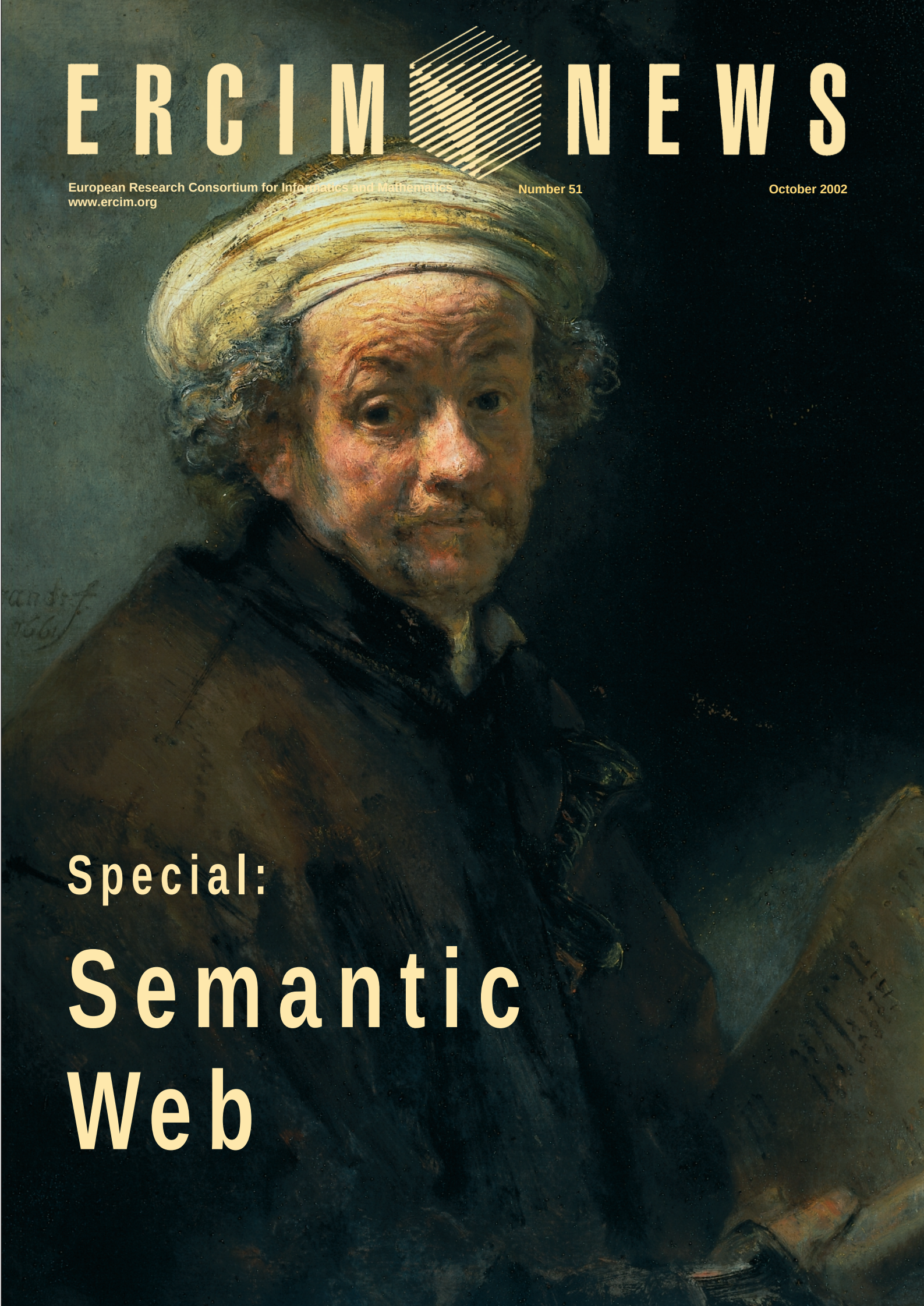
ERCIM NEWS

European Research Consortium for Informatics and Mathematics
www.ercim.org

Number 51

October 2002



Special:

Semantic Web

Contents

Cover: Rembrandt's self-portrait as the Apostle Paul. Photograph courtesy of the Rijksmuseum, Amsterdam.

Read about the relation between Rembrandt and the Semantic Web in Lynda Hartman's article 'Hypermedia Presentation Generation on the Semantic Web' on page 36.

Next issue:
January 2003

Special theme:
Embedded Systems

Photo: LeFevre Communications

Tim Berners-Lee and Eric Miller contributed the article 'The Semantic Web lifts off' on page 9.

Timothy J. Berners-Lee holds the 3Com Founders chair at the Laboratory for Computer Science at MIT, and directs the World Wide Web Consortium, an open forum of companies and organizations with the mission to lead the Web to its full potential. He is the author of Weaving the Web.

Eric Miller is the Activity Lead for the W3C World Wide Web Consortium's Semantic Web Initiative.

The FNR becomes ERCIM's Luxembourg Member

by Raymond Bausch

FNR, the Luxembourg National Research Fund, (Fonds National de la Recherche) joined ERCIM in July 2002.

The National Research Fund is a public establishment with scientific, financial and administrative autonomy, set up by the Law of 31 May 1999 in order to further stimulate research activities in Luxembourg. To fulfil this mission, the Fund develops multi-year research programs and assures their implementation by allocation of the financial means put at its disposal.

Five national multi-year priority programs are currently ongoing, among them 'Security and efficiency of new practices in e-commerce for all socio-economic actors (SE-COM)', which has a budget of 7 500 000 Euro for five years. In addition, the National Research Fund subsidises accompanying measures to strengthen the general framework of scientific research in Luxembourg, eg subsidising the organisation of scientific conferences and the mobility of researchers.

In ERCIM, the National Research Fund represents researchers in Luxembourg working in informatics and mathematics at the following institutions:

The CRP-Santé is a public research centre in Luxembourg whose main research and development activities lie in the fields of public health, health care, and biomedicine ('CRP' stands for Centre de Recherche Public). The CRP-Santé aims to promote scientific progress, technology transfer and innovation within the framework of international and European research and development programs.

The CRP-Gabriel Lippmann is one of Luxembourg's leading players in the fields of scientific research and technology transfer. The centre focuses its activities on some of the most crucial topics of interest of the 21st century, including nanotechnology, novel materials, sustainable management of natural resources, biotechnology, and last but not least technologies of the information society and their relevant legal frameworks. As a public research centre, it makes a point of keeping its commitment in line with the economic usefulness of its scientific achievements, however significant they may be.

The CRP-Henri Tudor, founded in 1987, is an integrated centre of applied research and technological innovation. Its main activities range from research to technology transfer including support for entrepreneurial activities and an incubator for start-ups. 170 employees are working at the centre, of whom 140 are engineers active in the fields of information and communication technology, and industrial and environmental technology.

The Centre Universitaire is a public institute in Luxembourg offering university studies. It consists of a number of departments: law and economic sciences, business administration and computer science, literature and humanities, science, professional training for lawyers, and professional training for teachers.

The Institut Supérieur de Technologie (IST) is a public institute for higher technical education composed of four departments: electrical, mechanical, and civil engineering, and computer science. Besides teaching, the IST also actively performs fundamental and applied research. Several collaborations exist with different national and international institutions, both in teaching and research. Close relationships are also maintained not only with numerous European universities, but also with industry in Luxembourg and especially with the CRP Henri Tudor for the purpose of technology transfer.

The implementation of a new University of Luxembourg with a major focus on research activities is currently in preparation and a new law will restructure the institutions of higher education in Luxembourg.

Links:
FNL: http://www.fnr.lu/
CRP-Santé: http://www.crp-sante.lu/
CRP-Gabriel Lippmann: http:// www.crpgl.lu/
The CRP-Henri Tudor: http://www. tudor.lu/
Centre Universitaire: http:// www.cu.lu/
Institut Supérieur de Technologie: http://www.ist.lu/

Please contact:
Raymond Bausch
Fonds National de la Recherche
Tel: +352 26 1925 31
E-mail: raymond.bausch@fnr.lu

FRAMEWORK VI

ERCIM submitted ten Expressions of Interest informing the Commission of its intention to propose a series of Networks of Exellence within Framework VI. Interested scientists should contact the ERCIM office at office@ercim.org.

Anaerobic Digestion: Multidisciplinary Networking for Better Understanding and Optimal Monitoring — understAnD
Objectives: This proposal aims at establishing Europe as the scientific leader in AD processes through the sharing of experiences and efficient networking between European leaders in multidisciplinary fields. The main goal will be to lower the fragmentation of European research and to increase the scientific understanding of the AD process, leading to a higher competitiveness of European industry in this field.
Co-ordinator: Jean-Philippe Steyer, INRA.

Applications of Computational Mathematics and Statistics in Science and Technology
Objectives: The aspiration of the network is to link together mathematicians working in national laboratories, universities, and industrial R&D departments.
Co-ordinator: Mario Arioli, CLCR.

Cross-Language Evaluation Forum — CLEF
Objectives: The network for multilingual cross-language system evaluation intends to stimulate and assist the development of mono- and cross-language information retrieval systems working with European languages and languages of importance to Europe.
Co-ordinator: Carol Peters, CNR.

NoE for Digital Libraries — DELOS
Objectives: Digital libraries should enable access to all human knowledge any time and anywhere, in a friendly, multi-modal, effective way, overcoming barriers of distance, language, and culture. The network will define a joint program coordinating the activities of European research teams with the aim of developing next generation digital library technologies, thus making feasible the implementation of this grand vision.
Co-ordinator: Costantino Thanos, CNR

European Excellence in Health Information Technology — E2Health
Objectives: This proposal intends to reinforce European expertise in the development of intelligent, mobile and secure systems and services aimed at providing patients with personalized healthcare and information services, and at stimulating health promotion and disease prevention in the general population. The network will include experts in key fields in order to strengthen and develop community scientific and technological excellence. The network will integrate, at European level, existing national research capacities and will establish Europe as the international scientific leader in the relevant R&D domains.
Co-ordinator: Stelios Orphanoudakis, FORTH.

Foundations, Software Infrastructures and Applications for Large-Scale Distributed, Grid and Peer-to-Peer Technologies — CoreGRID
Objectives: The network will gather together European institutions (mainly academics but with support from European industry) carrying out R&D activities in the area of large scale distributed computing using Grid and Peer-to-Peer (P2P) technologies.
Co-ordinator: Michel Cosnard, INRIA.

Image and Video Understanding: Extracting Semantics from Visual Data — XViS
Objectives: The network will act as a virtual lab in which the main actors will outline an ambitious vision for the future and prepare a corresponding roadmap. Support will be sought on a pan-European scale in order to establish the EU as the international scientific leader in this domain. Closer collaboration among the many European groups that are active in this area will be fostered by stimulating the exchange of results and ideas, by maximizing the dissemination of information, by pooling complementary expertise through joint research initiatives and by creating focussed task forces (either virtually or physically).
Co-ordinator: Eric Pauwel, CWI.

Information Systems for Environmental Degradation and Disaster Monitoring and Management – ISEM
Objectives: To conduct on a European scale a cooperative research effort defining knowledge-based systems with generic components and application modules dedicated to environmental degradation and disaster monitoring and management.
Co-ordinator: Isabelle Herlin, INRIA.

Intelligent Environments of Use – In-EU
Objectives: To mobilize Europe's R&D capabilities in those fields that constitute the discipline of Human Computer Interaction. A range of short-, medium- and long-term activities aimed at providing a new basis for designing and developing user interface software and technologies will be pursued in order to empower citizens' interaction with distributed and context-aware user environments in the emerging Information Society.
Co-ordinator: Costantine Stephanides, FORTH.

Semantic Web
Objectives: The aim of this proposal is to establish Europe as the international scientific leader in the next-generation Semantic Web.
Co-ordinator: Julius Stuller, CRCIM.

EURO-LEGAL

News about legal information relating to Information Technology from European directives, and pan-European legal requirements and regulations.

Metatags that cause Trademark Infringement
The UK has seen a recent landmark case that contains the first judicial analysis of trademark infringement where the use of the mark was made via metatags. Website owners rely on metatags to ensure that search engines bring visitors to their sites. The search engines use spider software to create indexes of catalogued websites by reviewing the sites' metatags for keywords. In some cases these metatags could include a trademark and website owners can use this as a method of redirecting internet users to their site. In order to establish trademark infringement, it is necessary to show that the infringing name or mark has been used 'in the course of trade'. This landmark case rules that despite the fact that these metatags are invisible to the website visitor they are visible when translated into a search result on the user's browser. This case provides further protection to registered trademark owners from other websites using their brands to attract visitors to their site.

Internet Sales — New Rules for EU Consumers
The Brussels Regulation, which came into force on 1 March 2002, introduced new rules for determining where claims arising from civil and commercial disputes involving parties from different countries can be heard within the EU. The rules apply to claims made by consumers arising out of sales of goods or services over the Internet.

Court proceedings must be brought within the Member State where the consumer lives. This right of jurisdiction cannot be waived even if the terms of the contract imply something else. Until now the contract must have been concluded while the consumer was in his or her Member State. The new regulation, however, introduces an additional jurisdiction that deals specifically with cross-boarder sales to consumers over the internet, where the trader "directs such activities to that Member State or to several Member States including that Member State". In other words, if a consumer from the UK purchases goods on the internet whilst on holiday in another Member State (eg France) he or she can bring a claim against the trader in France providing the trader "directed his activities" to the home State of the consumer (in this case, the UK).

Although this has not yet been tested in the courts it is likely to impact more heavily on traders who offer their goods or services in a number of languages and accept payment in different currencies, including the Euro. If traders wish to avoid claims in the future, they may have to consider accepting orders only from consumers residing in their own Member State and in the currency of that Member State. This is likely to impact on the value of trading on the internet. It will be interesting to see how case law develops in this area.

by Heather Weaver, CLRC
Tel: +44 1 235 446151
E-mail: H.Weaver@rl.ac.uk

Articles in this Section

A Few Words about the Semantic Web...

by Jérôme Euzenat

A first illustration of the direction of the Semantic Web consists in noting that the Web as it currently exists is very difficult to search. For instance, looking for a book about Agatha Christie is not easy, since many current search engines discard the word 'about' as meaningless, and may then return a plethora of pages referring to books by Agatha Christie. If we want machines to perform effective document searches, we must help them a bit by telling them what documents mean. The correct identification of relational information (such as relating a Book object to a Person object named 'Agatha Christie' using the 'about' or 'author' link) would be a step towards a more accurate search. It is thus natural that the first language made for the Semantic Web, RDF (Resource Description Framework), emphasises these relationships.

The idea of the Semantic Web (a term coined by Tim Berners-Lee, the creator of the Web) involves annotating documents with 'semantic markup', that is, markup that is not interpreted for display but rather as an expression of document content. This is often described as 'a Web for machines' as opposed to a Web to be read by humans. This idea was experimented with on a small scale during the nineties, both by the SHOE system, developed at the University of Maryland, and Ontobroker, developed at the University of Karlsruhe.

The development of the Semantic Web took off with the DARPA Agent Markup Language (DAML) initiative in the USA. Its goal was to provide the semantic markup language successor to SHOE. Shortly after, European researchers created the Ontoweb thematic network in order to federate the research in these fields, and W3C launched its long-awaited Semantic Web activity that took over the development of RDF and its extensions. These had previously been advanced by the two other groups. Research activities for building the Semantic Web are now active all over the world, and are central to the 'knowledge technologies' area of the European Union's 6th framework program. We invited several major players in Semantic Web research (W3C, the European Commission, INTAP and the Ontoweb network) to present their current and future activities in the domain.

The applications of the Semantic Web are limited only by our imagination. A more elaborate scenario for the Semantic Web involves not only the facility for improved searching, but also for selecting, assembling and triggering services found on the Web, matching interest profiles and resource descriptions, automatically rescheduling and reassembling services upon an unexpected event, notifying customers of a schedule change on their favourite communication medium or mining biomolecular databanks at night.

For instance, the 'travel agent scenario' deals with a software agent able to plan a complex trip involving several forms of transportation, hotel stays, and conference and entertainment registration by resolving heterogeneous constraints. These could include using trusted hotels in preferred areas, taking eating requirements into account, finding connections, using only prescribed airlines and frequent-flyer-affiliated car rental, and minimising the total cost of the trip. Such a scenario necessitates four main capabilities on the part of the agent:

1. a network of well-annotated resources, eg plane schedules and fares, hotel locations and facility descriptions, etc
2. a reservoir of available knowledge (or an ontology), eg a bus is a means of transportation, Sardinia is in Italy, etc
3. a description of user preferences, eg food preferences, accommodation needs, frequent flyer programs, agendas, etc
4. inference capabilities, eg taxonomical inference, temporal inference, trust propagation, etc.

In short, this scenario involves doing what a very good assistant could do, but in a more systematic and provable way (though perhaps in a less flexible and agreeable manner). However, the Semantic Web should not be confused with what has sometimes been called 'strong artificial intelligence' for three reasons. Firstly, there is no anthropomorphic claim to the Semantic Web – instead it is intended to complement humans in areas where they do not perform very well (ie, dealing quickly with large amounts of information, working continuously, analysing large texts for certain pieces of information, etc). Secondly, it takes into account many lessons learnt from the Web; that is, being large, inconsistent and spread, Semantic Web applications must be scalable, robust and decentralised. Lastly, it must be adapted to the context in which it will evolve.

The research aspects of the Semantic Web were debated at a recent ERCIM strategic workshop that was held in Sophia-Antipolis last year. It identified several important research directions to be investigated:

- identification and localisation, for the correct and universal addressing of resources

...and its Development in the ERCIM Institutes

- relationships between semantic models, for adequately transforming and interpreting the content of the Semantic Web
- tolerant and safe reasoning, for overcoming the inherent heterogeneity and inconsistency of an open world, and
- facilitating Semantic Web adoption, by imagining original knowledge-capture means and adequate growth models.

