ON THE ROAD TO BUSINESS APPLICATIONS OF SEMANTIC WEB TECHNOLOGY

Sematic Web in Business - How to Proceed

Kari Oinonen
Kiertotie 14 as 3, 40250 Jyväskylä, e-mail: karioinonen@kolumbus.fi

Abstract: This paper discusses potential usage of Semantic Web in business applications and provides one way to proceed faster.

Current situation in Semantic Web application area is discussed. General appropriate trends in technology development and in communication industry and industry general are reviewed in order to see how Semantic Web technology fits with these trends.

Finally this paper suggests that, to get the technology into use, a common application framework should be formed. This framework shall look not only the technology, but also application, ICT architectures and business models this technology makes relevant. Definition of this framework is proposed to be a part of a road map process for which guidelines are provided.

Key words: road map, Semantic Web , semantic technology, business models, ICT architecture

1. BUSINESS APPLICATIONS AND SEMANTIC WEB

The potential of semantic technology is far too wide to be covered fully. There exists Semantic Web applications like MuseumFinland in Semantic Web (MuseumFinland) or some e.g. diagnostic applications in production or preproduction stage. However it is a fact that Semantic Web applications have emerged much slower pace than was assumed by field or technology experts. Not in great numbers can we see application types and application

areas where the semantic layer in the Web can make formerly impossible things possible, formerly uneconomical things and applications economical. This may be explained partly by the fact that some successful applications are not publicly known. Meanwhile various kind of more traditional Web applications are being created in big numbers. Why is that. The next section discusses on this more deeply.

1.1 Why Semantic Web has progressed so slowly

The following list describes some of the facts that have or are affecting on the progress of Semantic Web and its applications in business. One of the main facts is that the problems related to Semantic Web are one or two orders in magnitude bigger compared to problems that were encountered with the Web. This is because of the complexity involved with the semantics and communication in general.

- Technology development has taken time and resulted in changes in tools and standards.
- Standardization: problems with standardization and parallel development, perhaps too many standards so users do not know which ones to use for what.
- Competition within technology.
- Distribution of development.
- Possibly the existing Semantic Web technology do not cover the problem or business needs area well enough.
- Real existing business needs are lacking or are not recognized.
- Too complex (technology).
- Too new – not proved.
- Most public examples are more or less simple ones and they may not reflect the technology potential.
- Existing applications and business models are naturally something newer technology and business models have to overcome.
- Semantic Web technology may be at best in areas that have not yet been tried enough. I mean communication between applications or humans who have not been able to communicate because too high costs or too complex integration and associated maintenance costs involved with the ICT systems.
- Currently - and probably in the future - user base only fraction of HTML user base.
- People who understand the HTML world in general are countless. Compared to the basic Web technology gaining corresponding competence and understanding level of semantic web technology and

related terms and concepts is a much bigger task. Only a fraction of people can navigate with these concepts or use corresponding tools.

- Another challenge for application of the technology is the lacking of application domain expertise that is needed to bring together with the semantic technology. From the technology point of view this may be the biggest cause of delay that is still affecting the technology progress to business applications.

Just looking the web pages of W3C (W3C) reveals part of the problem. Semantic Web is just one topic among more than 50 topics. It is not a straightforward task for a newcomer or potential future user or decision maker to find out relations between these topics and their applicability. The another part of the problem is not a technical one. A national research project RUBIC (RUBIC) pointed out that technology is not the main limiting factor in interoperability and networking between companies. The main challenge is how to get companies into open and co-operative development of business models, business processes and supporting technologies.

This paper aims to overcome some of these causes of delay or obstacles by suggesting a co-operation between relevant parties to form a road map on how best utilize possibilities the Semantic Web enables. The suggested road map will also serve as a means to rise the level of understanding the technology potential for decision makers in the businesses. The end part of this paper provides a process according to which a road map can be formed.

2. LOOKING HOW GENERAL TRENDS RELATE TO WEB AND ESPECIALLY SEMANTIC WEB POTENTIALLY INCREASING USE

Here I will cover some of the facts that have affected in the adoption of the Web in its basic form and trends that are affecting or can affect on the adoption of the semantic layer on the Web.

