Six Challenges for the Semantic Web

V. Richard Benjamins[1], Jesús Contreras[1], Oscar Corcho[2] and Asunción Gómez-Pérez[2]

[1]Intelligent Software Components, S.A.
www.isoco.com (Spain)
{richard,jcontreras}@isoco.com
[2]Technical University Madrid. Facultad de Informática.
Campus de Montegancedo, s/n. 28660 Boadilla del Monte. Madrid. Spain.
Tel: +34913367439 Fax: +34913524819
{ocorcho,asun}@fi.upm.es

Abstract. The Semantic Web has attracted a diverse, but significant, community of researchers, institutes and companies, all sharing the belief that one day the Semantic Web will have as big an impact on life as currently the WWW/Internet has. We share that vision, based on the ever-increasing need to reduce information overload, and to increase task delegation to software agents. However, there is still a long way to go before the Semantic Web dream comes true. In this paper, we identify some of the major challenges the community faces in the coming years, and we outline solution directions. The major challenges concern: (i) the availability of content, (ii) ontology availability, development and evolution, (iii) scalability, (iv) multilinguality, (v) visualization to reduce information overload, and (vi) stability of Semantic Web languages. We will also say some words on the economic impact of the Semantic Web.

1 Introduction

One of the biggest problems we nowadays face in the information society is information overload, a problem which is boosted by the huge size of the WWW. The Web has given us access to millions of resources, irrespective of their physical location and language. In order to deal with this sheer amount of information, new business models on the web have seen the light, such as commercial search engines (of which Google is by far the champion). With the expected continuous growth of the WWW, we expect search engines will have a hard time maintaining the quality of retrieval results. Moreover, they only access static content, and ignore the dynamic part of the web (pages generated from databases). It is our vision that the technology of current generation of search engines has its limits. To be able to deal with the continuous growth of the WWW (in size, languages and formats), we need to exploit other information. This is where the Semantic Web comes in.

The current Web is based on HTML, which specifies how to layout a web page for human readers. HTML as such cannot be exploited by information retrieval techniques to improve results, which has thus to rely on the words that form the content of the page; hence it is restricted to keywords. In the Semantic Web, pages not only store content as a set of unrelated words in a document, but also code their meaning and structure. Figure 1 illustrates that Semantic Web languages are based on XML and go up the Semantic Web Language Pyramid to RDF [Lassila et al], RDFS [Brickley et al], OIL [Horrocks et al], DAML+OIL [DAML+OIL], etc.

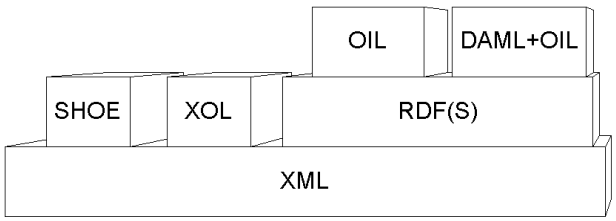

Figure 1: Web languages pyramid.

Those languages are much richer than HTML and allow –to a greater or lesser extent- to represent the meaning and structure of content (interrelationships between concepts). This contributes to make web content understandable and procesable by software agents, opening the way to a whole new generation of technologies, innovative knowledge-based services and business models, where we will see a gradual change from information retrieval support to task delegation and achievement.

The problem of Information Overload can be partly tackled by adding intelligence to the web. Software agents could manifest various levels of intelligent behaviour from simply reactive to adaptive and learning behaviour, where agents actually learn what users like and dislike [Etzioni, Perkovitz 1997]. This would shield users from irrelevant information, who would only be 'bothered' for information of real value. Instead of the current ubiquitous *pull*

paradigm that requires users to actively look for information and to execute programs, we should work towards delegating those tasks to autonomous software agents. Finally, the tasks users will want to execute will become increasingly complex. Software agents should 'learn' to function in 'social' environments and where necessary collaborate, compete or negotiate with other agents. The quality and usability of the Semantic Web infrastructure will depend on advances on all three dimensions

Recently, the US and EU governments have recognized the importance of the Semantic Web and have established dedicated programs (DAML and IST Action Line III.4.1) to fund research aimed at developing the core technology for the Semantic Web.

