Knowledge Web Technology Roadmap "The Technology Roadmap of the Semantic Web"

This Knowledge Web Technology Roadmap is meant for any actor who is interested in the Semantic Web.

The aim of this document is to describe Semantic Web challenges providing new visions for investments and adoptions.

Authors:
CUEL Roberta
DELTEIL Alexandre
LOUIS Vincent
RIZZI Carlo

With a special thanks to Ian BLACOE, Kris JACK, Alain LEGER, Lyndon NIXON and Pavel SHVAIKO

1. The Semantic Web Mainstream

Since 1998, researchers, in particular members of the W3C consortium, have been discussing the idea of a "Semantic Web", in order to overcome the simple way of classifying online data such as pictures, text, or database entries (data or document centred view), in favour of a system of relations among object categories and concepts (meaning centred view) [John Borland, 2007]. In order to develop semantic-based technologies, meanings can be managed autonomously, separating them from data, content and application code, so that machines as well as people can understand, share and work with them [Mills, 2004]. To this end, metadata, taxonomies, classifications, context, and ontologies have been largely studied and adopted, becoming the basic building blocks for the Semantic Web.

Considering the original definition, the Semantic Web is a web of actionable information such as information derived from data through a semantic theory for interpreting the symbols (for an in-depth discussion see [Shadbolt et al., 2006]). In other words, the semantic theory provides an account of "meaning" in which the logical connection of terms (i.e. objects concepts and categories) establishes interoperability between systems [Berners-Lee et al., 2001]. As a result, the Semantic Web can be considered as "the web of meaning" [MacManus, 2005]. Although this vision seems very simple and effective, it remains largely unrealised and maintains huge growth potential [Shadbolt et al., 2006].

A lot of researchers have been involved in Semantic Web studies for a long time, and practitioners are starting to be attracted by the Semantic Web. The Semantic Web will constitute one of the most challenging topics in all of the business sectors (particularly in the IT field), and will be developed and adopted in the near future. As described by a Merrill Lynch analyst, Norman Poire, we are only at the beginning of the distributed intelligence wave and the Semantic Web, as part of distributed intelligence technologies, will likely develop until the 2060's.

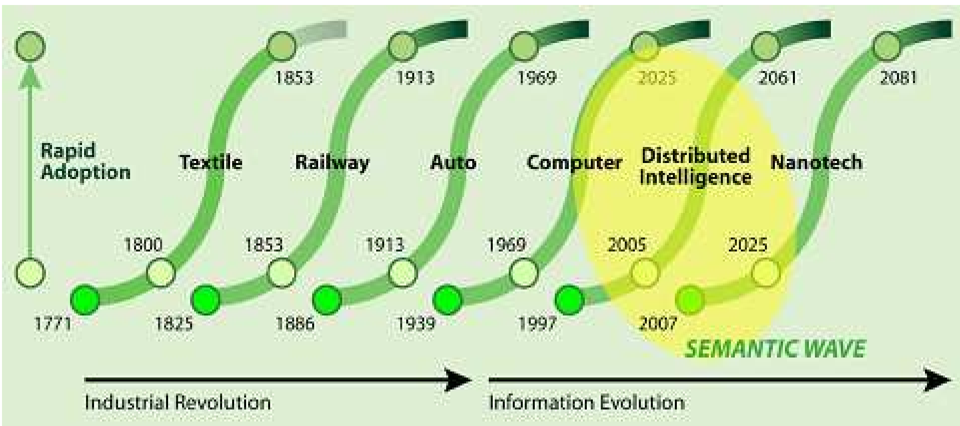

Figure 1. Semantic Wave (Source: Norman Poire, Merrill Lynch, 2006)

1.1. Positioning the Semantic Web – a Common Misunderstanding

In recent years, semantic-based technologies have been increasing their relevance in both the business and research worlds. Often non-experts may misunderstand the real meaning of the Semantic Web, confusing it with Web 2.0. This is due to the fact that the Semantic Web and Web 2.0 have:

- similar dates of birth: in 2004, the RDF/OWL recommendations were approved by the W3C and the term of "Web 2.0" was coined by Tim O'Reilly and MediaLive;
- similar technical inspirations: both the Semantic Web and Web 2.0 combine natural language, taxonomies and tools in an open environment (see the gray boxes for examples of both Semantic Web and Web 2.0 applications) [John Borland, 2007];
- similar goals: both the Semantic Web and Web 2.0 tend to improve web capabilities from different points of view. The Semantic web is more computer-oriented, while Web 2.0 is more user-oriented and socialization is one of the key factors [Graves, 2007].

Web 2.0 considers the web as a social platform [MacManus, 2005], adding a new layer of information interactivity based on tagging, social networks, and user-created taxonomies (often called "folksonomies"). The Semantic Web is based on artificial intelligence and knowledge representation theories. It develops language standards, such as RDF, OWL, rule-based languages, and the other topics described in the cake model, which focus on machine readable semantics (knowledge).

Examples of Web 2.0 Applications

There are a lot of Web 2.0 based applications aimed at supporting the management of various sources of knowledge such as text, photos, and videos. Some of these are:

Wordpress - *http://wordpress.org*
Wordpress is a weblogging application, a "semantic personal publishing platform", which collects articles and information about any specific subject (e.g. services and products) and allows people to discuss it. Each article reflects the opinions and interests of their owner/writer through post meta data, categorizations, labels and comments to the posts.

Digg - *http://digg.com*
Digg is a web content sharing solution where end-users are looking at information from a collective community point of view. It represents a Web space where people can jointly determine the value of content items and interact with other people to discover, select and share new ones. Furthermore, Digg provides a collaborative editorial process through which end-users discuss news, videos, podcasts and any kind of information available online.

Flickr - *http://www.flickr.com*
Flickr is an example of a Web 2.0 solution for sharing photos online. This open-end system manages a massive amount of pictures that are labelled by end-users. People can upload images, organise photos using collections, sets, and tags, share pictures, make photo books, frame prints, DVDs, etc. and also keep in touch with friends.

Del.icio.us - *http://del.icio.us*
Del.icio.us is an open-ended collection of favourites that people has already found. Each end-user keeps links, shares favourites and discovers new things, using tags to organise information (web pages, documents, videos and images online) in a flexible and personal way.

YouTube - *http://www.youtube.com*
YouTube is the leader in online video where end-users share their personal videos. YouTube allows people to create specific thematic channels and groups sharing them.

Key Benefits of Using Web 2.0 Technology

- Web 2.0 applications are usually end-user friendly and satisfy end-user needs (see the concept of "*consumeriz*ation" in paragraph 2.2.1);
- Web 2.0 applications take advantage of social connections among end-users;
- people can provide personal labels and tags in order to represent (through knowledge objects such as pictures, videos, and texts) personal visions;
- people can share personal points of view, taking advantage of other perspectives.

Examples of Semantic Web Applications

There are many Semantic Web-based applications which can support the management of various sources of knowledge such as text, photos, and videos. Some of them are:

Real Time Suggestion of Related Ideas in the Financial Industry
http://www.w3.org/2001/sw/sweo/public/UseCases/Bankinter
Bankinter, the fifth largest Spanish bank, allows its 4,200 employees to submit ideas for new products or services, cost reduction solutions, or innovative organizational processes. Bankinter is developing a Semantic Technology to effectively manage human resource assets. This solution helps employees to enter a new idea and the system analyses, in real time, the text and recognises the relevant concepts from a financial perspective. It is based on ontologies.

The swoRDFish Metadata Initiative: Better, Faster, Smarter Web Content
http://www.w3.org/2001/sw/sweo/public/UseCases/Sun
Sun Microsystems, Inc. uses Semantic Web technology to dynamically develop web content on several of its external sites. The swoRDFish Metadata Initiative provides a framework to address web content development. Thanks to RDF and ontologies the tool enables end-users to create, classify, and manage information about resources and their relationships better and faster within centrally managed vocabularies and taxonomies.

CHOICE project
http://ems01.mpi.nl/choice-wiki/CHOICE_Main_Page
The CHOICE project is part of the CATCH program of the Nethererlands Organization of Scientific Research (NWO). The CHOICE project focused on semi-automatic semantic annotation and context information. Semantic annotation involves the annotation of archived objects, such as video, images and books with semantic categories from standardised metadata repository. Using semantic annotation it helps people to search through a wide collection of objects.

mSpace Mobile
http://eprints.ecs.soton.ac.uk/11101/01/iswc-final.pdf
mSpace Mobile is an Semantic Web application that provides access to location-based information while on the move. Especially applicable to those unfamiliar with their surroundings, the application provides information about topics of chosen interest, based upon the location, as determined by an optional GPS receiver. This application combines an innovative interface and architecture to support exploration of rich information spaces.

Key Benefits of Using Semantic Web Technology
- Enhanced search and navigation of a complex corporate Web site;
- Improved visualization capabilities to guide end-users to rich information spaces;
- Improved ability to personalise a Web site.

Nowadays, some researchers and practitioners talk about Web 3.0 as the resulting convergence of both the Semantic Web and Web 2.0. In other words:

> *"The thing being called Web 3.0 is an important subset of the Semantic Web vision, (…) It's really a realization that a little bit of Semantic Web stuff with what's called Web 2.0, is a tremendously powerful technology" [Borland, 2007].*

Researchers, practitioners and venture capitalists have also started new collaborations to develop Web 3.0 tools and applications[1]. An example of such an application is the Garlick system[2], created by the internet bank, Egg. Thanks to ontology-based systems (semantic data-organization) and Web 2.0 features, a hybrid system gives people the power to take back control of their personal information and to protect them from any criminal intent [Lassila and Hendler, 2007]. Another example is Joost[3], a peer to peer TV that people can program to personalise virtual TV networks. The Joost creators, the same ones who developed Skype and Kazaa, have used techniques based on RDF.

---

[1] See http://www.investors.com/editorial/IBDArticles.asp?artsec=16&artnum=3&issue=20070601
[2] See https://www.garlik.com/index.php
[3] See http://www.joost.com/

1.2. Stakeholders and the Value Chain of the Semantic Web

Stakeholders are people and organizations, who are affected by and can influence the Semantic Web. Several main actors, which can be considered as both consumers and producers ("*prosumers*") of the Semantic Web, have been identified. The Semantic Web value chain represents groups of interests which co-create value through knowledge and experience exchanges.

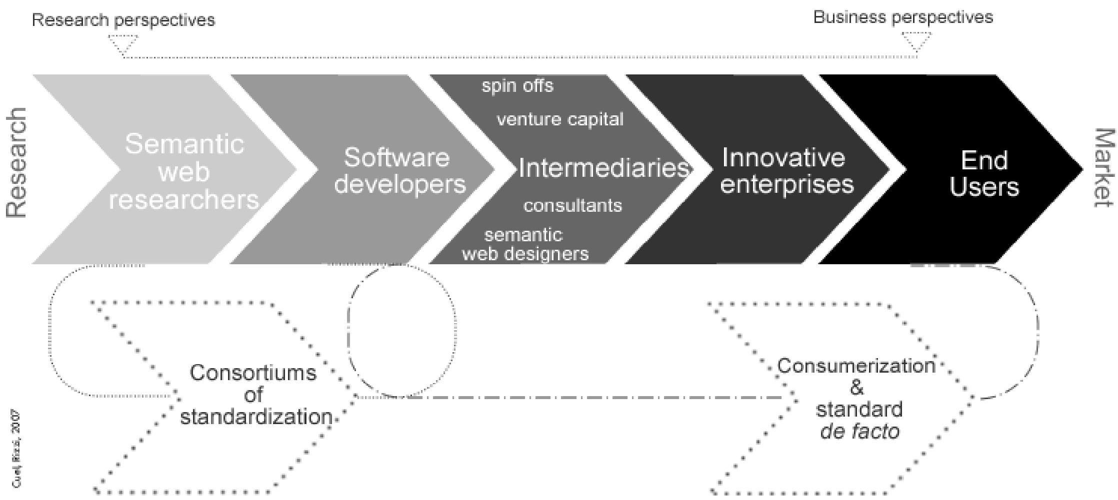

Figure 2. Stakeholders and the Value Chain of the Semantic Web.

Figure 2 shows the value chain of the Semantic Web's evolution. It connects researchers, who mainly produce Semantic Web theories and methods, computer science firms, which mainly produce solutions, and firms and end-users, which adopt and use Semantic Web technologies.
According to this value chain the following actors have been identified:

- Semantic Web researchers, who are directly involved in European and international projects, are developing and innovating theories and methods of the Semantic Web (i.e. Semantic Web languages, Semantic Web services or algorithms to deal with reasoning, scalability, heterogeneity, and dynamics). In order to validate the resulting theories, researchers often develop application prototypes which are tested within innovative firms.
- Consortiums of standardization are interested in sorting out new recommendations and standards for Semantic Web technologies, giving the basics for innovative tools.
- Software developers are interested in developing Semantic Web solutions and applications. They can directly sell the latter to firms and, at the same time, test innovative theories and methods.
- Intermediaries are interested in transferring technology and knowledge from researchers and developers to practitioners. Intermediaries can assume the form of Semantic Web designers, consultants, venture capitals, spin offs, etc.
- Innovative enterprises are interested in catching new opportunities from the Semantic Web and also developing new business models.

- End-users are interested in obtaining useful and effective solutions. One of the most important requisites is to use a transparent technology - "No matter what is behind, if it works". Although end-users are positioned at the end of the value chain, their needs are very relevant for all Semantic Web stakeholders.
- Standardization de-facto. End-users should be considered as a central element in the production of social and semantic-based technology, thus should be strongly connected with all of the other value chain actors. Also, the *consumerization* phenomenon is unveiling new standards, which are de-facto substituting for consortiums of standardization.

A few years ago, the Semantic Web became an interesting vehicle of innovation that only the Information and Communication Technology (ICT) sector considered as very relevant. Nowadays, thanks to several European projects developed in different fields, a lot of companies are considering Semantic Web technologies as very challenging, and are starting to test these technologies in real applications. In many sectors, such as the ones that the Knowledge Web NoE has explored, beneficial results are wide spread [4]:

- aerospace;
- automobile industries;
- banking and finance;
- consumer goods;
- distribution energy and public utilities;
- environment;
- government and public services;
- food industries;
- industry and construction;
- pharmaceuticals and health;
- service industries;
- sports;
- technology and solution providers;
- telecommunication;
- transport and logistics.

