

Semantic Web Patterns: A Guide to Semantic Technologies

Written by [Alex Iskold](#) / March 25, 2008 3:20 PM / [32 Comments](#)

[« Prior Post](#) [Next Post »](#)



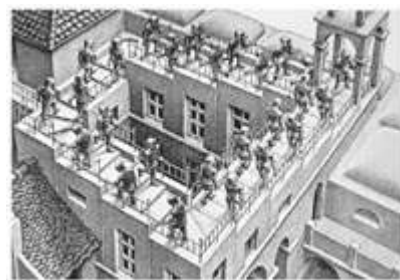
In this article, we'll analyze the trends and technologies that power the Semantic Web. We'll identify patterns that will emerge, classify the different trends, and peak into what the future holds.

In a recent interview [Tim Berners-Lee pointed out](#) that the infrastructure to power the Semantic Web is a complex one. ReadWriteWeb's founder, Richard MacManus, even [picked it](#) to be the number one trend in 2008. And while there are a few bits of infrastructure now in place, but we are also seeing startups and larger corporations working hard to build on top of this sophisticated set of technologies.

The Semantic Web means many things to different people, because there are a lot of pieces to it. To some, the Semantic Web is about data, where information is represented in RDF and OWL. Some people replace RDF with Microformats. Other people think the Semantic Web is about web services, while for many it is about artificial intelligence - computer programs solving complex problems that are out of our reach. And business people always redefine the problem in terms of end user value, saying they want to have simple and tangible applications for consumers and enterprises.

The disagreement is not accidental, because the technology and concepts are broad. Much is possible and much is not.

1. Bottom-Up and Top-Down



We have written a lot about the different approaches to the Semantic Web - the [bottom-up](#) approach and the new [top-down](#) one. The bottom-up approach is focused on annotating existing web pages, using RDF, so that it is machine readable. The top-down approach is focused on deriving information in existing web pages, as-is, to derive meaning automatically. Both approaches are making good progress.

A big win for the bottom-up approach was [recent announcement](#) from Yahoo! that they are going to support RDF and microformats. This is a win-win-win for publishers, for Yahoo! customers - publishers now have an incentive to annotate information because Yahoo! is taking advantage of it, and users will then see better, more precise results.

Another recent win for the bottom-up approach was the [announcement of the Semantify](#) web service from [Dapper](#) (previously [Dapper](#)) offering will enable publishers to add semantic annotations to existing web pages. The more tools like Semantify there will be for publishers to annotate pages. Automatic annotation tools combined with the incentive to annotate the the bottom-up approach more compelling.

But even if the tools and incentive exists, to make the bottom-up approach widespread is difficult. Today, the market can understand information as is, without asking people to fully comply with W3C standards of SEO optimization. Top-down semantic tools are focused on dealing with imperfections in existing information. Among them are the natural language processing tools that do entity extraction - such as the [Calais](#) and [TextWise](#) APIs that recognize people, companies, and places in documents; vertical search engines, like [ZoomInfo](#) and [Spock](#), which mine the web for people; technologies like [Dapper](#) which recognize objects in web pages; and [Yahoo! Shortcuts](#), [Snap](#) and [SmartLinks](#), which recognize objects in text.

[Disclosure: Alex Iskold is founder and CEO of AdaptiveBlue, which makes BlueOrganizer and SmartLinks.]

Top-down technologies are racing forward despite imperfect information. And, of course, they benefit from the bottom-up as well. The more annotations there are, the more precise top-down technologies will get - because they will be able to work with structured information as well.

2. Annotation Technologies: RDF, Microformats, and Meta Headers



Within the bottom-up approach to annotation of data, there are several choices for annotation. powerful, and in fact each approach is a tradeoff between simplicity and completeness. The most powerful approach is *RDF* - a powerful, graph-based language for declaring things, and attributes and relationships. In a simplistic way, one can think of RDF as the language that allows expressing truths (type expression), Alex HAS a brain (attribute expression), and Alex IS the father of Alice, Lilly (relationship expression). RDF is powerful, but because it is highly recursive, precise, and mathematically sound.