In spite of the big advantages that the Semantic Web promises, its success or failure will -as with the WWW- be determined to a large extent by easy access to, and availability of high-quality and diverse content. There are still several problems to solve before making this happen, including, but not limited to:

- **The availability of content**. Currently, there is little Semantic Web content available. Existing web content should be upgraded to Semantic Web content including static HTML pages, existing XML content, and dynamic content, multimedia and web services.
- **Ontology availability, development and evolution.** Ontologies will become a key piece, as they allow expliciting the semantics of Semantic Web content. A big effort must be made in the creation of common widely used ontologies for the Semantic Web, on the provision of adequate infrastructure for ontology development, change management and mapping, and, in this distributed web environment, on the adequate control of the evolution of ontologies and the annotations referring to them.
- **Scalability**. A significant effort must be made to organize Semantic Web content, store it and provide the necessary mechanisms to find it. All these tasks must be performed and coordinated in a scalable manner, as these solutions should be prepared for the huge growth of the Semantic Web
- **Multilinguality**. This problem already exists in the current Web, and should also be tackled in the Semantic Web. Any Semantic Web approach should provide facilities to access information in several languages, allowing the creation and access to SW content independently of the native language of content providers and users.
- **Visualization**. Intuitive visualization of Semantic Web content will become more and more important to solve the increasing amount of information overload, as users will demand the easy recognition of relevant content for their purposes. New techniques must be explored that differ from the usual hypertext structure visualization of the current web.
- **Stability of Semantic Web languages**. Finally, standardization efforts must be performed urgently in this emerging field, in order to allow the creation of the necessary technology that supports the Semantic Web

In this paper, we will explain those challenges[1] and provide some solution directions. But before that, we will sketch some relevant notions that have emerged in the short history of the

[1] The six challenges discussed in this paper will be dealt with in the new European project Esperonto Services (www.esperonto.net).

Semantic Web. At the end of the paper we will also say something about possible commercial applications exploiting the Semantic Web.

2 The Semantic Web in Retrospect

The Semantic Web promises to make web content machine understandable, allowing agents and applications to access a variety of heterogeneous resources, processing and integrating the content, and producing added value output for users. Current technology, mostly developed in research centres, already enables this type of scenarios, such as in the ISI/USC Ariadne (theatre, restaurants suggestions) and Heracles (travelling) projects[2] [Knoblock et al Ariadne], [Knoblock et al] We call these pre-Semantic Web applications because, from an application point of view, they allow many of the functionalities the Semantic Web promises: software agents accessing, manipulating, and integrating content from distributed heterogeneous web resources. However, under the hood of these applications there is an ad-hoc solution using wrapper technology. A wrapper is a program that accesses an existing website and extracts the needed information (e.g. the address of a location or the price of an article). Wrappers are screen scrapers in the sense that they parse the HTML source of a page, using heuristics to localize and extract the relevant information. Not surprisingly, wrappers have high construction and maintenance costs since much testing is needed to guarantee robust extraction, and each time the website changes (e.g. its design), the wrapper has to be updated accordingly.

The semantic web aims to provide an extra –machine understandable- layer, which will considerably simplify programming and maintenance effort for knowledge-based web services.

Two dimensions are especially important in this context:

- Programming effort to set up new added-value services
- Web preparation effort (in particular the annotation effort)

The programming effort is determined by the amount of programming required to build a new service. Figure 2 illustrates[3] this by means of some typical information retrieval (IR) applications like basic IR, Google[4] and CiteSeer[5], and contrasts this with the effort needed if the Semantic Web already existed (last bar).

IR is a standard technique currently not requiring a lot of effort. Advanced search engines such as Google exploit the web structure for ranking results and therefore, require more effort. CiteSeer, a scientific research index for Computer Science, additionally analyses the content of documents to find information such as authors and bibliographical references, and requires

[2] http://www.isi.edu/info-agents/index.html
[3] This example has an illustrative character and is not meant to reflect actual programming effort.
[4] http://www.google.com
[5] http://citeseer.nj.nec.com

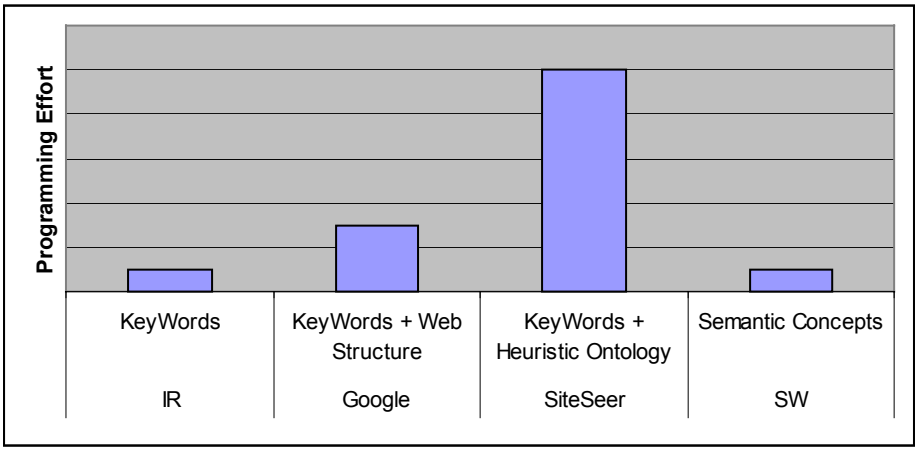

Figure 2: Programming effort for some typical web applications (search engines). Number on the Y axis is only to indicate relative effort [source: authors]

a larger programming effort. The point we want to make is that the more knowledge is exploited in a web service (an application), the larger is the programming effort. In the vision of the Semantic Web, the programming effort decreases significantly because there the knowledge would be available *in* the web rather than that it has to be 'scraped' from it.