As traditional technologies, it is not that one domain is more suitable than another, but the Semantic Web will spread all over the industry sectors, as an important building block of any sort of application. In other words, the Semantic Web is affecting most of the industry sectors in which technology and knowledge are relevant assets to be managed.

---

[4] The main objective of the Knowledge Web's outreach to industry area is to promote a greater awareness and a faster uptake of Semantic Web technology within Europe in full synergy with the research activity. In the project web site, the portal http://knowledgeweb.semanticweb.org/o2i is aimed at promoting the most important activities of the outreach to industry area, providing information on innovative results, requiring industry to give feedback to the results and deliverables of the technical committees which better suit companies' needs. It will help keep the scientific results aligned with the actual needs of industrial and governmental organizations and will also play a prominent role in disseminating these results in user communities.

2. Business Perspectives

One very important aspect is to improve the social awareness of how Semantic Web technologies could practically help organizations deliver new products and services and create new business value. In this way the technology roadmap should give a vision, such as correct information and understandable recommendations, show innovative solutions and show how firms might use them appropriately. For these reasons, experts (both researchers and practitioners) were asked to provide useful insights on what Semantic Web businesses lack, which should be overcome in the next few years, and what type of Semantic Web technologies are ready to be used in daily life. Both researchers and practitioners have filled in questionnaires, have participated in focus groups and workshops, and have provided visions, experiences and results about their research in actions (case studies, experiments, etc.). In this chapter some important Semantic Web gaps are depicted, a holistic approach that tends to solve these problems is proposed, and a Semantic Web Hype Cycle Curve is described. In particular, the hype cycle curve focuses on the practitioners' points of view, which are related to three main Semantic Web categories:

- research and basic components;
- business applications;
- and technologies.

2.1. Semantic Web Business Gaps

Thanks to many innovative projects and initiatives, the Semantic Web is evolving rapidly but, in Europe, it seems far from providing saleable products or services. Consider the fact that many organizations and start-ups declare using or selling semantic-based technologies, but often develop and adopt very light semantic annotation tools, XML, etc. Consequently, practitioners might believe that semantic-driven applications are not widely used in industry, because the benefits from new technologies will not cover the payout. This depends on the following factors:

- Semantic Web solutions are still too complex to be easily used and understood by practitioners and end-users. Technicians only can deal with them and end-users – without any technical skill – do not completely understand the usefulness of these technologies, thus, do not use them;
- European business in semantic technology is less mature than the American market. For instance, the number of firms which participated in the first European Semantic Technology Conference 2007 (200 firms) is quite low, in comparison with the huge participation of USA companies in SemTech 2007 conference (700 companies, almost 40 case studies, more than 35 sponsors) and the involvement of a huge number of venture capital companies in the American market. Although

this is an unstable situation, an increasing number of workshops and events on the Semantic Web and business applications, have been organised in order to share research results with industry[5];

- Most of the innovative solutions developed by researchers are not immediately suitable and sustainable within industrial settings. There are some case studies and experiments, but Semantic Web solutions are not on the plateau of productivity;
- Various Semantic Web developers/researchers are very focused on their technical problems and do not use a business vision to develop effective semantic solutions (e.g. few design techniques take into account business needs) [Singh et al., 2006]. Although researchers and developers tend to solve technologically open problems rather than business problems, several big organizations consider Semantic Web technologies as innovative tools for their core business.
- End-users value should be unveiled and profitably used by organizations. Even if end users might decree the success of innovative solutions and tools, organizations might develop new business models to propose and eventually push pioneering solutions and tools and reach new market niches.

The factors described above point out one weak element: the technology transfer process, which is defined as the successful process of converting research outputs into marketable products and services (developing Semantic Web solutions) [EIF, 2005]. There are some contextual needs and conditions (deeply studied in organizational, knowledge management, innovation, and information systems studies) that foster successful technology transfer processes and successful business solutions. In the case of the Semantic Web, these are data integration and data quality, search and information access, reconciliation and content evaluation [Gartner Report, 2007c].

A holistic approach to Semantic Web technology transfer processes is proposed, which take into account, at the same level of importance, innovative research outputs, foreseen end-user benefits, business earning opportunities, and business sustainability conditions [Singh et al., 2006].

---

[5] Some of the most important research in action workshops are:

FIRST 2007: First Industrial Results of Semantic Technologies (*http://www.disa.unitn.it/net-economy/first/*) in conjunction with the International Semantic Web Conference and the 2nd Asian Semantic Web Conference, in Busan, Korea, 2007 (*http://iswc2007.semanticweb.org*).

STAB 2007: First International Workshop on Semantic Technology Adoption in Business (*http://events.idi.ntnu.no/stab07*) and Workshop on Semantic Business Process and Product Lifecycle Management (SBPM - *http://sbpm2007.fzi.de*). In conjunction with the 4th European Semantic Web Conference in Vienna, Austria, 2007 (*http://www.eswc2007.org*).

MSWFB 2007: Making Semantics Work For Business (*http://www.ag-nbi.de/conf/MSWFB*) co-located with Business Aspects of Semantic Technologies (BAST). SWTS 2007: Semantic Web Technology Showcase (*http://www.csc.liv.ac.uk/semanticweb/KWeb_ESTC_WS.html*). In conjunction with the European Semantic Technologies Conference (ESTC) in Vienna, Austria, 2007 (*http://www.estc2007.com*).

SEBIZ 2006: First International Workshop on Applications and Business Aspects of the Semantic Web (*http://www.ag-nbi.de/conf/SEBIZ06*) in conjunction with the Fifth International Semantic Web Conference (ISWC 2006) in Athens, Georgia, USA, 2006 (*http://iswc2006.semanticweb.org*).

2.2. A Holistic Approach to the Semantic Web

Semantic Web technologies, as all technological solutions, should be shaped according to organizational processes, practices, and business needs. There are a lot of experiments in business and social sciences, which show how a technology might be effectively implemented within a firm, and what the impacts of the implementation and adoption phases are. Often, a new technology is deserted by end-users, who prefer to continue to use obsolete solutions. This might be caused by the fact that a solution does not suit end-user desiderata perfectly, thus results are worthless or very complex to understand and complicated to use (this is considered oppressive by end-users) [Bowker and Star, 1999]. All actors along the value chain (described in Figure 2) should be aware of key business opportunities. On one side, researchers, developers and intermediaries should satisfy end-user needs, and on the other side end-users and firms should better understand the innovative opportunities of Semantic Web technologies.

A methodology that enables this holistic approach is based on the analysis of three perspectives, which respectively focus on:

- people, who use the Semantic Web, i.e. end-users, consumers, "*prosumer*s", and companies;
- processes, which are the way in which people use the Semantic Web;
- products, which are innovative Semantic Web solutions such as technologies and applications. Technical aspects of Semantic Web technologies will be described Chapter 3.

These three components are used to understand how the Semantic Web might interact with culture, behaviours and practices of end-users, teams, and organizations.

2.2.1. People

As already tested in several e-commerce and Web 2.0 experiments, end-users take a very important role in product/service design, production, dissemination, etc. [von Hippel, 2002]. End-users move from a passive role, as simple readers, to an active role, as authors, regarding web blogs, wiki systems, and other collaborative tools. Namely, end-users are producers and consumers ("*prosumers*") at the same time, and are socially related to share and create knowledge. This social aspect is defined by Gartner as *"consumerization"* and focuses on the ongoing trend of new technologies and models appearing in the consumer world before evolving into enterprises. For instance, wiki systems, blogs, flickr, p2p tv and video games (typically developed in Web 2.0) have been spread over the consumer market as end-users' tools, before being adopted by companies, like knowledge sharing and marketing tools. *Consumerization* focuses, also, on how these technologies and models can be safely designed for enterprise consumption [Gartner Report, 2007b], due to the fact that Web 2.0 consumer markets continue to evolve and innovate.

The Semantic Web can be used to semantically integrate, reconcile and evaluate enterprise content, concerning corporate knowledge and data management. The need to share knowledge among individuals in a very complex organization, or among networked organizations, increases the importance of introducing new Semantic Web technologies. These allow practitioners to integrate individual knowledge into the organizational knowledge, according to semantic negotiation and alignment processes.
In mobile and distributed work settings (such as consultancy companies, team or project-based organizations, etc.) Semantic Web technologies, as social software, have a lot of potential. It can help workers to stay in contact with other remote colleagues enabling them to cooperate both asynchronously and synchronously [Koskinen, 2006].

2.2.2. Processes

In complex organizations composed of a constellation of units, which manage specialised processes autonomously, information communication technologies and knowledge management systems must take into account the distributed nature of knowledge, and should allow coordination among autonomous units [Cuel, Bouquet, Bonifacio, 2005]. Considering each unit as a community or an informal social group based on locally shared interests and practices [Lave and Wenger, 1991; Wenger, 1998; Starbuck, 1992], organizations can thus be viewed as constellations of communities of practices [Brown and Duguid, 1991]. In order to improve knowledge sharing, organizations often need to make changes to the way their internal and external processes are structured, and sometimes the way organizational structure is defined. For a long time technologies have been proposed and applied as neutral tools whose implementation does not have any impact on organizational processes nor consumer uses. Opposed to that point of view, studies focused on structuration theories [Giddens, 1984; Orlikowski, 1991] do not consider technology as a neutral asset, but focus on the fact that technologies and social structures (such as the de facto organizational models, worker uses, people behaviours, etc) are strongly related and interdependent. Thus, as demonstrated in [Cuel, Bouquet, Bonifacio, 2005] the introduction of an unsuitable solution (a centralised system in a very distributed organization, or vice versa) might cause unexpected reactions, forcing end-users to desert the system.

Semantic Web technologies [Fensel et al., 2001] might overcome this problem, offering more flexible and automated reasoning applications, allowing people to interact in a distributed system to share knowledge according to common corporate conceptualization. The Semantic Web, also, provides explicit meaning to the information available on the Web (Internet or Intranet) for automated processing and information integration based on underlying corporate ontologies.

2.2.3. *Products (Semantic Web Solutions)*

As described above, technologies should be shaped by the processes, practices, and organizational models in which they are implemented. Thus, Semantic Web solutions and products should be developed according to a holistic approach in order to be transparent to end-users. Semantic Web technologies should be integrated into solutions that are not only effective (in theories) but also valuable for end-users. As a matter of fact, Semantic Web technology is often a crucial enabler of knowledge management, content sharing and knowledge representation, but is not a solution itself. It should be integrated into a "killer application", which helps people to:

- organise data and access them in a very easy and personalised way;
- transfer and integrate information in a frictionless way [Pollock and Hodgson, 2004];
- reconcile and evaluate content, etc.

In order to facilitate the development of effective Semantic Web products, a lot of Semantic Web technologies and theories should be developed and standardised. On this track, researchers aim for the languages and techniques that they develop to be recognised by W3C or similar consortiums. See, for instance, the cake model which describes various Semantic Web languages, rules, and logics; processes and techniques for semantic annotation, matching, versioning, reasoning, etc. In turn, for instance, languages, rules and ontologies might be used to design innovative semantic solutions based on more traditional tools, such as blogs and wikis, grid tools, P2P platforms, mobile communication systems, Semantic Web Services and agents, which are implemented for e-Learning, B2B and B2C, human resource, knowledge management, e-government, bioinformatics and e-health, etc.

Some promising applications are based on:

- *Knowledge management.* Although information and knowledge are stored in various forms (e.g. documents, processes, policies, etc.) and are usually organised by experts according to their personal points of views (using personal keywords, paths folders, etc.), they should be retrieved, selected, and understood by anyone in the company. The Semantic Web might allow people to find the right information in dynamic, complex and distributed corporate knowledge bases, enabling the development of dynamic vocabularies of keywords and concepts.
- *Semantic Web Services.* Web services are self-contained, modular business process applications that web users or programs can publish, discover and invoke in a distributed computing environment. In practice, web services became very interesting tools to develop real-time interactions among information systems and business processes. Although a huge number of companies invested in very complex Internet and Intranet portals which support an increasing amount of web services, the latter are not semantically integrated. Therefore, innovative Semantic Web products are expected to play an important role in the semantic interoperability of web services.

- *Integration and interoperability among organizations* is growing rapidly and organizations spend a lot of resources (in term of money and time) in dealing with these new needs. The increasing product delocalization, business coalitions, and co-optations processes force organizations to share a growing amount of information and knowledge. These are organised and used very different ways, and should be integrated through semantic brokers, such as common ontologies, matching processes, etc..
- *Decision support systems.* Data warehouses are based on various forms of data and processes of extraction, transformation and reasoning should solve both syntactic and semantic heterogeneity (see the concepts of inclusion, similarity and relations).
- *Non-structured activities*. All non-structured information and knowledge (i.e. e-mail, messaging, phone calls, taxonomies, indexes, etc.) should be automatically managed by semantic technologies.

2.3. The Knowledge Web Business Perspective

A panel of experts and practitioners in Semantic Web applications were asked to define a Semantic Web hype cycle curve, aimed at unveiling what real world applications, products, and components are available in the market, what technology might be developed in the near future, and what solutions are already obsolete. A questionnaire was submitted to 21 experts during two international events such as:

- the workshop "Formal Ontologies Meet Industry"[6]; an international forum where researchers in different disciplines and practitioners of various industrial sectors met to analyse and discuss issues related to methods, theories, tools and applications based on formal ontologies. Held in Trento (Italy), December 2006.
- the 1st "European Semantic Technology Conference"[7]; a new European meeting ground for customers, developers and researchers to discuss the applicability and commercialization of Semantic technologies in corporate settings. Held in Vienna (Austria), June 2007.

Participants were asked to locate semantic based technologies and applications in the hype curve, using the appropriate signs of the timeframe dimension.

