Near-Term Prospects for Semantic Technologies

V. Richard Benjamins, *Telefonica R&D*

For the past few years, the Semantic Web has been enjoying significant investment, mostly through research but to a lesser extent through start-ups and commercial projects. A major topic of discussion is where we can see those investments' results, so I asked several experts to consider what semantic technology will accomplish in the near future. The experts are from academia, venture capitalist firms, and companies focused on semantic technology, telecommunication, and Web 2.0.

Current projects and applications

Before giving the floor to the authors, I'd like to mention a few initiatives relevant to the Semantic Web discussion. The World Wide Web Consortium maintains a Web page for its Semantic Web Education and Outreach Interest Group. The page lists known case studies and use cases of practical applications of Semantic Web technology (www.w3.org/2001/sw/sweo/public/UseCases).

The NeOn project aims "to advance the state of the art in using ontologies for large-scale semantic applications in the distributed organizations" (see www.neon-project.org). As part of this project, I participated in a survey to find out how many semantic applications were out there, both within corporations and on the public Internet. We looked at parameters such as application status (prototype, pilot, or deployed), origin of budget (for example, public funding or company budget), and application areas (such as knowledge management or enterprise application integration). We received about 30 applications from researchers and Semantic Web technology practitioners. Our main findings were that most applications deal with search, and only a few of them were deployed and financed with corporate budget. For more details, see www.neon-project.org/web-content/index.php?option=com_content&task=view&id=47&Itemid=1.

Web 2.0 refers to applications that focus on user communities and user-generated content, such as YouTube and Wikipedia, and thus enjoys high visibility. Semantic technologies, on the other hand, are infrastructure technologies on top of which we can build intelligent applications for end users. Semantic technologies should be invisible to users.

Mashups are becoming increasingly popular because they let people create new Web applications based on existing ones. Adding semantics to mashups (and thereby combining the Semantic Web and Web 2.0) lets final users create new applications even more easily by abstracting from the syntax of the applications involved. This development allows addressing the "long tail"—that is, it creates profitable business opportunities and specifically tailored solutions for many different customers (niches) that otherwise wouldn't be profitable owing to development costs.

In this installment

As a starting point, I asked the authors several questions:

- Where can we see or expect to see the visible impact of semantic technologies—for example, within corporations for consumers or with respect to the public Semantic Web?
- What are those applications, or what will they be? What will they look like? Will they address knowledge management? Business-to-business integration?
- What's the relationship between semantic technologies and Web 2.0? Are they independent, competing, complementary? What about Web 3.0?

In the first essay, John Davies discusses different types of semantic applications and explains how they differ from traditional applications. Next, Ricardo Baeza-Yates, Peter Mika, and Hugo Zaragoza analyze the Semantic Web's (non)role in search. Mark Greaves discusses the five-year prospects for core semantic technologies such as RDF, OWL, and RuleML. Jose Manuel Gómez-Pérez and Jesús Contreras focus their discussion about the Semantic Web's prospects on natural interaction, service-oriented architectures, and Web 2.0. Finally, John Domingue and Dieter Fensel discuss their belief that the integration of Semantic Web and service-oriented technologies will become a main pillar of the next-generation computing infrastructure.

Although predicting the future is difficult, hearing various leaders' different viewpoints might help us to develop an informed opinion about the future Semantic Web.

—V. Richard Benjamins

1541-1672/08/$25.00 © 2008 IEEE
Published by the IEEE Computer Society

Applications of Semantic Technology

John Davies, *BT Group*

Despite its explosive growth over the last decade, the Web remains essentially a tool to let humans access information. The next generation of the Web, dubbed "the Semantic Web," will extend the Web's capability through the increased availability of machine-processable information.

Currently, Web-based information is based primarily on documents written in HTML, a language useful for describing the visual presentation of Web pages through a browser. HTML and today's Web, however, offer limited ways of describing the content itself. For example, you can specify that a given string should be displayed in a large bold font, but you can't specify that the string represents a product code or product price.

Semantic Web technology aims to address this shortcoming, using the descriptive languages RDF and OWL and the data-centric, customizable markup language XML. These technologies, which are World Wide Web Consortium standards (see www.w3c.org/2001/sw), allow rich descriptions of Web documents' content. These machine-processable descriptions, in turn, let developers write more intelligent software systems, automating the analysis and exploitation of Web-based information.

Here, I briefly survey numerous business application areas for Semantic Web technology and discuss the relationship between the Semantic Web and Web 2.0. I conclude with a discussion of the current status of and prospects for uptake of semantic technology.

Applying Semantic Web technology

Semantic technology is being increasingly applied in the business arena. Here, I look at three specific application areas that are in the vanguard of semantic technology exploitation.

Knowledge management and business intelligence

Forrester Research estimates that around 80 percent of all corporate information is unstructured.[1] Knowledge workers are increasingly overwhelmed by information from a bewildering array of information sources—emails, intranets, the Web, and so on—and yet still find it hard to access the specific information required for the task at hand. This implies that knowledge worker productivity is being reduced and that organizations might be making decisions on the basis of incomplete knowledge. Furthermore, an inability to access key information can lead to failure to comply with legal or regulatory requirements.

Semantic technology is helping to address these issues by associating unstructured information with domain ontologies. This makes possible more intelligent information-access facilities by annotating documents (and parts of documents) with semantic meta-information—information, formally expressed, which tells the machine what the document or subdocument is about. This allows more sophisticated analysis of documents. For example, named-entity recognition is a language-processing technique that can identify particular locations, organizations, or people mentioned in textual documents with ontological descriptions of those entities (a technique known as semantic annotation—see figure 1). This offers numerous advantages:

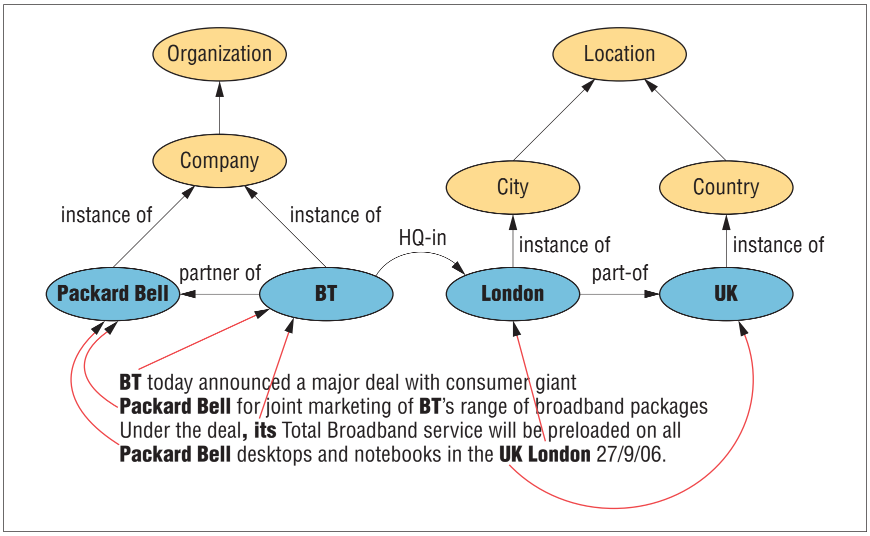


Figure 1. Semantic annotation.