2.1 General technology trends

According to Altchuller there are eight factors along which technology develops. These are (Altchuller, G. 1998): life-cycle, dynamization, multiplication cycle (transition from mono- to bi- and poly- systems, transition from macro to micro level, synchronization, scaling up or down, uneven development of parts, replacement of human (automation).

Looking the list above we can note that the basic Web technology has taken the development further by increased possibilities to manage life-cycles in business, there has been an increase in dynamization of some of the content on the Web, linking mechanisms have provided possibilities to combine file based information with information in other files. The content of Web pages have become more and more structured as XML-tagging is being taken into use widely. And what is important to note is that the amount of human work needed is decreasing in proportion to automation in areas like e-business and automated content configuration or filtering.

When comparing Semantic Web potential to the current Web technology, it is possible to note the following: semantic layer is promised to offer functionalities that increase dynamization, takes multiplication cycle further, enables management of increasingly more detailed information pieces, enables synchronization of content and events as time based content management becomes a possibility, and takes the automation level higher. In that respect the Semantic Web can be seen to lie in a natural development path of the Web.

2.2 Changes in communication technology

Looking to the past not log ago we can note that communication and Web based information systems are more and more common. Today it may be difficult to find a newer ICT system, which does not have a some kind of Web connection.

In the telecommunication business the technologies are converging. At general level this means that more and more of the voice and data contents are being delivered via Internet infrastructure. One of the key parameters in this convergence is the fact that Internet infrastructure is relatively cheap. Transmission speeds of Internet are increasing faster than in other technologies in telecommunication. This fact is a strong favor for more and more content being transferred between devices via Internet. Even mobile devices are starting to support Web.

The convergence taking place in communication industry opens up many doors and application areas to be utilized by semantic technologies. This means that Web technologies and Semantic Web technologies can be applied more and more easily on the content that is currently mainly managed in telecommunication networks.

General increase of information structuring – textual and audiovisual – is increasing. Not perhaps the proportion of structured compared to non-structured but still more and more content is being tagged in HTML and

increasingly in XML and this trend is continuing at least where content re-use and configurability are of importance.

2.3 Changes in business needs

One special feature that has changed recently in many business areas is the fact that there is an increasing need to be able to look complex topics from different points of view. Depending on the industry there are needs of information management supporting engineering which itself consists of many disciplines, asset management, maintenance management, quality reliability and environmental information management. Traditionally the solution has been several independent IT system solutions. However that kind of solutions have resulted in risks of local optimization and decrease on competitiveness. From the business point of view there exists heavy interdependencies and needs of data access between business functions or sub-businesses.

Related to this, knowledge and services based businesses, which are gaining favor, rely on capabilities to manage information which is essentially of network type and is originally created in a heterogeneous ICT environment.

3. ON THE NEED OF GENERAL APPLICATION AND BUSINESS FRAMEWORK

Benefits of creating a general application and usage oriented road map for the Semantic Web lie in two areas. One area where a road map is useful is interoperability of technologies and applications. Without general framework especially applications will not be inter operable. This means that that semantic layer of the Web will not be a semantic web. Instead, there will be a collection of application specific solutions and technologies which are poorly inter operable. That kind of situation will not support semantic web to the main road of technology. The other area where the road map is essential is to provide possible future users and decision makers means to provide general understanding of the technology, what the technology can do for them and when it can do that. These are the main reasons the general road map process is proposed.

The Web in the form most of people understand in the WWW started from the need. The need was to overcome the obstacle of being able to read fellow workers publications and project reports in CERN. To overcome this a simple page layout description language HTML was developed. Most of the additional technology needed was in place already. HTML proved

functional and turned to be a success – a bigger success than anybody could imagine. Now we must again recall what was the original need. It was the need to be able to read fellow worker's publications and reports that were originally produced by one of the several writing tools in use at CERN. Layout oriented HTML was the common layout language for documents, HTML viewers provided viewing capability, file access and transfer were made possible by http based on ip-protocol. There was no need for road maps or general application framework as the original problem was solved.