From this, it is evident that a number of fields can contribute to the development of the Semantic Web. Many groups are now applying their knowledge at various levels, from how to put category information on Web pages to how computers can be related to meaning. As a matter of fact, one of the good features of the Semantic Web is that it has encouraged the close cooperation of many people from different backgrounds. We can already see structured markup developers talking logic with knowledge representation researchers, database engineers talking protocols with multimedia designers, electronic commerce managers starting to talk to agent developers. Of course, the Semantic Web primarily takes advantage of the techniques developed on the Web and in knowledge representation, but other fields also have important contributions to make. These include document management for manipulating complex markups, databases for dealing efficiently with assertions, digital libraries for metadata and search schemes, logics for crafting language semantics and reasoning systems, interface design for offering new views of these marked documents, and agents for gathering knowledge and negotiating over the network.

This variety of topics in research and experimentation is well illustrated by the work presented in this special issue. The ERCIM people are working hard on the development of the Semantic Web, and we have organised their presentations into three sections:

- 'Tools and Experiments' gathers projects that are developing tools for manipulating and experimenting with semantic markup
- 'Metadata, Ontologies and Information Retrieval' is dedicated to projects on the design and alignment of vocabularies for use in marking up and retrieving documents and products
- 'Multimedia and Adaptation' deals with the modular design of multimedia documents, and the appeal they hold in terms of individual adaptation and retrieval as well as profiled assembly.

We are now at a moment in which the players are mobilised in order to provide the best languages, architectures and tools for experimenting with the concepts of the Semantic Web. In this special issue, you have the opportunity to read all about it.

Links:
Semanticweb.org: http://www.semanticweb.org
W3C/SW activity: http://www.w3.org/2001/sw/
Ontoweb: http://www.ontoweb.org
DAML: http://www.daml.org
The International Semantic Web Conferences:
http://iswc.semanticweb.org/

Further reading:
Tim Berners-Lee, James Hendler, Ora Lassila,
The Semantic Seb, Scientific American 284(5):35-43, 2001,
http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html

Isabel Cruz, Stefan Decker, Jérôme Euzenat, Deborah McGuinness (eds.),
The Emerging Semantic Web: Selected Papers from the First Semantic Web Working Symposium, IOS press,
Amsterdam, 2002, 300pp,
http://www.inrialpes.fr/exmo/papers/emerging/

Ian Horrocks, James Hendler (eds.),
The Semantic Web - ISWC 2002,
Lecture Notes in Computer Science 2342,
Springer-Verlag, 2002
http://link.springer.de/link/service/series/0558/tocs/t2342.htm

Dieter Fensel, James Hendler, Henry Lieberman, Wolfgang Wahlster (eds.),
Spinning the Semantic Web, The MIT press,
Cambridge, 2002, to appear

Jérôme Euzenat (ed.),
Research Challenges and Perspectives of the Semantic Web,
http://www.ercim.org/EUNSF/semweb.html

Please contact:
Jérôme Euzenat, INRIA
Tel: +33 476 61 53 66
E-mail: Jerome.Euzenat@inrialpes.fr

The Semantic Web lifts off

by Tim Berners-Lee and Eric Miller

Many researchers at ERCIM Institutes are aware that this is an exciting time to be involved in work done at the World Wide Web Consortium (W3C). Scalable Vector Graphics, Web Services, and the Semantic Web are but a few of the W3C Activities attracting media attention. This article focuses on the W3C's Semantic Web Activity and recent developments in the Semantic Web community. Although it is difficult to predict the impact of such a far-reaching technology, current implementation and signs of adoption are encouraging and developments in future research areas are extremely promising.

The Semantic Web is an extension of the current Web in which information is given well-defined meaning, enabling computers and people to work in better cooperation. The W3C Semantic Web Activity, in collaboration with a large number of researchers and industrial partners, is tasked with defining standards and technologies that allow data on the Web to be defined and linked in a way that it can be used for more effective discovery, automation, integration, and reuse across applications. The Web will reach its full potential when it becomes an environment where data can be shared and processed by automated tools as well as by people.

How might this be useful? Suppose you want to compare the price and choice of flower bulbs that grow best in your zip code, or you want to search online catalogs from different manufactures for equivalent replacement parts for a Volvo 740. The raw information that may answer these questions, may indeed be on the Web, but it's not in a machine-usable form. You still need a person to discern the meaning of the information and its relevance to your needs.

The Semantic Web addresses this problem in two ways. First, it will enable communities to expose their data so that a program doesn't have to strip the formatting, pictures and ads from a Web page to guess at the relevant bits of information. Secondly, it will allow people to write (or generate) files which explain — to a machine — the relationship between different sets of data. For example, one will be able to make a 'semantic link' between a database with a 'zip-code' column and a form with a 'zip' field that they actually mean the same thing. This will allow machines to follow links and facilitate the integration of data from many different sources.

This notion of being able to semantically link various resources (documents, images, people, concepts, etc) is an important one. With this we can begin to move from the current Web of simple hyperlinks to a more expressive semantically rich Web, a Web where we can incrementally add meaning and express a whole new set of relationships (hasLocation, worksFor, isAuthorOf, hasSubjectOf, dependsOn, etc) among resources, making explicit the particular contextual relationships that are implicit in the current Web. This will open new doors for effective information integration, management and automated services.

How is it being developed?

There are two places to look for Semantic Web progress: from the ground up, in the infrastructural and architectural work coordinated by the W3C, and from top down, in application-specific work by those leveraging Semantic Web technologies in various demonstrations, applications and products. This article provides an introduction to both views with a specific focus on those areas in which the W3C is directly involved.

Enabling Standards

Uniform Resource Identifiers (URIs) are a fundamental component of the current Web and are a foundation of the Semantic Web. The Extensible Markup Language (XML) is also a fundamental component for supporting the Semantic Web. XML provides an interoperable syntactical foundation upon which the more important issue of representing relationships and meaning can be built. URIs provide the ability for uniquely identifying resources as well as relationships among resources. The Resource Description Framework (RDF) family of standards leverages URIs and XML to provide an stepwise set of functionality to represent these relationships and meaning.

The W3C Semantic Web Activity's charter is to serve a leadership role in the design of specifications and the open, collaborative development of technologies focused on representing relationships and meaning. The base level of the RDF family of standards is a W3C Recommendation. The RDF Core Working Group is in the process of formalizing the original RDF Model and Syntax Recommendation which provides a simple yet powerful framework for representing information in the Web. Building on this work, the group is additionally defining a simple means for declaring RDF Vocabularies. RDF Vocabularies are descriptive terms (eg Service, Book, Image, title, description, rights, etc) that are useful to communities recoding information in a way that enables effective reuse, integration and aggregation of data. Additional deliverables include a precise semantic theory of these standards that will support future work, as well as a primer designed to provide the reader with a basic understanding of RDF and its application.

Simple data integration, aggregation and interoperability are enabled by these base level RDF standards. An increasing need for interoperability at a more expressive descriptive level is also desired. The Web Ontology Working

Group is charted to build upon the RDF Core work a language for defining structured, Web-based ontologies. Ontologies can be used by automated tools to power advanced services such as more accurate Web search, intelligent software agents and knowledge management. Web portals, corporate website management, intelligent agents and ubiquitous computing are just some of the identified scenarios that helped shaped the requirements for this work.

Advanced Development

Just as the early development of the Web depended on code modules such as libwww, W3C is devoting resources to the creation and distribution of similar core components that will form the basis for the Semantic Web. The W3C Semantic Web Advanced Development (SWAD) initiatives are designed to work in collaboration with a large number of researchers and industrial partners and stimulate complementary areas of development that will help facilitate the deployment and and future standards work associated with the Semantic Web.

SWAD DAML

The purpose of the SWAD DAML project is to contribute to the development of a vibrant, ubiquitous Semantic Web by building critical Semantic Web infrastructure and demonstrating how that infrastructure can be used by working, user-oriented applications.

SWAD DAML is designed to build on the DARPA Agent Markup Language (DAML) infrastructure to provide an interchange between two or more different applications. The first involves structured information manipulation required to maintain the ongoing activities of an organization such as the W3C. These include access control, collaborative development, and meeting management. The second application is focused on the informal and often heuristic processes involved in document management in a personalized information environment. Integrated into both environments will be tools to enable authors to control terms under which personal or sensitive information is used by others, a critical feature to encourage sharing of semantic content.

SWAD-Europe

SWAD-Europe will highlight practical examples of where real value can be added to the Web through Semantic Web technologies. The focus on this initiative is on providing practical demonstrations of how the Semantic Web can address problems in areas such as: site maps, news channel syndication, thesauri, classification, topic maps, calendaring, scheduling, collaboration, annotations, quality ratings, shared bookmarks, Dublin Core for simple resource discovery, Web service description and discovery, trust and rights management and how to effectively and efficiently integrate these technologies together.

SWAD-Europe will additionally concentrate on exploratory implementation and pre-consensus design in areas such as querying, and the integration of multiple Semantic Web technologies. It shall provide valuable input and experiences to future standards work.

SWAD Simile

W3C is additionally working with HP, MIT Libraries, and MIT's Lab for Computer Science on Simile, which seeks to enhance interoperability among digital assets, schemas, metadata, and services across distributed individual, community, and institutional stores and across value chains that provide useful end-user services by drawing upon the assets, schemas, and metadata held in such stores. Simile will leverage and extend DSpace, enhancing its support for arbitrary schemas and metadata, primarily though the application of RDF and Semantic Web techniques. The project also aims to implement a digital asset dissemination architecture based upon Web standards, enabling services to operate upon relevant assets, schemas, and metadata within distributed stores.

The Simile effort will be grounded by focusing on well-defined, real-world use cases in the libraries' domain. Since parallel work is underway to deploy DSpace at a number of leading research libraries, we hope that such an approach will lead to a powerful deployment channel through which the utility and readiness of Semantic Web tools and techniques can be compellingly demonstrated in a visible and global community.

SWAD Oxygen

The MIT/LCS Oxygen project is designed to enable pervasive, human-centered computing through a combination of specific user and system technologies. Oxygen's user technologies directly address human needs. Speech and vision technologies enable us to communicate with Oxygen as if we're interacting with another person, saving much time and effort. Automation, individualized knowledge access, and collaboration technologies help us perform a wide variety of tasks in the way we like to do them. In Oxygen, these technologies enable the formation of spontaneous collaborative regions that provide support for recording, archiving, and linking fragments of meeting records to issues, summaries, keywords, and annotations.

The Semantic Web is designed to foster similar collaborative environment and the W3C is working with project Oxygen to help support this goal. The ability for "anyone to say anything about anything" is an important characteristic of the current Web and is a fundamental principle of the Semantic Web. Knowing who is making these assertions is increasingly important in trusting these descriptions and enabling a 'Web of Trust'. The Annotea advanced development project provides the basis for asserting descriptive information, comments, notes, reviews, explanations, or other types of external remarks to any resource. Together with XML digital signatures, the Annotea project will provide a test-bed for 'Web-of-Trust' Semantic Web applications.

Applications - spinning upward

Though not the focus of this article, the deployment of RDF-based technologies is increasingly significant. The W3C Semantic Web Activity hosts the RDF Interest Group, which coordinates public implementation and shares deployment experiences of these technologies. Arising out of RDF Interest Group discussions are several public issue-specific mailing lists, including RDF-based calendar and group scheduling systems, logic-based languages, queries and rules for RDF data and distributed annotation and collaboration systems. These discussion groups are designed to

focus on complementary areas of interest associated with the Semantic Web Activity, each of which fosters cooperation and collaboration among individuals and organizations working on related Semantic Web technologies.

In addition to these Interest Group lists there are a variety of domain specific communities who are using RDF/XML to publish their data on the Web. These notably include the Dublin Core Metadata Initiative, the IMS Global Learning Consortium vocabularies for facilitating online distributed learning, XMLnews, PRISM, the RDF Site Summary (RSS 1.0) for supporting news syndication, Musicbrainz for cataloging and cross-referencing music, and Creative Commons for supporting a digital rights description to name but a few. The Topic Map (XTM) community has been finding increasing synergy with the RDF data model.

Early commercial adopters such as Adobe's eXtensible Metadata Platform (XMP), for example, leverage RDF/XML to enable more effective management of digital resources. Adobe applications and workflow partners through XMP can leverage the power of RDF/XML to provide a standardized means for supporting the creation, processing, and interchange of document metadata across publishing workflows. This in-turn reduces cost and makes for more effective management of digital resources possible both within and across organizational boundaries.

New Things opening up

The most exciting thing about the Semantic Web is not what we can imagine doing with it, but what we can't yet imagine it will do. Just as global indexes, and Google's algorithms were not dreamed of in the early Web days, we cannot imagine now all the new research challenges and exciting product areas which will appear once there is a Web of data to explore. Many existing fields for knowledge representation and data management have typically made assumptions regarding a conceptually or physically centralized system, and as such their application to the Semantic Web is not straightforward. Given a mass of rules relating data in different vocabularies, and an unbounded set of datasets in different vocabularies, what algorithms will efficiently resolve general queries? What conventions for the storage of tips and pointers will allow data to be reused and converted automatically? What techniques will allow a system to operate securely while processing very diverse data from untrusted agents? How can one represent — and then implement — personal privacy in such a world?

The Semantic Web starts as a simple circles-and-arrows diagram relating things, which slowly expands and coalesces to become global and vast. The Web of human-readable documents spawned a social revolution. The Semantic Web may in turn spawn a revolution in computing. In neither case did a change occur in the power of one person or one computer, but rather a dramatic change in the role they can play in the world, by being able to find out almost anything virtually immediately.

For more information on the Semantic Web, including additional projects, products, efforts and future directions check out the Semantic Web home page.

Link:
http://www.w3c.org/2001/sw/

Please contact:
Tim Berners-Lee, Director of W3C
E-mail: timbl@w3.org

Eric Miller
W3C Semantic Web Activity Lead
E-mail: em@w3.org

Semantic Web Technologies in Europe's IST Programme 1998 - 2002

by Hans-Georg Stork

The European Commission supports a broad range of research and development activities related to the Semantic Web. This support is expected to continue in the forthcoming 6th Framework Programme.

The European Commission's Information Society Technologies (IST) Programme was designed in 1997-1998 as part of the 5th Framework Programme for R&D in Europe, covering the period 1998 - 2002. One of its Key Action III ('Multimedia Content and Tools') modules was entitled 'Information access, filtering, analysis and handling' (IAF for short). Its objectives were to support the development of "...advanced technologies for the management of information content to empower the user to select, receive and manipulate ... only the information required when faced with an ever increasing range of heterogeneous sources." These technologies should lead to "... improvements in the key functionalities of large-scale multimedia asset management systems (including the evolution of the World Wide Web) to support the cost-effective delivery of information services and their usage." (Quoted from http://www.cordis.lu/ist/b-oj-en5.htm#KA3).

Only a few months before the first IST Call was launched in early 1999, Tim Berners-Lee had published his informal 'Semantic Web Road Map' note on the

W3C website, outlining possible future directions for the evolution of the World Wide Web. These ideas, partly based on previous content and resource description activities, had met with growing enthusiasm of researchers and developers worldwide, both in academia and in industry. They provided a new vision for integrating efforts that had been ongoing for some time in many R&D communities, involving specialists in various computer science disciplines. These efforts were aimed at capturing the semantics of digital content of all sorts, and at devising ways of acting sensibly upon the formal knowledge representations thus gained.

Picking up on this vision the European Commission dedicated a specific action line of its IST Work Programme 2001 to 'Semantic Web Technologies' (Action Line III.4.1), thereby underlining the importance (in terms of research challenges and expected impact) of 'semantics issues' for achieving the declared goals of the IAF module of IST.

It offered four broad interrelated R&D tracks as an orientation for submitting project proposals:

- creating a usable formal framework in terms of formal methods, models, languages and corresponding tools for semantically sound machine-processable resource description
- fleshing out the formal skeletons by developing and applying techniques for knowledge discovery (in databases and text repositories), ontology learning, multimedia content analysis, content-based indexing, etc
- acting in a semantically rich environment, performing resource and service discovery, complex transactions, semantic search and retrieval, filtering and profiling, supporting collaborative filtering and knowledge sharing, etc
- making it understandable to people through device-dependent information visualisation, semantics-based and context-sensitive navigation and browsing, semantics-based dialogue management, etc.

This agenda provided some continuity with respect to previous Key Action III activities (notably on 'media representation and access' and digital libraries) and activities supported by other IST departments (for instance under action line 'Methods and tools for intelligence and knowledge sharing' of Key Action IV - Essential Technologies and Infrastructures, or under the Open Domain of FET — Future and Emerging Technologies).

But it also provided a sharper focus on the problems of creating and using knowledge representations, in the context of large-scale distributed systems, such as the World Wide Web.

