiently similar to SW tech, such that if they are in play, then SW tech is in play, too.

that many changes in their internal data models are simply opaque to integration partners.[4]

Further, automated reasoning tools are used to demonstrate conceptual consistency between the various ontologies, or data domain formalizations, as well as to provide other reasoning services useful to integration efforts.[5] One such reasoner for OWL DL ontologies is Clark & Parsia's open source tool Pellet.[6]

Decision Support & Analysis In addition to its role in information integration problems, SW tech also plays an increasingly important role in another class of advanced application: decision support and analysis (hereinafter, "DSA").

A DSA application has as its chief use case aiding a human decision maker to comprehend a complex body of information and, based on that understanding, to guide or suggest an appropriate course of action.

SW tech is useful for DSA applications because its core standards allow people to model complex data and problem domains in such a way that users can gain actionable insights into complex problem areas, including configuration problems, automated planning and logistics, risk management and assessment, support for basic and applied scientific research, etc.

Again, automated reasoning tools based on OWL, like Pellet, play a crucial role in both verifying that conceptual formalizations of problem domains are logically consistent and expressive enough to capture reality accurately, as well as providing other reasoning services that are required to build DSAs.

Synergies Between Integration and Analysis Finally, it should be noted that in large enterprises, there is very often a feedback loop between these two sorts of application: DSA applications require access to the relevant data, which is provided by bootstrapping integration efforts. And, contrariwise, one the initial investment has been made to integrate some data domains, that opens up the possibility of providing new and improved sorts of DSA applications over the integrated data.

In that regard, using SW tech as a substrate for both kinds of application often leads to a kind of network effect whereby the sum of the whole is greater then the sum of the parts. New DSAs that would not have been feasible under some budget or schedule constraints become possible because of existing information integration efforts. And the success of some new DSA provides reason to integrate even more unintegrated data sources in order to extend the reach of the DSA tools.

[4] For example, this is the approach taken with Clark & Parsia's POPS, a workforce analytics app, that we developed for NASA's expertise location service. See `http://www.w3.org/2001/sw/sweo/public/UseCases/Nasa/` for more details.

[5] For example, realization of instances, schema-level reasoning, expressive query answering using SPARQL and SPARQL-DL, as well as higher level services like ontology alignment and fusion.

[6] See `http://pellet.owldl.com/`.

1.3 O&G IT Challenges

As it turns out, O&G companies — like most complex enterprises — have both sorts of problem described above in no short supply. Over the lifecycle of products produced in the process industry, for example, integrated data flows decreases costs and increases productivity. And the outputs of process plants, for example, are sufficiently complex that, as we discuss below, cottage industries to support and manage them are underway, and SW tech, particularly OWL, is playing an important role in those efforts.

In short, O&G players face integration, analysis, configuration management problems at large scales of both data and complexity, as parts of the global economy that are not only strategically vital, but also under increasing pressure to deliver increasing output with what seems to be decreasing input. O&G is ripe for a spike in productivity based on information technology, and it appears that SW tech will be a key player in that spike.

2 Our Focus

Our focus at Clark & Parsia LLC — as a boutique firm active in SW tech for several years — is on semantic infrastructure, focusing primarily on OWL software and tools to empower developers to successfully use OWL and SW tech in integration and DSA efforts.

2.1 Semantic Infrastructure for ISO 15926

We anticipate many position papers to the present workshop will include lengthy, detailed discussion of ISO 15926, of EXPRESS, and of existing efforts to use SW tech to build solutions for this domain. Given our expertise in SW tech, and our comparative lack of expertise in ISO 15926, we won't say very much about the latter.

We will say, however, that, at both the level of software systems and business partnerships, we are planning adaptations of our existing infrastructure tools — including a variety of OWL reasoners, APIs, and associated tools for ontology development and maintenance — to support O&G efforts in the ISO 15926 ecosystem to adapt SW tech, for both integration and DSA efforts.

In particular, we think the ongoing W3C Working Group to update the OWL family of standards provides a market inflection point where new players and new participants will have a chance to apply OWL and SW tech to new markets, including O&G. And, more specifically, several aspects of OWL2, including the profiles and new features, are applicable to both ISO 15926 and, generally, contribute to the maturation and industrialization of the SW tech such that it will be a cornerstone of O&G uptake.

2.2 On the Horizon

Finally, while OWL2 — together with other parts of the SW tech stack, to be sure — will empower the first generation of O&G SW applications, given our role as commercializers of early-stage R&D, we're looking forward to future standardization efforts to extend the reach of OWL even further, and we anticipate all of these "on the horizon" aspects will be directly relevant to O&G efforts:

- integrity constraints semantics for OWL[7]
- finer-grained modes of closed world reasoning
- concrete domains, equational reasoning, and quantities and units in OWL[8]
- O&G-specific SWRL builtins, and some consideration of SWRL's standardization status
- additional expressivity in OWL DL relevant to O&G like description graphs
- increasing cooperation between OMG-UML and W3C-OWL at the level of interchange and conversion

It is premature to contemplate the next generation of OWL standardization; but it is not too early to build into standards-compliant systems features that solve customer needs and are potential candidates for future standardization efforts. We consider the aforementioned extensions of OWL2 will not only be useful in ISO 15926 efforts but in other problem domains as well.

3 Conclusion

To summarize:

- The O&G industry is poised to realize big productivity gains based on SW tech
- ISO 15926 and OWL infrastructure support thereof will be a key enabler
- O&G needs integration and DSA application support from SW tech
- Clark & Parsia LLC adapting its infrastructure offerings for O&G use cases and requirements

[7] See `http://clarkparsia.com/files/pdf/ic-owled2008-eu.pdf` for more about integrity constraints and OWL.

[8] See `http://clarkparsia.com/files/pdf/units-owled2008-eu.pdf` for more about equational reasoning and datatype modeling in OWL.