At present, most use of RDF is for interoperability. For example, the medical community uses many different databases. Because the information is normalized, the databases that were previously silos can now be brought together and correlated. In general, in addition to semantic soundness, the major benefit of RDF is interoperability, particularly for enterprises, as we will discuss below.

Microformats offer a simpler approach by adding semantics to existing HTML documents using specific CSS styles that are compact and is embedded inside the actual HTML. Popular microformats are hCard, which describes personal and business information, hReview, which adds meta information to review pages, and hCalendar, which is used to describe events.

Microformats are gaining popularity because of their simplicity, but they are still quite limiting. There is no way to express complex hierarchies, which the classic semantic community would say is critical. The other issue is that microformats are limited because the focus is to keep the annotations to a minimum. This, in turn, brings up another question of whether the current view (HTML) is a good idea. The question is: what happens if the underlying data changes when someone modifies the HTML document? Nevertheless, despite these issues, microformats are gaining popularity because they are simple and currently used by Flickr, Eventful, and LinkedIn; and many other companies are looking to adopt microformats, per a recent Yahoo! announcement.

An even simpler approach is to put meta data into the meta headers. This approach has been around for a while but has not been widely adopted. As an example, the New York Times recently launched extended annotations for its pages. The benefit of this approach is that it works great for pages that are focused on a topic or a thing. For example, a news article with a set of keywords, geo location, date, time, people, and categories. Another example would be for book pages where you've been putting book information into the meta headers, describing the author, ISBN, and category of the book.

Despite the fact that all these approaches are different, they are also somewhat complimentary; and each of them adds to the annotations there are in web pages, the more standards are implemented, and the more discoverable and powerful the web becomes.

3. Consumer and Enterprise



Yet another dimension of the conversation about the Semantic Web is the focus on consumer and enterprise applications. In the consumer arena we have been looking for a *Killer Application* that delivers tangible and simple consumer value. People simply do not care that a product is based on the Semantic Web, all they are looking for is utility and usefulness.

Up until recently, the challenge has been that the Semantic Web is focused on rationalizing information to make it machine readable. The promise was that once the web is annotated and the web becomes one big giant RDF database, then exciting consumer applications will come. The skeptics, however, have been pointing out that first there needs to be a compelling reason for consumers to use the Semantic Web.

Some consumer applications based on the Semantic Web: generic and vertical search engines, social networking shortcuts and previews, personal information management systems, semantic browser extensions. These applications are in their early days and have a long way to go before being truly compelling for the average web user. Still, even if these applications succeed, consumers will not be interested in the underlying technology - so there is really no marketing play for the Semantic Web in the consumer space.

Enterprises are a different story for a couple of reasons. First, enterprises are much more used to techno speak. The marketing speak of semantic technologies translates into being intelligent and that, in turn, is good marketing. 'Our products are better because we use the Semantic Web' sounds like a good value proposition for the enterprise.

But even above the marketing speak, RDF solves a problem of data interoperability and standards. This "Tower of Babel" has been in existence since the early days of software. Forget semantics; just a standard protocol, a standard way to communicate between two programs, is hugely valuable in the enterprise.

RDF offers a way to communicate using XML-based language, which on top of it has sound mathematical elements. This sounds great, and even the complexity of RDF is not going to stop enterprises from using it. However, there might stop it - scalability. Unlike relational databases, which have been around for ages and have been optimized, databases are still not widespread. In general, the problem is in the scale and querying capabilities. Like object-oriented technologies of the late nineties, XML-based databases hold a lot of promise, but we are yet to see them in action.

4. Semantic APIs



With the rise of Semantic Web applications, we are also seeing the rise of Semantic APIs. These web services take as an input unstructured information and find entities and relationships. To think of these services is mini natural language processing tools, which are only a small subset of the language.