This focus was largely retained in Work Programme 2002 as part of Key Action III's 'Preparing for future research activities' action line (AL III.5.2). Moreover, Work Programme 2002, in one of its 'Cross Programme Activities (CPA)', took account of a new trend that has surfaced over the last couple of years:

Some IST Projects on Semantics in Distributed Systems ('Semantic Web Technologies')

COG
Corporate Ontology Grid
http://www.cogproject.org/

ESPERONTO
Application Service Provision of Semantic Annotation, Aggregation, Indexing and Routing of Textual, Multimedia, and Multilingual Web Content
http://www.isoco.com/isococom/whitepapers/files/ SemanticWeb-whitepaper-137.pdf

FF-POIROT
Financial Fraud Prevention-Oriented Information Resources using Ontology Technology
http://www.starlab.vub.ac.be/research/projects/default.htm#Poirot

GRACE
Grid Search and Categorisation Engine
IST-2001-38100

IBROW
An Intelligent Brokering Service for Knowledge-Component Reuse on the World Wide Web
http://www.swi.psy.uva.nl/projects/ibrow/

InDiCo
Integrated Digital Conferencing
IST-2001- 34306

MONET
Mathematics on the Net
http://monet.nag.co.uk/

MOSES
A modular and Scalable Environment for the Semantic Web
IST-2001-37244

On-To-Knowledge
Content-Driven Knowledge Management through Evolving Ontologies
http://www.ontoknowledge.com

ONTO-LOGGING
Corporate Ontology Modelling and Management System
http://www.ontologging.com

SCULPTEUR
Semantic and Content-Based Multimedia Exploitation for European Benefit
IST-2001-35372

SEWASIE
Semantic Webs and Agents in Integrated Economies
http://www.sewasie.org/

SPACEMANTIX
Combining Spatial and Semantic Information in Product Data
IST-2001-34159

SPIRIT
Spatially-Aware Information Retrieval on the Internet
http://www.cs.cf.ac.uk/department/posts/SPIRITSummary.pdf

SWAD-Europe
W3C Semantic Web Advanced Development for Europe
http://www.w3.org/2001/sw/Europe/

SWAP
Semantic Web and Peer-to-Peer
http://swap.semanticweb.org/

SWWS
Semantic-Web-Enabled Web Services
http://www.cs.vu.nl/~swws/

VICODI
Visual Contextualisation of Digital Content
IST-2001-37534

WIDE
Semantic-Web-Based Information Management and Knowledge-Sharing for Innovative Product Design and Engineering
http://www.cefriel.it/topics/research/default.xml?id=75&_euro=1&old=1&dc=1&tid=27

WISPER
Worldwide Intelligent Semantic Patent Extraction & Retrieval
IST-2001-34407

WonderWeb
Ontology Infrastructure for the Semantic Web
http://wonderweb.semanticweb.org

the application of Grid technologies to "knowledge discovery in ... large distributed datasets, using cognitive techniques, data mining, machine learning, ontology engineering, information visualisation, intelligent agents... (quoted from WP2002, CPA9)", all more or less directly pertinent to the Semantic Web vision.

Calls for submission of proposals to these action lines were published in July (AL III.4.1) and November (AL III.5.2 & CPA9) 2001, respectively (Calls 7 and 8). Both calls met with considerable interest in relevant R&D communities across Europe and drew altogether nearly one hundred submissions involving several hundred participating organisations. They resulted in a significant growth (by 17 projects) of a portfolio of projects that are all poised to contribute in one way or another, to making the Semantic Web a reality (see the references in the box, including a few earlier and concurrent projects in Key Actions II, IV and in FET).

While at the time of writing the new projects have only just commenced, some of the older ones have already produced noteworthy results. It may suffice to mention project On-To-Knowledge (http://www.ontoknowledge.com), which has been one of the birthing grounds of OWL, the proposed new Web ontology language, currently under discussion at the W3C.

Moreover, in recognition of the central role ontologies are likely to play in building the 'Semantic Web', the European Commission, through its IST Programme, supports the 'Thematic Network' OntoWeb, a platform for fostering collaboration between industry and academia, on creating a 'semantic infrastructure' for applications in many different areas (e-business, Web services, multimedia asset management, community webs, etc). Through OntoWeb, European researchers and practitioners also have an opportunity to make more targeted contributions to international standardisation activities and to the W3C process.

It is expected that EU support of Semantic-Web-related R&D will continue under the forthcoming 6th Framework Programme within the broader context of 'Knowledge Technologies', as part of the 'Priority Thematic Area' IST - Information Society Technologies.

The views expressed in this article are those of the author and do not necessarily engage the European Commission.

Links:
http://www.cordis.lu/fp6/
http://www.cordis.lu/ist/fp6/workshops.htm
http://www.cordis.lu/ist/ka3/iaf/iaf_workshop.htm

Please contact:
Hans-Georg Stork
European Commission,
DG Information Society
E-mail: Hans-Georg.Stork@cec.eu.int

Activities of INTAP Semantic Web Committee in Japan

by Tomihiko Kojima

The Interoperability Technology Association for Information Processing (INTAP) drives forward researches and developments related to the Semantic Web and its promotion in Japan.

INTAP is an association authorised by the Ministry of Economy, Trade and Industry (or METI) of Japan, established in 1985. The Association conducts R&D, surveys, and publicity activities for promoting interoperability of information and communication technologies. The Association stimulates the progress of information processing and related technologies, which should contribute to the creation of a sound and healthy information society.

INTAP Semantic Web Committee

INTAP has established the Semantic Web Committee, based on recognition that the Semantic Web is a critical technology that will affect widely economic activities and social life. INTAP drives forward research and development related to the Semantic Web and its promotion in Japan through the Committee, in cooperation with Keio University and the World Wide Web Consortium (W3C).

A brief history of the Semantic Web Committee:

- established in 2001
- Semantic Web Conference (October 2001, at Keio University)
- Semantic Web Workshop (December 2001, at DBWeb2001 in Kyoto)
- presentation in panel session by the committee chairman, Noboru Shimizu at ISWC2002 (June 2002)
- Semantic Web special report of the technical magazine issued by the Information Processing Society of Japan (IPSJ) in July 2002: What is the Semantic Web?; Meta-Data on the Semantic Web and its Usage; Web Ontology language for the Semantic Web; Tools for the Semantic Web; Application Systems on the Semantic Web
- Semantic Web Conference 2002 in 18th September 2002.

Members of the Committee

The committee is chaired by Mr. Noboru Shimizu of NEC Corporation and technically led by Dr. Nobuo Saito, Vice President of Keio University and Dr.

Tatsuya Hagino, Professor of the same University. Dr. Saito is working for W3C as the Director and Deputy Director of W3C. The committee members are comprised of NEC Corporation, Hitachi Ltd, Mitsubishi Electric Corporation, Oki Electric Industry Co Ltd, Sumitomo Electric Industries, Ltd, Toshiba Corporation, IBM Japan Ltd, Matsushita Electric Industrial Co Ltd, Fujitsu Ltd, Fujitsu Laboratories Ltd, NEC Software Kyushu Ltd, NEC Planning Research Ltd, and Nippon Telegraph and Telephone Corporation.

Activities of the Committee

- Survey on trends on Semantic Web technology in Europe and the US: RDF, OWL, DAML, DAML-S, RSS, Dublin Core, etc.
- Research on the successful story of Semantic Web applications, such as next-generation e-government.

Blueprints

- *Development of Japanese Content Description System using RDF:* We will develop a Japanese Content Description System using RDF that can automatically generate RDF metadata on a certain Japanese content based on pattern formats which are used to identify particular letter strings as metadata for that content. We will also develop a Japanese RDF Metadata Standard to describe summaries of various Japanese contents.
- *Development of Integrated Management System of Distributed Information (License Repository):* Based on the assumption that the Internet is a huge repository, we will develop an Integrated Management System of Distributed Information that enables attached documents of a written application to be held on an applicant-side system and enables the staff of administrative organs who receive the application to browse the attached documents when they want to see them.
- *Research and Development of Intelligent Search System using RDF and Ontology:* We will research RDF, ontology description language including OWL and intelligent agent technology and we will develop an Intelligent Search System using RDF and Ontology based on the research.

Link:
http://www.net.intap.or.jp/INTAP/s-web/

Please contact:
Tomihiko Kojima, INTAP, Japan
Tel: +81 3 5977 1301
E-mail: info-semanticweb@intap.or.jp, kojima@intap.or.jp

OntoWeb: The Thematic Network for the Semantic Web

by Ying Ding and Dieter Fensel

OntoWeb: Ontology-based information exchange for knowledge management and electronic commerce is an EU-funded three-year IST Thematic Network project within the 5th Framework Programme.

The goal of the OntoWeb is to bring researchers and industry together to enable ontologies to improve information exchange and provide better services. Meanwhile, OntoWeb also aims to strengthen the European influence on standardisation efforts in areas such as Web languages (XML, RDF, OWL), upper-layer ontologies, and content standards.

OntoWeb: The Hub for a Semantic Web

The World Wide Web is currently moving from the first generation Web to the second generation Web — the Semantic Web. Tim Berners-Lee coined the vision of a Semantic Web that provides more automated services based on machine-processable semantics of data and heuristics that make use of these metadata.

Ontologies, which provide shared and common domain theories, are a key asset for such a Semantic Web. They can be seen as metadata that explicitly represent semantics of data in a machine processable way. Ontology-based reasoning services can operationalise this semantics for providing various services. Ontologies therefore play a crucial role in enabling content-based access, interoperability, and communication across the Web, providing it with a qualitatively new level of service — the Semantic Web. It weaves together a net linking incredibly large parts of the human knowledge and complements it with machine processability.

Currently, there exist five Special Interest Groups (SIGs) within OntoWeb:

SIG1 on content standards: SIG1 aims to coordinate cooperation and participation in current initiatives related to ontology-based content standardization and content harmonization across different standards. For more details, please visit: http://www.ladseb.pd.cnr.it/infor/ontology/OntoWeb/SIGContentStandards.html

SIG2 on ontology language standards: SIG2 focuses on ontology language standards through coordinating cooperation, communication, and participation in related initiatives and relevant standardization efforts. For more details, please visit: http://www.cs.man.ac.uk/~horrocks/OntoWeb/SIG/

SIG3 on enterprise-standard ontology environments: The mission of this SIG is to support a dialog to develop quality requirements for enterprise-standard ontology environments and steer future

ontology development towards addressing these requirements. For more details, please visit: http://delicias.dia.fi.upm.es/ontoweb/sig-tools/index.html

SIG4 on industrial applications: SIG4 offers an international forum for the exchange of ideas, practical experiences, and advances in research into industry and business applications of ontology and Semantic Web technology, especially the Web services powered by Semantic Web technology. For more details, please visit its homepage: http://www.cs.vu.nl/~maksym/sig4/sig4.html

SIG5 on language technology in ontology development and use: The Motivations of SIG5 is to provide the connections between language technology and ontologies. Being the platform for cooperation and standards, SIG5 conveys the multidisciplinary expertise between natural language technology and Semantic Web. For more details, please visit: http://dfki.de/~paulb/ontoweb-lt.html

OntoWeb is the European focal point for bringing together activities in the area of ontology-based methods and tools for the Semantic Web. It will significantly secure a competitive position in the emerging business field by:

- envisioning and describing the current state-of-the-art in ontology-related fields in Europe and world-wide, providing a strategic guide to industrial and commercial applications
- co-operating with different standardisation bodies to promote the development of ontology-based content standardisation and ontology language standardisation
- providing non-discriminatory access to services by individuals and businesses and aiding identification and extraction of information in a rapidly growing network.
- organising dissemination workshops, special interest groups, a scientific journal, and training or education courses with special emphasis on Semantic-Web-Based applications, e-commerce, knowledge management and information integration. It addresses aspects of knowledge transfer between academia and industry for the emerging ontology-based technologies, enabling competitive advantages in these fast developing markets.

OntoWeb currently has about 120 members from Europe, America, and Asia. The ratio of academic and commercial members is roughly 2:1. We invite companies, research institutes, and universities active in the Semantic Web area to join the OntoWeb project. We offer the opportunity for project members to:

- share the enriched resources of the latest ongoing research, cutting-edge knowledge, and innovative technologies through the OntoWeb portal
- attend various OntoWeb conferences, workshops, and seminars to meet possible future collaborative partners or potential customers and to exploit new academic or business opportunities
- join the OntoWeb Special Interest Groups (SIGs).

To become a member, please visit http://www.ontoweb.org/member.htm
Let us work jointly to bring the Semantic Web to its full potential.

Link:
http://www.ontoweb.org/

Please contact:
Ying Ding, Dieter Fensel
Vrije Universiteit Amsterdam,
The Netherlands
E-mail: {ying, dieter}@cs.vu.nl

Semantic Web Advanced Development in Europe

by Brian Matthews, Michael Wilson and Dan Brickley

The Semantic Web Activity of the World Wide Web Consortium (W3C) aims to augment the existing Web architecture by adding communicable information about resources which can then be processed by automated agents. This initiative has attracted much interest in the last few years, and led to the creation of a European project — SWAD-Europe or Semantic Web Advanced Development in Europe which aims to highlight practical examples of where real value can be added to the Web through the Semantic Web.

The initial inspiration for the SWAD-Europe project was a response to questions which frequently arise when considering Semantic Web technology, such as:

- How do I use RDF with XML Schemas?
- Web Services with Web Ontologies?
- MathML with RDF-rules?

These technology-oriented questions often mask an overall application-oriented need. It is not clear which, if any, of these technologies are most appropriate to use, nor what the relationship between them is. A common theme recurs: technology integration. (Semantic) Web technology has a daunting array of tools, specifications and techniques — there is too much to choose from!

SWAD-Europe aims to address these concerns by tackling simple things first, to provide real examples of how the Semantic Web can address problems in areas such as: sitemaps, news channel syndication ; thesauri, classification, topic maps ; calendaring, scheduling, collaboration ; annotations, quality ratings, shared bookmarks ; Dublin Core for simple resource discovery; Web service description and discovery; trust

and rights management ; and how to mix all these together. Our general approach is not to focus on the AI-related technology from the higher levels of the Semantic Web architecture, but rather to exploit what has been developed so far to have it used in practice, thereby answering some of the queries we hear. SWAD-Europe will concentrate on exploratory implementation and design in pre-consensus areas such as querying, and the integration of multiple Semantic Web technologies. In this way, we intend to show the utility of the Semantic Web to the European developer community.

An outreach programme is a key component of the project which will be approached via showcasing Semantic Web for industry applications, guidance on how to integrate Semantic Web, XML and other Web standards, and facilitating technology maturation, eg Semantic-Web-enhanced Web Services. Further, the Semantic Web will be encouraged and deepened for all parts of the industrial and academic communities within Europe.

The project will develop scenarios for demonstrating the practical utility of Semantic Web technology. We would like to give developers both a 'big picture' understanding of the Semantic Web and a detailed knowledge of the tools and techniques and thus provide feedback for standardisation work. The project will be looking at demonstrator projects and developments in a variety of technical areas:

- *XML / Semantic Web Integration:* Demonstrating that SW technologies are compatible with the broader family of XML languages (especially SVG and MathML).
- *Web Services Integration:* Convergence between RDF and SOAP serialisations.
- *Scalability, APIs and Query Languages:* Advances in technical knowledge with the rationalisation and documentation of deployed best practice for APIs and simple query languages (eg SquishQL), including testing on large triple bases.
- *Trust, Annotations and Information Quality:* Specific applications to augment the Web Services with tools from the Semantic Web.
- *Thesauri and Knowledge Management:* Simple classification tools to aid knowledge management, with an emphasis on multi-linguality in a European context.

We intend to make technical innovations in all these areas, and combine them to provide a practical demonstration of the utility of the Semantic Web.

SWAD-Europe is a collaboration between the World Wide Web Consortium in Europe, hosted by INRIA, research institutions at the Central Laboratory of the Research Councils (CLRC) and the University of Bristol (ILRT), and the companies Stilo International Plc and Hewlett-Packard Ltd.

SWAD-Europe is an EU-funded project, part of the Information Society Technologies (IST-7) programme. The first deliverables are already publically available at http://www.w3.org/2001/sw/Europe/reports/intro.html).

Links:
http://www.w3.org/2001/sw/Europe/
http://www.w3.org/2001/sw/

Please contact:
Michael Wilson, CLRC
E-mail: M.D.Wilson@rl.ac.uk

Dan Brickley, W3C
E-mail:danbri@w3.org

Calendars, Schedules and the Semantic Web

by Terry R. Payne and Libby Miller

The emergence of the Semantic Web has simplified and improved knowledge reuse across the Internet. By committing to various ontologies, an agent can now understand and reason about published information such as calendar events and schedules to meet its user's needs and provide assistance. We illustrate the benefits of garnering schedules from the Semantic Web for agent-based assistance, and introduce other initiatives being pursued by the RDF Calendar Taskforce.

Scheduling meetings can be tedious and time consuming, and involves collating information from multiple sources — colleagues, family, friends, Web pages and organisations — about events which may or may not happen in multiple locations and timezones, an arbitrary period ahead. Many scheduling applications require up-to-date knowledge of each user's schedule in order to automate the process of finding mutually agreeable appointments for several attendees. However, the methods used to publish Web-based schedules (eg train timetables or conference schedules) renders them opaque to machine comprehension, thus forcing users to enter these events into their calendars manually.

The Semantic Web supports the creation of Web pages containing structured, meaningful knowledge. These pages can be parsed to elicit relevant information, and through reference to published ontologies, agents can understand and reason (using DAML/OWL-based description logics) about this knowledge to provide user assistance. Thus, by defining schedules on the Semantic Web, agents can reason about not only when events occur, but where they occur and which individuals are attending them.

Access to users' calendars, including their location and availability, is pivotal for providing assistance in many domains. For this reason, various ontologies have been defined for representing schedules, and several agents and applications have been developed to explore how this resultant knowledge can be utilised. Whilst most calendaring applications use the iCalendar format (RFC 2446), a self-contained, text-based representation suitable for sharing and presenting event information, the lack of semantics limits the level of reasoning that can be performed. For example, using iCalendar, Jim's agent cannot reason about whether or not Jim's friend 'Bob' is the same 'Bob' who is presenting at a conference. Likewise, it cannot determine if the conference is close to the user's current location. Thus, it is unable to make decisions that reflect Jim's preferences, such as scheduling lunch with Bob at a nearby Thai restaurant.

```
<foaf:Person rdf:ID="terrypayne">
   <foaf:name>Terry Payne</foaf:name>
   <foaf:mbox rdf:resource="mailto:terry@acm.org"/>
   <foaa:RCalendarAgentName>terry_acm.org-
CalAgent</foaa:RCalendarAgentName>
</foaf:Person>

<ical:VCALENDAR rdf:id="TAC01">
   <dc:title>Trading Agent Competition 2001 Workshop</dc:title>
   <dc:contributor rdf:resource="#terrypayne"/>
   <dc:date>2001-10-03</dc:date>

   <ical:VEVENT-PROP
rdf:resource="http://www.tac.org/2001event.rdf#PainInNEC"/>

   <ical:VEVENT-PROP>
      <ical:VEVENT rdf:id="RetsinaTrading">
         <ical:DTSTART>
            <ical:DATE-TIME><rdf:value>20011014T134500</rdf:value></ical:DATE-TIME>
         </ical:DTSTART>
         <ical:DTEND>
            <ical:DATE-TIME><rdf:value>20011014T140000</rdf:value></ical:DATE-TIME>
         </ical:DTEND>
         <ical:LOCATION rdf:resource="#HRTampa" />
         <ical:ATTENDEE rdf:resource="http://www.daml.ri.cmu.edu/people.rdf#yn" />
        <ical:DESCRIPTION>Presentation: Retsina</ical:DESCRIPTION>
      </ical:VEVENT>
</ical:VCALENDAR>
```

Figure 1: A Hybrid RDF schedule containing two events.

