PREFACE

This publication evolved from the Austrian research programme FIT-IT (acronym for Research, Innovation, Technology – Information Technology). FIT-IT supports co-operative research addressing most challenging innovation and technology development in the area of Information Technology (IT) by focusing on visionary and interdisciplinary projects. The main reason to further invest in Information Technology is that they belong to one of the most prosperous fields of today's economy and that they will continue to do so.

Since 2002 the Austrian Ministry for Transport, Innovation and Technology (BMVIT) has been running the FIT-IT programme with a total budget of 45 million Euro. The programme claim is to operate at the highest international quality level and as a consequence to leverage Austrian research profile to a level which would not be possible without FIT-IT. One of the five programme lines of FIT-IT deals with "Semantic Systems and Services". Unlike "non-semantic" systems, semantic systems utilize meaning and associations and thus leading to a radically new paradigm to deal with content of any type. Over the last years Semantic Systems have become an important and integral part of both technological IT advancements and applied research of non-IT branches. According to recent studies of internationally recognised analysts the market potential for these new technologies is huge and it is expected that it will continuously grow over the years to come. Within this context, the Austrian government invests approximately three million Euro per year in pre-competitive projects addressing open and still unanswered research questions in "Semantic Systems". And many of the funded projects have already achieved impressive results.

Semantic Systems can only be brought to their full potential, if researchers from different disciplines, including computer science, linguistics, knowledge management, etc. collaborate in a synergetic way. And first results are already reaching the market with a significant impact on the quality of services provided. Within this context the FIT-IT project sem'base aims to increase the awareness of the potential and the resulting benefits of Semantic Systems in Austria – and this not only for the scientific community, but also for industry and the wider public.

One measure to increase awareness are publications as the one provided here. This publication perceives itself as a snapshot of the current state of the art in Semantic Systems. Many different perspectives are taken to offer the readers a most comprehensive picture: A scientific contribution sketches the topic from a technical and non-technical viewpoint; an online survey with about 70 participants provides a deep insight into the topic, carefully evaluated interviews of about 20 key players in semantic technologies indicate future trends, the funding agency describe their intentions to invest in these technologies and finally the personal assessments of many individuals complete the picture provided by this book.

We hope that the results of this book will serve as a unique starting point for further qualified and focused discussions about Semantic Systems in Austria. Those readers who even want to access more information and who want to participate in a hopefully lively discussion are invited to visit the website http://wiki.sembase.at of the sem'base project. This site also compiles information about key players in Austria and about other related activities and projects. We hope that our activities are an important stimulus to trigger and to further develop the Austrian Semantic Systems community. If you are not part of this community yet – take your chance, read the book and join!

Reinhard Goebl
Federal Ministry of Transport, Innovation and Technology
Head of Department "Information, Nano and Industrial Technologies and Space"

October 2006

TABLE OF CONTENTS

Part I

Approaching Semantic Technologies

sem'base

Andreas Blumauer (Semantic Web School), Gisela Dösinger (Know-Center), Thomas Fundneider (tf consulting), Max Harnoncourt, Paul Meinl (beide factline)

1. Motivation

Topics related to semantics have been of interest to scientists across a wide range of professions, such as linguists, philosophers, social/law sciences and others, for a long time. Findings in each of these fields have an impact on the other professions, so that semantics origins are highly trans-disciplinary. Lately, semantic issues have increasingly started to attract attention. Furthermore, interest in semantic topics has become more specialised, in the sense of the semantic web, focusing on Internet-related aspects. The currently most contributing discipline is computer sciences. However, due the legacy of the “borderless” character of semantics, a demarcation towards other disciplines has not yet been established.

On the other hand, semantic technologies may have an enormous potential to stimulate innovation and knowledge creation in other disciplines. Thus, money spent on organisations involved in semantic systems has a likely transfer effect into other areas, as key researchers in the semantic systems field may trigger new developments in other research and industry fields. This major value-added process must be seen in the wider macro-economic context as semantic systems can be considered as a bridging technology providing input and value for many other professions.

2. About sem'base

sem'base is an awareness measure in the programme FIT-IT Semantic Systems. It started in February 2006 and finished its activities in November 2006. The project embraced five main activities to reach its goals:

- Survey and interviews with key players
- Awareness building for semantic technologies at the European Forum in Alpbach 2006 and at the topic related conferences I-KNOW 06 and Semantics 06
- Gap analysis indicating the opportunities and weaknesses of the semantic systems sector in Austria
- Development of a domain specific glossary
- An online catalogue of semantic systems players in Austria

A consortium of four partners contributed to the project, which was co-ordinated by Thomas Fundneider from tf consulting:

Factline	www.factline.com
Know-Center	www.know-center.at
Semantic Web School	www.semantic-web.at
tf consulting	www.tfc.at

3. Objectives of the project

The main aim of sem'base was to create awareness and visibility for the semantic systems sector with the geographical focus on Austria. To this end, sem'base pursues the following objectives:

- Increase the visibility of semantic systems through an active engagement and commitment at the European Forum Alpbach 2006, at I-KNOW 06 and at Semantics 06
- Find out who is doing what in the Austrian semantic systems sector
- Undertake a gap analysis illustrating the potential and shortcomings of the Austrian semantic systems sector
- Strengthening the competitiveness of Austrian semantic systems players through a triggered dialogue between researchers and industrial partners
- Provide a common language through a glossary in order to facilitate communication between users, researchers and industry

sem'base addresses the current challenge that industrial users pretty well know what they need/want, but cannot formulate it in a way that is understood by researchers and developers. Dissemination and presentation activities will support this challenge by increasing awareness and visibility of semantic systems in general and of Austrian actors in this area in particular.

4. Project results

sem'base successfully achieved its objectives. Besides the publication at hand, the project´s main results can be summarised under three headings.

Cross-over project at the European Forum in Alpbach
In co-operation with the Austria Press Agency (APA) and the artist Max Frey, sem'base placed a cross-over project (science, economy and art) at the heart of the main conference building, thus making it highly visible for every participant. The art installation (termed LIVING SEMANTICS) visually transformed knowledge maps that have been created by means of semantic technologies.

Illustration 1: Opening of LIVING SEMANTICS (speech by Reinhard Goebl) at the European Forum in Alpbach

Semantic Wiki representing the semantic systems sector in Austria
An online semantic wiki (wiki.sembase.at) has been created in order to represent the semantic systems sector in Austria. The approach of this portal is bottom-up, enabling users to edit the appearance and the content according to their needs. It is foreseen that the semantic wiki will stay active also after the end of sem'base. Further, the wiki also hosts a domain specific glossary.

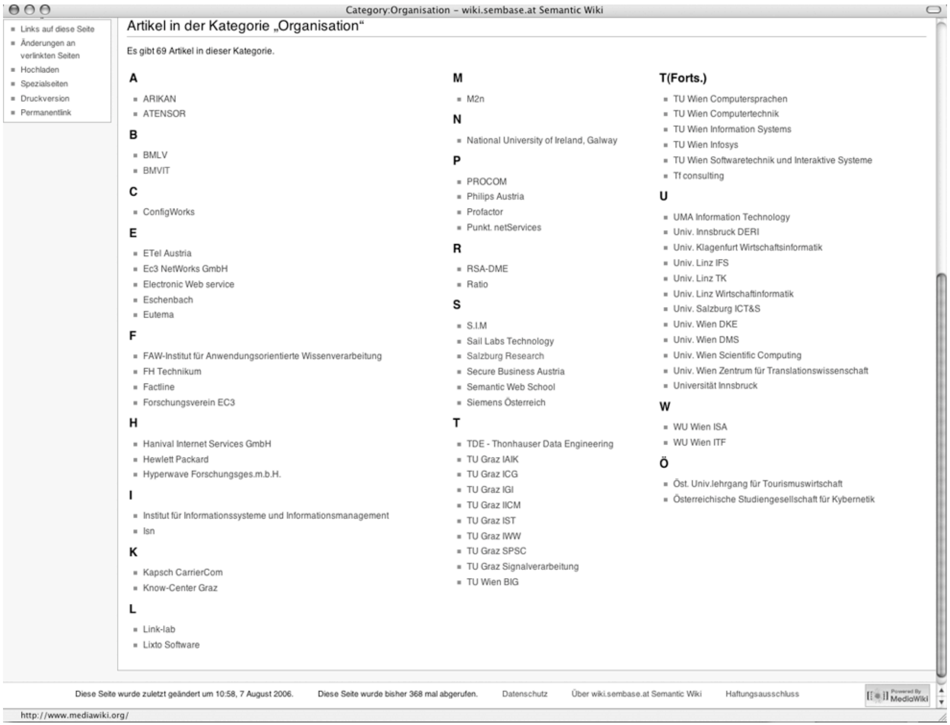


Illustration 2: Screenshot of the semantic wiki displaying key players in the semantic systems sector in Austria

Report on the Austrian situation regarding the semantic systems sector
Based on a web-based survey and on individual interviews with key players, the project deduced a report that highlights the challenges, strengths, and weaknesses with respect to the semantic systems sector in Austria. In this context, demands and offers, educational issues, the way semantic technologies are approached and technology awareness are highlighted. The detailed report can be found in this book starting on page 29.

SEMANTIC TECHNOLOGIES – AN INTRODUCTION

Klaus Tochtermann (Know-Center Graz, Graz University of Technology), Hermann Maurer (Graz University of Technology)

Abstract
This paper introduces semantic technologies from two different perspectives. The first part takes a non-technical perspective, explains what semantic technologies are all about and motivates by example why more semantics in today's applications would leverage their usefulness. But Semantic Technologies do not only affect stand-alone computers and applications, they have already heavily affected the World Wide Web and will continue to do so. A recent study by the University of California at Berkeley estimates that the amounts of available information will double every 70 days by 2020. Computational power, by contrast, doubles only every 18 months according to Moore's Law. Consequently, traditional information systems will not be able to cope with this flood of information. Semantic Web Technologies are considered a key technology having the potential to counterbalance this development. The most important Semantic Web Technologies are briefly described from a much more technical perspective in the second part of this paper.

1. Semantic Technologies – The non-technical Perspective

The recently published Semantic Wave 2006 Report [2] deals with many different perspectives one can have on Semantic Technologies. It defines the semantic wave – the current hype we observe for semantic technologies, as *a long wave of investment involving fundamental shifts in paradigm, technology and economics*. The fundamental shift in paradigm becomes possible due to a new concept of better integrating semantics and technologies. This new concept if often referred to as "Semantic Technologies". Semantic Technologies can be defined as *functional capabilities that enable both people and computers to create, discover, represent, organize process, manage, reason with, present, share, and utilize meanings and knowledge to accomplish business, personal, and societal purposes* [2]. What makes semantic technologies different from "non-semantic" technologies is the utilization of meaning and associations and their separation from data and program code.

To give an example: Media companies which sell automated media analysis about the history and development of an organization, are often confronted with the fact that organizations change their name over time or that different variations of the company name exist within different contexts. For example, a media analysis for the consultancy Accenture using "non-semantic" technologies could only find information after 2001 when the company name Accenture replaced the former name Anderson Consulting. But the time span from 1989 (founding year of Anderson Consulting) to 2001 would not be considered resulting in an incomplete analysis. Also, important information from the New York Stock Exchange (NYSE) could not be included in the analysis – simply because Accenture is listed there under the NYSE symbol ACN. The reason for this incompleteness is that due to the lack of semantics traditional technologies could not capitalize on "meaning" and thus classify Accenture, Anderson Consulting and ACN as instances of the same concept. Semantic technologies, however, could explicitly model Accenture and Anderson Consulting as instances of the same concept "Consultancy" and ACN as instance of concept "NYSE symbol". With their capability to model associations, the relationship between Accenture, Anderson Consulting and ACN could easily be expressed.

There is another important trend which started its success story in the mid-90ties: the development of Internet technologies and the establishment of the World Wide Web. Computers and the information stored therein are no longer isolated but tightly networked with countless many other computers around the world. This opens up radically new opportunities for exchanging information and knowledge not only within an organization but also across organizational boundaries. However, this success can only be brought to its full potential, if the networked computers and people collaborate with each other using the same meaning or semantics for the information and knowledge to be exchanged. For example, if one performs a Google search for "Apache" with the current Google technologies, the result will include hits about the US army helicopter, the Apache web server technology and the Indians "Apache", even though the user maybe had only the web server technology in mind when he/she started his/her search. This example illustrates well the limitations of the current WWW: The lack of semantics does not enable Internet applications, such as Google, to separate between the three concepts helicopter, web server technology and Indians which all have instances of the same name "Apache".

These limitations are the starting point to extend semantic technologies towards the WWW to provide more intelligent access to WWW resources and to better mediate between the wants of the users and the available WWW resources. This development is referred to Semantic Web, implying the development and application of Semantic Technologies for the WWW. Tim Berners-Lee defined in 2001 the Semantic Web as *is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperations* [1]. In the same year, the standardization committee for the World Wide Web, the WWW Consortium (W3C), argues that the Semantic Web is about two things *…about common formats for interchange of data…* and *…about language for recoding how the date relates to real world objects…* [6].

The scope of today's research in this field concentrates mainly on the following four areas [2]:
(Web) Services put the emphasis on computers' ability to discover, compose, orchestrate and manage services. The idea is simple: For example, if one wants to organize his/her holidays he/she should be able to start with one service (e.g. flight booking); once the flight is booked, the flight booking service should automatically trigger another service which organizes the hotel booking at the holiday location, once the hotel is booked, yet another service should automatically suggest day-trips starting from the hotel etc.
Content should be made interoperable, better searchable and accessible thanks to the combination of Semantic Technologies with traditional retrieval methodologies. The above example with the media analysis illustrates the needs of this research direction.
Cognition should augment the capabilities of knowledge workers. This requires that knowledge can be better analyzed, processed and executed at the workplace. An important trend in this field is context-aware support of knowledge workers which provides the knowledge worker *ad-hoc* with the knowledge they need to perform a given task. The challenge here is to identify the working context of a knowledge worker, semantically associate it with supporting resources and finally offer these resources to the knowledge worker. For example, in a situation in which a knowledge worker needs support to continue his/her task such resources could include a link to a related virtual community of experts, collaboration support, or business process support etc.
Networking should become semantic-enhanced to better enable computers to work together. This relates very much to Grid computing (i.e. distribution of the process execution across different machines on the Internet) and peer-to-peer computing (i.e. a computer network which parameters such as bandwidth relies on the parameters of the participants, for example telephony over IP).

In order to make all these visions become reality languages are required which are powerful enough to express well defined meaning and associations. The most important languages will be sketched in the next section.

2. Semantic Technologies – The technical Perspective

The way how information is structured is the key for expressing semantics in documents. This is why different document markup languages have evolved over the past years, all of them with different capabilities to express meaning. The markup helps to add structure and meaning to a language. The idea of markups goes back to the 60-ties, when SGML as meta-language for markup languages for documents was defined. Since then different descendants of SGML emerged, including HTML and XML. Due to the limitations of such markup languages in expressing meaning further description languages appeared, including RDF (Resource Description Framework) as metadata model and

OIL+DAML/OWL as full-fledged knowledge representation languages. These languages will be further described in the next sections.

2.1. Well-defined meaning in HTML

HTML (Hypertext Markup Language) is probably the most used and best known description language on the WWW. Its markups define how the information of a document has to be rendered. Take the following example:

```
<html>
<body>
     <h1> Semantic Technologies </h1>
     <p> This book, edited by the sem'base project, presents recent
         developments of semantic technologies.
         It can be purchased for the price € 29,-  </p>
</body>
</html>
```

The text between `<h1>` and `</h1>` is the first headline while the text between `<p>` and `</p>` is the paragraph containing the text of the document. The problem with well-defined meaning in HTML can be best explained by the question "How does an algorithm looks like which extracts the editor, title and price of the book?" Humans can easily derive this information from reading the text and it might be possible to develop one corresponding text parsing algorithm for exactly this document. But this algorithm would fail if it is applied to another HTML document which uses different words between the tags (for example "cost" instead of "price"). To overcome this situation of poorly formalised well-defined meaning it is necessary to further specify the meaning of the tags.