- The system can associate different strings ("Mr. Brown," "Gordon Brown," "the Prime Minster," or "the PM") with the same individual, so users can retrieve relevant information regardless of which string they use in the query or document at hand.
- Reasoning over the ontology allows more focused searching than a simplistic keyword approach. The query "Show me all documents about British politicians" could return a set of documents, none of which contain the words "British" or "politician."
- Natural-language-generation techniques can be used to provide summaries about entities taken from the formal ontological data ("Gordon Brown is a politician and is the Prime Minister of the UK. The UK is part of Europe, a member state of the European Union, and has a population of 61 million …").

This kind of semantic annotation lets semantics-based knowledge management systems exploit semantic links between concepts for information access. Consider the following query:

"telecom company" Europe director

The user might be seeking, for example, documents concerning new board appointments of telecom companies in Europe. Note, however, that a document containing the following sentence wouldn't be returned using conventional search techniques:

> At its meeting on the 10th of May in London, the board of O2 appointed John Smith as CTO.

To be able to return this document, a semantic search engine could exploit the following semantic relationships:

- O2 is a mobile operator, which is a kind of telecom company.
- London is located in the UK, which is a

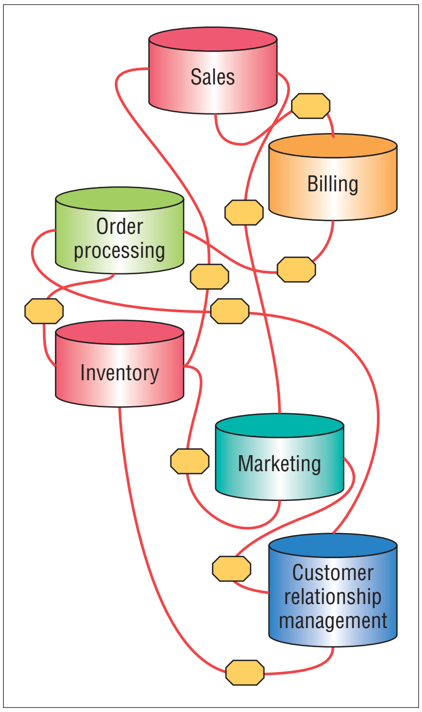


Figure 2. Corporate data silos.

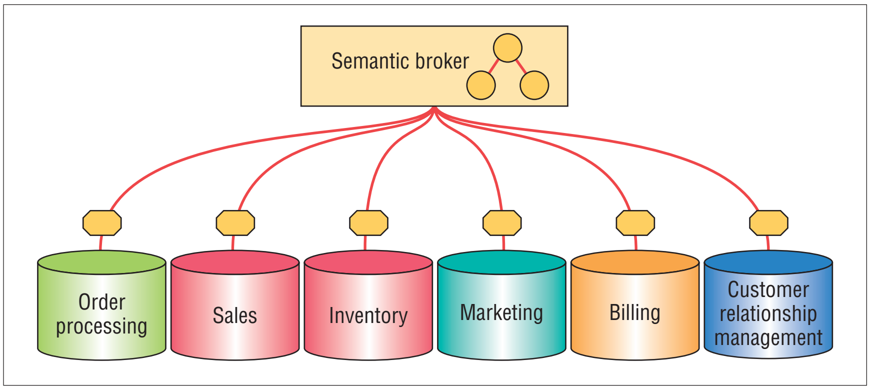


Figure 3. Use of a semantic information broker.

part of Europe.

- A CTO is a kind of director.

The SEKT (Semantic Knowledge Technologies) European collaborative project (www.sekt-project.com), for example, involving 12 industrial and academic partners from across Europe, developed semantic information access tools based on the meaning of documents (rather than the particular words used to present that meaning). These included, for example, a tool that generates a digest of relevant information from a set of documents and highlights and offers summaries about people, organizations, or locations of specific interest. The opportunity for enhanced information access and business intelligence tools is clear.

Information integration

A typical large organization in an information-intensive sector (such as finance) will have numerous data silos (an HR system, a customer-relations-management system, one or more billing systems, and so on). Each such system has its own data model, no two of which are typically the same, making information exchange and integration difficult. Analysts report that the majority of businesses resort to expensive new software or manual processes when confronted with a need to integrate information from multiple sources (see figure 2).

Ontologies can help to address this issue by providing a uniform access layer to heterogeneous data sources. Linking multiple structured, semistructured, and unstructured information sources using a consistent vocabulary makes it easier to build applications that pull data together from across the enterprise. It also facilitates the introduction of new systems and databases (see figure 3).

The semantic-integration approach has several advantages:

- There's no need to reengineer legacy data sources; existing data sources are instead "wrapped" by a semantic description.
- It's based on lightweight, open W3C standards.
- It's inherently extensible. RDF and OWL were designed to make it relatively straightforward to integrate concepts and relations from more than one ontology.
- It supports reasoning with the information. Because OWL is a logical language, formal reasoning is supported, allowing the inference of new facts from the explicitly stored information and allowing the definition of business rules.

Semantic service-oriented environments

Selecting products or services using the Web can be time consuming and onerous. You might visit multiple sites to book a holiday—for example, to compare hotel and flight prices, to rent cars, and so on. Users would prefer a software agent to automatically search the Internet on their behalf, comparing multiple products or services until it matches the requirements specified on price, accommodation quality, location, and so forth. For this to become possible, Web services must be described in a language that computers can process—a semantic Web language. A survey of recent work on languages and ontologies for Semantic Web service description appears elsewhere.[2]

In the corporate domain, Semantic Web services can enable faster discovery and composition of services in a service-oriented architecture (SOA). By describing system components using Semantic Web languages, system integration inside the organization and with partners becomes faster and more efficient. This enables faster responses to changing customer requirements and regulatory pressures.

Furthermore, as organizations increasingly expose their capabilities via Web-based services, the SOA concept is being exported to the public Web. Some envisage a "service Web" of millions of services exposed by thousands of organizations of all kinds and sizes. Organizations such as Amazon, Google, BT, and AOL have already made the first steps in this direction. Scalability will be achieved in such an environment only by describing services and their functional and nonfunctional properties at a semantic level to enable the automation of service discovery, composition, repair, and so on.

The Semantic Web and Web 2.0

In parallel with Semantic Web developments, a group of technologies have emerged in recent years that enable Web-based community and participatory systems such as Flickr and del.icio.us. Known collectively as "Web 2.0," Web 2.0 systems are complementary to the Semantic Web:

many are based around the notion of user tagging, involving annotating data (such as photographs and Web pages) with user-supplied descriptive words or phrases (tags). This is reminiscent of the Semantic Web's tagging of data with metadata. An important difference is that, in the case of the Semantic Web, this metadata is associated with a formal ontology, ensuring the consistent usage of particular tags across users and applications and formalizing the data representation.

It could be that the Web 2.0 approach will prove most useful in systems where ease of use is key, whereas the more formal ontological approach will tend to be adopted when representing mission-critical information. Certainly, it's hard to imagine a health application whereby clinicians can supply their own tags, which other health professionals might or might not share and understand. Conversely, it's doubtful that information-sharing sites such as Flickr would have attracted as many participants if they had constrained each user to a shared ontological vocabulary.

Building on the popularity of wiki technology, a number of "semantic wiki" systems (such as Semantic Mediawiki[3]) have recently emerged. In such systems, users can create semantic links between pages, but, importantly, there's no restriction on what the semantic links are called (that is, there's no predefined ontology). One user might create a link between Cardiff and Wales annotated with "is capital of," whereas another might create a link between Paris and France annotated with "capital city of." This informal approach, borrowed from the world of Web 2.0, is crucial to encouraging these links' creation. Of course, it's preferable if users reuse existing semantic terms rather than invent their own. Users can also create equivalences between links where appropriate and to indicate where one link type is a special case of another (for instance, "is capital of" is a special case of "is located in").

