



The Future of the Semantic Web

By
Franz Inc.



Talk Outline

- This is a business and marketing oriented talk.
- Other speakers are covering technical issues in great detail.
- In particular the presentations from Mr. S. Koide, Dr. K Hashimoto and Mr. R Hodgson are quite comprehensive and valuable.
- In this talk we are making tentative and speculative predictions about the future – and all of them might be wrong....



The Semantic Web – What is it anyway?

- Many descriptions have been given. We like this one:
- The New York Times reporter – John Markoff - in a November 11, 2006 article, writes:

SAN FRANCISCO, Nov. 11 — From the billions of documents that form the World Wide Web and the links that weave them together, computer scientists and a growing collection of start-up companies are finding new ways to mine human intelligence. Their goal is to add a layer of meaning on top of the existing Web that would make it less of a catalog and more of a guide — and even provide the foundation for systems that can reason in a human fashion. That level of artificial intelligence, with machines doing the thinking instead of simply following commands, has eluded researchers for more than half a century.



Semantic Web – What is it?

- Another view – Almost everyone is familiar with Google – and understands how Google’s keyword search works.
- Google searches can be surprisingly frustrating – too often, too much information comes back – most of which is just wrong.
- Google may be the first example of a very highly-valued company whose products mostly return the wrong answer!
- So here is our first speculative prediction – With the Semantic Web, you will be able to “ask a question” – at first by writing it, and still later by speaking it – and expect to get an accurate response.



Semantic Web – What is it?

- Web 2.0 – today describes web-based applications sharing data and results – see: <http://www.paulgraham.com/web20.html> and http://en.wikipedia.org/wiki/Web_2 for more details of how this came about.
- Recently, the term Web 3.0 has come into being – and describes a (future) World Wide Web coupled with next generation search capabilities – enabling “smart search” – in contrast to “keyword search” provided today.
- Realizing this idea means finding ways to add meaning to existing web pages – a difficult and possibly impractical task – and developing new software based on Artificial Intelligence. And these are just a few of the problems.
- Debates are furiously raging on how these goals might be achieved



Semantic Web – What Can We Do Today?

- Today tools exist capable of extracting meaning from some types of web pages.
- So one approach might be this: Crawl the web, extracting meaning from web pages, while duplicating the web (on disk) but this time with structured meaning along with the web page content, and finally offering a next generation search engine.
- What would this get you? Say 90% of search today is keyword search – yet 80% of the value might lie in the remaining 10% not amenable to keyword search – let's call this structured search.
- Keyword search (Google, etc.) simply fails when search must be accurate, repeatable and auditable – so the market could begin to emerge from the access point of enabling structured search.



Semantic Web – Who needs it Now?

- Which markets might first adopt the Semantic Web?
- Government Intelligence Communities really need it now – Relational Databases lack the ability to analyzed linked information easily.
- Life Sciences – are currently working hard to integrate semantic information with biological data
- Speculative Predictions – At least two other markets are ready:
 - Regulatory and Compliance applications (Stock market regulation, insurance regulation, financial regulation – any kind of compliance application can use this today)
 - Enterprise Application Integration – This is potentially really big – and database companies are unable to solve it - and the technology exists now to solve problems in this market.



Semantic Web – What will happen next?

- Here is another speculative prediction – and this is one that Franz has made a large technological bet on....
- The underlying infrastructure for storing and manipulating information seems likely to require a database infrastructure referred to as a “Triple Store”.
- So what is a Triple Store? Many complicated descriptions are offered – with complicated terms like RDF Triple Store, graph database, named graphs, etc. There is a simple way to understand it though.
- Conceptually, any sentence in any language can be mapped into a series of three pieces of information – Subject, Object and Predicate.



Semantic Web – What will happen next?

- For example – Consider the very simple sentence: Franz is a company. Franz is the Subject, “is a” is the Predicate and the object is “company”.
- The important point is this: Virtually any written information can be mapped into a series of “triples” – in the form of Subject, Object and Predicate.
- Once written information gets represented as triples, we can easily add meaning to the triples.
- Once meaning is added, programs manipulating the meaning become possible.
- Franz has invested in the development of a huge “triple” store – one that can potentially contain all of the World Wide Web’s information.



Semantic Web – Conclusions

- It is very clear that 5 years from now much more intelligent search capabilities will be available. We are already drowning in data – and data comes online in ever larger streams.
- Business opportunities are plentiful – if history is any guide, it is very unlikely that any of today's major search engine companies will invent the next generation technologies. It is far more likely they will buy it from others.
- The Semantic Web and the term Web 3.0 are likely to be used interchangeably. And these terms will appear everywhere in the press.
- Triple Stores are likely to be the underlying representational structure of the semantic information.



Semantic Web – Conclusions

- Keyword search fails for at least two markets capable of supporting the Semantic Web – Intelligence Agencies and Biological Sciences.
- Two other potential markets are Regulation and Compliance and the Enterprise Application Integration markets