SIGSEMIS Bulletin

Volume	1
Issue	1

April 2004
Theme: "SW Challenges for IS"

The Official Bimonthly Newsletter of AIS Special Interest Group on Semantic Web and Information Systems

Inside Bulletin

TOP NEWS

Announcing Int. Journal on Semantic Web and Information Systems
Announcing Encyclopedia of Semantic Web Research

Visit Our Portal at http://www.sigsemis.org

EDITORIAL…

The Special Interest Group on Semantic Web and Information Systems is a newly formed SIG of Association for Information Systems. In our portal site at www.sigsemis.org you can find valuable information concerning our activities and up to date material (this will be evident in the forthcoming months).

It's a great pleasure for us to welcome you to the first issue of our official bimonthly Bulletin. This AIS publication intents to become an open tribune of dialogue for important themes that cultivate the semantic web vision in IS. We plan in the first year of publication to initiate interesting discussions concerning the changes that SW poses to IS research. This inaugural issue is dedicated to the discussion of the Challenges that Semantic Web poses to Information Systems Research Community. We would like to thank all the contributors for their support. Without their help this issue would be of limited value. More over we would like to stress your interest in an excellent piece of news: Our SIG agreed with IDEA Group Publishing for the publication of a new journal entitled *International Journal on Semantic Web and International Journal.* More information on this can be found in page 5. The first issue of the journal will discuss a critical theme: *Semantic Web Challenges for Information Systems Research: Towards the Knowledge Society.* In the second issue of AIS SIGSEMIS Bulletin we will provide an open call for papers concerning our journal.

In the first issue you will find two excellent interviews: Prof. James Hendler from University of Maryland and Prof Amit Sheth from University of Georgia comment on the role of Semantic Web and provide exciting views. I invite you to read these interviews. In the forthcoming issues of Bulletin we will have a special column entitled "**Meeting of the Minds"** which will host top academics and practitioners interviews. Finally we would like to communicate one more interesting piece of news: SIGSEMIS Board members undertook the project to develop (and to publish in early 2006) the *Encyclopedia on Semantic Web Research* in IDEA group Publishing. We will provide very soon more information on this through our portal and newsletter.

We are looking forward to your active participation and collaboration to our initiative: SIG SEMIS is an open forum: We invite you to join us (http://www.aisnet.org/sigs.shtml) and to share your thoughts and perspectives. We would like to thank AIS Council for their support and their approval. We have to promise that we will do our best for the promotion of the Semantic Web Vision in IS. Join us!!!

On behalf of SIG SEMIS Board
Miltiadis D. Lytras
ELTRUN, The E-Business Center, Department of Management Science and Technology, Athens University of Economics and Business, Greece

An Interview with Professor James Hendler, Univ. of Maryland, USA

"I often use the example of a dive shop's portal as the ideal place for Semantic Web technology…"

James Hendler

Full Professor, Dept. of Computer Science, UM Institute for Advanced Computer Studies,
Institute for Systems Research, and Dept. of Electrical Engineering, University of Maryland at College Park, USA
Director, Semantic Web and Agent Maryland Information and Network Dynamics Lab – MINDSWAP
http://www.cs.umd.edu/users/hendler

Jim Hendler is a Professor at the University of Maryland and the Director of Semantic Web and Agent Technology at the Maryland Information and Network Dynamics Laboratory. He has joint appointments in the Department of Computer Science, the Institute for Advanced Computer Studies and the Institute for Systems Research, and he is also an affiliate of the Electrical Engineering Department. He has authored close to 150 technical papers in the areas of artificial intelligence, robotics, agent-based computing and high performance processing. Hendler was the recipient of a 1995 Fulbright Foundation Fellowship, is a member of the US Air Force Science Advisory Board, and is a Fellow of the American Association for Artificial Intelligence. He is also the former Chief Scientist of the Information Systems Office at the US Defense Advanced Research Projects Agency (DARPA), and is a prominent player in the World Wide Web Consortium's Semantic Web Activity.

Miltiadis: Dear Prof. Hendler, we are really honored for your kindness to provide us this interview. You are considered a leader in SW community. What is your opinion for the evolution of SW and what do you answer to all those people who claim "Ok..Semantic Web. But where is it"?

Prof Hendler: The Semantic Web is well on its way. The languages needed, RDF and OWL, are now recommendations of the World Wide Web Consortium (W3C) and many companies and universities have been investing in Semantic Web work. The investment in the technology is growing and there are a large number of users. There's a long way to go, but we're off to a great start.

> *The Semantic Web is well on its way…. There's a long way to go, but we're off to a great start.*

Miltiadis: Your article with Sir Berners Lee and Ora Lassila in Scientific American (http://www.sciam.com/issue.cfm?issueDate=May-01) is considered an important paper in this area. Its contribution to the communication of what SW means was excellent. If you write (this or next year) a follow up on this article which will be the main points?

Jim: When we wrote the article, OWL was still just an emerging vision, RDF schema was not yet a recommendation, and industry was just starting to show its first interest in much of this work. In the three years since it came out, Tim and I have written articles about the business case for the Semantic Web. I suspect now, however, if we were asked for an article, we would write about next steps -- how to approach the higher levels of the Semantic Web and what some of the exciting possibilities are for this technology.

Miltiadis: I have seen your excellent work in MindSwap Lab. What do we have to expect from your research in the near **future?**

Jim: I have a great group of students and programmers that are doing really great work in the Semantic Web area. We are exploring a number of different things: how to integrate semantics with Web services, how to use the Semantic Web to support researchers doing work in interdisciplinary scientific areas, and how to use Semantic Web technology to expand

the capabilities of current Web applications. To see what we're working on, come see http://www.mindswap.org -- which is actually a web site powered by RDF and OWL.

Miltiadis: Several times recently we hear more and more about next generation web research. I would like you to share with us your personal view on this next generation web research. What is it all about?

Many people seem to think the only role of the Semantics is to improve search, but that doesn't make sense to me -- there's way more power in bringing new sorts of resources to the web, linking them together with current resources, and in exploring how the use of semantics will create new applications on the Web.

Jim: In my opinion, the Semantic Web provides an infrastructure on which many exciting research problems can be explored. Many people seem to think the only role of the Semantics is to improve search, but that doesn't make sense to me -- there's way more power in bringing new sorts of resources to the web, linking them together with current resources, and in exploring how the use of semantics will create new applications on the Web. Consider the many new applications that were made possible by Web technology, and imagine how many new and different ones we can create if the Semantic Web is even half as successful.

Miltiadis: A few days ago I had a conversation with a colleague in the university. He wanted to learn more concerning Semantic Web and Next Generation Web Research. And basically he put me a dilemma: "Technologies or Theories?". What is your advice to a newcomer in the field? Many student members of our SIG would be interested on it.

I think that isn't an "exclusive-or" -- that is, we need to explore both of these directions. There is a tremendous amount of work to be done in figuring out how current technologies (such as ontology tools) are changed by the differences between the Semantic Web and traditional AI approaches (for example, how does the fact that ontologies can be linked together and refer to terms in other ones change our reasoners?). There's also work to be done figuring out new ways businesses can use these new technologies and what sort of applications are enabled. What new business models can emerge when we explore, for example, Web-based application integration or simple to run Web portals? New information sources coming to the Web, new models of use of these sources, and new technologies taking advantage of these models are all exciting new things to explore on the Semantic Web.

New information sources coming to the Web, new models of use of these sources, and new technologies taking advantage of these models are all exciting new things to explore on the Semantic Web.

Miltiadis: Prof Hendler would it be possible to outline a few (lets say 3-5) hot research areas in the field of Next Generation Web Research? In other words assume that a PhD candidate ask you to indicate to him a few possible titles for interesting research. What would you recommend to him?

Jim: I think there are many exciting possibilities. Here's a few.

1. *On the Semantic Web the ontologies are linked together and can use terms from other ontologies and change them. The system is open and distributed and there is no way to guarantee consistency. How do we do reasoning in this kind of distributed and inconsistent system?*
2. *Social networks are becoming very popular on the Web, and it is clear that Semantic Web technologies help support large, distributed networks of people who know other people (like the Friend of a Friend, FOAF, work). What new and exciting things can explicit semantics add to these?*
3. *How are traditional technologies (information retrieval, artificial intelligence, etc.) changed by Web semantics and Semantic Web languages?*
4. *One of the promises of the Semantic Web is that it will let us bring databases and structured information sources (like spreadsheets) to the Web. How will query and search engines for this kind of information work?*
5. *How will semantics function in the emerging world of mobile and ubiquitous computing and other emerging IT trends?*

Ok, that's five, there are more, but you have to let me save a few for MY graduate students.

Miltiadis: Lets say that you share your thoughts with a student. How do you describe him/her the next generation IT world? A world of less creativity, less imagination and more IT control?

Jim: No, No, No! Exactly the opposite - a world of more creativity, more imagination and more human control over our IT. The systems are here to work for us, not the other way around!

Miltiadis: Would you like to share your vision for semantic web for the next decade. Would it be possible to share with as a few examples of real world SW enabled tools and applications for citizens?

Jim: I think we are going to see continually evolving capabilities based on the Semantic Web infrastructure. I think the first application area where we see it deployed is in enterprise application. That will let us see the creation of "islands" of semantic web functionality. We will also see the Semantic Web allowing the creation of easier to build and run web portals. These will also give us areas of content to link together, and as all these things do link together, the web of metadata will grow, and we will see the Semantic Web really emerge.

I think the first application area where we see it deployed is in enterprise application. That will let us see the creation of "islands" of semantic web functionality. We will also see the Semantic Web allowing the creation of easier to build and run web portals.

Miltiadis: The student I mentioned you earlier concluded his exercise for your work by stating: If I could ask Prof Hendler something I wouldn't concentrate on SW. I would rather have asked him to share with me a Scuba Diving experience.

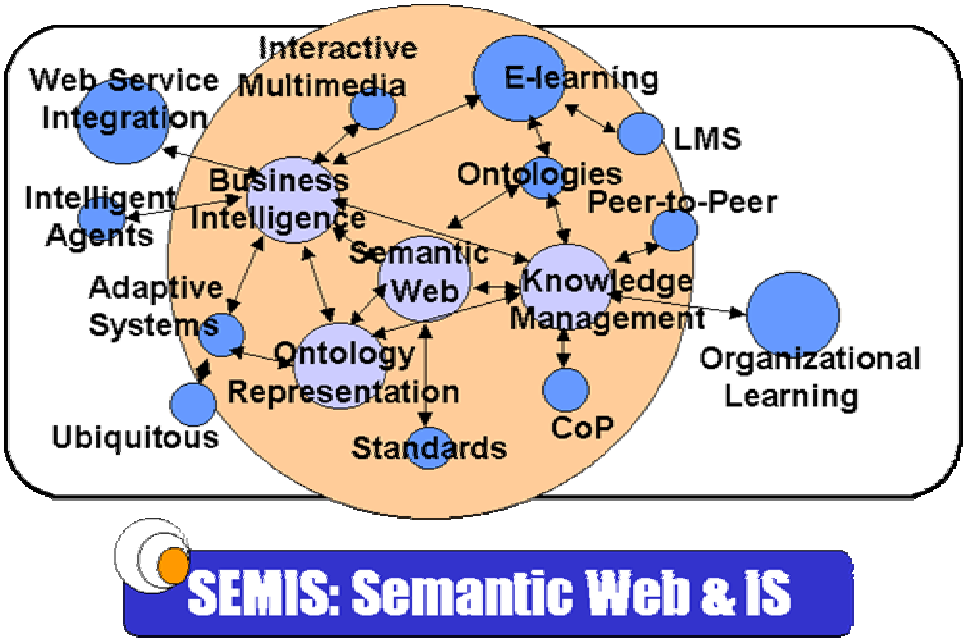
Jim: Well, I guess once again I would just wonder if this is an "exclusive-or" -- I don't dive as much as I used to, but I do it whenever I get a chance. And, I often use the example of a dive shop's portal as the ideal place for Semantic Web technology, so maybe someday we can do both at the same time.

Miltiadis: Dear Jim thank you for your time. It was an excellent talk.

SIGSEMIS: A new AIS Special Interest Group on Semantic Web and Information Systems

By SIG Board

Figure 1. SIGSEMIS Positioning



Introduction

SIGSEMIS is a new AIS Special Interest Group approved in the December 2003 meeting of AIS Council in USA. In figure 1, a graph summarizes the main areas of emphasis in our SIG: Seven critical pillars provide the basic backbone for the main contributions. Namely: Semantic Web, Knowledge Management, Information Systems, E-Learning, Business Intelligence, Organizational Learning and Emerging Technologies. This basic backbone and the relevant branches of the graph also provide the ultimate objective of the SIG. To create knowledge capable of supporting high quality knowledge and learning experience concerning the integration of the 7 depicted pillars. We are deeply ensured that this integration provides interesting and motivating understanding of the relevant issues and delivers critical value to the participants of the SIG. We would like to state some ideas before explaining our strategic view of the SIGSEMIS.

Semantic Web poses new challenges to Information Systems. A first comment concerning the current situation is that the field of SW is dominated by rather technical approaches where there is a lack of multidisciplinary contributions. From this perspective the SIGSEMIS is attempting to fill this gap, with a special emphasis to be paid on Information systems. SIG SEMIS aims to develop an extensive web of relations with several critical stakeholders. We distinguish six specific stakeholders that can gain critical value from the operation of SIG and its services: Academia, IS students, IT Industry, Public Bodies, Consultants, Standards Communities and several others. The central positioning of SEMIS in this web implies its critical role for the knowledge transfer. In other words all the services provided by SEMIS will have a clear orientation in knowledge creation and diffusion concerning the issues of Semantic Web in the IS research Community. Our main distinction is that semantic web is not treated as a technological evolution but we concentrate on the translation of SW within the IS research community.

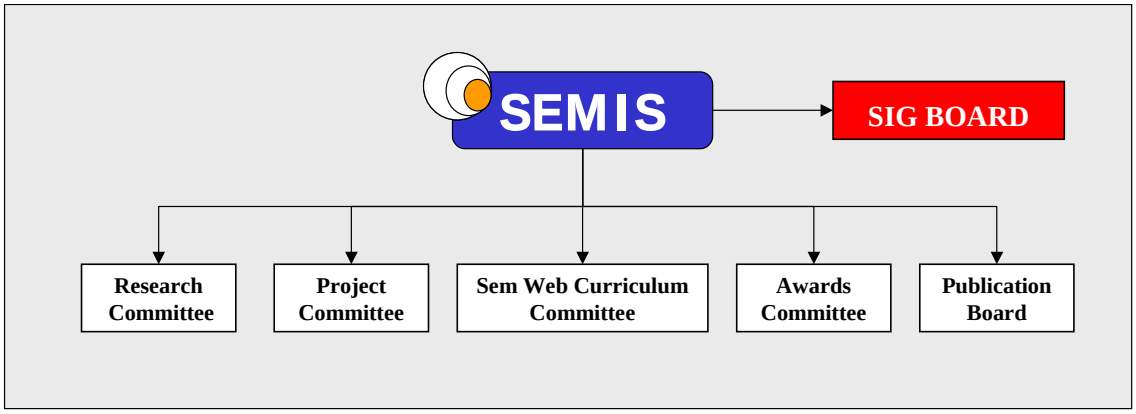
Services Offered

The Semantic Web And Information System Special Interest Group (SEMIS SIG) is an initiative, which promotes the establishment of several value adding services to AIS members. More Specifically the SIG Board will be responsible for the following services:

- A **Portal site** which will have the following subsections: Mailing List, Forum, Newsletter, Library, Download Area, Bulletin Board, Projects Announcement & Deliverables, Research Areas, PhD students corner, Master & PhD thesis available,
- A **Bimonthly Electronic Newsletter** (AIS SIGSEMIS Σigma Newsletter) covering the recent developments in Semantic Web and relevant areas
- A **Quarterly Edited Journal** (The International Journal on Semantic Web and Information Systems) publishing research papers concerning Semantic Web
- Two **Annual International Awards** for the "Best" PhD Student's Research Work concerning Semantic Web (SIG board will decide for the winner) and the **Outstanding Achievement in the Field** awarded to an internationally recognized researcher for the promotion of the discipline.

SIGSEMIS Organization

Figure 2. SIGSEMIS Organization



SIGSEMIS will exploit further an internal organization through the establishment of 6 committees:

- The overall management of SIG will be made through the **SIG board**. Currently thirty well known academics and researches constitute the SIG Board
- The **Research Committee** will undertake the critical mission to define, develop and monitor the evolution of the discipline and to concentrate on mappings, taxonomies, best practices, frameworks, adoption models, theories and other research-oriented issues of Semantic Web
- The **Project Committee** will concentrate on the integration of the research and business community. The ultimate objective of this convergence is to disseminate project outcomes (deliverables) as well as to promote project opportunities at an international level.
- The **Curriculum Committee** has a critical role in the SIG. The development of analytical curriculums concerning the education in the subjects of the semantic web will be a critical success factor for the SIG. This orientation implies that the SIG SEMIS has a clear interest for the promotion of the Semantic Web at the Academic Level (from an IS Perspective) where there is a tremendous demand for emerging courses and new text books.
- The **Awards Committee** will be responsible for the development of a strategic plan for awards. At least two annual International Awards will be announced, one concerning PhD students (a review board will be responsible for the judgment) for their research work and one for Outstanding Work in the Semantic Web Field (SIG Board will decide the winning person and possibly this award will be announced in a SIG track in an AIS conference).
- The **Publication Board** which will manage all the electronic and print publications of SEMIS SIG. In case of acceptance the SIG will develop an analytical five-years strategic plan concerning critical milestones and intended publications.

Announcing Our SIG sponsored International Journal: International Journal on Semantic Web and Information Systems

By SIG Board

Headline

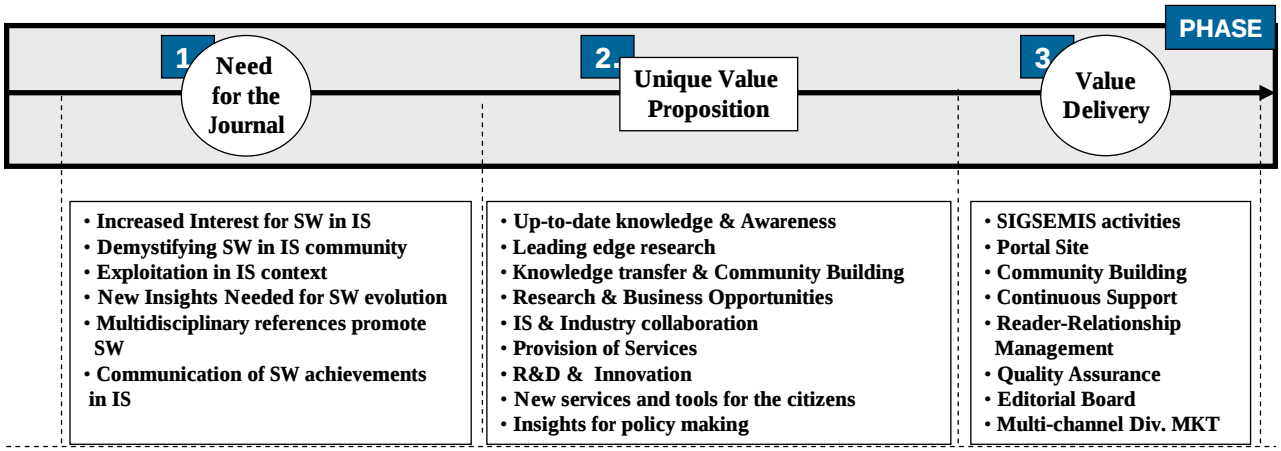
In last week IDEA Group Publishing accepted our proposal to publish in print our SIG Sponsored Peer Reviewed International Journal entitled *International Journal on Semantic Web and Information Systems*. With a clear publication strategy and a solid interest in the promotion of Knowledge Society through semantic web, our journal will be our main communication channel for research outcomes. The cultivation of the Semantic Web Vision in Information Systems Research Community will require diversified means. The IJSWIS journal, through its capable editorial team and its direct linkages to IS community will substantially initiate scientific discussions on critical issues that correspond to the ultimate objective to express and exploit meaning through information systems. In the next pages we outline the purpose, the objectives and the topics covered by the new journal and we introduce the editorial team. In the 2nd issue of our newsletter more information will be provided concerning our Journal.

The overall mission of this journal

The International Journal on Semantic Web and Information Systems is an open forum aiming to cultivate the Semantic Web vision within the Information Systems research community. In the common practice of anticipating Semantic Web as a technology- driven phenomenon, we provide a scientific insight, which reveals the practical implications and the research challenges of SW in the context of Information Systems. Our approach goes beyond the traditional research agenda of Information Systems and critical themes are analyzed through a Semantic Web perspective in horizontal and vertical pillars.

The main idea is to communicate high quality research findings in the leading edge aspects of Semantic Web and Information Systems convergence. This statement distinguishes our journal and differentiates our publishing strategy from other publications: Traditionally Semantic Web is treated as a technological phenomenon with the main emphasis on technologies, languages and tools without similar attention given to theoretical constructions or linkages to multidisciplinary references: *Our focus is on the Information Systems Discipline and we are working towards the delivery of the main implications that the Semantic Web brings to Information Systems and the Information/Knowledge Society.*

Figure 3. The motivation for the IJSWIS as a value chain



A new journal definitely has to answer three critical questions:

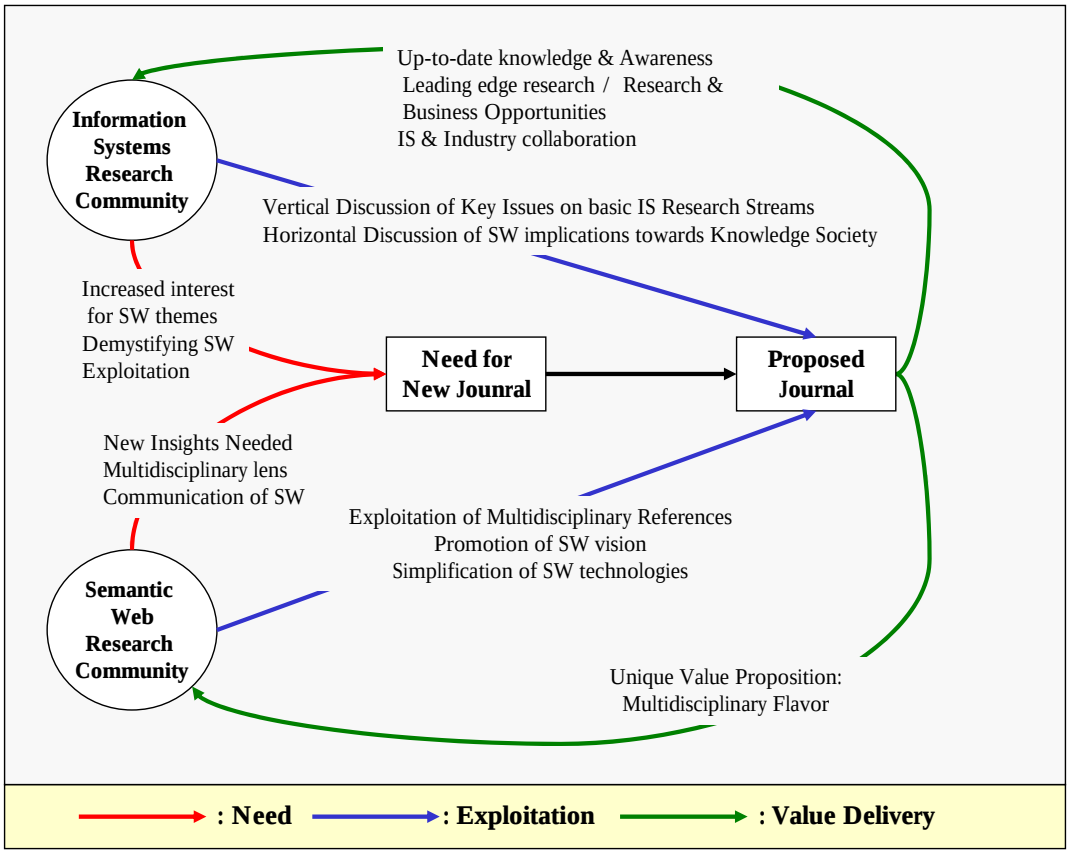
1. What is the need for the new journal?

2. What is its unique value proposition for relevant target audiences?
3. Which strategy will support the publication process towards the development of a branded, recognized and high ranked journal in the major scientific area that contributes?

We will try in the next paragraphs to provide a thorough argumentation with respect to the above questions. Figures 3 & 4 depict the main points and summarize key issues.

The motivation for the new journal is derived from two facts: First of all, during the last years there has been a tremendous evolution of the Semantic Web, which has developed a great potential for its future role. Additionally, the IS research community requires a research forum and a publication outlet which will pursue the simplification and the promotion of SW for its specific characteristics. The convergence of SW and IS provides an excellent research context which is not covered from other initiatives that pay an enormous attention to themes related to Technology or Artificial Intelligence. *The starting point for the justification of our journal is that the IS research community can play a much more active role in the research discussion of the Semantic Web and our journal intends to be the main and dominant dialogue channel for this process.*

Figure 4. The Unique Value Proposition of the journal



Our journal targets on the convergence of Semantic Web and Information systems and this strategically interesting synergy is going to be exploited for the provision of a unique value proposition to Academics, Industry and Government. High Quality Assurance, Readers/authors-Relationship Management, Fully Integrated IT support as well as communication of leading edge research and opportunities will contribute to a powerful value mix.

The third question refers to the critical strategy that will be diffused in several actions and practices. The key issue is the differentiation of our proposition in comparison with other initiatives. An integrated publication strategy is pursuing the high quality aspect of our proposition and an integrated communication marketing process works towards a 1-1 relationship management. The Association for Information Systems provides an excellent starting point for the building of a recognized and branded journal.

Our goal is to contribute in Theory, Practice and Methodology of Information Systems through the following integrated approach:

(i). **THEORY:** The development of theory in the convergence of IS and SW is organized on Key Themes such as:

- Information Systems Discipline
- E-business
- Knowledge Management
- E-learning
- Business Intelligence
- Organizational Learning
- Agents
- Adaptive Systems
- Enterprise Application Integration
- E-government
- Mobile and Wireless Technologies
- Decision Making
- Database systems
- Impact of XML
- Standards such as RDF and OWL
- Ontologies, their design and exploitation
- Process orientation
- HCI
- Multimedia
- Semantic interoperability
- Human-Machine semantic mediation

Our Goal and strategy is to prepare for all these themes, special issues that will discuss how Semantic Web poses new challenges and a new research agenda for each vertical IS theme and how it shows promise to deliver impact. Moreover the goal for each special issue will be to promote innovative propositions.