We should contrast this favourable characteristic of the Semantic Web with its current state (January, 2002) and observe that there is hardly any Semantic Web content available. So, exploiting the knowledge in the Semantic Web currently requires a significant web preparation effort. In Figure 3 below we illustrate this with the same typical applications.

A simple keyword-based IR system, Google and CiteSeer do not need any web preparation effort to work. Any current Semantic Web application, on the other hand, requires a significant annotation effort to put the knowledge into the web to turn it into the Semantic Web [Luke et al] [Decker et al] [Benjamins et al (KA)2].

One of the main challenges the Semantic Web community faces for the construction of innovative and knowledge-based web services is to reduce the programming effort while keeping the web preparation task as small as possible. In the next section, we will elaborate this challenge, as well as the other challenges.

3 The Challenges

3.1 Challenge 1: The Availability of Content

Semantic Web content is web content annotated according to particular ontologies, which define the meaning of the words or concepts appearing in the content. Before the notion of the Semantic Web existed, we were involved in an experiment ([Decker et al], [Benjamins et al

(KA)2]) aimed at knowledge-based retrieval of information in the Knowledge Engineering community. A simple extension to HTML was used to annotate web pages with ontological information. This experiment taught us that the annotation effort is one of the major factors for potential failure. Creating Semantic Web content is therefore a serious challenge for the Semantic Web. Since the infrastructure of the Semantic Web is still being built (RDFS, OIL, DAML+OIL, etc.), currently, there is little Semantic Web content available. Apart from the infrastructure, researchers are currently building tools to support semantic annotation of web content. Such tools are important and critical to the success of the Semantic Web. However, they have two limiting characteristics: 1. Most of them annotate only static pages, and 2. Many of them focus on creating new content. This leads to the following not optimal situation:

- Dynamically generated content is not considered. 'Dynamic' content is generated from databases, and according to a study of [Lawrence & Giles 1999] referring to it as the 'Deep Web'- in March 2000, [Michael K. Bergman] its size is estimated to be 400 to 550 times larger than the commonly defined World Wide Web, which includes more than one billion static web pages.
- Existing content is running the risk to be excluded from the Semantic Web, even though XML content is gaining ground on content share. Although we agree that in the end the whole web will be semantic (10 years horizon?), it would be a pity if current high-quality web content were not available on the Semantic Web.

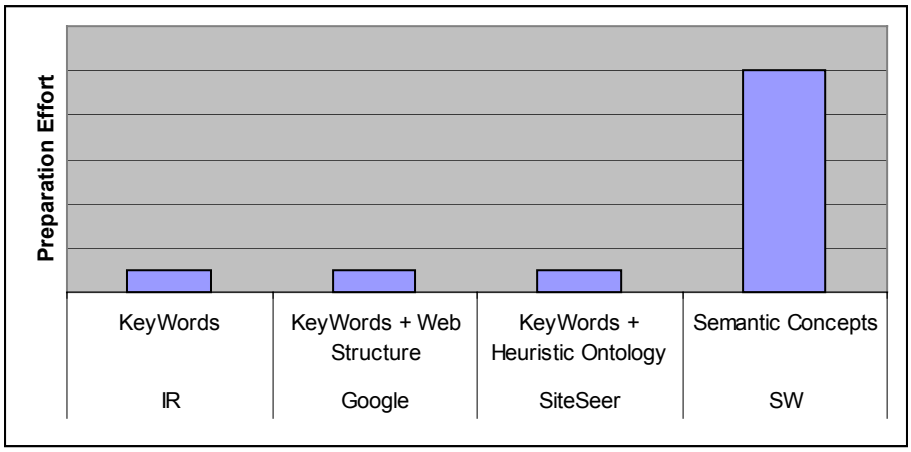

Figure 3: Web content preparation effort in typical web application [source: authors]

We see the solution direction for dealing with this challenge as follows. We need to create a set of annotation services (middleware) concerning static and dynamic web documents, which may include multimedia, and web services. The output of these annotation services will be generated according to the language pyramid of the Semantic Web, so that different agents understanding different languages of the Semantic Web might always understand and handle as much content as might be represented in the different languages. This approach to the annotation of resources is rather innovative, in the sense that current efforts have proposed the

annotation of web resources in just one of these languages, and this layered approach has just been taken into account in the development of languages.