How the previous paragraph relates to the Semantic Web development. There is an analogy: Semantic Web seems to be developed to add WWW capabilities to actively search and retrieve and access data on the Web by active agents by utilizing semantic meta data. This in turn would increase resource use and interoperability of the data in the Web. Looking this initiative and comparing it with the discovery of WWW reveals that: The problem is analogous, but far more general. The problem itself is or seems to be much more complex than the problem of read-accessing fellow workers file contents. The problem itself is not so clear – it seems – even to the developers and this leads to a situation where the Semantic Web is being developed more on the technology aspect and the real world business needs are somewhat in the background.

The development of the Semantic Web and related tools are technology and solution driven. This is not not necessarily a handicap, if like was the case with HTML, the technology is targeted to solve the real and right problems. The problem for business applicability rises in this case from the fact that there exists some basic semantic technology, some of the technology needed may even be missing, the time span of the development is wide and still there seems to be no common understanding how to use the technology. There exists fully functioning semantic and Semantic Web applications, but often these applications are specially built for a purpose leading to a situation where these Semantic Web applications are in essence not inter operative.

Much of the technology and application development is taking place in distributed and even competitive environment. Continuing this route leads to non-convergent technology and slowing down of the technology adaptation and use. The fact that various semantic web technologies and applications are not inter operative is one of the key issues this paper aims to overcome. If development resources are needed to be utilized effectively and if the technology potential is needed to take into use, then a plan or road map is useful. The road map can act as a rough plan or guideline for the semantic technology adaptation and development too. In essence a road map makes the field more visible to the technology developers, financial decision makers and future users. Road map can provide also guidelines on what

technology or standards may best be applied and where and describe some future application areas that do not currently exist.

What makes the need for the semantic technology and especially Semantic Web application development and usage even more challenging is the combination of the following three facts:

- The definition of ontology: ontology is a formalization of conceptualization (Heimbürger, A. et al. 2004). Formalization itself is a straightforward task. The conceptualization process is instead a very challenging one. Challenges in the conceptualization process get bigger as the potential ontology user group gets bigger. The conceptualization needs agreement of potential users either in the ontology development phase or later phase when the ontology is taken into use. In both phases mapping between existing more or less local or proprietary concepts and concepts in the ontology needs to be done.
- Ontology usage in the Semantic Web: ontologies are treaded as global. This is fine if we can agree on the ontologies on some global forum. In general this is not possible although there exists limited area globally applicable ontologies. The need to be able at first stage to agree on ontologies is a challenge to overcome in many real world business cases. Of course the communication and automated resource discovery is eased if there is an agreement on terms concepts and their relations agreed beforehand. In general communication case this is not possible.
- The intention of the Semantic Web is: "The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in co-operation" (Berners-Lee, T. et al. 2001), then the way ontologies are created and used in Semantic Web only provide capability to very limited resource discovery and access.

If the limitation imposed by centralized architecture of ontologies management is of importance or not depends entirely on the business case – how the Semantic Web needs or is wanted to be used. In a machine diagnostic case it is well possible to create suitable ontologies mainly for own purposes and arrange limited but functioning interoperability between devices to be monitored on the field and a diagnostic service provider. For that particular service provider it is enough to use their own ontology. On the other hand the same equipment or device or plant specific information is used in documentation creation and maintenance, at design and maintenance stages in life-cycle. The problem of interoperability and information integrability is partly transferred to the ontology level.

Trust in the technology is essential for wide and general business adoption. Trust is related to the technology itself, its applicability, costs and benefits that are gained. The other part of trust is related to information

security. In the Web much of the data is available to be browsed. Password protection and similar means are used to limit the accessibility of the Web pages to only those that are allowed. In Semantic Web the data, information and knowledge is in a more structured and thus generally more detailed form compared to a Web page of file level. With the exception that in semantic environment more meta data may be available. To take the decision of how inter operative we want the semantic layer of the Web to be, we have to consider what data, information or knowledge is wanted to be inter operative. Are there needs in favor of general inter operativeness, or is it enough that Semantic Web will be used mainly for point-to-point information integrations purposes. If interoperability is wanted to cover, say a certain pool of data, then the technology should allow this even if this pool of data was formed grouping together data described by several independent ontologies.