The first example is the Open Calais API from Reuters that we have covered in two previous articles. This service accepts raw text and returns information about people, places, and organizations in a document. The output not only returns the list of found matches, but also specifies the location where the information is found. Behind Calais is a powerful natural language processor developed by Clear Forest (now owned by Reuters), which relies on algorithms and machine learning to extract entities out of text. According to Reuters, Calais is extensible, and it is just a matter of time before new entities will be added.

Another example is the *SemanticHacker API* from TextWise, which is offering a *one million dollar prize* for the best competing application developed on top of it. This API classifies information in documents into categories called semantic spaces. For a document, it outputs entities or topics that the document is about. It is kind of like Calais, but also delivers a topic hierarchy where actual objects are leaves.

Another semantic API is *offered by Dapper* - a web service which facilitates the extraction of structure from unstructured data. Dapper works by enabling users to define attributes of an object based on the bits of the page. For example, a book page can define where the information about author, isbn and number of pages is on a typical book page and the Dapper API can create a recognizer for any page on the publisher site and enable access to it via REST API.

While this seems backwards from an engineering point of view, Dapper's technology is remarkably useful in the real world scenario, for web sites that do not have clean APIs to access their information, even non-technical people can build applications with Dapper. This is a powerful way of quickly *turning web sites into web services*.

5. Search Technologies



Perhaps the first significant blow to the Semantic Web has been the inability thus far to achieve a semantic search. The premise that semantical understanding of pages leads to vastly better search results is still unproven. The two main contenders, *Hakia* and *PowerSet*, have made some progress, but not enough to compete with Google. That Google's algorithm, which is based on statistical analysis, deals just fine with complex queries for people, cities, and companies. When asked *What is the capital of France?* Google returns Paris as the answer.

There is a growing realization that marginal improvement in search might not be enough to declare search the killer app for the Semantic Web. Likely, understanding context is not sufficient to build a better search engine. A combination of semantics, innovative algorithms, and memory of who the user is, will be necessary to power the next generation search engines.

Alternative approaches also attempt to overlay semantics on top of the search results. Some search engines venture into verticals by partitioning the results into different categories. The consumer can then decide which type of results they are interested in.

Yet search is a game that is far from won and a lot of semantic companies are really trying to raise the bar. There is a lot of work to be done. The whole search play - contextual technologies, as well as semantic databases, could lead to qualitatively better search engines in the future.

6. Contextual Technologies



We are seeing an increasing number of contextual tools entering the consumer market.

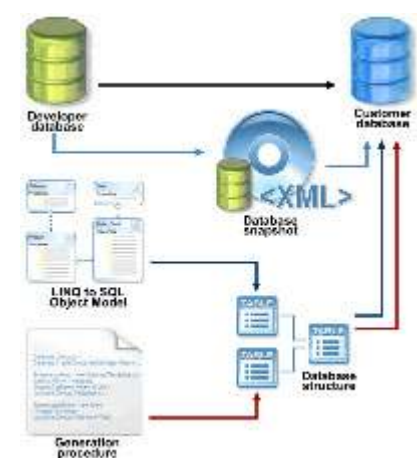
navigation does not just improve search, but rather shortcuts it. Applications like SmartLinks "understand" the objects inside text and links and bring relevant information right into the user's context. The user does not need to search at all.

Thinking about this more deeply, one realizes that contextual tools leverage semantics in a much more interesting way. To parse what a user types into the search box, contextual technologies rely on analyzing the content. So the mechanism is much more precise way - or rather, there is less guessing. The contextual tools then offer the users relevant choices to a correct result. This is fundamentally different from trying to pull the right results from a myriad of possible choices in a web search.

We are also seeing an increasing number of contextual technologies make their way into the browser. Top-down search engines need to work without publishers doing anything; and so to infer context, contextual technologies integrate into the browser. A recommended extensions page features a number of contextual browsing solutions - *Interclue*, *ThumbStrips*, *Coolio* (from my own company).

The common theme among these tools is the recognition of information and the creation of specific micro contexts that allow users to interact with that information.