In knowledge management, the overall paradigm shift that is enabled by Semantic Web technology can be described as a transformation from knowledge-push to knowledge-pull. This is achieved by opening up previously hidden sources of information that can be cross-searched and combined in well-defined and machine-processable contexts as well as configured and controlled by each individual user.

This paradigm shift is manifesting itself as a fundamental transformation of many different fields, e.g. from teacher-centric to learner-centric e-learning, from producer-centric to consumer-centric e-business, from doctor-centric to patient-centric e-health and from authority-centric to citizen-centric e-government. Our journal will focus on making the implications of these structural changes visible and understandable to a non-technical audience, so that they can take a more active part in the discussion on how these new possibilities should be exploited in order to optimize the benefits for society as a whole.

(ii). **PRACTICE:** A core strategy in our journal is to pay close attention to the simplification of what Semantic Web means in practice. According to this, published articles will provide added value to readers by answering key questions for new semantic web enabled information systems.

(iii). **METHODOLOGY:** The research orientation of AIS as well as our SIG SEMIS research role requires the development of methodological guidelines for conducting research in Semantic Web related themes. Thus issues such as epistemology and Research Methods for Semantic Web and Information Systems research are of critical importance for our journal.

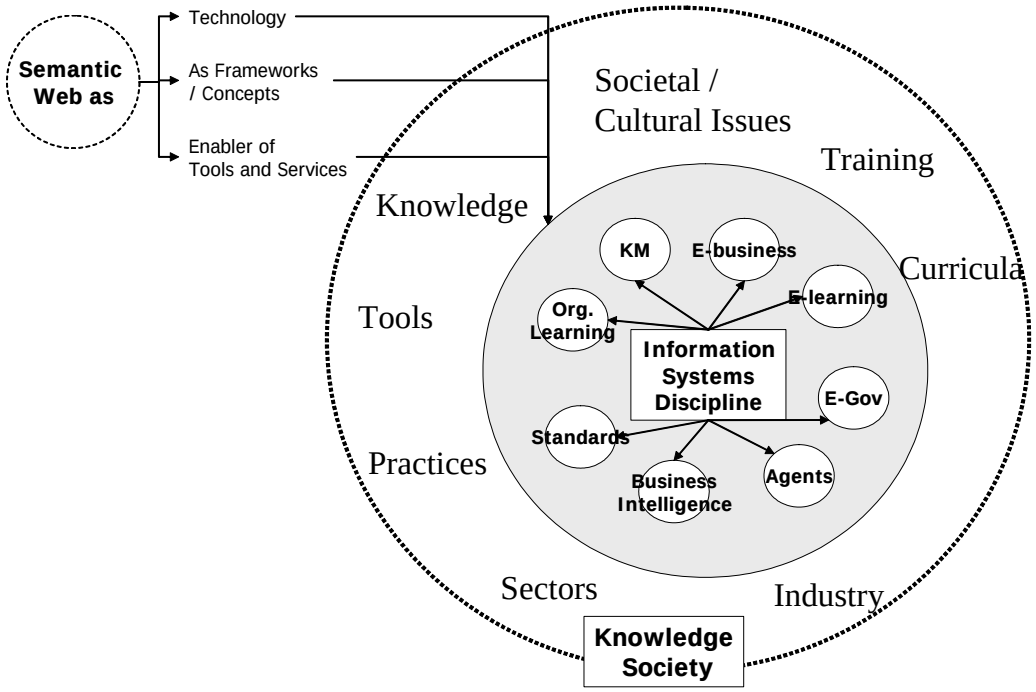
As a concluding remark our journal has a clear ultimate goal:

To provide awareness, new knowledge and significant insights within the IS research community for the current and the future role of Semantic Web, recognizing that technology is the facilitator and not the ultimate goal. Thus concepts, models, theoretical propositions as well as real world examples and innovative case studies bring forward the capabilities of the Semantic Web and the evolution of the New Generation Information Systems.

The overall scope of this journal

International Journal on Semantic Web and Information Systems promotes a knowledge transfer channel where academics, practitioners and researchers can discuss, analyze, criticize, synthesize, envisage, realize, communicate, elaborate and simplify the more than promising technology of the Semantic Web in the context of Information Systems. In the ancient Greeks' rhetoric semantic was the ultimate milestone in the quest of human mind to create and to communicate meaning.

Figure 5. The overall scope of the IJSWIS



In the Knowledge Society the exploitation of knowledge requires integration, intelligence, flexibility as well as accuracy and reference layers that are held together by a simple and ultimate goal: To utilize available technologies towards the "effective" knowledge representation and retrieval. But this fact is the basis of Information Systems: From an ontology perspective we develop conceptualizations and from these we build systems, services and practices that require an effective mix of information technology, processes, people and applications. The scope of our journal is to discuss the Semantic Web as an indissoluble whole of Technologies, Frameworks, Concepts and Practices that enable tools and services capable of supporting new innovative, effective and feasible information systems.

Figure 5, provides a graphical overview of the overall scope of IJSEMIS. Three areas are in a continuous interchange and provide value through dynamic flows and exchanges, namely the Information Systems Discipline, the IS Research Streams and the emerging Knowledge Society.

The inner circle provides a main research objective for the AIS Special Interest Group on Semantic Web and Information systems. The extensive discussion of issues and the production of new knowledge is related to the semantic web implications in main IS research streams such as E-business, Knowledge Management, E-

learning, Business Intelligence, Organizational Learning, Agents, Adaptive Systems, Enterprise Application Integration, E-government, Decision Making, HCI, Multimedia etc. We will encourage special issue proposals for most of the above 3 themes and we believe that the capacity of the SIG Board members as well as the contribution of excellent Guest Editors will provide a competitive advantage for our journal. This internal web provides a solid orientation for our publishing strategy but represents only the basic layer of our scope and potential contribution.

The outer circle represents another interesting aspect of the journal's overall scope. Our vision for the cultivation of the Semantic Web vision in IS, is closely related to the so-called Knowledge Society, where transparent technology provides new tools, services for the citizens, learners, disabled people as well as for businesses, organizations and government.

From these perspectives several more pillars outline the scope of our journal:

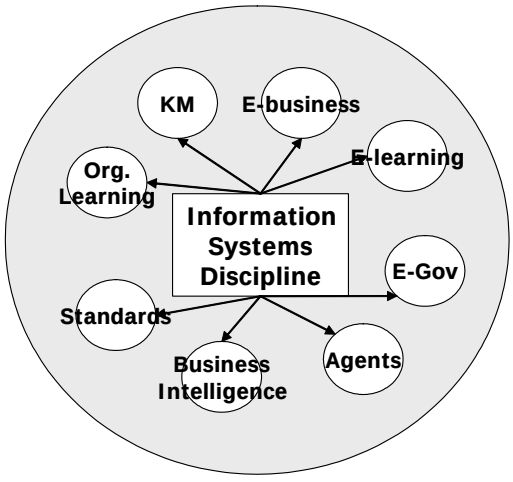
- Presentation of new semantic web enabled tools depicting the power and the capabilities of the Semantic Web.
- Sectoral/Industry Analysis of Semantic Web enabled IS (in Tourism, Knowledge Intensive Organizations, Commerce, etc).
- Discussion of innovative training activities concerning SW (community building, knowledge transfer) as well as thorough discussion on the critical theme for the inclusion of SW in IS curricula.
- Forum for epistemological, societal and cultural issues that are affected by the Semantic Web.
- Discussion of leading edge research in knowledge representation and retrieval onfive levels: Artifact, Individual, Team, Organization, Network

The above description indicates the multidisciplinary flavor of our approach. We don't believe that the SW is a solid AI issue nor that it is a "Database"-driven phenomenon. In our perception, the Semantic Web is a milestone on the road towards the ultimate human quest for efficient knowledge heritage. From philosophy we derive the axioms of dialectic and we will work hard to develop a multidisciplinary and evolutionary journal of high quality.

Possible topics to be covered by this journal

The Main themes covered in the journal include:

Figure 6. Main topics in the IJSWIS



i. Semantic Web Issues, challenges and Implications in each of the IS research streams (some indicative are presented in figure….
ii. Real world applications towards the development of the Knowledge society
iii. New Semantic Web enabled Tools for the citizen/ learner/ organization/ business
iv. New Semantic Web enabled Business Models
v. New Semantic Web enabled Information systems
vi. Integration with other disciplines
vii. Intelligent Systems
viii. Standards
ix. Semantic enabled business intelligence
x. Enterprise Application Integration
xi. Metadata-driven (bottom-up) versus ontology-driven (top-down) SW development
xii. From e-Government to e-Democracy.

Paper Submission guidelines and Evaluation process

The integrated Publication strategy is depicted in figure… Four pillars and a stream of enablers collaborate towards a high quality journal.

The first step in the process is the securing of **Qualitative Contributions**. For this purpose, 4 main streams of activities will facilitate a push strategy:

- Personal invitations to top academics and researchers will be used in the first issues in order to communicate our journal to top researchers and potential contributors.
- Special issues from SIGSEMIS Minitracks in international conferences of IS (ECIS, AMCIS, ICIS).
- Relationship building with Competence - Research centers and Leading IT companies in the semantic web field.

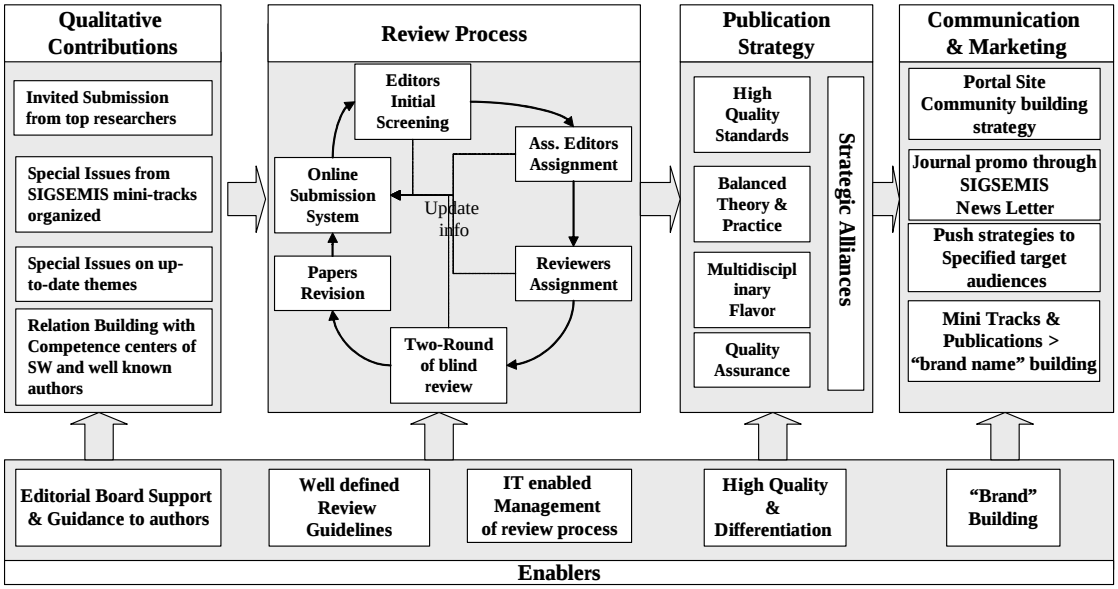


Figure 7. The integrated Publication Strategy of the IJSWIS

The Review Process will be another critical part of our strategy for competitive advantage and value adding services. A life cycle with distinct phases and an integrated support from an integrated IT infrastructure will secure a neutral, and constructive review process. All submissions will be made through the integrated Review System and will be forwarded to one of the editors. An initial screening will be used in order associate editors with relevant papers according to their main themes.

Associate Editors will use the system to assign 3 reviewers, which through a two round blind review will provide constructive comments. Associate editors in close collaboration with the editor(s) will decide on the issue of final acceptance.

Publication Strategy: The SIGSEMIS board consists of high quality academics and an excellent advisory board, which will support the journal in every stage of its preparation and development. The publication strategy has been crafted from a close understanding of the demands in the IS community for accurate and leading edge research on Semantic Web. The main emphasis is paid on quality assurance and a balanced discussion of theoretical and practical discussion of SW for IS.

Communication and Marketing: Several communication channels and a diversified marketing mix will be used in order to reach the relevant readership segments. Four main "vehicles" will be the carriers of our value proposition. First of all the Portal Site at www.sigsemis.org will support a community building strategy for all the readership audiences as well as authors and key people in the SW. A significant role in the promotion of the journal will be played by the SIGSEMIS Σigma Newsletter that will not only advertise the forthcoming

journal issues but will also host interviews with forthcoming contributors aiming to foster the interest in upcoming journal issues. Finally several other activities of SIGSEMIS such as Mini Tracks organizing and book editing and publishing will pursue the "brand" building process for our journal.

The contributions will fall into 6 different categories:

Table 1. Contribution Types

Contribution Type	Length	Key Objective	Evaluation factors
Full Research Papers	*4000-5000*	*Presentation of Research outcomes*	*Theoretical Background (20%)* *Significance of propositions (40%)* *Quality of writing (20%)* *Discussion of implications (20%)*
Research in Progress papers	*3000-3500*	*Outlining interesting future research outlets*	*Theoretical Background (30%)* *Methodology outlined (30%)* *Research Problem Description (20%)* *Quality of writing (20%)*
Literature Review Papers	*5000-7000*	*Intensive Critiques of literature / Gaps for possible research*	*Theoretical Background (40%)* *Critical thinking (20%)* *Discussion of gaps in theory (20%)* *Quality of writing (20%)*
Case Studies	*4000-5000*	*Discussion of real world implementations*	*Research Issues (30%)* *Promotion of theory & Practice (30%)* *Discussion of outcomes (20%)* *Quality of writing (20%)*
Critique of Clusters of SW projects	*5000-7000*	*Evaluation of outcomes*	*Methodologies used (50%)* *Discussion of Performance Gaps (30%)* *Quality of writing (20%)*
Visioning papers	*4000-6000*	*Crafting roadmaps for the future*	*Innovation (50%)* *Theory and Technology exploitation (20%)* *Quality of writing (20%)*

An Interview with Amit Sheth: The Information Systems Perspective on Semantic Web Research

"Semantic technology is here to stay..."

Amit P. Sheth

Professor, Computer Science, University of Georgia
Director, Large Scale Distributed Information System Lab

Sheth is an educator, researcher and entrepreneur. He joined the UGA and started the LSDIS lab in 1994. For nine years before that, he served in R&D groups at Bellcore, Unisys, and Honeywell. In August 1999, Sheth founded Taalee, Inc., a Venture Capital funded enterprise software and internet infrastructure startup based on the technology developed at the LSDIS lab. He managed Taalee as its CEO until June 2001. Following Taalee's acquisition/merger, he serves as the CTO and co-founder of Semagix, Inc (formerly Voquette, Inc). His research has led to two companies, several commercial products and many deployed applications. He has published over 150 papers and articles, given over 130 invited talks and colloquia including 19 keynotes, (co)-organized/chaired ten conferences/workshops, served on 90 program committees, etc.

Miltiadis: Dear Prof Sheth, we are really honored for your kindness to provide us this interview. I would like to start by asking you what do you think for the so-called Next Generation Web Research. Do we have to wait for much more time before we will have an evidence of real world applications?

Amit: Thanks for talking to me. There are a number of exciting "next generation" technologies on the horizon that will drastically change how we see the Web, internet, computing and communication in general. Use of semantics (as in semantic technology or in a more focused perspective taken by some in SW community) is only one of the important component technologies. Here again, semantics is only being exploited by a majority of SW researchers for a relatively narrow purpose of automating simpler- or shall we say well defined- things, although some commercial semantic technologies are being used for addressing broader objectives. In the larger scheme of things, more automated travel reservation or scheduling is not something humans care so much about – I mean it is interesting for pedagogy but given that there are so many subjective elements in making a travel arrangement, not something that would be used in the way the problem is formulated. What really matter is the real purpose of any technology – e.g., for businesses it is gains in productivity and competitiveness, for a common man it is often about luxury, relaxation or entertainment, and so on. Here automation is only a part of the equation.

As for your question on evidence of real world applications – for semantic technology (as also relevant to the vision of the SW), applications are here already – paid for and deployed! Because of business considerations, most of these are Enterprise applications rather than pan-Web applications.

As for your question on evidence of real world applications – for semantic technology (as also relevant to the vision of the SW), applications are here already—paid for and deployed! Because of business considerations, most of these are Enterprise applications rather than pan-Web applications

Miltiadis: Dear Amit, you are one of the most active people in SW and the most interesting is your IS perspective. What is your opinion for the role of SW in IS research and vise versa? What do you answer to all those people who claim that OK...SW? But where is it?

Amit: Semantics has long been recognized to be very important in IS, Databases, AI, Linguistics and many other fields. From the IS/DB perspective, I remember talking about "So Far (Schematically) yet So Near (Semantically)" in 1992, but lots of smarter people have talked about semantics for some time. More recently however, two things have happened – one positive, one potentially not so positive. The positive thing is that we have now been able to engineer semantic technology that supports large scale semantic applications, and use large populated ontologies to provide semantic underpinning. At

the same time a questionable development is a rather overwhelming importance attached to "formal semantics". Tom Gruber who brought limelight to the term ontology in early 1990s last year talked about limited value and success of formal ontologies (something he had worked on) and underscored the importance of semi-formal ontologies. I have further explored this theme, along with real world observations, in my article in Data Engineering Bulletin in December 2003. Underlying to my albeit personal view is that I must have the expressiveness in representation to meet my application requirements, rather than starting at the other end -- with computability and computing concerns – for example of an inferencing technique- and then set out to determine expressiveness and my modeling capabilities. Thankfully, query processing works for a broad variety of data and knowledge representation, and we have learned to implement that efficiently. This is a key relevance of IS/DB for semantic technology. As for SW for IS, we get to look at challenging applications where multidisciplinary approaches are necessary and hence learn to leverage techniques from allied areas. For example, to develop solutions for some of the complex problems in integration analytics that we deal with now, one needs to bring together automatic classification using machine learning, NLP techniques for document processing, and IS/DB for semi-structured and structured data management and query processing.

"*Underlying to my albeit personal view is that I must have the expressiveness in representation to meet my application requirements, rather than starting at the other end -- with computability and computing concerns—for example of an inferencing technique- and then set out to determine expressiveness and my modeling capabilities*".

Miltiadis: I have seen your excellent work in **the Large Scale Distributed Information Systems (LSDIS) Lab at the University of Georgia**. *What do we have to expect from your research in the near future?*

Amit: At the LSDIS lab, our research on semantic technology and SW have two foci. One area of focus is on techniques for knowledge discovery and semantic analytics, where we are learning to analyze a large amount of information in the form of populated ontologies and automatically extracted semantic metadata from heterogeneous content. An intermediate outcome of this work has been the development of SWETO, an ontology testbed with a large public-use populated ontology for evaluating SW tools and benchmarking semantic application.
Our second area of focus is on Semantic Web Services and Processes, where we have taken a comprehensive view of semantics (describing data, function, execution and QoS semantics) to improve all steps in the Web Process lifecycle (annotation, discovery, publication, composition, orchestration). Intermediate outcomes from this research has been WSDL-S (for semantic annotation of WSDL), and tools for semantic annotation, publication and discovery of Web services, with more to come. Domain specific knowledge is critical understanding and applying semantics, so the LSDIS Lab has collaborations in the areas of bioinformatics, geographic information systems, and homeland security.

Miltiadis: You are also the Chief Technical Officer of SEMAGIX, one of the most significant companies in SW applications. Would you like to share with us your experience related to SEMAGIX?

Amit: In August 1999, I licensed a technology resulting from research at LSDIS to start Taalee and managed it as a CEO for two years. Taalee is now Semagix after its acquisition/merger and renaming. Most companies that we thought could be our competitors have disappeared for one reason or the other, and new ones have come up. Although the bubble-bust was not easy for anyone, us included, strength and uniqueness of our technology was the primary reason for our survival. Now the stability in our development team has allowed us to engineer a terrific technology. Our CEO Larry Levy's ability to match the semantic technology and Freedom product with real market needs and his ability in signing up marquee customers in these markets have landed us several Fortune 500 customers as well as opportunities for developing incredibly exciting and demanding semantic solutions.

Although the bubble-bust was not easy for anyone, us included, strength and uniqueness of our technology was the primary reason for our survival. Now the stability in our development team has allowed us to engineer a terrific technology.

Miltiadis: Could you provide us some sound cases of SEMAGIX applications? How does industry (medical, banking sector, etc) reflects on new semantic web applications and tools?

Amit: Semagix has focused on Financial Services, Media & Entertainment, Government/Defense, and Pharmaceuticals. Most stand-out applications of Semagix Freedom are in semantic analytics involving analysis of heterogeneous

documents and databases, although semantic integration and semantic search are often involved too. Primary enabling technical capabilities utilized in these solutions involve creation of large populated domain ontologies (from multiple knowledge sources and using techniques for disambiguation), automatic semantic metadata extraction, use of main memory techniques for demanding computation such as discovering complex relationships, and being able to do all these with structured, semi-structured and unstructured data. I believe some of the domains such as bioinformatics and pharmaceuticals are inherently well suited for SW applications as domain knowledge plays an important role.

Miltiadis: Dear Amit, I know that you are a visionary with an active participation in the field. Would it be possible to outline a few (lets say 3-5) hot research areas in the field of Next Generation Web Research?
Amit: Its hard to pick a few, but here I have a few favorite ones. In research arena, these include: increasing automatic extraction/annotation of newer forms of digital media, including streaming media, broadcast TV, and sensor generated data streams; complementing semantic or thematic metadata (and corresponding domain ontologies) with spatial and temporal metadata and ontologies, and providing comprehensive spatio-temporal thematic reasoning; and extending semantics description of static aspects (such as data input and output) of resources or Web services to descriptions related to functional and execution behavior and quality of service, along with increasing semantic support for dynamic nature of Web Processes In commercial and application arena, some of the favorites include automated literature search and mining for pharmaceutical R&D; business intelligence applications of opinion and brand management for marketing; and increasing use of semantics in Web search especially as more major players compete with Google.

Miltiadis: Would you like to share your vision for semantic web for the next decade? Would it be possible to share with as a few examples of real world SW enabled tools and applications for citizens that you do believe that we will see in the next decade?

Amit: More exiting and important goals of the next generation are improving the human experience and enriching the living, and I can now see a possibility of a major shift from focus on computing to improving human experience – not only with better ability to use heterogeneous content and apply knowledge, but also to incorporate perception and pervasive computing. My view is an amalgamation of what I have seen on "experiential computing" by Ramesh Jain, "computing with words" by Lotfi Zadeh, and "humanist computing" by Jonathan Rossiter. For those focused on semantics and IS, we still need to address the difficult and fundamental problem of identifying entities (from unstructured text), semantic disambiguation and discovering (potentially fuzzy, inexact or probabilistic) relationships. And while formal representation and techniques certainly have a role, we need to find much better way for involving humans – much more than in human interfaces and visualization issues-- in any approach supporting semantics and knowledge management

"My view is an amalgamation of what I have seen on "experiential computing" by Ramesh Jain, "computing with words" by Lotfi Zadeh, and "humanist computing" by Jonathan Rossiter.

I consider myself to be among the very few lucky guys who has an opportunity to simultaneously work with entire span of research, prototyping, technology transfer, commercialization and real world application deployment.

Miltiadis: Do you think that the emergence of wireless technologies will influence our traditional approach of designing IS? Do you feel that we will soon be talking about a "ubiquitous semantic WEB"?

Amit: Two of the interesting issues that arise with wireless data applications are location awareness and increasing value of metadata due to the inability to transfer and display large amount of text. So wireless is about as different as interactive TV, streaming media or sensor data. Supporting these will provide new applications where semantics and metadata will play increasing role compared to text on terminals of wired computers.

Miltiadis: Amit, your career path in Industry and Academia is amazing. What can you glean from these experiences? What is in other words the added value on you from industry and academia?

Amit: I consider myself to be among the very few lucky guys who has an opportunity to simultaneously work with entire span of research, prototyping, technology transfer, commercialization and real world application deployment. At LSDIS, I can work with colleagues and a large group of PhD students on long term and conceptual research which allows me to collaborate with industry and provide inputs to standards activities. We have twice licensed technology resulting from our research, leading to start ups, including Semagix (earlier Taalee). At Semagix, I get to work with smart engineers –

some of whom are LSDIS alumni-- to develop a leading product in SW and architect customer specific solutions. On the same day I can work on research papers and prototypes, as well as deal with challenges of a deployment at a Fortune 500 customer. It has been incredibly exciting.

Miltiadis: Information Systems affect everybody's life. Nowadays we all discuss about the Knowledge Society? Is this utopia? What Knowledge Society means from you?

Amit: No, it is fairly certain that nothing we are seeing is a utopia. We all have tendency to get unduly excited with every new trend and fad, and after a field matures, we find out that instead of them being major life-changing technology or science, they are a step towards a continuing evolution. As to your question about the role of knowledge, we are now able to automate processes involving large knowledge bases. An average domain ontology developed by Semagix is populated by well over a million entity instances and well over a million relationship instances, typically created by integrating parts from multiple knowledge sources (some of which may be created and maintained by humans). Such a populated ontology at a center of a semantic technology is well suited for many interesting applications.

Miltiadis: Dear Amit, I asked from a student to prepare for you a profile for your work and career. In fact I asked him to summarize you (if I may say so) in few pages. He concluded his exercise by stating: "Amit Sheth delivers trust in his eyes… So I would ask you to share with as your work and life style.

Amit: I am a type A person-- Students at LSDIS and employees at Semagix do find me quite demanding ☺☹ Life style is pretty hectic, especially while directing both the LSDIS lab and Semagix R&D, sometime compromising social and family life. On the other had, when dealing with a challenging problem, such as developing a scalable technology that support and utilize semantics, one needs to have patience.