In the Semantic Web environment new kinds of security mechanisms are needed to guarantee access to those allowed and at the same time protect valuable business information and knowledge. One of the solutions may come from the separation of security issues from the actual information content or service provision server.

Data and information encryption is one key topic that a road map should look at. This is because instead of accessing data for reading we are entering to a more dynamic situation where data integrity must be able to guarantee. Password protection will not be flexible enough on situations where interoperability takes place over several independent protected and unprotected systems.

A road map plan is provided for the Semantic Web. Road map process is presented with topics. The motivation for making the road map is to speed the pace the semantic technology is taken to the business use. Besides the main topics and propagation order suggestions there are suggestions who should contribute to the road map. In essence the road map definition process itself is of great importance because the process produces not only the road map but also creates understanding and wider view to the technology and its business use among the parties involved in the process

4. ROAD MAP PLAN

The Semantic Web road map plan consists of two main sections or parts. The first part defines the road map definition process. The second part helps in reaching agreement on the Semantic Web vision in terms of business, business models, information and knowledge management requirements and

technology usage in the area. The road map process provides the Semantic Web vision definition and will be a definition of the framework for Semantic Web business usage.

4.1 Road map definition process

There are several sources to get an applicable road map process and associated other information needed. The following description is based loosely on (Naumanen, M. 2001) and a process used in national TEKES technology development program (ÄLY 2004). It can be stated that road maps have been formed in Finland lately in many technology areas and there are some clear benefits of defining a road map for a field. One of the main benefits is a more unified view of the field that is created among stakeholders. To achieve this the Semantic Web road map should be published in appropriate forums.

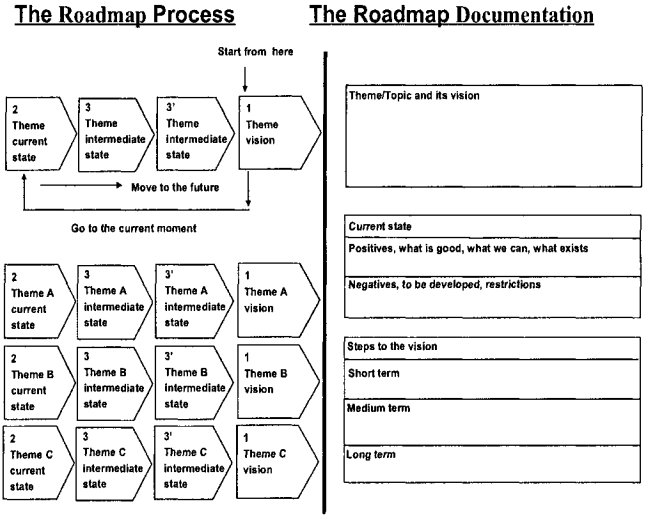

Figure 1. The Road map process and documentation

Generally the topics for the road map can be selected freely but often topics or views of market and business, applications and services, technology and R&D projects and science are used. The major task is to form development paths of the topics and to define what are the interactions of these topics.

What will make the road map process actualization a challenging task in the case of Semantic Web , is the fact that the area to be covered is huge. So,

in that respect there exists no single body that can form the road map and run the whole road map process taking the following list into account: the road map should cover many different business areas, it should cover many business models from single part producers to service business, the technology aspects and related existing standards and models usage and finally the one of the essential thing for communication: common language - ontologies management. Therefore it is suggested that there will be a suitable division of small dynamic work groups that contribute to the road map. Business and industry should take part in the work for two essential reasons: correct requirements and buy-in.

4.2 Road map process

In the following a collection of some essential Semantic Web road map topics are presented. For each topic there are presented some subtopics that I feel has to be covered in the semantic web road map process. Together the topics and their subtopics form the basic requirements that a road map for Semantic Web in business use shall cover.

As the road map process presented earlier describes, it is suggested that the vision is defined first. The Semantic Web vision in the W3C (Berners-Lee, T. et al. 2001) can be used as the basis. To support the Semantic Web vision statement it is beneficial to form a description of what can be done with the technology when the vision is true. Additionally it is needed to have technology, use cases, business and user needs and ontology management defined at vision level. These vision definitions can then guide the path definition during the road map work.