7. Semantic Databases



Semantic databases are another breed of semantic applications focused on annotation. They tend to be more structured. *Twine*, a product of Radar Networks and currently in private beta, is a personal knowledge base. Twine works by absorbing unstructured content in various formats into a personal database of people, companies, things, locations, etc. The content is stored in a bookmarklet or via email or manually. The technology *needs to evolve more*, but once mature, semantic databases can be useful once the kinks are worked out. One of the very powerful applications that can be built on top of Twine, for example, is personalized search - a way to filter the results of a search engine based on a particular individual.

It is worth noting that Radar Networks has spent a lot of time getting the infrastructure in place. The underlying representation is RDF and is ready to be consumed by other semantic applications. A chunk of the core algorithms, the ones that are dealing with entity extraction, are implemented in Twine. Semantic Web APIs. Reuters offers this as an API call, for example, and so moving forward, we need to be concerned with how to do that.

Another big player in the semantic databases space is a company called Metaweb, which created *Freebase*. In its current form, it is just a fancier and more structured version of Wikipedia - with RDF inside and less information in total. The overall goal, however, is to build a Wikipedia equivalent of the world's information. Such a database would be enormously powerful. It can be queried exactly - much like relational databases. So once again the promise is to build much better search.

But the problem is, how can Freebase keep up with the world? Google indexes the Internet daily and grows together with it. Freebase currently allows editing of information by individuals and has bootstrapped by taking in parts of Wikipedia. But in order to scale this approach, it needs to perfect the art of continuously taking in unstructured information from the web and updating its database.

The problem of keeping up with the world is common to all database approaches, which are effectively silos. In the case of Freebase, there needs to be continuous influx of user data, and in the case of Freebase there needs to be influx of data from the web. These are far from trivial and need to be solved successfully in order for the databases to be useful.

Conclusion

With any new technology it is important to define and classify things. The Semantic Web is offering an exciting promise of information discoverability, automation of complex searches, and innovative web browsing. Yet the Semantic Web is being used in different ways by different people. Indeed, its definition in the enterprise and consumer spaces is different, and there are different patterns - end - top-down vs. bottom up and microformats vs. RDF. In addition to these patterns, we are observing the rise of contextual browsing tools. All of these are in their early days, but hold a big promise to fundamentally change the way we access information on the web.

What do you think about Semantic Web Patterns? What trends are you seeing and which applications are you working with semantic technologies in the enterprise, please share your experiences with us in the comments below.

Comments

Subscribe to comments for this post OR Subscribe to comments for all ReadWriteWeb posts

1. Thanks. This is one of the most thorough semantic web posts I've seen...really appreciate the effort that must have gone in

Posted by: [Eric Willis](#) | [March 25, 2008 4:32 PM](#)

2. This was a great post, Alex. I'll keep this one in my back pocket for awhile -- definitely a great overview. I learned a lot about (as I usually do from your posts on this topic).

Posted by: [Josh Catone](#) | [March 25, 2008 4:40 PM](#)

3. "Microformats offer a simpler approach by adding semantics to existing HTML documents using specific CSS styles."

This is false; microformats do not use CSS styles. They use the class attribute (among other HTML semantics), which - as a classification for information. Visual presentation is only one of many applications of such classification, and is not at all relevant to microformats.

Posted by: [Scott Reynen](#) | [March 25, 2008 5:56 PM](#)

4. No discussion of Semantic technologies is complete without at least mentioning Openlink Software and its monster CEO, [Kingsley](#), as he said. He has assembled an international team of really talented engineers, and they have delivered a fully integrated solution for an ISP to carrier class Semantic Server architecture.

I can't do it justice, and seemingly, even Openlink is having a hard time getting the message out. I worked very briefly with them on so many levels. I am flabbergasted that they have not as yet been snapped up by the likes of Oracle or SAP. <http://openlink.com>

Im not a big fan of OWL / RDF, but then again, I'm not a tools vendor - for only tools which sufficiently abstract out the complexity of the mess will get us to where we want to be.

It's a bandied about term, this ole Semantic thing - see what Kingsley and company have done with it, and you will get a glimpse of what systems are needed to get the revolution underway.