2.2. Well-defined meaning in XML

XML (eXtensible Markup Language) [7] is a subset of the above mentioned SGML. XML has the ability to express data formats of structured documents in a more formal way than it is possible with HTML. The above example could be expressed in XML as follows:

```
<book> Semantic Technologies </book>
<editor> sem'base project </editor>
<description>  This book presents recent developments of semantic
               Technologies. </description>
<price> € 29,-  </price>
```

Based upon the named entities it becomes easier for machines to interpret the XML code. XML documents can be seen as ordered labelled trees. Of course, this facilitates the development of an algorithm for extracting the editor, the title and the price of the book. Such an algorithm could be a simple traversal of the labelled tree defined by the XML tags. Still, there are some limitations of XML: Firstly, the well-defined meaning is not yet fully formally specified – there is no non-ambiguous definition that the price-tag really describes the price of the book. Secondly, the elements can be named or structured arbitrarily which in turn complicates the automated extraction of well-defined meaning. Some of these limitations can be overcome by using so-called Document Type Definitions (DTD). A DTD specifies the allowed vocabulary and the allowed combination of tags. The strength of XML lies in its informal approach to structure data; the weakness is that it cannot formally express well-defined meaning. What is required is a model which allows for the representation of data about other resources to give meaning to them in a clear and unambiguous way.

2.3. Well-defined meaning in RDF

RDF (Resource Description Framework) [8] is a language based upon XML to model data about other data (i.e. metadata). With RDF is becomes possible to make statements about resources on the WWW. Such statements are always structured in triples consisting of subject, predicate and object. To continue with the above example, the RDF statement for the information about which editor edited the book “Semantic Technologies” would look like follows

Illustration 1: Example of a RDF-statement

In RDF statements, the subject denotes the resource (the book on semantic technologies in the example), the predicate denotes aspects of the resources (the editor in the example) and expresses the relationship between subject and the object of the RDF statement. Composed RDF statements allow expressing data about more complex resources.

Similar to XML and XML DTDs it is possible to define RDF Schemas. A RDF Schema defines domain-specific properties and classes of resources which are associated with each other through the properties. The class primitives declare classes of resources (e.g., “Semantic Technologies” is resource of class “book”); they might be composed of subclasses which results in class hierarchies. The same applies to properties. To associate resources as instances with a specific class, type statements can be applied. The valid combinations of properties and classes are defined by domain and range statements. For example, these restrictions make it possible to declare that a resource of type “book” can only be associated with a resource of type “project” through the relation “editor” but not with resources from other classes like for example the funding agency of the sem'base project.

The advantages of RDF are that it is unambiguous and that it well separates between the semantic and the syntax of resources. The language modelling concepts are powerful enough to model well-defined meaning for many application domains. Still, the modelling possibilities are limited and what is even more restrictive is that it is not possible to express various logical statements. For example, the statement `editor is disjoint from publisher` can not be expressed semantically unambiguous in RDF and RDF Schema.

2.4. Well-defined meaning in OWL

OWL (Web Ontology Language) [5] is an ontology description language which has been derived from the American DARPA agent markup language (DAML) [3] and the European Ontology Interchange Language (OIL) [4]. OWL provides additional vocabulary together with a formal semantics. The claim of OWL is to be an extension of RDF/RDF Schema and XML/XML DTD both in terms of expressiveness and in what can be validly concluded. This of course requires more powerful language concepts. Since this in turn increases its complexity, OWL is divided into three increasingly expressive classes (OWL Lite, OWL DL for description logic, and OWL Full). Each class is an extension of its simpler predecessor. This means that OWL DL and OWL Full constructs are either additional to or expand constructs of OWL Lite. OWL DL and OWL Full offer the same language constructs, but OWL Full is more liberal about how to combine these constructs. The above claim of OWL in comparison to RDF/RDF Schema only applies to OWL Full and not to OWL Lite and DL, these two are just extensions of a restricted view of RDF. OWL DL and OWL Lite restrict RDF for the sake of using description logics and making decidable logical inferences through reasoners.

Unlike RDF, OWL is based - at least in part - on concepts from description logics. With this different basis, it becomes possible to express logical statements in OWL. For example, OWL Lite can state two individuals to be the same by using the feature `sameAs`. In contrast the feature `differentFrom` denotes that an individual is different from other individuals. In the above example of the media analysis, the feature `sameAs` could express that Andersen Consulting and Accenture are considered to be the same companies for the analysis. In addition, OWL DL and OWL Full offer constructs such as `unionOf, intersectionOf` or `complementOf`. For example, it would be possible for the media

analysis to define a new entity being the union of the company names Accenture, Andersen Consulting and the NYSE symbol ACN.
With OWL well-defined meaning can be much better expressed as with the other description languages. However, the trade-off is the high language complexity which makes it more difficult to apply OWL.

For the sake of completeness but without going into further detail, Topic maps should also be mentioned at this stage. Topic Maps [9] are an ISO standard aiming at an improved representation and interchange of knowledge. The interchange syntax of Topic Maps is based on XML.

3. Summary and Outlook

This paper provided a motivation for the emerging semantic technologies and semantic web technologies. It also presented the main language concepts of different and commonly used description languages. OWL certainly is the most expressive language, however, at the price of high complexity which might delay its application outside the scientific world. But OWL has also its limitations, particularly if it comes to proof (prove whether a statement is true) and trust (assess whether a resource is valid). These are the next steps to be addressed in current and future Semantic Web Community.

The Semantic Wave has by far not reached its maximal potential. As the preparation of the 7th Framework Program of the European Commission indicates, Semantic Technologies will play a crucial role in future research. For example, the unit DG INFSO/E2 convened an expert panel in March 2006 investigating research directions in "semantic based knowledge technologies and applications, with special emphasis on networked information". And also the market prospects are promising: The earlier mentioned Semantic Wave Report foresees that Semantic Technology is a $2 billion per year market and likely to grow to over $50 billion by the year 2010.

4. Acknowledgement

The Know-Center is a Competence Center funded within the Austrian Competence Center program K plus under the auspices of the Austrian Ministry of Transport, Innovation and Technology (www.kplus.at), by the state of Styria and by the City of Graz

5. References

[1] BERNERS-LEE, T. HANDLER, J., LASSILA, O: The Semantic Web, Scientific American, May 2001

[2] DAVIS, M., Semantic Wave 2006, A Project10X Special Report, www.project10x.com, January 2006

[3] DAML www.daml.org, November 2006

[4] OIL www.ontoknowledge.org/oil, November 2006

[5] OWL Web Ontology Language Overview, www.w3.org/TR/owl-features, November 2006

[6] The W3C Semantic Web Activities, www.w3.org/2001/sw, November 2006

[7] The W3C XML homepage, www.w3.org/XML/, November 2006

[8] The W3C RDF specification, www.w3.org/rdf/, November 2006

[9] TopicMaps www.topicmap.com, November 2006

[10] ftp://ftp.cordis.europa.eu/pub/ist/docs/kct/report-knowledgemeetin-march_en.pdf, November 2006

PUBLIC FUNDING FOR SEMANTIC SYSTEMS RESEARCH IN EUROPE AND AUSTRIA

Georg Niklfeld (FFG - Austrian Research Promotion Agency),
Bernd Wohlkinger (eutema Technology Management)

Abstract
Research in Semantic Systems is of strong interest to policy makers both at the European level and particularly in Austria. Semantic Systems is seen as a key area of innovation in software technology, supporting a wide variety of applications in a horizontal manner. Semantic Systems technology promises solutions to urgent societal and business problems in the fields of search and storage, knowledge management and interaction of humans and machines. In this article we review the available funding schemes for Semantic Systems research on the European and national level, and attempt to look ahead at the key opportunities and challenges for this technology in the next few years.

1. Introduction

Searching for digital information has become a ubiquitous task that is at the core of many business processes as well as the private or professional activities of IT users. Yet search is just one of the applications where future developments will depend heavily on further progress in Semantic Systems research. Web 2.0 is not just a buzz-word, but refers to new, more productive and more satisfying ways of using the available web infrastructure. Web 2.0 hinges on Semantic Systems research. Ontologies have been built and adapted for the last few years, now they build the basis for applications and technology research. In the next few years Semantic Systems will thus stop being a slightly elusive software technology developed in fancy research labs, and start to find their place in IT-infrastructures and networks.

Where Semantic Systems have already entered real applications, the process so far has been demand-driven by consumers and industry needs, and not so much technology-driven. Today many users and companies still have little understanding of Semantic Technologies – they have problems to be solved, e.g. large archives of multimedia material to be screened. The awareness that solutions to their problems are to be found in Semantic Systems is often lacking – awareness of Semantic Systems as a technology remains low, sometimes hindering the coordination and consequently the visibility of actors in the field.

Technological progress in the last years has been significant: a wide variety of applications are today enabled and strengthened by Semantic Technologies. Horizontal linkages to IT fields have been established. Yet, before Semantic Systems can enter volume markets, important technical challenges still need to be overcome. In the next years key problems to be solved are trust, interoperability, availability, correctness, maintenance, upgradeability, and interfaces of solutions that employ semantic technologies.
Often also the business processes and procedures are still not in place for the implementation of the new Semantic Technologies. Thus, commercial applications are still far away in several parts of Semantic Systems, and a further big effort in research will be needed. As just one example, semantic web service composition is today probably 2-3 years away from the market. Emerging application fields for Semantic Systems can today be seen in B2B and ISP. More showcases would be needed to raise awareness to increase the take-up of the technology by high-potential clients resp. users such as insurance companies or banks. Another problem is a shortage of trained researchers, experts and young professionals in several sub-areas of Semantic Systems. There are few dedicated programmes on Semantic Systems and large demand.

In sum, Semantic Systems technology today provides exciting opportunities for innovative economies, but it also includes a lot of unsolved issues that call for public intervention to overcome long-term risks and bottlenecks of supply and fledgling demand. In the next sections, we review the current public funding situation for Semantic Systems first at the European and then at the Austrian national level, before addressing some challenges for the next few years.

2. Semantic Systems Research in Europe

The European Commission funded R&D on Semantic Technologies already in the 5th Framework Programme for Research and Development (FP5, 1998 – 2002). The strategic objectives of the IST Programme („Information Society Technologies") as part of FP5 included topics on interactive publishing, human language technologies, as well as information access, filtering, analysis and handling. However, it was only in the IST Work Programme 2001 when "semantics" were explicitly mentioned. The strategic objective "Semantic Web Technologies" included the following contents:

- Methods and tools for coding and structuring digital content, for defining and declaring its semantics;
- Methods and tools for the derivation of semantic attributes of Web-based content (in particular video, audio and images);
- Semantics based tools for knowledge discovery and intelligent filtering and profiling such as information agents and specific query languages. Semantics based tools for collaborative filtering and knowledge sharing in specific or general user communities;
- Information visualisation: intelligent and visual interfaces which take advantage of semantic information structures to provide users with radical new ways to navigate and search naturally through unknown and complex information spaces.

In the following 6th Framework Programme (FP6, 2002-2006), resp. in its IST Programme, a concrete unit of the European Commission's DG INFSO devoted parts of its focus on Semantic Technologies: Unit E2 „Knowledge and Content Technologies". Moreover, "Semantics" and "Web Services" are found in several strategic objectives of IST in FP6, like:

- technology-enhanced learning and access to cultural heritage,
- networked business and governments,
- eSafety of road and air transports,
- eHealth,
- multimodal interfaces, and
- networked audiovisual systems and home platforms.
- Semantic Technologies became more and more integrated in other areas of R&D, thus spanning several topics and areas of ICT.

The coming 7th Framework Programme (FP7, 2007 – 2013) follows this trend as seen in the actual ICT draft "Work Programme 2007-2008". The former "IST" part will be called "ICT – Information and Communication Technologies", the third topic within FP7. Whereas IST in FP6 had a budget of 3.6 billion Euro for 4 years, ICT in FP7 will have funds of 9.11 billion Euro available for 7 years. A clear comparison of thematic budget distribution between FP5, FP6 and FP7 is impossible because of structural changes.

A specific objective within FP7-ICT is again devoted to "Semantic Technologies": "3.4.2.1 Intelligent content creation and semantics". This objective calls for environments to ease content sharing and management like "Semantic Web Technologies" did 5 years ago. However, the aims have changed over the years, from FP5 to FP7. Partly because content has become more interactive, expressive, self-aware and adaptive. The changes are also due to merging areas like content creation, digital libraries, eLearning and semantic web technologies. Furthermore it takes into account current trends in content production and consumption and particularly the move from few-to-many to many-to-many models. This objective targets results in the medium and in the longer term to solve the main challenges of Semantic Systems research.

Besides the specific objective 3.4.2.1, several other objectives indicate relations to the Semantic Systems field, e.g.:

- 3.4.1.1 Digital libraries and technology-enhanced learning
- 3.5.2.1 Advanced ICT for risk assessment and patient safety
- 3.6.1.1 ICT for the intelligent vehicles and mobility services
- 3.8.3.3 ICT forever yours
- 3.7.1.1 ICT and ageing
- 3.7.2.1 Accessible and Inclusive ICT
- 3.2.1.1 Cognitive systems, interaction, robotics
- 3.1.1.4 Networked media
- 3.3.2.3 Networked embedded and control systems

At present on the European level several technology fields are being promoted as foci and core technologies: they are given special attention and activity through the establishment of European Technology Platforms (ETP). ETPs provide a framework for stakeholders, led by industry, to define research and development priorities, timeframes and action plans. The main objective is the formulation of a joint „Strategic Research Agenda", defining strategically important issues where achieving Europe's future growth, competitiveness and sustainability objectives is dependent upon major research and technological advances in the medium to long term.

The ETP with closest connection to the Semantic Systems area is "NESSI" (Networked European Software and Service Initiative). NESSI has been established by 13 „promoters": ATOS Origin, British Telecom, Engineering Ingegneria Informatica, IBM, Hewlett-Packard, NOKIA, ObjectWeb, SAP, Siemens, Software AG, Telecom Italia, Telefonica, Thales. NESSI envisages the wide adoption of services offered through networks and devices as a distributed system which in turn provides virtualisation. A crucial piece is a software infrastructure middleware facilitating a seamless and cost-effective composition of services.

Semantic Systems are not only strongly driven by standardisation bodies, but also by initiatives like ETPs and national programmes. It is important that the area not only spans related technologies or applications, but also keeps pushing to establish "Semantic Systems" as known and prominent technology field.

3. Semantic Systems Research in Austria

Since 2004, "Semantic Systems and Services" has been a thematic priority of the Austrian national funding programme for IT research "FIT-IT", which is funded by the Federal Ministry for Transport, Innovation and Technology (BMVIT). FIT-IT targets the co-operation of industry and academia and addresses projects with a time-to-market in the range of three to eight years. The programme aims at pushing enterprises towards a longer-term horizon in their research activities for developing their innovative products and services. Further aims are to increase the competitiveness of Austrian IT research actors – both companies and research institutions, to train young researchers, and to increase the international visibility of Austrian IT research.

The Austrian research scene in Semantic Systems is characterised by a dominant position of the universities on the one hand, and by the predominantly small and medium-sized structure of enterprises carrying out research on the other – only few large companies are active in Semantic Systems research. In general, Austrian universities can raise funds for projects in basic research from the Austrian Science Fund (FWF), while another growing part of research funding comes from contract research for the public or the private sector (external funding).

Until around the year 2000, there was a weakness in the Austrian research funding system that consisted in a lack of funding schemes for co-operative projects involving both university researchers and private companies. Another shortcoming concerned funding for longer-term company research and longer-term co-operative activities. Therefore several initiatives and programmes of the last years were designed to address this issue. The first such programme was FIT-IT, which was started by BMVIT in 2002 with the goal to support industry-academia co-operation in selected areas of information technology and to promote high-quality research with a longer market perspective.

