Hurdles in the Business Case for the Semantic Web

by

David Provost

Master of Business Administration, Boston University (1990)
Bachelor of Science in Business Administration, Boston University (1983)

Submitted to the Sloan School of Management in Partial
Fulfillment of the Requirements for the Degree of

Master of Science in the Management of Technology

at the

Massachusetts Institute of Technology

June 2004

Signature of Author:____
MIT Sloan School of Management
May 7, 2004

Certified by:____
Brian Subirana, Thesis Advisor
Visiting Associate Professor of Information Technologies, MIT Sloan

Certified by:____
James M. Utterback, Thesis Advisor
David J. McGrath jr. Professor of Management and Innovation, MIT Sloan

Certified by:____
Timothy J. Berners-Lee, Thesis Reader
Principal Research Scientist and 3Com Founders Chair,
MIT Computer Science and Artificial Intelligence Laboratory

Accepted by:____
David A. Weber
Director, MIT Sloan Management of Technology Program

Hurdles in the Business Case for the Semantic Web
by
David Provost

Submitted to the Sloan School of Management in Partial
Fulfillment of the Requirements for the Degree of
Master of Science in the Management of Technology

at the
Massachusetts Institute of Technology
May 7, 2004

Abstract

The nuclear winter that filled the vacuum created by the Internet implosion was characterized by highly conservative investments in new technologies. This was particularly true for Internet- and Web-oriented technologies since after all, being a believer just wasn't as popular as it used to be. However, life, business, and science go on, and the Web is no exception.

This thesis will examine hurdles in the business case for the Semantic Web. In one sense, the Semantic Web is an extension or enhancement of the existing World Wide Web (Web). As we know it today, the Web is a rich medium that allows humans to express themselves, learn, interact, and reach an audience that was a pipe dream just a decade ago.

At the same time, the Web is of limited utility to computers (machines). For example, a human being could easily recognize a postal address or the specifications of an order for steel; a machine could not. To a machine, these data would simply be elements to be rendered and displayed on a monitor, with no intrinsic or cumulative meaning. In this sense, one of the goals set for the Semantic Web is to create meaning and utility for machines that allows for interpretation and action with far less human intervention.

Issues related to the challenges, practicalities, theories and opportunities of the Semantic Web will be discussed. In the process, hopefully, this thesis will identify some of the stepping stones in building a business case for this evolution. Notably, today's comments regarding the Semantic Web sound very similar to what was once said about the practicalities of eBusiness and the likelihood of its adoption.

Thesis Advisor: Brian Subirana
Visiting Associate Professor of Information Technologies, MIT Sloan

Thesis Advisor: James M. Utterback
David J. McGrath jr. Professor of Management and Innovation, MIT Sloan

Thesis Reader: Timothy J. Berners-Lee
Principal Research Scientist and 3Com Founders Chair,
MIT Computer Science and Artificial Intelligence Laboratory

Table of Contents

1 The Semantic Web and Its Components

The bursting of the Internet bubble in 2000 shifted the focus of popular opinion elsewhere. Despite the "Chicken Little" commentary, the World Wide Web has continued to evolve and is now poised to enter a new era – one defined by what has been termed by the creator of the Web, Tim Berners-Lee, as the Semantic Web. The technical underpinnings of this next phase are based upon years of rigorous research in a range of computing disciplines. The Semantic Web has begun the process of leaving research labs and is entering commercial development and production environments. In the simplest sense, through the Semantic Web, computers can begin to understand the meaning of what they presently merely project onto displays for the benefit of human comprehension.

This thesis examines the hurdles associated with the adoption of Semantic Web technology from a corporate, utilitarian, and managerial perspective. While not an exhaustive list, issues considered include:

- How the Web infrastructure reduces investment requirements and may serve as a model for adoption.
- Understanding the increased reach of the Semantic Web to include real, material objects - not just virtual objects.
- Whether the Semantic Web is an added layer of complexity or a tool for simplification.
- The Semantic Web's role in the management and evolution of enterprise IT infrastructure.
- The soundness of the underlying logic and research supporting the Semantic Web.
- Reasons for undertaking the adoption of the Semantic Web.

The emergence of the Semantic Web may represent a remarkable opportunity to study the adoption of a new and compelling technology. Given the early stages of the commercialization of this technology, it's not surprising to find a lack of real-world case histories. This thesis is bound to leave some issues unaddressed, but hopefully it will serve as a stepping stone to more exhaustive study in the future. While this represents an opportunity for further research, it's also unfortunate that these kinds of experiences can't be included here.

1.1 The Semantic Web Defined

The Semantic Web refers to the next evolutionary phase of the Web. Presently, the Web is easily read and understood by humans. For instance, if you were planning a visit to a museum and were looking at its address and schedule online, you would readily understand where the museum is located and when it is open. Conversely, your computer would only know that this information was alphanumeric text (and possibly graphics) called from one location and displayed at another. As defined by Berners-Lee, Hendler and Lassila, "The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation."[1]

In a Semantic Web environment, your browser might let you know that you could import the address information into your Personal Information Manager (PIM) and if you indicated a particular day, it could highlight the museum's hours of operation on your desktop calendar. Furthermore, if you had access privileges to a friend's calendar who also wanted to go, you could identify the times and dates where you were both free to visit together. As a result, the Semantic Web eliminates the need to cut and paste the address information into your PIM, contact your friend to begin the negotiation process of selecting a mutually agreeable time to meet at the museum, and then finalize these details on your calendar. Instead, you could specify a time range in which you'd like to visit this museum with that particular friend and within moments, have a clear idea of when this visit would be feasible. Human intervention would be required for the final confirmation of these plans and little else.

The aforementioned example could be expanded to demonstrate how corporate travel plans can be constructed that would allow a work group to travel to various locations and meet at mutually agreeable times. A more sophisticated example might comprise database integration. For instance, in assigning meaning, semantic technology uses "ontologies," which can be thought of as a catalog of terms, a glossary, or a thesaurus.[2] They allow disparate databases to specify the meaning of displayed fields in a declarative manner. Once each database links to the appropriate ontological term, a common understanding can be reached and data integration can occur.

The above example is based on a declarative model. For instance, there are two ways to convert temperature data from Fahrenheit to Centigrade. One is procedural, requiring that a value be input, a conversion procedure executed, and a value output and then re-incorporated into a destination file. Conversely, the conversion Fahrenheit = 9/5 Centigrade + 32 can be written in a declarative fashion and stored in a Web file. This declaration is then easily re-usable by anything that requires the use of the value in question.[3] An example of this re-usability might be if the relationship between Kelvin and Centigrade were also defined in a declarative fashion. Knowing these two facts, it would be possible to combine the two declarative statements above to convert Fahrenheit data to Kelvin without explicitly knowing the formula.

There are several other examples of how the Semantic Web might be put to use within corporations, including data integration (per the example above), application interoperability, computational service composition, knowledge representation and

[1] Tim Berners-Lee, James Hendler, Ora Lassila, "The Semantic Web", *Scientific American*, May 2001.
[2] Deborah L McGuinness,., "Ontologies Come of Age", in *Spinning the Semantic Web*, eds. Dieter Fensel et al. (Cambridge, The MIT Press, 2003).
[3] Tim Berners-Lee, "The Semantic Web", Presentation, March 5, 2004, MIT Sloan School of Management, Cambridge, MA.

management, utility computing[4], enhancing the adaptability of enterprise IT infrastructure and Web service discovery.

1.2 XML's Continued Adoption is Critical

XML 1.0 (eXtensible Markup Language) was issued as a recommendation (global standard) by the W3C (World Wide Web Consortium) on February 10, 1998. RDF (Resource Description Framework), and OWL (Web Ontology Language) were approved as recommendations on February 10, 2004. These three standards represent the basic building blocks of the Semantic Web, with XML being perhaps the most critical. Let's begin with a discussion of XML.

"XML lets everyone create their own tags—hidden labels such as or that annotate Web pages or sections of text on a page."[5] These tags are out of sight to the casual Web surfer and instead designed to be acted upon by scripts, or programs. One very visible example of a Web site using XML is CNET.[6] In this case, CNET is using XML for Really Simple Syndication (RSS), which allows non-CNET Web sites to publish CNET RSS feeds related to product reviews, news, downloads and more.

An even more visible example is eBay and its planned use of XML to link with Microsoft applications via Web services. Specifically, the intent is to allow sellers to enter items into an Excel spreadsheet, save the file in XML format and then send it to eBay for listing. The reverse is intended as well – auction results could be downloaded from eBay in XML format, analyzed, and stored for record-keeping purposes.[7]

The other half of this commitment to XML is highlighted on Microsoft's own Web site, where the company promotes its Office 2003 software. Obviously, Office is used in professional as well as home environments. It is the standard in personal productivity suites. It's interesting to note that Microsoft dedicates a significant amount of space to its XML capability when highlighting "what's new" in the Office 2003 software.[8] It would seem that introducing the use of XML in software as ubiquitous as Microsoft Office has the potential to significantly advance the adoption of XML and in turn, further lay the foundation for RDF and OWL.

Of course, there's always the possibility that Microsoft motives for promoting XML aren't entirely altruistic. Through its .Net (codename: Hailstorm) initiative, Microsoft has

---

[4] Steve Ross-Talbot, interview by author, Cambridge, MA, February 12, 2004. The term "utility computing" refers to a software licensing model whereby customers pay only for the extent to which they use the software as opposed to an outright (and traditional) license that provides access regardless of usage levels.