The semantic wiki is an example of the fusion of the methods of the Semantic Web, with its formally defined ontologies, and Web 2.0, with its informally defined folksonomies. We will likely see more such examples in the future, along with attempts to automatically analyze folksonomies and either associate them with existing ontologies or derive more formal ontologies from them.

The Semantic Web: Status and prospects

The Semantic Web won't emerge overnight. Indeed, areas of the Web will likely remain "unsemantic." Regions of the Semantic Web already exist, and it's likely that, in general, Web-based data on intranets and extranets will be semantically described first, followed by the public Web. Indeed, numerous organizations are already applying semantic technology in the areas I outlined in the previous section.

Already, however, millions of Web pages are annotated with the popular RSS (http://inamidst.com/rss1.1) and FOAF (Friend of a Friend; see http://xmlns.com/foaf/0.1) lightweight ontologies, which describe news feeds and personal social networks, respectively. The Unified Medical Language System ontology (http://umlsks.nlm.nih.gov), devised by the US National Library of Medicine and used worldwide, contains more than 700,000 concepts. BT is using semantic information integration to link customer data from multiple data silos. It's also using an ontology of more than 400,000 concepts in the health sector with one of its customers.[4] Similarly, a range of ontologies for other sectors has been created. (See, for example, http://protegewiki.stanford.edu/index.php/Protege_Ontology_Library.) The W3C's Semantic Web Education and Outreach working group has published a set of use cases from industrial organizations using semantic technology today (see www.w3.org/2001/sw/sweo/public/usecases).

> Users would prefer a software agent to automatically search the Internet on their behalf, comparing multiple products or services until it matches the requirements specified.

An important recent development that will affect the uptake of semantic technology is a W3C initiative called Gleaning Resource Descriptions from Dialects of Languages (see www.w3.org/TR/grddl-primer). GRDDL aims to provide a link between the Semantic Web and microformats communities. With GRDDL, software can automatically extract information from structured Web pages to make it part of the Semantic Web (essentially by converting it to RDF). Those accustomed to expressing structured data with microformats in XHTML, for example, can thus make their existing data more valuable by porting it to the Semantic Web at a very low cost.

In parallel with this activity by end-user organizations, there has been increasing activity in the venture-capital sector. There are many new start-ups in this area, including

- Hakia, Powerset, Siderean, and Ontotext in semantic search,
- Metatomix and Ontoprise in the information and process integration space, and
- Radar Networks' recently announced Twine semantic-social-networks offering.

Larger, more established vendors are also active, including Oracle with RDF support in Oracle 10g (www.oracle.com/technology/tech/semantic_technologies), IBM with Omnifind, and Xerox with FactSpotter.

In short, as Gartner recently put it, "metadata … will drive interoperability, automation, cost cutting, better search capabilities, and new business opportunities."[5]

Clearly, a potential barrier to uptake of Semantic Web technology is the effort required to mark up Web information with semantic annotations. However, projects such as SEKT and IBM's Unstructured Information Management Architecture are addressing this, deploying techniques from disciplines such as computational linguistics and data mining to derive ontologies and metadata automatically or at least semiautomatically.

It won't be the case that only one ontology will be developed for a given domain, making it necessary to mediate between different (but related) ontologies. Ontology mediation is an ongoing research area. Ontologies will typically evolve over time, and managing this process is a current research issue.

Ontologies will be developed for different domains and applications. Importantly, however, RDF and OWL (recommended W3C standards for creating ontologies) allow importing whole or partial ontologies into new ontologies. This encourages reuse

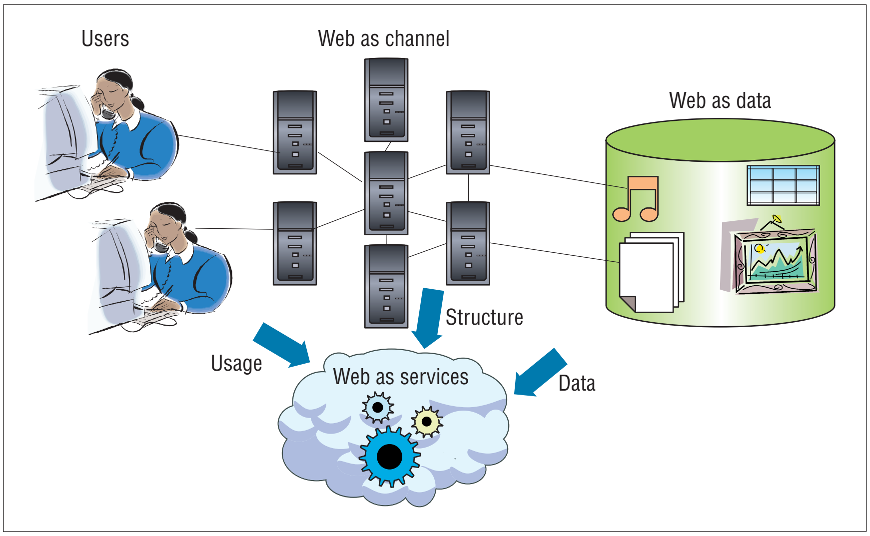


Figure 4. The Web and its interrelations.

of existing ontologies where appropriate and helps avoid the proliferation of similar ones.

References

1. "The World of Content Management Is Changing," *Information InDepth*, Oracle, Feb. 2007, www.oracle.com/newsletters/information-insight/content-management/feb-07/forrester-future.html.
2. D. Roman et al., "2006 Semantic Web Services—Approaches and Perspectives," *Semantic Web Technologies*, J. Davies, R. Studer, and P. Warren, eds., Wiley, 2006, pp. 191–234.
3. M. Voelkel et al., "Semantic Mediawiki," *Proc. 15th Int'l Conf. World Wide Web*, ACM Press, 2006, pp. 585–594.
4. C. Wroe, "Is Semantic Web Technology Ready for Healthcare?" *Proc. 3rd European Semantic Web Conf.*, Industry Forum, vol. 194, A. Leger et al., eds, CEUR-WS, 2006, http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-194/paper2.pdf.
5. A. Linden, *Semantic Web Drives Data Management, Automation, and Knowledge Discovery*, Gartner, Mar. 2005.

Search, Web 2.0, and the Semantic Web

Ricardo Baeza-Yates, Peter Mika, and Hugo Zaragoza, *Yahoo! Research*

With the 10th anniversary of RDF approaching, it might be time to reflect on the Semantic Web's successes and failures. Successes are easy to point out: it's enough to glance through the winning applications of the yearly Semantic Web Challenge (see http://challenge.semanticweb.org) to see the many areas where semantic technologies have made a difference. These include applications in a variety of well-defined expert domains, where the power of semantics has been successfully exploited in interconnecting heterogeneous data sources and finding new insights through reasoning. However, because failures can be more telling than successes, we'll focus on one particular area where the Semantic Web failed to achieve the impact envisioned a decade ago: search on the World "Wild" Web.

The importance of search

Why is search important in general, and what's the role of semantics? As figure 4 illustrates, the Web can be largely conceived as the interplay of three major components: a channel, data, and services. We can view the Web as a channel that connects people to each other and to a wealth of data and services. Data ranges from the weakly structured, hardly processable textual contributions of Web users to highly structured and curated databases. Many services are available on the Web, most of them building on some data or other, but chief among them is search.

