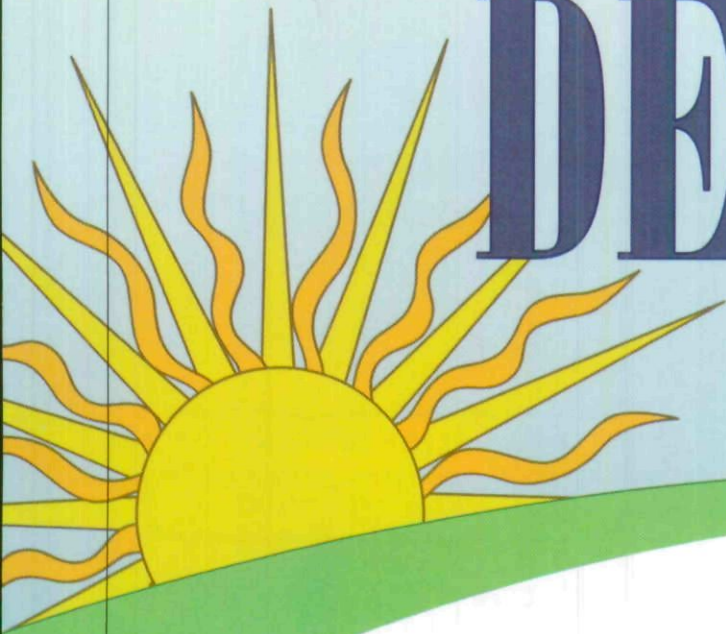


# CONTENT DELIVERY



Imagine going to a website that delivers content just for you—based on your unique needs and interests—instead of a one-size-fits-all, generic page. The intelligent systems behind the site don't just display content because the markup language says to do it. The systems begin to analyze the meaning of the content they are displaying, they can interact with other systems, and they can relate it to other data they handled in the past. What's more, these intelligent machines are able to interpret and understand the information they are handling in a similar fashion to the way we as humans understand the complexities and subtleties of human speech.

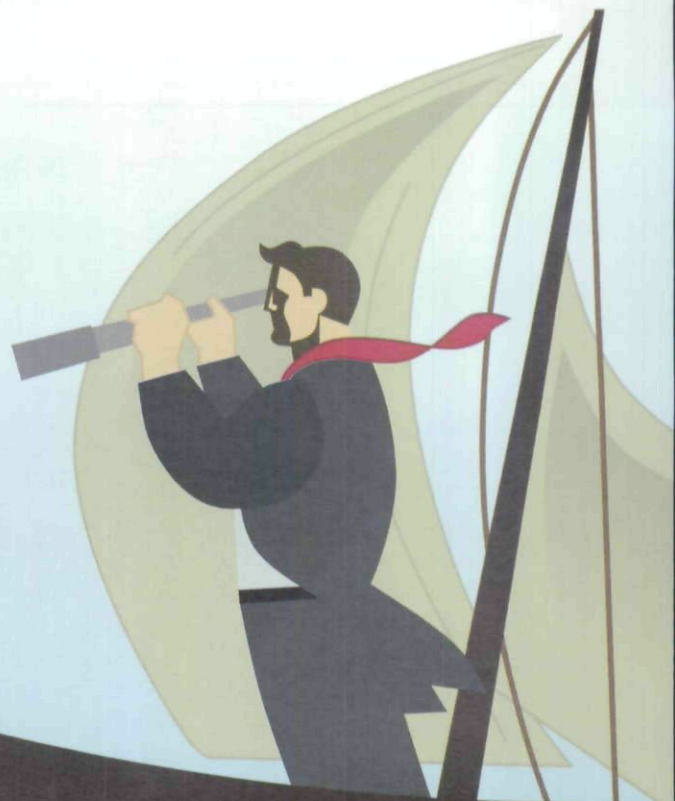
This scenario may sound far-fetched, but it is, in fact, Tim Berners-Lee's vision of the future of the World Wide Web, which he has termed the "Semantic Web."