[5] Tim Berners-Lee, James Hendler, Ora Lassila, "The Semantic Web", *Scientific American*, May 2001.

[6] CNET Networks, Inc., "CNET RSS Home", CNET Networks Web site, <http://www.cnet.com/4520-6022-5115113.html>, accessed March 11, 2004.

[7] David Becker, "XML get eBay and Microsoft into bed", Silicon.com, March 2, 2004, <http://www.silicon.com/management/itpro/0,39024675,39118802,00.htm>, accessed March 11, 2004.

[8] Microsoft, Inc. "Office 2003 Editions: Compare the to Previous Versions", Microsoft Inc. Web site <http://www.microsoft.com/office/editions/prodinfo/compare.mspx>, accessed March 11, 2004.

been actively positioning itself as a serious player in XML Web Services. Including this capability in the Office suite only reinforces this effort. The ultimate irony of this effort is that Microsoft will be able to claim that instead of promoting proprietary standards, it's actually embracing open standards.

XML's role in industrial production environments also seems well established as a common means of "…distributing data from one application to another."[9] In fact, some assert that instituting a standard for XML and its resulting adoption represents an achievement that is more than half the battle in transitioning the Semantic Web from high concept to material reality.[10] If so, it would seem that the utility of XML is well accepted and its adoption will continue.

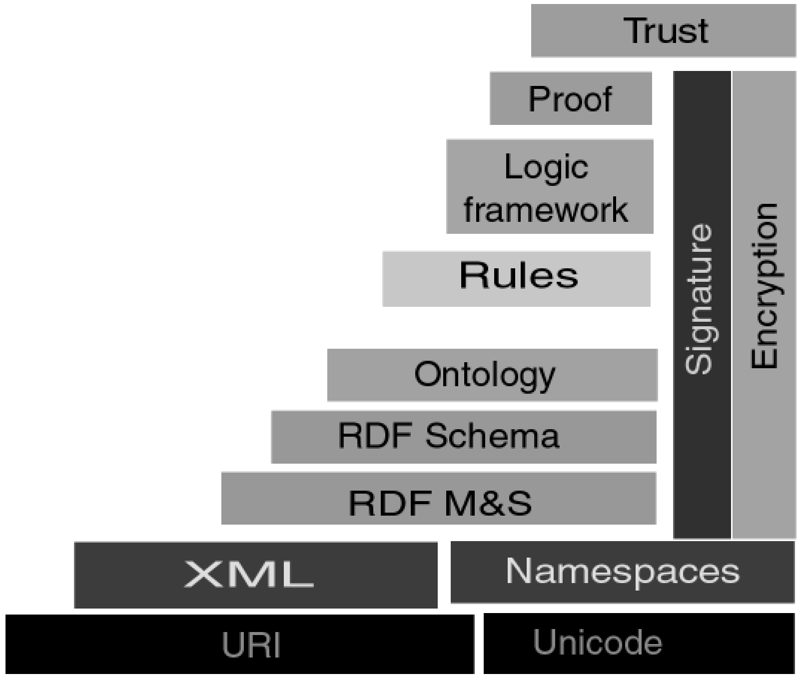
At the same time, occasionally the questions arises "If I use XML, why do I need RDF?" The answer is simple enough – XML describes the structure of documents "but says nothing about what the structures mean."[11] Instead, this meaning is conveyed through the use of RDF and to "the RDF purist, the documents and datasets being expressed in XML and XML Schema can anchor their models with interoperable data."[12] With RDF it's entirely possible to achieve inter-operability and readily exchange data.

1.3 RDF and OWL as Next Steps

It would be an achievement in itself if XML's adoption continued and its use became routine. eBay users and CNET readers would have the ability to perform a range of tasks that presently, they can do only with great difficulty, if at all. But XML has greater utility if additional layers are added to the Semantic Web stack in the form of RDF and OWL. Through the addition of these two layers, it becomes possible to identify the meaning of XML tags in a way that can be read by machines in a consistent fashion and without the confusion that might arise through variations in nomenclature, terminology, etc.

[9] "Future of search – When the web starts thinking for itself", *Information World Review*, (December 2002).
[10] Alex Linden, interview by author, Cambridge, MA, February 16, 2004.
[11] Tim Berners-Lee, James Hendler, Ora Lassila, "The Semantic Web", *Scientific American*, May 2001.
[12] James Hendler and Bijan Parsia. "XML And The Semantic Web: It's Time, Stop Squabbling -- They're Not Incompatible." XML Journal 3 n10 (October 2002): 30.

Figure 1. The Semantic Web Stack



LCS Seminar 2002-09-19, [*http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/slide19-0.html*](http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/slide19-0.html)*, accessed April 8, 2004.*

By introducing RDF, a Web publisher can begin taking further steps in formally relating the various data elements that have been tagged with XML. Doing so allows the meaning of the data as well as the relationships between different elements to begin to take form (as discussed in the earlier temperature conversion example). However, the Web and its elements are authored by a wide range of individuals who may use the same term to mean different things or different terms to mean the same thing. This is where ontologies (created according to the OWL standard) come into play.

In the case of terminology conflicts or confusion, an ontology, occasionally described as a thesaurus, can serve to relate disparate or confusing identifiers that may appear on different Web pages or different databases. Used properly, ontologies define equivalence relationships between terms, e.g. "…term A on my web page is expressing the same concept as term B on your web page."[13] This assertion suggests hope for the integration of disparate data and the composition of multiple software programs into a functioning (and effective) whole.

Putting it all together, XML tags can be applied to any data element. If these tags are related consistently through RDF, the identification and maintenance of relationships between the various tags is ensured. Ontologies then allow differing XML tags organized in a variety of structures through RDF to be unified after identifying the range of equivalencies that exist. Furthermore, since the data reside in all of their original

---

[13] "Future of search – When the web starts thinking for itself", *Information World Review*, (December 2002).

locations (they're not imported into a data warehouse where they're inaccessible to the originating application), when they're updated, these updates become instantly available through the use of consistent XML tags structured via RDF and equilibrated via an ontology(s).

1.4 Lead User Application Scenarios

In the course of researching this thesis, a number of examples were suggested to demonstrate the rationale for the use of Semantic Web technology. Of course, the common element in each of these examples is that the application of Semantic technology is so far unproven. So, without providing solutions, some of these examples are described below:

- **Data Integration and Application Interoperability**

  These two uses of Semantic technology are usually the first mentioned in any discussion of the Semantic Web's utility. There may be small and efficiently operating examples in research laboratories around the world. The real challenge lies in ensuring integration and interoperability in large-scale computing environments. For instance, the value of the Semantic Web will be proven when a Fortune (or Global) 50 corporation finds it more economical to use this technology to ensure the integration and interoperability of its infrastructure. Alternately, if the United States (US) government can do the same across its broad range of agencies, branches, and divisions, then definitely, the Semantic Web can be considered successful. Conveniently, this example includes the issue of scalability when considering the sheer size of the entities in question.

- **User-Configurable IT Infrastructure**

  This example is a bit more abstract, but closely related to the securities trading desk given earlier. Assuming that a user (in this case, a trader) has become sophisticated in the application of IT, he might compose his own set of IT capabilities, designed to provide a specific Quality of Service (QoS). This application would in its entirety be tailored to the requirements of a given customer and the demands of the customer's investment portfolio. Interestingly, this example could easily begin to resemble the habits of a traditional craftsman – that is, while certain fundamentals are almost always universal, the fine tailoring of necessary implements allow that particular professional to enhance his strengths, defend his weaknesses, and in turn provide a truly unique service.

- **Utility Computing**

  In another time and place, this concept might have been known as time sharing, service bureau operation, or application service provision. Some assert strongly that software licensing models are evolving away from outright licenses (or leases for unlimited usage) to those that account for utilization. This "utility" concept meters customer usage and charges on this basis, plus maintenance and any additional services. In the extreme, a customer might choose from several different applications and employ them for the duration of a given project, business cycle, or as otherwise needed. A utility model would allow a customer to avoid substantial capital expenditures, freeing up financial resources to be productively deployed elsewhere in the enterprise. When considering the skills required to provide these types of services, some suggest entities such as large systems integrators will be the first movers. Others suggest existing (surviving) hosted application service providers will lead the charge, as they evolve with market requirements for Web services. Still others offer telecoms (as a means of escaping the constraints of bandwidth provisioning), although these companies may lack the necessary experience with software applications. Mainframe manufacturers have been suggested (as an escape from the binds of big iron) as likely propagators of Semantic Web services, which would build on their revenue momentum in delivering services to complement their hardware. In any case, the current evolution of software licensing may make Semantic technology an attractive means for satisfying customer demands.

2 Method

The method applied to researching this thesis comprises a thorough literature review and field interviews. More specifically, the literature review examined published work in refereed journals, news sources, conference proceedings, Web sites, and presentations. The bulk of the literature addressed the technical aspects of the Semantic Web, typically with some references to business problems or situations.

Universally, there was a lack of hard analysis of Semantic technology applied in live, production environments in a business context. The reason for this paucity is simple enough: the technology is in its early stages and corporate adoption is only just beginning.