The co-operation of industry and universities in research projects is particularly important for longer-term research projects: university research becomes informed by market signals, companies gain access to high-risk long-term projects and research expertise that would be too expensive if kept in-house. Consequently by now several funding schemes have been created in Austria that specifically target research co-operation, such as the competence centre programmes ("K-programmes", relaunched in 2006 under the label COMET), the translational research programmes (including the BRIDGE-programme of the FFG), the research studios programme and various smaller schemes. Various aspects of IT and of Semantic Systems research are treated in individual projects in these funding schemes, although none of the schemes has a comparable thematic focus on the topic as FIT-IT.

The largest share of direct Austrian public subsidies for RTD in IT is however disbursed not by the funding schemes for co-operative research or by the FWF, but rather by the so-called General Programme (Basisprogramm) of the Austrian Research Promotion Agency FFG. The General Programme funds technologically innovative projects with a clear market perspective in the 2-3 years time-frame. These are usually single-firm projects with a somewhat lower research risk. Accordingly the level of public subsidy is significantly lower (typically around 22% of project costs) than in the co-operative programmes. The General Programme is a bottom-up programme that accepts projects from all branches of industry, not just IT.

When it was proposed in 2004 to the Austrian Council for Research and Technology Development (Rat-FTE)[1] to establish "Semantic Web" as a new thematic priority in the FIT-IT programme, this was done in conformance to the following criteria for a suitable topic for such a thematic priority:

- Trend-setting: The priority should focus on topics which can be estimated to remain core themes of research in information and communication technologies for at least another 10 years.
- Excellence: Austria should possess high-quality researchers in the topical area with a significant impact at an international level.
- Economic relevance: The programme should define areas with an existing economic potential based on already established companies in Austria.
- Impact: The budgetary amounts available for FIT-IT should allow for a significant contribution to a topic with respect to the anticipated number of projects and the project funds necessary.

At the time of the initiative, it was judged that the topic possessed strong potential in terms of research capacity, opportunities for co-operation (both companies and universities were active in the area), impact and market potential. These findings were confirmed further by a questionnaire study, in which 41 experts participated.

FIT-IT offers a range of instruments to reach the programme goals. The most important ones are two types of co-operative research projects with a medium (3-8 years) time-to-market perspective: one has a funding ratio of 75% of project costs, the other offers a funding ratio 50% in cases where the company partners consume the larger part of the project budget. The other available instruments are dissertation scholarships and accompanying measures (e.g. networking activities).

During the years 2004-2006 the FIT-IT programme has launched three competitive calls in the thematic priority "Semantic Systems and Services". The projects were evaluated by international experts. A total of 31 have been selected for funding with total funding of 8.85 million Euro and total project costs of 11.88 million Euro.

[1] The Austrian Council has the task to advise the federal government, the ministers and the "Länder" in all matters related to research, technology and innovation, and defines long-term national RTD strategy.

Until 2006, 125 different institutions have participated in co-operative proposals for funding in FIT-IT Semantic Systems. Of these, 79 have been successful at least once: 59 institutions were funded once, 15 were funded twice, four were funded three times, and one institution has been successful with eight applications. This pattern is consistent with the goal to involve new institutions in the programme and to promote new co-operations, while at the same time recognising and promoting excellence. As Illustration 1 shows, with stable call budgets of approx. 3 million Euro per year the total number of funded institutions has remained fairly constant over the three calls, but the proportion of companies in the programme has increased (2004: 11, 2005: 14, 2006: 19).

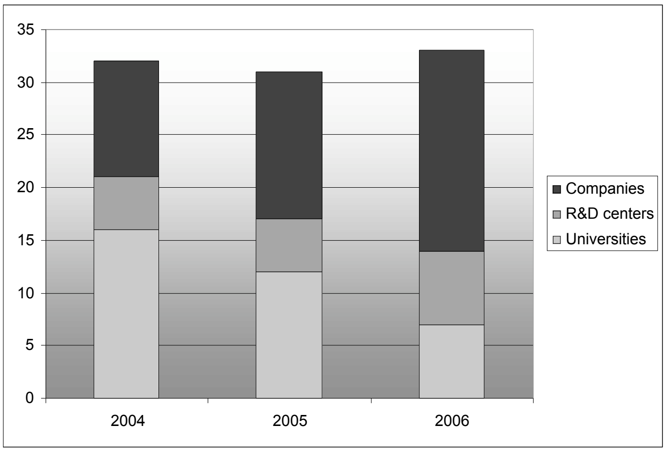


Illustration 1: FIT-IT Semantic Systems - Number and type of funded institutions

In terms of project content, the projects that have been funded in FIT-IT Semantic Systems can be categorized in 4 thematic areas:

- ***Organisation and access to unstructured data***. Semantic technology in these projects helps users to deal with unstructured information on the web, provides foundations for semantically aware file-systems or makes it easier to manage ontologies.

- ***Integration of distributed information***. Numerous challenges that arise from information that is distributed among different locations, channels or media types are addressed with new semantic approaches.

- ***Human-machine-interaction***. Semantic technologies at the human-machine interface make it easier to interact with applications, and they make it easier to develop user interfaces that are flexible and adaptive – but often, hard problems remain to be solved and motivate the research undertaken here.

- ***Automation of information processes***. In these projects, semantically aware components enable new automation steps that make it possible to build previously unconceivable network applications such as Semantic Web Services. Projects in this group develop technologies e.g. for semantic machine-machine-interfaces, semantic filtering, or semantic modelling of business processes.

By autumn 2006, first mid-term reviews of projects in FIT-IT Semantic Systems have shown promising progress in technology research, and some projects exhibit a promising outlook on the market. However, none of the projects has been finalised yet, and it will take at least another year to have results at hand.

4. Looking ahead

Innovation in software technologies is a broad phenomenon of high importance for satisfactory economic performance and growth in the future. Public funding of research in software technology comes in many shapes and at different levels, among them the European and the national level. As we have seen in the preceding sections, present reality is characterised by a mix of subsidies that are unspecific to any particular technology topic, and of targeted top-down programmes that focus on a selected set of technological or societal challenges. For the foreseeable future Semantic Systems and semantic technologies will find their place in a number of different top-down initiatives on the European level due to their horizontal characteristic and generic applicability. On the Austrian national level, the FIT-IT programme line in Semantic Systems and Services provides a distinct centre of gravity for research, complemented by national and regional bottom-up funding schemes.

At the time of writing, when the FIT-IT programme line has completed its second year and its third call for project proposals, the perspectives for Austrian Semantic Systems research are bright, while important challenges lie ahead.

To start with, the topic “Semantic Systems and Services” has been successfully established as a brand in the Austrian research funding system. This provides visibility and access to audiences in research policy at the national level. Continued efforts will be necessary to position the topic prominently in the European discussion, requiring among others a further strengthening of collaboration and coherent awareness measures by Austrian actors. The establishment of a successful international conference in the area of Semantic Systems creates welcome opportunities here.

The second major element for a successful Semantic Systems strategy in Austria continues to be excellent research, both in depth and breadth. The top-performing research actors that fortunately exist in Austria must be strengthened further so that their visibility on the European and global scale is ensured. It will be mainly their own responsibility however to identify co-operating partners both in the national and international Semantic Systems community that can increase the reach of Austrian Semantic Systems research in terms of both technology and long-term business opportunities. In addition it will be necessary to foster a broad base of smaller innovative actors, in particular start-up companies, to keep and further increase the dynamic in the sector. Support for SMEs must be unequivocal, while public intervention in the field should not go so far as to obscure market signals that are vital for the optimal strategic orientation of small firms in particular.

Successful exploitation of opportunities at the European level will be another key element. Beyond the chances offered by the 7th Framework Programme, the European Technology Platform (ETP) NESSI is an important instrument for defining the topic and future development, but also for generating measures and support at the level of research policy. It needs to be kept in mind however that NESSI is interdisciplinary and that Semantic Systems is only one of several topics. According to the definition of an ETP, thematic priorities and technology strategies shall be determined and reinforced with the participation of all stakeholders. The industry representatives in NESSI have strong affinity to Semantic Systems and will promote the topic with the European Commission and with national policy-makers.

In addition to the research community and European commission, three additional types of actors have a strong influence on Technology Policy: national governments and their representatives, national research promotion agencies and their initiatives, as well as supranational interest groups and NGOs.

In the next three to five years, the ultimate test – and the ultimate challenge – for all public and private initiatives in the area of Semantic Systems will lie in the transfer of the technology into systems, services and applications that will enter into the public sphere and business. This has not yet been accomplished except in small niches and calls for a combination of activities in several dimensions. The profile of the topic Semantic Systems will have to be accentuated further in the conceptual, economic, legal and socio-cultural realm. Awareness measures will only succeed if they can be complemented by convincing business cases. Remaining technical issues in the areas of trust, interoperability, availability, correctness, maintenance, upgradeability, and interfaces of the technology will have to be solved. This will require a persistent strong engagement of the research actors – companies and university researchers – but also of research policy actors and public research funding programmes.

5. Summary

Semantic Systems today is not just an important and promising technology field, it also calls for public intervention in the innovation system to help research actors to overcome several challenging issues. The 7th Framework Programme of the European Commission will provide funding opportunities for Semantic Systems research, but they are not very focussed and will require energetic efforts by research actors to fulfil their ambitions. At the national level, the thematic priority of the IT-research programme FIT-IT on Semantic Systems and Services provides favourable conditions for companies and university groups. Yet, also here a lot still needs to be done if the ultimate test for Semantic Systems is to be passed: in several years time, Semantic Systems technology shall be not just an exciting technology field, but a massive IT-market.

ON THE SITUATION OF SEMANTIC TECHNOLOGIES IN AUSTRIA

Andreas Blumauer (Semantic Web School), Gisela Dösinger (Know-Center), Thomas Fundneider (tf consulting)

Abstract
The contribution at hand presents the results of one central work package within the project Semantic Knowledge Base, in short sem'base, carried out by tf consulting, Know-Center, factline, and Semantic Web School. The work package consisted of a web-based mass survey and expert interviews, both aiming to highlight the situation of semantic technologies in Austria. The main purpose of the web based mass survey was to identify the need for action concerning information related challenges and market coverage, utilised information sources, and the familiarity with semantic technologies from the users´ point of view. In the expert interviews, representatives from key organisations within the sector were, amongst others, asked about education & training facilities, ways of knowledge acquisition, maturity, application, and applicability of semantic technologies, and funding initiatives. To summarise the results: even though the potential of semantic technologies is not fully exploited yet – especially knowledge of potential customers/users about possible applications and their applicability is lacking – organisations within the sector have prepared a sound and fruitful basis on which further activities can build on.

1. Introduction

One of the goals of the project sem'base, amongst others, was to assess the characteristics of the semantic systems sector concerning its activities, achievements and failures, its strengths and weaknesses, and its demands and offers. In order to reach this goal, two instruments were applied: key organisations were interviewed and a web based mass survey was carried out. On the basis of the thereby achieved results, conclusions were drawn which might serve as the foundation for further activities advancing semantic technologies in Austria.
In the following section, the web based mass survey and the interviews are described alongside method, samples, results, and conclusions.

2. Web based mass survey

2.1. Method

In order to obtain quantitative data about organisational and technological challenges that organisations will face, an online questionnaire was developed. It contained 22 questions out of which 14 covered issues directly related to the subject matter of the project. The other 8 questions asked for information about the respondents and the organisations they represented. In the following, an overview of the topics which were covered by the questions is presented.

- Perceived need for action concerning information related challenges and rating of market coverage
- Utilised information systems and their potential for optimisation
- Used information sources and rating of their adequacy
- Institutions exploited for finding solutions/systems and rating of their adequacy
- Familiarity with semantic technologies and definition of this term

For the main survey altogether about 800 persons were directly contacted via email and asked to fill in the questionnaire. The questionnaire was accessible from July 10, 2006 to August 31, 2006. The rate of return was about 20%, but finally only 68 questionnaires were able being analysed, since they have been completed.
In the following sections sample characteristics and results are given. They are illustrated, not interpreted. Conclusions will be drawn thereafter.

2.2. Sample characteristics

To start with, it should be mentioned that the characteristics of the sample – which in the following are reported in detail – describe the group of respondents. The portrayed demographic structure should not be misinterpreted as highlighting certain groups which are most interested and engaged in information issues and technologies, since it might be an artefact of the composition of the mailing lists biased towards the reported characteristics.

2.2.1. Region, age, and sex

According to the used mailing lists all respondents were German speaking, most of them coming from Austria [77%]. The focus on Austria was intended, since the project mainly aimed at clarifying questions concerning semantic technologies in Austria. About three fourths of the respondents were male [72%] and the age group of 30-39 years was most strongly represented [49%].

2.2.2. Sectors, organisational size, areas and roles

As can be seen from Figure 1, most of the respondents belonged to the sectors of research and science [22%], management and IT consulting [18%], education and further training [10%], and transport and logistics [10%].

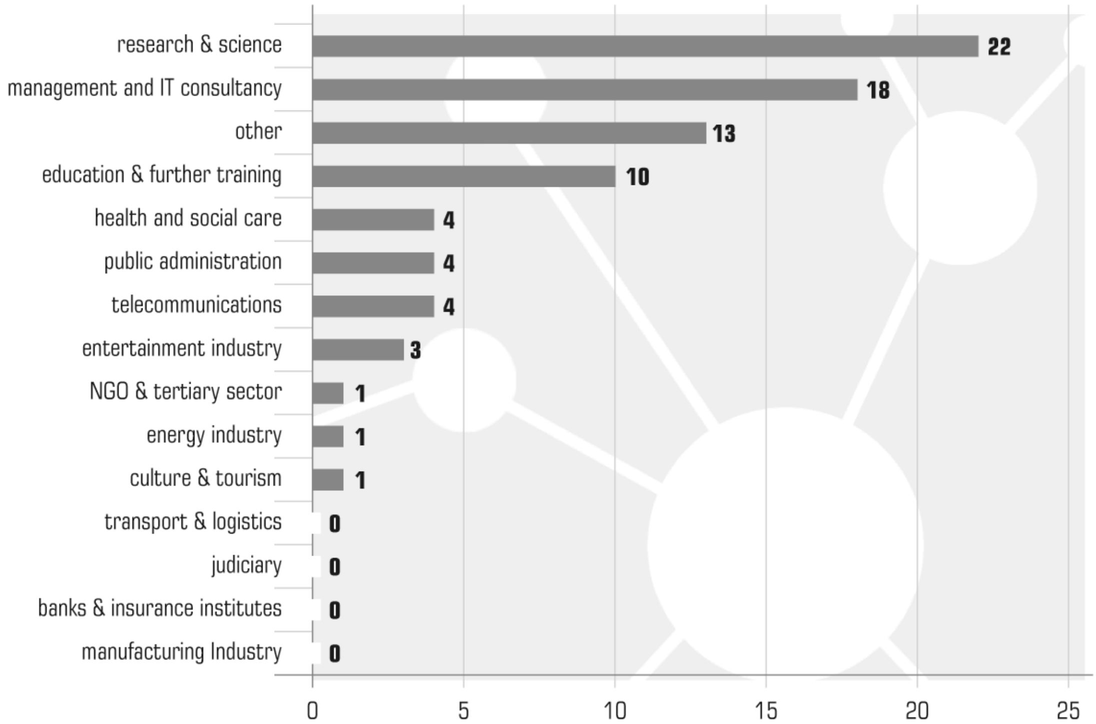


Illustration 1: Distribution of sectors [%]

About half of the respondents represented organisations with 2 to 49 employees [46%] and the other half represented organisations which employ between 50 and more than 250 employees [50%].