Basically, automatic annotation of static content uses existing wrapping technology. Annotation of dynamic content is much more complex. There are several alternatives, which can be explored, for example:

- Extract the dynamic content from its source, annotate it (as if it were static) and store it. The problem here is the almost infinite amount of static pages that can be generated from a dynamic (database powered) site, and the almost continuous updates, creations and removals of generated static pages when data changes in the databases.
- Leave the content in the database and annotate the query that retrieves the concerned content. This option is less space-consuming, and consistency in the annotations is guaranteed with respect to the underlying sources of information, since the content is dynamically generated/annotated when retrieved from its sources.

Annotation of *web services* also needs to be considered, so that complex services will be able to use existing services over the Semantic Web. This work needs to build on emerging standards such as SOAP [SOAP] (simple object access protocol), WSDL [WSDL] (Web Service description language), UDDI [UDDI] (Universal Description, Discovery and Integration), and WSFL [WSFL] (Web Service Flow Language). For more complex services, DAML-S[6] [DAML-S] and UPML [UMPL] are interesting candidates, however more research is needed to defined the exact characteristics of such language. DAML-S is a DAML-based Web service ontology, which supplies Web service providers with a core set of mark-up language constructs for describing the properties and capabilities of their Web services in unambiguous, computer-interpretable form. UPML is the language framework to mark-up high-level services for task achievement as used in the IBROW project[7].

Related to the annotation of multimedia content, there are at least two alternatives: image and video processing (on the base of so-called "low level features" detected in the shots they try to extract content from) and information extraction using NLP techniques of so-called "collateral" documents, including textual explanations of images, transcripts of automatic speech recognition that report on the events present in video sequences, etc. Independently from what alternative is selected, the results would be annotated following the same ideas as in the static and dynamic cases presented above.

3.2 Challenge 2: Ontology Availability, Development and Evolution

Ontologies are key to the Semantic Web because they are the carriers of the meaning contained in the Semantic Web, that is, they provide the vocabulary and semantics of the annotations. There are three main issues to be solved regarding this challenge, the first two issues are related to traditional ontology development problems that haven't been solved

[6] http://www.daml.org/services/

[7] http://www.swi.psy.uva.nl/projects/ibrow/home.html

completely until now, and the last one is much more related to the new context of the Semantic Web:

The first is the construction of *kernel* ontologies to be used by all the domains. Initiatives exist for the construction of some of these kernel ontologies in different domains: the IEEE Standard Upper Ontology Group[8] aims to create a common unified top level ontology); initiatives in the e-commerce domain also exist, such as UNSPSC[9], eclass[10], RosettaNet[11], NAICS[12], etc.

The second is to provide methodological and technological support for most of the activities of the *ontology development process* [Fernandez-1997], including:

- Knowledge acquisition, conceptual modelling and ontology coding in Semantic Web languages (RDFS, OIL, DAML+OIL), and new languages that may arise in the coming years [Maedche, Staab – 2001].
- Ontology alignment and mapping, ontology integration, ontology translation tools, and ontology reengineering tools if existing ontologies are going to be reused [Fensel et al, 2001], [Noy, Musen 2000].
- Consistency checking tools for the ontologies to be reused [Gomez-Perez 1996].

The third is the evolution of ontologies and their relation to already annotated data. Configuration management tools are necessary to keep control of the versions of each ontology as well as the interdependencies between them and annotations. All of those points are far from trivial, but need to be solved before a realistic and scalable Semantic Web becomes possible.

3.3 Challenge 3: Scalability of Semantic Web Content

Once we have the Semantic Web content, we need to worry about how to manage it in a scalable manner, that is, how to organize it, where to store it and how to find the right content. We have detected two main issues underlying this challenge:

The first one is related to the storage and organization of Semantic Web pages. The 'basic' Semantic Web consists of ontology-based annotated pages whose linking structure reflects the structure of the WWW, that is, pages connected to others by means of hyperlinks. This hyperlinked configuration does not fully exploit the underlying semantics of Semantic Web pages. We foresee the use of **semantic indexes** to group Semantic Web content based on particular topics. This is a necessary step to make applications able to aggregate content in order to provide added value services. Semantic indexes will be generated dynamically using ontological information and annotated documents.