As an example we could assume that the vision is true after 10 years of time in a board context. Starting from the vision and its description we come to the current state. Excellent review depicting current state and needs in business communication and possibilities for semantic technology applications was formed by national project RUBIC (RUBIC). This project concentrated on interfaces between companies or between applications. Not all possible application areas for the semantic web were covered. However RUBIC project results can be used as a good starting point. As the first intermediate state we can choose two years and for the next intermediate we can choose five years from now. For these intermediate steps that lead to the vision we need to define and describe what is possible with the technology and how it is used in each particular state.

4.2.1 Technology topic

This topic in the road map covers technology issues. What Internet and Semantic Web technologies are ready and when, what technologies may be best applied on what application areas and how to guarantee interoperability.

- What Semantic Web standards to use on what purpose : SUO (standard upper ontology), DAML (www.daml.org), OIL (http://www.ontoknowledge.org/oil/), OWL, PSL , KIF, etc.
- Relations and connect ability to e.g. STEP (ISO 10303) standards, national or international industry specific standardization initiatives on procurement item management and on electrification projecting (PSK 7401) data exchange.
- What kind of information management functionalities and capabilities are needed in e.g. adaptability of business networking, life-cycle support of products and services – from the requirements management to the usage and service phases, innovation process, competence management, e-business and relations to standards in e-business management – ebXML and RosettaNet.
- Semantic Web relation and support of modeling activities like CIM-OSA, or MFM.
- How to manage the fact that industry areas are at different development stages.
- There should be a review to check usability of existing ontologies and the need to create ontologies from needs basis.
- How to guarantee future compatibility if the technology and ontologies are developed concurrently and parallel.
- User interfaces for the semantic multidimensional information.
- How to guarantee information integrity and security in a distributed heterogeneous environment.
- What other technology can be beneficial to join together with semantic technology or semantic information for a certain kind of task or application. As an example Web Services provide handy means for semantically marked information access.

4.2.2 Use cases -topic

A list of basic functionality Semantic Web technology is presented. Some of the listed functions are basic functions of the technology and some are built on the basic functions. The question is of communication between IT systems or applications, but with the aid of suitable devices this can be extended to man-to-man communication and to other combinations.

Initiatives like the W3C Speech Interface with a suite of markup definitions are aiming to realize this kind of communication combinations (W3C Speech Interface).

- Finding of information and resources.
- Building of collections of interdependent Web pages.
- Classification of information.
- Information collection for analysis or decision purposes.
- Linking of information in Web pages between different information locations.
- Maintenance of the links between pieces of information.
- Support of network type of information models.
- Information analysis based on other information or knowledge.
- Knowledge management based on the basic functionality.

4.2.3 Business and usage -topic

This topic shall act as needs and requirements for other topics. General and industry specific trends should be reviewed as well as industry and business representatives shall be interviewed to get better insight of business requirements and the change of requirements during 10 years time span. The road map should clearly point out how business needs relate to Semantic Web technology and its matureness at current state, at each intermediate state and at the target state 10 years from now.

- The meaning of Semantic Web usage to the customer or end user in terms of costs, speed and benefits in information and knowledge management.
- The effects of Semantic Web on knowledge and service based businesses.
- To be taken into consideration: different industry and business areas have their own needs and preferences that can have affect on the development and the pace Semantic Web is taken into use within that industry or business.
- What kind of functionality and capabilities are needed in e.g. adaptability of business networking.
- Life-cycle support of products and services – from the requirements management to the usage and service phases, innovation process, competence management, e-business and relations to standards in e-business management – ebXML and RosettaNet (RosettaNet).
- In life-cycle support and in related knowledge and service business it is a need to be able to support and manage network type of information models.

- Current and future needs of industry or business that are potential users. Find out at general level needs change in 3, 5 and 10 years perspective to help targeting.
- The challenges of managing effectively ever increasing number of documented and non-documented information is increasing continuously.
- Definition of the most promising application areas from the business point of view.

4.2.4 Ontology topic

Ontologies and their management is one of the key components of the Semantic Web. How ontologies are created, managed and maintained and how easy it is to integrate and refer other ontologies defines ontological limitations to the success and interoperability of the whole semantic web.