In the earliest days of the web, sites were created by design gurus, and visitors simply came, read some content, clicked a few links, and maybe bought a book or a T-shirt. Several years later, the social web began to develop, and suddenly the tools were

more accessible. We as users not only interacted with the content, we could create it too. This era, which we are still in the midst of, has been coined Web 2.0. Yet we find ourselves perched on the crest of a third wave, one that takes this social web further still by adding a semantic layer to the web, in which the underlying systems themselves begin to have an inherent semantic understanding of the content, without us having to explicitly tell them what we want to do.

Today, the World Wide Web Consortium (W3C) has created a series of standards to help pull this vision together, but we are still very much at the beginning of the evolution of this technology. As you would expect, there are plenty of nonbelievers who think the semantic web concept is far too complex and will never happen. What's more, while people may have some notion of a semantic web, it's extremely difficult to understand, even on a high level. Yet despite these obstacles, there are companies out there working with these technologies and trying to solve these issues, one small puzzle piece at a time.

# Rides the Semantic Wave



## BOILING THE OCEAN

At its most basic level, the semantic web involves machines communicating with other machines in a more intelligent way than they do today by adding a semantic layer to the existing web. Mills Davis, founder and managing director of Project10X, a Washington, D.C.-based consulting firm specializing in semantic technologies, recently authored a comprehensive report on the state of semantic technologies called “The Semantic Wave 2008 Report: Industry Roadmap to Web 3.0 and Multibillion Dollar Market Opportunities.” The report’s executive summary describes the semantic web as follows:

“Semantically modeled, machine executable knowledge lets us connect information about people, events, locations, times—in fact, any concept that we want to—across different content sources and application processes. Instead of disparate data and applications on the Web, we get a Web of interrelated data and interoperable applications.”

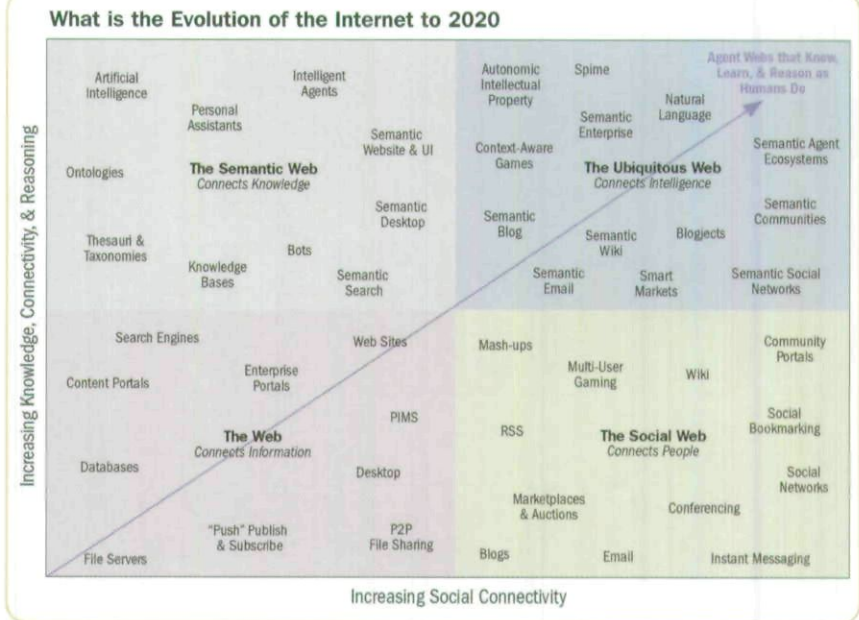
Howard Greenblatt, CTO at Metatomix, who has been working with semantic technologies since 2000, says it was around that time that Tim Berners-Lee began talking about creating the semantic web as an extension of the existing web: creating a web of meaning instead of a web comprised of simple hyperlinks. Greenblatt explains that when we visit a site, we see a link and we read it and understand the words and where it’s going to go, but the machine doesn’t understand anything beyond the underlying HTML that identifies it as a hyperlink. “There’s nothing machine-processable about that link at all,” he says. However, he explains that if you were to build a semantic connection to that link, the link itself would have intelligence and the system could understand it. Thus, when you click it, the underlying computer system could process it and send more meaningful information to the next system.