Posted by: [abm](#) | [March 25, 2008 5:57 PM](#)

5. Fantastic post Alex, thanks for putting such thorough entries. My organization has been slow in adopting web services which is a shame since now RDF will be a requirement. A 3 step guide to making your API RDF compliant would be oh so welcome ;)

Posted by: [PG](#) | [March 25, 2008 5:59 PM](#)

6. What a well-written and thorough article. I've long been interested in the advances in the semantic web, and agree with Moore's trend in the web at the present time.

What I'm most interested in is seeing the tools that become available for web authors. Adoption will really take off once it becomes compliant content, and it's this area that I'll be following the most.

Thanks for such an in-depth review of the current "State of the Semantic Union". A great read that I've passed on to my friends.

Posted by: [honest ape](#) | [March 25, 2008 6:12 PM](#)

7. Great summary of current Semantic Technologies. I think we'll see a lot of powerful stuff come out of the juxtaposition of Content Management Systems and Semantic DBs.

Another tool to add to the list is Orchestr8's AlchemyGrid (grid.orch8.net) -- it offers a web-based facility to "scrape" structured content from a webpage, and other unstructured content sources as well. The entire process is visual, top-down, and zero-code.

AlchemyGrid lets you provision your own REST/RSS/ATOM/JSON/etc APIs, get email/SMS/twitter/AOL IM alerts when we and "enrich" scraped content with information gleaned from 3rd-party DBs and web services.

"Scrapes" can be shared, tagged, commented-on, turned into RSS feeds, and so on. Several thousand APIs have already sites on the web, giving a great "initial push" to those wanting top-down access to structured data already out there "in the

Posted by: [Elliot](#) | [March 25, 2008 10:00 PM](#)

8. Interesting Post. I get confused by a lot of these technologies but Microformats are really useful.

I have recently released a WordPress plugin that looks for hReview Microformats when a site is pinged so that better inform opinion of the reviewing author.

I am hoping that this kind of use will increase in future for better connectedness between blogs.

Posted by: [Andrew](#) | [March 26, 2008 1:11 AM](#)

9. Hey you have briefly explained and analyzed the trends and technologies that power the Semantic Web.nice information...

Posted by: [search engines](#) | [March 26, 2008 3:03 AM](#)

10. Great post.

One of the questions I find most fascinating is how marketing will evolve to take advantage of the semantic web, whether it I think this is more than a linear extension of how marketers have been optimizing the web today, but something qualitative SEO + Semantic Web = SEO++ (after all, it is sort of an object-oriented paradigm shift).

Here are 7 possible missions for "semantic marketing":

1. Marketing becomes the champion of generating the underlying data.
2. Marketing views categorization, metadata, RDF graphs, relevant microformats, etc., as a new kind of market positioning branding", if you will.
3. Marketing takes a much broader view of distribution and promotion of its semantic web data in search engines and vertical including the sponsorship or creation of new niche semantic networks.
4. Marketing comes up with new ways to incentivize the conversion of semantic web interactions in real business objective
5. Marketing will have a real challenge with tracking and attributing distributed data in the semantic web to measure its impact marketing to micro-touch marketing. Hard problem but entrepreneurial ingenuity will prevail.
6. Marketing will want to leverage other people's data in their own value-add mash-ups (interesting "joint venture" semantic as for internal-only apps focused on market research and competitive intelligence.
7. Marketing will need to be concerned with brand protection in the semantic web: quality control to watch for bad data, copyright misuse, etc.

If you're interested, <http://www.chiefmartec.com/2008/03/marketing-in-th.html> is the full post. Would love feedback from other web aficionados.

Posted by: [Scott Brinker](#) | [March 26, 2008 3:39 AM](#)

11. Great post, RRW is definitely the cutting edge blog when it comes to semanticweb. Thanks for that.

Posted by: [Mr Boin](#) | [March 26, 2008 3:50 AM](#)

12. Alex, thanks a lot for the post. This kind of posts keep me reading RWW every day (well, ok, and night also due to the time

The trends I'm seeing right now from the user's perspective are not that bright. I'm a bit tired of tagging the same info all the

twine, bloglines, etc... On the one hand, yeah - helping to make better web, on the other - why can't the systems scan head ease my life? Why RSS feeds are not already bundled with microformats? Why powerset sucks at answering natural-langu

It seems that semantic web is just a colorful painless _future_ for an _end user_. Right now only context semantic applicati you going to offer non-firefox browser support?).