- There should be a review to check usability of existing and applicable ontologies and the need to create ontologies from needs basis.
- Different industry and business areas have their own terms and concepts in use. There are limitations on how well it is possible to force a certain ontology into use.
- Ontologies management: the top-down approach is not a feasible except in limited areas. There are industry or industry area specific initiatives that have produced applicable basis for the formation of ontologies. These can be utilized. There are existing generally accepted standards and ontologies that are widely used. Example of these is Dublin-Gore (Dublin-Core) in publication business and RosettaNet in e-business (RosettaNet).
- Currently ontologies are treated as global, but in practice there is a need for smaller, more local, more dynamic and at the same time inter operable ontologies.
- If it is concluded, during the Semantic Web road map process that a more dynamic way of ontologies management is needed, then there must be mechanisms to update ontologies and state the ontology owner. As these are very important meta data concerning the ontologies them self.
- Complexity can be managed by organizing the elements into local networks or modules which, because of their connectivity, have strong, well defined behavioral characteristics (Tossavainen, T. 2002). This reduces the global burden of producing coherent behavior, since the internal behavioral co-ordination of the modules is substantially handled locally. For ontologies this principle means distributed nature of ontologies development and management. Also this means that there should be mechanisms that guarantee good enough interoperability between ontologies where a need exists.

4.2.5 Who should contribute

As the road map definition process for the Semantic Web and its business use is a challenge because of the nature of the task, it is not possible that a small group of people can define the road map. Instead there is a need for co-operation between technology developers in research institutes, universities and companies; ICT managers and business developers as well end users in business; experts in standardization bodies and technology integrators.

5. HOW TO CONTINUE

One possibility to proceed is to form a small group of experts who could at first stage predefine a suggestion of the vision together with industry and research representatives. After that a more formal and wider road map definition process could be started. The road map process should be finalized and results published in less than a year.

6. REFERENCES

Altchuller, G. 1998. *40 Principles; TRIZ keys to technical innovation.* Technical innovation center, Worchester, MA, 1997.

Berners-Lee, T., Hender, J., Lassila, O., The Semantic Web, Scientific American, May 2001.

Dublin-Core, Dublin-Core Metadata Initiative, http://dublincore.org, (referred 29.4.2005).

Heimbürger,A et al.(2004) . *XLinkTime: Developing time-sensitive linking mechanisms for the future Web.* Julkaisusarja A – Tampereen teknillinen yliopisto, Porin yksikkö. Nro A4/2004.

ISO 10303, http://www.tc184-sc4.org/SC4_open/SC4_Work_Products_Documents/STEP_(10303)/.

MuseumFinland – Finnish Museums on the Semantic Web, March 8, 2004. University of Helsinki, Department of Computer Science, Helsinki Institute of Information Technology, http://museosuomi.cs.helsinki.fi (referred 29.4.2005).

Naumanen, M., *Roadmap – Kartta Menestykseen*, published by Teknologiainfo Teknova Oy, 2001, ISBN 951-817-775-9.

PSK 7401, Information Exchange in Electrification and Automation Project, http://www.psk-standardisointi.fi/Standardiluettelo/tekstiSS.htm. (referred 29.4.2005).

RosettaNet, http://www.rosettanet.org (referred 13.6.2005)

RUBIC, http://www.tieke.fi/mp/db/file_library/x/IMG/13108/file/Rubic_Loppuraportti_V3.0_02_07_04.pdf , in Finnish, (referred 13.6.2005)

Tossavainen, T., *System Usability of Complex Technical Systems,* Acta Polytechnica Scandinavia, Industrial Management and Business Administration Series No. 12, Espoo 2002, 236 +17 pp. Published by the Finnish Academies of Technology. ISBN 951-666-597-7, ISSN 1459-9426.

ÄLY, TEKES technology program, http://webserv2.tekes.fi/opencms/opencms/OhjelmaPortaali/Kaynnissa/ALY/fi/julkaisutjalinkit.html. (referred 29.4.2005).

W3C, http://www.w3.org/ (referred 13.6.2005)

W3C Speech Interface, http://www.w3.org/Voice/Activity.html (referred 13.6.2005)