Marc Strohlein, a consultant at Outsell, Inc., thinks it’s important to look beyond the theoretical viewpoint associated with the W3C and Tim Berners-Lee to the semantic technology that

*As the web evolves, the underlying machines will grow more intelligent and be able to process and understand content. Right now we are in the middle of the development of the social web and at the edge of the development of the semantic web.*  
 Source: Nova Spivak, Radar Networks, John Breslin DERI, & Mills Davis, Project10X.

underlies it. He says of the W3C approach, "It's like boiling the ocean. I think Berners-Lee set the bar pretty high." Strohle explains that with semantic technology, you have three layers—meaning, code, and content—and the semantic technology separates these three layers from one another. It also provides a way to establish models of relationships between entities and concepts, he says, so you go up a level of abstraction. "The semantic web is essentially the web enabled with these technologies."

Trying to communicate this concept to people is a challenge. "It's a pretty complex topic with a complex nomenclature around it," Strohle says, "and this can be a barrier for people. The first reaction is 'I can't deal with a technology that I need to be a rocket scientist to understand.' Most people can understand content management and search, but when you talk about ontologies, people's eyes glaze over." That's why companies selling semantic technology products are generally taking small steps, and talking



more about the results than the underlying technology.

**WHERE DOES CONTENT FIT?**

What this means from a content standpoint, says Project10X's Davis, is that we are going to see the ways we use content on the web transformed, as the social web interacts and intertwines with the semantic web. "The whole realm of content has several ways it's going to move," he says. For instance, "From the standpoint of communities, a key trend is the collective knowledge system and

that's the intersection of Web 2.0, collaboration and Web 3 or the semantic knowledge-intensive stuff."

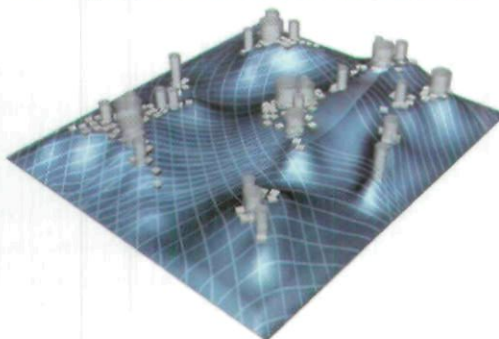
He explains that this is dependent upon intelligent interfaces and interfaces that can actually learn, and that this learning can come from members of the community interacting with the system. He uses an example of a travel site where people post their travel experiences, and this community information then becomes the basis for the website to provide more accurate and in-depth recommendations. "The learning can come from people's interaction with the system, so you get a concept of systems that improve their operations without having been explicitly programmed to process the new information."

This means that knowledge becomes an operational piece of the process. With content delivery, he explains, the system can use this information in new ways based on the task, interest, or the point of view of the visitor to the website, and it is the semantic layer that makes this possible. "This can only be done in a computer era where [the system] knows something about the reader and is able to learn and

**The Metatomix Approach**

Metatomix simplifies the way in which an application uses data from ANY data source

- Embrace the Distributed Enterprise Landscape
- Integrate your Application Specific Data Silos
- Leverage our High Performance Semantic Rules Engine to apply a common understanding across applications



Result: A semantic real-time virtual integration layer that unifies the applications across your enterprise landscape and allows you to see relationships amid 1000's of seemingly unrelated events

metatomix

*Metatomix tools can cross data silos to process data while applying reason and inference and embedding logic into the data itself to get you the information you need.*

observe and adapt its behavior to that person, and also knows something about the content and the meaning of what it's dealing with in terms of subject matter in order to be able to organize and project and adaptively present content that is tailored to the needs of that reader."

#### LET'S GET REAL

While this sounds good, getting from the realm of the theoretical to real-world application takes a substantial technological leap and we are only at the beginning of the journey. There are, in fact, plenty of companies out there developing technologies to deliver specific web content, to help distill it from a vast collection, to help people find the right information (semantic search), or to communicate across disparate data repositories.