Why is search crucial? Although users can access data directly through the channel (for instance, by browsing or sharing links Web 2.0-style), they can't synthesize the data without machine processing. On the other hand, synthesizing is what search engines do: in a fraction of a second, they pull together pieces of information that are relevant to any ad hoc need the user might have. They do this not only on the basis of the data, but also on their observation of Web usage patterns as represented by query logs and linking patterns, for example.

Given this image, we can easily identify the current focus of Semantic Web research: recasting the Web by providing methods to add semantics to data, either manually or automatically, thereby moving the Web toward easier machine processing. In contrast, Web 2.0 has focused on shaping the Web as a channel by bringing in new modalities of Web usage, which has also resulted in more user-generated content. Missing, however, is the exploitation of these improvements in terms of services—in particular, improving the synthesizing power of search engines and building additional capabilities on top of data, such as analysis, planning, design, and prediction. In search parlance, this would let search engines move away from document retrieval to directly addressing the anomalous state of knowledge that triggered the user's action.

Adding structure to the Web

We at Yahoo! Research identify the emergence of structure as one of the four overarching trends that will shape the future of online interactions:

> We believe that capturing and exploiting such structure is a key to next-generation Web applications, including search and advertising ... We see a growing stream of user-contributed content becoming a key resource. In many cases, it conforms to a large and growing set of schemas meant to incorporate specifics of either data type [...] or domain data ... As a variant of user-contributed content, we see user-contributed metadata in the form of reviews, tags, bookmarks, ratings, or even clicks.[1]

An original aim of semantic technologies is to structure information on the Web, so we expect these technologies to play an important role in the development of next-generation search services. In particular, we think it's very probable that the Semantic Web community's data representation efforts will affect search and advertising technology. In the short term, crawlers will likely begin to accept information directly in RDF-like formats to enable search services to take a more active role in the presentation of results.

Traditionally, search services have simply presented snippets of the found objects to the user. But more and more, search engines are tailoring their results to the query's intent so the user can act on them. For example, consider the results we obtained when issuing the query "Barcelona hotel" in Yahoo! Web search. The search engine presents lists of hotels by price and class, and one click lets us enter a particular hotel's reservation site. One way to achieve this integration, of course, is by agreeing on a particular data schema for this transaction. But this greatly limits the number of queries that can be resolved and the number of resources that can be catalogued and searched. Standardized semantic technologies are letting numerous providers make their resources available for search services to act upon. We believe that ultimately authors will be motivated to adopt these technologies to make it easier for users to find and use their resources.

As we predicted three years ago, Web 2.0 provides opportunities that complement those offered by semantic technologies.[2] Tasks that are difficult for machines on the Semantic Web (in particular, understanding natural language descriptions and creating structured summaries and annotations) are effectively solved in Web 2.0 ecosystems through the aggregate work of users. Such users are motivated by the desire for self-organization, social factors, and other rewards such as the enjoyment of games. Sites such as Flickr or del.icio.us have shown that the right incentives and interfaces can seduce people into describing the site's resources using tags. These tags, as well as queries in search engines, represent the Web's language in the form of explicit and implicit folksonomies.[3]

If there's any sense in talking about the Web that will come after Web 2.0, we believe that it will integrate and use all the experience that can be captured from both traditional content (Web 1.0) and user-generated content and metadata (Web 2.0), as well as data about this content's usage. Although this last part still receives less attention from the Semantic Web community, we believe that browsing, clicking, and, in particular, querying, encode semantic information. Relating all this data will let us find the implicit wisdom or knowledge of the social networks that create, synthesize, and consume Web content. Extracting new information and new relations lets us improve the user experience, creating a virtuous circuit of continuous user-driven design. However, we can go further and bootstrap the inferred information into important applications such as search or generate pseudosemantic resources that other applications can use later.[4]

The Semantic Web and information retrieval: Bridging the gap

Let's return to the main question: Why has the Semantic Web had so little effect on search services? And, even more worrying, why does it have such little presence in the agendas of fundamental information-retrieval research programs, search engine designers, and search startups? What's stopping researchers in the IR and Semantic Web communities from moving more bravely in the direction of the promised next-generation search engines?

> What's stopping researchers in the IR and Semantic Web communities from moving more bravely in the direction of the promised next-generation search engines?

We put forward three possible reasons: First, this integration is an extremely hard scientific problem. Second, the Web imposes hard scalability and performance restrictions. Third, there's a cultural divide between the Semantic Web and IR disciplines.

First, realizing the Semantic Web would require solving extraordinarily difficult problems in the areas of knowledge representation, natural language understanding, and semantics—problems that have been recognized and studied for years in IR and other fields, with partial success. It's important to revisit old problems and try to tackle them from novel perspectives. In this sense, we find the Semantic Web program extremely beneficial. Nevertheless, scientific breakthroughs come in small doses, and it's important to understand our limitations.

Second, the Web is for all intents and purposes infinite, with wide use of dynamic page generation. At the same time, users are becoming more impatient. This imposes on any massive Web application—in particular, search—scalability and performance constraints that increase by an order of magnitude every few years. Few technologies have been able to keep up this pace, although the fast improvement of hardware has helped.

Third, a clear cultural divide exists between the IR and Semantic Web communities. IR conferences prove difficult for researchers from other fields to attend, owing largely to a strong emphasis on methodology and, in particular, evaluation. Consequently, existing work on ontology-based IR focuses on smaller subtasks such as query expansion using ontologies or improving search result presentation, but there's little work on reshaping the IR core. In the commercial space, innovation requires courage in the towering presence of the large Web search providers. Also, as Semantic Web research continues to experience considerable growth and attract significant funding for basic research, members of the community feel less compelled to even attempt breaking these barriers, real or imagined, despite the significant economic motivators. IR research is strongly driven by a problem, whereas Semantic Web research is driven by a solution. Metaphorically speaking, Semantic Web researchers are like the hobbyist toolsmith who has the idea for the perfect tool and resents compromises on the design. However, IR, as a customer, is interested in buying a hammer that might not be perfect but can drive nails quickly and precisely.

The role of relevance evaluation exemplifies this divide. At the heart of IR lies the notion of relevance: only a human judge can inform us of a document's true degree of relevance with respect to his or her query in a given context. Because IR systems try to predict this relevance, research ideas in IR can (and should) be evaluated by comparing such predictions against human judgments. For this, standard data sets and baselines have painstakingly been created so that new ideas can be tested and evaluated.[5,6]

As it turns out, in terms of relevance, existing IR technology is very hard to beat. Well before the Semantic Web was formulated, IR researchers tried to apply natural language processing and knowledge representation techniques without much success. However, many of the early failures should be revisited because NLP technology has changed a lot in the last 10 years.[7]

More importantly, the notion of

"relevance" itself is changing. In the last few years, the IR field has been undergoing an important change in focus, and it has become increasingly interested in new problems, new formulations, and new forms of search and evaluation. Trends include, for example, understanding the intention behind a query, shifting from explicit retrieval to implicit information provision based on the task context, and integrating text and multimedia. So, not only are the tools improving, but there are needs for new ones. Sooner or later, these changes will necessarily bring the Semantic Web and IR closer together than in the decade past.