---

[8] http://suo.ieee.org/
[9] http://www.unspsc.org
[10] http://www.eclass.de
[11] http://www.rosettanet.org/
[12] http://www.census.gov/epcd/www/naics.html

The second issue is related to the easy finding of information in the Semantic Web, in other words, to the coordination among semantic indexes. A mechanism must be provided for the easy finding of SW content taking into account the semantics of web resources. In this context, a peer to peer (P2P) architecture could be explored, similar to the current configuration of routers in the WWW, in what we could call a "semantic TCP/IP protocol", the new European Semantic Web (SWAP, Semantic Web and Peer-to-peer) project is dedicated this topics (Semantic Web and Peer-to-peer). Indexes could be considered as active agents that know what topics they can handle (i.e. find content for). Topics that do not occur in the index are **semantically routed** to neighbour indexes. The use of agents should be explored for negotiation techniques in order to obtain the semantic routing of topics. Also the notion of 'semantic distance' used in WordNet should be investigated [Budanitsky et al 2001]. In this way, there would be no need for a central register of semantic content making the whole architecture more scalable in order to accommodate continuous addition of content to the Semantic Web. Again, such organisation would be in the spirit of the original WWW.

3.4 Challenge 4: Multilinguality

Studies on language distribution over the WWW content show that even if English is the predominating language for documents, there exist a important resources written in other languages: English 68.4%, Japanese 5.9%, German 5.8%, Chinese 3.9%, French 3.0%, Spanish 2.4%, Russian 1.9%, Italian 1.6%, Portuguese 1.4%, Korean 1.3% ,Other 4.6% [Source: Vilaweb.com, as quoted by eMarketer]. The diversity of languages is much more acute for European WWW resources. Multilinguality plays an increasing role at the following levels: at the level of ontologies, of annotations and of user interface.

At the **ontology** level, ontology builders may want to use their native language for the development of the ontologies in which annotations will be based. Since not all users will be ontology builders, this level has lowest priority. Existing multilingual and linguistic resources, such as WordNet [wordnet], EuroWordnet [eurowordnet], etc., might be explored to support multilinguality at this level.

At the **annotation** level, annotation of content can be performed in various languages. Since more users (especially content providers) will rather annotate content than develop ontologies, proper support is needed that allows providers to annotate content in their native language. To make Semantic Web content generation as low an effort as possible, we cannot require from a French to annotate content in the German language nor vice versa.

Finally, at the **user interface** level, millions of people would like to access relevant content in their native language irrespective of the source language in which annotations are presented. Although currently, most content is in English, we expect that more content will appear in other languages. Any Semantic Web approach should include facilities to access information in several languages. Internationalisation and localization techniques should be explored to personalize information access based on the native language of the user.

3.5 Challenge 5: Visualization

With the increasing amount of information overload, intuitive visualization of content will become more and more important, as users will be increasingly demanding easy recognition of the relevance of content for their purposes. Additionally, the use of semantic indexes and routers for the storage, organization and finding of information, will require a major step forward in visualization, compared to traditional site maps that represent link structures, as presented in section 3.3. Techniques should allow for three-dimensional and new visualisation techniques to visualise SW content in any of the current SW languages (RDFS, OIL and DAML+OIL). Technologies to be considered include, but are not limited to X3D (of the Web3D Consortium), Java3D (API for writing programs to display and interact with three-dimensional graphics, Shockwave3D (new technology introduced by Macromedia). By developing adequate 3D real-time graphic representation technology and exploiting semantic relationships, an innovative three-dimensional interface could be generated automatically. This way, more information can be represented in less space, and users can interact with the site in a realistic and comfortable way [Van Harmelen et al 2001].

3.6 Challenge 6: Semantic Web Languages Standardization

The Semantic Web is an emerging field and the WWW consortium will produce recommendations on the languages and technology that will be used in this area. In order to advance the state of the art in the Semantic Web, it is important that such standards appear within this year. As already presented above, a layered approach to ontology languages' creation and annotations has been adopted by the community.

Tool support is also essential to make a significant step forward in the construction of the Semantic Web, and tools are partly dependent on the Semantic Web language they are supposed to work with. Important sources to arrive at a Semantic Web language standard and the necessary tool support include the W3C Semantic Web Activity (Web Ontology group: www.w3.org/2001/sw/WebOnt), the DAML (www.daml.org) initiative and results of key projects funded by the European Commission, such as OntoKnowledge (www.ontoknowledge.org), IBROW (www.swi.psy.uva.nl/projects/ IBROW/ibrow-home.html). Esperonto (www.esperonto.net), etc.

4 Towards a Semantic Web Architecture

It is too early to come up with an overall architecture that accommodates all challenges mentioned in this paper, and others to be put forward. However, a first high-level draft is illustrated in Figure 5, which will be the basis of the European IST Semantic Web project Esperonto.