These technologies and applications are divided into 3 main categories:

- basic Semantic Web components, which constitute the building blocks of Semantic Web such as semantic annotation, automatic alignment,

[6] See http://www.loa-cnr.it/fomi

[7] See http://www.estc2007.com

scalability, versioning, standardization, security, trust and legal issues, authoring and collaborative tools;

- business applications, which represent Semantic Web-based technologies that can be introduced and applied to improve intra- and inter-organizational processes: B2B, B2C, knowledge management, e-government, e-learning, bioinformatics;
- technology trends, which represent innovative solutions (hardware and software) that through the Semantic Web offer "smart" services: semantic wiki, semantic blogs, semantic grid, mobile communication, peer to peer, Semantic Web Services, agents.

Figure 3 depicts the results. Some considerations seem interesting. Most of the basic Semantic Web components such as trust, scalability, versioning, and alignment are positioned at the beginning of the technology trigger mostly with a range from 5 to 10 years to mainstream adoption. Also technology trends are positioned in the first part of the curve but closer to the peak of inflated expectation. In particular, P2P, semantic grid, and mobile communication benefit from a very high visibility in the markets (through articles, journals, etc.). Surprisingly, bioinformatics and standardization (regarding web languages such as RDF and OWL) are positioned at the beginning of the slope of enlightenment with a range from 0 to 5 years to mainstream adoption. Notice that no item has been positioned in the plateau of productivity, probably because practitioners believe that no Semantic Web technology (component, application, etc.) is good enough to be immediately adopted by companies and introduced in the market. Although practitioners believe that no Semantic Web technologies and solutions are mature enough, some experiments and case studies that have been developed by researchers and innovative companies have reached a competitive position in the market (see the examples of Ontoprise and Software AG).

Ontoprise *http://www.ontoprise.de*

Founded in 1999, Ontoprise is a leading software provider of solutions based on ontologies.

Products:

The basis of Ontoprise applications are:

- *OntoBroker* processes ontologies and the logic represented inside them. So expert knowledge and business logic can be modelled independently from the execution logic.
- *OntoStudio* is the professional developing environment for ontology-based solutions. It combines in a unique way modelling tools for ontologies and rules with components for the integration of heterogeneous data sources.

Applications:

Some of the Ontoprise applications are:

- *SemanticMines* is a search engine which supports end-user search.
- *SemanticGuide* is a recommendation system for service management.
- *SemanticIntegrator* is an integration platform which facilitate a unified view of heterogeneous information.

Source: *http://www.ontoprise.de*

Software AG http://www.softwareag.com
Founded in 1969, six young employees at the consulting firm AIV (Institut für Angewandte Informationsverarbeitung) established Software AG in Darmstadt. An adaptable and extremely versatile database management system is the basic concept of all the Software AG products. **Products & Applications** The firm provides several products in data management solutions, business application development, and Service-Oriented Architecture. In particular, the "Enterprise Information Integrator" uses Semantic Web technology to dynamically combine meanings and contexts of business data with rules that manage their uses.
Source: http://www.softwareag.com/corporate/company/pressroom/pressreleases/20050309_enterprise_information_integration_page.asp

For each of the items described in the hype curve there are interesting case studies which are mainly presented in the research community. Some tools are developed by large and established companies such as, Adobe, IBM, Cisco, Nokia, Hewlett Packard, and Oracle as well as many small but pioneering companies such as (Unicorn, Network Inference, and) ETeCH, Sirma Group iSOCO. In addition, there is a number of open source and publicly available tools created by public and private research institutions and organizations, such as SEKT, KnowledgeWeb and W3C[8].

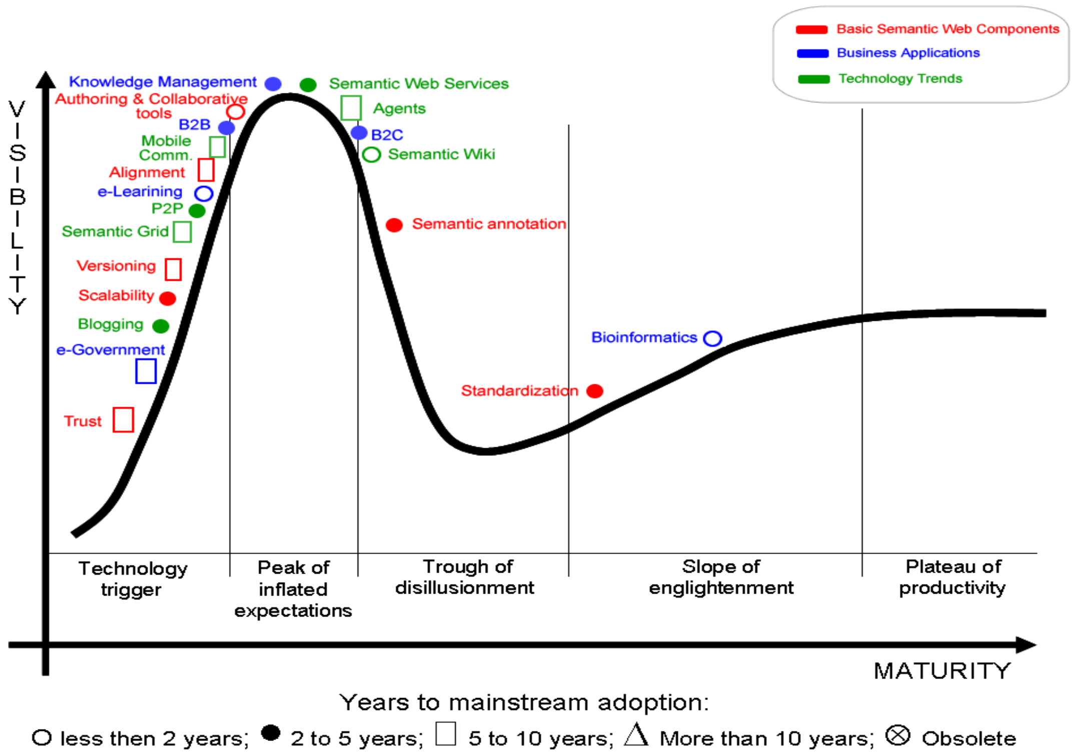

Figure 3. Semantic Web Hype Curve from a Business Point of View

[8] See http://esw.w3.org/topic/CommercialProducts

In the gray boxes, some case studies are presented, referring to:

- basic web components such as trust, authoring and collaborative tools, and semantic annotation;

Improving the Reliability of Internet Search Results Using Search Thresher
From: Semantic Web Use Cases and Case Studies of W3C
(http://www.w3.org/2001/sw/sweo/public/UseCases)

General Description
Search Thresher is a plug in for the Firefox browser that has been developed to improve reliability (trust) of Internet web pages available from a search engine. It's based on the Semantic Web and it runs on metadata, such as Content Labels. It makes some information available about trust in search results.

Key benefits claimed by the developers
For the end-user

- End-users can find what they are looking for and trust what they find.
- End-users will be warned if they are browsing on a site that has made fraudulent claims.

For business

- Sites using Content Labels will get highlighted in search results. Depending on the end-user settings, they may visit only trusted web site.
- End-users will understand which web sites care about standards and codes of conduct.

See more: http://www.w3.org/2001/sw/sweo/public/UseCases/Segala

Geographic Referencing Framework
From: Semantic Web Use Cases and Case Studies of W3C
(http://www.w3.org/2001/sw/sweo/public/UseCases)

General Description
Ordnance Survey is Great Britain's national mapping agency that manages the largest vector geospatial database in the world. Ordnance Survey invested in semantic technologies in order to cut the costs and improve the accuracy of data management by semantic data integration, ontologies, and RDF language.

Key benefits claimed by the developers
For business

- Data integration costs are reduced.
- Accuracy of data is improved.
- The quality of control, classification, and decision-making is improved.

See more: http://www.w3.org/2001/sw/sweo/public/UseCases/Segala

Enhancing Content Search Using the Semantic Web
From: Semantic Web Use Cases and Case Studies of W3C
(http://www.w3.org/2001/sw/sweo/public/UseCases)

General Description
Oracle improved its developer community by Semantic Web technologies (http://otnsemanticweb.oracle.com). Thanks to a partnership with Siderean Software, Oracle have worked on semantic data discovery and navigation by integration of Oracle Secure Enterprise Search and Siderean's Seamark Navigator. Using data stored in RDF and enabling SPARQL queries, Oracle

have aggregated many sources of content and improved the information access processes.

Key benefits claimed by the developers

- The search and navigation processes and the results of a complex corporate web site are enhanced.
- Visualization capabilities are improved enabling the system to address end-user data interests.
- The ability to personalise the web site is improved.

It allows for a quick and simple implementation.

See more: Link: *http://www.w3.org/2001/sw/sweo/public/UseCases/OracleSiderean*

- Business applications such as e-government, knowledge management;

An Intelligent Search Engine For Online Services For Public Administrations, Municipality Of Zaragoza, Spain

From : Semantic Web Use Cases and Case Studies of W3C (*http://www.w3.org/2001/sw/sweo/public/UseCases*)

General Description

The Public Administration of Zaragoza has invested in Semantic Web technologies to provide useful on-line services. Developers improved their web portal and the search engine with a combination of natural language processing and ontological reasoning (using the *Knowledge Index* software).

Key benefits claimed by the developers

For the city of Zaragoza the system:

- Provides a better service to citizens, improving the reputation of public administration.
- Stimulates e-Government processes by providing easier access to relevant information.
- Reduces cost of call center and physical offices because people find information on the web, and can execute semi-automatic services.

For citizens the benefits are:

- Easy interaction, either through keywords or natural language.
- High precision, without loosing recall.
- Concrete answers instead of long lists of documents.
- Innovative services suggested by the relations with the requested ones.
- High level of satisfaction, the system is perceived as a search engine which "understands" the citizen.

Link: *http://www.w3.org/2001/sw/sweo/public/UseCases/Zaragoza*

Use of Semantic Web Technologies in Natural language interface to Business Applications

From: Semantic Web Use Cases and Case Studies of W3C (*http://www.w3.org/2001/sw/sweo/public/UseCases*)

General Description

Tata Consultancy Services Limited (TCS) developed the NATAS system that enables end-users to interact with a business application in natural language by posing questions and invoking tasks. It's based on Semantic Web technology, especially RDF, OWL and SPARQL.

Key benefits claimed by the developers

- Distinct semantics for various concepts in the domain is enabled defining multiple schemas.
- A crisp and simple mechanism to represent an ontology using the

RDF structure is provided.
- Mechanisms to create parts of the ontology and query it seamlessly, are made available.
- Rule evaluation and execution mechanisms enable the creation of *derived facts.*
- Mechanisms to link in *external concepts* with existing concepts of the domain are provided through simple RDF structures.

Link: http://www.w3.org/2001/sw/sweo/public/UseCases/Tata

Semantic Web Technology for Public Health Situation Awareness
From: Semantic Web Use Cases and Case Studies of W3C
(http://www.w3.org/2001/sw/sweo/public/UseCases/)

General Description
SAPPHIRE stands for Situation Awareness and Preparedness for Public Health Incidents using Reasoning Engines. It is part of an Integrated Biosurveillance System developed by the Center for Biosecurity and Public Health Informatics Research at the University of Texas Health Science Center at Houston. SAPPHIRE is a distributed and collaborative system based on the Semantic Web, which helps end-users to contextualise clinical knowledge and tasks by developing abstractions and models on top of integrated data, unstructured text (e.g. doctors' notes), and patient-structured electronic medical records.

Key benefits claimed by developers
The system architecture has the following features:
- Distributed collaboration and interoperability: disparate and heterogeneous data can be exchanged, integrated and utilised seamlessly and dynamically between remote systems.
- Dynamic adaptability: new requirements, data, functions and tasks can be introduced to the system without major rewrites or reprogramming to address novel situations or new tasks.
- Multidisciplinary reuse of information: existing data in the system can be repurposed to address unprecedented use cases.
- Human computer interaction: systems interact intelligently with human users, allowing for more effective, intuitive and easy communications.

Link: http://www.w3.org/2001/sw/sweo/public/UseCases/UniTexas

- Technology Trends such as Semantic Wiki, Agents and Media;

Semantic Media Wiki,
From: http://en.wikipedia.org/wiki/Semantic_MediaWiki

General Description
Semantic MediaWiki is an extension to MediaWiki (a wiki application on which Wikipedia and other sites run), that allows for the encoding of semantic data within wiki pages, thus turning a wiki that includes the extension into a semantic wiki. Data that has been encoded can be used in semantic searches, used for aggregation of pages, and exported to the outside world via RDF.

Key benefits claimed by the developers
- End-users can find what they are looking for, organise, browse, evaluate, and share the wiki's content.
- End-users will help to produce content faster.

See more: http://en.wikipedia.org/wiki/Semantic_MediaWiki

Rcal - Retsina Calendar Agent, From: http://www.cs.cmu.edu
General Description The RETSINA Calendar Agent (RCAL) is a distributed meeting scheduling agent based on the Semantic Web. The Agent runs on Semantic Web content to act as a useful meeting scheduling assistant. Thanks to this application, meeting and events are automatically scheduled in a better way for its user, without continually requesting additional information. **Key benefits claimed by the developers** • End-users can use an automatic schedule function to arrange meeting end events.
See more: http://www.cs.cmu.edu/~softagents/papers/payne_terence_2002_2.pdf

A Digital Music Archive (DMA) for the Norwegian National Broadcaster (NRK) using Semantic Web techniques, From: Semantic Web Use Cases and Case Studies of W3C (http://www.w3.org/2001/sw/sweo/public/UseCases)
General Description Digitising the complete radio and television broadcasting production process is a major undertaking in many public and commercial broadcasters. The Norwegian National Broadcaster (NRK) wants to increase the efficiency and effectiveness of the production process in public broadcasting through Semantic Web technology. The Digital Music Archive is based on metadata representation by Xml, RDF and uses SPARQL. **Key benefits claimed by the developers** **For the end-user** • End-users can find what they are looking for. • End-users will help to faster produce multimedia content. **For business** • Highly efficient music archive, combining multi-channel access with a fully automatised ordering and production flow. • Better integration across multiple archives and resources.
See more: http://www.w3.org/2001/sw/sweo/public/UseCases/NRK

3. Research Perspectives

The Semantic Web vision given by Tim Berners-Lee and others, which is currently supported by the World Wide Web consortium, is quite ambitious and has to be gradually realised (and in particular outreached to industry) in the long term. Thus, this grand vision both represents and stands on an ongoing research framework, which has early roots in computer science, more precisely in formal logics, knowledge representation and reasoning, and databases.