I have a few principles or guiding rules. First my own motto and guidance to my students is "learning to learn", as learning a skill or technology is not sufficient. New issues and challenges constantly arise, new techniques and technologies become relevant, so being able to learn new things, being adaptive is crucial if one wants to continue to be relevant. Second, I see opportunities as an educator, a researcher, and an entrepreneur (something I have adapted from Ramesh Jain). Correspondingly, I select activities to achieve impacts through training my PhD students to become among the best in our area of work, seek to achieve research impact through quality (as reflected in citations) rather than quantity (as reflected in number of publications), and try to achieve impact of research in terms of commercialized technology, deployed applications, and increased commerce.

First my own motto and guidance to my students is "learning to learn", … "Second, I see opportunities as an educator, a researcher, and an entrepreneur"… "I select activities to achieve impacts through training my PhD students to become among the best in our area of work, seek to achieve research impact through quality (as reflected in citations) rather than quantity (as reflected in number of publications).."

Miltiadis: You have organized several workshops, tracks and special issues on SW related themes. Do you plan anything for the near future? We heard that you would have a leading role in new International Journal on Semantic Web and Information Systems of AIS SIGSEMIS.

Amit: Berner-Lee's coining of SW was quite timely. It was time to take help of semantics for bringing more value to the huge amount of information Web grew to contain. But semantics has been investigated in IS and DB for quite some time; in 70s and 80s in the context of semantic data models, and in 80s and 90s in the context of schema and database integration (e.g., in 1993-94, Vipul Kashyap and I talked about Semantic Information Brokering involving Web and non-Web content). AI and some other fields have embraced SW readily. So I am a bit surprised that IS and DB communities have not been very enthusiastic about SW. We had an excellent workshop on DB/IS Reseach for SW and Enterprise in Amicalola to discuss just this. And in the Data Engineering Bulletin article I tried to respond to some of the hesitations and caveats that DB and IS communities have expressed regarding SW and have perhaps kept them from embracing it. May be the timing is better for SIGSEMIS to build a more vigorous community, but I hope that you will involve both formal and non-formal approaches, and both automated and human-in-the-loop approaches.

Miltiadis: Amit, I would like to share with us your thoughts for the future of SW. Is it one more bubble? What are the critical milestones for its evolution?

Amit: When thinking about SW, I focus on semantics rather than Web – semantics has been around for a long time, and Web is one context of its application. If SW is narrowly defined, ruling out anything that does not involve formal representation and inferencing, then it may not be in vogue too long. That is why, we included enterprise in our Amicalola workshop name, and that's why Ramesh Jain and I have named the book series we just started "Semantic Web and Beyond: Computing for Human Experience."

When thinking about SW, I focus on semantics rather than Web—semantics has been around for a long time, and Web is one context of its application. If SW is narrowly defined, ruling out anything that does not involve formal representation and inferencing, then it may not be in vogue too long.

Miltiadis: A few days ago I have a conversation with a colleague in the university. He wanted to learn more concerning Semantic Web and Next Generation Web Research. And basically he put me a dilemma: "Technologies or Theories?". What is your advice to a newcomer in the field? Many student members of our SIG would be interested on it.

Amit: Fundamental concepts, multidisciplinary approach (not limiting to just one of AI, KM, IR, DB, or IS), and smart engineering that scale with enterprise and Web scale data and large knowledgebases.

Miltiadis: Amit, you teach several semantic web courses on your university. Do you think that IS curriculum in the next years will incorporate such courses and in which direction?

Given the high level of student interest, I would venture to predict that this topic will be increasingly added to graduate curricula. Textbooks on SW and semantic information systems will likely be published this year, which should help.

Amit: Absolutely. I think I was the second person to teach a course on Semantic Web (and the first one that took a multidisciplinary view as the first one had a distinct AI perspective; Google: Semantic Web Course). I have already taught it thrice, and increasing number of faculty has offered a course on this topic world wide. Recently, I introduced a course on Semantic Web Services and Processes. Given the high level of student interest, I would venture to predict that this topic will be increasingly added to graduate curricula. Textbooks on SW and semantic information systems will likely be published this year, which should help.

Miltiadis: I would like to comment your perception for the role that IS community can play in SW. At the current moment the field is dominated by researchers from the AI and DB field. Do you thing that can be a convergence?

Amit: This has got to be a multi-disciplinary approach. For example Semagix Freedom uses to a varying degree techniques generally considered to be part of KR, DB, IR, and lexical analysis.

Miltiadis: How do you find the formation of the New Special Interest Group on Semantic Web and Information Systems on AIS? What role do you expect from SIG SEMIS in the forthcoming years?

Amit: It all has to do with the enthusiasm of the community and adaptive direction by its leaders. You have invested considerable energy and have managed to get broad participation in a short time. Some of us would be delighted to be available as guide or resources, but it is on the shoulder of your team.

Miltiadis: Dear Amit thank you for your time. It was an excellent talk. Would you like to state something to our readers?

Amit: We live in an exciting time. It is now possible to engineer solutions of real world value – I know this as enterprise customers are paying for these semantic solutions. The problem remains challenging – a lot has to be done and can be done. Semantic technology is here to stay.

We live in an exciting time. It is now possible to engineer solutions of real world value – I know this as enterprise customers are paying for these semantic solutions. The problem remains challenging—a lot has to be done and can be done. Semantic technology is here to stay.

Short Vita

Dr. Amit Sheth is a Professor of Computer Science and the director of the Large Scale Distributed Information Systems (LSDIS) Lab, at the University of Georgia. His primary current research interests include

- **Semantic Interoperability, Semantic Web and Global Information Systems** with applications to Digital Libraries, Video and Digital Media applications utilizing Broadband, and E-commerce, and
- **Enterprise Integration** with emphasis on Semantic Web Services, their discovery and composition; Multi-organizational Business Processes with emphasis on exception handling, and adaptation.

Sheth jointed the UGA and started the LSDIS lab in 1994. In August 1999, Sheth founded Taalee, Inc., a Venture Capital funded enterprise software and internet infrastructure startup based on the technology developed at the LSDIS lab. He managed Taalee as its CEO until June 2001 when it was acquired by Voquette, Inc. Subsequently he has been the CTO/co-Founder for Voquette, Inc. and Semagix (resulting from merger of Voquette and another company). He has also founded another high-tech company—Infocosm, Inc. He has provided consulting to a wide variety of companies involving real world solutions to the real world problems. His research has led to several commercial products and applications (including AdaptX/Harness from Bellcore based on InfoHarness project, METEOR EAppS from Infocosm based on the METEOR project, and Freedom from Semagix or Semantic Engine from Taalee based on SCORE technology and VideoAnywhere project), as well as two patents (6,311,194, 5,241,675). From 1985 to 1994, he served in R&D groups at Bellcore (now Telcordia Technologies), Unisys, and Honeywell.

Prof. Sheth has given over 130 invited talks, tutorials and professional courses, including nineteen keynote and plenary invited talks at international conferences and meetings,. He has over 150 publications, including two outstanding conference papers, and some of the most cited papers in federated databases, workflow management and semantic interoperability. He is on the editorial board of five journals. Recently he has served as a Program co-chair/Co-organizer of 2004 IEEE International Conference on Services Computing, NSF-EU Workshop on DB/IS Issues for Semantic Web and Enterprises, IEEE Digital Library Conference and WWW10 Workshop on Semantic Web. Earlier he has led six international conferences and workshop as a General/Program (Co-)Chair in the area of information system cooperation/interoperability, workflow management, and parallel and distributed information systems. He has also served twice as an ACM Lecturer, was one of the two international advisors of Japan's TARA program, and has served on over 75 program and organization committees. He received his B.E.(Hons.,1981) from B.I.T.S., Pilani, India and M.S.(1983) and Ph.D.(1985) from the Ohio State University.

Prof. Sheth enjoys mixing activities that span from conceptual research to entrepreneurship involving commercial products and real-world applications. His teaching activities have included introduction of innovative courses on Semantic Web, Semantic Enterprise, Enterprise Integration and Global Information Systems. These courses combine latest academic research with commercial state of the art research, as well as conceptual frameworks to hands on projects with practical applications. He is also involved in developing an interdisciplinary program in Masters of Internet Technology and has been on the steering committee of New Media Institute at UGA..

LSDIS lab (http://lsdis.cs.uga.edu) maintains very active collaboration with industry, and has been awarded six significant projects in the above mentioned areas with funding of over $4.5 million by NIST, NRL, NSF and industry. Along with colleagues that jointly direct some of the projects-- Profs. Krys Kochut, John Miller, and I. Budak Arpinar, the lab has usually comprised of several research assistants and visitors.

Special Issue Theme: Semantic Web Challenges for Information Systems Research Community

Table of Contents

Semantic Web and Information Systems: Exploiting the value synergy

Miltiadis Lytras
ELTRUN, The E-Business Center
Department of Management Science and Technology
Athens University of Economics and Business
mdl@eltrun.gr

Abstract

The discussion on the importance of Semantic Web for the promotion of Information systems is extremely interesting. In this first newsletter we tried through personal invitations to outline only few aspects of the convergence of SW and IS. As we already mentioned in the editorial of NEWS, the Inaugural Issue of the International Journal on Semantic Web and Information Systems Journal (at the end of 2004) will promote further this discussion. In the next pages, 7 interesting contributions provide an excellent overview for the new era of Semantic Web enabled Information Systems Research.

Introduction

In the last years the role of semantic web as the new milestone towards the ultimate objective of expressing and exploiting meaning for value adding services has been emphasized. In this context two critical issues require a multifold approach. From the one hand the required technical specification and from the other hand the strategic fit of these technologies in new business models and exploitation scenarios. The convergence of technologies and strategy will be the basis for the judgment of the importance of Semantic Web for everyone.

The transparency of Semantic Web services and their usability will support a "black box" of pervasive information provision. In the relevant literature several interesting approaches can be found for both critical issues. In our perception the communication of the importance of semantic web very often fails due to a difficulty to simplify SW technologies for the mass, and a rather soft explanation of new services and models enabled by the semantic web.

The role of SIGSEMIS is to work hard in order to communicate, to simplify, and to develop new knowledge on the Semantic Web and Information Systems convergence. This objective can be informed from different and complementary approaches.

The notion of Semantics is not a novel idea. In fact the expression and exploitation of meaning is as old as the history of the human beings. In this special issue Kim Veltman from Maastricht McLuhan Institute (MMI), European Centre for Digital Culture, Knowledge Organization and Learning Technology, Netherlands provides an excellent discussion of the historical background and outlines six challenges for SW and IS convergence. In his contribution "*Challenges for the Semantic Web and Information Systems from Culture" the following excerpt* summarizes his interesting point of view:

"…During the 20th century there were many breakthroughs in computers and information systems. Paradoxically this quest to achieve automated forms of communication also led to a great simplification of semantic problems and a serious dumbing-down with respect to solutions. As a first challenge, this paper argues that we need to integrate what we left behind if we are to go ahead…"

Benjamins, Contreras and Corcho from Intelligent Software Components, S.A., Spain and Gómez-Pérez from Universidad Politécnica de Madrid., Spain discuss the critical theme of information overload and claim that Semantic Web provides a catalytic solution to this problem. In their contribution "*Six Challenges for the Semantic Web",* they provide a three-dimensional space where Information Overload, Complexity and Autonomy define a justification framework for SW deployment. They conclude that the promotion of SW is closely related to the solution of 6 problems namely *Availability of Content, Ontology availability, Development and Evolution, scalability, Multilinguality, Visuability and Stability of Semantic Web languages*.

Chebotko, Shiyong Lu and Fotouhi, from the Department of Computer Science, Wayne State University, USA emphasize on the importance of annotation management and ontologies as two critical enablers of SW. Their involvement in SW is critically expoited and they provide several more challenges towards the exploitation of the potential of SW. In their short article *"Challenges for Information Systems Towards The Semantic Web"* they build upon the comment of Tim Berners-Lee "*It's not unlike the early days of the Web, when once people saw how it worked, they understood its power. We're entering that phase now, where people can see the beginnings of the Semantic Web at work*". More specifically they outline four more challenges:

(1) the challenge of the development of a domain ontology; (2) the challenge of ontology mapping, alignment and merging; (3) the challenge of annotation management; and (4) the challenge of ontology-based information retrieval.

Jeffrey Nickerson from Stevens Institute of Technology, USA provided kindly a very short article but full of interesting ideas. His critical point of view triggers reader to think about the underpinnings of semantic web. He summarizes in an excellent way the ultimate objective of SIGSEMIS: "*Those engaged in work on the semantic web will need to look hard at how meanings are actually grounded. On one pole are those who seek to create a fully general array of machine-discoverable and machine-compassable services. On the other are those who think that practical systems come from people hashing out provisional agreements - and implementing them. There is surely a promising territory in between, in which we call on people to ground meaning, and use the semantic web to allow agreements - and services - to be more easily modified*". Our SIGSEMIS is willing to play a critical role in this territory.

The second part of this "special issue" consists of articles that discuss the importance of semantic web in specific IS domains like E-government, E-learning, Multimedia Management etc.

Hüsemann and Vossen from the European Research Center for Information Systems at the University of Muenster in Germany reflect on the critical theme of multimedia object management. In their article "*Ontology-Driven Multimedia Object Management for Private Users –Overview and Research Issues"* they provide their experience from an interesting project and they outline their research strategy and contribution towards an ontology-driven multimedia framework. This area of research in the next years will gain an extraordinary interest since meaning cannot be considered as resident only in textual formats of content.

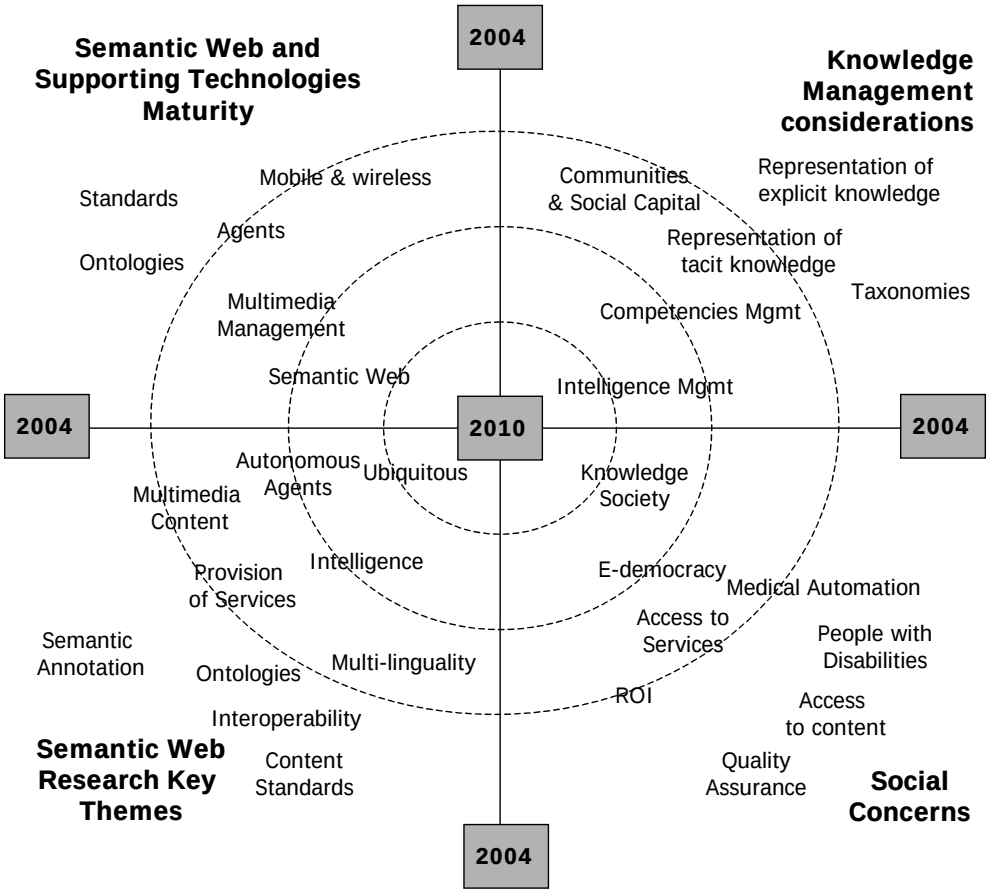
Learning is another extremely interesting application field for SW. Miguel-Angel Sicilia from the Computer Science Department in University of Alcalá, Spain provides a thorough discussion for the Automation of E-learning through SW. In his article "*The Role of the Semantic Web in the Automation of e-Learning"* emphasizes on the importance of SW towards effective specification of learning content and meaning metadata and consequently on value processes of Learning Management Systems.

Finally Ralf Klischewski from the Informatics Department at Hamburg University, Germany, in his article "*Semantic Web for e-Government – a Research Agenda*", provides interesting insights for the expoitation of SW in the context of E-government. Copying for his article: "*Semantic Web has started out with a document-oriented approach; the basic idea was to annotate Web pages with semantic markup. But the informational resources on e-government websites can comprise a variety of electronic "things": fine-grained information elements (e.g. opening hours), multimedia items (pictures, maps, videos), files (for downloads, e.g. documents, forms, client applications), interfaces to transactions (processed by backend applications), links (to other sites and their resources), services (e.g. authentication), user related objects (e.g. user profiles) etc. All of these are meaningful for site users, all of these are searched for and being used – thus all of these are candidates for semantic markup*". This variety of information resources undoubtedly poses significant challenges for IS and SW convergence.

The limited time for the preparation of this special issue didn't permitted, to incorporate and more contributions from other interesting application areas of SW in the context of IS (e.g. Medical IS, Interactive Television, Mobile and Wireless Applications, Semantic E-business etc). In the forthcomings newsletters several interesting approaches will be presented in detail.

Next figure summarizes initial perceptions for the role of semantic web but we will discussed them in detail in issue 2 of AIS SIGSEMIS Bulletin.

A Road Map towards maturity in SW perceptions of IS



Six Challenges for the Semantic Web

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

[1]Intelligent Software Components, S.A.
www.isoco.com (Spain)
{richard,jcontreras,ocorcho}@isoco.com
[2]. Facultad de Informática. Universidad Politécnica de Madrid
{asun}@fi.upm.es

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

"The problem of Information Overload can be partly tackled by adding intelligence to the web. Software agents could manifest various levels of intelligent behaviour from simply reactive to adaptive and learning behaviour, where agents actually learn what users like and dislike".

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

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

Figure 1 illustrates three dimensions relevant in this context: Information Overload, Task Delegation, and Complexity. The problem of Information Overload can be partly tackled by adding intelligence to the web. Software agents could manifest various levels of intelligent behaviour from simply reactive to adaptive and learning behaviour, where agents actually learn what users like and dislike. This would shield users from irrelevant information, who would only be 'bothered' for information of real value. Instead of the current ubiquitous *pull* paradigm that requires users to actively look for information and to execute programs, we should work towards delegating those tasks to autonomous software agents. Finally, the tasks users will want to execute will become increasingly complex. Software agents should 'learn' to function in 'social' environments and where necessary collaborate, compete or negotiate with other agents. The quality and usability of the Semantic Web infrastrucure will depend on advances on all three dimensions

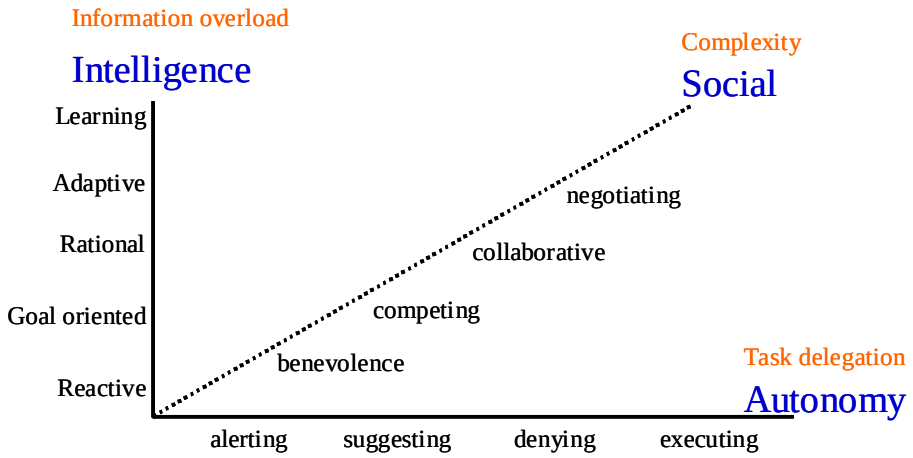


Fig. 1. Three dimensions relevant for the Semantic Web

Figure 1 illustrates three dimensions relevant in this context: Information Overload, Task Delegation, and Complexity. The problem of Information Overload can be partly tackled by adding intelligence to the web. Software agents could manifest various levels of intelligent behaviour from simply reactive to adaptive and learning behaviour, where agents actually learn what users like and dislike. This would shield users from irrelevant information, who would only be 'bothered' for information of real value. Instead of the current ubiquitous *pull* paradigm that requires users to actively look for information and to execute programs, we should work towards delegating those tasks to autonomous software agents. Finally, the tasks users will want to execute will become increasingly complex. Software agents should 'learn' to function in 'social' environments and where necessary collaborate, compete or negotiate with other agents. The quality and usability of the Semantic Web infrastrucure will depend on advances on all three dimensions

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

- The **availability of content**. Currently, there is little Semantic Web content available. Existing web content should be migrated to Semantic Web content including static HTML pages, but also dynamic content, multimedia and web services.
- **Ontology availability, development and evolution.** Ontologies will become a key piece, as they allow expliciting the semantics of Semantic Web content. A big effort must be made in the creation of common widely used ontologies for the Semantic Web, on the provision of adequate infrastructure for ontology development, change management and mapping, and, in this distributed web environment, on the adequate control of the evolution of ontologies and the annotations referring to them.
- **Scalability**. A significant effort must be made to organize Semantic Web content, store it and provide the necessary mechanisms to find it. All these tasks must be performed and coordinated in a scalable manner, as these solutions should be prepared for the huge growth of the Semantic Web
- **Multilinguality**. This problem already exists in the current Web, and should also be tackled in the Semantic Web. Any Semantic Web approach should provide facilities to access information in several languages, allowing the creation and access to SW content independently of the native language of content providers and users.
- **Visualization**. Intuitive visualization of Semantic Web content will become more and more important to solve the increasing amount of information overload, as users will demand the easy recognition of relevant content for their purposes. New techniques must be explored that differ from the usual hypertext structure visualization of the current web.
- **Stability of Semantic Web languages**. Finally, standardization efforts have already been performed in this area, which have recently provided as a result the W3C recommendations of RDF Schema and OWL. And new standardization efforts are being currently performed for the provision of rule-based support on top of these languages.

Challenges for Information Systems Towards The Semantic Web

Artem Chebotko, Shiyong Lu, Farshad Fotouhi
Department of Computer Science
Wayne State University, Detroit, MI 48202
{artem, shiyong, fotouhi}@cs.wayne.edu

The Semantic Web [1, 2, 3] is the next generation Web, in which information is structured with well-defined semantics, enabling better cooperation of machine and human effort. The Semantic Web is not a replacement, but an extension of the current Web.

The Extensible Markup Language (XML) [4] will play an important role as a syntactic foundation for the Semantic Web. XML provides a set of constructs for representing data on the Web, but XML only supports limited semantics via its nested structure. To illustrate the limitations of XML, consider the two XML documents shown in Figure 1. Suppose we would like to retrieve all the *persons* from these two documents. To return the complete results, the query processor needs to understand that each "student" or "professor" is also a "person". In addition, to integrate the search results from both documents, the query processor also needs to understand that "familyName" in document 1 is equivalent to "lastName" in document 2, but "date" in document 1 is different from document 2 — while "date" in document 1 means a date of birth, "date" in document 2 means the date when a student or a professor enters a university. These semantic relationships between different elements are not supported by XML, and current XML query languages such as XQuery and XPath do not support the correct processing of the above query.

```
<person id="1">
        <firstName>Emily</firstName>
        <familyName>Lu</familyName>
        <date>2001-01-15</date>
</person>
<person id="2" >
        <firstName>Shiyong</firstName>
        <familyName>Lu</familyName>
        <date>1972-08-14</date>
</person>
```

```
<student id="3">
        <firstName>Artem</firstName>
        <lastName>Chebotko</lastName>
        <date>Fall 2003</date>
        <gpa>3.0</gpa>
</student>
<professor id="4">
        <firstName>Farshad</firstName>
        <lastName>Fotouhi</lastName>
    <date>Fall 2003</date>
</professor>
```

Figure 9: XML document 1 and 2

To overcome the limitations of XML, ontologies [5] will be used to represent various concepts and the relationships between these concepts for a domain. For example, one can represent that a *"student" is subClassOf "person"*, and *"familyName" equivalentProperty "lastName"* in an ontology and use the ontology to guide the query processor to process the above query properly.

"…We conclude our article with a comment from Tim Berners-Lee on the state-of-the-art of the Semantic Web: "*It's not unlike the early days of the Web, when once people saw how it worked, they understood its power. We're entering that phase now, where people can see the beginnings of the Semantic Web at work****""***

What challenges does the Semantic Web pose to information systems? In this article, we share our perspective to this question based on our experience with building the Semantic Web for endangered languages [6]. We focus our discussion on the following challenges.

The challenge of the development of a domain ontology. Ontologies now play an important role in enabling the Semantic Web. Semantic Web communities develop ontologies in their domains, which involves many experts in the same domain and each of them might have his own perspective (a social challenge). This requires a collaborative ontology development environment (a technical challenge) that will enable version control, proposal and release control, and coordination and collaboration support. The development of such an environment is a challenge. Most of

the ontology development tools today, like Protégé-2000 [7], are personal ontology editors and they lack these functionalities. To the best of our knowledge, the only tool that has the collaboration support is OntoEdit [8], but it needs further improvements, such as rights- and user- management layers.

The challenge of ontology mapping, alignment and merging. Over the past few years, people have come to the consensus that even in one domain, it is very difficult to enforce a single ontology to be used for each data source. Instead, people should have the full freedom to use their own proprietary ontology to annotate their data source, and then, if they are willing, provide additional mapping to map them to a standard (central) domain ontology to support data interoperability and queries across data sets. This mapping is a challenge because there might exist heterogeneities between ontologies: syntactic, schematic and semantic. The mapping process might include not only ontology alignment to make ontologies coherent, but also ontology merging to add new terms in a central ontology. For further information on available algorithms and tools in this research area, please refer to [9, 10].