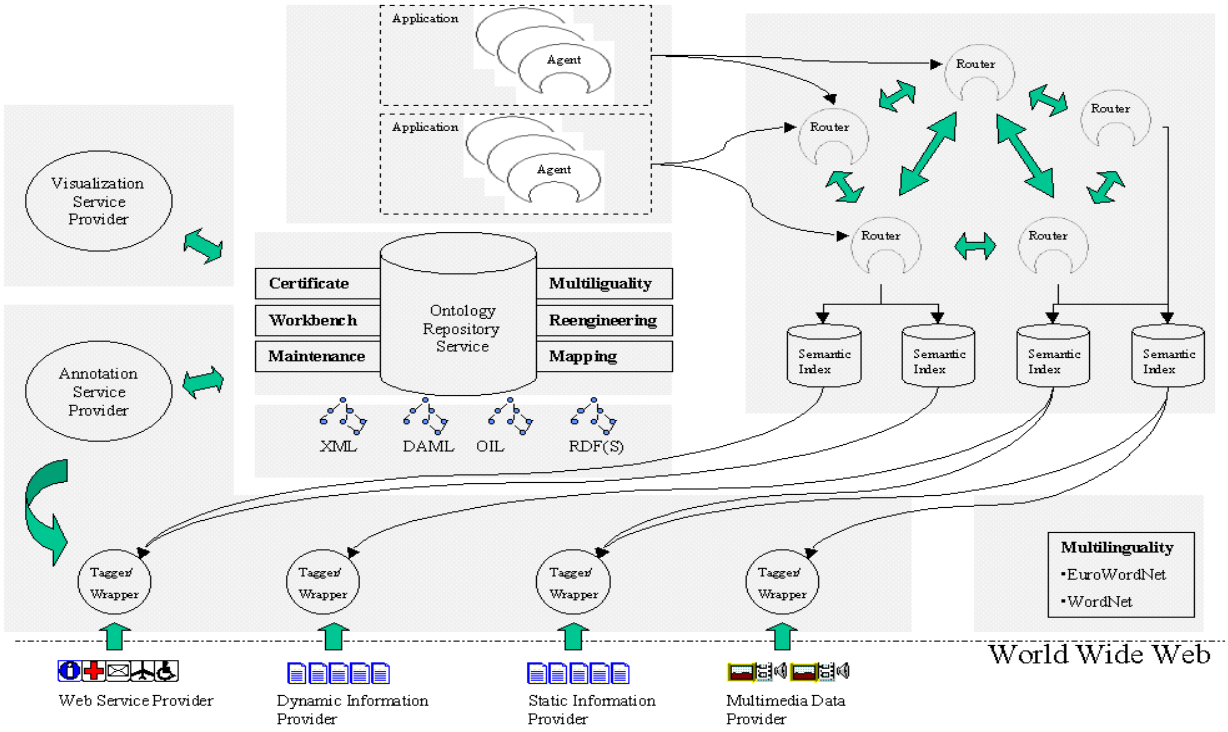

Figure 4: High-level architecture of the Semantic Web construction technology

- The bottom side of the figure shows the current web with its different types of resources: static pages, dynamic pages, web services and multimedia, all of them can be in different languages. Dedicated editors and wrappers aggregate the information in those sources into semantic indexes. A routing mechanism establishes and maintains the relation and communication between the various indexes. Software applications (agents) access Semantic Web content through the routing mechanism. Since semantics are represented using ontologies, an ontology lifecycle model forms a central component of the architecture.

5 Economic Impact of the Semantic Web

When thinking of economic impact of the Semantic Web, one has to think of applications, services and business models made possible with Semantic Web technology, that are not possible without such technology. We foresee several types of products and/or services that can be thought of, some of them within reach on a short term, whereas others on a mid or long term.

On a short term, we can think of a service where content providers can upgrade their existing content to the Semantic Web. Such service could operate in ASP mode and could be exploited on a commercial basis (e.g. based on the amount of data to upgrade, or on a

periodical fee, etc.). Why would content providers be willing to pay for such as service? For example, because their content then becomes available for other business (one step higher in the information food chain) that use their information (Semantic Web content) to sell added value services. Another reason to pay for such kind of service would be to have a place in the global 'Semantic Yellow Pages'. We think however, that currently the feasibility for such commercial scenarios is limited, not in the least place because such service can only work properly when sector/area specific ontologies are in place. The development of such ontologies, however, is still an ongoing activity where a lot of improvement is expected in the next 5 years.

On a midterm, we envision that ultimately semantic annotation service provision will disappear as a stand-alone service, and will be incorporated instead in all kinds of end user applications. Compare that for instance to utilities as 'save as html' in MSWord, 'Export to HTML', 'Import HTML'. In the future we envision similar utilities for the Semantic Web: 'save as SWML[13] ' in MSWord, 'Export to SWML', and 'Import SWML'. Notice however that such utilities for the Semantic Web are much more complex than for HTML since the Semantic Web deals with semantics rather than with syntax. A critical factor (condition), again, is advancing ontology technology.