Finally, most of the literature search applied to articles published since 2002. According to Lassila and Khushraj[14], this timeframe would be satisfactory to cover the relevant literature with only a few exceptions. The most notable exception is the article by Berners-Lee, Hendler, and Lassila titled "The Semantic Web", published in *Scientific American* in May, 2001.

Field interviews were conducted in person whenever possible, but most often by phone. The individuals represented vendors, researchers, and standard setters. The theme of these interviews centered around reasons corporations should consider the Semantic Web as a technology that might contribute to business performance. Below is a chart of the interviewees:

Figure 2. Interviewees

Name	Title	Organization
Tim Berners-Lee	Principal Research Scientist	Massachusetts Institute of Technology
Frank Careccia	Vice President of Engineering and Chief Technology Officer	Brandsoft, Inc.
David Clark	Senior Research Scientist	Massachusetts Institute of Technology
John Gardner	Partner	Nokia Venture Partners
James Hendler	Professor, Department of Computer Science	University of Maryland
Deepali Khushraj	Manager	Nokia Research Center
Paul Kogut	Principal Investigator	Lockheed Martin Management and Data Systems
Ora Lassila	Research Fellow	Nokia Research Center
Alexander Linden	Vice President, Emerging Trends and Technologies	Gartner Research
Brian McBride	Researcher	HP Laboratories
Eric Miller	W3C Semantic Web Activity Lead	World Wide Web Consortium
Irene Polikoff	Executive Partner	TopQuadrant, Inc.
Steve Ross-Talbot	Chief Scientist	Enigmatec Corp.

[14] Ora Lassila and Deepali Khushraj, interview by author, Burlington, MA, January 29, 2004.

International representation was ensured by the fact that a considerable amount of activity in this realm is based in Europe. There is evidence that a migration to Silicon Valley is beginning, as in the case of Network Inference, a United Kingdom (UK) company. The reason why Europe is the home to activity in this area appears related to the technology's conceptual relationship to Artificial Intelligence (AI). While a great deal of foundational research was performed in the United States (US), commercial, or industrial interest in the field waned when AI failed to live up to the hype that set unrealistic expectations. So while US investment in AI was reduced due to a lack of corporate interest, the European research community benefited from continued or possibly expanded governmental funding[15] As a result, with some exceptions, the knowledge most relevant to the Semantic Web continued to grow in Europe while it tended to lag elsewhere.

The body of this thesis is organized into a framework that identifies and discusses the various hurdles confronting the adoption of Semantic Web technology. Some of these obstacles are common to any sort of change, such as the normal human reluctance to venture into the new and unknown. Other hurdles are much more specific, such as comprehending the blurred distinction between real and virtual objects. Regardless, this discussion centers on identifying a hurdle to the broad adoption of this technology and in some cases, why and how these hurdles might be overcome.

[15] Alexander Linden, interview by author, Cambridge, MA, February 17, 2004.

3 The Semantic Web's Hurdles

This section will describe a series of hurdles related to the adoption of the Semantic Web. By no means is this section intended to be exhaustive – as with any new technology there are bound to be encouraging discoveries and those that are quite the opposite. Regardless, the hurdles presented may be among the more significant, meaning that if they can be cleared, any remaining hurdles will be more easily surmounted.

The hurdles chosen for discussion resulted from their tendency to arise during field interviews or to appear with some regularity in both technical and managerial literature (managerial literature in this case refers primarily to published work on the study of innovation). As a result, there seems to be good reason to believe in the relevance of these hurdles and the importance of successfully addressing them.

3.1 Achieving Critical Mass

The Internet and all that resides upon it comprises a network. An obvious statement, but networked environments have certain characteristics that can increase their value and utility far beyond the obvious. In this case, the operative concept is the "network effect." This term refers to the idea that if one member of a network enhances their online presence, while it's clearly beneficial to them, it's of limited value to any other member of the network. However, if many members of the network do so, it's very likely that the enhancement enacted individually will compound in a networked environment to produce an effect far greater than the cumulative sum of the individual actions, hence, the "network effect."

Implicit to achieving this enhanced network state is reaching the tipping point, or "critical mass," in adoption of the enhancement. Once reached, the incentive for all members to follow suit increases, since the benefits of doing so are likely to accrue to each new member as well as any existing members (although likely to a lesser degree).

In discussions related to the future of the Semantic Web, there is a common misperception that the entire Web must be changed. This comment is usually closely coupled with the confident assertion that "It'll never work." This raises some intriguing issues about human perception and memory, but that work is best left to behaviorists. Instead, these assertions are notable in their assumption that adoption of the Semantic Web must be universal for it to work at all.

This argument is punctured fairly easily. First, the Semantic Web community envisions the adoption process occurring along two paths with rough simultaneity. One path will follow closely defined, highly targeted projects that have the effect of embedding the technology within a corporate environment. The second path will be broader and Web oriented, as the realization sets in that data exchange, interoperation, and resource sharing has the potential of being easier and more efficient in inter-corporate settings.[16]

[16] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.

In fact, there are several ways to approach this issue. For instance, it's conceivable that a Web author could embed appropriate tags (XML, in the case of the Semantic Web) on a site that would enhance the likelihood of discovery during a Semantic Web-based search. The reasons for taking the extra step of defining XML tags for this purpose have everything to do with "discovery," or making it easier for information to be found. The music industry provides a reasonable example.

Suppose a fan of mid-19th century American folk music were to use the major search engines to find new titles to add to his collection. Presently, this kind of search yields hundreds of thousand of entries, which is far too many to sift through. Now, suppose a Semantic Web search engine were available in the near future, which is not an unreasonable assumption. Knowing this, a musician specializing in this genre might hire a thoughtful Web author to embed tags that define his music with greater precision than currently possible.

As a result, the music seeker is much more likely to find the music maker and in turn, a mutually satisfactory exchange or relationship may ensue. Of course, once the musician in question experiences success under these conditions, other musicians will likely follow suit. The entire music industry, realizing that a growing percentage of consumers' music dollars are being spent on an arcane genre as a result of superior search results, might follow suit. If this grassroots pattern were to appear, it would closely resemble early Web pages, published by individuals who wanted to express themselves in a very public and discoverable way.[17]

The existence and growth of (presumably successful) small corporate projects and grassroots adoption will eventually lead to a critical mass Semantic Web sites, content, and tools. Reaching this point is a critical factor in the growth and use of this technology[18] because it is at this point that different Semantic Web properties can combine and collaborate in new and unique combinations. Interestingly, since the announcement of the RDF and OWL standards, several corporations have come forward to highlight their work in this field.[19]

In many respects, the tagging of objects, both virtual and real, will be optional and based on the desire to be discovered. Years ago, pioneering individuals published their personal pages in a barren Web landscape hoping they would be discovered. With the advent of search engines, it was possible to register a site for the purpose of increasing the likelihood of being discovered. So just as the Web developed, when the desire to be discovered becomes great enough, appropriate tagging will likely follow.

The lifespan of Web sites and Web objects must also be taken into consideration. During the explosive growth phase of the Web (e-commerce in particular), it was common to update and even completely re-design Web sites on a seasonal, quarterly or certainly

[17] David Clark, interview by author, Cambridge, MA, January 21, 2004.
[18] Frank Careccia, interview by author, Cambridge, MA, January 28, 2004.
[19] Eric Miller, interview by author, Cambridge, MA, March 31, 2004.

semi-annual basis. Now that the medium has matured, the urgency to overhaul a given Web site has diminished, but not disappeared. So with these thoughts in mind, one could assume periodic re-designs of thriving Web sites driven by seasonal factors, specific events, or fundamental technical changes.

Obviously, Web sites are not synchronized to the same eighteen month schedule, and the maturity and ease of use of the tools sets for XML tagging is an issue that will be dealt with later in this thesis. However, making reasonable assumptions in the evolution of Web sites and the tool sets for their development, it seems quite reasonable to envision a gradual, but steady, evolution to the Semantic Web.

Analysis

Christensen offers a helpful insight when he states "…products that do not appear to be useful to our customers today (that is, disruptive technologies) may squarely address their needs tomorrow."[20] As the Semantic Web community waits for critical mass to be achieved, it might best serve itself by continuing to develop its technology and seek instances for testing. This may well be an instance where waiting for market validation is an empty exercise since, as Christensen goes on to argue, "…we cannot expect our customers to lead us toward innovations that they do not now need."[21]

In his work, Utterback details the adoption and evolution of the typewriter by describing Samuel Clemens' experience of purchasing and using one during the summer of 1874[22] . This early machine was difficult to use, and had unacceptable limitations by today's standards (e.g., it produced text in upper case, only, and typists couldn't actually see their work in real time since the paper was deep in the machine). Because items such as handbills (advertisements) were produced only in upper case the machine was also considered culturally offensive. Despite these hurdles, the industry persevered, and as Utterback points out, thirty years later the typewriter industry had become well established and populated by a number of competitors.

Dismissing Semantic technology at this embryonic stage may prove to be a serious error. Even assuming that technological adoption cycles are faster now than they were in the late nineteenth century, it's still early in the Semantic Web game. If today's researchers are addressing computer needs that are only now arising (or will shortly arise), then patience on their part and experimentation on the part of customers might be a prudent course of action.

---

[20] Clayton M. Christensen, *The Innovator's Dilemma* (New York, NY: HarperBusiness Essentials, 2003). p. 258.
[21] Ibid.
[22] James M. Utterback, *Mastering the Dynamics of Innovation* (Boston, MA: Harvard Business School Press, 1996), p.1.