Two organisational areas were most strongly represented. About one fifth of the respondents is employed in research departments [21%], another fifth is employed in the area of organisational development or business management [21%]. About one fourth could not assign oneself to the predefined categories. Concerning their organisational roles, the majority of the respondents takes higher or leading positions [63%]. About one third of the respondents is involved in project management [29%], another third distributes almost equally to department management [16%] and business management [18%].

2.3. Results

In the following section, the results which directly relate to the subject matter of the project are reported in a descriptive way. An evaluation alongside certain characteristics of the sample was not possible, since the sample was too small for performing statistical tests. Nevertheless, tendencies can be observed and conclusions can be drawn. But these assertions always must be understood against the background of the earlier described sample characteristics.

2.3.1. Challenges concerning information issues, technologies and market coverage

In the questionnaire, it was assessed in which aspects regarding informational and technological issues the respondents saw a need for action. Further, it was intended to identify how well the market answered the challenges according to the opinion of the respondents [1].
Figure 2 lists aspects focusing on organisational challenges, while Figure 3 stronger relates to technological challenges. The aspects of increasing the utilisation and reuse of information, providing a better overview on information sources, and improving the relevance of search results were attested a high need for action. At the same time, exactly these aspects were said to be covered insufficiently by the market.

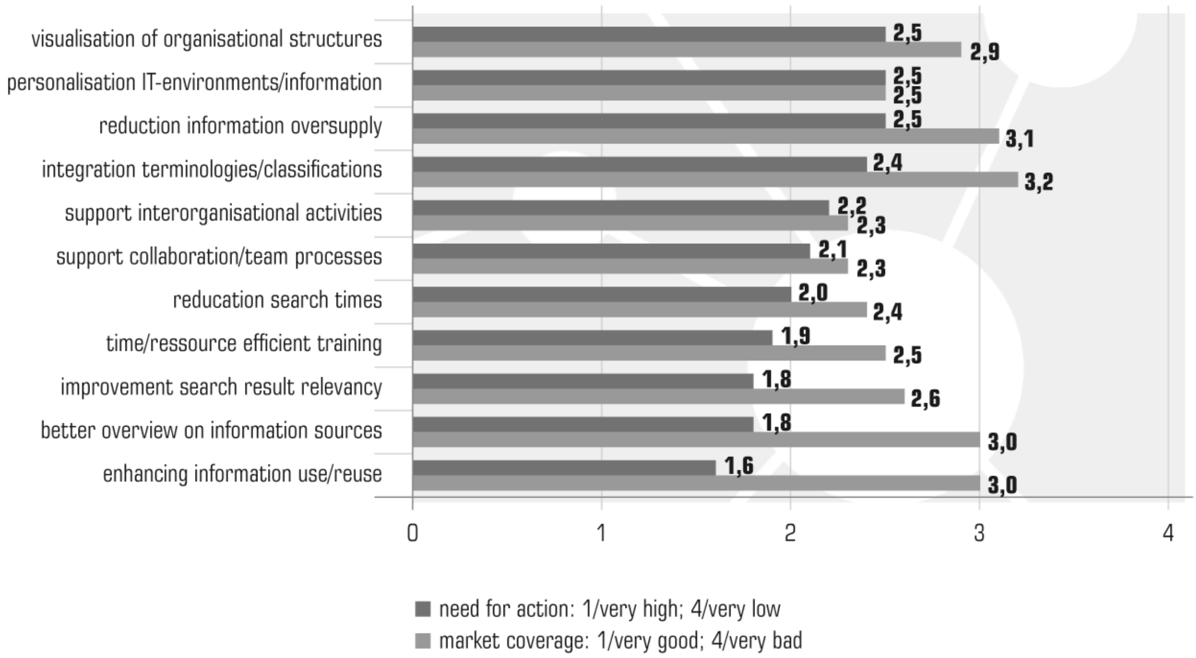


Illustration 2: Perceived need for action and market coverage concerning organisational challenges [means]

[1] The response scales reached from 1 to 4 with the extremes representing a high/low need for action and very good/bad market coverage, respectively.

Concerning the more technological challenges, a high need for action was assigned to the aspects of organising and discovering unstructured data, integrating distributed information repositories and applications, decreasing redundancies in data management, achieving data compatibility and system interoperability, as well as automating information processes and reducing the efforts for maintaining databases and portals. Out of these, only the aspects of automating information processes and integrating distributed information repositories and applications were attested a good market coverage.

A comparison between the more organisational challenges and the more technological ones shows that the need for action was judged slightly higher for organisational than for technological challenges.

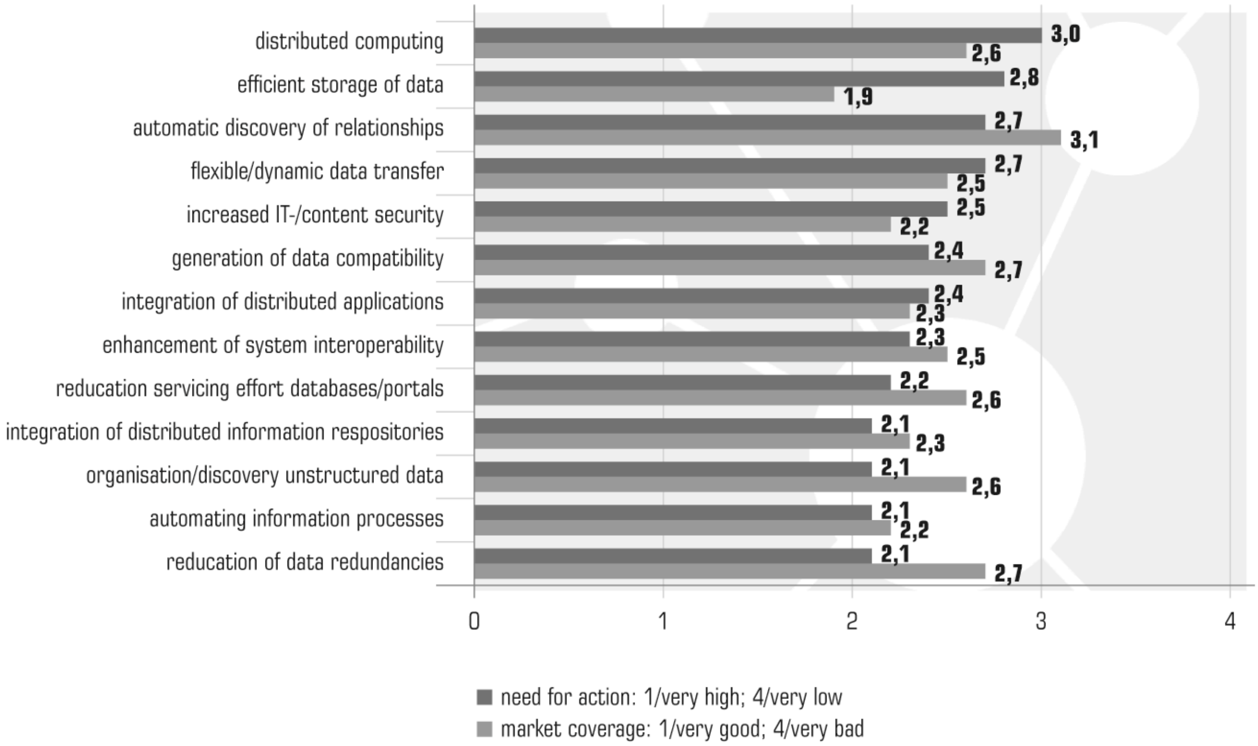


Illustration 3: Perceived need for action and market coverage concerning technological challenges [means]

2.3.2. Information sources, utilised institutions and their adequacy

As can be seen from the previous section, the respondents expressed a need for action concerning a number of aspects. In this section, it is shown how people satisfy their information needs and which institutions they contact for support.

91% of the respondents stated that they talk to colleagues and experts in their organisation. About three fourths search the Internet for information [77%] in a traditional way, meaning via websites, and about two thirds consult books and professional documents [62%]. These sources are also judged to cover information needs well. The means were 1.8, 2.0 and 1.7, respectively[2]. On the contrary, distance learning/eLearning, in-house training, and case studies are said to cover information needs insufficiently. These sources are exploited only by a minority of respondents - 12%, 13% and 21% use these sources, respectively.

In general, training and information offers were judged sufficient by more than two thirds of the respondents [37%]. 29% of the respondents see the offers as insufficient, and the same percentage cannot decide about adequacy.

[2] The response scales reached from 1 to 4 with the extremes representing a very good/bad coverage.

When people were asked for institutions to which they revert when looking for solutions and systems, most of the respondents mention informal personal contacts. As Figure 4 shows, between 35% and 40% state that they exploit university research institutions, software providers, management and IT consultants, and open source. University research institutions and integrators are not so prominent sources. Correspondingly, informal personal contacts cover the needs best. The means are 1.5 and 2.9, respectively[3].

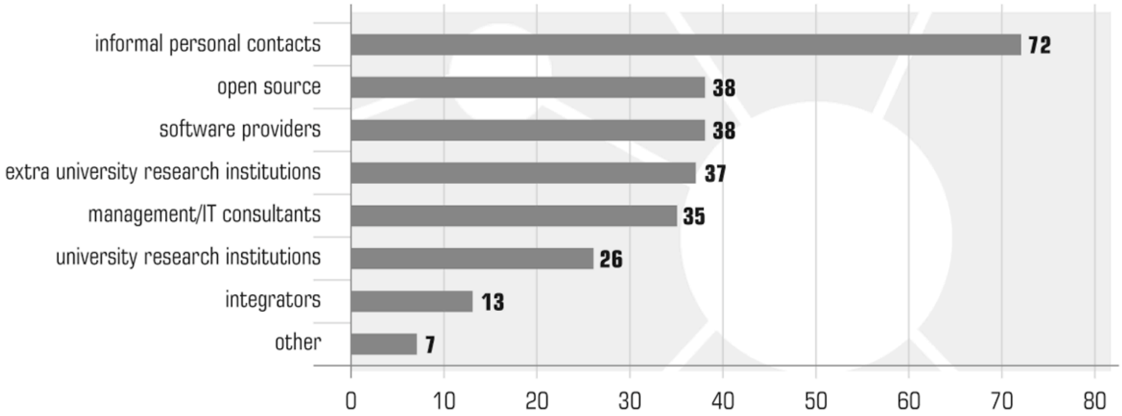


Illustration 4: Utilised institutions [%]

2.3.3. Used information systems

In the questionnaire, the respondents were also asked which information systems they use. They were presented a list from which they could choose, and the questionnaire also allowed naming further systems. In the list, the most common systems were mentioned. Even though this has nothing to do with semantic technologies at first sight, this information is important, since the commonly used systems are the basis on which further technologies will have to build upon.
Content management systems [68%] and document management systems [62%] as well as search engines and desktop search facilities [79%] are used most often.
When asked about the optimisation potential of their information systems, respondents mentioned a broad range of aspects of improvement. Next to anticipated demands, also specific demands that could be met by semantic technologies were mentioned. To point out a few: integration and consolidation of systems, integration of data bases, data exchange between data bases, cross organisational data structures, automatic deduction of relationships.

2.3.4. Familiarity with and knowledge about semantic technologies

As the previous section shows, most of the questions did not directly aim at semantic technologies. This was intended, since the authors wanted to highlight the conditions and expectations which will have to be met when semantic technologies will be put into practice. Nevertheless, in this context it is also important to know about familiarity of people with the term and how they specify semantic technologies.
When asked for familiarity, nearly 60% stated that they are confident with the term. 25% showed no familiarity, and 12% could not really decide.
In order to obtain a more precise picture about the association of the respondents with semantic technologies, they were presented a list covering standards, methods, and applications. Among those which per definition represent semantic technologies, the authors also mixed terms from other domains such as social software, service oriented architectures, or grid computing. In doing so, the aim was to find out about the respondents understanding regarding semantic technologies.

[3] The response scales reached from 1 to 4 with the extremes representing a very good/bad coverage.

As Figure 5 illustrates, the majority of the respondents, namely 74%, connect intelligent search engines/optimisation of search engines with the term semantic technologies. For 62% of the respondents, metadata modelling/annotation represent semantic technologies. About half of the respondents consider text analysis/natural language processing, ontologies/ontology modelling, knowledge visualisation/knowledge maps, reasoning/automatic inference/artificial intelligence, as well as classification systems to be the main domain of semantic technologies.

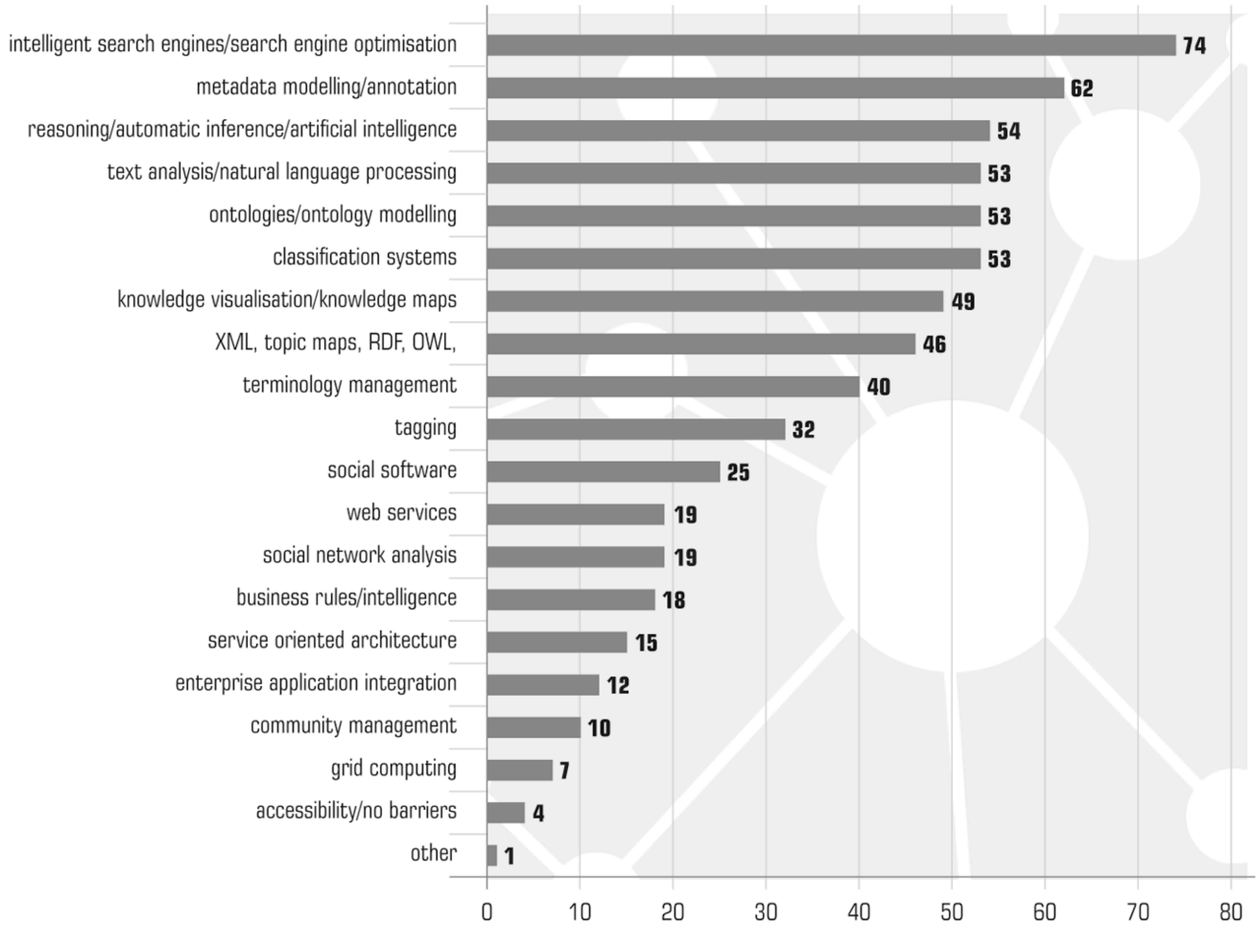


Illustration 5: Terms associated with semantic technologies [%]

2.4. Conclusion

Since the sample was small and because of its composition, the results must not be overrated. Nevertheless, useful conclusions can be drawn from the findings. For organisations which are engaged in the development and provision of semantic systems, it is important to know what the demands are. For tailoring their offers with the demands of potential customers/users, they must know whether companies have problems to which semantic technologies are the answer. The answers to the questions asking for need for action and market coverage – see section 1.3.1. – provide corresponding information. So the highest need for action is perceived concerning the following areas.