References

1. Yahoo! Research Team, "Content, Metadata, and Behavioral Information: Directions for Yahoo! Research," bulletin, IEEE Computer Society Technical Committee on Data Eng., 2006.
2. P. Mika, "Social Networks and the Semantic Web: The Next Challenge," *IEEE Intelligent Systems*, vol. 20, no. 1, 2005, pp. 82–85.
3. P. Mika, *Social Networks and the Semantic Web*, Springer, 2007.
4. R. Baeza-Yates and A. Tiberi, "Extracting Semantic Relations from Query Logs," *Proc. 13th ACM Sigkdd Int'l Conf. Knowledge Discovery and Data Mining*, ACM Press, 2007, pp. 76–85.
5. R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley, 1999.
6. C.D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge Univ. Press, 2008.
7. R. Baeza-Yates, "Challenges in the Interaction of Natural Language Processing and Information Retrieval," *Proc. 5th Int'l Conf. Intelligent Text Processing and Computational Linguistics* (CICLing 04), LNCS 2945, Springer, 2004, pp. 445–456.

Prospects for Semantic Technologies

Mark Greaves, *Vulcan*

One of the most exciting things about semantic technologies is the recent acceleration in the pace of innovation. Knowledge representation and reasoning (KR) technologies had a long, slow, and sleepy gestation in the AI world of the 1980s and 1990s. In parallel with the Web's growing sophistication, though, the KR world has been completely transformed over the past five years. The development and deployment of usable Web-oriented KR languages, coupled with the availability of Web-scale data and the massive success of search engines, have now made it possible to reasonably predict the type and degree of impact that semantic technologies might have in the next five years, as well as the reasons for this impact.

Core technologies

Modern semantic technologies exist in a cluster. At the center of the cluster are the traditional logic-oriented technologies supporting knowledge representation and reasoning with RDF, OWL, RuleML (Rule Markup Language), SCL (Simplified Common Logic), and other major KR languages. I view these as the core semantic technologies, and I think they represent some of the best work that the AI community has produced. The semantic technology cluster also encompasses an outer layer of supporting technologies, including lightweight taxonomies, tag systems, latent semantics, shallow parsing, information extraction, and data mining technologies. These all in turn depend on a group of more fundamental Web and database technologies. Each of these technologies is important in processing human meanings by computers and so could correctly be called semantic. For this essay, though, I will confine myself to considering the core semantic technologies' five-year prospects.

> AI was limited for decades because the knowledge systems tended to be hand-built systems that, in the final analysis, didn't actually know very much.

Assessing these prospects requires us to consider the total volume of useful knowledge that will be available for these technologies to exploit. After all, even the most sophisticated reasoning technologies won't have an impact if the necessary prior knowledge isn't available. AI was limited for decades because the knowledge systems that were created tended to be hand-built systems that, in the final analysis, didn't actually know very much.

As I write this, though, the Digital Enterprise Research Institute's Sindice engine (see www.sindice.com) indexes over 3 billion total RDF triples. The Linking Open Data project claims about 2 billion triples, virtually all of which come from high-quality sources. These raw numbers are certainly impressive, but the implied growth rate is even more impressive—in 2006, Swoogle found only about 400 million triples. The rapidly improving performance of information extraction, Web mining, and XML/database translation technologies continues to make expressing knowledge in semantically usable formats easier and cheaper.

Finally, Web 2.0-style application design has shown that a significant volume of high-value knowledge can be rapidly authored by individuals responding only to the social urges to show off, share, and connect. Semantic wikis' growing success suggests that this Web 2.0 experience is generalizable—that a great deal of useful semantic information can be cheaply authored and maintained using Web 2.0 techniques. With the costs of encoding knowledge quickly dropping, it's a solid bet that adequate knowledge will be available to support major semantic applications in several verticals over the next five years.

Near-term prospects

I can now address the five-year impact of these semantic applications. Let's assume that the fundamental reasoning power of the available semantic technologies won't change significantly in this period. One simple way to measure semantic technologies' impact is to consider the portion of the global spending decisions they could be responsible for. People make conscious resource decisions when they purchase applications, and these decisions are a powerful clue to the value that they place on the technology. (Of course, you can gauge a technology's impact in many other ways. For example, you could describe the possible new scientific advances that would be enabled. The basic economic impact of the semantic technologies, further assuming no additional scientific breakthroughs, is therefore very much a lower bound on their total impact.)

A simple analysis would go like this: Global software spending has been increasing at a rate of about 10 percent per year since 2001 and in 2006 was around US$317 billion.[1] Let's assume that a new type of information technology is economically significant if it can command at least 1 percent of global software spending. This means that to be economically significant, applications powered by semantic technologies would need to account for about $5.1 billion of worldwide software spending per year by 2012. This is a big number, but I believe that the overall prospects for semantic technologies reaching this level of economic activity within the next five years are quite good.

Corporate customers

Two basic types of customer can drive the adoption of semantic technologies. The most visible customers are large corporations and government, who already possess large amounts of well-curated electronic data and sophisticated ideas for how they would like to use it. Their slow-but-steady uptake of semantic technologies mainly supports internal knowledge management and the integration of corporate and supply chain information. Most current semantic technology start-ups target these customers.

However, customers of this size are inherently bureaucratic and risk-averse. Their sales cycles are drawn out, and persuasive use cases demonstrating large value-adds have proven difficult to find. There are large sunk costs and significant legacy issues with existing technologies, the market already has several established technology vendors, and, overall, the whole new technology uptake process in this space is painfully slow. Finally, a major attraction of semantic technologies for this class of customer—that the basic corporate data integration and fusion task can be made significantly faster or cheaper—has turned out to be problematic. In corporate data integration, the largest fraction of time is spent in documenting the semantics of existing data schema and relating them to business workflows. This includes dealing with all the complex real-world messiness of data uncertainty, data cleaning, and the fact that the significant amounts of operational data semantics are often buried in applications—areas in which semantic technology has no particular advantage. Nevertheless, semantic technology has been making inroads in several large KR-intensive verticals, including finance, communications, media, compliance, and government/defense operations.

Consumer-oriented applications

To my mind, a much more promising, high-growth area for semantic technologies is the prospect of a new generation of semantically enhanced, consumer-oriented applications. Obvious targets for these new applications to support are social networking, advertising delivery, shopping, mobile personalization, identity management, and control of personal information. The aggregate economic value of semantic technology in this area derives from these applications' ability to support the decisions and desires of millions of people. Unlike in the large company space, thousands of individual consumers are willing to be early adopters, and consumers have relatively fewer technology commitments and sunk costs. Moreover, semantic technology's strengths are a good match for the kinds of broad, dynamic information spaces that characterize shopping, advertising, and social networking. Navigation, pivoting, consumer-oriented analytics, and processing long-tail queries in these domains is difficult using the traditional RDBMS (relational database management system) technology of fixed data schemas and join tables. Semantic technologies can provide an important advantage in these areas.[2] For these reasons, I expect that the largest component of the increasing valuation of semantic-technology applications will come from consumer-oriented companies, rather than those that concentrate on the enterprise.

> Even the most attractive commercial opportunities will lie fallow without driven, risk-taking entrepreneurs and the venture capital to transform ideas into new companies.

Moving forward

Finally, even the most attractive commercial opportunities will lie fallow without the presence of driven, risk-taking entrepreneurs and the venture capital that can transform their ideas into new companies. Here, again, I'm encouraged. The growing vigor in the commercially oriented conferences in the area (such as the US's SemTech Conference and the European Semantic Technology Conference), coupled with the success of Web 2.0 companies and the emerging Web 3.0 meme, has resulted in a level of excitement and commercial interest that I haven't seen with other knowledge-related technologies. The energy level at these conferences, and especially on their trade show floors, is amazing.