3.2 A Standard Taxonomy That Addresses Real and Virtual Objects

A critical aspect of the Semantic Web is its extension beyond the virtual realm of Web pages, into components and callable procedures (Web services). The Semantic Web can include real objects such as textile looms, trucks, cell phones, physicians, managers, buildings and anything else that can be described by some set of attributes. Distinguishing between the two a bit more sharply, URL's (Uniform Resource Locators)[23] refer to Web pages or the components that are called to create a Web page such as images, text, or perhaps a media file. URI's (Uniform Resource Identifiers) can refer to Web pages as well as anything else, including those mentioned above.

Currently, any number of Web pages are rendered by calling content, applications data and graphic elements from different locations. To accomplish this, each of these components is assigned its own unique URL. Once called, the components are assembled in a coherent and pre-determined fashion defined by the publishing site's author. On the Web, this works fine, but in expanding the boundaries through the Semantic Web, it quickly becomes clear that users can't actually navigate to a home's refrigerator or a factory's textile loom and take into account the unit's attributes such as cubic capacity, cooling ability, input requirements, or output capacity.

The example below highlights some simple distinctions between Web URI's and Semantic Web URI's. Note that each URI can be broken into two parts (at least). In the Web version, the document URI refers to a Web page, which in this case is the "About" page for "bigbiz.com." However, the document anchor is more precise, and refers to a specific paragraph.

Figure 3. URIs Compared

Web URIs and Semantic Web URIs Compared
World Wide Web URI Document Anchor Document URI http://bigbiz.com/about.html#para1
Semantic Web URI URI of a "Real Thing" (Loom 6) Document URI http://bigbiz.com/loom.rdf#loom6

[23] Tim Berners-Lee, Weaving the Web (New York, NY: HarperCollins, 2000), p.62, "We argued, but at the IETF [Internet Engineering Task Force] the universal resource identifier became URL, the *uniform resource locator*. In years ahead the IETF community would use the URL acronym, allowing the use of the term *URI* for what was either a URL or something more persistent. I use the general term *URI* to emphasize the importance of universality, and of the persistence of information."

The first part of the Semantic Web URI, or document URI, refers to a document that contains information about looms and the entire URI refers to a specific loom, in this case, Loom 6. In its combined form, the Semantic Web URI refers to a specific loom and can also take into account and describe its attributes.

Using the loom as a working example (although it can't actually be reached through the Web today with any real meaning), the Semantic Web allows for the definition of attributes in a re-usable fashion. For instance, a given loom might be able to weave 100 feet per minute, produce fabric of a certain thickness, do so at a consistently high quality, require nylon of a certain gauge range for input, and be available according to a schedule updated in real time. So, in an ideal Semantic world, a manager (*not* a developer) could create a solution to a client requirement for nylon fabric delivered on a specific schedule that includes a loom with defined attributes. In fact, anyone with access (determined by security levels) could include this particular loom if needed, as long as it was available.

The above reference to a business manager devising a Semantic Web solution is deliberate. Prior to the common use of office productivity tools, programming was a skill confined to those people willing to learn new vocabularies and syntaxes contained in languages such as COBOL, FORTRAN, Basic, and others. In fact, some assert that using a computer language is fairly easy, just as long as you're willing to think in terms that suit that language. With the advent of productivity tools like spreadsheets, database tools and others, end users (business managers and other non-programmers) can create software and applications without knowing a computer language. In other words, they can engage in programming, only no one calls it that.[24] Instead, these activities might be referred to as creating bookmarks, or a table of contents in a Word document, or creating formulas or macros in an Excel spreadsheet.

There are software solutions that can solve the loom production task now, so why bother with the Semantic Web? The issue becomes the proprietary nature of the current solutions and the fact that any data entered into such a system becomes a captive of that system and can only be re-used by that system. Semantic technology seeks to not only obviate these bastions of propriety, but due to its universality, make all objects, virtual and real, easily re-usable when needed.[25]

An example of the Semantic Web's universality becomes clear through a continued examination of URIs and URLs. In the aforementioned loom example, let's assume that a variety of supporting resources are required. This would include operating personnel, personal communication devices (for these people), and forklifts, trucks and other transportation equipment for the finished goods. As described earlier, URLs refer to Web pages or their components. URIs can be assigned to anything (including Web pages), which means they can be assigned to people, personal communication devices, forklifts, trucks, etc. Going even further, URIs can be assigned in the place of serial numbers, part numbers, or any other kind of unique identifier presently in use.

---

[24] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.
[25] Zavisa Bjelogrlic, Dirk-Willem van Gulik, and Alberto Reggiori, "Making Business Sense of the Semantic Web", Presentation, 2003 International Semantic Web Conference.

When solving a production problem, all aspects of the supply chain can be addressed and accounted for, as in the loom example. But the flexibility and re-usability of the Semantic Web goes far beyond this example. For instance, communication devices can be re-assigned to other personnel while forklifts and trucks can be assigned to move completely different items. Moreover, the original personnel can be included in payroll, benefits, and training systems. In each case, a "best of breed" application could be used for the given situation, as long as suitable XML tags were used and appropriate RDF and OWL structures were in place.

Additionally, as discussed at the conclusion of Section 1, data in a Semantic environment can reside within their original environments, eliminating the requirement for export to a standardized data warehouse. Since the data are retained by the host application, they can be updated constantly, which is not always the case in a warehoused environment. Up-to-date data in a Semantic environment allows changes to the data to become immediately available to any other application with access privileges. Obviously, in a world where speed has proven to be a competitive advantage, the business advantages of this timeliness seem quite apparent.

Analysis

Managers of contemporary organizations can choose to adhere to known solutions or at least their evolutionary progeny. Doing so lets these managers remain within a comfort zone of sorts, even in a business environment that appears in which the rate of change seems to be constantly accelerating. In their role as observers (and *not* innovators), Thomke[26] points out many business managers' willingness to "…wait for changes to be induced and then carefully study what has been presented to them."

The preceding discussion describes unknown terrain and emerging tools and standards such as URIs, XML, RDF, and OWL. But then it's hardly unknown for corporations to establish experimental facilities. In fact, experimentation -- even in a resource-constrained world -- would seem increasingly necessary for companies seeking to create or sustain competitive advantage. Whether the Semantic Web is ultimately successful or not, there may well be at least aspects of it, such as XML, that prove to be of great value in current and future computing environments.

In this light, taking initiative now and gaining institutional knowledge may become a competitive asset. Of course, if Semantic technology fails to live up to its hype, as laser technology and gallium arsenide microchips have, then it may be relegated to niche applications where the needs are more sharply defined. The trial and error process can be expensive and time consuming, but in the "fluid phase…in which a great deal of change is happening at once and in which outcomes are highly uncertain…"[27], experimentation

[26] Stefan H. Thomke, *Experimentation Matters* (Boston, MA: Harvard Business School Press, 2003), p.91.
[27] James M. Utterback, *Mastering the Dynamics of Innovation* (Boston, MA: Harvard Business School Press, 1996), p.92.

may be necessary to the success of the technology and the organizations considering its use.

3.3 The Ever-Present Hurdle: Resistance to Change

It's a fact that most people don't like change, even software developers. But change is inevitable, particularly in mediums such as the Web and Semantic Web that are still in their infancy, when any change is likely to occur at the greatest pace and divergence. Even highly qualified people may object to the difficulties associated with a new technology, and in the case of the Semantic Web, make comments like "RDF…Who wants to write or read that stuff?"[28]

Conversely of course, there are those that disagree and in the case of writing in RDF, state "I did it in the back of the room. RDF is very elegant. Don't confuse it with the XML syntax."[29] Others point out that in the brief history of the Web, there was a time when developers didn't like reading, writing, or learning HTML either and that today's HTML tools often produce code that is not intuitively readable by humans, asserting that Semantic Web coding may well follow a similar pattern.[30] Regardless, there is learning process that needs to occur, and the tools associated with authoring Semantic Web code need to achieve a higher degree of maturity to render moot this issue.

There are legitimate business concerns when contemplating the implementation of Semantic Web technology such as the:

- Time required to develop production grade applications.
- Knowledge base within a given organization or even on a global basis.
- Reliability of the code.
- Devising of an effective scale/scope of implementation.

These concerns are compounded by perception, given recent history where technologies related to Customer Relationship Management (CRM), Enterprise Resource Planning (ERP), and Supply Chain Management (SCM) have been over-hyped, generating skepticism if not outright cynicism among corporate decision makers.

This questioning attitude may actually benefit Semantic technology. For instance, some would argue that once a corporation's data has been stored in SAP or an Oracle database, then that particular vendor "controls" how this information is structured and applied. While this control is of a passive nature, it's very likely that the corporate owner of the data will be required to convert or somehow re-format the data to make it broadly usable,

---

[28] Steven Pemberton, World Wide Web Consortium "Minutes: W3C Technical Plenary, March 5, 2003", World Wide Web Consortium Web site http://www.w3.org/2003/03/Plenary-Minutes.html#TownHall, accessed March 15, 2004.
[29] Brian McBride, World Wide Web Consortium "Minutes: W3C Technical Plenary, March 5, 2003", World Wide Web Consortium Web site http://www.w3.org/2003/03/Plenary-Minutes.html#TownHall, accessed March 15, 2004
[30] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.

or at least usable outside the vendor's storage architecture. In stark terms, "AOL is to the Web as proprietary database and application schemes are to the Semantic Web".[31]

With this assertion as a working proposition, then the morass of applications, databases, and the related maintenance requirements could be greatly reduced through the use of Semantic technology. In this case, through appropriate representation via XML and structuring with RDF, ontologies created under the OWL standard could be created once, and maintained on a comparatively occasional basis. This scenario also ensures live access to the latest, most up-to-date data, suggesting access to superior information in the shortest possible timeframe. The evolving business environment increasingly seems to favor organizations that successfully use the latest information in the most rapid manner, meaning the Semantic Web has the potential to contribute meaningfully to corporate performance.