The challenge of annotation management. One challenge of annotation management is the integration of a domain ontology and an annotation tool. For each domain, an annotation tool should be customized based on a given ontology and the requirements of the users. For example, linguists require the support of International Phonetic Alphabet in an annotation tool and this requirement is not specified in the domain ontology. A second challenge is the development of a universal but customizable annotation tool for several domains since it requires an open architecture and the support of common annotation functionalities needed for various domains. A third challenge is the development of a shared annotation capability, where the support of concurrency control is needed. Shared data annotation is especially efficient when users simultaneously annotate data segments that do not intersect. For example, one person annotates the gestures of a speaker, while another person adds linguistic characteristics of the speech. Two different ontologies can be used for this purpose and each person may be familiar with only one of them. User rights play an important role in client co-work coordination. One of the projects in the shared annotation direction is Annotea [11], which implements a client-server architecture, where server stores annotated data and manages its access. Finally, the consistency maintenance between annotations and the corresponding ontology is also a challenge. When ontology evolves, the annotations might need to change accordingly.

The challenge of ontology-based information retrieval. Annotated data is not useful if one cannot search through it. One of the promises of the Semantic Web is high precision. Search engines now should exploit available semantics and ontology reasoning to return not only precise results, but also specify meaningful relationships between them. New opportunities also require new approaches to query refinement and user interface tactics [12]. But the major challenge is searching across data sets annotated using different ontologies. Generally, there can be several ontologies for one domain since each domain can be modeled by several ontologies or a domain may require the usage of several ontologies. As a result, not only ontology mapping is required, but also user query mapping will be needed.

Conclusions and outlook

In this article, we discussed some challenges that we experienced working on the Semantic Web for linguistics. These are (1) the challenge of the development of a domain ontology; (2) the challenge of ontology mapping, alignment and merging; (3) the challenge of annotation management; and (4) the challenge of ontology-based information retrieval. We conclude our article with a comment from Tim Berners-Lee on the state-of-the-art of the Semantic Web: “It's not unlike the early days of the Web, when once people saw how it worked, they understood its power. We're entering that phase now, where people can see the beginnings of the Semantic Web at work.”

References

1. T. Berners-Lee. “Semantic Web Road Map”. IW3C Design Issues. Cambridge, MA: W3C. Available at: http://www.w3.org/DesignIssues/Semantic.html. 1998.
2. T. Berners-Lee, James Hendler and Ora Lassila. “The Semantic Web”. Scientific American, May 2001.
3. S. Lu, M. Dong and F. Fotouhi, "The Semantic Web: Opportunities and Challenges for Next-Generation Web Applications". International Journal of Information Research, 7(4), 2002.

4. T. Bray, J. Paoli, C. Sperberg-McQueen, E. Maler, and F. Yergeau, "Extensible Markup Language (XML) 1.0 (Third Edition) W3C Recommendation 04 February 2004". 2004.
5. T. Gruber, "A translation approach to portable ontology specifications." Knowledge Acquisition, 6(2), 199-221. 1993.
6. S. Lu, D. Liu, F. Fotouhi, M. Dong, R. Reynolds, A. Aristar, M. Ratliff, G. Nathan, J. Tan, and R. Powell, "Language Engineering For The Semantic Web: A Digital Library For Endangered Languages", International Journal of Information Research, 2004.
7. Protégé-2000. Protégé Project, http://protege.stanford.edu/index.html. 2004.
8. Y. Sure, M. Erdmann, J. Angele, S. Staab, R. Studer, D. Wenke, "OntoEdit: Collaborative Ontology Development for the Semantic Web". ISWC 2002.
9. D. McGuinness, R. Fikes, J. Rice, and S. Wilder, "An environment for merging and testing large ontologies". KR-2000.
10. N. Noy and M. Musen, "PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment". AAAI 2000.
11. M. Koivunen, R. Swick, E. Prud'hommeaux. "Annotea Shared bookmarks". KCAP 2003.
12. E. García and M. Sicilia, "User Interface Tactics in Ontology-Based Information Seeking". PsychNology Journal, 2003.

Ideas on the challenge of the semantic web

By Jeffrey V. Nickerson, Stevens Institute of Technology, USA
jnickerson@stevens.edu

Meaning comes through the process of interpretation, not the process of parsing. So can a Semantic web ever live up to its name? The current web works because web pages are read by humans. One can argue that interpretation can be grounded in the people defining the XML, the rules, and the parsers, yet this argument is seldom made - the general, and mistaken perception is that the machines will just talk to each other, culling meaning from definitions and rules.

"…Those engaged in work on the semantic web will need to look hard at how meanings are actually grounded. On one pole are those who seek to create a fully general array of machine-discoverable and machine-composable services. On the other are those who think that practical systems come from people hashing out provisional agreements - and implementing them. There is surely a promising territory in between…"

Also, there are indications that visions of automated discovery of services may be overblown. We do have DNS - which allows discovery of sites through names. But the UDDI standard - which allows discovery of services - is not being used. How is commerce really happening? It is happening through industry groups, which agree on shared processes, events, messages - and implement something. General discovery and arbitrary composition may not be necessary after the humans agree what they are going to do.

Those engaged in work on the semantic web will need to look hard at how meanings are actually grounded. On one pole are those who seek to create a fully general array of machine-discoverable and machine-composable services. On the other are those who think that practical systems come from people hashing out provisional agreements - and implementing them. There is surely a promising territory in between, in which we call on people to ground meaning, and use the semantic web to allow agreements - and services - to be more easily modified.

Ontology-Driven Multimedia Object Management for Private Users – Overview and Research Issues

Bodo Hüsemann, Gottfried Vossen
European Research Center for Information Systems
University of Muenster, Germany

The rapid evolution of digital technology continues to connect or even substitute most of the analogous devices found in private households these days. Digital consumer electronics (e.g., DVD recorders, MP3 players, digital cameras etc.) are broadly available in the mass-market and produce tremendous amounts of digital images, music, movies, and other (including traditional) types of media. In the same instant, home networking technology interconnects multimedia devices and provides omni-present access to the Internet. Thus, the management of multimedia objects in a private and personal environment is a task of increasing importance especially for private users who need to archive, organize, and search their multimedia collections in an appropriate fashion. Moreover, users want to be able to access these collections from a variety of devices, and seek to integrate them with information available on the Web, or simply share documents with other users.

"… the management of multimedia objects in a private and personal environment is a task of increasing importance especially for private users who need to archive, organize, and search their multimedia collections in an appropriate fashion. Moreover, users want to be able to access these collections from a variety of devices, and seek to integrate them with information available on the Web, or simply share documents with other users".

At present, users typically arrange their multimedia collections in file systems, which provide poor naming mechanisms and hierarchical directory structures for organization and searching. Although this approach looks sufficient at first sight there are many problems involved which complicate the management of large multimedia collections:

- The logical organisation is bound to the underlying physical storage system;
- categorisation is limited to strict classification hierarchies;
- document-inherent metadata (e.g., ID3 tags in MP3 files) remains unused and is only available to media-specific applications (e.g., an MP3 player);
- identification based on file names alone is often not globally consistent (e.g., duplicates are possible).

Current multimedia management applications (e.g., ThumbsPlus[1]) provide databases to index collections with user-defined keyword annotations and media-specific metadata import. Although a step in the right direction, these systems use proprietary databases without access to the stored organisation structures or to metadata outside the application.

In this context the *Semantic Web* provides data models to describe Web resources (e.g., Web pages) with application-independent languages suchas the Resource Description Framework (RDF) [MM04] which are also applicable to multimedia resources. The vocabulary used to describe documents can be specified in terms of ontologies, where each description term and its semantic relation to other terms are defined. "*Ontologies* provide a shared and common understanding of a domain that can be communicated between people and application systems" [KFHH00], and thus ontologies facilitate the sharing and exchange of information about multimedia between applications and users. Accordingly, a *multimedia ontology* comprises a shared vocabulary

[1] http://www.cerious.com/

to describe multimedia documents and their organisation in a structured way such that users and applications can process the descriptions with reference to a common understanding specified in ontologies.

In Figure 1 we propose an ontology-driven multimedia framework where different digital devices are connected in a private home network. Each of the displayed devices (e.g., PDA, laptop computer, digital camera, etc.) has its own storage volumes A-E such that multimedia documents can be distributed over different network locations. A central Ontology *Multimedia Database* (OMDB) stores all information about the available multimedia documents in the network. The OMDB provides Web-based interfaces to browse and search for multimedia documents in the network independently of their actual physical location. Using access to other information resources of the Semantic Web it is possible to integrate additional information sharing a common ontology for multimedia.

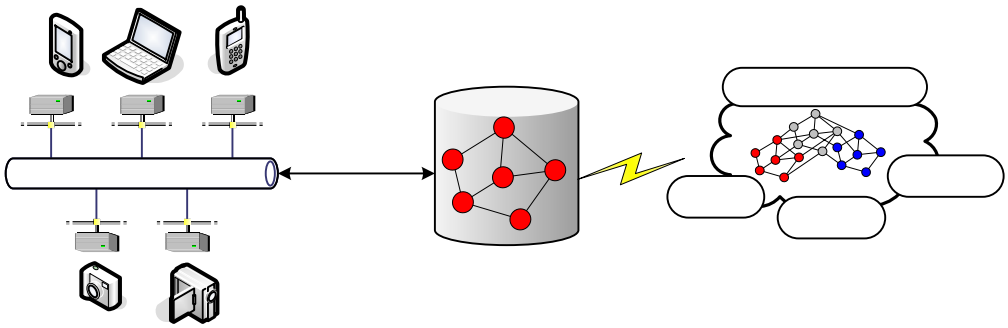


Figure 1 Overview of an ontology-driven multimedia framework.

To come up with an ontology-driven multimedia solution for private users we are currently focusing our research efforts in the following areas:

1. Development of a *reference multimedia ontology* which comprises basic terms for the semantic description of multimedia metadata and its local organisation. The ontology is targeted at the special requirements of private users.
2. Development of an *ontolog- driven multimedia framework* to support the management of large multimedia collections with Semantic Web technology.
3. Evaluation of available *ontology design methods* and their implications for the definition of a well-formed multimedia ontology.

The literature concerning multimedia ontologies so far focuses on professional application domains typical to commercial users. An example is [Hun02], where the authors propose an ontology for the description of MPEG7 metadata, which provides technical information common to professional video material. Our ontology, on the other hand, will describe multimedia metadata relevant to private contexts which includes information about music genres, film actors, or the content of digital photos. Thus we aim at incorporating existing metadata available in MP3, PDF, or JPEG files commonly found in private multimedia collections. The ontology will represent a reference for private users, which can be extended with own concepts and categorisation information.

Ontology design is a nontrivial task and requires methodological foundation based on proper design criteria and methods. Much research has been done in this area starting 1990 with Lenat and Guha [LG90], who published general guidelines for ontology design involved in the CYC project. Other approaches have been described (for an overview, see [CFP03]), of which METHONTOLOGY [LGS99] and ONTO-KNOWLEDGE [SS02] represent the state-of-the-art in ontology design based on techniques used in knowledge management. The existing approaches focus on the design of heavy-weight ontologies which are subject to advanced reasoning techniques based on fully-fledged logical theories. These requirements are uncommon for light-weight multimedia ontologies described in our application domain. Here, ontologies are used as advanced semantic data models for the description of multimedia metadata and its organisation. As a consequence, we are working on a database-oriented approach to ontology design the follows traditional methods of database

design as described, for example, in [BCN92, Vos00]. This development process is based on the common four sequential steps of database design starting with requirements analysis and specification, followed by conceptual design, logical design, and physical design. The aim of this research area is the development of a design method for light-weight ontologies which leads to well-formed ontologies according to testable criteria.

[BCN92] Batini, C., Ceri, S., Navathe, S.B.: "Conceptual Database Design: An Entity-Relationship Approach". The Benjamin/Cummings Pub. (1992)

[CFP03] Óscar Corcho, Mariano Fernández-López, Asunción Gómez-Pérez: Methodologies, tools and languages for building ontologies: Where is their meeting point? Data Knowl. Eng. 46(1): 41-64 (2003)

[FGJ97] Fernandez, M., Gomez-Perez, A., Juristo, N.: METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Ontological Engineering: Papers from the 1997 AAAI Spring Symposium. Technical Report SS-97-06, AAAI Press (1997)

[KFH+00] M. Klein, D. Fensel, F. van Harmelen, and I. Horrocks: The relation between ontologies and schema-languages: Translating OIL-specifications in XML-Schema. In Proceedings of the Workshop on Applications of Ontologies and Problem-solving Methods, 14th European Conference on Artificial Intelligence ECAI'00, Berlin, Germany August 20-25, 2000.

[Hun02] J.Hunter, "Enhancing MPEG-7 and MPEG-21 Semantic Interoperability through a Core Ontology", submitted to IEEE Transactions on Circuits and Systems for Video Technology, Feb 2002

[LG90] D.B. Lenat and R.V. Guha. Building Large Knowledge-based Systems, AddisonWesley 1990.

[LGS99] M.F. Lopez, A. Gomez-Perez, J.P. Sierra: "Building a Chemical Ontology Using Methontology and the Ontology Design Environment", IEEE Intelligent Systems, Vol. 14, No. 1, Special Issue on Ontologies, January/February 1999

[MM04] Manola, F., Miller E.: RDF Primer. W3C Recommendation Series Editor: Brian McBride (Hewlett-Packard Laboratories). 10-February-2004. Latest version URL: http://www.w3.org/TR/rdf-primer/

[SS02] Y. Sure, R. Studer. On-To-Knowledge Methodology - Final Version. Institute AIFB, University of Karlsruhe, On-To-Knowledge Deliverable 18, 2002.

[Vos00] Vossen, G.: Data models, database languages, and database management systems (in German). 4th edition, Oldenbourg, 2000.

Webs, Grids and Knowledge Spaces ? Programmes, Projects and Prospects ?

Hans-Georg Stork
European Commission, DG Information Society
DG Information Society
Hans-Georg.Stork@cec.eu.int

Abstract: Many believe that today's Web has not yet reached the full potential which globally distributed systems may achieve in terms of information access and use. Realizing this potential may indeed turn the Web into a vast knowledge and service space. We discuss some of the issues involved and present a number of activities initiated and supported by the European Commission that are likely to make significant contributions towards attaining this goal.

Full paper at: http://www.jucs.org/jucs_8_9/webs_grids_and_knowledge

The Role of the Semantic Web in the Automation of e-Learning

Miguel-Angel Sicilia
Computer Science Department, University of Alcalá
Alcalá de Henares (Madrid) Spain
msicilia@uah.es

Semantic Web technologies can be seen as an opportunity to enhance the metadata associated to learning materials, expanding the possibilities of current *e-Learning* specifications and standards (Duval et al., 2002). The key element of such kind of enhancements is the provision of *learning object* metadata records connected to or integrated inside formal *ontologies*, thus enabling certain forms of "reasoning" inside rich knowledge representation structures. As a matter of fact, a considerable amount of recent research exists about the integration of ontologies in e-learning settings – e.g. a survey can be found in (Lytras et al., 2003).

"… The reformulation of current metadata schemas in ontology description languages requires also the provision of semantic interpretations oriented to providing higher degrees of "machine-understandability". For example, mandatory and "recommended" conditions on target users should be clearly separated, and the intended outcomes of a learning object should be expressed through ontology elements in a way that enables by itself the automated design of personalized learning paths"

Nonetheless, the use of ontologies by itself is not enough to effectively produce a shift in the level of standardized automation provided by *Learning Management Systems* (LMS). The reason is that effective standardization oriented to support automated or semi-automated behaviours requires not only formal, unambiguous and rich languages, but consistent *description practices* oriented to enable reuse through automated means. In fact, recent empirical studies (Pagés et al., 2003) point out to the necessity of providing more disciplined approaches to metadata specification to come up with metadata records that have enough quality to be used by software systems as part of the automation of LMS processes.

The reformulation of current metadata schemas in ontology description languages requires also the provision of semantic interpretations oriented to providing higher degrees of "machine-understandability". For example, mandatory and "recommended" conditions on target users should be clearly separated, and the intended outcomes of a learning object should be expressed through ontology elements in a way that enables by itself the automated design of personalized learning paths.

In addition, an LMS is a particular type of information system that is intended to give support to individual and organizational learning needs. In consequence, the scope of processes that are supposed to be totally or partially LMS-controlled include business processes of search and acquisition of learning objects, and also their selection, personalization, composition and publishing, among others, which clearly exceeds the current scope of e-learning standards. Semantic Web research is now working on enabling technologies that would be able to solve such complex automation requirements. An example of such enabling technologies is the *Web Service Modelling Ontology* (WSMO)[2] project, which provides the architectural framework for the expression of needs and capabilities and their matching according to logic languages. But these technologies require that (a) learning objects are properly and consistently annotated, and that (b) complex processes are specified in a flexible and unambiguous way.

[2] http://www.wsmo.org/

The first need (a) calls for further studies and proposals about the design and evaluation of metadata. A recent example of such specification approaches is learning object *design by contract* (Sicilia & Sánchez, 2003), which uses pre- and post-conditions as a way to specify requirements about the learner and her context, and about the intended outcomes of the object, respectively. Design by contract also addresses learning object relationships (Sánchez & Sicilia, 2004), and includes several levels of strength of requirement and run-time considerations explicit to learning object composition processes. Such run-time semantics can be expressed through ontologies that explain the intended behaviour for metadata.

The second need (b) entails an intensive specification effort in levels of conformance, with details about required meta-information, and conditions and effects of the enactment of each type of process. The concept of *semantic conformance profile* (Sicilia et al., 2004) has been recently proposed as a flexible specification technique for processes that entail the participation of a LMS. These profiles determine explicitly where ontologies play a role in each profile, and what requirements are needed to enable each level of automation. For example, if categorizations in the form of ontology concepts are used to describe the expected outcomes of the learning object (along with other conditions), compatibility with the composition profile C-CMP is possible, resulting in a concrete degree of "intelligence" in learning object composition. More sophisticated profiles may *extend* or *use* C-CMP to add, for example, the consideration of learning styles or abilities in the composition process. As a result, LMSs would be characterized by the profiles they are compliant to, and Semantic Web extensions to e-learning could evolve in an incremental, stepwise fashion.

The ideas described so far are only a fraction of the hard work ahead to make evolve current LMS technology towards consistent Semantic Web behaviors that automate as much as possible the design and evaluation of learning experiences. In the current e-learning context of increasing standardization and convergence, Semantic Web technologies should become an essential tool in the specification of both content metadata and LMS processes.

References

Duval, E., Hodgins, W., Sutton, S.A., & Weibel, S. (2002). Metadata principles and practicalities. *D-Lib Magazine*, 8(4).

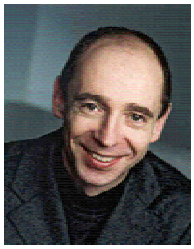
Lytras, M., Tsilira, A., Themistocleous, M.G. (2003). Towards the semantic e-Learning: an Ontological Oriented Discussion of the new research agenda in e-Learning. In *Proceedings of the Ninth Americas Conference on Information Systems*.

Pagés, C., Sicilia, M.A., García, E., Martínez, J.J., Gutiérrez, J.M. (2003) On The Evaluation of Completeness of Learning Object Metadata in Open Repositories. In *Proceedings of the Second International Conference on Multimedia and Information & Communication Technologies in Education* (m-ICTE 2003), 1760-1764.

Sicilia, M.A., Sánchez, S. (2003). On the concept of learning object "Design by Contract". *WSEAS Transactions on Systems*, 2(3), 612-617.

Sicilia, M.A., Pagés, C., García, E., Sánchez, S and Rius, A. (2004). Specifying semantic conformance profiles in reusable learning object metadata. In *Proceedings of the 5th International Conference on Information Technology Based Higher Education and Training*: ITHET 2004. Istambul, Turkey (forthcoming).

Sánchez, S. and Sicilia, M. A. (2004). On the semantics of aggregation and generalization in learning object contracts. In *Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies* - ICALT 2004. Joensuu, Finland (forthcoming).



Semantic Web for e-Government – a Research Agenda

Ralf Klischewski
Hamburg University, Informatics Department
Vogt-Kölln-Strasse 30, D-22527 Hamburg
klischewski@informatik.uni-hamburg.de

The domain of e-government is unique because of its enormous challenge to achieve interoperability given the manifold semantic differences of interpretation of e.g. law, regulations, citizen services, administrative processes, best-practices, and, not least, many different languages to be taken into account within and across regions, nations and continents. These semantic differences are related to a great variety of IT solutions (on local, regional, inter-/national level) which will have to be networked (despite any effort of standardisation). In consequence, some of the key obstacles for networked computer applications in administrative processes and services are those kinds of barriers in which the different meanings of the data objects and interfaces in focus cannot be automatically mediated. Technically, those problems refer to semantic mismatches or "interoperability clashes, caused by differences in the conceptual schemas of two applications attempting to cooperate" (Missikoff 2002, p. 33). Typical problems are the use of different names, structures or scales for the same kind of information, as well as information represented at different levels of granularity, refinement, or precision.

Semantic Web has started out with a document oriented approach; the basic idea was to annotate Web pages with semantic markup. But the informational resources on e-government websites can comprise a variety of electronic "things": fine-grained information elements (e.g. opening hours), multimedia items (pictures, maps, videos), files (for downloads, e.g. documents, forms, client applications), interfaces to transactions (processed by backend applications), links (to other sites and their resources), services (e.g. authentication), user related objects (e.g. user profiles) etc.

These problems are not of the same kind as the ones being addressed in the area of knowledge management in e-government (e.g. Wimmer 2003). The point of departure in Semantic Web applications is to enable machines to find their way around in a world of meaningful objects. The semantic markup and semantic links are to "allow machines to follow links and facilitate the integration of data from many different sources" (Berners-Lee & Miller 2002). On the contrary, knowledge management is concerned primarily to enable humans to find their way in the world of knowledge – with the help of machines. However, it is foreseeable that both areas will be closely interlinked (if not merged) within the next couple of years.

Semantic Web has started out with a document oriented approach; the basic idea was to annotate Web pages with semantic markup. But the informational resources on e-government websites can comprise a variety of electronic "things": fine-grained information elements (e.g. opening hours), multimedia items (pictures, maps, videos), files (for downloads, e.g. documents, forms, client applications), interfaces to transactions (processed by backend applications), links (to other sites and their resources), services (e.g. authentication), user related objects (e.g. user profiles) etc. All of these are meaningful for site users, all of these are searched for and being used – thus all of these are candidates for semantic markup. In the visions of Semantic Web, humans as Web users are accompanied by machine actors searching for "meaningful" resources and seeking to use/compose seamless services:

1. *end users*, such as individual citizens or employees of some organisation, using an internet browser
2. *electronic agents* on behalf of a third party (end user, organisation)
3. interoperability components of *e-government sites* (e.g. searching for additional resources on other sites to reply to a primary user or agent request)

All of these actors may search for and/or use (a) public resources (which are of the same kind for every potential requestor) or (b) "private" resources, i.e. parts of or the whole assembly of the resources are related to a specific user case (i.e. incorporating or drawing on specific information about the user, usually concealed from the public).
Setting up seamless e-government services requires information integration as well as process integration involving a variety of objects with specific semantics. How can research help to bring out the semantic interoperability as prerequisite for integrating the related technical and social aspects?

Initially, research must acknowledge that the roll-out and application of IT in public administration will continue to be based on individual organisational and technical strategies. Organisational units (or even individuals) decide on the application of standards (if available) and how existing and new meaningful "business" objects are represented by data structures and processed by computerised applications. There will always be individual implementation and markup strategies, and no standardisation will avoid semantic mismatches or interoperability clashes. This semantic variety and related interoperability problems and inefficiency will even increase with the complexity of applications and data structures. But if the individual semantics can be encapsulated in distinct objects and it is known how those objects are constructed – then there is hope for machine interpretation and for paving a common ground for true seamless services. To make this vision become reality, it is claimed that research should support two directions of development (see also Klischewski 2003): (1) the readiness of information providers and consumers for the Semantic Web, and (2) a common infrastructure for semantic interoperability.

Ad (1): applied research needs to address the following issues to make administrations successful in their Semantic Web projects, or at least to get them started:

- *Organisational cost/benefit:* it requires time and effort to professionally construct semantic-based systems and to create or contribute to those virtual knowledge spaces which are intended to be exploited through the future application – how can administrations evaluate and justify their investment? It needs "business cases" for administrations on all levels to show why this direction of information management is beneficial.
- *„User" involvement:* new semantic-based applications can only deliver enhanced value when the domain experts are actively involved in creating new services – but who are the experts? How can they be supported? What are the prerequisites that primary information providers and consumers (e.g. administrative staff or clients who are not technology or ontology experts) may work effectively with or on top of the new technologies? The critical issue is mainly the production of semantic expressions, therefore this demands understanding and reference models of editorial processes and tools for the domain experts.
- *Technical integration:* what is the centre of gravity for integrating Semantic Web technologies – front-end, middleware or back-end? How do new tools for ontology construction and knowledge representation relate to given IT environments? There is a need for shared or standardised ontologies and metadata suggesting basic elements, for relations and structures to be embodied in application specific ontology production as well as for middleware components and/or services to support local activities and interoperability (e.g. mapping of or translating between ontologies, resolution of semantic conflicts).
- *Implementation strategy:* what are the best practices and methods for implementing semantic based systems? What are the success factors? We surely need more empirical research and project evaluation to provide guidelines for administrations.