The first commercial applications of Semantic Web technology are foreseen in intranets, rather than on the Internet. The major reason for this is that not all companies will like to see their essential business information included in the Semantic Web, if this implies that it will be accessible by any software agent understanding Semantic Web languages. Public organizations offering services for citizens are probably the first entities ready to exploit Semantic Web technology on the Internet. This vision is sustained by the On-To-Knowledge project where Semantic Web technology is used to help improve companies' intranets, or the project Esperonto, where in two of our three test cases public organizations are involved, willing to disseminate their information. There are still a lot of issues (technical, business, legal) to resolve (and discover!) before Semantic Web technology can and will be commercially exploited at the same level as the current web.

Having said this about full-fledged Semantic Web applications, for so-called 'pre-Semantic Web' applications these caveats do not apply. (Remember that pre-Semantic Web means that from an application point of view, many of the functionalities the Semantic Web promises are enabled: software agents accessing, manipulating and integrating content from distributed heterogeneous web resources.) For example, in iSOCO we are currently commercialising a 'Financial Aggregator[14]' through which one can access and manipulate multiple online accounts for all major Spanish banks and telecom operators through one web page. The software includes wrappers (one for each site involved) to extract the user-related (and password protected) information automatically from the online accounts in order to present it in a coherent way to users, showing their total balances of money and telephone expenses, and allowing them to perform financial transactions between accounts. Construction of such

[13] Non-existing acronym for Semantic Web Markup Language.
[14] See www.getsee.com for the aggregator in ASP mode.

advanced pre-Semantic Web applications requires significant investment (about a year for a medium-sized team). It is our vision that mature Semantic Web technology will significantly lower the barriers to offer such advanced services over the web, opening the way for much more players on the web, and thus raising the number of added value services for people.

6 Conclusions

In this paper we identified what in our opinion are some of the main challenges the Semantic Web faces in the coming years. The major challenges include: (i) the availability of content, (ii) evolving ontologies, (iii) scalability, (iv) multilinguality, (v) visualization to reduce information overload, (vi) stability of Semantic Web languages. Current research projects specifically funded, for developing Semantic Web technology, by the European Commission, DARPA and other funding agencies, should provide answers and solutions to most of those issues over the coming 3 years.

In our vision, the expected scientific, technological, economic and social implications of the Semantic Web are large. The recent history of computing sciences has shown the disruptive nature of some technologies and their implications on many aspects of life. Internet technology changed the way people communicated. WWW technology changed the way people work, thereby making other technologies obsolete. We are moving from EDI to XML, from CORBA to SOAP. All these technological changes have impact on the industry providing services and products in these areas. High-tech companies focusing completely on one technology, run the risk to become obsolete when such technological discontinuity occurs.

Likewise, the Semantic Web technology will contribute to the change of HTML to XML to SWML (OIL, DAML+OIL, RDFS, etc.). Companies currently focussing on HTML run the risk to run out of business when the Semantic Web comes. There is, however, still a very long way to go before the Semantic Web dream becomes true, and will change the information society and the information economy. But, ultimately this is what the Semantic Web envisions: from syntax to semantics, from information retrieval to task delegation, from information overload to infomediaries.

Bibliographic References

[Benjamins et al (KA)2] Benjamins R., Decker S., Fensel D., Gómez-Pérez A. "(KA)2: Building Ontologies for the Internet": A Mid Term Report. International Journal of Human Computer Studies (IJHCS). 51(3). September 1999. 687-712.

[Brickley et al] Brickley, D., Guha, R.V. Resource Description Framework (RDF) Schema Specification. W3C Proposed Recommendation. March, 2000. http://www.w3.org/TR/rdf-schema/

[Budanitsky et al 2001] Budanitsky, Alexander and Hirst, Graeme.`Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures.' Workshop on WordNet and Other Lexical Resources, Second meeting of the North American Chapter of the Association for Computational Linguistics, Pittsburgh, June 2001.

[DAML-S] Ankolekar A., Burstein M., Hobbs J., Lassila O., Martin D., McIlraith S., Narayanan S., Paolucci M., Payne T., Sycara K., Zeng H. DAML-S: Semantic Markup for Web Services, Proceedings of the International Semantic Web Working Symposium (SWWS). July 30-August 1, 2001.

[Decker et al]. Decker, S.; Erdmann, M.; Fensel, D.; Studer, R. Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information. In R. Meersman et al. (eds.): Semantic Issues in Multimedia Systems. Proceedings of DS-8. Kluwer Academic Publisher, Boston, 1999, 351-369.