Some of these companies, such as Metatomix, use semantic concepts to deal with specific business problems. "What we have done is focus on some specific areas where we could actually leverage semantic technology to solve real-world business problems, as opposed to trying to implement the Semantic Web," Greenblatt says. The Metatomix products, which they call 360° Solutions, integrate a traditional business rules engine with semantic data to enable enterprise users to cross data silos and pull information together in intelligent ways, including the ability to use reason and inference over the data and embed logic in the data model. "The ability to have that gives you the ability to discover things about your information that you wouldn't normally be able to do and that's the advantage of using semantic technology," Greenblatt says.

Metatomix has applied this technology to specific vertical solutions in finance and law to help pull information from different data sources. For instance, using the judicial product you can cross many

different databases to build a criminal profile about a person that would normally take 20 different log-ons (while still adhering to privacy rules). Using semantics, the tool looks for information in each database that enables it to get the information it needs to access another database until the end result is a complete view of the person's record pulled from a range of criminal database sources.

Meanwhile, Semanticator LLC is bringing semantic technology to marketing by trying to mimic the live sales experience to deliver content based on the individual visitor's needs. "It's all about recognizing the context of an internet visitor and being able to respond relevantly based on that context," says Chris Hewitt, VP of technology at Semanticator. "What we are talking about on an internet scale is really kind of revolutionary, but from a human perspective it's not." What Hewitt means is that when a person walks into a store and talks to a sales person, the two people involved recognize contextual clues and begin to interact based on that.

Instead of face-to-face human clues, Semanticator uses information available by querying the browser you are using. "When we open a web browser, we inherently share some information about ourselves, totally preserving privacy and keeping everything anonymous, so we're not intruding on anyone's personal space,

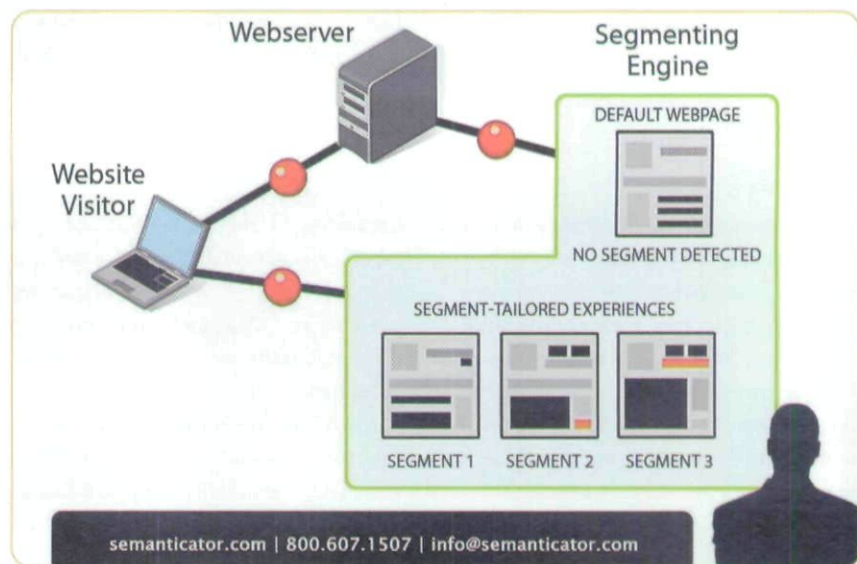
#### Companies Featured in This Article

Elsevier, Inc. www.elsevier.com	Project10x www.project10x.com
Illumin8 http://illumin8.com	Semanticator LLC www.semanticator.com
Metatomix www.metatomix.com	Trovix, Inc. www.trovix.com
Outsell, Inc. www.outsellinc.com	

but we can use that information to understand the context of the visit and how to respond more relevantly."

Semanticator focuses on a user's needs by building one or more personas of a typical visitor. These personas define the characteristics of these visitors and help the Semanticator engine deliver the right content based on that persona along with other contextual clues it finds from querying the browser. In fact, Hewitt says that many clients can reach a majority of their visitors with just a couple of personas. For example, a hotel site might have a meeting planner persona consisting of what types of websites and blogs they might visit, their location, and so forth, and Semanticator uses the underlying semantic technology to pull this and other information together to display the page that makes most sense for a meeting planner.