As part of the RDF Calendar Taskforce, we have developed several tools and vocabularies to address this problem. The development of the hybrid RDF calendar ontology was based on Tim Berners-Lee's analysis of an RDF adaptation of iCalendar. Several conference schedules have subsequently been published using this ontology (see Figure 1), and can be queried and viewed using simple RDF query tools. As RDF supports the creation of concepts (such as contact details or locations) as distributed resources indexed by a URI, schedules can reuse existing concepts defined elsewhere. Thus, agents can reason about the resulting schedules (coupled with the ontology definitions and associated axioms) and hence perform tasks such as the above lunch-scheduling task.

The RCAL agent exploits the structure and semantics of schedules defined using the hybrid RDF calendar ontology by providing browsing, importing, reasoning and distributed meeting scheduling functionalities. Users can import schedules into Microsoft's Outlook 2000(TM), and compare them with existing appointments to identify conflicts. Meetings can be automatically scheduled between several users, using the Contract Net Protocol to locate mutually agreeable times. Notifications associated with events can be sent to mobile devices to alert users when they are about to start. Imported events are also monitored, and the user is notified if these events change.

Other tools and ontologies have emerged from the Calendar Taskforce. A vocabulary similar to that used by Palm PDAs has been developed in tandem with tools that can synchronise these devices with RDF data. Conversion tools that map between the iCalendar and xCalendar formats (an XML format for iCalendar) have been developed and adapted to generate the modified Palm format from the Mozilla calendar. Translation Services that convert schedules defined by other ontologies (eg DARPA Meeting Agenda Ontology) into hybrid RDF calendar markup have been developed; these Web services are located and invoked automatically when RCal encounters new markup. Members of the EC-funded SWAD-Europe project are experimenting with RSS 1.0 and its events module, a simple subset of iCalendar in RDF. For planning purposes we require information about deliverables, meetings, and related events (eg conferences) to be used in scheduling. We use various tools (eg the XHTML-RSS converter) to store, transform, merge and view event information to simplify planning. Our aim is to demonstrate the utility of these tools and the Semantic Web by 'eating our own dogfood', ie using the very tools our work is promoting.

Once people make their events information available in RDF, we can not only show why it is now useful to do this but we can also use that data to create more advanced research scheduling applications. In this sense, calendaring provides a partial answer to the question — how do we get from here (HTML web) to there (Semantic Web)?

Link:
http://ilrt.org/discovery/2001/04/calendar/

Please contact:
Terry R. Payne
University of Southampton
Tel: +44 23 8059 3255
E-mail: terry@acm.org

Libby Miller, University of Bristol
Tel: +44 117 928 7113
E-mail: Libby.Miller@bristol.ac.uk

The ICS-FORTH RDFSuite: High-level Scalable Tools for the Semantic Web

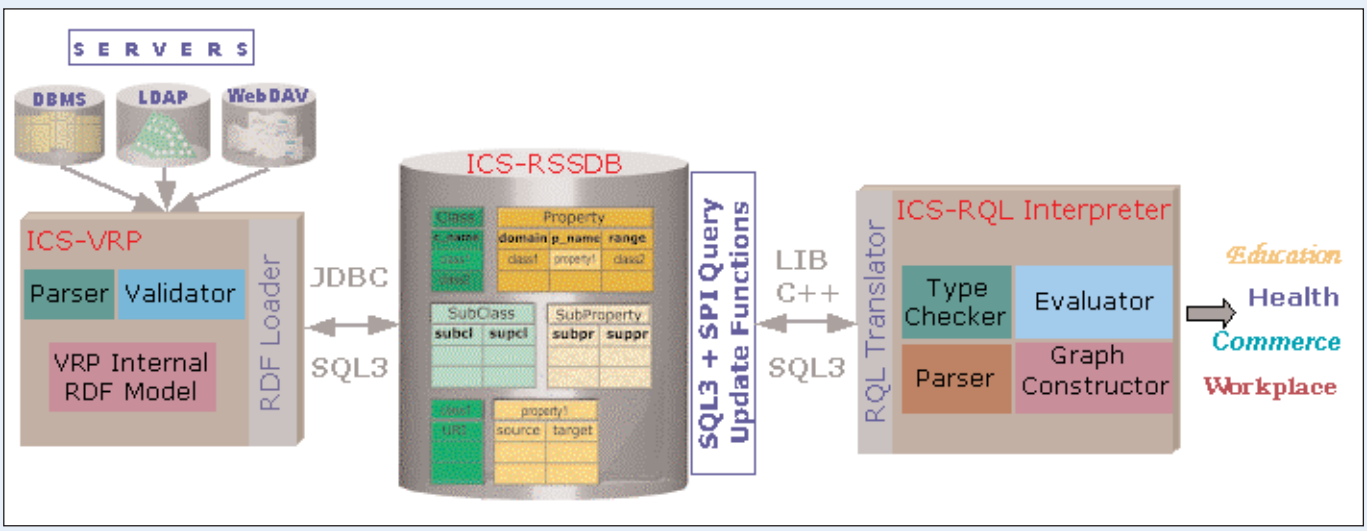
by Sofia Alexaki, Nikos Athanasis, Vassilis Christophides, Greg Karvounarakis, Aimilia Maganaraki, Dimitris Plexousakis and Karsten Tolle

RDFSuite addresses the need for effective and efficient management of large volumes of RDF metadata as required by real-scale Semantic Web applications.

In the next evolutionary step of the Web, termed the Semantic Web, a vast number of information resources will be made available along with various kinds of descriptive information, ie metadata. The Resource Description Framework (RDF) enables the creation and exchange of resource metadata as normal Web data. To interpret these metadata within or across user communities, RDF permits the definition of appropriate schema vocabularies (RDFS). Managing voluminous RDF description bases and schemas with existing low-level APIs and file-based implementations does not ensure fast deployment and easy maintenance of real-scale Semantic Web applications (eg Knowledge Portals and E-Marketplaces). Still, we would like to benefit from database technology in order to support declarative access and logical and physical RDF data independence. This is the main design choice of RDFSuite, which comprises the Validating RDF Parser (VRP), the Schema-Specific Data Base (RSSDB) and an interpreter for the RDF Query Language (RQL). It was developed in the context of the IST projects C-Web and MesMuses, both coordinated by INRIA. (http://cweb.inria.fr). The Figure depicts the architecture of RDFSuite, which is available for download under an Open Source Software License from http://139.91.183.30:9090/RDF/.

The Validating RDF Parser (VRP) is a tool for analysing, validating and processing RDF schemas and resource descriptions. The Parser syntactically analyses the statements of a given RDF/XML file according to the RDF Model & Syntax Specification. The Validator checks whether the statements contained in RDF schemas and resource descriptions satisfy the semantic constraints derived by the RDF Schema Specification (RDFS). Unlike other RDF parsers, VRP is based on standard compiler generator tools, namely CUP and Jflex, which ensure good performance when processing large volumes of RDF descriptions. VRP has been successfully used with RDF schemas from existing Semantic Web applications, hosted by the Schema Registry at http://139.91.183.30:9090/RDF/Examples.html. VRP supports embedded RDF in XML or HTML documents, XML Schema Data Types and Unicode character set. Users can activate or deactivate semantic constraints against which validation is performed and configure the parser according to their needs. Another substantial feature is the fetching of remote namespaces and their integration in VRP's internal RDF model. VRP provides various options for debugging, serialisation under the form of triples or graphs and complete statistics of validated schemas and resource descriptions.

The RDF Schema-Specific Data Base (RSSDB) is a persistent RDF Store for loading resource descriptions in an object-relational DBMS by exploiting the available RDF schema knowledge. It preserves the flexibility of RDF in refining schemas and enriching descriptions at any time, and can store resource descriptions created according to one or more associated RDF schemas. Its main design goals are the separation of the RDF schema from data information and the distinction between unary and binary relations holding the instances of classes



RDFSuite architecture.

and properties respectively. The experiments we have carried out illustrate that RSSDB yields considerable performance gains in query processing and storage volumes as compared to triple-based RDF Stores.

RSSDB has been implemented on top of the PostgreSql ORDBMS. It comprises Loading and Update modules, both implemented in Java. Access to the ORDBMS relies on the JDBC interface. Its most distinctive feature is the customisation of the database representation according to the employed meta-schemas (RDF/S, DAML-OIL), the peculiarities of RDF schemas and description bases and the target query functionality. The RSSDB Loader supports incremental loading of distributed namespaces by automatically detecting changes in already stored RDF schemas or data.

RQL is a typed language, following a functional approach. RQL relies on a formal graph model that captures the RDF modeling primitives and permits the interpretation of superimposed resource descriptions by means of one or more schemas in various application contexts (recommendations, content ratings, push channels, etc). The novelty of RQL lies in its ability to seamlessly combine schema and data querying.

The RQL Interpreter has been implemented in C++ on top of PostgreSql using a standard client-server architecture for Solaris and Linux platforms. It consists of three modules: (a) the Parser, which analyses the syntax of queries; (b) the Graph Constructor, which captures the semantics of queries in terms of typing and interdependencies of involved expressions; and (c) the Evaluation Engine, which accesses RDF descriptions from the underlying database via SQL3 queries. It supports XML Schema data types, grouping primitive aggregate functions and recursive traversal of class and property hierarchies. It is easy to couple with commercial ORDBMSs as well as to integrate with various Web Application Servers. Finally, it provides a generic RDF/XML form of query results that can be processed by standard XSL/XSLT scripts. An online demo is available at http://139.91.183.30:8999/RQLdemo/.

Links:
http://139.91.183.30:9090/RDF/
http://139.91.183.30:8999/RQLdemo/
http://cweb.inria.fr/

Please contact:
Vassilis Christophides, Dimitris Plexousakis
ICS-FORTH
Tel: +30 810 3916 28, 3916 37
E-mail: {christop, dp} @ics.forth.gr

Corporate Semantic Webs

by Rose Dieng-Kuntz

The ACACIA multidisciplinary team at INRIA-Sophia-Antipolis is working towards offering methodological and software support (ie models, methods and tools) for knowledge management (ie for building, managing, distributing and evaluating a corporate memory) for an organisation or community. This corporate memory can be materialised in a 'Corporate Semantic Web' consisting of ontologies, resources (such as documents or persons) and annotations, possibly with modelling of multiple viewpoints.

We make an analogy between the corporate memory and the open World Wide Web: both are heterogeneous and distributed information landscapes and share the same problem of information retrieval relevance. In contrast, however, corporate memory has a context, an infrastructure and a scope limited to the organisation in question.

As research on the Semantic Web aims to make the semantic contents of the Web interpretable by software, we propose to materialise a corporate memory as a 'Corporate Semantic Web', constituted of ontologies, resources (ie documents or humans) and semantic annotations on these resources (ie on the document contents or the person features/competences), where these annotations rely on the ontologies.

We are studying the main research problems raised by such a 'Corporate Semantic Web':

- how to detect organisational needs, and whether to build the Corporate Semantic Web and each of its interrelated components from several human experts or from texts (cf methodology and architecture)
- how to diffuse and use the memory (information retrieval vs proactive dissemination)
- how to evaluate and evolve the memory (from both designer and user viewpoints).

Ontology and Annotation Construction

For the construction of ontologies, the CoMMA method integrates:

- a top-down approach relying on a synthesis of existing ontologies
- a bottom-up approach so as to exploit the (possibly partly automated) analysis of corporate documents or of human experts' interviews guided by scenarios
- a middle-out approach relying on generalisation and specialisation of key concepts of the organisation.

We used this hybrid method to build the O'CoMMA ontology (see below), dedicated to organisational memory and to scenarios such as newcomer insertion or technological monitoring.

As several corporate information sources can help to build an ontology (eg human experts, textual or multimedia documents, databases with possibly textual

data), the SAMOVAR method and system rely on partly automated construction and enrichment of ontologies and of annotations from heterogeneous information sources, by applying natural language processing tools on textual sources. SAMOVAR was applied to project memory in the automobile industry.

Ontology and Annotation Representation

As several languages can represent ontologies and annotations in a Corporate Semantic Web, the cooperative project ESCRIRE with the EXMO and ORPAILLEUR teams enabled us to compare three knowledge representation (KR) formalisms (conceptual graphs (CG), object-based KR and description logics), from the standpoint of the representation and the handling of document content. Our team focused on the CG formalism.

In order to palliate some limitations of RDFS for ontology representation, we proposed DRDFS, an extension of RDFS which enables the expression of explicit definitions of concepts and of relations, as well as contextual knowledge. We also proposed the GDL language, a description logic inspired by conceptual graphs.

CORESE Semantic Search Engine

RDF, recommended by W3C, allows the contents of documents to be described in the form of metadata resting on an ontology represented in RDF Schema (RDFS). We developed CORESE (COnceptual REsource Search Engine), an RDF(S) engine based on conceptual graphs. It enables an RDFS ontology and RDF annotations to be loaded and translated into CG, and then, thanks to the CG projection operator, allows the base of annotations to be queried. CORESE also reasons on properties of relations (symmetry, transitivity, etc) and on XML-based inference rules, so as to complete the annotation base. CORESE was used in the projects Aprobatiom, CoMMA, ESCRIRE and SAMOVAR, and allows ontology-guided information retrieval in a 'Corporate Semantic Web'.

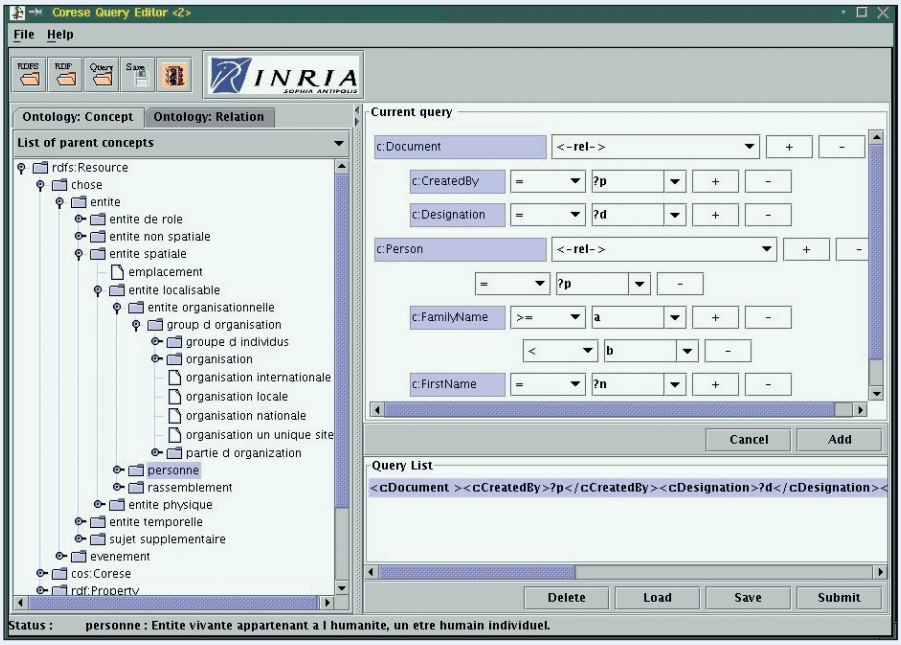


Figure 1: CORESE semantic search engine interface.

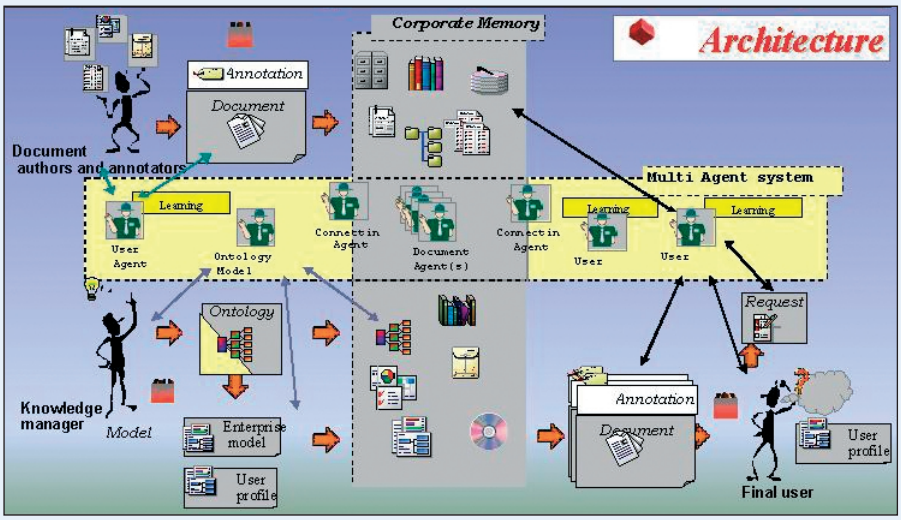


Figure 2: Corporate Memory Managements through Agents — CoMMA architecture.