In this chapter, the Semantic Web stakes are analysed at this technical research framework level. First a few prominent functionalities provided by the current state of semantic technologies are described. Then, the vision of the Knowledge Web network experts on the evolution of some topics related to the Semantic Web is presented. Finally, the current research directions, which aim at supporting the scaling up of semantic technologies from closed Intranets to the open Internet, are discussed.

3.1. Main Semantic Web Components

A decade of research on Semantic Web technology has led to a large spectrum of robust semantic components that have been developed and are now available to be integrated in applications. An illustration of this diversity is the Semantic Web Framework that is currently elaborated in the Knowledge Web project[9]. It is intended to provide a toolkit enabling developers to easily enhance their applications with semantic features. It allows them to reuse modules providing a semantic functionality and avoids them to recode these semantic capabilities from scratch.

The functionalities provided by existing tools and components cover almost all the dimensions of the ontology and metadata lifecycle and include in particular:

- ontology authoring, generation and maintenance;
- metadata authoring, extraction, storage and querying;
- metadata and ontology querying and reasoning;
- Semantic Web Services;
- ontology matching;
- semantic metadata display and visualization.

3.1.1. Ontology and Metadata Authoring and Generation

Ontology editors: these assist end-users and system designers to manage ontologies by providing a graphical interface for the definition and the description of concepts, relations and individuals and by offering export

[9] See http://knowledgeweb.semanticweb.org

functions for encoding them in a given formal language. For instance Protégé supports standard Semantic Web languages like RDF(S), OWL and SWRL. The integration of editors with reasoners enables end-users to check whether an ontology contains a logical contradiction and can thus help them to debug their ontology by pointing to and resolving possible sources of inconsistencies.

WebODE From: *Ontology Engineering Group, UPM*
Main Aim Ontology Engineering Workbench. **General Description** Ontology engineering tool with a plugable set of services that support the methodology METHONTOLOGY. The list of services in the workbench are: • Import services from XML, RDF(S), OWL, DAML+OIL; • Export services to XML, RDF(S), OWL, OIL, DAML+OIL, Prolog, Java with Jess and UML; • Merge functionality between two ontologies; • Inference; • Evaluation of an ontology using OntoClean method.
See more: *http://webode.dia.fi.upm.es*

Ontology generation: as the creation of ontologies by hand is still a costly and time consuming task, automatic ontology generation (ontology learning) has been studied as a way to support end-users in the ontology development process. Ontology learners (like OntoLearner [Missikoff, 2002] or Text2Onto [Cimiano and Völker, 2005]) combine statistical and linguistic analysis to detect semantic knowledge like synonyms, concept hierarchies, relation signatures and instances of concepts. For example, an ontology learner should detect that "drive" is a relation between "person" and "car" and that "compact" is a subconcept of "car". However the current techniques still cannot replace a user. The results generated by these systems usually need human validation before being reused in other systems and the acquisition of more complex forms of knowledge, like formal definitions of concepts, is still out of range of current technologies [Buitelaar, 2005].

Semantic annotation: tools for the manual semantic annotation of resources, which provide a graphical interface for selecting instances (highlighted portions of text), associating them with concepts, linking them by using relations and displaying the resulting annotation in RDF format. For example, SMORE [Kalyanpur, 2003] allows users to select concepts and relations in several ontologies and to use them in order to form triples describing instances appearing in a web page or in an email. It also features a multimedia annotator where a multimedia user can select a photo or video segment and annotate it with triples.
Furthermore some tools support the end-user annotation by semi-automatically generating annotations. For example, S-CREAM [Handschuh,

2002] exploits an information extraction component to generate semantic annotations from web pages through the use of knowledge extraction rules.
Automatic annotation of multimedia resources can also be done by exploiting computer vision techniques like low-level features extraction, face recognition, and motion tracking. However there exists a semantic gap between the low level annotations generated from the pixel data and the high level semantic concepts that are interesting for the user. For example, high level concepts like "wedding" or "vacation" are still hard to detect in general settings. However in restricted domains, high level concepts, like highlights in a soccer match (penalties, corners, shots on goal, etc.) can be extracted [Assfalg, 2003; Stamou et al., 2006; Petridis et. al., 2006].

GATE

From: *Department of Computer Science, University of Sheffield*

Main Aim

Freely available, highly customisable and extendable tool for language processing.

General Description

GATE (a General Architecture for Text Engineering) is a well-established infrastructure for the customization and development of Natural Language Processing (NLP) components. GATE allows users to handle a variety of linguistic formalisms in a common framework by means of a theory-independent annotation format for encoding metadata associated with documents. The annotations associated with each language resource (e.g. document) are a structure central to GATE, because they encode the language data read and produced by each language processing module. GATE also supports import and export back to the resource's original format (e.g. SGML/XML/HTML).

In addition to providing annotations, GATE also provides support for importing, visualising, and accessing ontologies, and connect the to NLP tools. GATE supports the process of ontology learning and population. Concerning ontology learning, GATE provides the functionalities for linguistic preprocessing of data, and it enables ontology population by harvesting instances automatically from text. GATE is developed as part of the FP6 project SEKT.

See more: *http://gate.ac.uk/documentation.html*

QuestSemantics

From: *SWLab, University of Liverpool*

Main Aim

Semi-automatic semantic annotation, and simplified semantic search.

General Description

QuestSemantics is a software platform supporting the semi-automatic discovery, annotation, filtering and retrieval of semi-structured resources (web, documents, databases, etc.), on the basis of fine-grained business knowledge. QuestSemantics is designed to maximise decoupling of the different types of knowledge represented - such as business domain, task specific and application knowledge. This decoupling aids reuse of both software and knowledge, and easy customization of the platform components, and positions QuestSemantics as a generic framework for semantic annotation and retrieval in a variety of task and domain scenarios.

The QuestSemantics platform is comprised of two main components: a general framework for the (semi-) automatic annotation of resources, based upon a detailed ontological model of the domain, and a search interface for the user friendly formulation and execution of knowledge-based queries over the generated metadata.

See more: http://www.csc.liv.ac.uk/semanticweb

Semantic browsing: semantically enriched browsing and visualization of annotations, like instances of concepts and relations between instances, can be done by using a Semantic Web browser such as MagPie [Domingue, 2004]. MagPie allows users to highlight instances of concepts inside a document and display semantic relations by right clicking on a given instance. Thus this allows users to access to contextual knowledge.

Magpie – Semantic Web Browser

From : KMI, Open University

Main Aim

Supports the navigation through Web and Semantic Web resources within standard web browsers by means of semantic layering.

General Description

Web and Semantic Web are usually seen as two fairly independent technologies. Magpie uses KMi's ontology infrastructure and expertise in handling ontologies to semantically markup web documents on the fly.

Magpie technology is lightweight, yet flexible and providing sufficiently robust and open features for semantically enriched web browsing. Magpie as a web browser plugin aims to identify and filter out the concepts-of-interest from any webpage it is given. The current set of concepts can be influenced by a selection of a particular ontology of concepts and relations.

In addition to identifying the concepts that are relevant from the perspective of a particular ontology, each such concept may provide an applicable set of relations or commands that can be executed. These are accessible via contextual semantic menus. Magpie is available for Internet Explorer and Mozilla/Firefox, and has been deployed in several commercial scenarios, the most recent one being semantic browsing support in the Food and Agriculture Organization of the United Nations.

See more: http://kmi.open.ac.uk/projects/magpie

Semantic Wikis: these can be of two sorts: either traditional wikis that are used to create collectively an ontology (in this case web pages denote concepts and relations) or semantic-enhanced wikis where web pages are annotated by metadata (more expressive than simple Web 2.0 tags), where semantic links are included in web pages and where reasoning is exploited to answer queries. For instance, SweetWiki [Buffa and Gandon, 2006] makes it possible to describe pages with concepts of an ontology. SweetWiki contains an ontology editor that enables users to manage the ontology by adding concepts, declaring subsumption relationships between concepts and assigning signatures to relations.

3.1.2. Ontology and Metadata Processing

Querying and reasoning: metadata can be saved in databases for efficient storage and access. For example, Oracle 10*g* has an RDF management platform that provides persistence and indexing for RDF graph data in a similar way to relational data.

SPARQL is one of the query languages designed for RDF. It enables the formulation of queries on an RDF database. SPARQL engines exploit the knowledge inside of RDF Schemas to compute the answers to the queries.
When ontologies are expressed in a more expressive language than RDF(S), a reasoner is used to draw the inferences needed to decide if one description is more specific than another one (i.e. if a resource description matches a query), or if one description is non contradictory (i.e. satisfiable in the reasoners jargon). Reasoners thus allow the infererence of new facts from knowledge (expressed in ontologies) and already known facts. Current technologies exploit highly optimised reasoning algorithms. For instance the reasoners Pellet, Racer and Fact support the full expressivity of OWL-DL, as well as reasoning on datatypes (integers, real numbers, strings…). Rule languages (like SWRL) are also partly supported, as their combination with OWL-DL makes the resulting language too expressive.

Aduna Metadata Server From: *Aduna Software*
Main Aim Provide a powerful and scalable (out-of-the-box) metadata extraction and indexing server that can be used by user-tools. **General Description** The Metadata Server is based on Sesame, an open source RDF-based storage framework. Techniques like crawling and parsing are used by the Metadata Server for metadata extraction. You can write your own applications that make use of the Aduna Metadata Server. The server is accessible with standard (Sesame) protocols.
See more: *http://www.aduna-software.com*

KAON2 From: *http://kaon2.semanticweb.org*
Main Aim Scalable reasoning with ontologies and rules. **General Description** KAON2 is a scalable reasoning tool for the Semantic Web, which enables practical reasoning with reasonably large ontologies. It is based on the type of description logic (SHIQ(D)) that provides the logical foundation of OWL. Furthermore KAON2 explores a completely new approach, based on the relationship between description logics and disjunctive datalog. More concretely, given a description logic knowledge base KB, our algorithms derive a disjunctive datalog program DD(KB) which entails the same set of ground consequences as KB. In this way, query answering in KB is reduced to query answering in DD(KB). KAON2 also supports the so-called DL-safe fragment of the Semantic Web Rule Language (SWRL). The DL SHIQ(D) and function-free rules are integrated as usual, by allowing concepts and roles to occur in rules as unary and binary predicates, respectively. It allows concepts and roles to occur in rule heads; but to achieve decidability, it requires that each variable in the rule to occur in a body literal with a predicate outside of the DL knowledge base. DL-safe rules provide means to circumvent certain expressivity drawbacks of OWL-DL without losing decidability of reasoning. KAON2 combines DL-safe rules by simply appending the rules to the program DD(KB).
See more: *http://kaon2.semanticweb.org*

Jena Semantic Web Platform From: *HP Labs*
Main Aim A comprehensive toolkit for RDF and OWL processing in Java. **General Description** Jena provides a core API for loading, storing, querying, processing and generating RDF, RDFS and OWL ontologies and instance data, and is the basis for a range of companion tools and technologies. Jena is widely used, and is actively supported by the developers on the jena-dev email list.
See more: *http://jena.sf.net*

Ontology matching and alignment: a matching tool finds an alignment, i.e. a set of correspondences, between entities of different ontologies. For instance, it tries to discover that "car" and "automobile" in two different ontologies denote equivalent concepts. Matching tools take advantage of various properties of ontologies, such as structures, data instances, semantics, or labels, and use techniques from different fields, such as statistics and data analysis, machine learning, automated reasoning, and linguistics [Euzenat and Shvaiko, 2007].

For instance, the OLA tool supports matching between two ontologies and further features like the visualization of correspondences and the formulation of bridge axioms between the ontologies [Euzenat and Valtchev, 2004]. S-Match takes as input two graph-like structures, e.g. classifications, XML schemas, and returns as output logic relations, e.g. equivalence, subsumption, which are supposed to hold between the nodes of the graphs [Giunchiglia et al., 2007].

3.1.3. Semantic Web Services

Enriching web services with semantic descriptions can help to automate many tasks concerning web services, like discovery, composition and invocation. Indeed the current standards for web service descriptions, like WSDL, enable only the description of the functionality of a service (what its inputs, outputs, preconditions, effects are) in a syntactic way: terms used in the descriptions are not linked to any ontology and the matching is made by comparing character strings and does not use any inference.

Semantics can be exploited for matching a service with a requester's need by using the inferences contained in an ontology and for translating some information provided as the output of a service into the input of another service by using ontology correspondences. In order to automate all of these tasks, the semantic descriptions concern various aspects of a web service, such as its function (which inputs it takes and which outputs it computes) and its process (which successions of messages and exchanges are allowed when interacting with the service).

Reasoning techniques can then be used to exploit the semantic descriptions for several tasks. For instance, discovery can be enhanced by finding the most appropriate service satisfying a set of requirements (for instance, a given concept as input and another concept as output). Composition can use the inputs and outputs to find a succession of web services, which provides required information given what is already known to the user.

Several standards have been proposed to semantically annotate web services. For example, OWL-S is an ontology expressed in OWL that can be used to annotate three aspects of a web service: a service profile, which describes the functionality that the web service provides (i.e. its preconditions, effects, inputs and outputs); a service model, which describes the sets and sequences of interactions that an end-user or an agent can use to interact with the services; and a service grounding, which describes how to concretely send messages to the service.

WSMO is another standard making it possible to associate semantic descriptions with web services by using ontologies. Furthermore, WSMO introduces the notions of goals and mediators. A goal corresponds to an end-user's need and mediators enable the resolution of mismatches between entities, like semantic, protocol or process mismatches. Typically, in this setting, an end-user's request will be transformed into a web service and a mediator will solve semantic mismatches among these web services, finding a solution.