- Utilisation and reuse of information
- Better overview on information sources
- Relevance of search results
- Organisation and discovery of unstructured data
- Decrease of redundancies in data management
- Achievement of data compatibility and system interoperability
- Reduced effort for maintaining databases and portals

As this list shows, there are actually a number of problems which could be removed by semantic technologies. This result can be seen as an affirmation of activities in this sector and also gives hints into which direction these activities should go.

Concerning the search for information in connection with semantic technologies, the following sources are most often named and also judged sufficient.

- Colleagues and experts in the organisation
- Internet research
- Books and professional documents

Already a number of offers use these channels. An extensive overview of communities, websites and publications can be found on the website of the Semantic Web School. An expert search can easily be performed via http://wiki.sembase.at – the platform which was initiated by the project sem'base.

Information sources such as distance learning/eLearning, in-house training, and case studies are not used often and are also judged as not being adequate sources. It is not yet clear, whether the first two sources would be used more often if the offer is more attractive – this would require a further examination, but from the interviews which are described in the next section it could be concluded that there is a strong demand for case studies and application scenarios.

Concerning institutions which are reverted to when looking for solutions and systems, next to informal personal contacts, the following institutions are most often mentioned.

- Extra university research facilities
- Software providers
- Management and IT consultants
- Open source

University research institutions and integrators are not so prominent sources, and together with management and IT consultants they have a poor rating concerning adequacy. Reasons were not expressed, but presumably these institutions do not match client needs and wants properly, maybe in the sense of not knowing their business environment or using another language. As a consequence, one could say that these institutions may invest efforts for better meeting the market and its clients. The pressure to do so might not be as strong for university research institutions, since their focus is on basic research, which anyway is reused and transferred to industry by extra university research facilities.

So far, we have drawn some conclusions from the more general results of the survey. Now, we present the results which are directly related to semantic technologies. As it was said earlier, 60% of the respondents mentioned that they are familiar with the term semantic technologies. When asked for terms they associate with semantic technologies, between half and three fourth of the respondents name the following.

- Intelligent search engines/optimisation of search engines
- Metadata modelling/annotation
- Text analysis/natural language processing
- Ontologies/ontology modelling
- Knowledge visualisation/knowledge maps
- Reasoning/automatic inference/artificial intelligence
- Classification systems

From this list it can be concluded, that the knowledge about semantic technologies is rather good: all mentioned applications and methods are associated with core semantic technologies. But it should be emphasised that even though these results imply a positive picture about the respondents understanding, at the same time it is reported that – this information stems from the interviews – organisations do have only a very vague idea (if at all) about what semantic technologies are.

3. Expert interviews

3.1. Method

In order to obtain qualitative data about the situation of the semantic systems sector in Austria, interviews with selected key organisations were conducted. For each type of organisation, a set of questions was prepared. The interviews were conducted as semi standardised interviews which means that interviewees had the possibility to go beyond the questions when talking about semantic technologies. All the interviews – each taking about 30 to 60 minutes – were conducted either by phone or personally over a three months period, starting June 1, 2006. From each interview, key statements were extracted.

3.2. Sample characteristics

Before preparing the interview questions, the targeted organisations were identified. The following institutions were found to be relevant to the sector: educational institutions, research institutions, integrators, public institutions, software providers, operating departments, IT departments, and management of corresponding institutions. Supplementing, also organisations were considered which have not yet actively used semantic technologies.

3.3. Results

The result of the interviews consisted in a number of key statements. These key statements were analysed and conclusions were drawn. They are reported in the following section. It should be mentioned here that the conclusions are based on 20 interviews and hence do not claim for generality and representativeness. Nevertheless they give a good impression about the situation of semantic technologies in Austria.

3.4. Conclusion

3.4.1. Education & training

Regarding semantic technologies in particular, there is no topic specific academic education available. There exists a deficit even though first efforts for establishing specific studies are accomplished. Training offers for decision makers are seen in a better state, even though in Austria there is only one institution especially focussing on the topic. However, there is a number of training offers emphasising on information and knowledge management through which interested people try to by-pass the gap. Information and training needs which are already salient will become most probably much stronger in the future when semantic technologies penetrate the market.

3.4.2. Knowledge acquisition

Organisations utilise a broad palette of methods for acquiring knowledge about semantic technologies. Among these methods are: events, conventional desktop research, funding programmes, diploma and doctoral thesis, lectures, advanced training, learning-by-doing, and corresponding software. Even though people utilise a variety of methods for satisfying their information demands, it should insistently be emphasised that a lack of show cases, success stories, and application scenarios is perceived.

3.4.3. Approach to semantic technologies

Primarily, organisations approach semantic technologies in a requirement and problem driven way. Companies do not seem to consciously search for semantic technologies, but stumble across them because of corresponding problems or even by chance, for example by personal contacts or in the context of trainings. Interviewees said that they do not really know that semantic technologies exist and their idea about the topic is rather diffuse. Even though it is said that the topic is interesting, knowledge about it is rare and the potential is only intuitively grasped. Users do not approach providers in an active way. It seems that problems are not matched with the already available technologies. It can be concluded, that consultants and integrators still have a long way to communicate the benefits of semantic technologies and to build awareness among users/customers. Further, it is repeatedly stated that there is a lack of business cases. Organisations, which are not or only modestly familiar with semantic technologies, report difficulties with mapping semantic technologies to business challenges. Among these there is the requirement of structuring data, overcoming and connecting knowledge isles, concerting decentralised systems and integrating distributed data, or uncovering relationships. However, especially small and medium sized enterprises need stable, inexpensive systems and there must be a positive value/effort ratio. Currently, semantic technologies do not seem to meet these conditions.

3.4.4. Maturity of semantic technologies

The perception is two-fold. On the one hand, the opinion rules that much is talked about semantic technologies but little has been realised yet. Semantic technologies are assigned a lack of market diffusion, only a small number of integrators can handle semantic technologies. In brief, the technology is only in its beginnings. On the other hand, it is reported that a number of software providers deal with semantic technologies and that there are mature technologies in the open source domain as well as in the commercial domain. In brief, semantic technologies are said to be reality. Nevertheless, the first perspective – semantic technologies are only in their beginnings – seems to be more prevalent.

3.4.5. Applicability of semantic technologies

The following can be said according to the responses from the interviewees. There are no definite sectors or organisational areas which especially benefit from semantic technologies. Rather, the technology can be broadly applied, a wide range of sectors is mentioned. This would indicate a high potential for value generation of semantic technologies, resulting from their broad applicability. To name a few, the following issues have been highlighted in connection with semantic technologies´ exploitation: intellectual property reuse, unstructured information, lacking connections, employee turnover, intersection of professional knowledge and knowledge about customers, citizen portals, information integration, plagiarism.

3.4.6. Experienced barriers

Since there are certain barriers, people still have to be informed further. So, the term semantic technologies implies something unfamiliar which is not graspable. Additionally, semantic technologies have the flair of artificial intelligence. People worry that this technology is still in its research stage. Reference projects, show cases and specific applications which could convince customers are lacking. Moreover, experts often cannot illustrate their domain in an understandable way .

3.4.7. Networking activities

Companies, which work in the analysed sector perform active networking, seeing networking as one core principle, presumably because of the newness of the sector which requires to bundle forces. It is reported that in Austria specific organisations that offer the required expertise can be found quickly. For developing tailored solutions, complementary competencies are integrated across organisations. There is also intensive contact with research institutions. The value of research institutions is seen in their function of providing input. By cooperating with them, companies can identify current and future trends and developments more convenient. Altogether, networking generates advantages of location and contributes to the fitness in international competition.

3.4.8. Funding initiatives

Even though on the organisational level semantic technologies are no strategic topic yet – as mentioned, the approach to the topic is requirement driven – relevance is lent to the topic via the FIT-IT

action line Semantic Systems. This initiative is carried out by the Federal Ministry of Transport, Innovation and Technology together with the Austrian Research Promotion Agency, with feedbacks to the Austrian Council for Research and Technology Development.
Concerning funding, it is mentioned that only technologies but not solutions of knowledge transfer activities are funded. FIT-IT, especially the action line Semantic Systems, is positively judged. This well advised funding initiative contributes to the fact that Austria is quite strong on the sector of semantic technologies, which are also thought to contribute to the innovativeness of Austria.

4. Final statement

Even though the investigation yielded a number of interesting results which will incite discussions and considerations as to further steps, there were two most striking results which might imply future activities. The two facts are given in the following.

- Even though organisations report a number of problems which can be solved by semantic technologies and they seem to know very well what semantic technologies are all about – see sections 2.3.1. and 2.3.4. – they do not understand well about the practical application as well as applicability of semantic technologies. Therefore, the necessity for case studies and application scenarios was repeatedly and insistently claimed.

- For acquiring knowledge about semantic technologies, organisations utilise a broad range of information sources which altogether cover the information needs. Methods reach from conventional desktop research to participation in funding programmes. Nevertheless, one major gap exists: there is no topic specific academic education yet.

Part II

Individual Expert Stories

Part II – Chapter A

Concrete Applications & Scenarios

A POLICY AWARE MOBILE WEB SCENARIO

Joachim Zeiß

FTW actively participates in Magnet Beyond, an EU-funded research project, dealing with personal networks and their federations. A ubiquitous mobile computing scenario shows a user equipped with mobile devices in her vicinity. Those devices are connected to each other while talking to distant home servers. As people meet each other, their personal networks may join predefined or ad-hoc federations to share resources and services. Within this setup we focus on semantic modeling of distributed policies for access control to resources and services, and privacy protection of user data.

To illustrate the possible use of these technologies, we have chosen as scenario the management of ftw. tutorials on research topics of interest for our partner companies. Our participation policy says that each partner company may delegate a certain number of employees for participation without paying additional fees. Interested ftw. employees not participating as speakers may attend the tutorials as long as there are places left.

Once the tutorial has been setup by the research team, the back office is responsible for announcing the event offers a registration procedure, writes a participation list, prints handouts, etc. In the mean time, the research team and IT support setup network connections and access rights on the demo servers. Finally, on the day of the event, a concierge verifies and enforces participation policies, fills in lists, counts people, distributes badges etc.

Well, couldn‘t we all have spend our precious time with something more exciting than formatting web pages, fixing wrong access rights or creating and - once the party is over - deleting temporary accounts? A pragmatic solution is to define only one account for all participants granting access usually beyond what is really required or even acceptable from security point of view. One ring to rule them all? Well, we know how the story ends.

Based on our work in the Magnet Beyond project, we could express polices by means of the Semantic Web, a machine readable set of distributed ontologies, facts and rules, instead of tediously predefining access roles and identities, an error prone procedure that has to be revoked after demo has finished. If our partner companies provide their data in the same way (in a trusted policy aware manner) a reasoning program could handle access rights and participation policies automatically. After all, the companies should know their employees better than FTW.

But what happens to our concierge at the reception desk? She is left alone to govern arriving visitors. In a future scenario, participants would bring their Semantic Web, policy aware mobile devices. When entering the meeting room devices will communicate with a reception computer to authenticate their owners. The hosting program, in turn will publish printer and networking services and grant access to the demo machines. Preferences and user context will be protected by policy engines in the same way as the access to services. While socializing with other participants our devices automatically exchange vcardsor create ad-hoc networks to enrich social interactions.

Organisation	ftw. Forschungszentrum Telekommunikation Wien Betriebs-GmbH
Core competencies	Semantic Web for Policy Aware Mobile Networking
Address	Tech Gate, Donau-City-Straße 1, Vienna
Contact	+43-(0)1-505 2830-62 ǀ zeiss@ftw.at ǀ www.ftw.at

ENRICHMENT OF CLINICAL PRACTICE GUIDELINES BY SEMANTICS

Katharina Kaiser, Silvia Miksch

Medical science is an area, which offers a great potential for the application of computer-support. Amongst others a huge amount of free text documents exist that could be used in a far better and more efficient way, if they are better structured, categorized, and computer-interpretable.

In our group we deal with so called clinical practice guidelines (CPGs), ”systematically developed statements to assist practitioners and patient decisions about appropriate healthcare for specific circumstances“ [1]. To optimally support the medical personnel approaches exist to transform these documents in a computer-interpretable (and thus -executable) format – a cumbersome and time-consuming task: it demands for detailed medical knowledge and knowledge of the complex formal methods.

We try to develop methods that automate parts of this task. Therefore, we also use semantic technologies, such as ontologies, in order to automatically process the documents, for instance, by using Information Extraction techniques. Due to the complexity of the medical area the structuring and description of the medical knowledge has already started very early. In the meantime, endless medical ontologies exist used by both physicians and computer scientists. We utilize these ontologies to semantically enrich CPGs in order to apply rules to structure and formalize them.

Thereby, ontologies provide us with the domain knowledge so we (as non-physicians) can understand the texts. Without this domain knowledge we would not be able to process the documents. A number of methods for formalizing CPGs exist that still fall back on the assistance of physicians for that part. But they are expensive and often manual semantic annotation is less accurate and it relies on the experts’ skills. Therefore, using semantic systems implicates a great advantage and facilitation. Text data evolves to information. Due to the semantic enrichment of documents they not only become computer-interpretable, but also comprehensible for non-experts. Furthermore, errors may be easier detectable.

Today, we are able to automate the modeling of CPGs in a computer-interpretable format using semantic technologies. In the near future, we will probably be able to unambiguously annotate very context-specific text data by semantics. This is especially important in such highly complex domains as the medical science. The next great milestone will be an almost automatic processing of text data to “understand” textual information. For our work that may mean the integration of computer-interpretable CPGs in the daily routine and more precisely in patient data management systems (PDMSs). Furthermore, an almost automatic generation of computer-interpretable CPGs in terms of “living guidelines” may be possible. A “living guideline” is one that remains under scientific review on an ongoing basis, with updates published at set intervals (e.g., annually), to present up-to-date and state-of-the-art knowledge. To always apply the latest version of a CPG it is important to have it implemented immediately after publication.

Organisation	Danube University Krems - Department of Information and Knowledge Engineering
Core competencies	Information Extraction, Information Transformation, Ontologies
Address	Dr.-Karl-Dorrek-Straße 30, Krems
Contact	+43-(0)2732-893-2450 | +43-(0)1-58801-18839 silvia@ifs.tuwien.ac.at | kaiser@ifs.tuwien.ac.at www.donau-uni.ac.at/ike | ieg.ifs.tuwien.ac.at

SEMANTIC WEB TECHNOLOGIES – A RADICAL SHIFT FOR E-COMMERCE

Markus Linder

Semantic Web Technologies enable machines to interpret data published in a machine-interpretable form on the web. At the present time, only human beings are able to understand the product information published online. However, semantic web technologies give digital assistants and agents the ability to search the web for products that correspond best to the specific needs of a certain user. While consumers today have to rely on the limited number of offers available on centralized e-commerce portals when looking for products, future applications will be able to provide users with a search process based on product attributes, which will include all products published in this form on the Internet. Furthermore, in a next step semantic web services will enable digital assistants to handle business processes like selling and buying or even negotiations automatically.