It's clear that the growing availability of valuable semantic data, the promising capabilities of semantic technologies, and the semantic needs of modern customers have produced a fertile field for new companies. The news of this opportunity is spreading among entrepreneurs, and venture capital is now available to back great ideas. In the next five years, I expect to see these factors combine into a group of successful semantic-technology companies with total revenues in the billions of dollars, built on commercializing today's semantic technologies.

References

1. World Information Technology and Services Alliance, *Digital Planet 2006: The Global Information Economy*, 2006, www.witsa.org/digitalplanet/2006/DP2006_ExecSummary.pdf.
2. M. Greaves, "Semantic Web 2.0," *IEEE Intelligent Systems*, vol. 22, no. 2, 2007, pp. 94–96.

Paradigm Shift and New Bets for the Semantic Web

Jose Manuel Gómez-Pérez and Jesús Contreras, *Intelligent Software Components, S.A.*

In the last few years, the Semantic Web—widely envisioned as the next step in the evolution of the Internet toward a meaningful web of content and services—has created growing expectations about the ways in which it will improve our daily lives. The time for promises of a technical and, eventually, social revolution is gone, and we're entering a period where the Semantic Web must live up to those expectations. Mainstream actors must identify compelling applications of the technology that can guarantee a concrete return on investment.

One of the most difficult challenges for

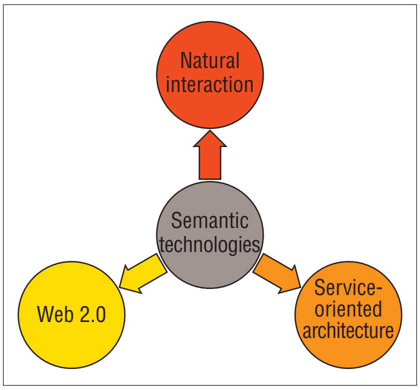


Figure 5. Future trends in semantic technologies.

the Semantic Web is content acquisition.[1] Producing semantically annotated content for real applications is very expensive, making impossible any commercial exploitation of this technology. The recent Web 2.0 initiative has discovered new, collaborative approaches to content annotation and creation. Web 2.0-like applications allow for creation of so called "semantic islands," well-defined and limited domains expressed using semantic languages that might constitute the first version of the next-generation Web. Meanwhile, domain-focused Semantic Web technologies are demonstrating this initiative's feasibility. We introduce two examples in the business-to-consumer (natural interaction) and business-to-business (enterprise application integration using SOA) industries.

Current scenarios

Imagine a world where cars autonomously attend to drivers' requests according to the driving context, where citizens and public administrations speak the same language, and where people can collaborate at work or at home, supported by virtual desktops that intelligently anticipate the services the user requires at each moment. Think of a world where we can produce and collect relevant information from vast heterogeneous sources anytime, anywhere.

These scenarios are becoming part of our daily lives, and semantic technologies are decisively contributing to it. As a matter of fact, semantic technology providers are experiencing an unprecedented pull from the market, which demands solutions to these problems. On the other hand, technological requirements including annotation, natural language processing, multimodality, context awareness, customization, and service and data integration are reaching the required level of maturity to satisfy this demand.

Combined, these drivers largely characterize the so-called Web 2.0 as the read-write Web, which involves more than a billion users interacting all over the globe. These factors converge in a new paradigm for work and leisure, representing new business opportunities for semantic technologies.

From available market studies and personal experience, we've determined that recent and future trends in semantic technologies can be represented as a 3D space with three main intertwining axes: natural interaction, service-oriented architectures, and Web 2.0 (see figure 5).

The user, the treasure

Natural interaction is the most immediate application field for semantic technology. It aims to improve the user experience during interaction with online information systems. The most distinctive feature of natural interaction is the extreme relevance of user interfaces over the remaining dimensions of software.

Natural interaction uses natural language techniques to interpret user requests. This lets users express themselves naturally regardless of domain-specific jargon, while at the same time guaranteeing high precision and recall of query answers. Multimodality also helps natural interaction improve the user experience by letting users move seamlessly between different modes of interaction—from text to voice, or even touch—according to context or preference changes.

An additional dimension in multimodality, customization, lets users adapt the system to better suit their goals. Context management plays a fundamental role here. Context-aware systems increase user satisfaction, producing more accurate answers to the specific circumstances in which the query was formulated. Additionally, context-sensitive systems can "guess" query targets in case of ambiguity and anticipate user requests on the basis of previous experiences.

Natural interaction systems focus on user perception, and explanation features give users feedback about the system's actions. Producing a comprehensive explanation of how a system obtained a result is at least as important as the correctness of the result itself. In this direction, Web development technologies such as Ajax help make Web pages feel more responsive and dynamic, increasing interactivity, speed, functionality, and usability.

Applications of semantic technology to natural interaction span across domains such as e-government and the automotive industry. Currently, most public administrations provide citizens with online access to their services. However, a big gap remains between the language that public administrations use and the way citizens refer to these services, making it difficult to match citizen requests. Using semantic technology in this scenario accomplishes two things: it lets citizens express themselves naturally rather than forcing them to use administrative jargon, and it lets them easily access precise information that satisfies their queries.

In this regard, current developments belong to the first stage of the natural-interaction roadmap (see figure 6). Intelligent search engines use semantic technologies to detect the services that the user's query refers to, such as "I want to move to a new house" or "What guided tours are available?" According to the roadmap, the next step will introduce mobility and voice interfaces to facilitate user interaction, including avatars that will improve user friendliness. In the roadmap's final step, semiautonomous virtual assistants will not only recognize services but also execute them for the user according to predefined policies. Thus, business-to-business integration will occur, involving third-party services such as banks and telecommunications companies.

In the automotive domain, context awareness is becoming a key issue. Current speech control systems require the driver to express commands with high precision to obtain the right answer from the system. This is a source of ambiguity and imprecise system responses. Additionally, a context-aware speech dialogue system can enable the driver to express more complex commands. For example, if the driver asks the navigation system to take him or her to a good restaurant according to the driver's tastes, a context-aware system can recommend a list of nearby restaurants that are open and serve the type of food that the driver has preferred previously.

Driver commands can trigger different actions depending on the context. For example, regular speech-recognition systems can't extract a given command or display useful dialogs from questions such as,

Positioning	Intelligent search engine	Attention to user	User self-service
Interface	• Natural language • User feedback • Question comprehension and answer explanation	• Multimodality • Voice question/answer • Avatars • Mobile access, PDA	• Virtual assistant • Chatbot
Functionality	• Information retrieval based on domain understanding • Synonyms, morphology • Interleaving with traditional search engine	• Dialogue management • User profiling • Context awareness	• Execution of online services
Content scope	• Noncollaborative content development	• Collaborative content development	• Electronic transactions
Technology	• The Semantic Web and ontologies • Natural language processing • Ajax	• Voice recognition • Text2Speech • Dialogue systems • Personal identification • VOIP • Mobility	• Digital signature • Business-to-business integration • Messenger

Figure 6. The roadmap to natural interaction.

"What is this?" Without context information, this expression is meaningless for the car system. However, if context is appropriately formalized, it's possible to ground the question in the domain and provide a useful response. For example, if the user sees a flashing light on the dashboard and asks, "What is this?" the system will be able to answer, "Please, stop immediately. You have run out of oil." Or, it could provide a set of suitable actions because this event has been previously logged and an ontology describes what actions to take under such circumstances.