For right now keep believing in semantic web and tag, and tag, and tag... for the best of everybody.

Posted by: [Sasha Kovaliov](#) | [March 26, 2008 5:23 AM](#)

13. There is no such thing as "Guitar Hero" for the Semantic Web, there never will be. Reading the tone of this post and some wonder if some of the (super smart) people who are the leaders in this space need to go be "rock stars" somewhere else.

Making the world's data more easily shared is a noble pursuit but far too geeky to ever yield any kind of "killer app" that will

Posted by: [Todd](#) | [March 26, 2008 8:34 AM](#)

14. Interested in what caused the idea that Microformats have anything to do with CSS? Until recently you couldn't even style rel=nofollow in IE, and XOXO is basically a formulation of how to use lists in (X)HTML...

Posted by: [Stephen Paul Weber](#) | [March 26, 2008 9:03 AM](#)

15. Hi Alex,

Our approach at Imindi (An application built by a team of Phd computer and neuro scientists) is to take a "mind" approach construct their own "mind maps" of connected thoughts and information on any subject. These Mind Maps are little semantic think" and the "Thought Engine" (Semantic Graph) at the core of Imindi enable Like Minded people (Social Graph) to connect Thoughts, Information and even Create Knowledge. At the core, all the semantic linkages from every body`s public mind map to form essentially one global mind map.

We agree with you that the named entity analysis functionality is going to be a commodity and indeed Imindi will be providing uses via an integration with an existing semantic repository.

However, the closest vision to what we have built was not that of the Semantic Web, but that of the Memex by Vannevar Bush of thought and meaning over the information on the web...I would suggest that we are also very much a Semantic App work

<http://www.imindi.com/journeys/382-semantic-web/maps/3195155-semantic-web>

Posted by: [Adam Lindemann](#) | [March 26, 2008 9:06 AM](#)

16. I started to disagree with the suggestion in paragraph two of this post that the semantic web is the number one trend for 20 information overload is. But then it occurred to me that the two need not be in competition.

Perhaps semantic web technologies should look to the problem of information overload as they search for that so far elusive ask me how though, but perhaps a good place to start would be to filter and de-duplicate the information we consume and the tools out there today. Imagine a feature in Google Reader that filters my feeds based on the semantics of items that I h removes or hides any duplicates.

At any rate, it is vital for the adoption and success of semantic technologies that they help ease the problem of information compound it.

Posted by: [IdeaTagger](#) | [March 26, 2008 11:05 AM](#)

Incredible post, on a topic that really needed to be cleared out. Thanks R/WWW.

Posted by: **Cedric** | **March 26, 2008 11:07 AM**

18. Surely Google will settle the top-down vs. bottom-up debate for us. There's no way it's ignoring semantic/web3.0 - it's the in web. But there's no way, either, it's going to favour bottom-up, for three reasons (see here: <http://tinyurl.com/2rvyhh>)

Mainly, a bottom-up web would threaten Google's search monopoly, by making it easier for rivals to get into the semantic s better placed to tackle the bigger tech challenge of giving webspiders and their search engines semantic abilities, than Google allows it to cement its market position.

Google will settle this debate for us (<http://tinyurl.com/2rvyhh>), just you wait!

Posted by: **Phil Bradley** | **March 26, 2008 1:10 PM**

19. Alex,

You should have a pingback/trackback from a post I penned earlier today.

The Semantic Web is a Web endowed with Linked Data URIs (Object IDs or Entity IDs).

If I can reference your data from my data space via URIs, we end up with a Web that offers transparent access to the data web pages.

It isn't really a complex value proposition. We are simply moving to a Web where perspectives aren't fixed. We all have the consume analysis etc..