WSMX – Web Service Execution Environment From: DERI International
Main Aim Execution Environment for Semantic Web Services. **General Description** WSMX is an execution environment which enables discovery, selection, mediation, and invocation of Semantic Web Services. WSMX is based on the conceptual model of WSMO, being at the same time a reference implementation of it. It is the scope of WSMX to provide a testbed for WSMO and to prove its viability as a mean to achieve dynamic interaperatability of Semantic Web Services. The components in WSMX are structured in three horizontal layers (Problem Solving Layer, Application Service Layer, Base Service Layer) augmented by vertical services (e.g. Security and Execution Management). Besides those functional components, WSMX offers features such as a plugging mechanism that allows the integration of various distributed components, an internal workflow engine capable of executing formal descriptions of the components behavior or a resource manager that enables the persistency of WSMO and non-WSMO data produced during run-time. Towards and based on WSMX studies, there is a standardization group in OASIS so called SEE-TC (Semantic Execution Environment) aiming at providing guidelines, justifications and implementation directions for an execution environment for Semantic Web Services.
See more: http://www.wsmx.org

3.2. The Knowledge Web Research Perspective

In order to support the research perspective on semantic technologies, a panel of researchers (selected from the Knowledge Web European network of excellence) was asked to locate some Semantic Web-related topics on a Gartner-like hype cycle curve (see Figure 4), similarly to the elaboration of the business perspective. In this section, the results of this collective work are analysed and some trends on the evolution of semantic technologies are sketched from the Knowledge Web network of excellence viewpoint.

Contrary to the business perspective (depicted by Figure 3), which is drawn by practitioners developing real applications, the research perspective reflects the technical point of view of people who settle generic tools for building Semantic Web applications. At first glance, one notices that topics on this researchers' curve are evenly distributed from the technology trigger phase (on the left) to the plateau of productivity phase (on the right). Topics on the previous practitioners' curve are mostly positioned around the first two (and least mature) phases. Essentially, this means that the research community considers that developments from the past ten years have resulted in some tools (see previous section) and standards, which are reliable and mature enough to be transferred to industry and successfully integrated into Semantic Web applications. The developers' community however is not yet fully aware of the availability of such tools, which, consequently, has to be promoted further, together with the innovative functionalities they can provide to software applications. Today, the semantic technologies that have been developed in research have been mastered and deployed mainly by early adopters and specialised start-ups, both in the USA and in Europe. This trend is confirmed by the growing number of such SMEs represented at the last semantic technology conferences, respectively SemTech 2007 and ESTC 2007.

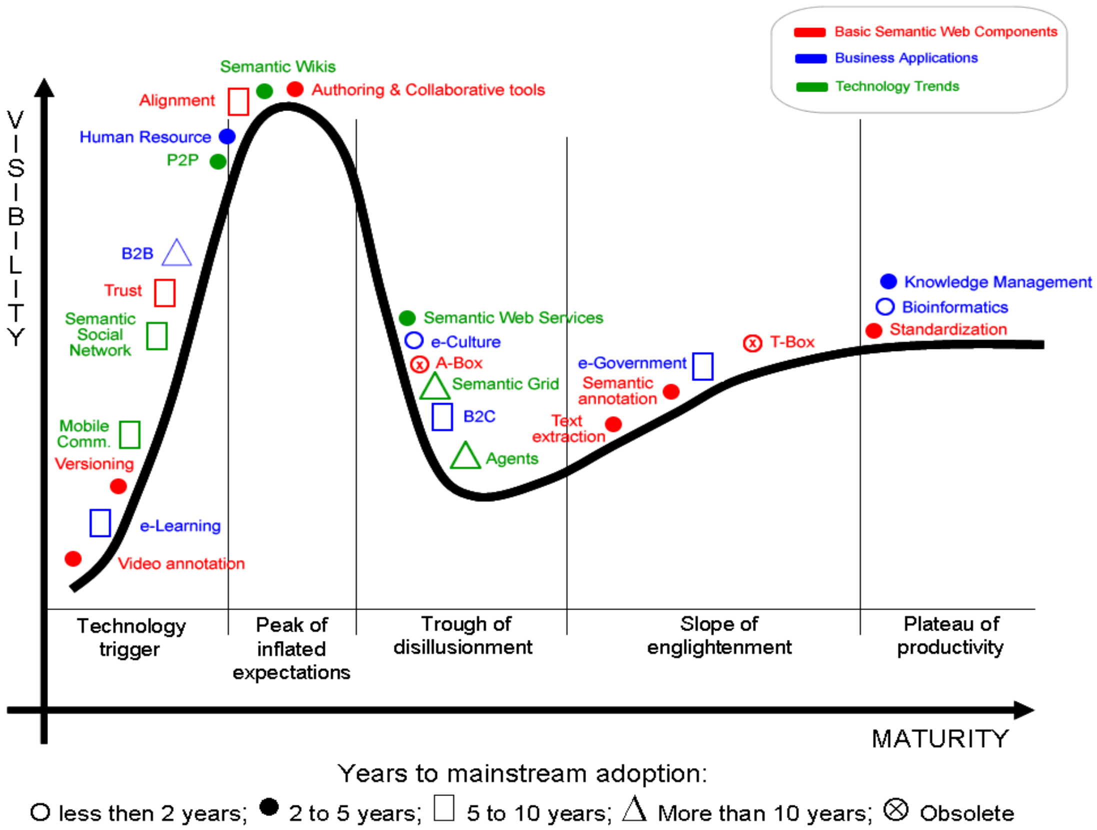

Figure 4. Hype Cycle Curve of Semantic Web-related Topics from the Researchers' Point of View

In addition to this first analysis, the previous curve, drawn by researchers, have been compared to some of the reference hype cycle curves produced by Gartner in 2006 [Gartner, 2006b, 2006c, 2006d, 2006e, 2006f, 2006g, 2006h]. This comparison has been done on topics that are closest to the ones considered by Knowledge Web. Consequently, Knowledge Web's topics that are too much specific to the Semantic Web domain, and so cannot be easily compared to any Gartner's counterpart, are not taken into account. Figure 5 presents the results. In Figure 5, Gartner's reference topics appear in orange. Each of them is linked thematically to a corresponding Knowledge Web's topic by a thick line. For example, the Knowledge Web's "B2B" topic is compared to three of Gartner's topics, namely "auto-trading grid", "B2B gateway software" and "B2B web services".

The link between two topics is light gray when their duration estimate for mainstream adoption (measured in years) is similar. It is dark gray when the duration estimated by Knowledge Web researchers is longer than the one estimated by Gartner. For example, there is a light link between the Knowledge Web's "B2B" and Gartner's "auto-trading grid" topics because both are estimated to take more than 10 years before mainstream adoption (represented by two triangle symbols). The links with the other two Gartner's "B2B gateway software" and "B2B web services" topics are dark because Gartner's estimation is shorter-term than the Knowledge Web's one (2 to 5 years, represented by two filled circle symbols).

In general, most of Gartner's topics currently involve no semantics at all ("B2B gateway software", "B2B web services", "social network analysis", "web services trust", "advanced web services", "grid computing"). Unsurprisingly, these topics are estimated equal to or more mature than their Knowledge Web counterparts (that is, they are located more on the right of the curve) with a smaller duration to mainstream adoption. This means than turning these technologies into semantically-enabled ones needs some more time, in average, approximately 5 to 10 more years. For example, Gartner's "B2B gateway software" and "B2B web services" topics are considered to be sliding into the trough with a duration to mainstream adoption of 2 to 5 years. The "B2B" topic, interpreted by the Knowledge Web community, however is only on the rise of the technology trigger with a duration to mainstream adoption longer than 10 years. When considering topics involving intrinsically (and often implicitly) more semantic capabilities (auto-trading grid, information extraction, intelligent agents, KM), the difference between Gartner's and the Knowledge Web's analysis is far smaller. For example, the "auto-trading grid" topic, which covers interoperability and integration between the business processes and sensor networks of heterogeneous partners, is located in the same phase of the curve (technology trigger) with the same duration to mainstream adoption (more than 10 years).

Interestingly, Gartner's "advanced web services" topic is estimated quite similar (sliding into the trough, with 2 to 5 years to mainstream adoption) to the Knowledge Web's "Semantic Web Services" topic, although Gartner's topic includes BPEL and security aspects but no semantics. This may mean that, specifically for the web service technology, semantics-related extensions appear to be at a comparable level of maturity and adoption than other non semantic extensions. This indication is quite promising for the Semantic Web community and reinforces the idea that semantic technologies will first penetrate the market of enterprise applications before the one of end-user applications (see the comments in section 3.3). However, this enthusiasm should also be tempered by the more noticeable difference between Gartner's "web services trust" topic and the Knowledge Web's "trust" one. Both are on the rise of the technology trigger, but the Knowledge Web's topic has a slightly longer duration to mainstream adoption (5 to 10 years) than Gartner's one (2 to 5 years). This shows that semantics incorporation into the trust level might be less straightforward.

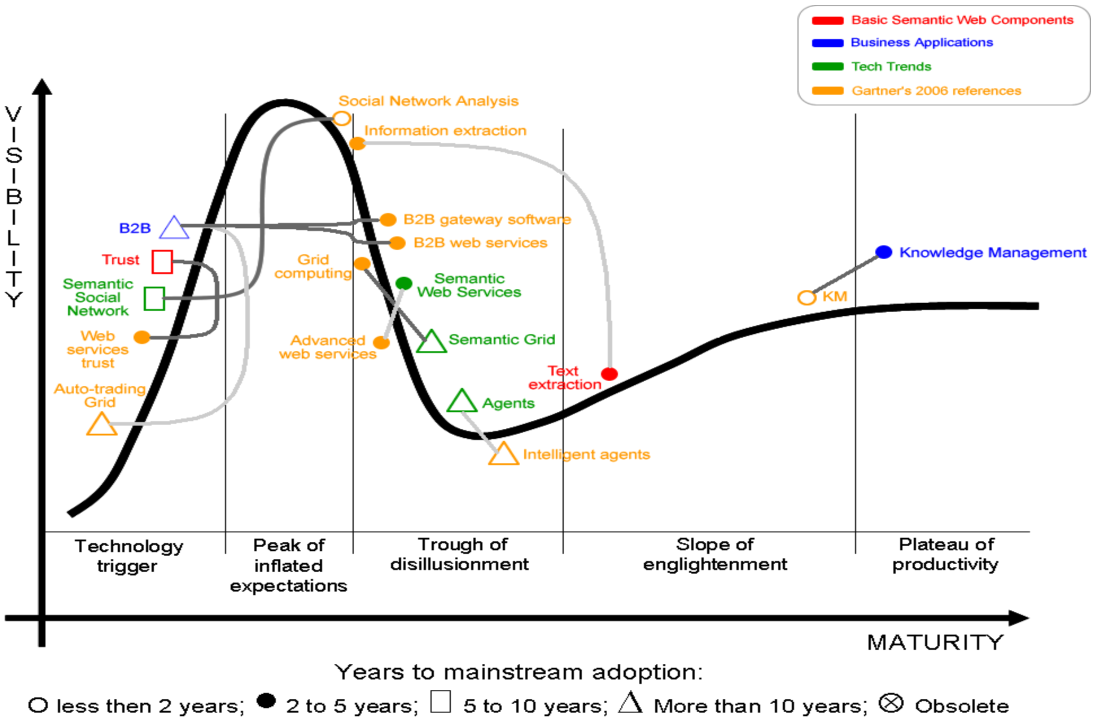

Figure 5. Comparison between the Knowledge Web Project Perspective and Gartner's References

3.3. Research Trends in Semantic Technologies

Although the ultimate vision of the Semantic Web has not yet been realised and is likely to take several more years before being so, some significant steps have already been achieved. On the standardization side, some major standards have been completed, in particular RDF (Resource Description Framework) and OWL (Web Ontology Language). On the technology side, some commercial, as well as open source tools, are available to efficiently implement these standards (see section 3.1 for an overview of the main ones). This current state of affairs already provides satisfying tools and products to deal with bound, *closed*, application domains, where the content of knowledge bases and their conceptualization are controlled and mature[10], such as *intra*-enterprise knowledge management, content management, and e-commerce applications. A number of specialised start-ups are investigating this market by providing products and services to enhance management of enterprise data and metadata (e.g. at the European level, Empolis, Hanival Internet Services, iSOCO, metatomix, Ontoprise, Ontos, Smart Information Systems[11]).

---

[10] "It is in the comparatively narrow context of enterprise and industry vocabularies, taxonomies, classification schemes and ontologies, as opposed to the World Wide Web at large, that we recommend enterprises pursue value from the full Semantic Web" [Gartner Report, 2007a, section 3.0], see also [Gartner Report, 2007a, figure 4].

[11] Alphabetical list of the Small or Medium Enterprises, which has supported the 1st European Semantic Technology Conference (ESTC2007, 2007, Austria. http://www.estc2007.com).

Some practitioners consider these first steps of semantic technologies (and in particular the use of the very basic Semantic Web standards) within industry already flourishing with a high potential for business, still to explore or to invent[12]. Nevertheless, the research community works all the same on future extensions of the related standards and technologies. The main challenge is currently to make possible and efficient the move to *open* domains, typically real web applications (that would implement the grand vision of the Semantic Web) and *inter*-enterprise applications. The underlying research roadmap could be summarised by the following items (mainly inspired by the invited talks given at *ESTC2007*[13]):

Ontology Construction. Large scale applications need significant amounts of data to be described and formalised, typically in the form of ontologies in the Semantic Web vision. In realistic environments (ranging from closed intra-enterprise applications to open web applications), it is very costly, if not impossible, to build them manually: the amount and variety of data and contents are quite huge, data and contents are continually evolving, knowledge experts, who are the people able to produce suitable conceptualisations, are rare (and expensive) and cannot have a global understanding of all possible topics, etc. Researchers have to therefore setup processes to generate adequate semantic representations of existing data (to be used by semantically enabled applications). When such processes are targeted to be automatic (or semi-automatic), they are called "*ontology construction*"[14]. In this case, most approaches are based on text-mining and natural language processing techniques, which extract information from textual resources (web pages, documents, textual annotations, tags, metadata, etc.). There are major challenges to reach an industrial level, such as the handling of non textual resources (in particular multimedia resources like pictures, sounds or video), the scalability of techniques (to be useful, such techniques must process huge amounts of data) and the improvement of result quality (how to evaluate the results and ensure their relevance). Researchers are also thinking of more collaborative processes to build ontologies, being explicitly inspired by the Web 2.0 trend and taking implicitly into account the people component of the holistic approach described in section 2.2.1. For example, one recent and promising way is the use of *semantic wikis*[15], which support end-user generated ontologies in the same way as traditional wikis support end-user generated web pages.