Using W3C standardized languages like RDF and OWL, manufacturers can publish their product and service descriptions in a standardized form based on product- and service ontologies on their web space. Similarly, it allows vendors to describe their offers in a machine interpretable form by using offer ontologies. Consequently, all search engines and digital assistants based on Semantic Web technology are able to find, retrieve and interpret products and offer information for their users. As these engines are able to "understand" this structured information, they are able to compare the products and offers based on specific attributes for the individual user.

Today, consumers can make use of centralized e-commerce platforms to compare prices of various dealers for a certain product. Only a few of these platforms allow consumers to compare products of different manufacturers based on detailed product features. In contrast, Semantic Web Technologies provide the opportunity to include all offers which are published in a machine-interpretable form online into a product attribute-based search process. This will increase the market transparency and maximise the consumers' chances to find the best and cheapest offer. At the same time the new technology will offer manufacturers and vendors a major opportunity to target niche customers in a more efficient way because digital assistants can evaluate and choose the offers that fit the customer's requirements best.

In the past, several technological and organizational innovations leading to a better aggregation of offer and demand on the markets could be witnessed. Market innovations like the introduction of stock exchanges or the World Wide Web have given the market participants a better overview of the current market situation. The latest step towards a higher degree of efficiency in the global market place was the introduction of centralized e-commerce platforms. The next generation market will allow customers to get the complete picture of supply and demand and will enable them to take effectively every single offer on the market under consideration with the help of digital assistants. These technological changes will lay the foundation for a more efficient market which is a key variable towards increased growth and employment.

Organisation	Smart Information Systems GmbH	
Core competencies	Semantic Web based E-Commerce, Digital Assistants	
Address	Rudolf Sallinger Platz 1, Vienna	
Contact	+43-(0)1-7157267-711 ǀ ml@smart-infosys.com www.smart-infosys.com	

SEMANTICS ON-THE-ROAD

Bernhard Lamprecht, Simone Fuchs, Kyandoghere Kyamakya

We think that semantic knowledge is necessary for intelligent vehicles. Present-day driver assistance systems, like automated braking system or automatic cruise control, are quite dumb. They are measuring speed, distances, etc. but they aren't able to act intelligent in the driving environment. They were designed for improving safety or comfort. They are not concerned with the overall driving task. Therefore, they know little about the overall driving environment, except what's absolutely necessary to fulfil their tasks. We do expect more from a modern driver assistance system (DAS), especially more intelligence. We would like to have an assistant (a driver co-pilot), which is able to support us like a driving instructor, or even take over-driving for us completely, if we want it to.

In our opinion, exploiting semantics is the enabling factor for such improved advanced DAS and autonomous vehicles. The analysis and fusion of different sensing technologies, followed by context-sensitive data interpretation on a semantic, descriptive level holds huge potential for safe navigation of intelligent vehicles. The introduction of semantic knowledge to advanced DAS will enable cars to drive nearly autonomous in the next century. User intervention will only be required in exceptional circumstances. Many accidents could be avoided by intelligent driving agents. We see the challenge in teaching semantics to a computer system. What is done intuitively by humans all the time (perceiving environment, assessing necessary actions…) is a difficult task for a computer. Making a machine intelligent enough to recognize and correctly interpret a dynamic environment is a fascinating task.

We will emphasize this statement by giving an example, how semantic influences decisions of intelligent driving agents: The system detects an obstacle on its lane. To act intelligent it must know what the object really is. By recognizing the object type and only its presence, we are able to derive suitable behaviour. For example, we recognize a small obstacle on our lane. If the object is a small human being, then we will brake immediately and try to stop before overrunning him/her, even if this abrupt braking maneuver may provoke a rear-end collision. However, if the recognized obstacle is, for example, a small animal, like a dog or a squirrel, it would be better not to brake, because saving a squirrel's life would not justify provoking an accident with possible personal injury of a human. As one can see from this simple example, semantic knowledge is the only way, to bring more intelligence into driving assistants. Perhaps, in some decades in the future, humans will only drive by themselves to enjoy the fun of it. In case people will not be in the mood to drive by themselves they will just delegate driving to the little semantic-based helper.

Organisation	Klagenfurt University - Research Group Transportation Informatics	
Core competencies	Intelligent Systems in Transportation, Multi-agent Decision Support Systems	
Address	Lakeside B02, Klagenfurt	
Contact	+43-(0)463-2700-3548 I kyandoghere.kyamakya@uni-klu.ac.at www.uni-klu.ac.at	

THE APA-ONLINEMANAGER - SEMANTIC RETRIEVAL AT AUSTRIA'S LARGEST DATABASE HOST

Manfred Mitterholzer

With currently over 82 million documents, the Austrian Press Agency group operates the largest database host in Austria and one of largest in Europe. To date the APA database host is the only host providing semantic retrieval features like topic clustering, similarity search and different kinds of search result visualization in a real-world application.

The host offers 250 multimedia sources continuously updated in real time accessed concurrently by more than 1.800 professional users at the peak hour. Among these sources are all Austrian daily newspapers and periodicals, major German publications and the APA general news service as far back as 1955. Version 6 of the APA-OnlineManager (AOM) released in 2005 provides a vast array of next-generation semantic tools supporting the user at every stage of the retrieval process. At the backend of AOM is the high performance database system AOM PowerSearch, developed by APA-IT.

In 1988 APA already dealt with the subjects of electronic media archives. Since then APA could build up substantial expertise and gained a broad level of practical experience. Based on increasing requirements and consistently expanding volumes of data by the end of the millennium the software supplier of the program licensed could not match the needs anymore. Response times were increasing and it was getting harder and harder for the user to find relevant documents satisfying their information needs. No significant enhancements appeared within the software program used. Therefore in 2002 APA was confronted with a create-or-by-decision. After extensive market research, trials and test series the company decided to develop its own full-text retrieval engine. The first version of AOM PowerSearch developed by APA-IT went into production in September of 2004. PowerSearch is a full-featured state-of-the-art search engine. Beyond its “traditional” text capabilities it offers powerful semantic retrieval functionality.

One of the essential semantic features is topic-clustering. This approach does not rely on sophisticated underlying structures, but calculates relevant taxonomic clusters from text on-the-fly. The throughput of clustering is extremely high and achieves nearly 2000 documents per second. Particularly interesting is the fact that knowledge not recognized yet first becomes visible by clustering documents. For deeper insights large clusters can be drilled-down forward. User can directly step into relevant topics. Visualization of search results is another key functionality of PowerSearch. Based upon similarity, clusters and timelines expressive 2- and 3-dimensional diagrams are generated in real time allowing the user to get a visual overview of the results and to navigate deeper into the knowledge space represented by the result.

In 2005 with dpa Deutsche Presse Agentur the first international media corporation joined and recently the publishing houses Gruner & Jahr and Süddeutsche Zeitung/DIZ moved their media archives to the AOM PowerSearch platform.

Organisation	APA-IT Informations Technologie GmbH
Core competencies	Large Scale Fulltext Retrieval, Topic Clustering, Named Entity Recognition, Search Result Visualization
Address	Laimgrubergasse 10, Vienna
Contact	+43-(0)1-360 60-6320 ꟾ manfred.mitterholzer@apa.at www.apa-it.at

Part II – Chapter B

Semantic Technologies viewed from different Perspectives

MEDIA SEMANTICS:
WHAT YOU MEAN IS WHAT YOU GET

Werner Haas

In the 1980s the "What You See Is What You Get" (WYSIWYG) principle – namely that you can see on the screen a one-to-one copy of what you later get on the printed page - was a major challenge for desktop publishing tools. This has subsequently had a major impact on professionals and home users. Since then, audiovisual media have become ubiquitously and permanently present, flooding the Internet, radio, TV, handheld devices and desktop computers. Finding, selecting and re-combining the "right" pieces of media are becoming more and more time consuming tasks, frustrating the consumers and making them a significant cost factor for media production.

So far our standard methodologies to describe and search specific content, e.g. an image, video or piece of music are mostly utilizing "piggy-back" text technologies based working on metadata. Text and metadata may be manually entered, gained from optical character recognition (OCR) or from automatic speech recognition (ASR). Content Based Indexing & Retrieval (CBIR) methods are extracting meaning directly from multimedia objects. While this is relatively easy for low level features like colour, texture, pitch or volume, it is extremely difficult to extract objects or genres, to name just a few real world concepts. At this point we have to resort to knowledge. Semantic Web technologies are offering a way to formalize the knowledge available and help us in describing – and finding later on – our content in a much more user oriented way. Even more, single content objects knowing about their meaning will on the long run be able to combine themselves on the fly into meaningful sequences, according to domain needs, following established drama rules.

As a user, I am looking forward to a situation in which – e.g. searching for videos – I will get results according to my cultural and scientific background, to my mood and current situation; automatically taking into account the environment I have got at home, in the car, my mobile devices and my other social interactions on the Web.

From a research perspective we will have to integrate the best of all worlds, namely existing methods in language and speech engineering, CBIR and semantic technologies in order to satisfy the users' demands. Main challenges will be to reduce cost of semantic enrichment, how to take advantage of „Web 2.0" and how to deal with other (fundamentally new) approaches (user feedback, statistics, social science). The search for the next generation "Media Google" has just started. From my point of view, semantic technologies will help us in achieving the "What You Mean Is What You Get"-principle for the media and entertainment area. It is yet to be seen if the impact is bigger than what was achieved by "WYSIWYG" for home users and professionals.

Organisation	JOANNEUM RESEARCH Forschungsgesellschaft mbH - Institute of Information Systems & Information Management
Core competencies	Media Semantics, Integration of low-level and high-level Features, Applications for Media Production
Address	Steyrergasse 17, Graz
Contact	+43-(0)316-876-1145 | werner.haas@joanneum.at www.joanneum.at/iis

FILE SYSTEMS AND THE SEMANTIC WEB

Bernhard Schandl

The File System – Still State of the Art?

The file system remains the backbone of users' personal data storage. However, the permanently increasing volume of data requires new methods to organize, search and retrieve digital objects. Common file systems provide only very restricted support for "metadata" (folders, file names, extensions, and a minimal set of technical attributes, like modification date). One possible method to enrich file systems is the inclusion of semantic metadata annotations that are tightly coupled and stored with file content. There are numerous tools available which support users in managing specific kinds of data (e.g. pictures or videos); also, full-text search engines allow users to find text-based content quickly. However, in most cases both classes of tools lack the support for semantic annotations, and metadata is mostly stored in tool-specific closed formats. Evidence for the need of semantic annotations in file systems is provided by the efforts that operating system vendors invest in this topic, e.g. the recently suspended Microsoft WinFS or Apple Spotlight.

Extending the File System with Semantic Web Technology

The Semantic Web technology family provides a set of tools and standards that allow the formulation of metadata graphs for all kinds of objects. Although originally designed to extend the HTML-based World Wide Web, we believe that Semantic Web technology would be an ideal candidate to be applied to the problem of metadata management for file systems, a topic which we will address in the course of the SemDAV project. The files in a user's workspace can be regarded as resources, and RDF may be used to relate and annotate them in a platform- and implementation-independent way. Such relations and annotations can be defined manually by the user or automatically by applications, and could also be implicitly determined by analyzing user interactions. Moreover, the formulation of file metadata using Semantic Web technology could eliminate the tight coupling of the technical and organizational aspects, since a file would in this case be identified by an URI instead of the combination of path and file name.

The Coalescence of File Systems and the Semantic Web

In a time perspective of 5 to 10 years, we envision the realization of the Semantic Desktop, which is currently the driving force for several projects. We envision a complete replacement of current file system structures by semantic data repositories. We envision applications that serve as metadata generators (by storing metadata as the user works with files) and metadata consumers (by using metadata to find and retrieve data). Finally, we envision the replacement of common file system management tools by semantically enhanced browsers and editors that allow the user conveniently to define and query his/her personal and shared data repositories.

Organisation	University of Vienna - Faculty of Computer Science Department of Distributed and Multimedia Systems
Core competencies	Semantic Desktop, Semantic Wikis, Metadata Visualization, Collaborative Classification, Ontology Evolution
Address	Liebiggasse 4/3-4, Vienna
Contact	+43-(0)1-4277-39621 ǀ bernhard.schandl@univie.ac.at www.cs.univie.ac.at/dms

SEMANTIC TECHNOLOGIES AS MELTING POT FOR KNOWLEDGE

Michael Granitzer

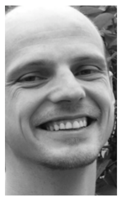
The noosphere – introduced by the philosopher and theologian Theilhard de Chardin's – indicates that phase in the biogenesis of mankind in which single minds of mankind coalesce to one. This points at a possible ultimate aim in human history.

But - how does the noosphere relate to semantic technologies?
Since the invention of speech and the written word, man has the chance to transfer knowledge and information, also beyond generations. While then the information amount was comparatively well manageable, with the development of digital information carriers and their worldwide aggregation we are provided with a practically inexhaustible information pool. A lot of knowledge – even though within reach – partly lies idle. The knowledge is hardly manageable by a single person and bringing it to the surface is hard work. Efforts for easing that situation focus on searching for relevant information. While current search technologies can handle large amounts of information, users are most often confronted with overwhelming amount of search results returned and, again, lost in information instead of deducing and creating new knowledge.

At this point semantic technologies are called. Ontological descriptions and the standardised processing of information allow for combining knowledge elements meaningfully and for deducing new knowledge. Moreover, they facilitate the non ambiguous knowledge transfer by considering the context of a user and - for the first time in human history - transfer of knowledge goes beyond interpersonal transfer by occurring also between man and machine. Intelligent agents – developed on top of ontologies and by using techniques from machine learning, information retrieval, and artificial intelligence - will help us in the near future to handle the huge and steady growing amount of information in a purposeful way. While today we are mainly engaged in information search and management, in the future our personal digital agent will overtake these and more tasks. Users no longer will search for documents or websites, but will have the possibility to ask questions to the systems about the interrelation of knowledge in a given context. The system will answer in natural language. So, unlike to the past, when information was searched and organised, future systems will be able to produce conclusions and deduce new knowledge – helping humans to cope with the steadily growing amount of knowledge in the world.

That said, it is clear that semantic technologies initiate the next big development step of society: today's information society will be transformed into a real knowledge society. In this society, semantic technologies constitute the melting pot for knowledge similar to the melting pots for iron in the industry age. Even though the noosphere indicates the melting of individual spirits to a single one, the fusion of human knowledge by semantic technologies maybe the first approximation to this.

Organisation	Know-Center Graz
Core competencies	Text Mining, Information Retrieval, Machine Learning
Address	Inffeldgasse 21a, Graz
Contact	+43-(0)316-873-9263 I mgrani@know-center.at I www.know-center.at



SEMANTICS AS THE HOPE IN PANDORA'S IT-BOX

Martin Povazay

Like the achievement of fire which was brought by Prometheus, the Internet and the omnipresent information technologies are a mercy for our society. We have experienced an evolutionary boost, which has been manifested in the expansion of the collective knowledge and the subjective possibilities. But as it was also told from Hesiod, every enlightenment has its darker side. By establishing the classical Information Technology and opening the modern-day "Pandora's IT-Box" many new plagues have afflicted our Information Society: information overload, unclear structures, incompatible systems, and data for which it requires too much time and effort to be placed in a meaningful context.

We believe that the introduction of semantics to the new generation of information technologies will provide a cure for this digital plague. Nowadays we are accumulating more and more data, without acquiring knowledge about its context. Only subjective expertise, which is expensively acquired, is delivering us the information we need for our business. In the old-fashioned IT-paradigm, permanent variation of data pools and structures results in growing operating expenses but, like Sisyphus, never seems to lead anywhere. The only alternative is to know more about the semantics: by connecting data with a subjective ontology we are able to obtain qualified information, which can provide each user a basis for valuable decisions in the twinkling of an eye. Indeed informatics alone cannot make the grade, but classical professions like cultural science must be involved in creating systems which meet human demands. This approach is particularly interesting because these divergent professions are brought together to make something new.