Ad (2): When developing a common infrastructure for semantic Web in e-government it needs a common semantic reference for the electronic exchange and interrelation of domain specific resources such as structured data, documents, Web information, e-services, Web services, etc. The meaning of these IT-based resources may be explored by evaluating the references to different models and meta data based on the respective languages, concepts, and domain knowledge. It is state-of-the-art to use more or less formalised ontologies for capturing the semantics (see e.g. Smith & Welty 2001 for the variety of approaches) and to use them for markup and other interoperability purposes. However, the aim is not to build a centralised ontology-driven information system, but to create a web of electronic objects and relations incorporating a semantics which may be "understood" by social actors (individuals, organisations) as well as machine agents. The key elements should be:

- e-government domain ontology as a common reference for all terms and relations relevant in administration (anchored in some common upper ontology and supported by general metadata standards)
- mappings to all local e-government domain ontologies
- actor and role model as a reference for management of e.g. identity, ownership and access rights
- additional concepts for markup (hopefully standardised somewhere else) to serve interoperability needs on the level of network operation (e.g. security/trust, ownership, service level) or which provide domain specific semantics for specialised application areas (e.g. markup of legal text)\
- interfaces to informational resources each with global identifier and markup based on some or all of the above globally accessible specifications

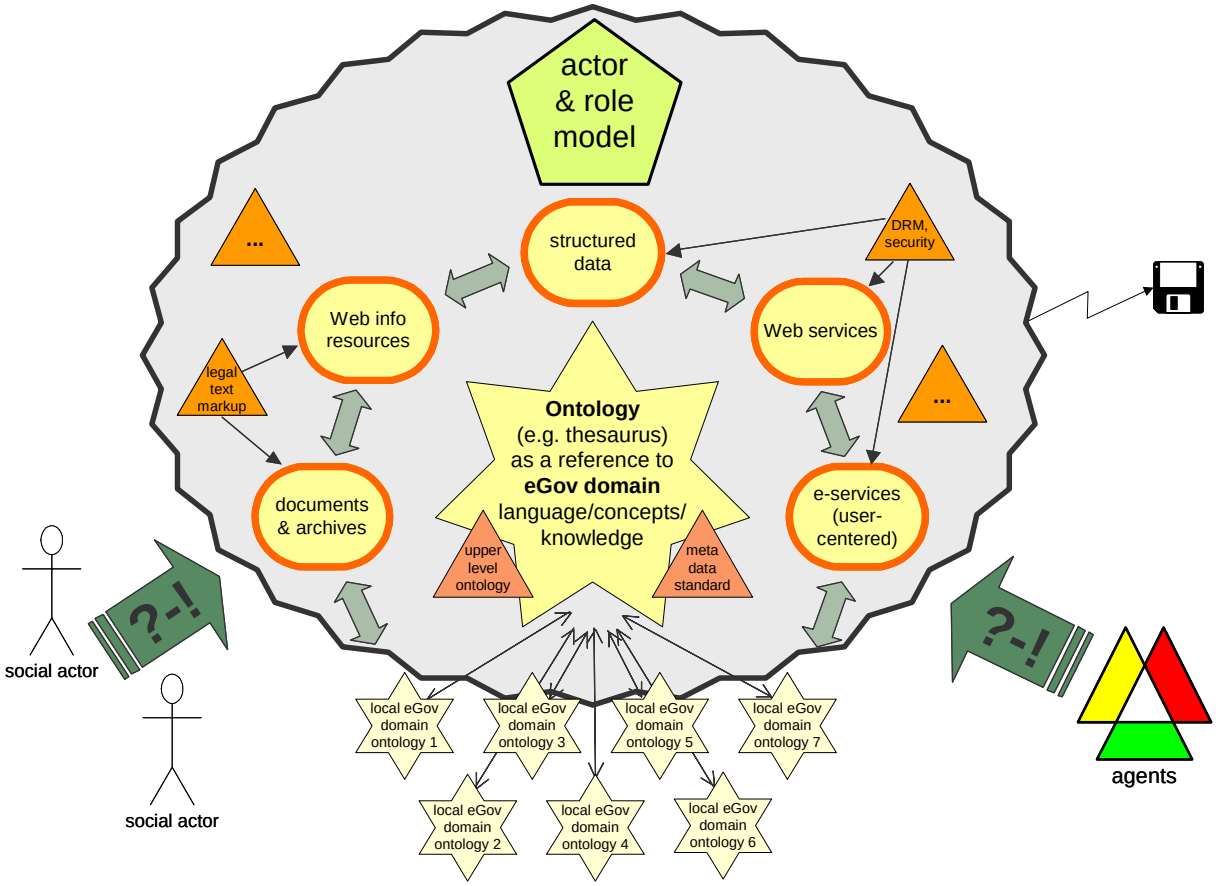


Figure 10. e-government domain ontology as a common reference to interrelate IT-based informational resources[*]

The readiness of information providers and consumers and the common infrastructure are both prerequisite for exploiting the potentials of Semantic Web, i.e. to enable Web-based applications which explicitly process semantic expressions. It will be a long time before e-government can fully exploit these potentials, but it is reasonable to start preparing for it now and making use of the new technologies. And research should do its best to support these efforts.

References

Berners-Lee, T., Miller, E. (2002) The Semantic Web lifts off. Special issue of ERCIM News, No. 51, October 2002, pp. 9-11

Klischewski, R. (2003) Semantic Web for e-Government. Proceedings EGOV 2003, Springer, Berlin, pp. 288-295

Missikoff, M. (2002) Harmonise – an ontology-based approach for semantic interoperability. ERCIM News, No. 51, October 2002, pp. 33-34

Smith, B., Welty, C. (2001) Ontology: Towards a new synthesis. In: Welty, C., Smith, B. (eds.): Formal Ontology in Information Systems. Ongunquit, Maine; ACM Press, pp. iii-x.

Wimmer, M. (ed.) (2003) Knowledge Management in e-Government. Proceedings KMGov 2003, Springer, Berlin

[*] This figure was presented at the research clustering workshop on "eGovernment Interoperability" by the European Commission (Brussels, March 1st, 2004).

Challenges for the Semantic Web and Information Systems from Culture

Kim H. Veltman
Scientific Director
Maastricht McLuhan Institute (MMI)
European Centre for Digital Culture, Knowledge Organization and Learning Technology
Maastricht, The Netherlands
E-mail k.veltman@mmi.unimaas.nl

At the turn of the 20th century, Semantics was a "science that studies the content (meaning) side of linguistic signs." It was linked with five other fields: 1) Semiotics a "science of general properties of sign systems"; 2) Lexicology, a "science that studies vocabulary of language also called Lexical Semantics"; 3) Lexicography, a "science of dictionaries and their creation"; 4) Semasiology, a "branch of semantics that seeks meaning departing from expression side of language that studies the development and changing meaning of words" and 5) Onomasiology, a "branch of semantics which departs from a meaning side of language and asks what expressions exist in other languages." [1]

"…During the 20th century there were many breakthroughs in computers and information systems. Paradoxically this quest to achieve automated forms of communication also led to a great simplification of semantic problems and a serious dumbing-down with respect to solutions. As a first challenge, this paper argues that we need to integrate what we left behind if we are to go ahead".

During the 20th century there were many breakthroughs in computers and information systems. Paradoxically this quest to achieve automated forms of communication also led to a great simplification of semantic problems and a serious dumbing-down with respect to solutions [2]. As a first challenge, this paper argues that we need to integrate what we left behind if we are to go ahead.

In the 1940s, Claude Shannon, who worked with Vannevar Bush on the ENIAC (Electronic Numerical Integrator and Computer), chose Boolean logic [3] as his point of departure. This reduced logic to choices between three logical operators: and, or and not. In terms of the semantic primitives [4], these operators dealt effectively with Exclusion, simplistically with Inclusion, and in a very limited way with Intersection. In terms of basic questions this focused on Who? and What? questions, ignoring entirely Where?, When?, How? and Why? questions.

In the 1950s, Curry and Feys developed typed combinatory logic [5]. This introduced a distinction between three levels of language: 1) Phenotype, which describes natural language as order of words; 2) Genotype, which expresses grammatical invariants and structures and 3) the Cognitive level which deals with lexical predicates as represented by semantic cognitive schemes. This linked cognitive science with semantics[6] and separated clearly perceptual and conceptual dimensions [7]. It also focused so much on mental dimensions that a recent survey of trends could ask: Does representation need Reality [8] ? Hence, a second challenge is to remember that our models need to explain the physical world and to bridge physical and mental worlds [9].

In the 1980s, the influential work of Langacker established the foundations of a new cognitive grammar [10]. This introduced a fundamental distinction between nominal predications used to designate things (i.e. nouns) and relational predications used for states (a-temporal relations) and processes (adjectives, adverbs, prepositions and verbs). In terms of basic questions, this established a distinction between Who? and What? (nouns) and relations entailing Where?, When?, How? and Why? (adjectives, adverbs, prepositions and verbs). This prompted new attention to relations using verbs: troponymy [11]. It also obscured very clear relations that had been established between division and partition. Since the 1980s, traditional distinctions have been even further eroded. For instance, Eduard Hovy [12], one of the authors of the influential, WordNet recently noted (2002):

We define an ontology rather loosely as a set of terms, associated with definitions in natural language (say English) and, if possible, using formal relations and constraints, about some domain of interest, used in their work by humans, data bases and computer programs. We view a set of semantic relations organized into collections and perhaps related in a generalization hierarchy as a special instance of an ontology [13].

These definitions are important because they establish new links between ontology and semantics that help to account for the meteoric rise of these two concepts in the past decade. Meanwhile, Christiane Felbaum (2002) claimed that: "If one examines the lexicalized concepts in relation to one another, it becomes clear that they differ in systematic ways that are characterizable in terms of similarities and contrasts. These consistent differentiations among concepts are what we call semantic relations"[14]. Interpreted narrowly this definition would mean that semantics has nothing to do with who or what a thing is, or what it means. Rather, semantic relations would be limited to differentiations in terms of similarities and contrasts [15].

To understand the larger challenges a brief excursus in history is necessary. In Antiquity the study of language was divided into the structure of language (grammar); the logic of language (dialectics, later logic) and the effects of language (rhetoric). This trivium of grammar, dialectic and rhetoric formed the language side of the seven liberal arts [16]. What? questions were shared by Grammar and Dialectic (Logic, figure 1). How? questions were the domain of rhetoric. Why? questions were relegated to philosophy. This discipline based approach meant that it was not until the 19th century that linguistic variation and dialect studies (Where? questions) and etymology (When? questions) came into focus. The 20th century precursors to the semantic web returned the focus to logic (What? questions). As a result, semantics as a study of the meaning of words, language and sign systems narrowed to become a study of relations and relationships between static words and terms. A third challenge which grows out of the second is to expand the scope of semantics to cover all the basic questions.

The focus on logic has had a more subtle consequence of assuming that the semantic web is/should be about a single logic, namely, the truth system with which we happen to use today. In the realms of science and business this is completely 'logical'. When making transactions, we need assurances that our customers use the same rules as we do.

Discipline	*Grammar,*	*Dialectic*	*Geography*	*History*	*Rhetoric*	*Philosophy*
Function	*Structure,*	*Logic*			*Effects*	*Causes*
Question	*Who?,*	*What?*	*Where?*	*When?*	*How?*	*Why?*

Figure 11. Basic questions and basic disciplines.

Inherent in this approach is the assumption that an entity is substantially one thing with fixed properties: e.g. the sun is a body around which planets rotate as the heliocentric theory claims. Prior to Copernicus people believed in a geocentric universe where the sun revolved around the earth. For culture, we need semantic models that cope with both systems: that allow us to trace how paradigm shifts and changes in world views change our understanding of entities. Hence a fifth, more elusive challenge is to integrate different world-views into our plans for the semantic web and information systems. We need to develop the plural meaning of information systems.

The official national view is often quite different from a regional or local view. The explanations of Indian New Delhi concerning Kashmir are typically different than the explanations of Pakistan's Karachi or Kashmir's explanations of itself. Their language, claims and even their maps may be different. This leads to a fourth challenge whereby our semantic webs and information systems need to become more multi-lingual [17] and multi-cultural.

"…The focus on logic has had a more subtle consequence of assuming that the semantic web is/should be about a single logic, namely, the truth system with which we happen to use today. In the realms of science and business this is completely 'logical'. When making transactions, we need assurances that our customers use the same rules as we do".

From a larger historical viewpoint there was a brief time in the 19th and early 20th centuries when simply exporting one's own political solutions abroad via colonialism and imperialism seemed a viable solution to differences. Today, however, when Europe is less than 5% of the world population, how can we legitimately assume that our version of politics is necessarily the only model when countries such as China, now the 8th largest economy in the world, have other systems? This is in no way to claim that we should abandon democracy, but simply to urge for multiple logics, belief systems and models.

These variations go far beyond questions of politics. They entail different world views, which Kuhn has called 'paradigm shifts' both within and across cultures. For instance, if we decided that today's heliocentric explanation was not only a valid theory but the sole admissible theory, then we would theoretically be unable to understand all proponents of a geocentric universe prior to Copernicus in the West and even more recently in some isolated regions of the world.

One of the important insights of the late Marshall McLuhan was that every medium for communicating knowledge also shapes and limits the knowledge that it communicates. McLuhan argued that new media typically used the previous medium as their content. Hence, the telegraph used print as its content, print used manuscripts as its content; manuscripts used oral communication as their content and so on. This was one of the reasons for McLuhan's claim that: "the medium is the message."

Digital communication is not simply another step in the tradition of a new medium adopting the content of its predecessors. Once materials are digital they can be translated back to other media without difficulty. Potentially, a digital test can be output as a printed text, as oral-audio text or, using stereo-lithography, even as cuneiform text [18]. Implicit in this insight is a sixth challenge. How can we create a more comprehensive semantic web with information systems that both reveal these new potentials and allow us to recognize explicitly the differences (both qua strengths and limitations) of earlier media?

The nexus of ontologies and the semantic web mentioned earlier led understandably to a focus on metadata. The scope of such metadata needs to be expanded into virtual reference rooms (which link classification systems, dictionaries, encyclopedias and titles)[19], that provide new access to distributed digital libraries and fora for collaborative research and creativity: a new virtual agora. These can evolve into a Distributed European Electronic Resource (DEER) [20] and lead to global efforts in the form of a World Distributed Electronic Resource (WONDER).

Notes

1. *See: http://www.uni-bonn.de/~dbuncic/14ung/semantik.pdf,*
2. *The ideas in this paper are discussed in more detail in the author's "Towards a Semantic Web for Culture," Journal for Digital Information, Summer 2004 (in press).*
3. *Boolean logic diagrams are themselves a simplification of Euler's diagrams. For Euler and Quantified Expressions, See: www.rci.rutgers.edu/.../Deduction/ EulerDiags.html*
4. *John Sowa, "Ontology, Metadata, and Semiotics": Presented at ICCS'2000 in Darmstadt, Germany, on August 14, 2000. Published in B. Ganter & G. W. Mineau, eds., Conceptual Structures: Logical, Linguistic, and Computational Issues, Lecture Notes in AI #1867, Springer-Verlag, Berlin, 2000, pp. 55-81.John F. Sowa, Ontology Metadata and Semiotics, See: http://users.bestweb.net/~sowa/peirce/ontometa.htm. For a more detailed understanding of his work see: John F. Sowa, Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks/Cole Publishing Co., Pacific Grove, CA, 2000. For a fuller bibliography cf. John F. Sowa. See: http://www.jfsowa.com/pubs/index.htm. As noted by Sowa, these semantic primitives go back to Peirce in his semiotics. See: http://users.bestweb.net/~sowa/peirce/ontometa.htm. Peirce claims that there are five semantic primitives: Existence, Coreference, Conjunction, Negation and Relation. As will become clear in this essay, this overlooks that the other semantic primitives are also formal relations.*
5. *H. Curry and R. Feys, Combinatory Logic, Amsterdam: North Holland, 1958.*
6. *For a thoughtful essay on this subject cf. Mihailo Antović, "The Position of Semantics within Contemporary Cognitive Science," UDC 81'37:165.19: in: Facta Universitatis, Series Linguistics and Literature, vol. 2, no. 10, pp. 415-424. See: http://facta.junis.ni.ac.yu/facta/lal/lal2003/lal2003-06.pdf, For linguistics and cognitive science See: http://www.press.uchicago.edu/Subjects/virtual_linguistics.html*
7. *S. Pribbenow, "Meronymic relationships", in Semantics as in note 8, p. 39.*
8. *Conference in Vienna May 14-16 1997. See: http://www.univie.ac.at/cognition/conf/ntcs97/*
9. *For an historical discussion see: "Visualization and Perspective. Visualizzazione e prospettiva" Leonardo e l'eta della ragione, eds. Enrico Bellone e Paolo Rossi, (Scientia, Milan, 1982), pp.185-210 (English), pp. 211-224 (Italian).*
10. *R. W. Langacker, Foundations of Cognitive Grammar, Stanford: Stanford University Press, 1987, (Volume 1: Theoretical Prerequisites) .*

11. *Cf. Christine Felbaum, "On the Semantics of Troponymy, in: The Semantics of Relationships, ed. Rebecca Green, Carol A. Bean, Sung Hyon Myaeng, Dordrecht: Kluwer Academic Publishers, 2002, pp. 23-34.*
12. *For Further publications See: http://www.isi.edu/natural-language/people/hovy/publications.html*
13. *Eduard Hovy "Comparing Sets of Semantic Relations in Ontologies," in: Semantics, as in note 8, 2002, p.92.*
14. *Christiane Felbaum, "as in note 8, p.23.*
15. *To be sure there are exceptions to this trend. Cf. Christophe Jouis, "Logic of Relationships," in: Semantics, as in note 8, 2002, especially pp. 129-130.*
16. *Along with the quadrivium of mathematics, arithmetic, music and astronomy.*
17. *Cf the important work of Accès Multilingue au Patrimoine (AMP). AMP newsletter under Ministère de la Culture, France. See: http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f_01.htm. On this subject see also: Marc van Campenhoudt, Abrégé de terminologie Multilingue, See: http://www.termisti.refer.org/ theoweb1.htm#intro*
18. *Cf. the author's "Electronic Media and Visual Knowledge," Knowledge Organisation, Wurzburg, Vol. 20 (1993), No. 1, pp. 47-54.*
19. *See: http://www.i-massweb.org/*
20. *Cf. the study by Francesca Monti and Suzanne Keene on the DEER under E-Culture Net. See: http://www.eculturenet.org/FP5/publicPDF/deliverable11a.pdf.*

Ongoing Research Column: Real World SW Cases

The Research in Progress or Ongoing Research Column is dedicated to the presentation of interesting research works with important achievements and critical milestones towards the realization of Information systems that prove the value potential of semantic web. In this issue we present two short articles describing RIP.

Ambjorn Naeve, head of the KMR Group, at the Royal Institute of Technology in Sweden provides an excellent approach for the design and implementation of a Public e-Learning Platform (PeLP) through an integration of SW technology, open source and open international ICT standards. In his article, the reader can find many answers in conceptual issues as well as in technical ones. I found excellent the following excerpt form his article: "*…The Semantic Web initiative, as it looks today, does not provide such semantics. It provides descriptions of web resources, but no way to present them to the user in a contextually clear way. There are initiatives, such as topic navigation and visual history browsers, that try to address this problem, but they fail miserably in giving the necessary overview of the conceptual context. In order to solve this problem, we are working on ideas to extend the semantic web in order to provide not only semantic information for the machine, but also conceptual information for the human user. This form of extended (human) Semantic Web, which we call the* ***Conceptual Web****, is a long-term vision with many parts".* Moreover Prof. Naeve outlines the work done in KMR for the last years which is summarized in the so-called ***Knowledge Manifold***: "*The KMR group has developed an information architecture called a Knowledge Manifold [3], which supports a number of different strategies for suppressing information. A Knowledge Manifold consists of a number of linked conceptual information landscapes (context-maps), whose concepts can be filled with content".*

William Grosky from the *Department of Computer and Information Scienc at University of Michigan-Dearborn* and D.V. Sreenath and Farshad Fotouhi from the *Department of Computer Science*, *Wayne State University*, *Detroit*, *Michigan* summarize their excellent work on their article entitled "*Deriving Emergent Semantics"*. From their full of ideas contribution (more info you can find in their relevant journal articles mentioned in the reference section" I was triggered by a phrase: "*We believe that the author of a web page contributes to the initial semantics of that page, but that the semantics of that page varies over a period of time based on the users who browse through the collection of web pages. The actual semantics of a web page is the emergent semantics that evolves over a period of time, which depends on the browsing paths of all the users who visit that page*". Their dynamic semantics approach challenges the semantic web considerations. We are looking forward to hear from them soon the answers of their concluded questions: "*…We still do not know the full impact of the semantic web. Can the results of this research be used to provide that crucial next higher semantic layer? Or can the knowledge gained from this research be used to derive better semantics from the semantic web? Only time will tell".*

Using Semantic Web technology to build a Public e-Learning Platform based on open source and open international ICT standards.

Ambjörn Naeve
Head of the Knowledge Management Research (KMR) group
Centre for user-oriented IT Design (CID)
Royal Institute of Technology (KTH)
Stockholm, Sweden

Introduction

The semantics[3] of the term metadata is ”data about data”. Through the emergence of the Semantic Web, the metadata has acquired the potential to become just as distributed as the data it describes - while still remaining just as "searchable and combinable" as if it resided inside a single database. This metadata decentralization process - which is enabled by the representational power of RDF[4] - is bound to have a profound impact on the design and use of ICT-enhanced environments in general. In particular, it will enable a transformation

- from teacher-centric to learner-centric e-learning.
- from doctor-centric to patient-centric e-health.
- from burocrat-centric to citizen-centric e-adminstration.
- from producer-centric to consumer-centric business models in general.

As stated in [8], ”Metadata will no longer be restricted to something 'objective' that has to be downloaded from some central server. On the contrary, metadata will be allowed to consist of subjective views of resources that are distributed and shared in contexts that can evolve dynamically”. In combination with peer-to-peer networks this will lead to ”opinion networks” that can exercise a selection pressure for quality improvement within all of the above mentioned fields.

The KMR group [10] at KTH has focused its involvement with Semantic Web technology around the design of interactive learning environments. We are presently working to create a *Public e-Learning Platform* based on open source and emerging international ICT standards [1]. This PeLP is based on a combination of Semantic Web techniques and peer-to-peer services for search, retrieval, publication, replication and mapping of metadata.”

The Conceptual Web

The stated goal of the Semantic Web[5] is to enable machine understanding of web resources. The rationale behind this goal has been that deriving meaning from contemporary HTML or other web resources is nearly impossible due to the lack of a common metadata framework for describing resources. In fact, most resource descriptions today are in the form of HTML text in a containing document. While such semantic descriptions are meaningful only to the human reader, the semantic web will provide such descriptions in machine readable format.

However, it is not at all evident that such machine-readable semantic information will be clear and effective for human interpretation. The hyper-linked structure of the web presents the user with a totally fluid and dynamic relationship between context and content, which makes it hard to get an overview of the conceptual context within which the information is presented. As soon as you click on a hyperlink, you are transferred, helplessly, to a new and often unfamiliar context. This results in the all too well-known “surfing-sickness” on

[3] i.e. meaning.
[4] Resource Description Framework.
[5] See e.g. www.w3c/RDF

the web, that could be summarized as "Within what context am I viewing this, and how did I get here?" The conclusion we draw is that *extracting usable meaning from web pages is often as difficult for a human reader as it is for a machine*. This strongly suggests that there is a need for a human-understandable semantics for web resources as well.

This form of semantics becomes even more important within the emerging field of e-learning. In a learning context, the conceptual structure of the content is an essential part of the learning material. Losing the contextual information of the content means more than just "surfing-sickness". It means that you will not be able to contextually integrate the concepts that you are trying to learn, which is vitally important in order to achieve an understanding of any specific subject area.

The Semantic Web initiative, as it looks today, does not provide such semantics. It provides descriptions of web resources, but no way to present them to the user in a contextually clear way. There are initiatives, such as topic navigation and visual history browsers, that try to address this problem, but they fail miserably in giving the necessary overview of the conceptual context.

In order to solve this problem, we are working on ideas to extend the semantic web in order to provide not only semantic information for the machine, but also conceptual information for the human user. This form of extended (human) Semantic Web, which we call the *Conceptual Web*, is a long-term vision with many parts [5]:

The fundamental building block of the conceptual web is *conceptual modeling*, which provides a human-understandable semantics for both abstract ideas and concrete resources. We use the *Unified Modeling Language*[6] for conceptual modeling, which provides a well-proven and standardized vocabulary for this purpose.

The Conceptual Web as a Knowledge Manifold

The KMR group has developed an information architecture called a *Knowledge Manifold* [3], which supports a number of different strategies for suppressing information. A Knowledge Manifold consists of a number of linked conceptual information landscapes (context-maps), whose concepts can be filled with content. The Conceptual Web equipped with the Knowledge Manifold architecture is outlined in [1].

The traditional educational systems are based on teacher-centric, curricular-oriented knowledge push. When used for learning, the KM architecture enables the design of interactive learning environments that are based on learner-centric, interest-oriented knowledge pull, and which can be designed to support inquiry-based and customizable forms of networked learning. This is the reason why we have chosen the KM architecture for our PeLP.

Fundamental pedagogical assumptions

The KM architecture is based on the following fundamental principles 3:

- Nobody can teach you anything. A good teacher can inspire you to learn.
- Your motivation to learn is based on the experience of subject excitement and faith in your learning capacity from live teachers.
- Your learning is enhanced by taking control of your own learning process.
- No 'problematic' questions can be answered in an automated way. In fact, it is precisely when your questions break the pre-programmed structure that the deeper part of your learning process begins.

The seven Knowledge Roles

When used for learning purposes, the KM architecture supports the following seven different knowledge roles 3:

[6] www.uml.org

• the *knowledge cartographer*, who constructs and maintains context-maps.
• the *knowledge librarian*, who fills context maps with content-components.
• the *knowledge composer*, who constructs customized learning modules.
• the *knowledge coach*, who cultivates questions.
• the *knowledge preacher*, who provides live answers.
• the *knowledge plumber*, who directs questions to appropriate preachers.
• the *knowledge mentor*, who is a role model and supports self-reflection.

It is fundamentally important that all these roles should be available to both teachers and learners. "You learn as long as you are teaching" is the pedagogical principle at work here.

Infrastructure, frameworks and tools of the PeLP

The Public e-Learning Platform under construction by the KMR group makes use of the Knowledge Manifold architecture. Moreover, it consists of

• the *Edutella*[7] *infrastructure*: A democratic (peer-to-peer) network infrastructure for search and retrieval of information about learning resources on the Semantic Web [6], [7].

• the *SCAM framework*: (Standardized Contextualized Access to Metadata) A framework that helps applications to store and share information about learning resources [9].

• the *SHAME framework*: (Standardized Hyper-Adaptable Metadata Editor) An editor framework that supports an evolving annotation process of learning resources in a way that enables the growth of an "ecosystem" of quality metadata [1].

• the *Formulator SHAMEditorEditor*, a tool for editing metadata editors that is built on top of the SHAME framework [1].

• the *Conzilla concept browser*: A knowledge management tool that supports the construction, navigation, annotation and presentation of the information in a knowledge manifold [4].

• the *EduFolio networked portfolios*: An e-portfolio system built on top of SCAM, SHAME and Edutella, that supports collaborative and reflective learning techniques [1].

• the *VWE composer*: (Virtual Workspace Environment) An environment for composing learning resources and building customized learning modules [1].