Another, and emotional, line of reasoning has been suggested which alludes to the adoption of the Web. Throughout the 1990s any number of corporations assumed that the Web was a passing fad that would have no material impact on the conduct of their business. The reality has proved to be precisely the reverse – the Web has permanently altered the manner in which corporations do business. In fact, it has increased the range and nature of business opportunities available to contemporary organizations.

This line of reasoning goes further, suggesting that "You can let the Semantic Web just happen to you the way the Web happened to you, or you can get out in front and start working on it now."[32] While this rationale lacks a rigorous foundation, it's difficult to deny the intuitive substance of this message. Confirming the inattention of senior corporate managers, some assert that "the Semantic Web is just noise to most CIO's."[33] Others offer indirect confirmation by suggesting that the "…people in major corporations that are interested in the Semantic Web are the visionaries and technology zealots. These people are engaged in experimentation."[34]

The success of these experiments will be critical to the adoption process, and it's too early to be conclusive. There are some cases where certain aspects of Semantic technology have proven successful, such as Lockheed Martin's use of DARPA Agent Markup Language (DAML), which forms the basis for OWL, the W3C's standard for ontology development.[35] The article describing this use concludes by mentioning the challenges that were encountered. It states the belief that the Semantic Web will eventually become mainstream technology, at least in the aerospace industry. Paul Kogut, co-author of the article, indicates that the US government has substantial interoperability challenges[36] and that it, too, may find great utility in Semantic technologies.

[31] Tim Berners-Lee, "The Semantic Web", Presentation, March 5, 2004, MIT Sloan School of Management, Cambridge, MA.
[32] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.
[33] John Gardner, interview by author, Cambridge, MA, February 17, 2004.
[34] Irene Polikoff, interview by author, Cambridge, MA, March 2, 2004.
[35] Paul Kogut, Jeff Heflin, "Semantic Web Technologies for Aerospace", *Aerospace Conference, 2003 Proceedings, 2003 IEEE.*
[36] Paul Kogut, interview by author, Cambridge, MA February 25, 2004.

In light of the preceding discussion, it would appear that if the Semantic Web were to live up to at least a portion of its potential, issues related to data integration and application interoperability would be reduced or at least eased. Examples of production grade uses of Semantic technology are in very short supply although experimentation is clearly occurring. Not only that, but one of the most critical elements, XML, is seeing broader adoption and integration with the standard in office productivity tools, Microsoft Office. The existing problems of integration and interoperability in large-scale computing environments are becoming too great to leave any possible solution unexplored, which augers well for the application of the Semantic Web.

Analysis

A very thoughtful analysis is provided by Thomke[37], which argues in favor of "experimenting early and often." Adopting this practice at a pace the overall organization can tolerate can bring together (presumably around a technical innovation) a variety of functional groups to examine their processes, how they fit with overall corporate goals, and result in a shared learning experience that's likely to be highly beneficial. Given Semantic technology's potential for re-use across a wide portion of a given enterprise, ensuring the involvement of a diverse group of participants in the experimentation process could result in applications not presently contemplated by researchers and developers.

In addition, Thomke suggests that the experimental process begin early and operate with short cycle times. This pattern allows findings to be quickly incorporated into the ongoing experimental process, which can be very effective at substantially reducing overall implementation costs. With this framework, a corporation could experiment with Semantic technology using concise, well defined projects, timed to allow any findings to be incorporated into the next round. Engaging in this kind of "prototyping" could limit the risk inherent in adopting any new technology and might serve to sharpen the focus of its application and enhance benefits of using the new technology at all.

3.4 The Semantic Web's Role in Adapting IT Infrastructure

Japanese auto manufacturers proved with great effect the importance of reducing time-to-market. This lesson has been learned well by industries around the world: In fact, responding "just-in-time" to changing market conditions has become a significant concern to all levels of management. At issue here is that humans can change their behavior relatively quickly, while machines and the software written for them often cannot, particularly in the case of large enterprise applications.

Semantic technologies can be viewed as a vehicle to ease the perpetual tension between evolving business strategies and IT infrastructure. The motivating belief is that the traditional way of applying resources to solve a problem results in solutions that are

[37] Stefan H. Thomke, *Experimentation Matters* (Boston, MA: Harvard Business School Press, 2003), p.161.

outdated the moment they're complete.[38] Vendor interviews produced a particularly insightful example of Semantic technology's flexibility[39] and is paraphrased here:

> Investment banks may have equity and fixed-income trading units, each with a front, middle and back office operation. In many cases, they all have "reference data," which serves as ontologies. The problem is that they have these reference data in all six areas (two business units, three operational areas each). So in a live environment, an equities trader might balance off a trade with a reverse deal in fixed income. These trades would be recorded in two different silos in the bank, using two different sets of reference data, on behalf of the same counterpart (whom we will call Joe Hedgefund). The entry for Hedgefund in the fixed income unit is not the same for Hedgefund in the equities unit, yet semantically it is. So financial institutions a good deal of money to create and maintain "look-up tables." These tables can be massive and very expensive to build and maintain. A Semantic Web implementation that incorporates XML, RDF, and OWL is a substantially less costly approach than look-up tables. This approach is much more flexible, readily reusable, and it's also computationally very efficient. In a trading environment speed is critical and assuring sub-second performance can be decisive.

If the above is accurate, then the immediate advantages of reduced maintenance, increased speed, and flexibility are clear. Weill and Vitale describe "atomic business models" and how they comprise larger business initiatives[40]. Their discussion occurs in the context of e-business, and they describe an atomic, or highly granular approach to IT. This analogy suits the Semantic Web well and is just as applicable to traditional computing environments.

Analysis

Christensen[41] echoes Thomke's exhortation to experiment with rapid and inexpensive tests to determine viability. Additionally, Hendler[42] believes that adoption of the Semantic Web will follow two paths, the first being incremental (and experimental), and the second, big-bang implementations, once critical mass has been reached and the network effect validates the value of the technology. Both sources make a case for identifying highly specific and well-defined situations where Semantic technology can be introduced.

Aside from the financial services example above, other test formats might include:

---

[38] John Gardner, interview by author, Cambridge, MA, February 17, 2004.
[39] Steve Ross-Talbot, interview by author, Cambridge, MA, February 12, 2004.
[40] Peter Weill and Michael R. Vitale, Place to Space (Boston: Harvard Business School Press, 2001), p. 55.
[41] Clayton M. Christensen, *The Innovator's Dilemma* (New York, NY: HarperBusiness Essentials, 2003). p. 259.
[42] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.

- The process of combining small business units.
- Deployment in support of a limited manufactured product line (comprising all the various components).
- A corporate organizational chart (which might enable greater ease representing the changes made to business unit composition and associated reporting relationships).

In any case, while the pace of change can be difficult to overcome, coping with it successfully might be a competitive advantage. Semantic Web technology may have value in these circumstances.

3.5 Fact or Fiction?

At some point, corporate decision makers must decide if the Semantic Web presents enough value for them to dedicate resources to develop their capabilities in this realm. Separating fact from fiction can often be difficult, and successfully navigating the process of doing so can make the difference between a skyrocketing career and looking for another job.

Much of the world is recovering from the bruising of the past few years following the burst of the Internet bubble. As if attempting to make up for lost time or repenting for "irrational exuberance"[43], a new-found skepticism has crept into the corporate mindset, which has proven challenging to many emerging technologies. Understandably, IT executives will want to understand the basis for Semantic technology and its history in general.

The concepts that underlie the Semantic Web have not simply burst on to the scene. In fact, while the following point is often downplayed, the field of Artificial Intelligence (AI) has established a well defined body of work which in this instance, serves to pave the way for progress. One possible reason for the separation of Semantic Web technology from AI may be for presentation and promotional purposes; specifically, to avoid being associated with a "failed" technology.

Unfortunately, the failure of AI is a notion that exists only in the minds of people outside the field. Those within the AI community point out the various aspects of the technology have been commercially implemented with great success, including rules-based systems (expert systems), advanced analytics (including neural networks), voice recognition, credit card fraud detection and others.[44] Some in the field of AI suggest that when permutations of the technology become considered part of mainstream computing, their AI heritage is conveniently forgotten.

---

[43] Alan Greenspan, "The Challenge of Central Banking in a Democratic Society", December 5, 1996, Remarks At the Annual Dinner and Francis Boyer Lecture of The American Enterprise Institute for Public Policy Research, Washington, D.C., Board of Governors of the Federal Reserve System Web site <http://www.federalreserve.gov/boarddocs/speeches/1996/19961205.htm>, accessed March 29, 2004.
[44] Ora Lassila, interview by author, Burlington, MA, January 29, 2004.