A Web of service is out there

In the nearer term, semantic SOA is also on its way to wide uptake in the market. The Web is evolving towards a Web of services at different levels, appealing to both individuals and organizations. This Web of services focuses on the functionality of the applications exposed, abstracting from the underlying technology and favoring maximal uptake by potential users. However, these services need exposure to users outside their organizations. Semantic technologies are key to producing comprehensive, semantic descriptions of these services, providing for fine-tuned service discovery and combination capabilities.

SOA applications have had demonstrable effects in large corporations and will continue to do so in the future. SOA promises a flexible IT infrastructure that can react to business changes more quickly than the classic monolithic IT systems. However, surveys and experience show that SOA is used today primarily for internal rather than external integration. Using the Web to expose available services will make it easier for companies to interact, forming business ecosystems.

On the other hand, semantic technologies that aim to integrate distributed applications and services on the user side provide much more realistic business opportunities than those that aim to integrate large corporations' business processes. This kind of service mashup will be reflected on virtual desktops that support users both at home and work, helping them to manage tax payments, plan vacations, or arrange business meetings with the best available flight connections. Once again, as with natural interaction, customization based on context awareness and user profiling will be a key issue.

However, a trust issue must be solved. How can we be sure that these services actually do what they are meant to and behave like they are supposed to? Additionally, after executing a process comprising several complex services in our hypothetical context-based, self-customized service mashup, users will need to understand how their input data was transformed into the resulting output. Furthermore, in cases where different contexts produce different outputs for the same process and input—for instance, in the case of an online auction—producing such explanations is even harder.

Provenance will play a fundamental role in this regard, increasing the user's understanding of the interactions between services by explaining process execution in a way closer to how users perceive or reason about a given problem and monitoring the execution of such services. Semantic components such as problem-solving methods[2] and ontologies will support this goal. We will use PSMs as high-level, domain-independent knowledge templates that help interpret the execution of past processes composed by the aggregation of different services.

So, where does Web 2.0 fit in?

The Web 2.0 philosophy is simply orthogonal to natural interaction and SOA and intersects with both. Because it follows a "democratic" model rather than the "oligarchic" model of Web 1.0, Web 2.0 proposes a scenario where potentially everybody with Internet access can not only consume but also produce online content, both as data and services. This is reflected in both our natural interaction roadmap and in the future of semantic SOA, where flagships of Web 2.0-like customization, context awareness, collaboration, or mashups shine on their own.

Semantic technologies are experiencing a new boost, resulting from human beings' basic needs to collaborate and express their ideas. These recent innovations are combining with existing applications of the technology, now stimulated by this paradigm shift and resulting in new business opportunities. However, this will create new problems. For example, the appearance of microformats has proven how far a little semantics can go (see http://microformats.org). But what if the sheer amount of heterogeneous Web 2.0 resources—a blend of informal knowledge repositories such as tag clouds and folksonomies, stemming from collaboration among zillions of individuals—dramatically increases as expected? Will Web 2.0 devices scale appropriately? Probably not. The Semantic Web's formal methods of granting comprehensive information access and management will step in with renovated strength, and we'll be talking about a new paradigm shift: Web 3.0.

On the other hand, market potential for semantic technology is closely related to the actual commercial uptake of ontologies, and to working Semantic Web applications. Ontologies are at the heart of semantic technologies, because they're the instrument used for representing the meaning of the domains concerned. As we've seen, in a strongly interconnected world, key aspects of ontologies will include collaboration, context, dynamics, and human factors. Projects such as NeOn (see www.neon-project.org) are already addressing these factors.

Acknowledgments

We would like to acknowledge the IST projects OntoGrid (IST-2004-511513), NeOn (IST-2005-027595), and SUPER (IST-2005-026850).

References

1. V.R. Benjamins et al., "Six Challenges for the Semantic Web," *Proc. Semantic Web Workshop,* held at 8th Int'l Conf. Principles of Knowledge Representation and Reasoning (KR 02); www.cs.man.ac.uk/~ocorcho/documents/KRR2002WS_BenjaminsEtAl.pdf.
2. V.R. Benjamins and D. Fensel, "Editorial: Problem-Solving Methods," *Int'l J. Human-Computer Studies*, vol. 49, no. 4, 1998, pp. 305–313, http://citeseer.ist.psu.edu/benjamins98editorial.html.

Toward a Service Web: Integrating the Semantic Web and Service Orientation

John Domingue, *Knowledge Media Institute, Open University*
Dieter Fensel, *University of Innsbruck*

By a variety of metrics, the Semantic Web is doing very well. For example, the 2006 Gartner report included the "Corporate Semantic Web" as one of three key technology themes, and in 2007, Tim Berners-Lee briefed a hearing of the US Congress subcommittee on the topic. The current state of the Semantic Web is said to be comparable with that of the early days of the Web. Until very recently, growth has been linear, but signs of exponential growth are beginning to emerge.

Given this promising situation, we should ask ourselves where and how we expect the Semantic Web to continue to grow in the near- to mid-term future. Our view is that the Semantic Web will come to the fore through a comprehensive integration with service orientation. In particular, we envisage that the combination of Semantic Web and SOAs will lead to the creation of a "service Web"—a Web where billions of parties are exposing and consuming billions services seamlessly and transparently and where all types of stakeholders, from large enterprises to SMEs and singleton end users, engage as peers consuming and providing services within a network of equals.

Maintaining millions of services, let alone billions, to cope with environmental and context changes solely through human effort isn't feasible.

Service orientation today

A Web service encapsulates discrete functionality through an interface that's accessible via standard Internet protocols. As such, Web services have proven very popular in industry. The technology abstracts over legacy hardware and software platforms and enables the exposure of commercial services to a potentially wide audience. Web services have also proven to be useful in B2B scenarios. Consequently, service-oriented frameworks such as SAP's NetWeaver (see www.sap.com/platform/netweaver) now dominate IT solutions for large enterprises.

Web services have, however, failed to make a significant impact on the Web. Current estimates are that the Web currently contains 30 billion pages, with 10 million new static pages added each day. In contrast, only 12,000 true Web services—Web services with active endpoints described with a Web Services Description Language (WSDL) file—currently exist on the Web. SOA solutions are thus restricted in their application context to being an in-house solution for corporations. Realizing service orientation's full potential will require integrating SOA with both semantic and Web technologies. Here, we briefly outline several principles that must be followed to successfully achieve this.

Enhancing SOA with semantic technologies

Current standards for describing Web services use syntactic (XML-based) notations such as WSDL. Because these descriptions aren't machine readable, IT staff must carry out all of the tasks associated with creating and maintaining Web service-based applications. The requirement for specialist workers to be involved in all points in the Web service life cycle causes numerous problems, the overriding one being the lack of scalability. Maintaining millions of services, let alone billions, to cope with environmental and context changes solely through human effort isn't feasible.

Tasks such as Web service discovery, composition, and invocation can be automated to a certain extent, so we can overcome this problem by applying semantic technologies. For example, semantics enable programs to access services through a description of offered capability rather than as an end point. The use of semantics thus forms a scalable access layer over Web service data and processes. Key to this is the creation of a semantic service bus that supports service usage through semantic formulations.

More generally, combining service orientation with semantics will require developing a comprehensive framework and architecture following several principles.

The service-oriented principle

This principle captures a distinct approach to analysis, design, and implementation and introduces specific subprinciples that govern aspects of communication, architecture, and processing logic. These include service reusability, loose coupling, abstraction, composability, autonomy, statelessness, and discoverability. With respect to service orientation, which enables a service-level view of an organization, there's a need to further distinguish services along several dimensions.