Testing and ongoing work

The KMR group is presently coordinating a collaborative effort that involves the *Swedish National Agency for School Improvement* (MSU), the *Swedish Educational Broadcasting Company* (UR), and the *Swedish National Centre for Flexible Learning* (CFL). These three major Swedish public service e-learning players have teamed up and are now jointly contributing to the PeLP. The *Soft infrastructure for IT in education* project of MSU [13], the D*igital Media Library* of UR [14], and the *Learning Resource Centre* of CFL [15] are three of the important stakeholders projects in the PeLP.

Some of the tools of the PeLP (notably Conzilla and EduFolio) are also being used by the students of the media technology program at KTH [12] in order to enhance the capabilitites of the students to document and reflect on their learning process and to share this material with others in a learning community of practice. This

[7] developed under the coordination of Wolfgang Nejdl at the Learning Lab Lower Saxony.

project is still in its early stages, but in spite of some initial technical problems and obvious needs for improvements, the students experiences with these types of tools seem to be mainly positive [1].

References

1. Blomqvist, U., Handberg, L., Naeve, A., *New Methods for focusing on Students Learning Process and Reflection in Higher Education*. Proc. of the 28th IUT (Improving University Teaching) Conference. Växjö, Sweden: May 2003.
2. Naeve, A. & Nilsson, M. & Palmér, M. & Paulsson, F., *Contributions to a Public e-Learning Platform – Infrastructure, Architecture, Frameworks and Tools*, to be published in the International Journal of Learning Technology, October 2004.
3. Naeve, A., *The Knowledge Manifold - an educational architecture that Supports Inquiry-Based Customizable Forms of E-learning*, Proceedings of the 2nd European Web-based Learning Environments Conference (WBLE 2001), Lund, Sweden, Oct. 24-26, 2001, http://kmr.nada.kth.se/papers/KnowledgeManifolds/KnowledgeManifold.pdf.
4. Naeve, A., *The Concept Browser - a New Form of Knowledge Management Tool*, Proceedings of the 2nd European Web-Based Learning Environment Conference (WBLE 2001), Lund, Sweden, Oct. 24-26, 2001, http://kmr.nada.kth.se/papers/ConceptualBrowsing/ConceptBrowser.pdf.
5. Naeve, A. & Nilsson, M. & Palmér, M., *The Conceptual Web - our research vision*, Proceedings of the first Semantic Web Working Symposium, Stanford, Ju1y 2001, www.semanticweb.org/SWWS/program/position/soi-nilsson.pdf.
6. Nejdl. W., Wolf, B., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmér, M., Risch, T., *Edutella: A P2P Networking Infrastructure Based on RDF*, Proc. of the 11th World Wide Web Conference (WWW2002), Hawaii, May 7-11, 2002, (Modified version published in special issue of Computer Networks, jan 2003), http://kmr.nada.kth.se/papers/SemanticWeb/p597-nejdl.pdf.
7. Nilsson, M. & Palmér, M. & Naeve, A.,*The Edutella P2P Network- Supporting Democratic E-learning and Communities of Practice*, in McGreal, R. (ed.) Flexible education using learning objects, to be published in 2004.. http://kmr.nada.kth.se/papers/SemanticWeb/Edutella-chapter.pdf
8. Nilsson, M. & Palmér, M. & Naeve, A., *Semantic Web Meta-data for e-Learning - Some Architectural Guidelines*, Proceedings of the 11th World Wide Web Conference, Hawaii, May 7-11, 2002. http://kmr.nada.kth.se/papers/SemanticWeb/p744-nilsson.pdf.
9. Palmér, M. & Naeve, A. & Paulsson, F., *The SCAM-framework – helping applications to store and access metadata on the semantic web*, to be published in the proceedings of the First European Semantic Web Symposium, Heraklion, Greece, May 10-12, 2004.

Relevant web sites

10. The Knowledge Management Research group: http://kmr.nada.kth.se
11. Centre for user-oriented IT Design (CID): http://cid.nada.kth.se
12. Royal Institute of Technology (KTH): http://www.kth.se
13. Soft Infrastructure for IT in education (MSU): http://mjukis.skolutveckling.se
14. Digital Media Library (UR): http://www.ur.se/mb
15. Learning Resource Centre (CFL): http://www.cfl.se



Deriving Emergent Semantics

W.I. Grosky*, D.V. Sreenath**, and F. Fotouhi
*Department of Computer and Information Science, University of Michigan-Dearborn, Dearborn, Michigan 48128
**Department of Computer Science, Wayne State University, Detroit, Michigan 48202
wgrosky@umich.edu, sdv@cs.wayne.edu

It is well known that interpretation depends on context, whether for a work of art, a piece of literature, or a natural language utterance. Our research is addressing the dynamic context of a collection of linked multimedia documents, of which the web is a perfect example. While static content-based, author-defined, document semantics has been the standard approach to derive the semantics of web pages, we believe that dynamic contextual document semantics, which emerge through identification of various users' browsing paths though this multimedia collection, also play a very important role. We are developing techniques that use multimedia information as part of this determination. Some implications of our approach are that the author of a web page cannot completely define that document's semantics and that semantics emerge through use.

Our work has been influenced by [SGJ01]. In this paper, it is argued that images only have particular semantics when placed in the context of other images, or through user interaction. Image semantics are context sensitive. From the image database perspective, the semantics of an image can be extracted by interpreting the sequence of queries posed by the user. In the image database system of [SGJ01], semantics is not an intrinsic property captured during the image filtering process, but an emergent property of the interaction of the user and the database.

One of the goals of this research is to implement the reciprocal of a search engine: given a sequence of documents comprising a user's browsing path, generate a query that summarizes what the user was looking for. A specific application of this research can lead to terrorist trend detection.

Privacy laws ensure that a user's browsing history is not sold to advertising and marketing agencies. This restricts the availability of such browsing history for research purposes like ours. Most of the earlier work on relevance-feedback based approaches to gathering and analyzing users' preferences have failed primarily due to this lack of user confidence in the agencies that collect such information and the ways in which that information will be used. Secondly, users do not trust any application that is downloaded and installed on their system that captures and profiles their browsing patterns. Thirdly, most agent-based feedback systems do not focus on capturing the underlying purpose for which the user is browsing through collections of documents.

We, on the other hand, have developed and tested several approaches to derive the emergent semantics of web documents using user browsing paths without any explicit feedback. The information used for our analysis is the list of URLs visited by users, which typically can be obtained from an Internet service provider or from a reasonably large organization like a university or a corporation with diverse users. We have tested several approaches to capture, filter and analyze the users' browsing paths from such a large organization.

We believe that the author of a web page contributes to the initial semantics of that page, but that the semantics of that page varies over a period of time based on the users who browse through the collection of web pages. The actual semantics of a web page is the emergent semantics that evolves over a period of time,

which depends on the browsing paths of all the users who visit that page. This dynamic semantics approach is different from earlier works on deriving static semantics based solely on link analysis. Static analysis only captures the intended semantics of the authors of the linked web pages. Our dynamic analysis derives the semantics of web pages by deriving the semantics of user browsing paths. This analysis also helps capture the semantic profile of the user.

We have used a vector-space-based approach to represent browsing paths. Each path is represented by a vector, which captures both textual and visual keywords from pages occurring along the path. We compare various approaches, each based on latent semantic indexing [DDF90], to find what we call *semantic breakpoints* [GSF02]. These are used to decompose our original path set into a set of subpaths, which are semantically coherent. Using these semantically coherent browsing paths will then enable us to track the emerging semantics of individual web pages, as well as to characterize the browsing behavior of individual users. Each subpath, represented as a vector in the term-path matrix, can be visualized as a point in a reduced dimensional space. The semantics of a web page, w, can then be defined as the subset of the points in the reduced dimension space corresponding to the sub-paths that are within a threshold distance from page w. The semantics of a user browsing path is then the collection of concepts represented by the semantics of the pages traversed by the user.

In order to improve the effectiveness of our approach and to derive semantics at a higher level, we have been experimenting with WordNet® [Fel98]. WordNet® is an online lexical reference system where English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept. In our attempt to derive the emergent semantics beyond that resulting from the latent semantics obtained from the singular value decomposition, we are experimenting with the synonym sets and the hierarchical (hypernyms) grouping from the WordNet database. For example, the hypernym hierarchy for *dog* is a tree with its parent as *canine,* and its parents' parents as *carnivore,* etc., all the way to the root *entity*. We are currently studying the combination of semantics from distinct subpaths in order to derive higher-level semantics. We would then be able to combine the semantics such as interest in chemicals, interest in locations of nuclear power plants and driving directions into a high-level concept such as *terrorist alert*.

There are still many challenges to this research and some unanswered questions. We still do not know the full impact of the semantic web. Can the results of this research be used to provide that crucial next higher semantic layer? Or can the knowledge gained from this research be used to derive better semantics from the semantic web? Only time will tell.

REFERENCES

[DDF90] S. Deerwester, S. T. Dumais, and G. W. Furnas, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, Vol. 41 (1990), pp. 391-407.

[Fel98] C. Fellbaum, *WordNet: An Electronic Lexical Database,* MIT Press, Cambridge, MA, 1998.

[GSF02] W. I. Grosky, D. V. Sreenath, and F. Fotouhi, "Emergent Semantics and the Multimedia Semantic Web," *SIGMOD Record*, Vol. 31 (2002), pp. 54-58.

[SGJ01] S. Santini, A. Gupta, and R. Jain, "Emergent Semantics Through Interaction in Image Databases," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 13 (2001), pp. 337-351.

Semantic Web Calendar Column

Forthcoming Special Issues in International Journals

Educational Technology & Society, ISSN 1436-4522
Published by International Forum of Educational Technology & Society
Endorsed by IEEE Learning Technology Task Force

Special Issue (October 2004)

Ontologies and the Semantic Web for E-learning

The *Semantic Web* is the emerging landscape of new web technologies aiming at web-based information and services that would be understandable and reusable by both humans and machines. We argue that *Ontologies*, generally defined as a representation of a shared conceptualisation of a particular domain, is a major component of the Semantic Web. It is anticipated that Ontologies and Semantic Web technologies will influence the next generation of e-learning systems and applications. To this end, key developments such as

- *Formal taxonomies* expressed, e.g., with the help of the web ontology languages RDFS and OWL, and
- *Rules* expressed, e.g., with the help of the web rule language RuleML,

are expected to play a key role in enabling the representation and the dynamic construction of shared and re-usable learning content.

The aim of this special issue is to explore topics related with the new opportunities for e-learning created by the advent of Ontologies and the Semantic Web. We aim at a balanced composition of conceptual, technological and system evaluation work and invite submissions dealing with the following topics:

- Ontologies for e-learning systems
- RDFS/OWL-based educational metadata languages and technologies
- Architectures for ontology-based e-learning systems
- Rules and formal logic for e-learning systems
- Semantic web services for e-learning systems
- Supporting personalized and adaptive e-learning with Semantic Web technologies
- Supporting flexible e-learning systems with Semantic Web technologies
- Innovative Case Studies

Special issue guest editors

Demetrios G Sampson, Dept of Technology in Education and Digital Systems, University of Peiraias, and Informatics and TelematicsInstitute(ITI), Center for Research and Technology - Hellas (CERTH)
Email: sampson@iti.gr, URL: http://www.iti.gr/db.php/en/people/Demetrios_Sampson.html

Paloma Diaz, Computer Science Department, Universidad Carlos III de Madrid, Spain
Email: pdp@inf.uc3m.es or pdp@dei.inf.uc3m.es
URL: http://www.dei.inf.uc3m.es/english/members/pdp.html

Miltiadis D. Lytras, ELTRUN, the E-Business Center, Department of Management Science & Technology, Athens University of Economics and Business, Athens, Greece
Email: mdl@aueb.gr, URL: http://www.eltrun..gr

Gerd Wagner, Faculty of Technology Management, EindhovenUniversity of Technology
Email: G.Wagner@tm.tue.nl, URL: http://tmitwww.tm.tue.nl/staff/gwagner/

Semantic e-Business

Americas Conference on Information Systems (AMCIS)
August 5th – 8th, 2004, New York, USA

The emergence of collaborative processes as an effective means for organizations to deliver their value propositions to their customers, and ultimately to consumers, places an increased onus on organizations to develop systems incorporating emergent technologies. These systems should support the seamless availability of information and knowledge, content and know-how, among partners in the organizations' value chains. Rapidly increasing volume of available information and growing competition in the digital economy are forcing organizations to find efficient ways to gain valuable information and knowledge to improve the efficiency and effectiveness of their business processes.

The realization of representing these knowledge-rich processes is possible through the broad developments in the 'Semantic Web' initiative of the World Wide Web Consortium. But significant amount of research is needed to understand how conceptualizations that comprise business processes can be captured, represented, shared and processed by both human and intelligent agent-based information systems to create transparency in service and supply chains. The developments in on-demand content and business logic availability through technologies such as web-services offer the potential to allow organizations to create content-based and logic or intelligence driven information value chains enabling the needed information transparencies for semantic e-business processes.

Developments on these dimensions are critical to the design of knowledge-based and intelligence driven processes in the digital economy. Research is needed in the development of business models that can take advantage of emergent technologies to support collaborative, knowledge-rich processes in the digital economy. Equally important is the adaptation and assimilation of emergent technologies to enable business processes that contribute to organizations' value propositions. This mini track invites original research contributions that investigate the development of innovative business models to support knowledge-rich business models that enhance collaborations in the digital economy.

Mini Track Chair

Dr. Lakshmi S. Iyer
Information Systems and Operations Management (ISOM) Department
Bryan School of Business and Economics
The University of North Carolina at Greensboro.
Email: lsiyer@uncg.edu
Office Telephone: (336) 334-4984

Co-Chairs

Dr. Rahul Singh & Dr. A. F. Salam
Information Systems and Operations Management (ISOM) Department
Bryan School of Business and Economics
The University of North Carolina at Greensboro, USA
Email: rahul@uncg.edu , amsalam@uncg.edu

MAY 2004

First European Semantic Web Symposium
(ESWS 2004), 10-12 May 2004,
Heraklion, Greece
Conference Web Site: http://www.esws2004.org/

Organized by
Academic track: **Dieter Fensel** (DERI, Austria and Ireland), Rudi Studer (University of Karlsruhe,
Tutorial programme: **John Davies** (BT, UK)
Industrial track & Demo chair: **Christoph Bussler** (DERI, Ireland)
Local arrangements: **Jos de Bruijn** (DERI, Austria), **Martin Doerr** (ICS-FORTH, Greece)

In the European context, the 6th framework programme has demonstrated the EU's commitment to this technology area and a number of exciting new projects in the Semantic Web area have been launched, aggregated in the SDK project cluster - see http://www.sdk-cluster.org/ for further details. A tutorial programme, based around these projects, offers the opportunity to get up to speed with European and global developments in this exciting new area. All 6th framework projects, as well as other leading projects, in the Semantic Web area will be represented with posters. The conference will be co-located with the OntoWeb thematic network seminar and the Knowledge Web network of excellence meeting – see http://www.ontoweb.org/ and http://knowledgeweb.semanticweb.org/ for further details of OntoWeb and Knowledge Web events, respectively.

JULY 2004

First International Workshop on Semantic Web Services and Web Process Composition
In conjunction with the 2004 IEEE International Conference on Web Services (ICWS'2004)
(SWSWPC 2004), July 6-9, 2004, San Diego, California, USA
Web Site: http://dme2.uma.pt/~jcardoso/ICWS-SWSWPC04/SWSWPC_Workshop.htm

Workshop Organization
Jorge Cardoso, University of Madeira, Portugal, jcardoso@uma.pt
Amit Sheth, University of Georgia, USA, amit@cs.uga.edu
Leonid A. Kalinichenko, Russian Academy of Sciences, Russia, leonidk@synth.ipi.ac.ru
Francisco Curbera, IBM, USA, curbera@us.ibm.com

Second European Summer School on Ontological Engineering and the Semantic Web
(SSSW-2004), 19th-24th July 2004 - Cercedilla (Spain)
Web Site: http://babage.dia.fi.upm.es/summerschool/

The KnowledgeWeb Network of Excellence (http://knowledgeweb.semanticweb.org/) is pleased to announce the second European Summer School on Ontological Engineering and the Semantic Web. This summer school, presented by leading researchers in the field, is the successor to the well-received school held in 2003 as part of the activities of the OntoWeb Network Consortium (http:// www.ontoweb.org). The school represents an opportunity for postgraduate students to equip themselves with the range of theoretical and practical skills necessary for full engagement with the challenges involved in developing Ontologies and Semantic Web applications. Given the high degree of interest in this new field, candidates will need to book early to avoid disappointment. The school will be limited to 50 participants.

NOVEMBER 2004

Third International Semantic Web Conference
(ISWC2004), 7-11 November 2004,
Hiroshima Prince Hotel, Hiroshima, Japan
Conference Web Site: http://iswc2004.semanticweb.org
Organized by
The Japanese Society for Artificial Intelligence and The Semantic Web Science Association
Conference Chair: Frank van Harmelen, Vrije Universiteit Amsterdam, Frank.van.Harmelen@cs.vu.nl
Program Chairs
Sheila McIlraith, Department of Computer Science, University of Toronto, sheila@cs.toronto.edu
Dimitris Plexousakis, Department of Computer Science, University of Crete, And Institute of Computer Science, Foundation for Research and Technology (FORTH), dp@ics.forth.gr

Local Chair
Riichiro Mizoguchi
The Institute of Scientific and Industrial Research, Osaka University, miz@ei.sanken.osaka-u.ac.jp
Important Dates
Submission deadline for Research Track and Industrial Track extended by 2 weeks to Friday April 30, 11:59pm (Hawaii Time). The deadline for submission of abstracts is also Friday April 30, 11:59pm (Hawaii Time).

Semantic Web Selected Papers Available On-line

OntoWeb portal selection of publications:
http://ontoweb.aifb.uni-karlsruhe.de/Publications/

The website for the Semantic Web:
http://www.semanticweb.org/

Key People…
Professor James A. Hendler
http://www.cs.umd.edu/~hendler/

Professor Dr. Rudi Studer
http://www.aifb.uni-karlsruhe.de/Personen/viewPersonenglish?id_db=57

SIG SEMIS Web Site Launch

By Miltiadis D. Lytras and Nikolaos Korfiatis

→ WWW.SIGSEMIS.ORG…

Introduction

As a part of our proposal to AIS Council the Semantic Web and Information Systems Special Interest Group (SIG SEMIS) has already formed and developed a portal site (http://www.sigsemis.org) that will be used as the electronic mean of communication with the Semantic Web and Information Systems Community.

Portal Structure

Sigsemis.org website has as the main aim to disseminate information effectively in order to semis community. Furthermore we aim to create a collaborative authoring environment and publication platform that will allow everyone to post his ideas and work regarding the SEMIS area. The portal has a hierarchical structure, which is as follows:

- **Sigsemis:** Contains Information about the SIG's research positioning, Services Offered and SIG Organization. As far as the activity of the SIG Grows there will be also placements for the special committees that will be formed under the SIG (Advisory Board, Awards Committee etc.). This Section contains also the SIG Board along with the CV's of the Board Members and their SEMIS Related publications.
- **Articles:** This section is the key area of the portal as it contains articles related to the semis area. Mainly these articles come from contributions to SIGSEMIS Newsletter (SEMIS News) and in the future will be short versions of published research work in the IJSWIS Journal (International Journal on Semantic Web and Information Systems) which will be published in print by the IDEA Group Publishing. Along with the presentation of the article there is a 'comment' future that allows any member of the portal to add a comment or post a follow-up to an existing one.
- **Newsletter:** This Section Contains archive of the "Semis News" Newsletter published by SIGSEMIS and officially sponsored by the Association of Information Systems. We target an open content approach so the newsletter will be in html form along with a commenting and rating system for each one of the article/columns published. There will be also a batch version of the newsletter available for download in pdf form.
- **IJSWIS:** This is planned to be the official homepage of the ISJWIS Journal published by the IDEA Group. It will host call for papers, an author registry and the CV's of the Editorial Board. There will also be a table of contents version of each volume with links to the short version-abstract (if exists) in the articles section. Our current development efforts contain the implementation of the workflow engine that will support IJSWIS processes such us manuscript submission, peer reviewing, submission notifications etc.
- **PhD Students Corner:** The range of research work that is currently been done in research centres across the world in the SEMIS area forced us to create a special section in our portal in order to provide a reference area the will be used for exchange and dissemination of research work. This area will provide information such us *researchers registry*, PhD *Abstracts and Proposals* as along us with PhD outcomes.
- **Library:** This section aims to host a reference service for the current and future work that is published and related to the SEMIS Field. It will certainly contain references from the IJSWIS and Related Journals, Publications of the SIGBoard Members etc. It will follow the Bibtex reference format with a friendly front-end for referencing.

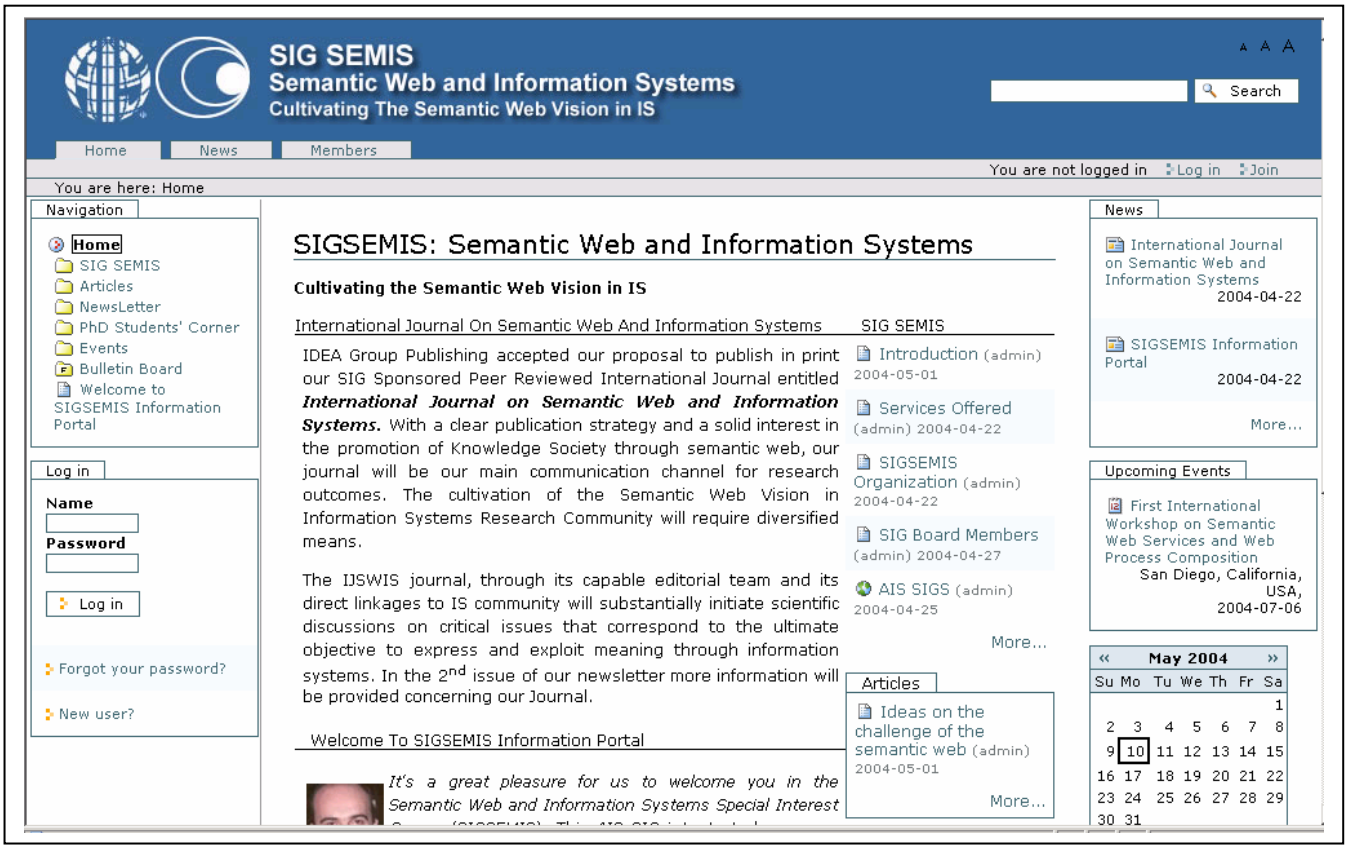
Our current development efforts support the following services

- Content Syndication through RDF/RSS chanels

- Open Publication approach. This service level gives the ability to each portal member to create his personal space and publish news.

Community Building and Knowledge Transfer

SIGSEMIS is an open forum and the incoming and outgoing stream of information is very important in order to make and maintain an active community in all levels of knowledge transfer and dissemination.



More info in the forthcoming issue of Bulletin.

Semantic Web Research Community: A column dedicated to presentation of Research Groups Worldwide

By Gerd Wagner and Lina Zhou

Key Research Centers

- The World Wide Web Consortium (W3C), http://www.w3.org/

  The World Wide Web Consortium (W3C) develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential as a forum for information, commerce, communication, and collective understanding.

- W3C SemanticWeb.org, http://www.w3.org/2001/sw/

  The ***Semantic Web*** provides a common framework that allows ***data*** to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (***RDF***), which integrates a variety of applications using XML for syntax and URIs for naming.

- Web-Ontology (WebOnt) Working Group, http://www.w3.org/2001/sw/WebOnt/

  The OWL Web Ontology Language is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. OWL has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full.

- Transatlantic Research Center for the Semantic Web and XML Technologies, http://www.semanticwebcenter.org.uk/

  The Centre provides leading European and American researchers and developers in the area of XML Technologies and the Semantic Web with unique opportunities for effective and flexible transatlantic collaboration aimed at achieving world-class results.

  The Center conducts research into a wide range of emerging leading-edge technologies. Specific research topics are defined in a particular Research Project. Each Project is being curried out by a Research Group, specially formed for this purpose. Every Research Project is aimed at publishing a world-class research monograph or research-based dictionary in order to make the major results of the project available to the world's scientific community.

- Competence Center Semantic Web (CCSW) at DFKI, http://ccsw.dfki.de/

  This site is part of the German research center for artificial intelligence. The focus of the center is on distributed information management with Web-based standardized object representations, ontologies, and rule systems.