Another far more concrete distinction has to do with how the Semantic Web will allow knowledge (data) bases to connect to each other. AI systems typically confine themselves to a specific class, which is described by an ontology. This model makes it very difficult to connect these systems together. The Semantic Web will allow disparate knowledge bases to connect at the "fringes" in a Web-like fashion and will not force a top-down, bottom-up, or element by element translation or equivalence. Furthermore, if each knowledge base connects to several other knowledge bases, and each one of the other knowledge bases connects to still more (and presumably, different) knowledge bases, the net effect is multiplied by a far greater factor than simply those knowledge bases that are directly connected to the original (or *your*) knowledge base.

Figure 4. Connecting at the Fringes

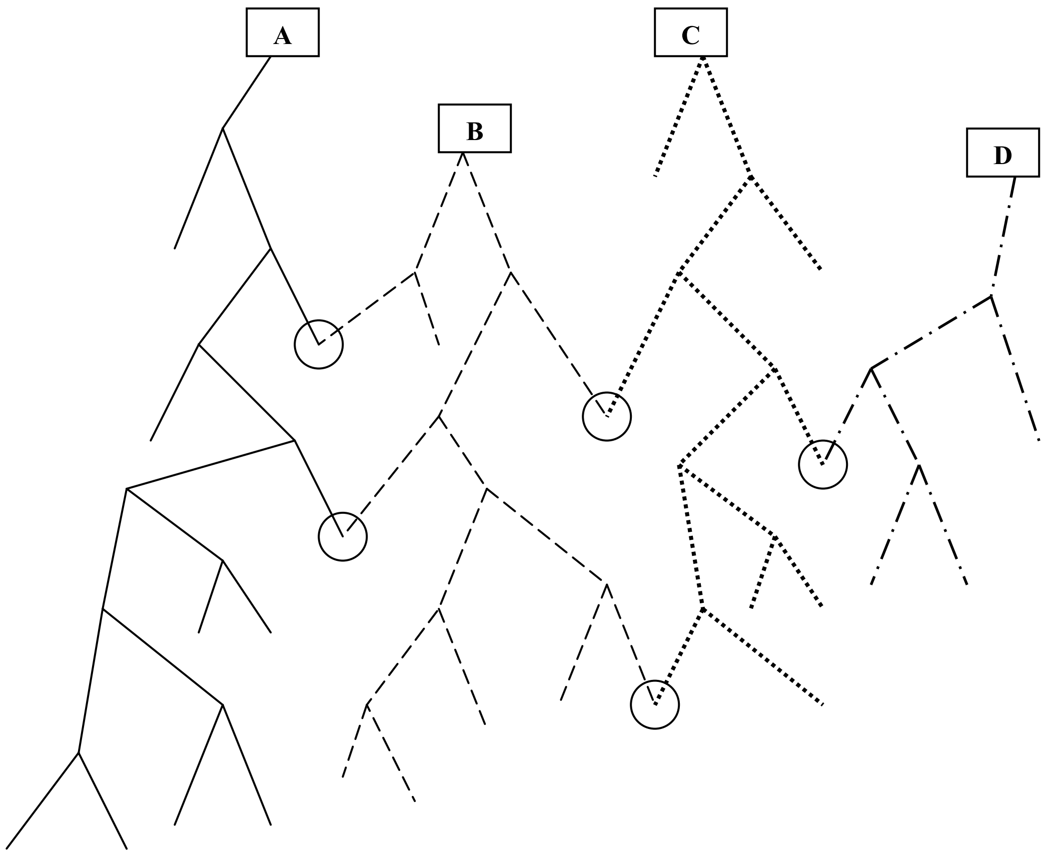


Figure 4 identifies four different knowledge bases. Each circle indicates a semantic connection, or rule. The level of the connection (in a hierarchical sense) does not matter – the connections are made just where they're needed. Also each of the four ontologies above can connect across if they so choose (A can connect directly to C or D, B directly to D) in addition to any connection indicated in the diagram.

In a very thoughtful article, Schwartz builds a convincing case for "…AI's development from a niche research area to the science that is now forming the Semantic Web core."[45] By citing numerous researchers and their work, he makes it clear that the concepts and challenges of the Semantic Web have been considered and at least identified, if not solved. In fact, Schwartz goes so far as to suggest that:

> " the universal connectivity, addressing, and transport mechanisms offered by combining TCP/IP, URIs, and RDF/XML wrappers might just provide the right glue to combine with Open Knowledge Base Connectivity, the Knowledge Query and Manipulation Language, or other such attempts to combine reasoning with ontology-based knowledge representation."[46]

Note that the final series of terms in the excerpt above all refer to aspects of AI, and are in the process of manifesting themselves within the Semantic Web.

Considering the preceding point, it would seem reasonable to take comfort in the research and grounding already in place for Semantic technology. There's still a fundamental problem however, which is directly related to the knowledge base among software developers. This issue has been hinted at earlier by references to learning "new vocabularies and syntaxes" and resistance to the idea of wanting "to write or read that stuff."

Without doubt, there will be a learning curve on the part of developers. The length of this learning curve will depend upon the scale and scope of the implementation in question, although a variety of projects are already underway. During this period, a number of people are willing to point out the relative scarcity of knowledge in this discipline. Assertions along these lines include the belief that there are not yet any gurus, and only a few genuine experts[47] and that the maturity of Semantic tools is lacking, requiring developers to work at a much lower level to which they might be accustomed.[48]

At the current rate of tool evolution and knowledge diffusion, it would seem that creating applications with Semantic technologies will become easier. Already, colleges and universities such as the University of Maryland and the Massachusetts Institute of Technology offer instruction either in the area of the Semantic Web or at least its components. When corporate expenses become too high and the investment climate warms up, it would seem reasonable to expect increased exploration of Semantic technology usage.

---

[45] David G. Schwartz, "From Open IS Semantics to the Semantic Web: The Road Ahead", *IEEE Intelligent Systems* (May/June 2003)
[46] Ibid.
[47] Alex Linden, interview by author, Cambridge, MA, February 16, 2004.
[48] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.

Analysis

The essential point is that any of the underlying technical aspects of the Semantic Web are probably known and reasonably studied in some form. As discussed earlier, Semantic technology draws upon established research in AI. This background serves to identify potential challenges in the logic and operation of Semantic technology. If the existing body of work doesn't identify solutions, then at least it may provide tools or frameworks that lend themselves to creating solutions, along with people who are familiar with the necessary solutions, tools and frameworks.

In fact, the issue of familiarity may lend itself quite nicely to the adoption of Semantic technology. With his "familiarity matrix" Roberts[49] graphically demonstrates how the migration process to newer technologies is facilitated by having at least some knowledge of related technologies. Viewed this way, it becomes clear that the human knowledge base exists, at least to a degree, and that it's possible that this knowledge base simply hasn't been oriented toward implementing Semantic Web solutions.

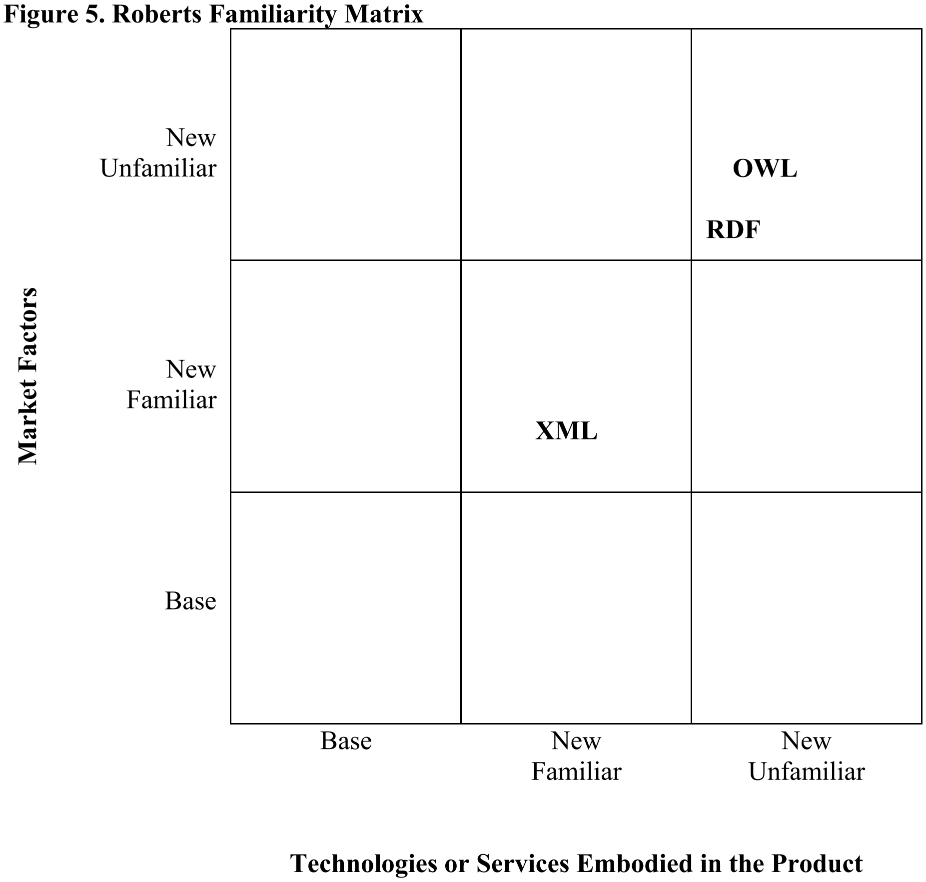
The chart below provides a framework through which the Semantic Web can be viewed in terms of familiarity, or the existing knowledge base in which the disciplines for Semantic technology are rooted. First, some definitions paraphrased from Roberts' work noted above:

- **Market Factors:** The environment, situation, or setting to which Semantic technology is applied.
- **Technologies or Services Embodied in the Product:** Refers to XML, RDF, and OWL.
- **Newness of a Technology:** The degree to which that technology has not been commonly employed.
- **Newness of a Market:** The degree to which a technology has not been applied to a given environment, situation, or setting.
- **Familiarity with a Technology:** The degree to which knowledge of the technology exists but is not necessarily in a Semantic context.
- **Familiarity with a Market:** The degree to a potential application of the technology is understood, but not necessarily through direct experience.
- **Base:** Common knowledge and common application of the technology.