Matching. Practical applications generally involve several data sources. It is quite an illusion to expect a universally shared and agreed format or

[12] "RDF [and the other rudimentary semantic technologies] solve meaningful problems, and it costs less than any other approach would. The entire remainder"–the more ambitious work with ontologies and artificial intelligence–"is completely academic", words of Dirk-Willem van Guilk, Joost chief technology officer, reported in [Borland 2007].
[13] See *http://www.estc2007.com*
[14] See *http://ontoworld.org/wiki/Category:Topic_ontology_construction*
[15] See *http://ontoworld.org/wiki/Category:Topic_semantic_wikis*

conceptualization for all sources, even in closed domains such as intra-enterprise applications. Hence a strategic challenge for semantic technologies to be used and useful in real environments is the ability to integrate (researchers would say "to match" or "to align") the structures and contents of data sources, in order to provide consistent relationships between identical elements of heterogeneous sources[16]. Some techniques have emerged and matured from academia, which are ready to be transferred to industry for current data integration needs (see for example [Bidault et al., 2007]). Most of these approaches are characterised by an offline computation of the correspondences between data sources, which are then used by applications. Researchers in the *ontology matching*[17] field are now focusing on new mechanisms that make dynamic discovery of correspondences possible. Such mechanisms are required by modern flexible middleware, such as peer-to-peer infrastructures, multiagent systems and web service platforms, where data sources (managed by peers, agents or web services) may appear, disappear or change at run-time. Another key issue for matching algorithms is to allow for multilingual data sources, where concepts, relations and individuals are described using different languages (for example, English, Italian, French).

Approximation. By definition, it is not possible to control all the entities taking part in an open environment, such as the World Wide Web. Consequently, there is no means to ensure a priori the quality, the completeness or the reliability of available information. Moreover, some pieces of information are naturally uncertain or approximate, for example the weather forecast or the temperature value, which is subject to sensor accuracy. Currently, human users rely on their natural capability of discernment in their usage of online documents, and this is even more crucial in the context of the user-contributed Web 2.0. The same kind of problem arises in a semantic world. However, as it is intended to be machine-understandable (and not just user-understandable as the current – non semantic – web), researchers have to devise proper mechanisms to automate human discernment. There are at least two issues to deal with. How to determine the reliability of a given piece of information and how to take into account the reliability of information while processing them (i.e. during the reasoning tasks). The first question includes various situations: the reliability of a piece of information may lie on the accuracy of a sensor, series of observations, the reputation of the information source, reasoning about objective facts, etc. The second question relates to finding appropriate rules to derive consequences from a set of approximate information items. One possible way currently studied in the context of the Semantic Web is that of *fuzzy reasoning*[18].

---

[16] From a business point of view, "Between 35 and 65% of the 300 billion dollars being spent per year on systems integration is attributable to resolving semantic mismatches between systems" [McComb 2005].

[17] See http://ontoworld.org/wiki/Category:Topic_ontology_mapping

[18] See http://ontoworld.org/wiki/Category:Topic_fuzzy_reasoning

Distribution. Today's open worlds are supported by distributed software architectures, such as the World Wide Web itself, Service Oriented Architectures (SOA), peer-to-peer infrastructures (P2P) and the Grid. Pushing semantics into such environments requires turning related algorithms in a distributed fashion. This involves both handling the knowledge bases (i.e. the semantic descriptions of information) and performing computations on knowledge (e.g. search, matching or *rule processing*[19] operations) in a decentralised way. Such constraints raise many problems to solve (see for example *P2P and Semantic Web*[20], and *semantic grid*[21]): how can knowledge pieces be optimally distributed in the nodes of a network? How can nodes be managed that hold related pieces of knowledge (such nodes may be totally or partially redundant, use different vocabularies to describe the same knowledge, be inconsistent, etc.)? How can reasoning algorithms be executed in a decentralised way on these nodes? How can a global consistency be ensured when knowledge and/or algorithms are distributed over nodes that evolve over time?

Semantic Web Services. In the *e-Business*[22] context, the dematerialization of business exchanges rests on the publication of an Application Programming Interface (API) that enables applications of external partners to access and invoke operations on the internal legacy system (e.g. to order goods, to notify the receipt of goods, to invoice). Currently, such APIs are implemented through the concept of services and deployed over Service Oriented Architectures (SOA). One promising way to promote business exchanges on a large scale is to make such services available on a wide spread network, such as the World Wide Web. The resulting standards, known as Web Services, are more and more widely used in B2B and B2C applications. For example, major services on the internet, such as the Google search engine, the Amazon electronic library and the eBay online market place, publish their APIs as web services. Semantic technologies may enhance usual web services, by turning them into *semantic Web services*[23], in the same way they enhance the usual Web by turning it into a semantic Web. Basically, the service descriptions are enriched with semantic relationships that specify in fine detail which types of input and output data are expected and returned, and how they relate to each other. Therefore, applications are able to automatically discover and suitably combine available services to fulfil their needs, without explicitly programming the services to use at design time. The major stake here is to find mechanisms that make business processes of business partners match automatically and more flexibly than current B2B approaches based on Web EDI or ebXML.

---

[19] See http://ontoworld.org/wiki/Category:Topic_distributed_rule_processing
[20] See http://ontoworld.org/wiki/Category:Topic_peer-to-peer_and_Semantic_Web
[21] See http://ontoworld.org/wiki/Category:Topic_Semantic_Grid
[22] See http://ontoworld.org/wiki/Category:Topic_ebusiness
[23] See http://ontoworld.org/wiki/Category:Topic_Semantic_Web_services

4. Final Remarks and Conclusions

4.1. Analysis of the Knowledge Web's Hype Curves

In this section, some conclusions on the future challenges of semantic technologies are drawn from the analysis of the differences between the hype curve issuing from Knowledge Web researchers (see Figure 4 in the section 3.2), which reflects a *technical* point of view, and its counterpart issuing from practitioners (see Figure 3 in the section 2.3), which reflects a *business* point of view. In oder to get a reliable comparison, only the topics that were evaluated by both researchers and practitioners (i.e. those appearing on both curves) are considered.

The results of this comparison are synthesised in Figure 6. The blue symbols, which are labelled with topic names, represent the estimations of the researchers (taken from Figure 4.). The corresponding estimation of the practitioners is represented by the linked red symbol, without label. For example, researchers locate the "B2B" topic on the rise of the technology trigger (see the corresponding blue symbol on the left of the figure), while practitioners locate it at the peak of inflated expectations (see the corresponding red symbol under the peak, which is linked to the previous one).

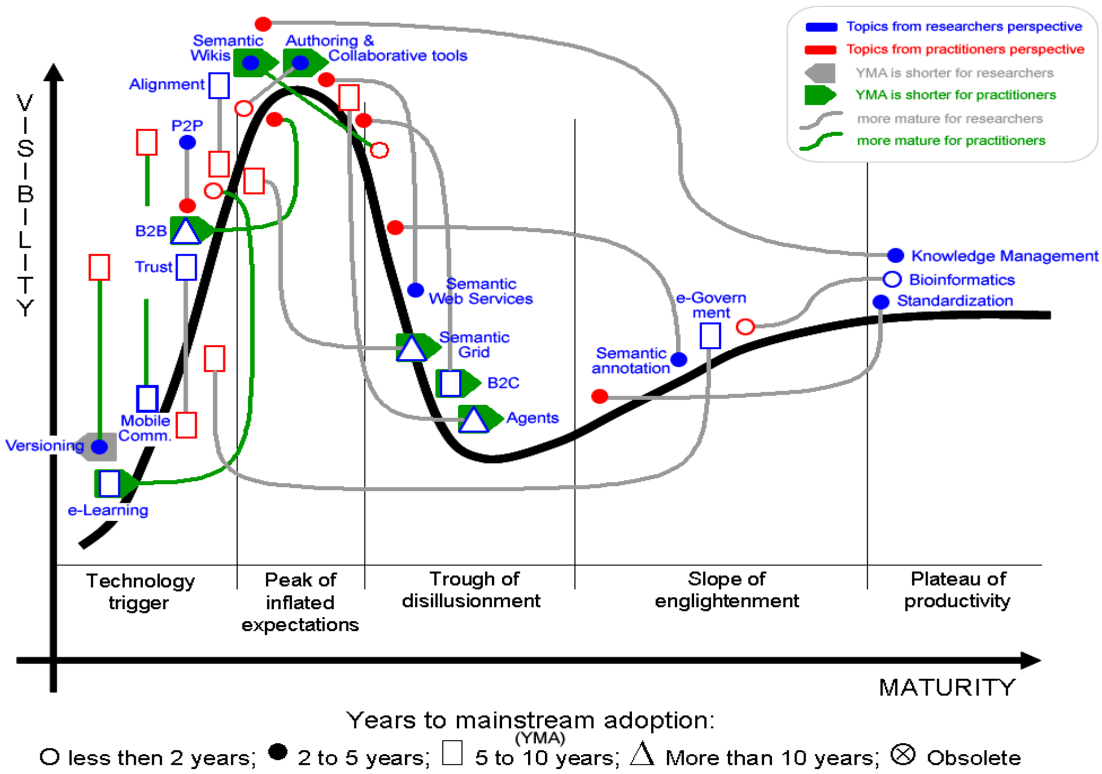

Figure 6. Comparison between the researchers' and the practitioners' views

A grey (respectively green) link means that the topic is more mature from the researchers' (resp. practitioners') viewpoint. This means that the symbol located by practitioners (resp. researchers) is earlier on the curve. For example, the link between the symbols associated with the "B2B" topic is grey because researchers estimate that this topic is more mature than practitioners consider. That is, the researchers' blue symbol is located earlier on the curve than the corresponding practitioners' red one. On the opposite, the link between the symbols associated with the "semantic wikis" topic is green, which means that practitioners consider this topic more mature (see the corresponding red symbol beginning to slide into the trough of disillusionment) than researchers do (see the corresponding blue symbol just before the peak of inflated expectation).
A grey left (respectively green right) arrow in the background of a labelled symbol means the duration to mainstream adoption of the corresponding topic is estimated shorter by researchers (resp. practitioners). Topics with the same duration to mainstream adoption have no arrow in their background. For example, the duration to mainstream adoption of the "B2B" topic is shorter according to practitioners (2 to 5 years, see the corresponding red filled circle) than according to researchers (more than 10 years, see the corresponding blue triangle), hence a green right arrow in the background of the blue triangle. Opposite, there is a grey left arrow in the background of the blue filled circle associated to the "versioning" topic because its duration to mainstream adoption is estimated shorter by researchers (2 to 5 years) than by practitioners (5 to 10 years).

Unsurprisingly, two general remarks emerge:

- Researchers estimate the various Semantic Web-related topics (except semantic wikis and, to some extent, e-learning) equally or more mature than practitioners. In other words, they locate the corresponding symbols more on the right part of the curve. This is represented on the figure by more grey links than green links. This may be interpreted by the fact that researchers are rarely directly involved in operationally deployed applications, so they know the maturity of "beta-prototypes", rather than the maturity of real applications and they consider technical maturity over business maturity. Considering that beta-prototypes are mature however also means trusting in a reliable integration into real applications. The difference in opinion between researchers and practitioners may result from the practitioners' lack of awareness of the technical maturity of the tools produced by researchers.
- Researchers estimate the duration to mainstream adoption of the various Semantic Web-related topics (except versioning), especially the basic Semantic Web components and the technical trends, equal or longer than practitioners. This is shown on the figure by more green right arrows than grey left ones. This may be explained by the fact that researchers have a finer awareness of technical difficulties to overcome

and that they usually target a "theoretically perfect" solution – necessarily longer to achieve – for mainstream adoption, whereas practitioners, once convinced by a technology, tend to put it on the market as quickly as possible, even in a simplified and imperfect – but nonetheless beneficial – fashion.

Interestingly, when considering the less technical standardization and business application topics, researchers and practitioners share close points of view, which seems to be quite promising in the perspective of increasing the market penetration of semantic technologies. Whereas practitioners consider standardization in Semantic Web technologies not completely mature (on the slope of enlightenment), as previously noted in remark #1, both researchers and practitioners estimate at 2 to 5 years the duration to mainstream adoption. This shows that the efforts of the W3C in the Semantic Web area are coming to fruition, despite less than ten years of work.
The duration to mainstream adoption for most business applications are appreciated equally by both researchers and practitioners, researchers being generally more optimistic on their maturity (see remark #1). The two noteworthy exceptions are the B2B & B2C domains and the e-learning domain. The former is analysed more deeply in the next paragraphs. As for the latter, researchers do not see e-learning applications on the market before 5 to 10 years, whereas practitioners expect them in less than 2 years. As a reference, in the last Gartner hype cycle on e-learning [Gartner, 2006a], a similar topic, such as e-learning suites, is estimated to be in-between 2 to 5 years to mainstream adoption, although this study does not explicitly lie within the context of semantic technologies. Consequently, it seems that practitioners gave their estimation rather for general e-learning applications than for the enhancement of e-learning applications with semantic technologies.

To provide a relevant analysis of the Knowledge Web researchers' point of view, only the technical topics (mainly among the basic Semantic Web components and the technical trends categories), which are directly related to the business cases promoted in the industry area of the project (work package 1.1), are taken into account. Indeed, researchers have been working on real case studies in cooperation with members of the Industry Board of the project. Therefore, the location of the corresponding symbols on the hype cycle curve should actually result from an accurate understanding of the technological needs of concrete applications, and should not reflect only theoretical ideas.

- *Multimedia Analysis & Annotation (Business Case #6):* this business case is mainly concerned with "semantic annotation", "text extraction" and "video annotation" topics. Only the first topic (semantic annotation) was evaluated by both researchers and practitioners, who equally envisage a duration of 2 to 5 years before mainstream adoption, researchers considering it a bit more mature than practitioners (in conformance with remark #1). Researchers estimate the same duration to mainstream adoption for the last two topics, "video annotation" being unsurprisingly emerging and "text extraction" being climbing the slope of enlightenment. As a conclusion, researchers and practitioners seem rather well synchronised on this application domain and products should appear on the market in the medium term.