In another case, science publisher Elsevier, Inc., a company that has had to find ways to expose vast amounts of



*Semanticator LLC can tailor the webpages you see based on information about who you are and other clues it gets from your browser. This gives you a more customized page, rather than a one-size-fits-all, generic page.*

semanticator.com | 800.607.1507 | info@semanticator.com

Elsevier's Illumin8 engine can process Elsevier's vast content store, as well as web content, then deliver a set of results formatted to simplify pinpointing the documents that make most sense for you.

content to its audience, has been working with semantic technologies to produce a search tool specifically geared for a research and development audience that includes Elsevier content as well as content from the open web. Its product, Illumin8, provides a way for customers to locate the information they need from mountains of information using natural language queries and an underlying semantic index built using NetBase.

"From our vantage point, we are content provider first and foremost," says Joe Buzzanga, product manager for Illumin8 at Elsevier, "and we have a lot of content. We are probably the largest scientific and medical publisher in the world. We are always looking for ways to exploit that content for our users' benefit, to make it more accessible in various ways, and we are actively exploring a number of different technologies," he says.

He points out that his company has been organizing its content for a long time using taxonomies and metadata, and taking that structure and applying semantic technology to the problem is the next logical step. "We feed our premium content with the metadata to the NetBase engine and they are able to extract and work with the metadata." The end result is a page of results organized in logical categories from which the user can quickly review the information that makes most sense for them.

#### HOLD ON A SECOND

Some people believe that developing a fully semantic layer on the web that can truly learn and pass knowledge from one computer to another is simply too good to be true. Or at least, that it's going to take an extreme level of cooperation, and doubters wonder if this level of cooperation would ever be possible to carry today's solutions out to the web at large.

One person who is skeptical about taking this whole vision to the web is Dan



Enthoven, VP of marketing at employment search vendor Trovix. Although his company uses semantic technologies, he has doubts about the Berners-Lee W3C vision of a semantic web. "If we have an axe to grind with the semantic web, it's that we think it takes two steps too far," says Enthoven. "One is the idea that you can have this universal knowledgebase that applies to all things. When you try to define a universal knowledgebase, it falls apart because different people have different opinions on what things are," Enthoven explains. "The other idea is that people are going to start tagging documents and creating this whole meta layer manually. Metatags went out of fashion because they were so widely abused. If metatags didn't work then, why would the semantic web work?"

However, Brooke Aker, CEO at Expert System, makers of semantic search technologies, doesn't buy that argument. "I think the W3C has put together a reasonable set of standards that will take us in the direction we need to go." He does point out that these standards can't provide an engine to enhance our understanding of the content on the web; they can only provide a conduit for systems to communicate. "Standards are just a means for interchanging, for adjudicating on how my technology might understand

something versus someone else's. If we process things with our technology, and we put that out with enhanced tags that meet the W3C standards, then someone else can absorb that and understand it," says Aker.

While many companies have clearly taken semantic technologies and built useful tools that can help you deliver more relevant content to your visitors, Tim Berners-Lee's vision of an overarching semantic layer has yet to be realized. In fact, Outsell's Strohleim believes that Berners-Lee's top-down approach, where the W3C dictates the direction, probably won't work. He thinks it actually more likely to work from companies such as the ones mentioned earlier, building a semantic web from the ground up.

For all that, Strohleim believes that the W3C standards are advanced enough that the foundation pieces are in place, but getting to a point where machines actually communicate intelligently with a semantic understanding of the content they are sending is still going to take a series of technological breakthroughs. But if we can get there, the possibilities for customized content delivered through smart interfaces is quite exciting indeed. **E**

**RON MILLER** (RONSMILLER@RONSMILLER.COM) IS A FREELANCE TECHNOLOGY WRITER BASED IN MASSACHUSETTS.  
**COMMENTS?** EMAIL LETTERS TO THE EDITOR TO ECLATTERS@INFOTODAY.COM.

Copyright of EContent is the property of Information Today Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.