The CoMMA project: Corporate Memory Management through Agents

Heterogeneity and distribution of multi-agent systems can be a solution to heterogeneity and distribution of a corporate memory. In the framework of the IST project CoMMA (with ATOS, CSTB, LIRMM, the University of Parma and T-Nova), we proposed a multi-agent architecture for the management of a Corporate Semantic Web, enabling ontology-guided information retrieval (the 'pull' approach) and proactive dissemination of information to users according to their profiles (the 'push' approach). We distinguish agents dedicated to ontologies, to documents, to users and to interconnection between agents. The agents are guided by the O'CoMMA ontology, a corporate model and user models. Some agents have learning capabilities and can adapt to the users.

We have also developed algorithms for learning ontologies incrementally from RDF annotations on (corporate) Web resources. This method can be used to build classes of documents automatically from their annotations.

Two scenarios were studied: the insertion of new employees and technological monitoring. The CoMMA system, implemented in JAVA on the FIPA-compliant multi-agent platform JADE, integrates our search engine CORESE.

Applications and Further Work

Our work on Corporate Semantic Webs was applied to project memory, technological monitoring, knowledge servers in the automobile industry (Renault), in the construction sector (CSTB), in telecommunications (T-Nova, TILAB) and in the biomedical domain. We also take part in the IST thematic network OntoWeb (Ontology-Based Information Exchange for Knowledge Management and Electronic Commerce).

In further work we intend to study a corporate memory distributed among several cooperating companies or communities. This memory will be materialised through a Semantic Web operating between organisations or communities, possibly handling multiple ontologies and multiple viewpoints and possibly requiring Web mining, for scenarios including project memory, technological monitoring and skills management.

The work described above was carried out by Olivier Corby, Alexandre Delteil, Rose Dieng-Kuntz, Catherine Faron-Zucker, Fabien Gandon, Alain Giboin, Joanna Golebiowska and Carolina Medina-Ramirez.

Links:
http://www.inria.fr/Equipes/ACACIA-eng.html
http://www.inria.fr/acacia/Publications

Please contact:
Rose Dieng-Kuntz, INRIA
Tel: +33 4 92 38 78 10
E-mail: Rose.Dieng@sophia.inria.fr

Joint Research Centre of the EC to use Semantic Web for New Privacy Agent

by Giles Hogben and Marc Wilikens

The Joint Research Centre of the EC is building an experimental privacy protection agent using Semantic Web technology and the Worldwide Web Consortium's P3P protocol. The agent automates the process of protecting a user's privacy through parsing and comparing privacy policies and user preference sets.

The Cybersecurity team of the EC's Joint Research Centre (JRC) in Ispra, Italy has recently completed work on a fully compliant implementation of the W3C's new privacy protocol, P3P, which uses XML files for the automated exchange and interpretation of website privacy policies and their matching against user preferences.

In the next stage of our research, the JRC will use this base platform to investigate how Semantic Web (SW) technology can improve this system. The project has been divided into a series of distinct stages.

The first stage is already under way and consists of the development and agreement of an ontology for data protection.

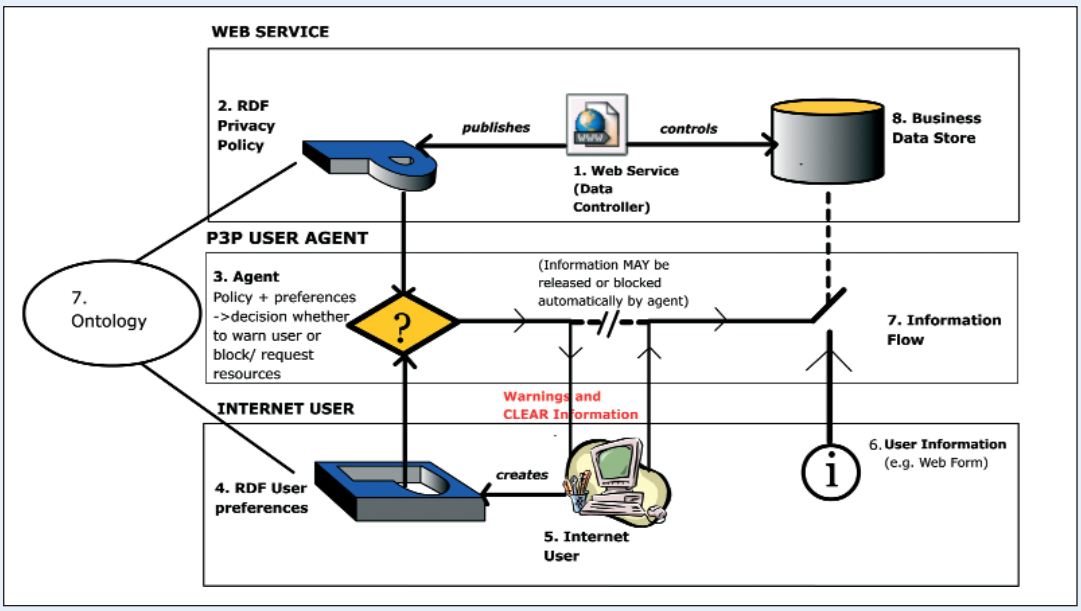
One of the most crucial improvements that the SW can make to the P3P platform will be to provide an interoperable, machine interpretable ontology for data protection. This will provide the following benefits:

- a common, interoperable vocabulary, which reduces misinterpretation of basic principles by technologists and legal experts and enhances interoperability between different systems
- clear understanding of terms allowing ease of translation between alternative ontologies
- clear separation of vocabulary and syntax meaning that the same vocabulary can be plugged into different data protection systems.
- a clearly documented development process offering clarity, authority and common agreement on terms by a large group of stakeholders.

An agreed vocabulary should be the basis for any privacy system. Significant problems were found in P3P's first version, mainly arising from inadequacies in its expression of data protection concepts such as those found in regulatory frameworks (eg the EU data protection directive and the OECD guidelines).

The JRC is developing a consensus process for capturing the knowledge of domain experts such as data commissioners. This is in conjunction with researchers from the University of Aberdeen, whose prototype ontology capture process will be used. This process includes input from psychologists and conceptual modeling research and has been put through several test cases.

Despite incompleteness in certain respects, P3P provides a sound foundation for such an ontology. Although P3P is not expressed in standardised ontology syntax, it represents, through the W3C processes that have underpinned it, a five year consensus process for a data protection vocabulary. Given a formally documented consensus process and improvements in the data typing schema, and purpose and recipient taxonomies, the existing version of P3P provides a very useful starting point.



Semantic-Web-enabled privacy agent.

The second stage will be to ensure that the proposed extensions are backward compatible. For this purpose, an XSLT stylesheet will be created to perform a translation between the two versions. The concepts of the ontology will also be placed within a clear W3C style specification.

The third stage will be to develop a rule system, which agents can use to make decisions based on the new framework. The rules, as in the previous system, look for a set of conditions in an RDF privacy policy, and the matcher executes a behaviour depending on the condition satisfied. The condition is identified using an RDF query language. The rule-matching process proceeds as follows:

- retrieve P3P privacy policy and load set of rules
- attempt to match conditions from each rule in turn
- when first rule matches, perform behaviour specified by that rule
- no rules fired is an error (therefore a catch-all rule must be included).

At this point, there will be a working system, as in the Figure.

The fourth stage will be to refine the rule system and policy language using test cases from European law. In the first phase of the JRC's P3P implementation, work was done on expressing the European Data Protection Directives using P3P vocabulary. This revealed that there are some inadequacies in the vocabulary, particularly when describing whether data is given to recipients outside of the EU and in the description of the purposes of data collection. The forthcoming SW version will, it is hoped, have sufficient conceptual flexibility to express the European Data Protection Directives. In any case, these will act as a litmus test for the system.

Finally, the research will cover ways of using the SW to make privacy information more accessible to other systems. At present, P3P functions only within the context of http transactions. The use of SW technology, as a standardised knowledge transfer medium, may help other areas of technology to use P3P. Examples might include ubiquitous computing, smart card technologies and IRC chat rooms. These developments put P3P in a good position as SW technology becomes more widespread. As knowledge sharing moves increasingly towards the use of the SW, the existence of an integrated privacy framework will facilitate association of privacy preferences and policies with data as they move between heterogeneous sources.

Links:
More information and a demonstration of our current platform: http://p3p.jrc.it

W3C's P3P Privacy protocol home page: http://www.w3.org/P3P

RDF schema for P3P: http://www.w3.org/TR/p3p-rdfschema/

OWL Web ontology home page: http://www.w3.org/2001/sw/WebOnt/

Please contact:
Giles Hogben, European Commission
Joint Research Center, Italy
Tel: +39 033 278 9187
E-mail: giles.hogben@jrc.it

Exmo: Assessing Semantic Properties of Knowledge Transformations

by Jérôme Euzenat

Sharing knowledge on the Semantic Web will require the transformation and adaption of knowledge to client needs. The complexity of the representations to be transformed can hide the effects of transformations, though the use of knowledge dictates them. Exmo aims to provide tools for characterising the action of these transformations in syntactic, semantic or semiotic terms.

Approaching semantically defined knowledge on the Web is not the end of interoperability problems. Indeed, knowledge providers use different languages, axiomatic or conventions for expressing knowledge. Importing knowledge in the context of different applications will require adapting it to the target context (language, axiomatic or convention). This adaptation can be processed a priori or on the fly and the transformation involved can be automatically or manually generated.

The action of these transformations ranges from minor syntactic formatting to complex processing such as deducing consequences and generalising them. It is important for applications to be confident of the properties of these transformations. These properties can be related to the syntax (eg correct expression of the result in a particular language), the semantics (eg preservation of the meanings of knowledge by the transformation) or the pragmatics (eg preservation of the reader's interpretation by transformation). They may also concern the transformation processing (eg correct termination) or its results (eg loss of information in the result).

One of the goals of the Exmo action is to assess the properties satisfied by transformations. This objective requires:

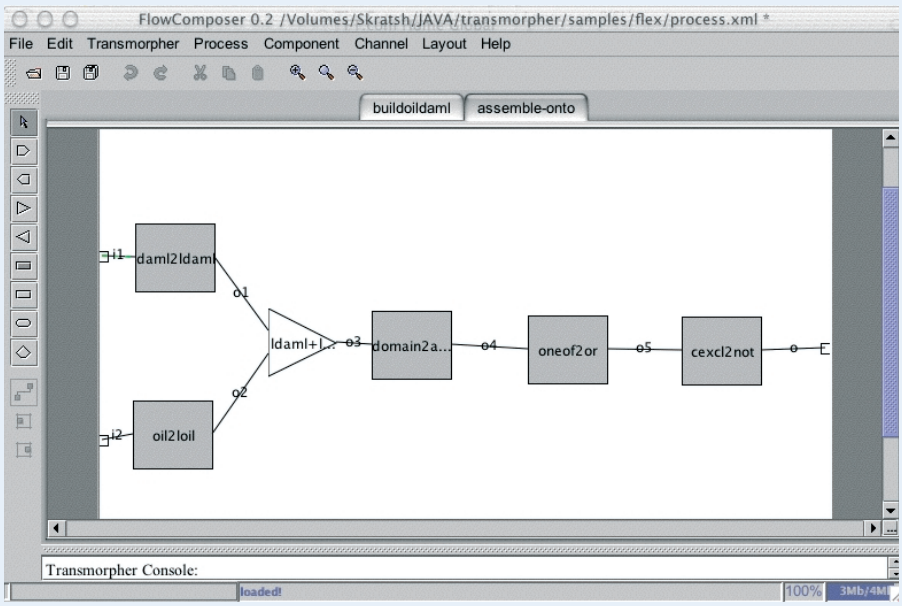
- the capacity to analyse transformations into units assembled by precise constructors
- the knowledge of properties that can characterize transformations and their behaviour with regard to the constructors
- the opportunity to prove properties from elementary transformations.

Exmo deals with these three problems and aims to build an environment in which users can compose transformations and formally characterise their properties.

In collaboration with the Fluxmedia company we developed the Transmorpher environment that allows the composition of XML transformations. The unit transformations are either expressed in XSLT (in which case the properties are asserted) or in a simplified rule transformation language (from which properties can be assessed). These transformations can be composed in subroutine, sequence, parallel or loops which constitute the main constructors of the system. The environment is portable and freely available (see Figure).

The transformations developed within Transmorpher will have to be annotated by the properties that they satisfy. In particular we are investigating semantic properties (tied to the model theoretic semantics of the languages involved) and their relations (eg interpretation-preservation implies consistency-preservation). These properties have been applied to an XML encoding of description logics and this effort should be reiterated in the next W3C-proposed languages.

The next step consists in providing Transmorpher users with tools for proving the properties of a particular transformation. Our current efforts use the relations between properties in order to establish the minimal properties enjoyed by the composition of Transmorpher constructors. We would also like to experiment with proof checkers for the inductive proofs often used for proving the equivalence of languages.



Screenshot of the Transmorpher graphic environment.

In the context of the Web, and of the Semantic Web infrastructure, transformations could be made available for many other parties that might want to integrate them within their own processes. In order to establish the properties satisfied by the resulting transformation, it will be necessary to be able to assess those of the gathered transformations. For that reason, we want to extend the system presented above with the capability of importing and exporting transformations, their properties and, if available, the proofs of these properties. The importers would then be able to check the proofs and thus trust the properties of the imported transformations. This is an instantiation of 'proof-carrying code' to transformations.

Besides classical properties tied to the syntax and semantics of the representations, the overall goal of Exmo is to be able to take into account properties tied to their pragmatics, such as the rhetoric of a discourse or the semiotic interpretations of readers. For instance, this would enable us to establish that the same impact is provided by a multimedia presentation before and after its transformation. This requires investigations into the field of semiotics.

Links:
http://www.inrialpes.fr/exmo
http://transmorpher.inrialpes.fr

Please contact:
Jérôme Euzenat, INRIA
Tel: +33 476 61 53 66
E-mail: Jerome.Euzenat@inrialpes.fr

CORES — A Forum on Shared Metadata Vocabularies

by Tom Baker

Can different standards communities agree on principles to help translate among a diversity of metadata languages?

The Semantic Web vision of merging information from a diversity of Internet sources presupposes methods for processing diverse types of descriptive metadata. The past decade has seen numerous initiatives to standardise metadata semantics for particular application communities, from multimedia and commerce to scholarly publishing. In practice, there is some overlap between these communities of practice, and implementors often draw pragmatically on available standards in designing metadata for particular applications, 'mix-and-matching' as needed.

Processing metadata scalably will require infrastructures to interpret, translate, and convert among various metadata languages automatically. The construction of such infrastructures will however require the availability of metadata dictionaries in forms usable by software agents, middleware applications, and indeed human beings. Such dictionaries should ideally be based on conventions, shared across metadata communities, for declaring metadata vocabularies machine-understandably.

The objective of the CORES Project, an Accompanying Measure under the Semantic Web action line of the EU's Fifth Framework Programme, is to facilitate the sharing of metadata semantics on several levels. At the level of standardisation initiatives, the project is bringing together key figures from major standardisation activities in a Standards Interoperability Forum to discuss principles and practicalities of 'playing together' in the Web environment. The participants include representatives from initiatives as diverse as DCMI, OASIS, DOI, CERIF, MPEG-7, IEEE/LOM, GILS, ONIX, and the W3C Semantic Web Activity (with apologies for the unexpanded acronyms). While undertaken with reference to Semantic Web technologies, the Forum includes communities that are pursuing different strategies for interoperability.

At the level of infrastructure, CORES is working with organisations that are designing or building value-added services for metadata-using applications. Building on the prior SCHEMAS Project (2000-2002), CORES is expanding a registry of application profiles and liaising with related activities on developing good practice for interoperability between specialised registries. This effort will include a hands-on 'profile-writing' workshop with implementors.

One goal is to test the SCHEMAS model for declaring application profiles as simple sets of RDF assertions — a form which in principle lends itself to merging into 'semantic landscapes' of information providers.

Project partners are Pricewaterhouse-Coopers (Luxembourg), UKOLN, University of Bath (UK), Fraunhofer-Gesellschaft, and SZTAKI.

Link:
http://www.cores-eu.net/

Please contact:
Thomas Baker, Fraunhofer-Gesellschaft
Tel: +49 30 8109 9027
E-mail: Thomas.Baker@bi.fhg.de

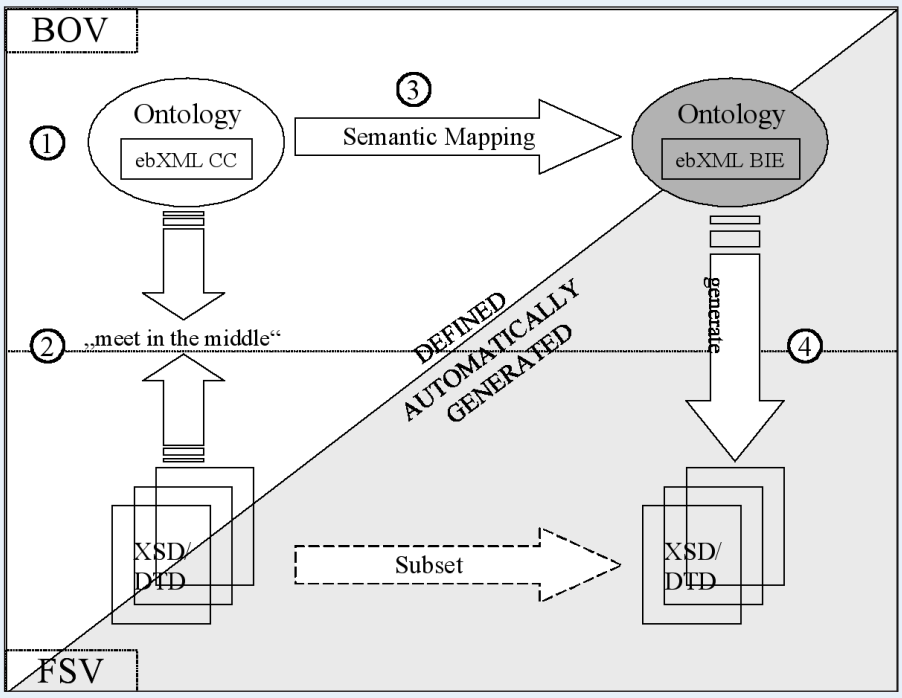
Towards Syntax-Independent B2B

by Birgit Hofreiter, Christian Huemer and Werner Winiwarter

When it was introduced, XML seemed to promise a solution to traditional B2B problems. Today, however, we are faced with a multitude of different XML-based e-business vocabularies that are not interoperable. Even in the case of business partners supporting the same e-business vocabulary, interoperability is not guaranteed since they may be implementing different subsets of the same vocabulary's document type. A research team at the University of Vienna addresses these problems with a framework that aims to align concepts known from the Semantic Web and the ebXML initiative.