- *Ontology versioning and lifecycle in Life Science scenarios (Business Case #3):* this business case is mainly concerned with "versioning", "semantic wikis" and "authoring and collaborative tools" topics. Although both researchers and practitioners locate the first topic on the rise of the technology trigger, researchers unusually consider it as a shorter term technology (2 to 5 years to mainstream adoption) than pratitioners (5 to 10 years). This discrepancy may reveal a potential misunderstanding between researchers and practitioners. At least researchers need to better promote and make accessible the corresponding technology to the practitioners. As for the last two topics, researchers and practitioners seem to see things in quite similar ways, both locating them around the peak of inflated expectations, the researchers considering them a little longer term (2 to 5 years) than the practitioners (less than 2 years), in conformance with remark #1. More generally, the Life Science domain is a pioneer in adopting semantic technologies and has been used to experimenting with them for some years. This is confirmed by the location of the "bioinformatics" topic that is very close to maturity, from both researchers' and practitioners' perspectives. Probably, this dynamic should continue and broaden into further emerging application domains.
- *Ontology alignment (Business Case #2):* this business case is trivially concerned with the "alignment" topic. Impressively, both researchers and practitioners agree on putting it just before the peak of inflated expectation, with the same long term duration to mainstream adoption (5 to 10 years). This is an example of quite a perfect synchronization between research and industry. It is expected that, as this technology is becoming more mature, practitioners will shortly increase their expectations and want to experiment with it in realistic conditions.
- *Enterprise Application Integration (Business Case #4):* this business case is mainly concerned with "alignment" and "B2B" topics. The enterprise domain, which provides a structured and closed world of data, is certainly the most promising one from which this technology can take off. The second topic (as well as the related "B2C" topic) reveals an unusual discrepancy in the duration to mainstream adoption between researchers and practitioners. Although this difference conforms to the introductory remark #2, it turns out to be quite significant (very long term – more than 10 years – against medium term – 2 to 5 years) and suggests that dialogue be improved between the research and industry on this topic. This discrepancy might be caused by both:

- researchers' vision: on one side they are very familiar with the Semantic Web and Web 2.0 applications and have invested a lot of resources in developing solutions which might be used in the real world. On the other side they don't foresee how Semantic Web technologies might improve business processes.
- companies' perspective: firms have a profound knowledge of business and B2C but do not know how the Semantic Web might be used to innovate B2C services. Therefore they have long term expectations on semantic technologies, also because the enterprise domain is mainly based on traditional marketing studies and strategies (such as [Gartner Report, 2007a]).

4.2. Research and Industry Roadmap – 2007 Take Away

In addition to the analysis carried out by Knowledge Web, based on various internal and public available studies, as reported up to now in this document, two prominent efforts have been undertaken in 2006-2007. One is a position paper carried out by a core team from the ICSOC Research community (Service-Oriented Computing) [Papazoglou et al., 2006]. The other is a survey conducted with a panel of selected experts from industry and academy in the field of Semantic Web Services [Bachlechner, 2007]. This section is tentatively trying to sum-up the main vision forged so far on these diverse and difficult fields of Semantic Web and Services technologies.

SWOT Analysis[24] is a strategic planning tool used to evaluate the Strengths, Weaknesses, Opportunities, and Threats involved in a project or in a business venture. It involves specifying the objective of the business venture or project and identifying the internal and external factors that are favourable and unfavourable to achieving that objective. Table 1 provides such an analysis for the Semantic Web Services technology.

INTERNAL	Strengths	Weaknesses
	• Improved service discovery capability • Facilitated interoperability • Facilitated reuse of services • Improved automated mediation between services • Explicit and formal definitions of conditions and functionalities	• Use of immature technologies • Description overhead • High initial learning costs • Labor-intensive service specification • Software engineers are not ontology, Knowledge Representation and Logic reasoning experts

[24] See http://en.wikipedia.org/wiki/SWOT_Analysis

EXTERNAL	Opportunities	Threats
	• Availability of business cases • Need for service interoperability • Availability of compliant middleware implementations • Availability of effective design and operational management tools • Availability of very efficient semantic technologies • Agility and Time to market effectiveness • Preceding agreement on standards	• Difficulty of describing semantics • Difficulty to master ontology evolutions • Lack of effective design tools • Lack of optimised KR technology for business scale use cases • Unclear benefits and cost effectiveness • Limited consideration of business needs and interest • Unavailability of convincing case studies

Table 1. SWOT analysis for Semantic Web Services

In addition, as often used in the Gartner's focused technology reports, a similar representation is used to summarise the Knowledge Web's vision of the key ordered factors that certainly will speed-up the take off of Semantic Web Services technology. See the next two tables.

	2 - 5 years	5 - 10 years
KEY	• Convincing pilot Use Cases • Availability of skilled practitioners • Industrial and Business key needs • Convincing Business ROI • Key Standards needs (W3C, OASIS)	• Scale-up to large business cases • Technology transparency to the final users • Mastering the fast-growing complexity • Mastering the IT semantic heterogeneity
2	• KR languages well adapted to industry and business needs • Industry grade design and deployment tools	• Robustness in very dynamic environment • Provision of trust and reputation mechanisms • KR and SWS Tool box • Open and public benchmark frameworks
3	• Mastering the Ontology evolutions • Open and compliant frameworks	

Table 2. Macro Roadmap for Semantic Web Services in a priority matrix

	State of Art	Grand Challenges
Service Foundations	Enterprise Service Bus • Open standard message backbone (SOAP) • Current industry ESB-SOA (*BEA*, *IBM*, *SUN*, *Microsoft*, *SAP*) • Semantic Execution environment OASIS[25] Ontology matching[26] SOA Reference Model OASIS[27]	Dynamically (re)configurable run-time architecture Dynamic connectivity capabilities Topic- and content-based routing capabilities Directory facilities Mediation facilities for data, application and process integration
Services facilities	Semantic integration (EAI ++): few deployed solutions on small scales (*Semagix*, *Ontoprise-Software AG*, *Contivo*, *WebMethods*/ *Cerebra*, …) Service discovery at the syntax matching level (UDDI): many solutions in research labs Service composition still ad-hoc and manual (EAI style): many solutions in research labs Services orchestration manually generated (BPEL) Semantic Web Services pilot platform[28]	Automated and accurate Service discovery Non-functional parameters (QoS) aware composition Business driven (SLA) composition Automated orchestration generation End-to-end security and transaction solutions KRR grand challenges: scale-up to large ontologies and annotations of real life needs; reasoning under uncertainty and inconsistency; finalization of KR, Query and process languages; mastery of expressivity *vs.* complexity for industry; matching and reasoning in P2P; Ontology learning and maintenance
Service management	Management of Web Services Web Service-based management Web service distributed management	Self-configuring services Self-healing services Self-optimising services Self-protecting services
Service Engineering	Port existing components using wrappers Component and OO analysis and design Tools include basic Web Service descriptions (WSDL) (*BEA*, *IBM*, *SUN*, *Microsoft*, *SAP*) Few tools include advance Web Service features (OWLS, WSML, SAWSDL)	CAD tools for semantic service engineering • Associate standard software development with business process modelling techniques • Automate service description • Make complexity very simple to use for practitioners

Table 3. Knowledge Web Industry's Vision on the State of the Art and Grand Challenges[29] in Semantic Web Service research[30]

[25] Developing guidelines, justifications and implementation directions for deploying Semantic Web Services in SOA: http://www.oasis-open.org/committees/semantic-ex

[26] See the Knowledge Web Ontology Alignment Evaluation Initiative (OAEI, http://oaei.ontologymatching.org) and [Euzenat and Shvaiko, 2007].

[27] See http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm

[28] See the Knowledge Web Dynamic Business to Business Integration Use Case (DERI, Bell Labs).

[29] This an adaptation of [Papazoglou et al. 2006].

[30] On the research side, see more details from the Knowledge Web deliverables in workpackages 2.1, 2.2, 2.3, 2.4 and 2.5. On the industry side, see the Knowledge Web Industry executive summaries in workpackages 1.1 (STC 06 and ESTC 07 main outcomes), 1.2 (technology benchmarks) and 1.3 (ontology re-use and recommendations in OAA) for landmarks.

4.3. Conclusion

From the comparison between the researchers' and practitioners' perspectives on Semantic Web technologies, the main lesson to be learned is that practitioners seem not to be fully aware of the maturity of the semantically-enabled tools produced by researchers. For example, noticeable discrepancies between researchers and practitioners have been observed in ontology versioning technologies, as well as in applications for the enterprise domain, in which industry expresses strong expectations. Therefore, the effort in outreaching concrete and practical results to industry must be sustained.

Meanwhile, initiatives such as the Knowledge Web project, which foster technology and knowledge transfers between research and industry, are definitely bearing fruits. In particular, both researchers and practitioners agree on the crucial topic of semantic technologies standardization, trusting in a middle term stabilization and maturity of useful standards for industry (2 to 5 years). Similar convergences have been identified in the short and mid-term adoption of semantic technologies in the domains of Multimedia and Life Science applications.

From a business viewpoint, if some key challenges remain to be overcome, it seems very plausible that there will be an increasing demand for Web Services and integration technologies. Indeed, businesses react very positively to the need for a very effective integration technology and for more agility in a very competitive worldwide economy. In the meantime, reducing interoperability problems will open opportunities for easier innovative solutions and for the increased cooperation between enterprises. This should result in re-combinations of the business activities supported by the technology and so in a profound impact on business and economic workflows (5-10 years). From a practitioner perspective, the technology must still mature to the point where it is as easy to make analysis, design and deployment at a higher and broader level – namely business – as it is today to develop applications in Java (5-10 years).

While the cooperation between industry and research has led to a good understanding of industry and professional user needs in a corporate setting, the effort to understand the needs of layman users and non-professional practitioners seems not to have reached a similar level. A lot of questions are still largely unanswered. Which usages of semantic technologies may the typical Web 2.0 users adopt? Which level of semantic language expressivity are they prone to use for building metadata and ontologies? In other words, market research, concerned with industrial needs, has been conducted, but mass market research, concerned with end-user needs, seems to be in a less advanced state. In fact, several visions of the realisation of the foreseen Semantic Web exist, two being particularly notable:

- A Semantic Web with strong semantic languages, making available high expressivity and inference capabilities to professional users who have the time and competencies to master these technologies. These languages will be used mainly in corporate settings or will be parts of services provided to non-professional end-users without the need to understand how they work nor to be aware of the use of such technologies.
- A Semantic Web with very light languages, which all users (even inexperienced ones) will use to describe resources. It can be seen as a continuation of the current usage of Web 2.0 tags and folksnomies. A little more effort from the end-users can lead to taxonomies and more elaborated facts (e.g. including equality statements). Due to the huge size of the Web, these semantics, which may seem at first hand to be too simple to be of interest, can actually represent a very large body of knowledge, and so be of tremendous value for anyone who is able to exploit it [Hendler, 2007].

These two visions are of course not at all contradictory and can coexist in Web 3.0. The path to the realization of the first vision seems to be well defined. Continuing the efforts to make industry and research collaborate will probably lead in a few years to industry ready solutions and products. This process is likely to be supported by firms themselves, non-for-profit organizations, and consortiums, which aim to develop strategies for Semantic Technology recommendation and standardization, and promote the Semantic Technology to industry[31].

As for the second vision (a semantic web where users are semantic contributors), things do not seem to have advanced as much. Therefore one major goal of the research should be to increase the expressivity of the user contributions and conceive methods to exploit a huge quantity of simple semantic add-ons. The success of turning users into semantic contributors relies on many factors, ranging from ergonomics to the perceived added-value for the user. In particular "webifying" the technology and integrating semantic technologies with the main current web technologies (widgets, social networks, blogs, etc.) may maximise the probability of appropriation of the semantic technologies by the users in their current usages and attitudes [Hendler 2007]. This vision for layman users could be realised step by step, by maturing technological solutions and services offering gradual innovative steps (2-10 years).

[31] For instance, the Ontology Outreach Advisory (OOA) is oriented towards the ontology technology. It is an international not-for-profit association that consists of industry, government, and research leaders and innovators who are promoters of ontology development, use, or education (http://www.ontology-advisory.org).

References

[Assfalg et al., 2003] Assfalg, J., Bertini, M., Colombo, C., Del Bimbo, A., Nunziati, W., Semantic annotation of soccer videos: automatic highlights identification. *Computer Vision and Image Understanding. Special Issue on Video Retrieval and Summarization.* Vol. 92, Issues 2-3, November-December, 2003.

[Bachlechner D., 2007], Relevance and Applicability of Semantice Web Services in E-Business – Executive Summary, Technical Report, University of Innsbruck, September 2007.

[Berners-Lee et al., 2001] Berners-Lee, T., Hendler, J., Lassila, O., *The Semantic Web*, Scientific Am., pp. 34–43, May 2001.

[Bidault et al., 2007] Bidault, A., Paulus, F., *Heterogeneous and Distributed Data Integration: The SWID Projet*, 1st European Semantic Technolgoy Conference (ESTC2007), Vienna, Austria, May 31st – June 1st 2007.

[Brown and Duguid, 1991] Brown, J.S., Duguid, P., *Organizational Learning and Communities-of-Practices: Toward a Unfied View of Working, Learning and Innovation*. Organization Science, 2. 1991.

[Buitelaar, 2005] Buitelaar, P., Cimiano, P., Magnini B., *Ontology Learning from Text: An Overview.*. In Paul Buitelaar, P., Cimiano, P., Magnini B. (Eds.), Ontology Learning from Text: Methods, Applications and Evaluation, pp. 3-12, IOS Press, 2005.

[Cimiano and Völker, 2005] Cimiano, P., Völker, J., *Text2Onto - A Framework for Ontology Learning and Data-driven Change Discovery*, in Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems, 2005.

[Cuel, Bouquet, Bonifacio, 2005] Cuel, R., Bouquet, P., Bonifacio, M. S., A distributed approach to knowledge management, the concept of knowledge nodes, and their implications. Cap. 16:Schwartz D. by Encyclopedia of knowledge management, Hershey, Pa: Idea group, 2005.