Many sectors of the economy are demanding improved innovative systems that, in addition to classical data accumulation, also enable the qualified preparation of information for specific domains and individual views. For example, introducing semantics to CRM systems will allow us to substantially combine data, making its context explicit, and is finally leading to profitable permutation of opportunities to profit. At the same time, semantic systems are making the realisation of user friendly interfaces possible. In addition to new interface approaches, Semantic Desktops facilitate individual user views: who would not wish to have documents and eMails processed and stored according to his own arrangement, which makes finding the necessary information easy?

"Computer, search all relevant databases. I'm looking for information about….".
In order to be able to use information systems this way, a great number of semantic projects must first reach the market, developments that will iteratively improve our life.

Starting with methods suggested by the web, more business and mobile systems will be enriched with semantic functions. In further succession, semantic services will change our working methods and society like never before.

Organisation	P.Solutions
Core competencies	SBS-Analyst, Semantic Systems Design, RDF-Engineering, Protocol Design (SemDAV)
Address	Daumegasse 5, Vienna
Contact	+43-(0)1-699 24 6621 | martin.povazay@psolutions.at www.psolutions.at

SELF-INNOVATING-SYSTEMS BY MEANS OF SEMANTIC TECHNOLOGIES

Reinhard Willfort

For years the technological process forces new business innovations. From user perspective, the availability and the employment of semantic technologies enables a new quality concerning innovations. Today a constantly grown information and communication infrastructure can be found in almost any enterprise, which often can only communicate in a difficult way with the other ones. The customer database has not any connection with project documents; each user has its own address database and collects his emails decentralized on his PC; the sharing of documents is based on a file server whose structure grows quickly and the stored data isn't up to date; processes and process description become more inconsistent every day; and especially in SMEs there is less time for innovation. The list of potential improvements could be continued endlessly.

With the availability of semantic technologies for the first time a new dimension takes place in business technology landscape: Now there is the possibility to link all resources including all data on a technology meta-level. If it succeeds that these linkages become part of a learning system that recognizes relations automatically, an optimal interaction between humans, processes and data can be ensured at any time. The result: current data at any time, arranged and understandable structures and expirations, which are actually lived from humans in that way they are documented and stored.

In the near future an Internet of all participants and resources within business processes can be developed and the topic "innovation" becomes a new meaning: In such a network innovation happens permanently enabled by semantic technologies. Innovation is facilitated by the purposeful connection between artefacts in the context of customers, suppliers, researchers and further important participants. In such a business-network, semantic technologies can favour the development of self-innovating-systems, which are held in a stable equilibrium for the benefit of all participants. Innovation becomes thereby a permanent institution in knowledge-based enterprise networks.

Organisation	ISN - Innovation Service Network GmbH
Core competencies	Innovation- and Knowledge Management Services
Address	Hugo-Wolf-Gasse 6a, Graz
Contact	+43-(0)316-919229-0 ǀ reinhard.willfort@innovation.at www.innovation.at

THE SEMANTIC WEB

Günther Fliedl, Christian Winkler

Semantic technologies, such as the languages of the Semantic Web (Resource Description Framework RDF, Web Ontology Language OWL) have been among our main interests for several years. They offer us the possibilities for filling up and complementiszing websites and existing texts on the web and thus facilitate targeted queries. Before that, however, words and phrases need to be annotated for their identification. With the help of the languages mentioned before, implicit information contained in not formally structured texts can be made explicit in structured form, which is also machine readable.

Having its roots in Artificial Intelligence Research of the Sixties and Seventies (an approach for helping people find decisions), Semantic Web today is one of the most challenging enterprises aimed at improving the interaction not only between machines but also between human beings and machines. Semantic Web is the logical development of today's Web in the direction of an intelligent and adaptable system for information and communication, turning existing data bases enriched with machine accessible metadata into future intelligence.

Nowadays, many domains such as telecommunication, public administration, education, banking, and insurance, represent the most ambitious users of semantic technologies.

In a Semantic Web, any piece of information should be represented not just as a readable document for human beings but also by means of machine readable code, thus enabling computer programs to operate on this code. As a result, queries could be processed with respect to their semantics instead of considering their textual representation.

Making information accessible to machines without doubt facilitates not only a targeted lookup but also the reusability of such information. Moreover, filtering texts without weighing their inherent real life information is much easier and thus less time-consuming. That way, at least to some extent human world knowledge can be linked with texts available on the Web.

It is one of the remaining problems up to now that schemata have to be filled manually with semantic markers. Therefore, it is our ambitious and visionary goal to handle the extraction of semantic features automatically with the help of computational linguistics tools (tagging, shallow parsing, co-occurrence analysis of words) or at least semi-automatically (interrupted by manually correcting data).

In the years to come, Semantic Web will enable us to formulate complex questions in natural language as the input for a special search engine, returning us exhaustive and meaningful answers in natural language. These answers will be the result of all efforts aimed at enriching knowledge as guided by semantic technologies.

Organisation	University of Klagenfurt - Department of Business Informatics and Application Systems (IWAS)
Core competencies	Computational Linguistics, Natural Language Processing
Address	Universitätsstraße 65-67, Klagenfurt
Contact	+43-(0)463-2700-3733 ǀ gfliedl@ifit.uni-klu.ac.at christian.winkler@uni-klu.ac.at ǀ nlp.ifit.uni-klu.ac.at/NLP

HOW THE EFFECTS OF THE CREATION OF THE TOWER OF BABEL CAN BE MANAGED TODAY

Elisabeth Harzhauser

Many years ago a united humanity was striving to make a name for themselves and therefore decided to build a tower higher than any other existing up to now. It should have been high enough that its top would reach heaven. But God was not very pleased about this provocation and open defiance. And that was the reason why he determined to come down to earth and confuse their languages and to scatter them to different parts of Earth. Since then we have all been suffering from the fact that communicating which each other is difficult or even impossible when we are not speaking the same language.

Since languages got bewildered many years have passed but the challenge is still the same. However, the problem of not understanding each other even occurs within one language. Let's imagine the following situation: You are looking for a nice birthday present for your friend who is a great fan of running. So you decide on buying some nice running shoes for him. To find what you want as quickly as possible you visit an appropriate Internet shop and use the search function of the site where you enter "running shoes" into the search field. Unfortunately, you get no results back. You leave the site frustrated and determine to buy something else instead. But it could have been so easy – the only problem was that the kind of sport shoes you were searching for is only known by the term "jogging shoes" in the web shop. Yet, a semantic web in the background of the search function would have helped you to find "jogging shoes" although you wrote "running shoes". This is made possible by the fact that in a semantic web terms are connected to each other through different kinds of relations. As a result, synonyms, broader and narrower terms; related terms and even the translation of the term into another language can be identified and searched for automatically. The great advantage of this new technology is obvious – searching for content becomes a lot easier and faster for the end-user.

Although the design of semantic web applications is still in its infancy, the future of the technology is certain. At present, it is researched how the construction of the terminology base which is needed for the design of semantic webs can be assisted by special software-tools. Of course, some terminology-tools have already been developed but for sure, this is only the first step into the right direction.

Organisation	ProCom-Strasser	
Core competencies	Design, Update and Usage of Semantic Webs in Professional Search Engines	
Address	Hauptstraße 157, Strasshof bei Wien	
Contact	+43-(0)2287-20068 | elisabeth.harzhauser@procom-strasser.com www.procom-strasser.com	

NO SEMANTIC WEB IN 2006

Kerstin Zimmermann

Challenges
From an information management perspective the Semantic Web would be great. All documents can be cross linked and referenced with meanings. When you search for a keyword or person you will get all relevant details in the right context and in the best case all online available. But the status quo is different.

First mentioned in 1989 by Berners Lee the Semantic Web became a hype around 2000 within the new economy. EU fundings and US activities spend mill of € and $ in international projects. Data with semantics should be generated automatically in order to create knowledge. The same approach has been done in artificial intelligence before. But then the dot bubble burst without much content. So the fundamental research went back to logics, describing things in a very formal way. New languages were developed and web services propagated. Companies did not see the market in this application because their businesses were not really integrated.

What are the problems regarding the full potential of online information? The add-on value of a worldwide semantic network lacks of ontologies. Everybody has another conceptual model in mind: in computer science concepts and relations are used but classes and attributes are given for different domains (e.g. library science). The history of ontology in philosophy deals with its fundamental categories. Semantic annotation of named entities shall give additional information on ‘plain’ web sites but the scheme behind is not public. In interdisciplinary research you can estimate the confusion.

Parallels
The activities to make the Semantic Web real remind very much on the introduction of plain Dublin Core (DC) meta data in the mid 90ies. The format and the technical procedure were quite clear but not spread out widely. The first interdisciplinary project defines a real case workflow of documents for the user as author and the information seeking person. This inclusion and the distributed task force gained international success.

Realizing the Semantic Web can not neglect the human factor and has to take the community aspect into account. Innovation only based on technology is not enough. User and expert have to meet and exchange their knowledge. Developers also have to reflect their own background, because each thesaurus includes implicit rankings. Ontologies with instances form a knowledge base. As multi-lingual thesauri are being developed e.g. they should also be coded as ontology and mapped to other ones and a meta one for the global context. This knowledge has to be made explicit because it contains also a paradigm shift in transferring methodology to other domains.

Organisation	Federal Ministry for Transport, Innovation and Technology (bmvit) Section III Innovation i5 Information- and industrial Technologies
Core competencies	Information Management, Portals, Classification Schemes, Ontologies, e-Learning and Gender Aspects
Address	Renngasse 5, Vienna
Contact	+43-(0)1-53464-3503 I Kerstin.zimmermann@bmvit.gv.at www.bmvit.gv.at

SEMANTIC TECHNOLOGIES GO FAR BEYOND ITS ENGINEERING SKILLS

Tassilo Pellegrini

What is information? What is knowledge? There are a vast amount of definitions of these terms. But how do they apply to everyday life? Generally speaking there are at least three types of knowledge: 1) the known known, 2) the known unknown and 3) the unknown unknown. Current IT is well suited to tackle the first type of knowledge. Semantic technologies will unfold their potential by improving the first and tackling the second. And it is relatively obvious that there will never be any technological means that might help us with the third.

The Semantic Web School offers applied knowledge on the use, benefits and risks of semantic technologies, investigating in the scale and scope of their application in everyday life from an individual, organisational and infrastructural point of view. Semantic technologies go far beyond engineering skills. Their roll out is accompanied by social, cultural, political and economic questions about the use and applicability of semantic technologies, especially as we are witnessing a rapid uptake in domains like education, health, life sciences, telecommunications, media, automotive and many more. This is a critical development as computational semantics is a prerequisite in the incremental development of future technologies - especially in the nano domain – providing mankind with new opportunities but also challenges. Hence semantic technologies must be seen as a fundamental development in the socio-technological transformation of advanced industrial societies in which knowledge and information are a crucial factor in international and local competitiveness.

So where are we heading?
At the current roll out stage semantic technologies are already creating an impact at the organisational level. Their application within existing IT infrastructures for purposes like searching, navigating and information integration creates benefits by providing single points of access to dispersed information sources, the qualitative improvement of search results, the efficient guidance through work flows and the leveraging of collaborative content creation. Especially the latter aspect should be watched carefully. The convergence of semantic technologies with social software (i.e. semantic wikis, structured blogging etc.) will have a strong impact on future knowledge work, laying the technological basis for a highly skilled information work force embedded within a knowledge society.
Within the next two to five years the broad diffusion of semantic technologies will have taken place on the organisational level. Semantic technologies and applications will have found their way into custom software products, leveraging the personalization of IT and bringing about new business models and working schemas. The integration of computational semantics into wireless communication tools (mobile phones, PDAs, RFID) and household appliances will increase the smartness of our technological environment, intensifying the spread of the ubiquitous computing paradigm. Technological questions about adequate trust and proof mechanisms will have been solved bringing Tim Berners Lee’s vision of a semantic world wide web just at our reach.
And in about 15 years nobody will talk about semantic technologies anymore. They will be an integral part of our everyday life and people will be pondering, how it ever was possible to live without them.

Organisation	Semantic Web School Center for Knowledge Transfer
Core competencies	Development of Semantic Web-technologies, Match-making Industry-Academia, Knowledge Transfer
Address	Lerchenfeldergürtel 43, Vienna
Contact	+43-(0)1-4021235-28 I t.pellegrini@semantic-web.at www.semantic-web.at

REINVENTING INFORMATION

Philipp Haindl, Olivia Dietzel, Doris Reisinger

Among the various efforts of the last years' technology attainments, semantic technologies are the recent cornerstone in the frame of information exploitation and discovery of associated knowledge. While in the past, information assets have been considered as atomic containers for information suspected to be sufficiently self-explaining, semantic technologies can be considered as a paradigm shift now regarding the whole life cycle of human comprehensible information. As such, information and even more its implications and further deductions are now perceived in conjunction with the knowledge and the context which they originated from. Semantic technologies afford to bridge the information gap between human comprehensible information and machine information systems and bring new means of unveiling the hidden but still existing knowledge between the lines.

Unveiling new knowledge

The current wealth and availability of information for everybody is accompanied by a set of controversial phenomenons for the individual. Finding the right information required to perform a specific task takes more time the more information is available, while concurrently and even worse the individual needs more information to successfully perform this task. Unfortunately also the continuously increasing possibilities to store knowledge have not yet provided the user with effective tools to exploit existing and discover new knowledge. Allowing the individual to make statements similar to natural language, semantic technologies provide innovative approaches to store and explore human knowledge as well as to reduce the information loss between human and machine. Developing information systems at the semantic level of the processed information allows leaving the constraints of traditional object-oriented programming languages and facilitating exploitation also of the ontology of the processed information, which allows development of sophisticated decision support systems and just-in-time modeling of software intelligence.

Why information can be interesting again

Until recently, information retrieval has been intimately connected with searching for text phrases either not or loosely coupled with additional semantic information. Thus, these queries not only yielded results originating from diverse and hence wrong ontologies, but also results not even related to the search topic. Querying any of the current state-of-the-art Internet search engines for a single search term in most cases yields a respectable amount of results regarded useless by the user. In this note, we all are invited to rethink Shannon's theory of communication demanding that information must be something new for the recipient. It should be the ambition of every professional computer scientist to provide supportive and comprehensive means for interacting with and retrieving information and we all are demanded to strengthen our efforts to pave the way for semantic technologies to arise to such a means.

Organisation	m2n - consulting and development gmbh	m2n Intelligence Management
Core competencies	Rule Based Engineering, Ontologies, Information Retrieval, Knowledge Discovery	
Address	Marienstraße 10, Linz	
Contact	+43-(0)732-711 987-0 | haindl@m2n.at www.m2n.at	

Part II – Chapter C

Systems, Models & Methods

SPATIO-SEMANTIC SYSTEMS FOR UBIQUITOUS ACCESS TO SPATIALLY STRUCTURED INFORMATION

Florian Ledermann

Our world is full of spatially structured information – road signs, signposts, warnings, but also some architectural features, paths and graffitis would be meaningless without the specific spatial context they are situated in. Especially in urban areas, real space is already overloaded with such chunks of information, while the complexity of information that can be expressed through conventional, physical signs remains somewhat limited. In an envisioned digitally enhanced space, such information could be adapted to user requirements and be shown only to those users who need it, while at the same time much more complex and dynamic information could be taken into account. In our research we attempt to create systems that are capable of displaying such spatially structured information to the user, adapting to her location, the state of her environment and her tasks and intentions. The utilized display techniques range from three-dimensional computer graphics, accurately registered to match the view of the user to provide an "augmented reality", to text-based output providing much coarser information.

To support this continuum of display techniques and provide a powerful mechanism to select and infer relevant information at runtime, we are exploring the capabilities of semantic systems for modeling the information space browsed by the user. One of the fascinating research problems in our field is the question of how to link real space – described by geometry – with such an abstract information space, possibly described in databases or semantic networks. Computer scientists and geographers have collected a lot of experience in the area of geometric description of space, and within the geography and GIS communities there is a huge body of work about spatial modeling. However, the border between geometry/geography on the one side and semantics/formal logic on the other has so far only rarely been crossed.