- The Information Management Group at University of Manchester, UK, http://img.cs.man.ac.uk/cgi-bin/index.pl?groupsGo=groupsShow&group=semweb&groupsType=Project&strReturn

  The group concerns with Ontologies Knowledge Representation Hypermedia. It uses knowledge representation language to represent conceptual models in machine-amenable formats, while allowing agents to reason and compute over those models. The group is linked to projects such as OilEd, OntoWeb, WonderWeb, and so on.

- The Knowledge Management Group at **University of Karlsruhe, Institute AIFB**, Karlsruhe, Germany, http://www.aifb.uni-karlsruhe.de/WBS/

The group has a strong focus on Semantic Web and related areas. Core Semantic Web infrastructure technologies such as Ontobroker, OntoEdit and KAON are developed in collaboration with other groups in Karlsruhe. The group is involved into projects such as SEKT, Knowledge Web, AceMedia, OntoWeb, WonderWeb, SWAP and so on.

- The Knowledge Management Group (WIM) at the **Research Center for Information Technologies (FZI)**, Karlsruhe, Germany, http://www.fzi.de/wim/eng/

  The research group develops techniques and applications for the acquisition, representation & modeling, extraction, storage, access and application of knowledge. A wide range of knowledge intensive systems are based on different core techniques. The group is involved in projects such as DIP, SWWS, KAON, and so on.

- On-To-Knowledge, http://www.ontoknowledge.org/

  On-To-Knowledge-Project aims to develop tools and methods for supporting knowledge management relying on sharable and reusable knowledge ontologies. The technical backbone of On-To-Knowledge is the use of **ontologies** for the various tasks of information integration and mediation.

- **Knowledge Systems Laboratory** at Stanford University, http://www.ksl.stanford.edu/projects/DAML/

  They are developing semantic markup and agent-based technologies to help realize the vision of semantic web. DAML-Enabled Web Services Project had the goal of developing next generation semantic web tools and technology.

- The MINDSWAP Group at the University of Maryland, http://www.mindswap.org/

  It is Maryland Information and Network Dynamics Lab Semantic Web Agents Project. Simple HTML Ontology Extensions (SHOE) is one of its first research projects on Semantic Web. It is also involved with trust and security on the Semantic Web and automatic ontology mapping.

- **eBiquity Research Group at University of Maryland, Baltimore County, USA,** http://ebiquity.umbc.edu/v2.1/research/area/id/9/

  The group has been involved with a variety of projects related to the Semantic Web. Among others, Spire, a Personal application for the Semantic Web, explores the use of semantic web technologies in support science in general and the field of ecoinformatics in particular. Securing the Semantic Web investigates distributed trust management as an alternative to traditional authentication and access control schemes in dynamic and pen computing environments such as multiagent systems, web services and pervasive computing. Semantic Discovery focuses on the design, prototyping, and evaluation of a system, called SEMDIS that supports indexing and querying of complex semantic relationships and is driven by notions of information trust and provenance.

- **OntoWeb** , http://ontoweb.aifb.uni-karlsruhe.de/

  Ontoweb is a thematic network funded by the European commission. Its goal is to bring together activities in the area of ontology-based methods and tools for the Semantic Web, bypassing communication bottlenecks between the various and heterogeneous groups of interest.

- Large Scale Distributed Information Systems Lab (LSDIS) at the University of Georgia, http://lsdis.cs.uga.edu/

  The LSDIS lab has extensive research, training, and technology transfer program in the areas of Semantic (Web) technologies. The SemDis project focuses on knowledge discovery and semantic analytics, and have developed a very large populated ontology testbed SWETO for evaluating (million object and relationship) that is being made available for all non-commercial usage. The METEOR-S project on Semantic Web Processes has researched and is developing tools/systems that utilize semantics in complete Web Service and Web Process lifecycle (annotation, discovery, composition, orchestration/execution). The Bioinformatics for Glycan Expression is applying semantic techniques for integration, analysis and discovery activities in the area of Glycomics, and has developed GLYCO, a comprehensive ontology covering some of the significant areas in the field. Example of

commercialization of LSDIS lab's research is Semagix Freedom that has been used to develop semantic web applications for some of the world's biggest companies.

- Semantic Web enabled Web Services (SWWS) at HP, http://www.hpl.hp.com/semweb/swws.htm

  HP Labs Bristol has overall responsibility for two of case studies, which will concentrate on different aspects of procurement, to support developing SWWS platform. SWWS (Web Web Services) is a European 5th Framework project whose goal is to demonstrate how Semantic Web technology can be used to enable an open and flexible approach to web services. More specifically its goals are: 1) provide a comprehensive web services description framework; 2) define a web service discovery framework; and 3) provide a scalable web service mediation platform.

- Protégé Research Group at the Stanford University, http://protege.semanticweb.org/

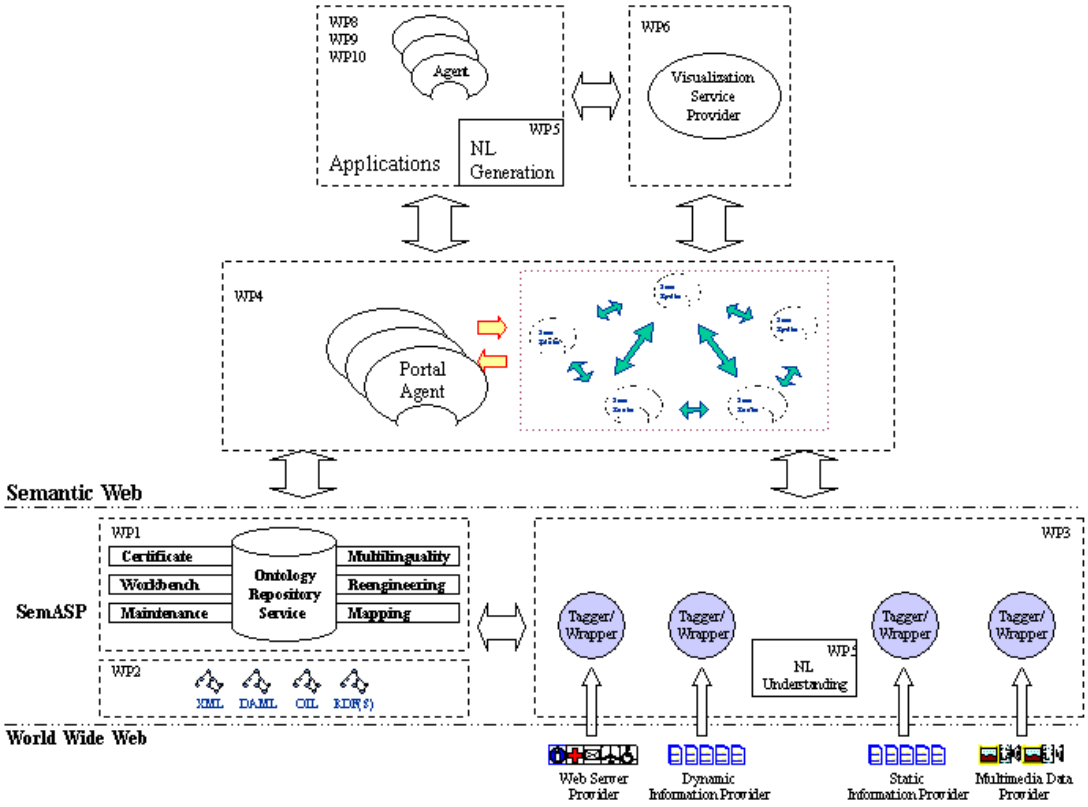
Protégé-20000 is on ontology editor and a knowledge-based editor. It provides support for editing Semantic Web ontologies.

Projects Corner: A column dedicated to dissemination of project outcomes

PROJECT ONE: Esperonto Project IST-2001-34373, provided by Oscar Corcho, ocorcho at isoco.com

URL: http://www.esperonto.net/
Short Description:

One of the biggest problems we nowadays face in the information society is information overload. The **Semantic Web** aims to overcome this problem by adding meaning to the Web, which can be exploited by **software agents** to whom people can delegate tasks. The aim of Esperonto is to **bridge the gap** between the actual World Wide Web and the Semantic Web by providing a service to "upgrade" existing content to Semantic Web content. **Ontologies** play a key role in this effort, aiming at unifying, bridging and integrating multiple heterogeneous, international and multilingual digital content.



Project Innovation:

- Wrapper technology and methodology to upgrade content to Semantic Web content
- Ontology lifecycle methodology
- Multilingual facilities
- Navigation and visualization techniques
- ASP to offer mentioned points to content providers
- To illustrate the technology, test cases are planned in the area of:
    - scientific discovery for biotechnology
    - online cultural archives
    - finding funding programs

Deliverables already finished:

- D1.1: State of the art in ontologies from the SW perspective
- D2.1: State of the art on Semantic Web languages
- D2.2: Report on SW languages evolution
- D3.1: State of the art on annotation tools and services
- D3.5: Annotation services for multimedia content
- D4.1: State of the art on indexation, routing techniques and negotiation techniques
- D5.1: State of the art on multilinguality for ontologies, annotation services and user interfaces
- D6.1: State of the art on visualisation technologies feasible for the Semantic Web
- D6.2: Ontology visualisation core services
- D6.3: Semantic Web content visualisation services

Future work: --

Partners: iSOCO (coordinator), UPM, UdS, UniLiv, IFI, CIDEM, Biovista, Residencia de Estudiantes

Position paper from Project Manager: --

Short Description of SEKT, DIP, KnowledgeWeb, and SDK Cluster Projects , By York Sure, Institut AIFB, Universität Karlsruhe (TH), Karlsruhe Germany

(In the forthcoming issues we will provide detailed presentation of projects)

SEKT – Semantically Enabled Knowledge Technologies (EU-IST integrated project)
http://sekt.semanticweb.org/

The vision of SEKT is to develop and exploit the knowledge technologies which underlie Next Generation Knowledge Management. We envision knowledge workplaces where the boundaries between document management, content management, and knowledge management are broken down, and where knowledge management is an effortless part of day to day activities. Appropriate knowledge is automatically delivered to the right people at the right time at the right granularity via a range of user devices. Knowledge workers will be empowered to focus on their core roles and creativity; this is key to European competitiveness. The SEKT strategy is built around the synergy of the complementary know-how of the key European centres of excellence in Ontology and Metadata Technology, Knowledge Discovery and Human Language Technology, a leading commercial exponent of semantic technology, together with a major European ICT organisation. The execution of SEKT is based on the integration of fundamental research, component development and integration driven by real world case studies in the public and private sectors. SEKT will provide new insights on knowledge technologies together with open reference implementations for others to build on.

DIP – Data, Information, and Process Integration with Semantic Web Services (EU-IST integrated project)
http://dip.semanticweb.org

The major mission of DIP is to further develop semantic web and web services, with the goal of enabling their combination into one. Web Services are the proper means for accessing semantically enriched data, and semantic enrichment of web services is essential for scalability and the maturation of the idea of Web Services. The new area that will be produced from the fruitful combination of the Semantic Web and Web Services may be called **Semantic Web Services**. Significant research work needs to be done before the infrastructure

underlying Web Services can make the vision behind the idea a reality. In a nutshell, it is the mission of DIP to make Semantic Web Service become a reality as the new infrastructure for eWork and eCommerce. The successful creation of such an infrastructure could potentially change the way electronic cooperation and business is conducted to the same extent that the original Web revolutionised electronic information access.

Knowledge Web – Realizing the Semantic Web (EU-IST network of excellence)
http://knowledgeweb.semanticweb.org

In a nutshell, it is the mission of Knowledge Web to strengthen the European industry and service providers in one of the most important areas of current computer technology: Semantic web enabled e-work and e-commerce. We will concentrate our efforts around the outreach of this technology to industry. Naturally, this includes education and research efforts to ensure the durability of impact and support of industry. Therefore, the main Objectives of Knowledge Web are: (i) outreach to industry, (ii) outreach to education and (iii) outreach to research.

SDK Cluster
http://www.sdk-cluster.org

It is the mission of SDK project cluster to strengthen European research and industry in the fields of Semantic Web and Semantic Web Services and to work towards international standardization together with the US-based DAML initiative. It does so through ensuring close cooperation between three major European research projects in the area of Semantic Web and Semantic Web Services: SEKT, DIP, and Knowledge Web. SDK promotes the projects' research results to both industry and academia through joint dissemination thereby strengthening the position of and providing a competitive advantage to European research and European industry and strengthening the world-wide research and standardization in Semantic and Semantic Web Services. Industry is supported both on the short-term and on the long-term, thereby enabling immediate implementation of Semantic Web and Semantic Web Services technologies and ensuring a lasting development of Semantic Web and Semantic Web Services technologies through standardization, education and research.

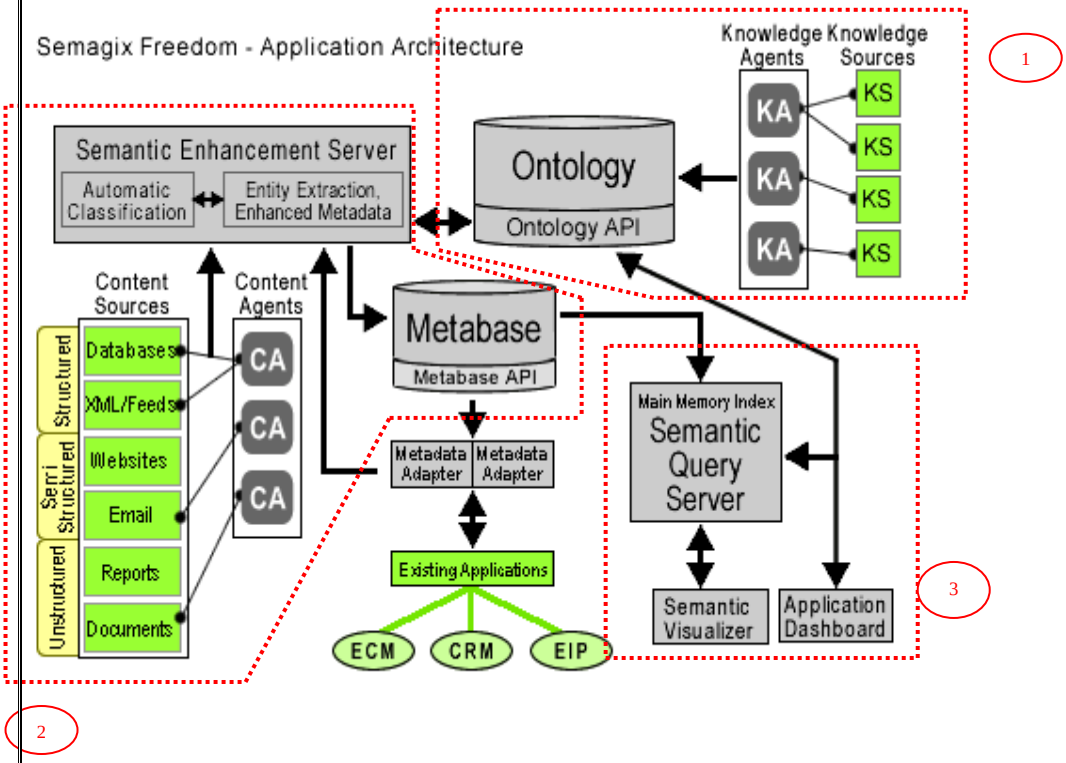
Industry Corner: A column dedicated to products presentation

By Miltiadis D. Lytras and Nikolaos Korfiatis

PRODUCT ONE:
Semagix Freedom, http://www.semagix.com

Enterprises are increasingly considering Semantic Technologies as a better and more comprehensive approach to exploiting large amounts of data they own or have access to, thus providing more insight and business intelligence, as well as actionable information that directly supports business-critical activities and decisions. Semantic Applications are capable of dealing with higher degrees of sophistication; they pose higher expectations for relevant and contextual search and browsing, integration of heterogeneous data/content, and text/content analytics. Semagix Freedom (http://www.semagix.com) is a pioneering technology that offers a high level of automation, high performance, and true scalability for building industry-strength and mission-critical Semantic Applications. Ontologies, not only the schema but populated with a huge number of facts or knowledge, are at the heart of Freedom's semantic technology. Critical capabilities Semantic Freedom offers include:

- graphical and user-friendly ontology schema design (see region 1 below)
- highly automated ontology population using ontology-driven knowledge extractors from trusted sources (with typical ontologies exceeding over one million instances) (region 1 of the Figure)
- automated metadata extraction from heterogeneous content including text (over 150 document types/formats), semi-structured data (html, XML), and structured data (relational databases), with support for a variety of delivery (content push) and access (content pull) mechanisms including dynamic and database driven web sites; the technology supports tens of millions of items and optional access to a Web index of over 4 billion pages (region 2 of the Figure)
- named entity identification, semantic ambiguity resolution and automatic text classification techniques for achieving a high degree of automation in ontology population and metadata extraction activities, with optional human curation support (region 2 of the Figure; semantic ambiguity resolution also used in regions 1 and 3)
- high-performance main-memory based complex queries that can involve both ontology and metadata (with the ability to perform 10 million queries per hour per server) with Java and HTTP APIs and XML/SOAP outputs (region 3 of the Figure)



Semagix Freedom is based on research in the LSDIS lab at the University of Georgia under the direction of Professor Amit Sheth. Technology based on the academic research was first licensed and commercialized by Taalee, Inc. founded and managed by Prof. Sheth. Taalee Inc., was acquired by Voquettee I nc., and then resulted in Semagix Ltd. In an earlier incarnation, this technology was used as a Taalee Semantic Search Engine [Townley 2000]. Semagix Freedom is currently used for information organization and integration in building knowledge-based e-Portals, analytical applications such as anti-money laundering and risk analysis, and sensitive government applications (Sheth et al 02, Sheth and Ramakrishnan 03, Sheth et al 04). Semagix has 11 installations of its Freedom product in the US, UK and Switzerland, with 5 of these installations being in Fortune 500 companies. These deployments of Semagix Freedom are across a variety of sectors, including two in financial services, four in Government, two in Pharmaceuticals and one in Manufacturing.

A. Sheth, C. Bertram, D. Avant, B. Hammond, K. Kochut, Y. Warke, Semantic Content Management for Enterprises and the Web, IEEE Internet Computing, July/August 2002, pp. 80-87.

A. Sheth and C. Ramakrishnan, "Semantic (Web) Technology In Action: Ontology Driven Information Systems for Search, Integration and Analysis," In *IEEE Data Engineering Bulletin, Special issue on Making the Semantic Web Real, December 2003, pp. 40-48.*

A. Sheth, B. Aleman-Meza, I. B. Arpinar, . Ramakrishnan, C. Halaschek, C. Bertram, Y. Warke, David Avant, F. S. Arpinar, K. Anyanwu, K. Kochut, Semantic Association Identification and Knowledge Discovery for National Security Applications, Journal of Database Management (to appear), 2004.

J. Townley, The Streaming Search Engine That Reads Your Mind, August 10, 2000.
http://smw.internet.com/gen/reviews/searchassociation/

Publications, presentation and further details can be found at: http://www.semagix.com/download.html and http://lsdis.cs.uga.edu/lib/

Semantic Web Technology Evaluation Ontology (SWETO): A test bed for evaluating tools and benchmarking applications

Boanerges Aleman-Meza, Amit Sheth, I. Budak Arpinar, Chris Halaschek and the SemDIS team
LSDIS Lab, Computer Science, UGA

The emergent Semantic Web community [SW] needs common infrastructure for evaluating new techniques and software which use machine processable data. Since ontologies are a centerpiece of most approaches, we believe that for evaluating and comparing tools for quality, scalability and performance, and for developing benchmarks for different classes of semantic technologies and applications, the Semantic Web community needs an open and freely available ontology with a large knowledge base (or description base) populated with real facts or data, reflecting real world heterogeneity of knowledge sources. If the use of tools is to be for advanced semantic applications, such as those in business intelligence and national security, then instances in the knowledge base should be highly interconnected. Thus, we present and describe a Semantic Web Technology evaluation Ontology (SWETO) test-bed [SWETO]. In particular, we address the requirements of a test-bed to support research in semantic analytics, as well as the steps in its development, including ontology creation, semi-automatic extraction, and entity disambiguation. SWETO has been developed as part of a NSF funded project using Freedom [Semagix], a commercial product from Semagix based in part on an earlier academic research [Sheth et al 2002], and is being made available openly for any non-commercial use.

Initially, SWETO was developed to be a large scale dataset for testing algorithms for discovery of semantic associations. The schema component of the ontology reflects the types of entities and relationships available explicitly (and implicitly) in Web sources. Given that we have available the use of Semagix Freedom, the selection of Web sources narrowed down to open, trusted sources, with metadata available having (semi-) structured layout for the viability of extraction and crawling. Essentially, with the Freedom toolkit, we created knowledge extractors by specifying regular expressions to extract entities from data sources. As the sources are 'scraped' and analyzed by the extractors, the extracted entities are stored in appropriate classes in an ontology. Given that we extracted semantic metadata from a variety of heterogeneous data sources, including Web pages, XML feed documents, intranet data repositories, etc., entity disambiguation is a crucial step. Freedom's disambiguation techniques were used for automatically resolving entity ambiguities in 99% of the cases, leaving less than 1% for human disambiguation (about 200 cases).

Given that SWETO is intended for ontology benchmark purposes, we continue to populate the ontology with diverse sources thereby extending it in multiple domains. Version is populated with well over 800,000 entities and over 1.5 million relationships, with the next larger release due out soon. SWETO access is available through browsing, XML serialization, and will soon be available though a Web service. SWETO has been used internally (LSDIS Lab) for discovery and ranking of semantic associations. Externally, our collaborators at UMBC are exploring trust extensions for SWETO, whereas within industry applications, Semagix uses it for evaluating fast semantic metadata extraction and enhancement in Marianas SDK.

SWETO is an effort of the SemDIS team, with significant effort in using Freedom by Gowtham Sannapareddy. It is partially funded by NSF-ITR-IDM Award # 0325464 and NSF-ITR-IDM Award # 0219649.

[Semagix] http://www.semagix.com/
[SemDis] http://lsdis.cs.uga.edu/Projects/SemDis
[Sheth et al 2002] Sheth, A., Bertram, C., Avant, D., Hammond, B., Kochut, K., Warke, Y.: Managing Semantic Content for the Web. IEEE Internet Computing, 6(4), 80-87. (2002).
[SW] http://www.w3.org/2001/sw/
[SWETO] http://lsdis.cs.uga.edu/Projects/SemDis/sweto

Semantic Discovery: Discovering Complex Relationships in Semantic Web

A NSF Medium ITR project

Semantic Web Technology Evaluation Ontology (SWETO)

The emergent Semantic Web community needs a common infrastructure for testing the scalability of new developments in software that makes use of machine processable data. One particular need is to have a large, high quality test ontology from which various ontology management tools can assess and test their scalability and other properties. Considering that there are somewhere between 20 to 50 ontology tools alone, the question arises: how do we compare them? Of particular interest is not just the schema of the ontology, but also the population (instances, assertions or description base) of the ontology. A populated ontology (ontology with instances or assertions) is critical for core semantic issues such as semantic disambiguation as well as being necessary for checking the scalability of tools and techniques, including e.g. reasoning techniques. An ontology of real-world scale is needed to build benchmarks for evaluating and comparing tools and techniques. An advantage of having access to both the schema of the ontology and the population of the ontology is that a class name can be understood by looking not only at the name of the class, but also by looking at the different instances that belong to that class (a class name by itself can be interpreted in different ways by different people). Many real word ontology have tens to few hundreds of classes and over one million objects (instances). An iterative process will be used to periodically extend SWETO description (schema) and description base (instances, assertions). By December 2003 end, we expect to have extracted at least 100,000 instances for SWETO v.1, and by end of 1Q 2004, we expect to have at least 1 million instances for SWETO v.2. A Semantic Web Technology Evaluation Ontology (SWETO) will serve the above purpose. It will be created (semi-)automatically by

1. designing the SWETO schema using an ontology design toolkit
2. identifying knowledge sources that can be used to populate parts of SWETO
3. utilizing extractor agents (written by humans without programming using a toolkit) to periodically and automatically extract parts of knowledge from various open and public sources,
4. semi-automatically disambiguating the knowledge (with limited human involvement)
5. integrating related knowledge to populate the SWETO
6. providing ability to export SWETO in RDF/RDFS and a version of OWL.

Semagix Freedom Toolkit will be used as the primary basis of the technology, suitably extending the results for standards compliance. All technical work will be done by the LSDIS Lab personnel working on the SAI and SemDIS projects. Semagix Freedom is based on licensing a technology developed at the LSDIS Lab and is available to the lab for such uses. W3C's SW Activity will help in securing knowledge sources (item 2 above). This ontology will be made available through W3C Semantic Activity for use of its active community members. We are looking at further funding resources to extend and continued maintenance of this ontology.

RuleML, SWRL and REWERSE: Towards a General Web Rule Language Framework

Gerd Wagner
Dept. Information & Technology, Eindhoven University of Technology, The Netherlands
G.Wagner@tm.tue.nl

We report on three important research efforts that attempt to develop languages and techniques for specifying and processing rules for the Semantic Web: the Rule Markup Language (RuleML) Initiative, the Joint Committee's Semantic Web Rule Language (SWRL) effort, and the European research project network REWERSE.

1 Introduction

Rules are considered to be a design issue for the Semantic Web (on top of the ontology layer in Tim Berners-Lee's Semantic Web layer cake) and have been a topic of discussion in the W3C Web Ontology Working Group, but have not been included in the Web ontology language OWL. It is expected that there will be a W3C Working Group for developing a W3C rule markup language, possibly starting in 2004.

Rule markup languages, that allow to express business rules as modular, stand-alone units in a declarative way, and to publish them and interchange them between different systems and tools, will play an important role for facilitating business-to-customer (B2C) and business-to-business (B2B) interactions over the Web.

In a narrow sense, a rule markup language is a concrete (XML-based) rule syntax for the Web. In a broader sense, it should have an abstract syntax as a common basis for defining various concrete sublanguages serving different purposes.