[49] Edward B. Roberts and Charles A. Berry, "Entering New Businesses: Selecting Strategies for Success." *Sloan Management Review*, 26 (Spring 1985): 3-17.

Now, an interpretation of XML, RDF and OWL through this matrix:

Figure 5. Roberts Familiarity Matrix



None of the technologies qualify for placement in the Base square simply because they aren't adequately pervasive for such placement. XML's degree of adoption and its inclusion in such common applications as the Microsoft Office Suite and its use in Web sites such as CNET, Yahoo!, and eBay easily qualify for placement in the central New Familiar square. Finally RDF and OWL are placed in the most "Unfamiliar" square, although not in precisely the same position.

Regarding this final point, these two technologies are placed here due to the lack of widespread knowledge related to their use, despite knowledge of progenitor technologies among a limited number of developers. And while the exact placement of these technologies might vary by author, the Roberts Familiarity Matrix provides a useful tool to assess the maturity of these technologies, which in two out of three cases seems to be in the early stages of maturity.

4 Conclusion

If the Semantic Web's promoters are right, this technology promises to simplify the management of increasingly complex and rapidly evolving computing environments. Semantic technology may also create opportunities for application interaction, synthesis, and extension that are not yet contemplated and can only exist online. Widespread evidence of these possibilities is in very short supply, but the possibility that the Web could one day become a live, planet-wide database available to machines and not just humans is quite intriguing.

Valid concerns remain, related to issues such as application or environment complexity and comprehensibility, scalability, and reliability. It's too early in the going to allay these concerns but perceptions are reality, regardless of the technology's evolutionary phase. Hendler[50] points out that the Semantic Web may suffer from "Younger Sibling Syndrome," i.e., unrealistically high expectations are immediately placed on a new and not fully developed technology. Hendler provides a reminder as well, stating that later versions of Mosaic were better not just because Mosaic itself evolved, but because the environment began changing to accommodate Mosaic. Today, Web sites are commonly optimized for popular browsers, a behavior that is easy to imagine being applied to the Semantic Web's evolution.

It's also difficult to underestimate the elusive quantity called "critical mass." The US government has been making substantial investments in Semantic technology for reasons related to national security and the integration of its vast bureaucracy. At the same time, major corporations such as Hewlett-Packard, Nokia, Lockheed Martin and others have been active investors in this technology. One benefit from making such investments is the possible discovery of greater efficiencies associated with operating their respective businesses. Another, and very intriguing benefit comes from the potential that may be realized when trading partners of all sizes suddenly realize the Semantic Web investments they've made independently within their corporate boundaries lead to new opportunities, efficiencies, and ways of working together that are yet to be discovered.

Research in the field of innovation and technology adoption strongly suggest the need to experiment, apply existing and related bodies of knowledge, and persevere through iterations of trial and error for the Semantic Web to gain a fair hearing in the marketplace. By some accounts, Thomas Edison tried more than 1,500 different materials during the course of developing a reliable light bulb filament. The sheer number of researchers working on Semantic technology should be able to determine within the foreseeable future whether the Semantic Web will be a niche technology like Gallium Arsenide microchips or a pervasive and revolutionary technology like its older brother, the World Wide Web.

[50] Jim Hendler, interview by author, Cambridge, MA, March 3, 2004.

5 Bibliography

Adams, Katherine. "The Semantic Web: Differentiating Between Taxonomies and Ontologies." *Online* 26 n4 (July/August 2002): 20-23.

Albanese, Andrew. “The Semantic Web We Weave.” *Library Journal* 127 n17 (Fall 2002): 8.

Allen, Joshua. “Better Living Through Software.” *Netcrucible.com*, January 04, 2003, <http://www.netcrucible.com/blog/2003/01/04.html#a265>

Amit Sheth, Clemens Bertram, David Avant, Brian Hammond, Krysztof Kochut and Yashodhan Warke. “Managing Semantic Content For The Web.” *IEEE Internet Computing* (July/August 2002): 80-87.

Anderson, Tim. “Prepare Now For Tomorrow's Web.” *Internet Week*, September 29, 2003.

Becker, David. “XML Gets Ebay And Microsoft Into Bed.” *Silicon.com*, March 2, 2004. <http://www.silicon.com/management/itpro/0,39024675,39118802,00.htm>, accessed March 11, 2004.

Berners-Lee, Tim. Interview by author. Cambridge, MA, December 18, 2004.

Berners-Lee, Tim. “The Semantic Web,” presentation, March 5, 2004. MIT Sloan School of Management, Cambridge, MA.

Berners-Lee, Tim. *Weaving the Web.* New York, NY: HarperCollins, 2000.

Bradbury, Danny. “The Search For 'Smart Data' Pays Off: Business Will Benefit: Customers Will Be Able To Find Data Much Quicker.” *Financial Post: FP Edge, Enterprise* (September 29, 2003).

Bradbury, Danny. “Unweaving The Tangled Web Of Dumb Data.” *ComputerWeekly.* United Kingdom: Reed Business Information, September 16, 2003.

Butler, Mark H. “Is The Semantic Web Hype?” HP Labs Bristol Working Paper, November 14, 2003.

Canter, Sheryl. “Getting The News Out.” *PC Magazine* 22 n17 (October 2003): 68.

Careccia, Frank. Interview by author. Cambridge, MA, January 28, 2004.

Cherry, Steven M. “Weaving A Web Of Ideas.” *IEEE Spectrum* 39 n9 (September 2002): 65-69.

Christensen, Clayton M. *The Innovator’s Dilemma*. New York, NY: HarperBusiness, 2000.

Clark, David. Interview by author. Cambridge, MA, January 21, 2004.

CNET Networks, Inc. “CNET RSS Home.” CNET Networks Web site. <http://www.cnet.com/4520-6022-5115113.html>, accessed March 11, 2004.

David Trastour, Chris Preist and Derek Coleman. “Using Semantic Web Technology To Enhance Current Business-To-Business Integration Approaches.” *Proceedings of the Seventh IEEE International Enterprise Distributed Object Computing Conference (EDOC’03)*, (September 2003): 222.

David Trastour, Claudio Bartolini and Chris Preist. “Semantic Web Support For The Business-To-Business E-Commerce Pre-Contractual Lifecycle.” *Computer Networks Amsterdam* 42 n5 (August 5, 2003): 661.

Dieter Fensel and Mark A. Musen. “The Semantic Web – A Brain for Humankind.” *IEEE Intelligent Systems* 16 n2 (March/April 2001): 24-25.

Dieter Fensel, Wolfgang Wahlster, Henry Lieberman and James Hendler. *Spinning the Semantic Web.* Cambridge, MA: MIT Press, 2002.

Douglas Tudhope and Daniel Cunliffe. “Semantically Indexed Hypermedia.” *ACM Computing Surveys* 31 n4 (December 1999).

Eric Miller and Ralph Swick. “An Overview Of W3C Semantic Web Activity.” *Bulletin of the American Society for Information Science and Technology Silver* 29 n4 (Spring 2003): 8-11.

“Evangelizing the Semantic Web.” *Dr. Dobb's Journal* 26 n9 (September 2001): 18.

Evans, Nick. “Warning: Colliding IT Trends Ahead.” *InternetWeek* 875 (August 27, 2001): 15.

Ewalt, David M. “Challenges: Speed Bumps Ahead For Semantic Web.” *InformationWeek* 910 (October 14, 2002): 44.

Ewalt, David M. “The Next Web.” *InformationWeek* 910 (October 14, 2002): 34-44.

Ewalt, David M. “World Wider Web.” *Asia Computer Weekly*, November 18, 2002.

Festa, Paul. “For W3C, It’s A Question Of Semantics.” *CNET* (November 18, 2002).

Fitzgerald, Michael. “Searching Through Babel.” *CIO,* March 15, 2003.

Frauenfelder, Mark. “A Smarter Web.” *Technology Review* 104 n9 (November 2001): 52.

“Future Of Search - When The Web Starts Thinking For Itself.” *Information World Review* 37 (December 1, 2002).

Gardner, John. Interview by author. Cambridge, MA, February 18, 2004.

Gates, Bill. (Keynote Speech, Comdex, November 26, 2003).