The success of the Internet and XML created a mistaken belief in a global B2B market where businesses would be able to communicate with each other without any prior agreements. These technologies seemed to overcome the most significant obstacles of traditional electronic data interchange (EDI) standards. After the initial excitement however, people realised that while XML provides a means to exchange data between applications, it does not guarantee interoperability. While XML provides a syntax that can be used for data transfer in B2B, it does not provide semantics, since tags have no predefined meaning. The meaning of XML vocabularies is defined by the document designer and is specified outside of XML itself. This resulted in a proliferation of XML-based e-business vocabularies within the first few years of XML's release. Having contributed to the work of UN/CEFACT in the Techniques and Methodology Working Group (TMWG) and ebXML, as well as conducting related research projects such as A-XML and Harmonise, we feel that a true B2B solution needs much more than a purely syntax-based approach. Here we elaborate on the open issues that we see on the road towards syntax-independent B2B.

In the future, we expect some vocabularies to disappear and others to merge. Nevertheless, a certain number of well-known 'standard' vocabularies will coexist. This means that business partners must either agree on a certain e-business vocabulary or map between their mutual vocabularies. This is not practicable with a large number of business partners. Consequently, a common syntax-neutral ontology capturing the business semantics independently of vocabulary is desirable. Furthermore, a clearly defined representation of the business semantics in each supported e-business vocabulary is necessary. This approach minimises the effort required since the number of mappings can be kept as low as the number of vocabularies. In this context, two key aspects are highly relevant for the research community, namely, how to build the ontology and the mapping between the ontology and the business vocabularies.



B2B Framework based on Ontologies and the Semantic Web.

A major failure of today's B2B approaches is the issue of overloaded document types. This is the result of the desire to create one single purchase order, one single invoice etc, for the whole world. In order to meet the requirements of any company in any business sector in any geographical region, document types are created by collecting (and structuring) all data elements that might be relevant for any company in any sector in any region. As a consequence, only a very small percentage (usually less than 5%) of data element types are instantiated in a document. To guarantee interoperability, partners have to agree on the relevant subset of the document type prior to the exchange. Thus, documents must be compliant not only with the document types (called valid documents in XML) but also with industry-, region- or partner-specific agreements. Accordingly, we identify a third key aspect for the research community: the automatic adjustment of general document types according to given context parameters such as industry, region, and business process.

In order to address these three aspects we suggest drawing on approaches from the Semantic Web and electronic business XML (ebXML). In the Semantic Web community, languages such as RDF and DAML have evolved that can be employed to develop the syntax-neutral ontology. Furthermore, there is ongoing research concerning the transformation of an ontology's semantics into XML representations of certain e-business vocabularies.

EbXML core components are oriented along the most extensive knowledge base in the business world, the trade data element dictionary UN/TDED. Therefore, ebXML core components, which will also be supported by BoleroNet, EAN/UCC, Open-Applications Group (OAG), Open Travel Alliance (OTA) and SWIFT, provide the best semantic foundation for a global business ontology. As ebXML also allows the specification of business processes, their activities and their activity-specific business data, this facilitates the adjustment of business documents to the needs of a business process on the basis of core components, rather than through the use of generic document types. Thus, ebXML provides a good starting point for our third identified aspect.

By aligning ontology and ebXML concepts we propose the framework depicted in the Figure. This framework is based on four major steps. In the first step the semantics of ebXML core components are defined in an RDFS-based ontology. The second step defines mappings between the ontology and RDF representations of e-business vocabularies, as well as between the latter RDF representations and the DTDs or XML schemas of the e-business vocabularies. Beyond traditional ontological approaches, we take on the ebXML idea to further restrict a general document ontology to the specific needs of a given business activity. This refinement will lead to a certain view of the general document ontology and must be specified by means of a constraint language. Steps one to three must be done manually and are depicted in the white triangle, whereas the last step depicted in the grey triangle will be derived automatically. Using the information from the previous steps, a more restrictive and appropriated XML schema or DTD for a specific activity will be created. The XSDs and DTDs are a subset of their corresponding general ones. Further restrictions expressed in a declarative language will accompany the document types. This framework focuses on a mapping between different e-business vocabularies and at the same time guarantees their semantic equivalence in support of a specific activity.

Links:
ebXML: Homepage of the ebXML Initiative:
http://www.ebxml.org
Semantic Web Portal:
http://www.semanticweb.org

Please contact:
Birgit Hofreiter, University of Vienna
Tel: +43 1 4277 38434
E-mail: birgit.hofreiter@univie.ac.at

Extended Faceted Taxonomies for Web Catalogs

by Yannis Tzitzikas, Nicolas Spyratos, Panos Constantopoulos and Anastasia Analyti

Which would be easier to remember: one thousand individual terms or three facets of ten terms each?

One way of creating a taxonomy for a knowledge domain is by identifying a number of different aspects or facets of the domain and then designing one taxonomy for each facet. Several studies in information, library and cognitive science have shown that in almost every knowledge domain we can indeed distinguish a number of facets, or planes of understanding. A faceted taxonomy is actually a set of taxonomies, called facets, each of which is a set of terms structured by a specialisation/generalisation relation. Using a faceted taxonomy, the indexing of objects is done by associating each object with a compound term, ie with a combination of terms coming from different facets. It was recognised long ago that a faceted taxonomy has several advantages over a single hierarchical taxonomy, including conceptual clarity, compactness and scalability. For example, consider two schemes for indexing the objects of a domain, the first using a single taxonomy consisting of 100 terms, and the second using a faceted taxonomy consisting of 10 facets each having 10 terms. The first scheme has 100 indexing terms while the second has 10^{10} (10 billion) compound indexing terms! Although both schemes have the same storage requirements, ie each one requires storing 100 terms, the second has tremendously more indexing terms than the first.

However, faceted taxonomies have a major drawback that prevents their deployment and use for real and large-scale applications like the Web. This drawback comes from the fact that it is possible to form a large number of invalid compound terms, that is, combinations of terms that do not apply to any object of the underlying domain. For example, refer to the faceted taxonomy for a tourist information application shown in Figure 1 and consider the terms WinterSports from the facet Sports, and Crete from the facet Location. The

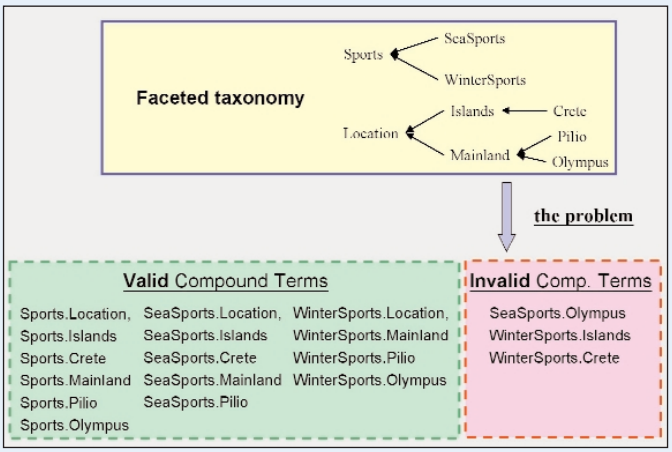


Figure 1: A faceted taxonomy consisting of two facets, Sports and Location, and the partition of the set of compound terms to the set of valid and the set of invalid terms.

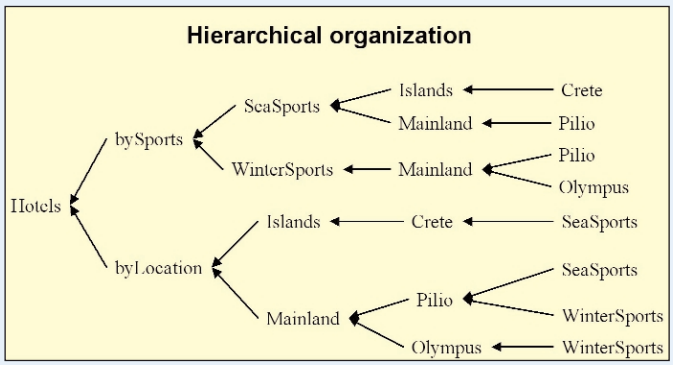


Figure 2: A hierarchical organisation of the valid compound terms of the faceted taxonomy of Figure 1.

compound term WinterSports.Crete is invalid, as there is never enough snow in Crete! In contrast, the compound term SeaSports.Crete is certainly valid. The inability to infer the valid compound terms may create problems in object indexing (laborious and/or erroneous indexing), and in browsing (an invalid compound term will yield no objects). Due to such problems, existing Web catalogues have a strictly hierarchical structure like the one shown in Figure 2. Such taxonomies suffer from several problems such as incomplete terminology, huge size (for example the taxonomy of Open Directory consists of 300 000 terms) and confusing structure.

Being able to infer the validity of a compound term in a faceted taxonomy would be very useful for facilitating indexing and for preventing errors, especially in cases where indexing is done by many human editors (indexers). For example, Web pages in the Open Directory are currently indexed by more than 20 000 volunteers. Moreover, if we could infer the valid compound terms in a faceted taxonomy then we would be able to generate navigation trees (like the taxonomies of Web catalogues) on the fly, consisting of nodes that correspond to valid compound terms. However, manually defining the set of valid compound terms even for facets of relatively small size would be a formidable task for the designer. For example, Figure 1 shows the partition of compound terms into sets of valid and invalid terms.

To alleviate this problem, we have defined two extensions of faceted taxonomies, which we call PEFT and NEFT, which allow the specification of valid compound terms in a flexible and efficient manner. A PEFT (Positive Extended Faceted Taxonomy) is a faceted taxonomy enriched with a set P of compound terms known to be valid, while a NEFT (Negative Extended Faceted Taxonomy) is a faceted taxonomy enriched with a set N of compound terms known to be invalid. The designer simply declares a small set of valid or invalid compound terms, and other (valid or invalid) compound terms are then inferred by a mechanism based on semantic implication. Figure 3 shows how we can specify the valid compound terms of the faceted taxonomy of Figure 1, ie the sets 'Valid Compound Terms' and 'Invalid Compound Terms' as enumerated in that figure, by employing either a PEFT or a NEFT. In each case we can dynamically derive a navigation tree such as the one shown in Figure 2 that can be exploited during object indexing and browsing.

Our approach can be used for developing Web catalogues that offer complete navigation trees, require less storage space, and are more comprehensive and scalable. Furthermore, taxonomies designed according to our approach can be integrated or articulated more easily than hierarchical ones. Further research includes extending this approach to define an algebraic system for taxonomies with operators that allow the specification of valid compound terms using both positive and negative sets of compound terms.

This work is the result of basic research conducted at ICS-FORTH over the last two years.

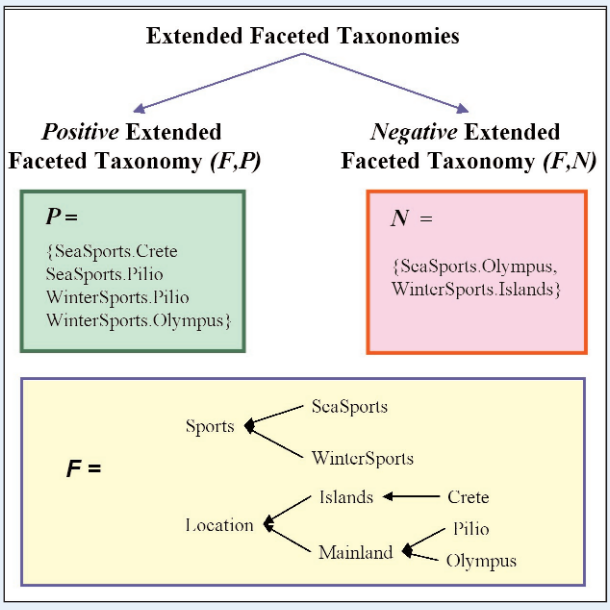


Figure 3: Two extended faceted taxonomies.

Please contact:
Yannis Tzitzikas, ICS-FORTH
Tel: +30 810 391 623
E-mail: tzitzik@csi.forth.gr
http://www.csi.forth.gr/~tzitzik

Migrating Thesauri to the Semantic Web

by Michael Wilson and Brian Matthews

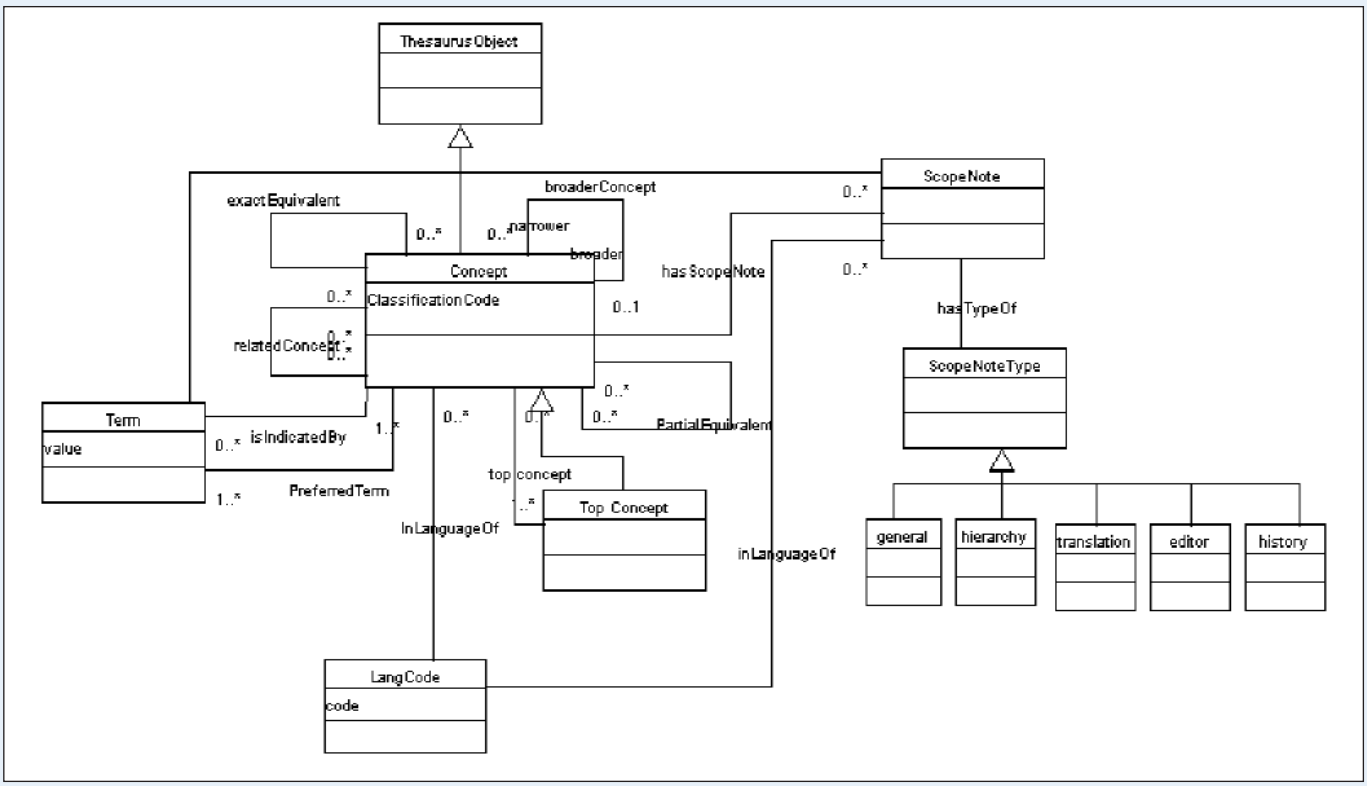
If Semantic Web technologies using RDF are going to be adopted and assimilated as HTML and XML have been, a clear migration path from present technologies to new ones is required. Thesauri are used throughout the information retrieval world as a method of providing controlled vocabularies for indexing and querying.

W3C is developing standards for the representation of ontologies to constrain the vocabularies of resource descriptions based on RDF. Such ontologies will allow distributed authoritative definition of vocabularies that support cross-referencing. Such ontology representations are planned to fulfil the role currently undertaken by thesauri. Therefore a migration path is required from current thesauri to ontologies, or support for their co-existence if those ontologies are to be adopted and assimilated into existing information retrieval infrastructure.

The structure of thesauri is controlled by international standards that are among the most influential ever developed for the library and information field. The main three standards define the relations to be used between terms in monolingual thesauri (ISO 2788:1986), the additional relations for multilingual thesauri (ISO 5964:1985), and methods for examining documents, determining their subjects, and selecting index terms (ISO 5963:1985). The general principles in ISO 2788 are considered language- and culture-independent. As a result, ISO 5964:1985 refers to ISO 2788 and uses it as a point of departure for dealing with the specific requirements that emerge when a single thesaurus attempts to express 'conceptual equivalencies' among terms selected from more than one natural language.

The ISO standards for thesauri (ISO 2788 and ISO 5964:1985) are developed and maintained by the International Organization for Standardization, Technical Committee 46 whose remit is Information and Documentation — not IT. ISO 5964:1985 is currently undergoing review by ISO TC46/SC 9, and it is expected that among changes to it will be the inclusion of a standard interchange format for thesauri. To facilitate the growth of the Semantic Web, it would be sensible to try to ensure that such an interchange format is as compatible with Semantic Web ontology representations as possible.

