Three Leading Arguments for Semantic Technologies

by Mike Bergman - Monday, January 14, 2013

http://www.mkbergman.com/1605/three-leading-arguments-for-semantic-technologies/

Part 2 in the Enterprise-scale Semantic Systems Series

Those involved with the semantic Web are passionate as to why they are involved. This passion and the articulateness behind it are notable factors in why there is indeed a 'semantic Web community.' Like few other fields -- perhaps including genomics or 3D manufacturing -- semantic technologies tend to attract exceptionally smart, committed and passionate people.

Across this spectrum of advocates there are thousands of pages of PDFs and academic treatises as to semantic this or semantic that. There is gold in these hills, and much to mine. But, both in grants and in approaching customers, it always comes down to the questions of: What is the argument for semantic technologies? What are the advantages of a semantic approach? What is the compelling reason for spending time and money on semantics as opposed to alternatives?

Fred Giasson and I at Structured Dynamics feel we have done a pretty fair job of answering these questions. Of course, it is always hard to prove a negative -- how do the arguments we make stack up against those we have not? We will never know.

Yet, on the other hand, we have found dedicated customers and steady and growing support from the arguments we do make. At least we know we are not scaring potential customers away. Frankly, we suspect our market arguments are pretty compelling. While we discuss many aspects of semantic technologies in our various writings and communications, we have also tended to continually hone and polish our messages. We keep trying to focus. Fewer points are better than more and points that resonate with the market -- that address the "pain points" in common parlance -- have the greatest impact.

It is also obvious that the arguments an academic needs to make to a funding agency or commission are

much different than what is desired by commercial customers. (Not to mention the US intelligence community, which is the largest -- yet silent -- funder of semantic technologies.) Much of what one can gain from the literature is more of this academic nature, as are most discussions on mailing lists and community fora. We distinctly do not have the academic perspective. Our viewpoint is that of the enterprise, profit-making or non-profit. Theory takes a back seat to pragmatics when there are real problems to solve.

Our three main selling points to this enterprise market relate to data integration and interoperability; search and discovery; and leveraging existing information assets with low risk. How we paint a compelling picture around these topics is discussed for each point below. We conclude with some thoughts about how and the manner we communicate these arguments, perhaps representing some background that others might find useful in how they may make such arguments themselves.

"Semantic Technologies Enable Data Integration and Interoperability"

As I have experienced first hand and have argued many times [1], the Holy Grail of enterprise information technology over the past thirty years has been achieving true data integration and interoperability. It is, I believe, the primary motivating interest for most all IT efforts not directly related to conventional transaction systems. Yet, because of this longstanding and abiding interest, enterprise IT managers react with justifiable skepticism every time new advances in interoperability are claimed.

The claims for semantic technologies are not an exception. But, even in its positioning, there is something in the descriptive phrasing of "semantic technologies" that resonates with the market. Moreover, to overcome the initial skepticism, we also tend to emphasize two bolstering arguments promoting interoperability:

1. Semantic technologies matched with natural language (NLP) techniques work to integrate unstructured data, finally incorporating the 80% of enterprise information locked up in documents and overcoming the limitations of manually assigned tags, and
2. The RDF data model is capable of capturing any existing data relationship, and ontologies are capable of capturing any existing information schema.

Since these are two of the core aspects to data integration and have heretofore been limited with conventional approaches, and since they can be demonstrated rather quickly, trust can be placed into the ultimate interoperability argument.

In the end, the ability of semantic technologies to promote rather complete data integration and interoperability will prove to be its most compelling rationale. Yet, achieving this with semantic technologies will require more time and broader scope than what has been instituted to date. By starting smaller and simpler, a more credible entry argument can be made that also is on the direct pathway to interoperability benefits.

"Semantic Technologies Improve Search and Discovery"

On the face of it, search engines and the search function are nearly ubiquitous. Further, search is generally

effective in eventually finding information of interest, though sometimes the process of getting there is lengthy and painful.

This inefficiency results because search has three abiding problems. One, there is too much ambiguity in what kind of thing is being requested; disambiguation to the context at hand is lacking. Second, there is a relative lack of richness in the kinds of relationships between things that are presented. We are learning through Web innovations like Wikipedia or the Google [Knowledge Graph](#) that there are many attributes that can be related to the things we search. The natural desire is to now see such relationships in enterprise search as well, including some of this public, external content. And, third, because of these two factors, search is not yet an adequate means for discovering new insights and knowledge. We see the benefits of serendipitous discovery, but we have not yet learned how to do this with purpose or in a repeatable way.

More often than not customers see search, with better display of results, at the heart of the budget rationale for semantic projects. The graph structures of semantic schema means that any node can become an entry point to the knowledge space for discovery. The traversal of information relationships occurs from the selection of predicates or properties that create this graph structure in the first place. This richness of characterization of objects also means we can query or traverse this space in multiple languages or via the full spectrum by which we describe or characterize things. Semantic-based knowledge graphs are potentially an explosion of richness in characterization and how those characterizations get made and referred to by any stakeholder. Search structure need not be preordained by some group of designers or information architects, but can actually be a reflection of its user community. It should not be surprising that search offers the quickest and most visible path to conveying the benefits of semantic technologies.