First, services can be distinguished according to the functionality they provide within an architecture, namely business and middleware services. Business services are services that various service providers supply through their back-end systems. Additionally, they're the subject of integration and interoperation within the architecture and can provide a certain value for users (such as booking a hotel room). On the other hand, middleware services are the main facilitators for the integration and interoperation of business services (such as discovery and mediation).

Second, services can be differentiated according to their abstraction level within an architecture—namely, Web services and services. A Web service is a general service that might take several forms when instantiated (such as purchasing a flight), whereas a service is an actual instance of the Web service that is consumed by and provides a concrete value for a user (such as the purchase of a particular flight from Innsbruck to Vienna). This distinction would be used for business services in the architecture.

The semantic principle

Combining semantics with service orientation lets you define scalable, semantically rich, formal service models founded on ontologies. This allows the total or partial automation of tasks such as service discovery, contracting, negotiation, mediation, composition, and invocation. A semantic service-oriented approach to the modeling and implementation of service-based applications will enable the scalable and seamless interoperation, reusability, discovery, and composition of the components to be used or reused.

The distributed principle

The distributed principle is the process of aggregating several computing entities' power to collaboratively run a single task, transparently and coherently, so that those entities appear as a single centralized system. Applying this principle to the architecture middleware system will allow the transparent distribution of components over the network so that executing processes running in a middleware can be scaled across numerous physical servers over a network. The distributed principle would also apply to business services, letting running processes span across enterprises distributed over a network.

The user-centric principle

The user-centric principle puts the user in the center of the architecture. It refers to concepts such as personalizing business services, facilitating service usability, promoting multichannel access and service delivery, building trust, and achieving efficiency, accountability, and responsiveness according to user requirements. It will also facilitate the seamless implementation of business processes across organizational boundaries.

Combining semantics with service orientation lets you define scalable, semantically rich, formal service models founded on ontologies.

SOA and the Web

The Web can be understood as a collection of principles and highly scalable means for electronic publication. We believe that applying these principles to service orientation will lead to a global, dynamically changing environment of services accessible for third-party usage. Within this environment, services will undergo many changes; for instance, they will

- appear, disappear, and change location;
- initially be free, then transform to pay-per-use; and
- occasionally be blocked, out of service, or inspected for antitrust violations.

The provision of Web-based lightweight integration infrastructures will facilitate openness and easy adoption for both the service provider and consumer. Moreover, the adoption of Web principles will open service orientation beyond the boundaries of single organizations. We advocate several main Web principles for widespread acceptance.

Openness

In principle, anybody can contribute to the Web as a provider or consumer of information. Openness is a major necessity for a Web-scale environment. Using the infrastructure must be as simple, smooth, and unrestricted as possible for both service providers and users.

Interoperability

Interoperability should be provided through the integration of heterogeneous proprietary and legacy solutions through a common interface. Interoperability on the Web is platform and vendor neutral to let all providers and requesters of information to participate on an even basis.

Decentralized changeability and dynamicity

Content can appear, be modified, or disappear in an ad hoc fashion. That is, the provisioning and modification of content is under the distributed control of peers rather than controlled by a central authority. Central control would hamper access and therefore scalability. In this respect, an element of chaos or "messiness" should be tolerated.

Automation of central mechanisms to route requests or responses

Manually generated repositories are inherently nonscalable and costly, and they quickly become outdated. You could argue that Web sites such as Google centralize control or access. However, what they actually implement is an abstraction process for accessing and caching information. In the Web's early days, sites were accessed by magic numbers (and later by magic names), and lists of bookmarked pages were considered valuable intellectual property. Through search engines, literal numbers and names have been replaced by keyword retrieval and ranking based on relevance factors. So, we can view the

Google search form as an abstraction over the browser address bar. (Still, search engines such as Google can be misused if central control is used to manipulate access to implement censorship or promote commercial interests.)

Enabling n : m relationships to maximize interaction

In contrast to email, where the content is targeted at specific receivers, the Web is based on anonymous distribution through publication. In principle, information is disseminated to any potential reader, which email can achieve only through spam-like processes.

Integrating these principles into SOA will require

- A worldwide addressing schema. A simple form for this is a unique name and, more elaborately, a description of the capability of a service (that is, the degree to which it can be used to achieve a certain goal).
- A transport layer (a protocol) to transmit requests for and the results of service invocations.
- A platform-independent interface to process (client) service requests and access to service implementations.

From a purely technological viewpoint, the mechanisms used in Web 2.0 are similar to the "standard" Web. However, Web 2.0 brings numerous Web-related concerns to the fore that can facilitate the creation of a service Web, including

- blurring the distinction between content consumer and content provider,
- providing a lightweight semantic mechanism through tagging,
- blurring the distinction between service consumer and service provider, and
- blurring the distinction between machine- and human-based computation.

Web 2.0 technology can thus provide a means to generate and access the semantic service layer outlined above. Incorporating human interaction and cooperation in a comprehensive fashion creates a route to solving tasks such as service ranking or mediation, which otherwise remain computationally infeasible. In several scenarios, Web 2.0 and human computing approaches, together with their underlying social consensus-building mechanisms, have proven the potential of combining human-provided services with services provided via automated reasoning. In our view, the transparent provisioning of services abstracting over whether the underlying provider is a human or a machine will significantly increase the overall quality of services available to the end user.

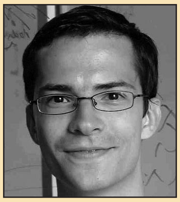
A promising future

From our point of view, the Semantic Web's future is very promising. Moreover, we believe that the successful integration of Semantic Web and service-oriented technologies will form the main pillar of the software architecture of the next generation of computing infrastructure. We envision a transformation of the Web from a Web of static data to a global computational resource that truly meets the needs of its billions of users, placing computing and programming within a services layer to facilitate computing's real goal: placing problem solving in the hands of end users through a properly balanced cooperative approach. ■

V. Richard Benjamins is the director of technological strategy at Telefonica R&D. Contact him at rbenjamins@tid.es.

John Davies leads the Semantic Technology research group at BT and is project director of the Active EU integrated project, co-organizer of the European Semantic Web Conference series, and chairman of the European Semantic Technology Conference (www.estc2007.org). Contact him at john.nj.davies@bt.com.

Ricardo Baeza-Yates is VP of Yahoo! Research for Europe and Latin America. Contact him at ricardo.baeza@barcelonamedia.org.



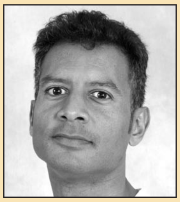
Peter Mika is a researcher at Yahoo! Research Barcelona. Contact him at pmika@yahoo-inc.com.

Hugo Zaragoza is a senior researcher at Yahoo! Research Barcelona. Contact him at hugoz@yahoo-inc.com.

Mark Greaves is the director of knowledge systems research at Vulcan. Contact him at markg@vulcan.com.

Jose Manuel Gómez-Pérez is the research manager at Intelligent Software Components in Madrid. Contact him at jmgomez@isoco.com.

Jesús Contreras is the R&D director at Intelligent Software Components in Madrid. Contact him at jcontreras@isoco.com.



John Domingue is the deputy director of the Knowledge Media Institute at the Open University, UK. Contact him at j.b.domingue@open.ac.uk.

Dieter Fensel is the director of the Digital Enterprise Research Institute at the Leopold Franzens University of Innsbruck. Contact him at dieter.fensel@deri.at.