Discourse discovery and participation becomes a cost-effective endeavor.

Databases provide Views and Lookup capability over structured data.

The Semantic Web provides exactly the same without any database platform lock-in.

Unfortunately this lack of lock-in is often misunderstood as being incongruent with business models (vehicles for providing

If you can just get your person entity a URI, it will all become a little clearer re. value proposition. I say this because Widgets simply Silos in another guise.

Silos can not deal with the imminent onslaught of "Information Overload" amongst others things like the warped "Attention

Kingsley Idehen

Posted by: **Kingsley Idehen** | **March 26, 2008 3:59 PM**

20. Though various movements such as Semantic Web may not be a panacea, they are stepping stones in bringing forth a more groupings of information, which can have many benefits including improvement of Search.

Posted by: **Peter T - webshop** | **March 26, 2008 4:05 PM**

21. Thanks Alex, for this thorough and informative article.

I'd say despite of being a great, revolutionary and highly useful concept, where Semantic web lacks is into posing immediate implementors. The major factor into wide adaptability of Web 2.0 (or Social-Web) is it's "right-now" benefits to adapters.

But as you mentioned, now Yahoo! search is going to appreciate semantic so we are getting some incentives in place. Hope incentives coming up along the way soon :)

Posted by: [ptc](#) | [March 26, 2008 6:01 PM](#)

22. Alex, a very interesting article that highlights some of the technologies available. It would be also interesting to learn how to further it is interesting to understand the future of linking semantic content to an object of interest. Assuming you create the web page which is embedded but what about the relation to the new information that arises from other web pages that have been annotated? How to keep up with the dynamic part and link always new content to an existing one? Hope to see also soon some progress in Europe for the new semantic technologies.

Posted by: [Daniel](#) | [March 27, 2008 1:08 AM](#)

23. Hi Alex,

This is a very nice article, an exciting overview of what has been done and what can be done in the area. However, it is pity about latest NLP technologies that are often in the core of a semantic web services/apps. E.g. in Ontos we are developing a service that has already been implemented as a web service. Look at our [Ontos Business portal](#), where the data coming from news articles is structured and appears on the portal in a structured form.

I've also written a post in our [Ontos blog](#) about your article. Really nice information, thank you!

Posted by: [Philip](#) | [March 27, 2008 3:23 AM](#)

24. @Daniel

Yes, you're right to highlight the importance of links in the Semantic Web - something which is frequently overlooked but for which a design is needed. You may be interested to have a look at the Linking Open Data project (link below) which is making great progress.

<http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData/>

As for progress in Europe, it's all around you! European institutes are powerhouses of research and development in semantic web technologies, and there are plenty of European companies (old and new) using these technologies already :)

Posted by: [Tom Heath](#) | [March 27, 2008 3:23 AM](#)

25. Hi Alex,

thanks for the nice overview! It's a pity that you didn't write much about NLP, it seems to be often in the core of semantic-web services. ClearForest and Calais are developing NLP web-services, look e.g. at [Ontos Business portal](#) which is also based on a powerful NLP engine. It can extract entities (persons, companies, etc.) and relations between them (employ, mention, found, etc.). There is also a **'try' mode** where you can test your own text with our NLP. I'll be happy to get any feedback if you are interested.

Thank you again! I've posted a short review on your article in our [NLP Ontos blog](#).

Posted by: [Philip](#) | [March 27, 2008 6:41 AM](#)

26. @Tom

Thanks Tom for the link which is interesting and of course your content on your web page. I have come across only a few examples of using RDF/OWL or semantic technologies in real commercial situations. Do you know any examples in Switzerland?

Posted by: [Daniel](#) | [March 27, 2008 7:44 AM](#)

27. One of the things I don't see associated with the Semantic web is a move to catalog and index everything. This would include photos, film, audio, video, etc. - including small business/enterprise and individual/personal data - wherever and whenever possible (rights respected, of course).

Before I became aware of the Semantic Web, I jotted down a few notes about where I believed the Internet was (or could be)

<http://www.squidoo.com/unstructuredinformation>

I have since added a few links related to the industry and emerging technologies.