[Domingue, 2004] Domingue, J.B., Dzbor, M., Motta, E., *Collaborative Semantic Web Browsing with Magpie*, In Proc. of the 1st European Semantic Web Symposium (ESWS) , Greece, May 2004.

[EIF, 2005] European Investment Fund, report Technology Transfer Accelerator (TTA), September 2005.

[Euzenat and Shvaiko, 2007] Euzenat, J., Shviako, P., *Ontology matching*. Springer-Verlag, Heidelberg (Germany), 2007, ISBN: 978-3-540-49611-3.

[Euzenat and Valtchev, 2004] Euzenat, J. and Valtchev, P., Similarity-based ontology alignment in OWL-lite, in *Proceedings of the 15th European Conference on Artificial Intelligence* (ECAI), pages 333-337, Valencia (ES), 2004.

[Fensel et al., 2001] Fensel, D., McGuiness, D. L., Schulten, E., W. K. Ng, E. P. Lim and G. Yan. *Ontologies and electronic commerce*. IEEE Intelligent Systems, 16, 1, 8 – 14, 2001.

[Gartner, 2006a] *Hype Cycle for Compliance Technologies*, Gartner Research report, ID Number: G00139223, 12 July 2006.

[Gartner, 2006b*] Hype Cycle for Web Services and Related Standards and Specifications*, Gartner Research report, ID Number: G00141033, 11 July 2006.

[Gartner, 2006c] *Hype Cycle for Emerging Technologies*, Gartner Research report, ID Number: G00141901, 18 July 2006.

[Gartner, 2006d] *Hype Cycle for Business Processes in Supply Chain Management*, Gartner Research report, ID Number: G00140122, 14 July 2006.

[Gartner, 2006e] *Hype Cycle for Application Integration and Platform Middleware*, Gartner Research report, ID Number: G00141124, 13 July 2006.

[Gartner, 2006f] *Hype Cycle for Web Technologies, Gartner Research report*, ID Number: G00141111, 7 July 2006.

[Gartner, 2006g] *Hype Cycle for Content Management,* Gartner Research report, ID Number: G00141105, 14 July 2006.

[Gartner, 2006h] *Hype Cycle for Human-Computer Interaction*, Gartner Research report, ID Number: G00141150, 4 July 2006.

[Gartner Report, 2007a] David, W. C., Gall, A., Gall, N., *Finding and Exploiting Value in Semantic Technologies on the Web*, ID Number G00148725, 9 May 2007.

[Gartner Report, 2007b] Smith, D.M., *Key Issues for Web 2.0 and Consumerization*, Gartner Research report, ID Number: G00146351., February 2007.

[Gartner Report, 2007c] Logan, D., Knox, R.E., *When to Use Semantic Technologies*, Gartner Research report, ID Number: G00147066, 11 May 2007.

[Giddens, 1984] Giddens, A., The Constitution of Society. Berkeley, CA: University of California Press. 1984.

[Graves, 2007] Graves, M., *The Relationship Between Web 2.0And the Semantic Web*, Invited Talk at the European Semantic Technology Conference, Vienna, 2007.

[Giunchiglia et al., 2007] Giunchiglia, F., Yatskevich, M. and Shvaiko, P., Semantic matching: Algorithms and implementation, *Journal on Data Semantics*, IX, pages 1-38, 2007.

[Handschuh, 2002] Handschuh, S., Staab, S., Ciravegna, F., *S-cream-semi-automatic creation of metadata.* Proc. of the European Conference on Knowledge Acquisition and Management, 2002.

[Hendler, 2007] Hendler, J., The Dark Side of the Semantic Web, *IEEE Intelligent Systems*, Vol. 22, No. 1, 2007.

[John Borland, 2007] Borland, J., *A Smarter Web*, Technology Review, Vol. 2 March/April 2007. http://www.mindswap.org/papers/2007/90-93.pdf

[Kalyanpur, 2003] Kalyanpur, A., Parsia, B. , Hendler, J., Golbeck, J., SMORE - semantic markup, ontology, and RDF editor, available from http://www.mindswap.org/papers/SMORE.pdf , 2003.

[Koskinen, 2006] Koskinen, T., *Social Software for Industrial Interaction,* OZCHI, Sydney, Australia November 20-24, 2006.

[Lassila and Hendler, 2007] Lassila, O., Hendler, J., *Embracing "Web 3.0"*, IEEE Computer Society, 2007. http://www.mindswap.org/papers/2007/90-93.pdf

[Lave and Wenger, 1991] Lave, J., Wenger, E., Situated Learning: Legitimate Peripheral Participation. New York: Cambridge University Press. 1991.

[MacManus, 2005] MacManus, R., *Web 2.0 and Semantic Web: Mars and Venus?* , September 18th, 2005. http://blogs.zdnet.com/web2explorer/?p=13

[Mills Davis, 2004] Davis, M., The Business Value of Semantic Technologies, TopQuadrant Special report 2004.

[Mills Davis, 2006] Davis, M., Semantic Wave 2006, Part 1: Executive Guide to Billion Dollar Markets, A Project10X Special Report, January 2006.

[Missikoff, 2002] Missikoff, M. Navigli, R. Velardi, P. , *Integrated approach to Web ontology learning and engineering*. Ist. di Analisi dei Sistemi ed Inf., CNR, Rome ; IEEE Computer Volume: 35, Issue: 11, Nov 2002.

[Orlikowski, 1991] Orlikowski, W. J., *Integrated information environment or matrix of control? The contradictory implications of information technology*. Accounting, Management, and Information Technology, 1(1). 9-42. 1991.

[Papazoglou et al., 2006], Papazoglou M., Traverso P., Dustdar S. and Leymann F., Service-oriented computing research roadmap. Technical report, 2006.

[Petridis et. al, 2006] Petridis, K., Bloehdorn, S., Saathoff, C., Simou, N., Dasiopoulou, S., Tzouvaras, V., Handschuh, S., Avrithis, Y., Kompatsiaris, I., and Staab, S. "Knowledge Representation and Semantic Annotation of Multimedia Content". *IEE Proceedings on Vision Image and Signal Processing, Special issue on Knowledge-Based Digital Media Processing,* Vol. 153, No. 3, pp. 255-262, June 2006.

[Pollock and Hodgson, 2004] Pollock, , J.T., Hodgson, R., Adaptive Information: Improving Business through Semantic Interoperability, Grid Computing, and Enterprise Integration Pollock, , J.T., Hodgson, R., John (Eds.), Wiley & Sons, 2004.

[Shadbolt at al., 2006] Nigel Shadbolt, N., Berners-Lee, T. , Hall, W., *The Semantic Web Revisited* IEEE Intelligent Systems pp. 96-101, May 2006.

[Singh et al., 2005] Singh, R., Iyer, L.S., Salam, A.F., *Semantic eBusiness*. International J. Semantic Web and Information Systems 1, 1 , 19–35, Jan–Mar. 2005.

[Singh et al., 2006] Singh, R. and Salam, A.F., *Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans*, IEEE, vol. 36, no. 3, may 2006.

[Stamou et al., 2006] Stamou, Giorgos, Jacco van Ossenbruggen, Jeff Pan and Guss Schreiber. (2006). Multimedia annotations on the Semantic Web. *IEEE Multimedia*, 13(1):86--90, January-March 2006.

[Starbuck, 1992] Starbuck, W. H., *Learning by knowledge-intensive firms*, Journal of Management Studies, 29, 713-740. 1992.

[von Hippel, 2002] von Hippel, E., *Horizontal innovation networks - by and for users*, MIT Sloan School of Management Working Paper No. 4366-02, 2002.

[Wenger, 1998] Wenger, E., Communities of Practice: Learning, Meaning and Identity, CUP, 1998.

[Wiig, 1997] Wiig, K., *Knowledge management: where did it come from and where will it go?*, Journal of Expert Systems with Applications, 13(1), 1–14, 1997.

Web Sites

http:// www.flickr.com
http://del.icio.us
http://digg.com
http://ems01.mpi.nl/choice-wiki/CHOICE_Main_Page
http://eprints.ecs.soton.ac.uk/11101/01/iswc-final.pdf
http://events.idi.ntnu.no/stab07
http://iswc2006.semanticweb.org
http://iswc2007.semanticweb.org
http://knowledgeweb.semanticweb.org/o2i
http://sbpm2007.fzi.de
http://wordpress.org
http://www.ag-nbi.de/conf/MSWFB
http://www.ag-nbi.de/conf/SEBIZ06
http://www.csc.liv.ac.uk/semanticweb/KWeb_ESTC_WS.html
http://www.disa.unitn.it/net-economy/first/index.html
http://www.estc2007.com
http://www.eswc2007.org
http://www.flickr.com
http://www.investors.com/editorial/IBDArticles.asp?artsec=16&artnum=3&issue=20070601
http://www.investors.com/editorial/IBDArticles.asp?artsec=16&artnum=3&issue=2070601
http://www.joost.com
http://www.loa-cnr.it/fomi
http://www.ontology-advisory.org
http://www.ontoprise.de
http://www.ontoprise.de/content/e1171/e1182/index_eng.html
http://www.ontoprise.de/content/e1171/e1196/e1197/index_eng.html
http://www.ontoprise.de/content/e1171/e1219/index_eng.html
http://www.ontoprise.de/content/e1171/e1231/index_eng.html
http://www.ontoprise.de/content/e1171/e1249/index_eng.html
http://www.softwareag.com
http://www.softwareag.com/corporate/company/pressroom/pressreleases/20050309_enterprise_information_integration_page.asp
http://www.w3.org/2001/sw/sweo/public/UseCases/Bankinter
http://www.w3.org/2001/sw/sweo/public/UseCases/Sun
http://www.youtube.com
https://www.garlik.com/index.php
http://en.wikipedia.org/wiki/Semantic_MediaWiki
http://www.cs.cmu.edu/~softagents/papers/payne_terence_2002_2.pdf
http://www.w3.org/2001/sw/sweo/public/UseCases/NRK

Appendix 1: the hype cycle curve

Since 1995, Gartner's hype cycle curves are widely accepted graphic representations, which characterise the over-enthusiasm and subsequent disappointment that typically happen with the introduction of new technologies.

Business applications of specific technologies evolve in a very different way from common products, and their life cycles are not based on the four traditional stages of evolution (introduction, growth, maturity and decline) but on the five stages defined by Gartner as:

1. technology trigger: the first phase, called also breakthrough, refers to the product launch process or other event that generates significant interest in the market and in the society;
2. peak of inflated expectations: the second phase is characterised by a frenzy of publicity that typically generates over-enthusiasm and unrealistic projections. There may be some successful applications of a technology, but there are typically more failures;
3. trough of disillusionment: in the third phase, technologies enter the "trough of disillusionment" because they fail to meet expectations and quickly become unfashionable. Consequently, the press usually abandons the topic and the technology, and only experts and some other passionate individuals work with that technology improving methods and theories;
4. slope of enlightenment: although the press may have stopped covering the technology, some actors experiment to understand risks, benefits and practical uses of technology applications;
5. plateau of productivity: in the fifth phase, the technology benefits become widely demonstrated and accepted in the market and in the society. The technology becomes increasingly stable and evolves in the second and the third generations. The final height of the plateau varies according to whether the technology is broadly applicable or benefits only a niche market.

As in the case of products, the five stages affect the technology diffusion, adoption, and profit, therefore different strategies have to be employed to ensure success within the market and the society, e.g. the plateau of productivity.

Another important factor described by Gartner's hype cycle curves is the timeframe dimension, which describes the probability technologies have to reach the plateau of productivity:

○	less than 2 years to achieve the productivity plateau's timeframe;
●	2 to 5 years to achieve the productivity plateau's timeframe;
□	5 to 10 years to achieve the productivity plateau's timeframe;
△	more than 10 years to achieve the productivity plateau's timeframe;
⊗	obsolete before the productivity plateau's timeframe.

Appendix 2: glossary

API Application Programming Interface is a source code interface that an operating system or library provides to support requests for services to be made of it by computer programs.

B2B Business to Business is a term commonly used to describe the electronic transactions between firms.

B2C Business to Consumer describes electronic commerce addressed to end consumers.

BPEL Business Process Execution Language is an executable business process modelling language.

ebXML Electronic Business using eXtensible Markup Language, commonly known as e-business XML.

EDI Electronic Data Interchange.

GPS Global Positioning System is the standard generic term for satellite navigation systems that provide autonomous geo-spatial positioning with global coverage.

IT/ICT Information Technology / Information Communication Technology. The development, implementation, and maintenance of computer hardware and software systems to organise and communicate information electronically.

KR, **KRR** Knowledge Representation, Knowledge Representation and Reasoning.

KW Knowledge Web is a 4 year Network of Excellence project funded by the European Commission under the 6th Framework Programme. Knowledge Web began on January 1st, 2004. Supporting the transition process of Ontology technology from Academia to Industry is the main and major goal of Knowledge Web.

NoE Network of Excellence is a particular form of financing networked research communities.

OWL Web Ontology Language is a W3C-endorsed language for defining and instantiating web ontologies.

OWL-DL OWL-Description Logic is an OWL sub-language. Its name stems from its correspondance with a highly epressive but still decidable description logic.

OWL-S OWL-Service is an ontology built on top of the Web Ontology Language (OWL) by the DARPA DAML program. It replaces the former DAML-S ontology within the OWL-based framework of the Semantic Web, and describes Semantic Web Services.

P2P Peer to Peer computer network.

RDF Resource Description Framework is a W3C specification originally designed as a metadata model, and now used as a general method of modelling information, through a variety of syntax formats.

SME Small and Medium Enterprises.

SOA Service Oriented Architecture is a software architecture that defines the use of loosely-coupled services to support the requirements of software users.

SPARQL SPARQL Protocol And RDF Query Language is an RDF query language. Its name is a recursive acronym.

SWRL Semantic Web Rule Language is a proposed standard which combines sublanguages of OWL (OWL-DL and OWL-Lite) with those of the Rule Markup Language (Unary/Binary Datalog).

W3C The World Wide Web Consortium develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential.

WSDL The Web Services Description Language is an XML-based language that provides a model to describe web services.

WSMO Web Service Modelling Ontology is an ontology currently developed to support the deployment and interoperability of Semantic Web Services.

XML Extended Markup Language is a general-purpose markup language standardised by the W3C.