Several proposals have arisen for thesauri interchange formats based on either RDF or DAML+Oil. The major problems with these is that either they cannot accommodate the multiple inheritance common in many multilingual thesauri or that the semantics of thesauri in the ISO standards are not as precise as these languages require. The links in thesauri hierarchies define the top term in the hierarchy, and the broader or



The thesaurus model in UML.

narrower coverage of terms down the hierarchy. There are also links between hierarchies to show equivalence in different languages, or similar meaning in the same language.

However, the hierarchical links in thesauri are semantically overloaded, and the potential exists using Semantic Web ontology representations to develop ontologies with less overloading. The terms 'broader', 'narrower', 'used for', 'related' and 'equivalent' are not defined by precise semantics. Therefore the proposals are either too precise to be compatible with some existing thesauri or include explicit statements of semantics which are seen to be unacceptable to other thesaurus developers and users.

We have developed a proposal for a thesaurus interchange format expressed in RDF to overcome these limitations, which has been applied to one large multi-lingual thesaurus for evaluation by users - ELSST. It is planned to represent many more thesauri in this representation, and to show how they can both be migrated into Semantic Web ontologies, and how such ontologies allow the thesauri from different domains to be related to each other.

ELSST: a multilingual thesaurus (English, French, Spanish, German) for the social science domain has been represented in the Thesaurus Interchange Format. The thesaurus has been produced containing 49 hierarchies, incorporating 1456 preferred terms. Following the initial development further translations of terms into Finnish, Norwegian, Danish & Greek are planned, as are the inclusion of terms related through inexact translations in addition to the exact translations already included. Also the CESSDA group of European Data Archives has agreed to adopt ELSST as the European Controlled Vocabulary for Social Science.

Link:
http://www.limber.rl.ac.uk/External/SW_conf_thes_paper.htm

Please contact:
Michael Wilson, Brian Matthews, CLRC
Tel: +44 1235 44 6619
E-mail: M.D.Wilson@rl.ac.uk

Semantic Characterisation of Links and Documents

by Silvia Martelli and Oreste Signore

Semantic characterisation can considerably enhance navigation possibilities and makes possible the presentation of documents for a real adaptive hypertext environment.

Documents on the Web are often deeply structured, and this can be the origin of incompatibility between different views of the same object. Links are the essence of hypertext, but they are meaningful only if their semantics is clear, and perceivable by the user. As a consequence, it can be useful to have lightly structured documents, where some elements can be seen as 'semantic items' that identify concepts chracterising the specific parts of the document. Links can share basic semantic categories with these semantic items, and can be implemented in the XLink framework. This quite simple approach leads to a considerable enhancement of navigation possibilities, and makes it possible to present documents apropriately for a real adaptive hypertext environment.

The Hyperlink Association Model

When reading, our attention is often captured by words (anchors) that lead our mind to other documents. In the Web context, documents, whatever their origin, are seen as resources. We can model the association process in the following way:

- the anchor leads to a concept
- the concept is related to other concepts
- the new concept is related to some resources.

This basic association mechanism (see Figure 1) is totally independent of the document structuring. In the data space, documents are connected by extensional links. In the concept space — a simplified version of the Semantic Web architecture's ontology level — associations among concepts implement intensional links among documents. Two questions now arise:

- how can we implement the link from resources to concepts
- how can concepts be linked together.

A simple and effective way to implement intensional links is to identify the semantic items. This can help in several cases (for example, 'The French emperor' can be an implicit reference to 'Napoleon'). We can also characterise each semantic item with a specific semantic category (eg person, location, date, taxonomy) in order to tailor a document to specific user interests, eg a reader interested in space-time associations will get location and date items emphasised.

The second question directly leads to the interaction metaphor issue. In addition to the case of taxonomic classifications, where we can make use of well-known thesaurus techniques, space and time can function as very powerful association mechanisms. A semantic item can point to a location, then, using an interaction metaphor based upon space, the user can either jump to other resources linked to the same location, or select a different location, and then find other resources related to this new location. This simple

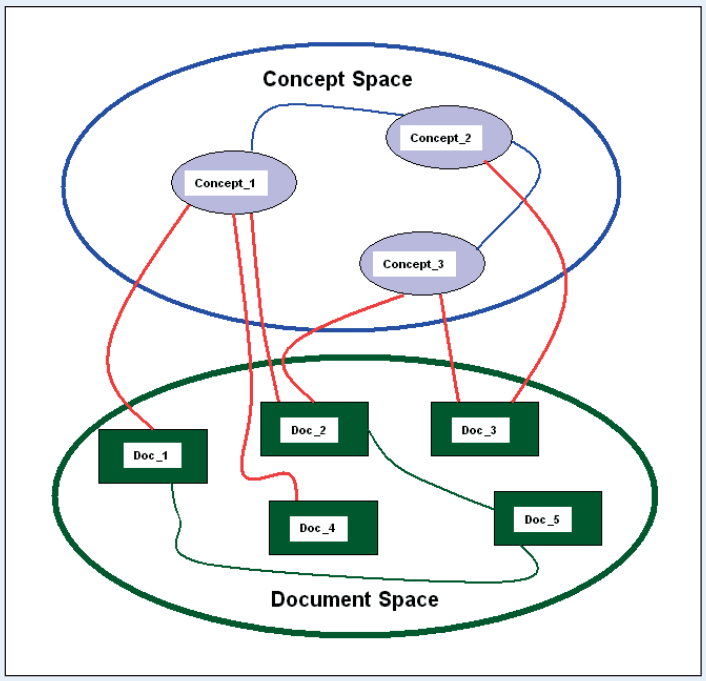


Figure 1: Document and Concept Space.

hyperlink association model can be implemented through a document, link and user model.

Semantic Model of Documents

In XML documents we must clearly distinguish between structural and semantic information, which can be associated with elements or parts of them. For the sake of simplicity and effectiveness, it is possible to define a limited and rough general set of semantic categories, that can be structured in a thesaurus-like fashion. These categories can be shared by a wide variety of users, can be used to define a user profile and can semantically characterise various parts of the documents and links. We can also specify a weight, stating the relevance of the concept in the document context.

Link Taxonomy and Model

Links allow navigation on the Web, and can implement the abstraction mechanisms needed to move from data space to concept space. Semantic qualifications of links explicitly identify their meaning in the document and the role of involved resources. The reason why the link has been inserted in the document, ie the nature of association (geographical, explicative, etc), can be explicated through the link's semantic type. Different types of links can suggest different and specialised interaction paradigms (time, map, classification, etc) having an enormous effect on the potential association mechanism: two documents can be linked through an intensional link existing in the concept space, even without the extensional link being specified in the document.

User Model

As a first approximation level, the user model is defined in terms of an essentially dynamic profile, tightly related to the semantic model of documents and links. The user profile is defined in terms of semantic categories, link types and link roles. For each of these, a degree of interest (weight) is stated.

A Simple Example

Take the following fragment of an XML document, containing some semantic items:

In <abwr:si st="date"> 1812 </abwr:si> <abwr:si st="person" canonicalName = "Napoleon"> the French emperor </abwr:si> invaded Russia ...

This document may have been entered as it is, or may be the result of a more complex process, involving a database search and data processing.

A software agent can take this document as input, producing a richer XML document, where the expression "the French emperor" becomes the anchor of an extended link (using XLink terminology).

When producing this document and also when displaying it, the software agent can examine the user profile, in order to produce a personalised document that will be displayed by the browser, as shown in Figure 2.

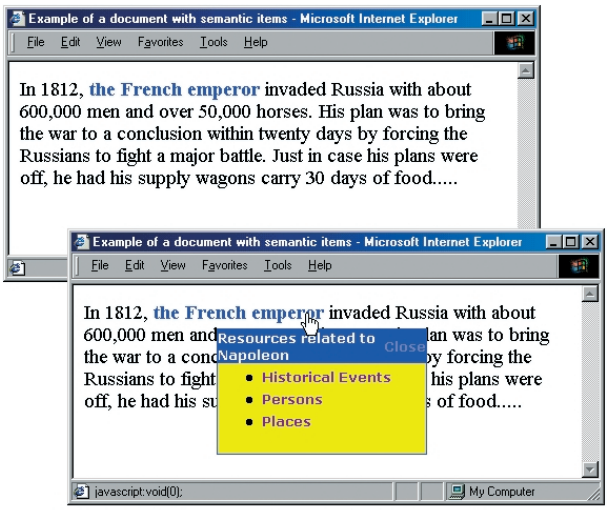


Figure 2: Document with an activated semantic item.

Please contact:
Oreste Signore, ISTI-CNR
Tel: +39 050 3152995
E-mail: Oreste.Signore@cnuce.cnr.it

An Ontology-Oriented Solution for Knowledge-Intensive Organisation

by Emanuele Della Valle and Maurizio Brioschi

Knowing what you know is increasingly a real problem in many organisations whose core competence is knowledge. We believe that using state-of-the-art Web-based Knowledge Management (KM) technologies will not be sufficient in the immediate future, since the lack of formal semantics will make it extremely difficult to make the best use of the massive amount of stored data.

Knowledge-intensive organisations are more common than you might expect: R&D departments and education centres are among those which first come to mind, but any community of interest, any team, any collaborative group that maintains a large amount of distributed and semi-structured data can fit in this category.

As Intranet solutions have increased in popularity, they have become populated with documents, forms, calendars of events, news, link collections and applications such as databases, newsletters and forums. As a consequence, most organisations have ended up with a huge repository of semi-structured knowledge distributed over the Intranet. This has resulted in some common maintenance and accessibility problems. On the one hand, the administration of these knowledge sources cannot be centralised, because they are maintained only as long as they are useful for those who write them. On the other hand, all members of the organisation, regardless of differences in cultural background and level of expertise, should be able to find what they need: the form for initiating an administrative procedure, the application to broadcast urgent news to a community of interest, someone with a specific set of skills to whom a task can be assigned, a tailored course to learn how to perform an unfamiliar operation, etc.

So far, the combination of Web technologies (XML, J2EE, SOAP, etc) and knowledge management methodologies has proved able to address some of these problems. Intranet portals, based on a collaborative and distributed editorial model, can provide a single point of access and a good content management solution. Moreover, the development of component-based applications, in particular with emerging Web services technology, provides the strong decoupling and easier composition required by these kinds of data-intensive Web applications. However, what our state-of-the-art technology can provide is good human access to this knowledge. While a good Intranet search engine can lead the user to a set of knowledge sources, because of its lack of common sense, it is unable to analyse them.

Semantic Web technology, with the explicit introduction of a formal semantics for each knowledge source, will make it possible for a machine to automatically process the source and in a way, to understand it. In particular, the use of ontologies, which are explicit conceptualisations of a shared understanding of a domain, can enable a reasoning service to analyse the knowledge stored in the sources. In addressing the common problems which arise in the Intranet environment of a typical knowledge-intensive organisation, the e-Service Technologies unit of CEFRIEL is working on two projects and will commence a number of further projects in the coming autumn.

In collaboration with the KM competence centre of Getronics, we are currently running the sKM (semantic Knowledge Management) project. The goal of sKM is to use ontologies as the core component of a methodology to select and configure the most suitable KM solution for a specific organisation. In order to achieve this, we validate the possibility of integrating into a framework a set of freely available tools for designing and maintaining ontologies. We monitor the available languages and tools and, so far, we have selected DAML+OIL as the language, OILed v3.4 as the editor and FaCT v2.2 as the reasoner. Moreover, we have conceived and implemented two tools: a graphic visualiser VisualDAML+OIL and an ad hoc report generator, ReportDAML+OIL.

We use ReportDAML+OIL to produce the information needed by a technician to configure a KM solution. The next step is to study the possibility of integrating current state-of-the-art KM solutions with reasoner-able services in order to address the common maintenance and accessibility problems. The options we are considering include a framework for automatic publication of content based on the semantics associated with the content, a mechanism to support the annotation which deduces implicit knowledge from that which is explicitly given, and a solution to empower the basic syntactic search mechanism widely available with the use of the given semantics. This means the search engine could learn to find what you mean, rather than simply what you say!

In the meantime, we're implementing COPPER (CEFRIEL Open Portal Project for Enterprise Resources). This is an industrial-strength portal solution built only with open-source technologies (Apache, Tomcat, Velocity, PostgreSQL, Java). We have already used it for building an Intranet portal and we plan to adopt it as our own testbed and training field for the prototypal implementation of those reasoner-able services described above. In the coming autumn we will start several new projects. In 'Semantic Enhanced e-Learning', we will study the problems which arise in populating and maintaining an open learning object repository with special attention paid to the

construction of ad hoc courses. Moreover, we will start a technology transfer project that concerns both the skill management and the accessibility of a CAD-drawing repository. Last but not least, we're looking at the possibility of using semantic-enabled Web services in an e-government Intranet application integration scenario.

People involved: Maurizio Brioschi, Stefano Campanini, Paolo Castagna, Emanuele Della Valle, Marco Riva, Nicola Simeoni.

Links:
http://etechdemo.cefriel.it/sw
http://www.cefriel.it/etech

Please contact:
Emanuele Della Valle
CEFRIEL, Politecnico di Milano, Italy
Tel: +39 02 23954 324
E-mail: emanuele.dellavalle@cefriel.it

Combining XML and Description Logics for Describing and Querying Documents

by Rim Alhulou and Amedeo Napoli

Representing and handling documents by their content is the core of ESCRIRE, a coordinated action involving three INRIA research teams.

In order to take advantage of the huge amount of information available on the Web and in document databases, it is necessary to design efficient techniques for retrieving, extracting and querying documents. Ontologies are playing an increasing role in these tasks, especially for content-based annotation and manipulation of documents. The objective of the ESCRIRE project is to use an ontology to annotate a set of abstracts of biological documents extracted from the NIH Medline public database, and then to query the annotated documents within a knowledge representation (KR) formalism. In the following, we restrict our attention to the use of a description logic, namely the RACER system, within the ESCRIRE project.

In order to manipulate documents by their content, annotations are attached to documents and a domain ontology has been designed for this purpose. The annotations and the ontology are described within a pivot language based on XML. This pivot language relies on a set of syntactic rules controlled by a DTD. The pivot language plays the role of a bridge between documents and the description logics (DL) formalism: every element in the ontology and every annotation have a corresponding element in the DL formalism. The pivot language is also used to describe SQL-like queries, which are in turn represented within the DL formalism to be handled by the DL classifier. The pivot language has been especially built for the needs of the application, and is not simply another XML-based language for document description.

Briefly, the ontology consists of a hierarchy of classes representing concepts, eg genes, and relations between classes, eg interactions between concepts. Each class in the ontology is described by a set of attributes and roles representing the properties of the class.

Two types of classes are available: defined classes with necessary and sufficient conditions, and primitive classes, with only necessary conditions.

A document is composed of three parts: (1) a textual abstract; (2) a set of classic metadata (Dublin core); and (3) a set of metadata concerning the content (annotations). The pivot language is used for representing the annotations according to the ontology, especially the objects and relations referenced in the documents, in this case, the genes and the interactions. Objects (instances of classes) and relations (instances of classes of relations) are described by their properties (names and values of roles and attributes), and the class to which they belong. The structure of a query respects the classical schema SELECT-FROM-WHERE, with some additional constructs being available.

The classes and relations of the ontology are translated into concepts within the DL system. Relations are also represented as concepts. All attributes and roles are translated into roles in the DL system. The two types of classes in the ontology — defined and primitive — are transformed into DL concepts according to their status: defined classes become defined concepts in the DL system, while primitive classes become primitive concepts. However, the properties of a relation, eg reflexivity, antisymmetry and transitivity, must be managed by a module that is external to the DL system.

Each document is then represented as an individual within the DL system. Objects and relations referenced in the ontology or in the documents are translated into individuals within the DL system. Moreover, an individual is related to all the individuals filling its roles and attributes. Individuals representing objects and relations are linked to those representing the documents where they are referenced.

Classification and subsumption are the main reasoning methods in the DL systems, and they are used for processing queries. A query Q can be translated into one or more query concepts C_i in the DL system. Each query concept C_i is then classified in the concept hierarchy. The answer to the query Q is constituted by the set of instances of each classified query concept C_i.

A number of problems have appeared during the development of the project. We may underline the following difficulties showing the extensions of a KR formalism for representing and handling documents by their content. The DL system does not provide any special constructor for taking into account binary as well as n-ary relations, especially for handling the properties of relations such as reflexivity, symmetry and transitivity. The possibility of working with or without the closed-world assumption was not available, and would have been very useful. The translations and the evaluations of queries were not simple nor always efficient. Actually, a more sophisticated module for handling queries based on the formalism of conjunctive queries has to be practically designed for a realistic document manipulation.

The first results of the ESCRIRE project show that a DL system such as RACER can be used with relative success for representing an ontology of a domain, and for describing and querying a set of documents in an XML-like form. More work must still be done to solve the problems mentioned above. A number of theoretical tools exist, but still haveto be made practical for an effective and realistic manipulation of documents by their contents.

Link:
http://www.inrialpes.fr/exmo/cooperation/escrire/

Please contact:
Rim Alhulou and Amedeo Napoli
LORIA, France
E-mail: {alhulou,napoli}@loria.fr

Harmonise: An Ontology-Based Approach for Semantic Interoperability

by Michele Missikoff

One of the services that we expect to find with the advent of the Semantic Web is semantic interoperability. The goal of semantic interoperability is to allow the (seamless) cooperation of two Information Systems that were not initially developed for this purpose. The European Project 'Harmonise' aims at building a technological infrastructure based on a shared ontology, to enhance the cooperation of European SMEs in the tourism sector.