In a ubiquitous information space, we have to relate the world as observed at run-time – represented by measurements and geometrical models – with our abstract model of the world – represented by a semantic network. One of the problems here is that the entities in the semantic network are discrete and sharp, while measurements are often provided on continuous scales and with various sources of error and uncertainty. Any framework that attempts to link real-time spatial information with semantic information must therefore provide means to support continuous scales and uncertainty – topics that are at the forefront of research in formal logics and semantic modeling.

The fusion of spatial and semantic information remains an exciting research topic that will open new possibilities for a wide range of applications in areas such as tourism, public service and personal information management – and especially the intersection of those three.

Organisation	Vienna University of Technology - Interactive Media Systems Group
Core competencies	Spatial Modeling, Spatial Information Visualization, Ubiquitous Computing
Address	Favoritenstraße 9-11/188/2, Vienna
Contact	+43-(0)1-58801-18855 I ledermann@ims.tuwien.ac.at www.ims.tuwien.ac.at

AN ONTOLOGY-BASED UNIFIED COMMUNICATION PLATFORM

Jürgen Falb, Hermann Kaindl, Roman Popp, Alexander Szep, Edin Arnautovic

Communication between humans happens on a very high semantic level, as compared with human-machine and especially machine-machine communication. The meaning of uttered information is heavily influenced by additional information like, e.g., intonation, context, and intention. Our basic assumption is that human-machine interaction and machine-machine communication should be also on a higher level and more similar to communication between humans.

We studied several theories of human communication from various fields. By integrating results from *Speech Act Theory, Rhetorical Structure Theory and Conversation Analysis*, we developed a new discourse model for describing dialogues. Starting from speech acts, abstracting from speech and natural language leads to *communicative acts*. The key idea is that speaking or, more generally, communicating is not only used to describe something or to give some statement but also to *do* something with intention — to act. Typical communicative acts are for example Request, Question, Informing and Offer. The same proposition uttered with different intention could be for example:

Sam smokes habitually. (Informing)
Does Sam smoke habitually? (Question)
Sam, smoke habitually! (Request)

Communicative acts can be seen as the atoms of our discourse model and link the content with an intention. Considering communicative acts of one communication partner, the atoms can be linked via *rhetorical relations* to build up *molecules*. Thus, rhetorical relations (e.g., Motivation or Elaboration) state how one utterance of a party links to another. For instance, an Informing can *motivate* a particular Request. In the dialogue dimension we use adjacency pairs to link the molecules of each party via their atoms. An adjacency pair states, for example, that a question has to be followed by an answer of the other communication partner.

We claim that such a discourse model contains enough semantic information to fully automatically generate useful and usable user interfaces (e.g., graphical or voice interfaces) for human-machine communication. In addition, such a discourse model enables semantically enriched machine-machine communication. For this purpose, we use ontologies for representing the discourse models and referencing domain knowledge.

A common discourse model for both types of communication allows us to unify them, eases its modeling, and makes communication independent of the type of communication partner (human or machine). Thus, the same dialogue can be carried out with either two machines or one machine and one human without requiring different discourse models.

Organisation	Vienna University of Technology Institute of Computer Technology
Core competencies	Modeling, Ontologies, Semantics in Communication
Address	Gußhausstraße 27-29, Vienna
Contact	+43-(0)1-58801-38417 | falb@ict.tuwien.ac.at www.ict.tuwien.ac.at

Part II – Chapter D

Projects & Institutions

SEMANTIC TECHNOLOGIES FOR SERVICES

Schahram Dustdar

Introduction

Worldwide, there is a growing recognition that Internet Systems enable new forms of human collaboration – for groups of private individuals, for companies, and for the public sector – and also for emerging Web services, which can be discovered, deployed, and enacted as workflows on the Internet. Current effort in research and industry is, therefore, aimed at transforming the World Wide Web from a network that makes information available and that facilitates transactions, towards an environment that provides software services and resources to facilitate the emergence of dynamic virtual organisations and collaborative working environments utilizing Service-oriented Information Systems on the Internet.

Problems with Semantic technologies

Semantic technologies claim to solve many of those problems, which were briefly outlined above. Do they deliver what they promise? Certainly not to the degree we all hope for. The reason is that attaching semantics to information is a hard problem. Sure, we have some initial standards for describing semantics; however, the process of having semantics attached to information (e.g., Web pages, Web services, and business processes) remains difficult and time consuming. In our research projects in this domain, we mainly focus on Service-oriented Computing and the related problems.

Some research projects

In the area of semantic technologies, we currently have three research projects at the Vita Lab which are aiming at contributing to the solution of some of the problems outlined above. In the EU Framework 6 project called inContext (www.in-context.eu) we aim at composing those Web services for teams and their members which are relevant to the work activities of the team members. Relevance and context information are certainly important ingredients for achieving higher levels of semantic importance. In our FIT-IT project Semantic Culture Guide (SCG) wiki.sembase.at/index.php/SCG we provide interaction pattern modeling techniques and tools for allowing construction of a guide to Austrian cultural events. In SCG a platform providing the dynamic relationships between providers and consumers of cultural information is the overall goal. In our FIT-IT project SemBiz (www.sembiz.org) we develop techniques and tools for semantic querying of business processes, thereby closing the gap between technical models of processes (workflows) and business views. More detailed information about these projects as well as our other projects and publications can be found at: www.VitaLab.tuwien.ac.at.

Organisation	Vienna University of Technology - Distributed Systems Group (DSG), Information Systems Institute	
Core competencies	Service-oriented Computing, Context-aware Computing, Process Technologies	
Address	Argentinierstraße 8/184-1, Vienna	
Contact	+43-(0)1-58801-18414 | dustdar@infosys.tuwien.ac.at www.VitaLab.tuwien.ac.at	

SEMANTICALLY ENABLED SERVICE-ORIENTED ARCHITECTURES (SESA)

Holger Lausen

“Semantics have the potential to radically alter the way we use computer and Internet technologies. It aims to enable computers to better understand information, and to cooperate “more intelligently” with each other. It is a very exciting time to be a researcher in this field” Univ.-Prof. Dr. Dieter Fensel, Director of DERI Innsbruck.

DERI Innsbruck was founded in 2002 by Univ.-Prof. Dr. Dieter Fensel. Current research involves an international network of other DERI institutes in Asia, Europe and the USA as well as several ties to global industrial partners. DERI Innsbruck research activities will establish semantics as a core pillar of modern computer engineering. Achieving semantic description of computing is a stepwise evolution, which started with the Semantic Web, moved on with Semantic Web Services, and it is currently in a transition period towards Semantically Enabled Service Oriented Architectures.

Practically every area based on IT, e.g. e-Business, e-Health, e-Government, eTourism, will benefit from this new technology. End-users will be able to search the Web much more efficient and successful for suitable offerings. For example, one will be able to search for “hotels in the north of London with French cuisine” and find also such offers that contain this fact implicit.

Semantically Enabled Services Oriented Architectures will support us in nearly all aspects of our daily life - making access to information as pervasive as access to electricity is today. They will be the next paradigm shift, which will lead us in a world of efficiency where, independent of different languages, human beings and computers can communicate smoothly and understand each other. Still, more work needs to be done before the Web service infrastructure can make this vision come true. Current web service technology provides limited support in mechanizing service recognition, service configuration and combination (i.e., realizing complex workflows and business logics with Web services), service comparison and automated negotiation. In a business environment, the vision of flexible and autonomous Web service translates into automatic cooperation between enterprise services. Any enterprise requiring a business interaction with another enterprise can automatically discover and select the appropriate optimal Web services relying on selection policies. Services can be invoked automatically and payment processes can be initiated. Any necessary mediation would be applied based on data and process ontologies and the automatic translation and semantic interoperation.

Organisation	Digital Enterprise Research Institute (DERI)
Core competencies	Semantically Enabled Service-oriented Architectures (SESA), Semantic Web Services
Address	ICT - Technologie Park, Technikerstraße 21a, Innsbruck
Contact	+43-(0)512-5076464 | holger.lausen@deri.org www.deri.org

HOW DO YOU MEAN? SEMANTIC SYSTEMS AND NATURAL LANGUAGE TECHNOLOGY

Alexandra Klein, Brigitte Krenn, Harald Trost

The ability to make sense of encountered data is the central cognitive skill. Therefore, several different disciplines, such as psychology, computer science, philosophy, and linguistics have investigated from their respective viewpoints how meaning is systematically assigned to what would otherwise be random pieces of data. During the last decades, Artificial Intelligence has contributed a more integrative perspective on how intelligent behavior, learning and adaptation is influenced by available information. In the 1980s, a multidisciplinary group of researchers, who were particularly interested in the distinctive human ability to process language data, was established at the Austrian Research Institute for Artificial Intelligence (OFAI). Since its formation, this Language Technology group has conducted and participated in numerous national and international research projects, approaching the relationship between spoken or written language and meaning from different angles, e.g.:

- VIECTOS: producing spoken utterances from symbolic representations
- MultiMod-WebAccess: accessing web pages by written and spoken utterances
- FASTY: assisting motor, speech, learning and language impaired persons to produce texts
- RASCALLI: developing cognitive agents that are adaptive and cooperative and support the user in finding information
- SPARC: integrating semantic knowledge bases into automatic speech recognition systems for dictation applications, funded by the FIT-IT "Semantic Systems" program
- SEMPRE: enhancing recommender systems with knowledge derived from semantic processing of web information, funded by the FIT-IT "Semantic Systems" program

Over the years, research has evolved from simple pattern matching techniques toward much more complex approaches. Multidisciplinary experience, better technology and the availability of vast amounts of textual data will let researchers gradually come closer to touching the core of what constitutes human cognition, thus making this one of the most exciting research fields we can think of.

Today, research has reached a point where its results are used in different applications. Consequently, applications dictate research angles. Users already benefit from Language Technology enhanced with semantic information in many areas of life, whether they realize it or not: e.g., search engines use natural language technology for a better precision and recall of the results, structured knowledge sources on the web such as Wikipedia rely on concept hierarchies for disambiguation of word senses, interfaces for spoken or written commands facilitate natural human-computer interaction. Such diverse tools as mechanisms for spelling correction, machine translation and spam filtering support users by employing semantic analysis. A lot of research will still be needed until semantic technology will enable machines to mimic the human language competence in all its complexity. Current systems focus on the extraction of factual information, future approaches will accommodate the many uses of language by humans and expand meaning analysis to the interpretation of opinions, intentions and emotions.

Organisation	Language Technology - Austrian Research Institute for Artificial Intelligence (OFAI)
Core competencies	Language Technology
Address	Freyung 6/6, Vienna
Contact	+43-(0)1-4277-63121 ǀ harald.trost@ofai.at ǀ www.ofai.at

USABILITY OF METADATA USER INTERFACES

Guido Kempter, Philipp von Hellberg

Motivation

We suppose, that manually assigning metadata to individual data input for semantic systems will be indispensable within the far future, despite all the widespread efforts in developing computerized alignment of these data inputs with predefined taxonomy (e.g. by computerized content analysis). This is of course not only a restriction in using semantic systems but for this purpose we will need, however, some user interfaces for which we have to apply the same usability requirements as for other ease-to-use human-computer interfaces. UCT Research, the User Centered Technologies Research Institute at University of Applied Sciences Vorarlberg, aims to discover the usability problems of individual and shared operation with metadata and taxonomies and works on new methods and technologies for optimising the usability of such user interfaces. We are dealing with semantic technologies in research and development, because users will disapprove semantic systems, if manually assigning metadata to individual data input is hardly to perform with the software system or even fails the user's needs. These user problems with semantic systems could derive from taxonomies which don't sufficiently reflect the mental information categories of the users or from user interfaces for manually assigning individual data input to empirically validated taxonomies which are not in line with common usability heuristics.

Realisation

UCT Research aims at developing new solutions for knowledge and information management systems giving specific support in optimising the usability of human-computer interfaces for the manual assignment of metadata to individual data input on the basis of individual work with digital information of these systems. For this purpose, we are performing off-line optimisation using common usability evaluation methods (e.g. inspection and testing) as well as on-line optimisation developing and validating intelligent and adaptive user interfaces. The latter approach makes use of alternative computer input devices for registering user behaviour (e.g. by psycho-physiological sensors), context of information usage (e.g by data logging software), and selected social interaction phenomena (e.g. by motion capturing) as well as real-time algorithms for user data analysing (e.g. genetic pattern recognition) to facilitate continuous adaption of the corresponding graphical user interface. The adaptivity follows intelligent decision rules in order to achieve the desired psychological effects within the user. This approach differs from other adaptive semantic systems as we don't modify information structures and taxonomies but are modifying specific background elements of graphical user interface for manually assigning metadata to individual data input.

Impact

As soon as adaptive human-computer interfaces of knowledge and information management systems are optimised in a way that users are regularly mapping their data inputs with the taxonomy of these systems this will have an essential impact on the daily life of users. This impact doesn't only concern the explicit individual and collective meaning of terms and ideas but also spontaneous decisions in rudimentary social situations.

Organisation	FH Vorarlberg - User Centered Technologies Research
Core competencies	Adaptive User Interfaces, Online Monitoring, Usability Optimisation, User Cognition and Emotion
Address	Hochschulstraße 1, Dornbirn
Contact	+43-(0)5572-792-7300 | guido.kempter@fhv.at | www.fhv.at

HTMLBUTLER - WRAPPER USABILITY ENHANCEMENT THROUGH SHARED ONTOLOGIES AND LARGE SCALE COOPERATION

Christian Schindler

Motivation & Benefits

The general idea behind the Semantic Web is to extend the current Web (which is basically an HTML web) with additional "meaning". The two main motivations are a.) data integration and b.) intelligent support for the end user. In more technical terms the Semantic Web extends the Web through standards, semantic mark-up languages (RDF, DAML+OIL, OWL), inference engines (F-OWL, RACER etc.) and various processing tools. The Semantic Web is tightly coupled to the term “ontology” - “An ontology is a formal, explicit specification of a shared conceptualization.” and provide a vocabulary for adding semantic information to content.

The benefit is that software can utilize these semantics to support the user in achieving his tasks through smart suggestions, heuristics and automatically completions.

The htmlButler Project

The htmlButler project aims at bridging the gap between the current web, mostly based on HTML, and the future Semantic Web which content is underpinned with ontologies and annotated. htmlButler supports the user to visually create wrappers - “specialized program routines that automatically extract data from Internet web sites and convert the information into a structured format”. For instance a user browses a certain last-minute travel portal, gets a list of different trips taking place within the next weeks, selects one row of the table and invokes the htmlButler service to inform him when certain changes happen to his selection. He can adjust thresholds to be notified when the price drops below a certain limit, the departure time, or the availability of the offer changes. The user can enter data which do not exist yet and be notified when an entry with the specified data appears. The htmlButler system understands the different cells of the row (e.g., “destination”, “price”, etc.) through the use of a basic ontology which can be enhanced by the user. Feedback is used for further suggestions - the more people contribute the more general the ontology becomes. Issues of quality can be addressed through a ranking system.

Future of the Semantic Web

No one is able to make precise predictions about the future - “You can only predict things after they have happened”. However, looking at the past evolution of the World Wide Web one can state that the development of the Semantic Web is only a question of time and that through the development of small and smart tools the idea can be made popular to the broad community of web users. The shift to the World Wide Semantic Web is not likely to happen in the next 5 years - there will be rather small different semantic webs within organizations and communities but it is and will be an interesting time.

Organisation	University of Technology Institute of Software Technology
Core competencies	Semantic enabled Webtool Development, Ontology Enhancement and Sharing
Address	Inffeldgasse 16b/II, Graz
Contact	+43-(0)316-873-5729 | cschindl@ist.tugraz.at www.ist.tugraz.at