Rules may be considered at three different abstraction levels:

1. At the business domain level, rules are statements that express (certain parts of) a business/domain policy (e.g., defining terms of the domain language or defining/constraining domain operations) in a declarative manner, typically using a natural language or a visual language. Examples are:

   (R1) "The driver of a rental car must be at least 25 years old"

   (R2) "A gold customer is a customer with more than $1Million on deposit"

   (R3) "An investment is exempt from tax on profit if the stocks have been bought more than a year ago"

   (R4) "When a share price drops by more than 5% and the investment is exempt from tax on profit, then sell it"

   R1 is an *integrity rule*, R2 and R3 are *derivation rules*, and R4 is a *reaction rule* (see below for explanations of these rule categories). These appear to be the major semantic categories of business rules. Actually, many business rules appear to be reaction rules, which specify policies for real-world business behavior.
2. At the platform-independent level, rules are formal statements, expressed in some formalism or computational paradigm, which can be directly mapped to executable statements of a software platform. Rule languages used at this level are SQL:1999, OCL 2.0, and ISO Prolog. Remarkably, SQL provides operational constructs for all three business rule categories mentioned above: *checks/assertions* operationalize a notion of constraint rules, *views* operationalize a notion of derivation rules, and *triggers* operationalize a notion of reaction rules.
3. At the platform-specific level, rules are statements in a specific executable language, such as Oracle 10g views, Jess 3.4, XSB 2.6 Prolog, or the Microsoft Outlook 6 Rule Wizard.

Generally, rules are self-contained knowledge units that involve some form of reasoning. They may, for instance, specify

- static or dynamic integrity constraints (e.g. for constraining the state space or the execution histories of a system),
- derivations (e.g. for defining derived concepts),

– reactions (for specifying the reactive behavior of a system in response to events)

Given the linguistic richness and the complex dynamics of business domains, it should be clear that any specific mathematical account of rules, such as classical logic Horn clauses, must be viewed as a limited descriptive theory that captures just a certain fragment of the entire conceptual space of rules, and not as the only definitive, normative account. Therefore, in RuleML, a family of rule languages capturing the most important types of rules is being defined.

2 RuleML

The main categories of rules considered in RuleML [1,2] are derivation rules, integrity rules (constraints), reaction rules, production rules and transformation rules, as depicted in Figure 2. The concepts of derivation rules, integrity constraints, and reaction rules are considered to be meaningful both as (computation-independent) business rule categories and as (platform-independent) computational rule categories, whereas the concepts of production rules and transformation rules appear to be only meaningful as computational rule categories.

Notice that those categories whose name is in italics, such as DerivationRule, refer to an abstract concept of rule, while the others (with non-italicized names), such as "SQL:1999 View", refer to rule concepts of concrete languages such as SQL:1999.

Each part of a rule is an expression of some type, playing a specific role within the rule. The different parts of RuleML rules named by the roles they play, and their types, are listed in Table 1.

The main link between the different types of rules is the notion of a LogicalFormula or of a LogicalSentence, one of which being used in all of them. Traditionally, logical formulas are expressed in a language based on a predicate logic signature. However, OCL is the language of choice for expressing logical formulas referring to the state of a system whose structure is defined by a UML class model.

Table 2: RuleML rule expression components and their corresponding types.

Rule Expression Component	*Type of Expression Component*
Constraint	Logical Sentence
Condition	Logical Formula
Conclusion	Logical Formula
Post-Condition	Logical Formula
Transformation Invoker	Term
Transformation Return	Term
Triggering Event	Event Term
Triggered Action	Action Term
Produced Action	Action Term

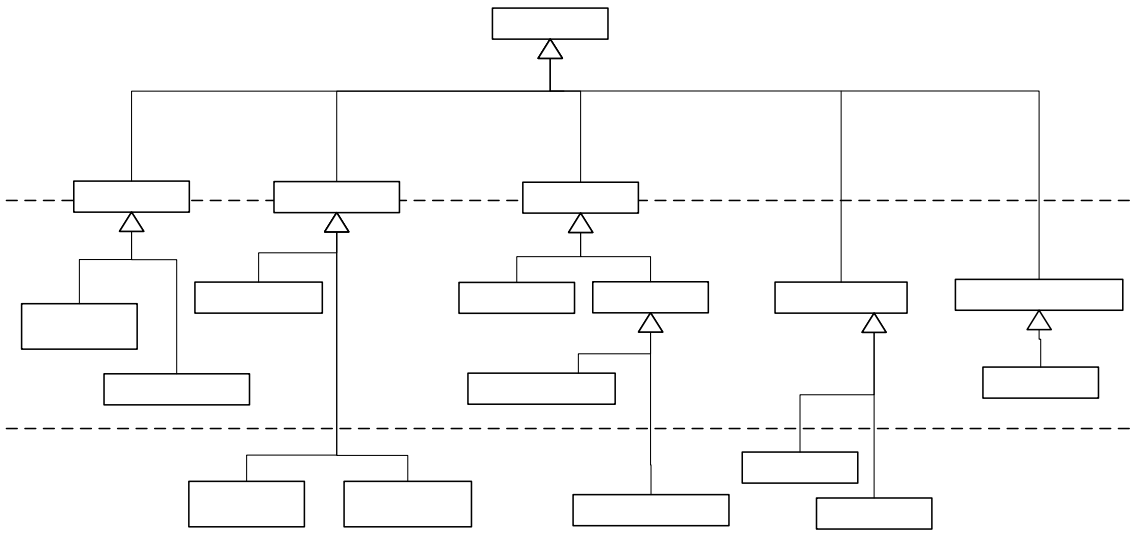


Figure 2: Rule concepts and rule expressions at different levels of abstraction.

The rules R2 and R3 above are examples of derivation rules. Another example, where a derived attribute in a UML class model is defined by a derivation rule, is the following:
A car is available for rental if it is not assigned to any rental contract and does not require service.
This rule defines the derived Boolean-valued attribute isAvailable of the class RentalCar by means of an association isAssignedTo between cars and rental contracts and the stored Boolean-valued attribute requiresService, as shown in the UML class diagram in Figure 3, where the derivation rule is expressed as an implicational OCL invariant that states that for a specific rental car whenever there is no rental contract associated with it, and it does not require service, then it is be available for a new rental.

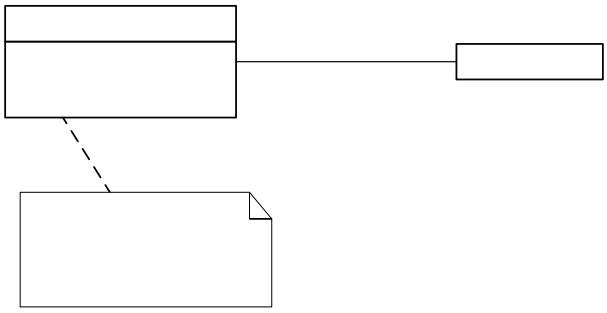


Figure 3: An OCL invariant defining the derived attribute isAvailable of the class RentalCar.

We now present the concrete XML syntax of this rule according to the simplified RuleML 0.85 syntax. The tag `<naf>` stands for *negation-as-failure*.

```
<imp>
      <_head>
            <atom>
                  <_opr>isAvailable</_opr>
                  <var>Car</var>
            </atom>
      </_head>
      <_body>
            <atom>
                  <_opr>RentalCar</_opr>
                  <var>Car</var>
            </atom>
            <naf>
                  <atom>
                        <_opr>requiresService</_opr>
                        <var>Car</var>
                  </atom>
            </naf>
            <naf>
                  <atom>
                  <_opr>isAssignedToRentalContract</_opr>
                  <var>Car</var>
                  </atom>
            </naf>
      </_body>
</imp>
```

RentalCar

RentalCarID : String
requiresService : Boolean
/isAvailable : Boolean

«invariant»
{ RentalContract->isEmpty
and not requiresService
implies isAvailable}

3 SWRL

Recently, the Joint US/EU ad hoc Agent Markup Language Committee, which had developed the influential DAM+OIL web ontology language proposal, has proposed the *Semantic Web Rule Language (SWRL)* [3], whose syntax is based on a combination of OWL DL and the Datalog sublanguage of RuleML. The semantics of SWRL is a straightforward extension of the first-oder model-theoretic semantics of OWL DL. Since, like RDF and OWL, SWRL is based on classical 2-valued logic, it cannot deal with partial information and with closed predicates. Therefore, it cannot encode the rule for available rental cars shown in Figure 3. But it can encode property-chaining rules like the rule for hasUncle:

```
<ruleml:imp>
    <ruleml:_body>
        <swrlx:individualPropertyAtom swrlx:property="hasFather">
            <ruleml:var>x1</ruleml:var>
            <ruleml:var>x2</ruleml:var>
        </swrlx:individualPropertyAtom>
        <swrlx:individualPropertyAtom swrlx:property="hasBrother">
            <ruleml:var>x2</ruleml:var>
            <ruleml:var>x3</ruleml:var>
        </swrlx:individualPropertyAtom>
    </ruleml:_body>
    <ruleml:_head>
        <swrlx:individualPropertyAtom swrlx:property="hasUncle">
            <ruleml:var>x1</ruleml:var>
            <ruleml:var>x3</ruleml:var>
        </swrlx:individualPropertyAtom>
    </ruleml:_head>
</ruleml:imp>
```

3 REWERSE

The acronym "REWERSE" stands for "Reasoning on the Web with Rules and Semantics". It is the name of a new research network funded by the EU Commission and Switzerland within the "6th Framework Programme" (FP6), Information Society Technologies (IST), see [4]. REWERSE aims at

- developing a coherent and complete collection of inter-operable reasoning languages for advanced Web systems and applications;
- testing these languages on context-adaptive Web systems and Web-based decision support systems selected as test-beds for proof-of-concept purposes;
- bringing the proposed languages to the level of open pre-standards amenable to submissions to standardization bodies such as the W3C.

4 Workshops

In 2004, there will be three international workshops on reasoning and rules in the Semantic Web:

- Business Rule Modelling and Markup for Intelligent Distributed Applications, 9/10 August 2004, PRICAI 2004 Workshop 3, Auckland University of Technology, New Zealand (http://www.pricai04.info/Workshop3.asp).
- The Second Workshop on Principles and Practice of Semantic Web Reasoning (PPSWR 2004) at the 20th International Conference on Logic Programming (ICLP) supported by the REWERSE Network of Excellence and the CoLogNet Network of Excellence. 6-10 September 2004, St Malo, France (http://www.pms.informatik.uni-muenchen.de/PPSWR04)
- Rules and Rule Markup Languages for the Semantic Web (RuleML-2004), 8 November 2004, Hiroshima, Japan, in conjunction with the 3rd International Semantic Web Conference (ISWC2004), http://iswc2004.semanticweb.org/CFParticipation/workshops.html

References

[1] Boley, H., Tabet, S., Wagner G.: Design Rationale of RuleML: A Markup Language for Semantic Web Rules. International Semantic Web Working Symposium (SWWS), June 2001, Stanford, USA.

[2] http://www.RuleML.org

[3] http://www.daml.org/rules/

[4] http://www.REWERSE.net

Students Corner: A column dedicated to dissemination of students work concerning Semantic Web

Title: A Flexible Approach for Ranking Complex Relationships on the Semantic Web
by Chris Halaschek

Recently, information retrieval over semantic metadata extracted from the Web has received an increasing amount of interest in both industry and academia. In particular, discovering complex and meaningful relationships among this metadata is an interesting and challenging research topic. Just as ranking of documents is a critical component of today's search engines, the ranking of complex relationships will be an important component in tomorrow's Semantic Web analytics engines. Thus, building upon our recent work on specifying and discovering complex relationships in RDF data, called Semantic Associations, I have defined a flexible ranking approach which can be used to identify more interesting and relevant relationships in the Semantic Web. Additionally, I have demonstrated the ranking scheme's effectiveness through an empirical evaluation over a Semantic Web test-bed (SWETO: available at http://lsdis.cs.uga.edu/Projects/SemDis/Sweto/).

Chris Halaschek is a Master's student at the University of Georgia working in the Large Scale Distributed Information Systems (LSDIS) Lab, directed by Dr. Amit P. Sheth. He is co-advised by Drs. I. Budak Arpinar and Amit P. Sheth. Halaschek is a research assistant funded by the lab's NSF-ITR-IDM Award#0325464 titled 'SemDIS: Discovering Complex Relationships in the Semantic Web' (details available at http://lsdis.cs.uga.edu/Projects/SemDis). More information regarding his work, including publications, is available at http://lsdis.cs.uga.edu/~ch.

Title: Semantic Web Languages and Ontology Mediation

by Jos de Bruijin

The current research interests of Jos de Bruijn include Semantic Web languages and Ontology Mediation.
In the context of Semantic Web languages, Jos' main interest lies in rule extensions of Semantic Web languages. In that context, he is involved in the WSML Working Group (http://www.wsmo.org/wsml/), where he investigates formal properties of ontology languages and possibilities for rule extensions of current Semantic Web languages. A subset of Description Logic-based languages, such as OWL, can be directly translated into the Logic Programming language Datalog. This fragment, identified by Raphael Volz, is called *Description Logic Programs* (DLP). Jos currently looks into ways to restrict OWL Lite to DLP and extend this variant into a more expressive Logic Programming language. In the context of Ontology Mediation, Jos is involved in the European Commission funded project SEKT (http://sekt.semanticweb.org/), which stands for Semantically-Enabled Knowledge Technologies. The goal of SEKT is to further the development of the Semantic Web by on the one hand improving the ontology management infrastructure and on the other hand applying Natural Language and Machine Learning technologies in order to simplify the processes of ontology construction and document annotation. Jos coordinates the work in the area of Ontology Mediation, where he brings together different partners to work together towards an Ontology Mapping Specification Language, as well as specific techniques for ontology alignment and instance alignment, along with a graphical tool supporting ontology mapping and run-time mediation components, providing the actual mediation. The Ontology Mapping Specification Language will use rule extensions of Semantic Web languages for the specification of ontology mappings. Furthermore, the mapping language will use elementary mapping patterns (similar to rule patterns) for the specification of ontology mappings. Elementary mapping patterns can be combined into more complex patterns, which capture frequently occurring types of mappings. These mapping patterns will be very helpful in finding and specifying mappings between ontologies.

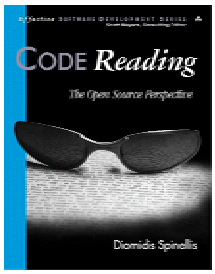
Jos de Bruijn finished his education in Computer Science at the Delft, University of Technology in 2003 with the degree of Master of Science (MSc). He is currently working as a PhD student at the Digital Enterprise Research Institute (DERI), University of Innsbruck, Innsbruck, Austria. His research interests include Semantic Web Languages and Ontology Mediation.

Book Corner: A column dedicated to the presentation of interesting books

Book title: Ontological Engineering
Book authors: Asunción Gómez-Pérez, Mariano Fernández-López, Oscar Corcho
Publisher: Springer-Verlag

Ontologies provide a common vocabulary of an area and define, with different levels of formality, the meaning of the terms and the relationships between them. *Ontological Engineering* refers to the set of activities that concern the ontology development process, the ontology life cycle, the methods and methodologies for building ontologies, and the tool suites and languages that support them. During the last decade, increasing attention has been focused on ontologies. Ontologies are now widely used in Knowledge Engineering, Artificial Intelligence and Computer Science; in applications related to areas such as Knowledge Management, Natural Language Processing, e-Commerce, Intelligent Information Integration, Bio-Informatics, Education; and in new emerging fields like the Semantic Web. The book presents the major issues of Ontological Engineering and describes the most outstanding ontologies that are currently available. It covers the practical aspects of selecting and applying methodologies, languages, and tools for building ontologies. Ontological Engineering is of great value to students and researchers, and to developers who want to integrate ontologies in their information systems.
Written for: Researchers, postgraduates, practitioners, libraries, institutions, industry, scientists, students
URL: http://webode.dia.fi.upm.es/ontologicalengineering/

Code Reading: The Open Source Perspective
Diomidis Spinellis. Addison Wesley, 2003. ISBN 0-201-79940-5.



Software source code is the definitive medium for communicating a program's operation and for storing human knowledge in a form a computer can process. Even though 40% to 70% of the effort that goes into a software system is expended after the system is first written, most programming books and courses focus on the skills required for writing programs from scratch. However, the activities following the initial program development invariably involve reading, understanding, and modifying the original code. In addition, the unrelenting, inevitable accumulation of legacy code; the increasing emphasis placed on software reuse; the high human turnover rates associated with the software industry; and the rising importance of open-source development efforts and cooperative development processes (including outsourcing, code walkthroughs, and extreme programming) make code reading an essential skill for today's software engineer. The book *Code Reading* provides background knowledge and specific techniques for reading code written by others. The book contains more than 600 concrete examples taken from important, real-life, open-source code systems covering most concepts related to code that are likely to appear before a software developer's eyes, including programming constructs, data types, data structures, control flow, project organization, coding standards, documentation, and architectures. A CD with 16 million lines of open-source code, accompanying the book, provides the complete context for all the presented examples. The book inaugurated Addison-Wesley's *Effective Software Development Series*, edited by Scott Meyers, and received the 2004 Software Development Productivity Award in the "Technical Books" category. Additional details can be found on book's web site http://www.spinellis.gr/codereading.

Diomidis Spinellis

Diomidis holds an MEng in Software Engineering and a PhD in Computer Science both from Imperial College (University of London, UK). Currently he is an Associate Professor at the Department of Management Science and Technology at the Athens University of Economics and Business, Greece. He has written more than 70 technical papers in the areas of software engineering, information security, and ubiquitous computing. His book "Code Reading: The Open Source Perspective" inaugurated Addison Wesley's "Effective Programming Series". He has contributed software to the BSD Unix distribution, the X Window System, and is the author of a number of open-source software packages, libraries, and tools. Dr. Spinellis is a member of the ACM, the IEEE, the Greek Computer Society, the Technical Chamber of Greece, and a founding member of the Greek Internet User's Society. He is a co-recipient of the Usenix Association 1993 Lifetime Achievement Award.

Forthcoming Book

Intelligent Learning Infrastructures for Knowledge Intensive Organizations: A semantic web perspective. Edited by Miltiadis D. Lytras, ELTRUN and Ambjorn Naeve, KTH. IDEA Group Publishing, Early 2005

Semantic Web for Beginners Corner:

Courses taught for SW:

By Amit Sheth

I teach two graduate courses that might be relevant here:

Semantic Web: http://lsdis.cs.uga.edu/SemWebCourse/index.htm

Semantic Web Services and Processes: http://lsdis.cs.uga.edu/SemWebProcess/

Cheers,
Amit

Two brief articles which give an introduction to Ontologies and Semantic Web:

Y. Sure. **Fact Sheet on Semantic Web.** In: KTweb -- Connecting Knowledge Technologies Communities, available at http://www.ktweb.org/doc/Factsheet-SemanticWeb.pdf

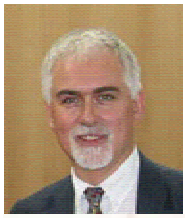
Y. Sure. **Fact Sheet on Ontologies.** In: KTweb -- Connecting Knowledge Technologies Communities, available at http://www.ktweb.org/doc/Factsheet-Ontologies-0306.pdf

SIG Board Members

Ambjorn Naeve
KMR Group, Interactive Learning Environments, Centre for user-oriented IT Design (CID)
Department of Numerical Analysis and Computer Science (NADA)
Royal Institute of Technology (KTH) Stockholm, Sweden.
E-mail: amb@nada.kth.se

Amit P. Sheth
Large Scale Distr. Info. Sys. Lab
Department of Computer Science
University of Georgia, Athens GA
Email: amit@cs.uga.edu (UGA)
Web Site: http://lsdis.cs.uga.edu/~amit/

William I. Grosky
Department of Computer and Information Science
University of Michigan, USA
Email: wgrosky@umich.edu
Web Site: www.engin.umd.umich.edu/~wgrosky



Gottfried Vossen
Institut für Wirtschaftsinformatik
Universität Münster, Germany
E-Mail: vossen@uni-muenster.de
Web Site: dbms.uni-muenster.de

Ram Ramesh
School of Management
SUNY at Buffalo
E-Mail: rramesh@acsu.buffalo.edu
Web Site: www.mgt.buffalo.edu/CFDOCS/Forms/faculty/bios/faculty.cfm?fac=rramesh

Miltiadis D. Lytras
ELTRUN, the E-Business Center
Department of Management Science & Technology,
Athens University of Economics and Business, Athens, Greece,
E-mail: mdl@aueb.gr
Web Site: www.eltrun.aueb.gr

Kinshuk
Advanced Learning Technologies Research Centre, Information System Department, Massey University, New Zealand
Chair, IEEE Learning Technology Task Force
E-mail: kinshuk@massey.ac.nz,
Web Site: infosys.massey.ac.nz/~kinshuk

Gerd Wagner
Department of Technology Management
Eindhoven University of Technology Eindhoven, The Netherlands
E-mail: G.Wagner@tm.tue.nl
Web Site: is.tm.tue.nl/staff/gwagner

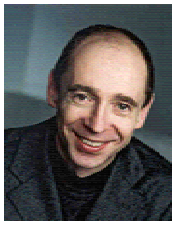
Demetrios Sampson
Dept of Technology in Education and Digital Systems, University of Piraeus, Greece
Head of Unit, Advanced e-Services for the Knowledge Society (ASK) Research Unit
Informatics and Telematics Institute (ITI)
Center for Research and Technology - Hellas (CERTH)
E-mail: Sampson@iti.gr
Web site: www.ask.iti.gr/

Lina Zhou
Department of Information Systems
University of Maryland, USA
E-mail: zhoul@umbc.edu
Web site: userpages.umbc.edu/~zhoul/

York Sure
Institut AIFB
Universität Karlsruhe (TH), Karlsruhe Germany
E-mail: sure@aifb.uni-karlsruhe.de
Web Site: http://www.aifb.uni-karlsruhe.de/WBS/ysu/

Henry M. Kim
Schulich School of Business
York University, Toronto, Canada
E-mail: hkim@schulich.yorku.ca,
Web Site: www.yorku.ca/hmkim

Richard Benjamins
Intelligent Software Components, S.A.
E-mail: rbenjamins at isoco.com
Web Site: www.isoco.com



Ralf Klischewski
Department for Informatics
University of Hamburg, Germany
E-mail: klischewski@informatik.uni-hamburg.de
Web Site: swt-www.informatik.uni-hamburg.de/people/rk.html

Asunción Gómez-Pérez
Laboratorio de Inteligencia Artificial (LIA), Facultad de Informatica (FI)
Universidad Politecnica de Madrid (UPM), Spain
E-mail: asun@fi.upm.es
Web Site: delicias.dia.fi.upm.es/miembros/ASUN/asun_CV_Esp.html

Rajiv Kishore
School of Management
SUNY at Buffalo, USA
E-mail: rkishore@buffalo.edu
Web Site: www.mgt.buffalo.edu/CFDOCS/Forms/faculty/bios/faculty.cfm?fac=rkishore

Shiyong Lu
Multimedia Information Systems Group
Department of Computer Science
Wayne State University, USA
E-mail:shiyong@cs.wayne.edu
Web Site: www.cs.wayne.edu/~shiyong/

Farshad Fotouhi
Department of Computer Science
Wayne State University, Detroit, USA
E-mail: fotouhi@cs.wayne.edu
Web Site: www.cs.wayne.edu/fotouhi

Oscar Corcho
Intelligent Software Components, S.A.
E-mail: ocorcho at isoco.com
Web Site: www.isoco.com

Lakshmi S. Iyer,
Rahul Singh
Al Salam
Information Systems & Operations Management
Bryan School of Business and Economics
The University of North Carolina at Greensboro
E-mail: lsiyer@uncg.edu
Web Site: www.uncg.edu/bae/people/lsiyer/

Ming Dong
Assistant Professor
Computer Science Department
College of Science, Wayne State University, USA
E-mail: mdong@cs.wayne.edu
Web Site: www.cs.wayne.edu/~mdong/

Jesus Contreras
Intelligent Software Components, S.A.
E-mail: Jcontreras at isoco.com
Web Site: www.isoco.com

Henrik Legind Larsen
Department of Computer Science
Aalborg University Esbjerg, Denmark
E-mail: legind@cs.aue.auc.dk
Web Site: http://www.cs.aue.auc.dk/~legind/

We welcome new members in our SIG Board:

Prof. James Hendler, University of Maryland, USA

Dr. John Davies,
Dr John Davies leads the Next Generation Web research group at BT Exact. Current research interests centre around the application of semantic web technology to knowledge management and web services. He is industrial chair of the Semantic Web Services Initiative, co-organiser of the 1st European Semantic Web Symposium, and is a Chartered Engineer. He has written and edited many papers and books in the areas of web-based information management, knowledge management, virtual communities and the semantic web. Earlier research at BT led to the development of a set of knowledge management tools which are the subject of a number of patents. These tools are now marketed through Exago Ltd (www.exago.com), of which Dr Davies is CTO.

SIGSEMIS News

Volume	1
Issue	1

April 2004
Theme: "SW Challenges for IS"

The Official Bimonthly Newsletter of AIS Special Interest Group on Semantic Web and Information Systems

IN THE FORTHCOMING ISSUE (JUNE):

- A Call for Papers for the *International Journal on Semantic Web and Information Systems*
- A Call for Contributions for the *Encyclopaedia of Semantic Web Research*
- Ten New Columns from well known international Academics and Researchers and an open call for Columnists
- Presentation of Three Top Research Centers / Interviews with Scientific Coordinators
- Announcement of the SIGSEMIS Sponsored Award for the **Best PhD Student Work on SW**
- Detailed Presentation of Projects in USA and EU

JOIN SIGSEMIS

Support our activities and Join our SIG in Association for Information Systems. A 10$ fee is required and the main reason for this is to sponsor the student awards.

http://www.aisnet.org/sigs.shtml

GET INVOLVED…

◘ Provide Content ◘ Post Articles ◘ Become a Columnist ◘ Announce Forthcoming Events ◘ Ask for guidance ◘ Initiate Discussions ◘ Browse Newsletter ◘ Enter to our Journal ◘ Browse Interviews ◘ Use E-learning Facility (soon) ◘ Find Members Details ◘ Learn about SIGSEMIS Sponsored Tracks, Books, Symposiums and Events ◘ Apply for the PhD student award (soon) ◘ Contribute to our Encyclopaedia of Semantic Web Research ◘ Contact SIG SEMIS Board and Members and Committees members

http://www.sigsemis.org

CONTACT AIS SIGSEMIS NEWS Editorial Board

Please provide any comments, inquiries, ideas, etc to Miltiadis D. Lytras at mdl@aueb.gr

DISTRIBUTE THIS DOCUMENT

Feel free to forward this issue to your colleagues, mailing lists and anyone who may be interested in. Thanking you in advance.

Thank you…