[Etzioni, Perkovitz 1997] Oren Etzioni, Mike Perkowitz: "Adaptive Web Sites: an AI Challenge", IJCAI-97

[eurowordnet] Vossen, P. (eds) (1998) EuroWordNet: A Multilingual Database with Lexical Semantic Networks, Kluwer Academic Publishers, Dordrecht

[Fensel et al, 2001] Fensel D., Ding Y., Omelayenko B., Schulten E., Botquin G., Brown M., and Flett A., "Product Data Integration in B2B E-commerce", IEEE Intelligent Systems (Special Issue on Intelligent E-Business), Vol. 16, No. 4, July/August 2001, pp. 54-59.

[Fernandez et al 1997] Fernández, M.; Gómez-Pérez, A.; Juristo, N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Symposium on Ontological Engineering of AAAI. Stanford (California). March 1997.

[Gomez-Perez 1996] Gómez-Pérez, A. "Guidelines to Verify Completeness and Consistency in Ontologies" The third World Congress on Expert Systems. February. 1996. Pags. 901-908.

[Van Harmelen et al 2001] Frank van Harmelen, Jeen Broekstra, Christiaan Fluit, Herko ter Horst, Arjohn Kampman, Jos van der Meer and Marta Sabou, "Ontology-based Information Visualisation", Proceedings of the workshop on Visualisation in conjunction with the 5th International Conference on Information Visualisation, London 2001

[Horrocks et al], Horrocks, I., Fensel, D., Harmelen, F., Decker, S., Erdmann, M, Klein, M. "OIL in a Nutshell". Proceedings of the 12th European Workshop on Knowledge Acquisition, Modeling, and Management ({EKAW}'00), R. Dieng

[Knoblock et al Ariadne]. Knoblock, C.A., Minton, S., Ambite, J.L., Ashish, N., Muslea, I., Philpot, A.G., Tejada, S. The ariadne approach to web-based information integration. International the Journal on Cooperative Information Systems (IJCIS) Special Issue on Intelligent Information Agents: Theory and Applications, 10(1/2), pp145-169, 2001.

[Knoblock et al]. Knoblock, C.A., Minton, S., Ambite, J.L., Muslea, M., Oh, J., Frank, M. Mixed-Initiative, Multi-source Information Assistants. The Tenth International World Wide Web Conference (WWW10), Hong Kong, 2001.

[Lassila et al], . Lassila, O., Swick, R. Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation. February, 1999. http://www.w3.org/TR/REC-rdf-syntax/

[Lawrence & Giles] S. Lawrence and C.L. Giles, "Accessibility of Information on the Web," Nature 400:107-109, July 8, 1999.

[Luke et al]. Luke, S., Heflin, J. SHOE 1.01 Proposed Specification, SHOE Project, Feb. 2000. http://www.cs.umd.edu/projects/plus/SHOE/spec1.01.htm.

[Maedche, Staab – 2001] Alexander Maedche, Steffen Staab: "Learning Ontologies for the Semantic Web", Semantic Web WorkShop 2001, Hongkong , China

[Michael K. Bergman] Michael K. Bergman "The Deep Web: Surfacing Hidden Value" White Paper http://www.brightplanet.com/deepcontent/tutorials/DeepWeb/index.asp

[Noy, Musen 2000] N. F. Noy & M. A. Musen "PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment". In the Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000), Austin, TX. Available as SMI technical report SMI-2000-0831 (2000)

[owl] "Language standardization for the Semantic Web:The long way from OIL to OWL" D. Fense1, Vrije Universiteit Amsterdam (VU): http://www.cs.vu.nl/~dieter/ftp/paper/dwc.pdf

[SOAP] SOAP: Don Box et al: "Simple Object Access Protocol (SOAP)" http://www.w3.org/TR/SOAP/ , May 2000

[UDDI] UDDI: "UDDI Executive White Paper" www.uddi.org , Nov 2001

[UPML] Fensel D., Benjamins V.R., Motta E., Wielinga B. UPML: A Framework for knowledge system reuse. Proceedings of the International Joint Conference on AI (IJCAI-99), Stockholm, Sweden, July 31 - August 5, 1999. Available at: http://www.cs.vu.nl/~upml/

[DAML+OIL] Reference description of the DAML+OIL language. http://www.daml.org/2001/03/reference.html Frank van Harmelen, Peter F. Patel-Schneider and Ian Horrocks, editors.

[wordnet] Miller, George A. ``WordNet: a lexical database for English." In: Communications of the ACM 38 (11), November 1995, pp. 39 – 41.

[WSDL] WSDL: E. Christensen (Microsoft), F. Curbera (IBM Research), G. Meredith (Microsoft), A. Weerawarana (IBM Research), "Web Services Description Language (WSDL)" http://www.w3.org/TR/wsdl , Mar 2001

[WSFL] "Web Services Flow Language" Prof. Dr. Frank Leymann , IBM Software Group: May 2001