“Global: Web Inventor's Semantic Web Is no Sure Success.” *ebusiness forum.com*, April 2, 2003, http://www.ebusinessforum.com/index.asp?doc_id=6431&layout=rich_story

Goble, Carole. “The Semantic Web: An Evolution For A Revolution.” *Computer Networks* 42 n5 (August 5, 2003): 551.

“Google Invests in the Semantic Web,” April 24, 2003, <http://www.high-search-engine-ranking.com/GoogleInvestsintheSemanticWeb.htm>

Greenberg, Jane. “The Semantic Web: More Than A Vision.” *Bulletin of the American Society for Information Science and Technology Silver* 29 n4 (Spring 2003): 6-7.

Greenspan, Alan. “The Challenge of Central Banking in a Democratic Society,” Federal Reserve System Web site, December 5, 1996 <http://www.federalreserve.gov/boarddocs/speeches/1996/19961205.htm>

Jeff Heflin and James Hendler. “A Portrait Of The Semantic Web In Action.” *IEEE Intelligent Systems* 16 n2 (March/April 2001): 54-59.

Hendler, James. “Science And The Semantic Web.” *Science* 299 n5606 (January 24, 2003): 520-521.

Hendler, James. Interview by author. Cambridge, MA, March 3, 2004.

“HTML in Semantic threat.” *The Nation,* May 26, 2003.

“Intelligent e-Biz Platform Technology Developed.” *The Electronic Times*, February 3, 2003.

James Hendler and Bijan Parsia. “XML And The Semantic Web: It's Time, Stop Squabbling -- They're Not Incompatible.” *XML Journal* 3 n10 (October 2002): 30.

Joanna J Bryson, David L Martin, Sheila D McIlraith and Lynn Andrea Stein. “Toward Behavioral Intelligence In The Semantic Web.” *Computer* 35 n11 (November 2002): 48-54.

Jorge Cardoso and Amit Sheth. “Semantic E-Workflow Composition*.” Journal of Intelligent Information Systems* 21 n3 (November 2003): 191.

Kabbaj, Mohamed Youssef. “Strategic And Policy Prospects For Semantic Web Services Adoption In Us Online Travel Industry” (M.A. diss., MIT, 2003).

Kenyon, Henry S. “Computer Language Seeks Deeper Meaning.” *Signal* 57 n10 (June 2003): 59-61.

Kettler, Brian. “Semantic Enabling & Exploitation (SEE) for Assured, Improvisational Workflows.” *DAML PI Meeting* (October 18, 2003).

King, Julia. “Taming Data Complexity.” *Computerworld* 37 n26 (June 30, 2003):31.

Kogut, Paul. Interview by author. Cambridge, MA, February 25, 2004.

Koprowski, Gene J. "The Future Of Human Knowledge: The Semantic Web. *TechNewsWorld* (July 28, 2003).

Krill, Paul. "Exec Touts Semantic Web." *InfoWorld Daily News*, June 7, 2002.

Krill, Paul. "IBM Execs Ponder Technology Plans; Semantic Web, Eclipse Improvements Eyed." *InfoWorld Daily News*, October 24, 2003.

Kulvatunyou, B. and Ivezic, N. "Semantic Web For Manufacturing Web Services." *Proceedings of the 5th Biannual World Automation Congress 2002*.

Kutler, Jeffrey. "Changing Rules." *Institutional Investor* 36 n12 (December 2002): 61.

Ora Lassila. Interview by author. Burlington, MA, January 29, 2004.

Linden, Alexander. Interview by author. Cambridge, MA, February 17, 2004.

Louise Crow and Nigel Shadbolt. "Extracting Focused Knowledge From The Semantic Web." *International Journal of Human-Computer Studies* 54 n1 (2001): 155-184.

McBride, Brian. Interview by author. Cambridge, MA, February 18, 2004.

McBride, Brian. "Minutes: W3C Technical Plenary, March 5, 2003." World Wide Web Consortium Web site. <http://www.w3.org/2003/03/Plenary-Minutes.html#TownHall>, accessed March 15, 2004

McGuinness, Deborah L. "Ontologies Come of Age." S*pinning the Semantic Web.* Cambridge, The MIT Press: 2003.

Menezes, Joaquim. "Web Not Up To Speed." *Computer Dealer News* 15 n23 (June 11, 1999): 35.

Michael C. Daconta, Leo J. Obrst and Kevin T. Smith. *The Semantic Web.* Indianapolis, IN: Wiley Publishing, Inc., 2003.

Microsoft, Inc. "Office 2003 Editions: Compare the to Previous Versions." Microsoft Inc. Web site. <http://www.microsoft.com/office/editions/prodinfo/compare.mspx>, accessed March 11, 2004.

Montalbano, Elizabeth. "Seeing The Light -- Touting ROI And Streamlined Business Processes, Solution Providers Look To Cash In On Web Services Opportunities." *Computer Reseller News* (December 16, 2002).

Miller, Eric. Interview by author. Cambridge, MA, December 15, 2003.

Neal, David. "Office 2003 To Integrate With Ebay." *Internet Week*, March 8, 2004.

Nigale, Bhushan Y. "The Web Is Not Enough, Semantic Web." *Smart News, Inc.*, May 1, 2003.

Ogbuli, Uche. “Sun Deploys Semantic Web Technologies For Knowledge Management.” *O’Reilly Network*, March 13, 2003, <http://www.oreillynet.com/pub/wlg/2903>

Ora Lassila and Deepali Khushraj. Interview by author. Burlington, MA, January 29, 2004.

Parsia, Bijan. “Semantic Web Services.” *Bulletin of the American Society for Information Science and Technology Silver* 29 n4 (Spring 2003): 12-15.

Paul Kogut and Jeff Heflin. “Semantic Web Technologies for Aerospace.” *IEEE Aerospace Conference* (March 2003).

Pemberton, Steven. “Minutes: W3C Technical Plenary, March 5, 2003.” World Wide Web Consortium Web site. <http://www.w3.org/2003/03/Plenary-Minutes.html#TownHall>, accessed March 15, 2004.

Perkins, Anthony B. “Betting On The Next Web.” *Red Herring*, May 1, 2002.

Peter Weill and Marianne Broadbent. *Leveraging the New Infrastructure*. Boston, MA: Harvard Business School Press, 1998.

Peter Weill and Michael R. Vitale. *Place To Space*. Boston, MA: Harvard Business School Press, 2001.

Polikoff, Irene. Interview by author. Cambridge, MA, March 3, 2004.

Port, Otis. “The Next Web.” *BusinessWeek online,* March 3, 2002 <http://www.businessweek.com/magazine/content/02_09/b3772108.htm>

Robin Peek and Paula J. Hane. “New Specifications For A New Web.” *Information Today* 19 n5 (May 2002): 42.

Roberts, Edward B. and Charles A. Berry, 1985, “Entering New Businesses: Selecting Strategies for Success.” *Sloan Management Review*, 26 (Spring 1985): 3-17.

Rosencrance, Linda. “E-Commerce On The Fly.” *Computerworld* 37 (May 26, 2003): 34

Ross-Talbot, Steve. Interview by author. Cambridge, MA, February 13, 2004.

Schwartz, David G. “From Open IS Semantics To The Semantic Web: The Road Ahead.” *IEEE Expert* 18 n3 (May 2003): 52-58.

Srini Narayanan and Sheila McIlraith. “Analysis And Simulation Of Web Services.” *Computer Networks Amsterdam* 42 n5 (August 5, 2003): 675.

Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein, Jeen Broekstra, Michael Erdmann and Ian Horrocks. “The Semantic Web – The Roles of XML and RDF.” *IEEE Internet Computing* 15 n3 (September/October 2000): 63-74.

Thibodeau, Patrick. “The Web's Next Leap.” *Computerworld* 37 (April 21, 2003): 34.

Thomke, Stefan H. *Experimentation Matters*. Boston, MA: Harvard Business School Press, 2003.

Thompson, Bill. “Finding What You Want on the Web.” BBC World Service, January 2, 2004.

Thuraisingham, Bhavani. “Building Secure Survivable Semantic Webs.” *IEEE International Conference on Tools with Artificial Intelligence (ICTAI ’02)*, (November 2002): 395.

Tim Berners-Lee, James Hendler and Ora Lassila. “The Semantic Web.” *Scientific American*, May 2001. <http://www.sciam.com/article.cfm?articleID=0004814410D2-1C70-84A9809EC588EF21>

Utterback, James M. *Mastering the Dynamics of Innovation*. Boston, MA: Harvard Business School Press, 1996.

Waddington, Paul. “Semantic Web Halted By Angst.” *Information World Review* 181 (June 2002): 36.

Warren, Paul. “The Next Steps For The Web: Putting Meaning Into The Web.” *Computing & Control Engineering* 14 n2 (April/May2003): 27.

Weaver, Belinda. "’Open the Pod Bay Doors, Please, HAL’: Here Comes The Semantic Web.” *Online Currents (ABIX abstracts),* September 11, 2003.

“What the Semantic Web Won't Do.” *Business Week* (April 8, 2002).

Wingfield, Nick. “Ebay To Test Technology Linking It To Other Sites -- Aiming To Be A New Platform For E-Commerce, Auctioneer Looks To Microsoft For Model.” *The Wall Street Journal*, November 20, 2000.

Zavisa Bjelogrlic, Dirk-Willem van Gulik and Alberto Reggiori. “Making Business Sense of the Semantic Web,” presentation, 2003. International Semantic Web Conference, Rome, Italy.