These arguments, too, are a relatively quick win. We can rapidly put in place these semantic structures that make improved search benefits evident. There are two nice things about this argument. First, it is not necessary to comprehensively capture the full knowledge domain of the customer's interests to show these benefits. Relatively bounded projects or subsets of the domain are sufficient to show the compelling advantages. And, second, as this initial stakehold gets expanded, the basis for the next argument also becomes evident.

"Semantic Technologies Leverage Existing Assets with Low Risk"

I have often spoken about the incremental nature of how semantic technologies might be adopted and the inherent benefits of the [open world](#) mindset. This argument is less straightforward to make since it requires the market to contemplate assumptions they did not even know they had.

But, one thing the market does know is the brittleness and (often) high failure rates of knowledge-based internal IT projects. An explication of these causes of failure can help, via the inverse, to make the case for semantic technologies.

We know (or strongly suspect), for example, that these are typically the causes of knowledge-based IT failures:

- Too broad a scope or the need to embrace too much of the information basis of the domain

- Changing knowledge and circumstances that causes initial design imperatives to change over the course of a project
- High visibility for multiple audiences and stakeholders, and no workable means for finding a common view or consensus as to objectives (let alone terminology) for the project amongst these stakeholders.

Getting recognition for these types of failures or challenges creates the opening for discussing the logic underpinnings of conventional IT approaches. The conventional closed-world approach, which is an artifact of using information systems developed for transaction and accounting purposes, is unsuited to open-ended knowledge purposes. The argument and justification for semantic technologies for knowledge systems is that simple.

The attentive reader will have seen that the first two arguments presented above already reify this open world imperative. The integration argument shows the incorporation of non-structured content as a first-class citizen into the information space. The search argument shows increased scale and richness of relationships as new topics and entities get added to the search function, all without adversely impacting any of the prior work or schema. For both arguments, we have expanded our scope and schema alike without needing to re-architect any of the semantic work that preceded it. This is tangible evidence for the open world argument in the context of semantic technologies applied to knowledge problems.

These evidences, plus the fact we have been increasingly incorporating more sources of information with varied structure, most of which already exists within the enterprise's information assets, shows that semantic technologies can leverage benefits from existing assets at low risk. At this point, if we have told our story well, it should be evident that the semantic approach can be expanded at whatever pace and scope the enterprise finds beneficial, all without impacting what has been previously implemented.

Actually, the argument that semantic technologies leverage existing assets with low risk is perhaps the most revolutionary of the three. Most prior initiatives in the enterprise knowledge space have required wholesale changes or swapping out of existing systems. The unique contribution of semantic technologies is that they can achieve their benefits as a capability layered over existing assets, all without disruption to their existing systems and infrastructure. The degree to which this layering takes place can be driven solely by available budgets with minimal risk to the enterprise.

Ambassadors and Archivists, as well as Entrepreneurs

There are, of course, other messages than can be made, and we ourselves have made them in other circumstances and articles. The three main arguments listed herein, however, are the ones we feel are most useful at time of early engagement with the customer.

Our messages and arguments gain credibility because we are not just trying to "sell" something. We understand that semantic technologies and the mindsets behind them are not yet commonplace. We need to be ambassadors for our passion and work to explain these salient differences to our potential markets. As later parts in this series will discuss, with semantic technologies, one needs to constantly make the sale.

The best semantic technology vendors understand that market education is a core component to

commercial success. Once one gets beyond the initial sale, it is a constant requirement to educate the customer with the next set of nuances, opportunities and technologies.

We acknowledge that vendors have other ways to generate "buzz" and "hotness." We certainly see the consumer space filled with all sorts of silliness and bad business models, But our pragmatic approach is to back up our messaging with full documentation and market outreach. We write much and contribute much, all of which we document on vehicles such as our blogs, commercial Web site, or TechWiki knowledge base. New market participants need to learn and need to be armed with material and arguments for their own internal constituencies. Insofar as we are the agents making these arguments, we also get perceived as knowledgeable subject matter experts in the semantic technology space.

I have talked in my Of Flagpoles and Fishes article of the challenges of marketing to a nascent market where most early sales prospects remain hidden. At this stage in the market, our best approach is to share and communicate with new market prospects in a credible and helpful way. Then, we hope that some of those seeking more information are also in a position to commission real work. If we are at all instrumental in those early investigations, we are likely to be considered as a potential vendor to fulfill the commercial need.

Of course, each new engagement in the marketplace means new lessons and new applications. Thus, too, it is important that we become archivists as well. We need to capture those lessons and feed them back to the marketplace in a virtuous circle of learning, sharing, and further market expansion. Targeted messages delivered by credible messengers are the keys to unlocking the semantic technologies market.

NOTE: This is part of an ongoing series on enterprise-scale semantic systems (ESSS), which has its own category on this blog. Simply click on that category link to see other articles in this series.

_____

[1] Simply conduct a search on http://www.mkbergman.com/?s=interoperability+integration to see how frequently this topic is a focus of my articles.

—————————————————————————

PDF generated by *AI3:::Adaptive Information* blog