Here's my observation: except for commercial archives, in many ways the Internet is a reflection of data collected from the words, perform a search for non-historical figures, small companies (that no longer do business) or local events from more and there's not much information to be found. It seems to me that whatever the Semantic Web is or could be, it will be much relevant when it captures everything that went before.

Posted by: Sid | [March 27, 2008 12:51 PM](#)

28. ... what is totally missing here are semantic web services (not semantic databases' APIs, but a way of storing meta information accessing web services); have a look onto what the Digital Enterprise Research Institute (DERI) is doing

Posted by: [Lars Ludwig](#) | [March 30, 2008 8:21 PM](#)

29. you have briefly explained technologies that power the Semantic Web. [greatinformation.....](#)

Posted by: [elena](#) | [March 31, 2008 7:58 AM](#)

30. I have been following Alex Iskold for some time, and have recently come to know and respect Scott Brinker (comment #10). [Marketing Google Group](#).

If we look back to an earlier post from Alex - "[Top-Down: A New Approach to the Semantic Web](#)". He says, "Here is what we want to do with the semantic web:

- * Spend less time searching
- * Spend less time looking at things that do not matter
- * Spend less time explaining what we want to computers

A consumer focus and clear benefit for businesses needs to be there in order for the semantic web vision to be embraced

So, by extension, another way to look at the mission for Semantic Marketing is to help people:

- * Spend less time searching for products and services that are relevant!
- * Spend less time looking at products or services that don't interest them, and more with those that do!
- * Spend less time explaining what they want to websites that market products or services!

We had been unknowingly working toward that mission for more than a decade. It was borne out of the observation that real conversion. Within the last two years, we began experimenting with the detection of combinations of visitor attributes - that location to language preference to operating system to search engine keywords to targeted websites visited. We now call them "[Semantic Personas](#)". They represent specific target segments within your market. Last year, we received patent pending for our business method (called [Semanticator](#)). The results from our oldest implementation are stunning - a 41% increase in time on site, a decrease in bounce rate and a 26% increase in conversion. All from making visitor experiences more meaningful.

So, if you buy into Alex's vision of the Semantic Web, and my extension of his concept to Semantic Marketing - we aren't just there, we have arrived! [See for yourself!](#)

Posted by: [John-Scott Dixon](#) | [March 31, 2008 8:17 AM](#)

31. Great article - I added [added some notes and semantic markup](#) to this article using our [a.nnotate.com](#) tool for discussing documents online. The idea is to start with plain text documents, let people highlight phrases using the browser, and attach sticky notes to the page. People can then link to people to discuss the same copy of the page. The annotations/links are stored independently from the page (an idea which

start of the web, but not often implemented in practice). Collaborative annotation / discussion of documents offers an immediate way to get people hooked adding semantic markup without having to know that's what they're doing. There's a variety of extensions of the interface for adding rdf-style connections too.

Posted by: [Fred Howell](#) | [April 9, 2008 12:14 AM](#)

32. With regards to contextual technologies, OpenURL is similar in concept to Yahoo Shortcuts and SmartLinks, only it is interestingly open source, as opposed to proprietary. It was created in the late 1990s.

"The OpenURL standard is designed to support mediated linking from information resources (sources) to library services (targets) through a "link-server", parses the elements of an OpenURL and provides links to appropriate services as identified by a library. A source is a bibliographic citation or bibliographic record used to generate an OpenURL. A target is a resource or service that helps satisfy user needs. Examples include full-text repositories; abstracting, indexing, and citation databases; online library catalogs; and other library services." [Wikipedia, <http://en.wikipedia.org/wiki/OpenURL>]

In other words, contextual linking based on a user's location to allow them access to information.

See also:

OCLC: <http://www.oclc.org/research/projects/openurl/default.htm>

ExLibris (commercial version): <http://www.exlibrisgroup.com/category/sfxopenurl>

Posted by: [Jewel Ward](#) | [April 14, 2008 10:56 AM](#)