The Semantic Web is expected to enhance the development of semantic interoperability among Information Systems. The aim is to allow Information Systems to cooperate without requiring modifications to their software or data organisation. In the IST European Project 'Harmonise', information interoperability has been addressed within a 'Local As View' (LAV) approach. The innovative issue of our LAV solution is represented by the use of a shared ontology to build a common view of the business sector in which the cooperation takes place.

The Harmonise Approach to Semantic Interoperability

The advantages of flexibility and openness provided by the Internet in the connection of computer systems are not matched in the connection of software applications. The primary techniques aimed at application interoperability are adapters (typically in Enterprise Application Integration) and exchange formats, such as Electronic Data Interchange (EDI) or Knowledge Interchange Format (KIF). Due to the limited success of existing solutions, Harmonise proposes an approach that would start from the above but be centrally based on a domain ontology. To this end, project activities have investigated three main areas.

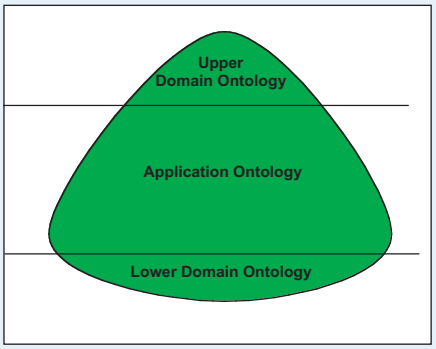
Interoperability clashes, caused by differences in the conceptual schemas of two applications attempting to cooperate. The possible clashes have been classified in two main groups:

- Lossless clashes, which can be solved with no loss of information. Examples include naming clashes, when the same information is represented by different labels; structural clashes, when information elements are grouped in a different way; and unit clashes, when a scalar value (typically an amount of money, or a distance) is represented with different units of measure.
- Lossy clashes, which include the clashes for which any conceivable transformations (in either direction) will cause a loss of information. Typical cases are information represented at different levels of granularity, refinement, or precision. For example, in expressing the distance of a hotel from the airport, one application simply reports 'near airport', while another expresses it in terms of miles, eg 'five miles from airport'. Another example is when one hotel specifies the presence of an 'indoor swimming pool', while another just says 'swimming pool'.

Ontology, which represents a common, shareable view of the application domain. This is used to give meaning to the information structures that are to be exchanged between applications. In Harmonise the ontology is based on the Object, Process, Actor Modelling Language (OPAL) representation method, which follows the Frame-Slot-Facet paradigm. It includes constructs such as ISA (with inheritance) and aggregation hierarchies, similarity and various kinds of built-in constraints

(such as cardinality constraints, enumeration). Based on OPAL, we have developed the ontology management system, SymOntoX.

The concepts in a domain ontology can be seen to be organised according to a (complex) hierarchical structure shaped like a chestnut (see Figure). In the top part (Upper Domain Ontology) we have generic concepts, such as 'process', 'actor', 'event' and 'goal'. In the bottom part (Lower Domain Ontology: LDO) we have elementary concepts, such as 'price', 'streetNumber', 'cost' and 'internet Address'. Generally, for two cooperating partners, it is relatively easy to reach a consensus on the concepts of these two parts. The difficult section is the middle part — the Application Ontology. Here concepts and definitions depend strongly on the specific application, the kind of problems addressed and the method used to solve them, not to mention the underlying technology (which often contaminates the conceptual model) and the cultural aspects. Typical concepts in this layer are 'invoice', 'customer', 'discount', 'reliableCustomer', 'approval', or more sector-dependent concepts such as 'hotelReception', 'confirmedReservation', 'advancePayment', 'lightMeal' or 'gymTrainer' (in the tourism sector, for example).



The ontology 'chestnut'.

Semantic annotation, achieved to represent the meaning of a local conceptual schema, expressed by using the concepts available in the ontology. These provide a unique semantic reference for each application wishing to expose an interface, referred to as a Local Conceptual Schema (LCS) for both exporting and importing information. In essence, every piece of information in the LCS will be annotated using the ontology content. For a given application concept (represented as a data structure in the LCS), annotation consists in identifying the LDO elements of the reference ontology and using them to define its meaning. These semantic annotations are used, at intentional level, to generate the mapping between the local conceptual schema and the reference ontology. Mapping rules (among concepts) correspond to the transformation rules at extensional level. Transformations are applied to local data to code them according to the Harmonise Interoperability Representation (HIR) used to actually exchange data..

Conclusions

The Harmonise approach to semantic interoperability is inherently different from the approaches that propose an interchange format, such as KIF. These latter approaches are 'neutral' with respect to the application domain, while HIR allows information to be exchanged by using ontology terms only. Another approach which is in the line of our proposal is PIF (Process Interchange Format). However, the main difference is that PIF proposes a format based on a predefined (limited) process ontology, which is given with the standard. Conversely, HIR is based on a rich domain ontology, the content of which is not provided by the method. It is initially built by the domain experts, and continuously evolves to keep pace with evolving reality. Accordingly, HIR evolves, since its vocabulary is (a subset of) the ontology.

Link:
http://www.harmonise.org

Please contact:
Michele Missikoff, IASI-CNR
Tel: +39 06 7716 422
E-mail: missikoff@iasi.rm.cnr.it

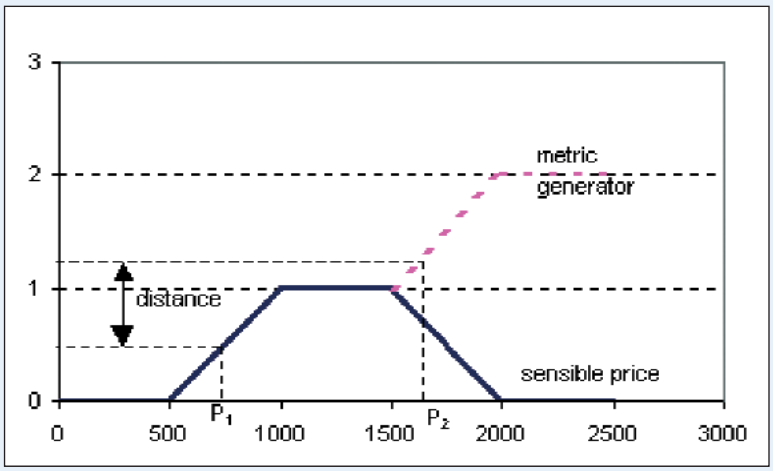
Reducing the Complexity of Information Retrieval by Weighted Querying

by Peter Vojtáš and Jan Vinař

This report describes a project which was originally applied to information retrieval from traditional (deductive) databases, and is now being extended to Semantic Web sources such as XML, unstructured text, and heterogeneous and distributed sources. The aim of the project is to find effective ways of answering vaguely formulated queries on crisp, vague, uncertain or incomplete data.

This project has been running for several years at the Institute of Informatics of Safárik University in Kosice, Slovakia (a member of SRCIM) and involves a group of institute researchers aided by PhD and Masters students.

The aim of the project was to find ways of increasing the expressivity of querying languages by various means: using background knowledge to process both the queries (formulated in language that is as natural as possible) and the source material, introducing a similarity measure on the data domains as well as metadata, and introducing new logical operators such as weighted aggregation, which in turn leads to fuzzy or probabilistic data models.



Metric generation from linguistic variables.

While largely successful on the level of formal models and experimentation, this method was soon found to create a double combinatorial explosion. First, there was the data explosion: where a deterministic database consists of a well-defined subset of the Cartesian product of data domains, the fuzzy or probabilistic database should contain practically every element of that product (since its probability or truth value will be positive, though very small).

The second problem deals with the weighting process itself: to assess the probabilistic measure of a composite event consisting of n elementary events, it is in general necessary to evaluate 2n probability assignments for these events.

It might seem that the first problem at least could be solved by assigning a threshold value for inclusion in the database. However, it is non-trivial to determine what partial threshold values will lead to a desired threshold for a composite query (which is necessary if we look, eg for k best answers or the so-called e - best answer). This is merely difficult for fuzzy operators (due to their special nature-truth functionality), but practically impossible for probabilistic operators, unless we assume the independence of the elementary events in the question, and certain other restrictions.

Two ways for solving the threshold problem were followed in our research. The first way involves the definition and study of the fuzzy join of several tables and different methods of evaluating it were examined. The phenomenon of threshold propagation through the computational process was studied and various heuristics for choosing the order of join evaluation are now being empirically tested. The second method is based on a metric induced on the original attribute domains by a fuzzy linguistic variable. Consider (see the Figure for a very simple motivating example) the trapezoidal function 'sensible price'. The metric generating function shown in the figure assigns distance to each pair of prices. This is easily generalised to a distance between sets. In this way each record (whether with crisp or vague values) is mapped to a set of distances with respect to a given elementary query, and from that to a point in n-space, where scales of different axes correspond to weights of different attributes (determined by, for example, a neural network). For choosing the best points (with respect to the 'ideal' point located in the origin) several methods of computational geometry are available.

Other methods of inducing a metric on the original data are also possible, such as that of fuzzy conceptual lattices.

The notion of metric and the corresponding similarity measure has other uses related to IR. One of these uses is the result of applying the notion to data structure. Similar data structures are mapped automatically to a unifying metastructure allowing the collection of data from heterogeneous sources to a whole. This was explored by a group working within the project and led to the notion of the 'metadata-engine', which may find uses in data mining.

An interesting special case is the mining of data from independently created XML documents. As every author may create his own ontology to describe the meaning of his XML tags, it is first necessary to find a unifying ontology on the basis of detected similarities.

At present, the loose collaboration of groups and individual researchers that we have been calling a project is passing from the stage of formal models and promising experimental results to the formation of prototype software tools. This activity is largely being carried out in cooperation with member institutions of CRCIM, such as the Institute of Computer Science of the Czech Academy of Science, the Faculty of Mathematics and Physics of the Charles University and the Masaryk University. Also participating are our Spanish colleagues from the University of Malaga.

The research reported here was partially supported by VEGA grant 1/7557/20.

Please contact:
Peter Vojtás,
P. J. Šafárik University, Košice (SRCIM)
Tel: +421 55 62 209 49
E-mail: vojtas@kosice.upjs.sk

Hypermedia Presentation Generation on the Semantic Web

by Lynda Hardman

The research goal of the Multimedia and Human-Computer Interaction group at CWI is to investigate the automated generation of Web-based hypermedia presentations tailored to the abilities, preferences and platform of the user. This requires the description and processing of different types of information in order to assemble semantically annotated media items into a coherent presentation, ie a presentation that communicates the intended semantic relations to the user.

In text, these semantics are often implicit in the text-flow. In multimedia, however, the semantic relations among the media items need to be communicated explicitly by choosing appropriate layout and style. The assumption underlying current style sheet technology (XSLT and CSS), stating that content and presentation are independent, is often an oversimplification for multimedia presentations. In addition, current style sheets operate on the syntactic XML level and are unaware of the new generation Semantic Web languages, such as RDF(S) and OWL.

Creating well-designed multimedia presentations requires an understanding of both the presentation's global discourse and interaction structure as well as the intricate details of multimedia graphic design. In addition, knowledge of a number of other factors is needed. A domain description allows the relationships among domain concepts to influence the layout and links within the presentation. Knowledge about the user's task or environment allow appropriate choices of media to be made. In conjunction with information stored in the domain model, presentations can be generated, for example, to skip things the user already knows and explain new concepts in terms of already known concepts. A description of the characteristics of the end-user platform (such as screen resolution, bandwidth, ability to display colour, audio capabilities) allows optimal use to be made of the capabilities of the device.

Our prototype generation engine, Cuypers, allows the specification of these different information types and incorporates them within the overall process of generating a presentation. In order to make the different knowledge sources explicit for use by the system we wish to use standard languages and tools. Given the range of tools available for the Web, incorporating Semantic Web tools and languages was a forgone conclusion. This allows us to re-use, for example, domain descriptions created by experts in the field, in a language with readily available processing tools. In addition, having generated a presentation, we are able to include semantic mark-up within it, thus capturing the knowledge used during its generation, eg which domain concepts media items correspond to, or platform profiles for which the presentation is suited. The current focus of our work is on creating a more realistic user model, developing a graphic design model and investigating the requirements for a discourse model.

Most work currently being carried out for the Semantic Web concentrates on the underlying semantics and is not concerned with presentation 'details'. On the other hand, most presentation tools on the Document Web operate only on the syntactic level. The long-term goal is the development of 'Smart Style', ensuring that both Semantic and Document Webs are integrated.

Based on the annotated media database of the Rjiksmuseum in Amsterdam, the Cuypers engine automatically generates multimedia presentations in SMIL in response to user queries. The figure illustrates a presentation about the painting technique 'chiaroscuro' in the context of the work of the painter Rembrandt van Rijn. The presentation consists of a slide show with examples of paintings by Rembrandt using the chiaroscuro technique, alongside a textual explanation of the technique itself. While consisting of a number of individual media items returned by the database, the presentation is intended to convey the semantic relations among them.

For example, domain-independent layout knowledge is used to place the presentation title centrally, because it applies to both the slideshow and the textual explanation. In contrast, the title of the textual explanation is left-aligned with the body text to indicate that it refers only to that paragraph. The close proximity of the text 'Self Portrait (1661)' to the image, is intended to convey that the text is a label referring to the painting. In addition, explicit knowledge of domain-specific presentation conventions allows the engine to display it in such a way that the user can interpret it as being the painting's title and the year of creation. This work is carried out in the context of the Dutch ToKeN2000 project.

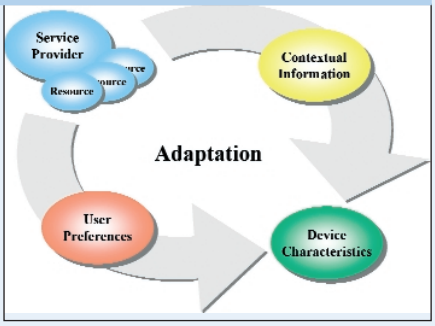
Links:
http://www.cwi.nl/ins2/
http://www.token2000.nl/

Please contact:
Jacco van Ossenbuggen,
Lynda Hardman, CWI
Tel: +31 20 592 4141
E-mail:
Jacco.van.Ossenbruggen@cwi.nl

Profile-Based Adaptability in the Semantic Web

by Santtu Toivonen

One of the key challenges of the Semantic Web is the adaptation of Web content according to diverse parameters. In addition to differences between various devices, contextual information and preferences of users should also be taken into account in order for a flexible and proactive Semantic Web to emerge. Kontti-project, carried out at VTT Information Technology, aims to contribute to this vision.



Adaptational relationships between the elements in the Semantic Web.

In the Semantic Web, content should be adapted to many levels. Some elements that influence multi-level adaptability can be identified:

- Service providers: parties that produce Web content, eg yellow pages
- Resources: pieces of content, eg electronic weather reports
- Users: parties that consume the content
- Devices: terminals that are used when accessing the content
- Context: situational details to be taken into account when accessing the content.

Service providers are any parties that produce content in the Web, such as companies, the public sector, and individuals. Resources are individual pieces of Web content offered by service providers. Resources can be described with a number of attributes, such as content type, creator, creation date, price, etc.

Users in turn are any parties that consume the resources. The most evident users are human beings, but some higher level services can also be identified as content consumers. For example, consider a software agent collecting weather reports from several information providers in the Internet. This agent consumes the content and can thereby be functionally identified as a user. However, the same agent can act as a service provider to yet higher level consumers, such as a Web portal providing, among other things, information about weather. As far as human users are concerned, typical attributes to be taken into account when selecting and adapting content are age, sex, nationality, language, hobbies, etc.

Devices are used by human beings when accessing the content. The physical details of devices such as display size and amount of memory can have an impact on the adaptation. Consider again the weather report. When delivered to a small device with black-and-white character-based display, there is no sense in trying to express the report graphically. At VTT Information Technology, the content adaptation based on device profiles has already been studied in earlier projects. Contextual details such as time and place might also affect the adaptation; for example, a mobile user would typically prefer the weather reports of the location he is at or heading to. Kontti-project concentrates particularly on the adaptation based on these contextual details.

The adaptational relationships between the above-mentioned elements are depicted in Figure 1. Resources offered by service providers are adapted to the characteristics of different terminals. Contextual details and user preferences are, however, considered before device characteristics. In this way users receive the content they wish, in an optimal format and with situational details having been taken into account.

In the Semantic Web, the most natural way to express the characteristics of the different elements shown in Figure 1 is in the form of profiles. Users, service providers, and device manufacturers can create profiles in a language, such as RDF (Resource Description Framework), which is recommended by the Semantic Web community. Depending on the usage, some selected attributes from the profiles of various elements can be matched in order to deliver personalised material to the users. Kontti-project aims to set up a framework where users can design profiles for themselves and different applications, as well as other users being able to utilise these profiles.

The intention in the Semantic Web initiative is that in addition to human beings, computer programs can process Web content, for example to perform profile matching. In order to ensure this machine-processability and interoperability between different independent parties, the profiles should be semantically bound to ontologies. Ontologies, expressed for example in DAML+OIL (DARPA Agent Markup Language + Ontology Interchange Language) or its forthcoming successor OWL (Ontology Web Language), specify the meanings of the concepts and statements given in profiles. The weather forecast provider might create a profile for the report service using DAML-S — a language for describing Web services - on top of WSDL (Web Services Definition Language), and store it in a UDDI (Universal Description, Discovery and Integration) registry for other parties to exploit.

There already exist some standardisation efforts that at least partially cover the elements of the Figure. For example, for devices there exists CC/PP (Composite Capabilities/Preference Profiles), for resources Dublin Core, for services the above-mentioned DAML-S, for users vCard, for context data Wireless Village Presence Attributes, etc. While the ontologies must be extensible and specialisable, these standardisation efforts naturally cannot cover all the features people wish to utilise in their profiles. In Kontti-project, an ontology for expressing the properties required for constructing contextual profiles is created.

Link:
http://www.vtt.fi/tte/projects/kontti/

Please contact:
Santtu Toivonen
VTT Information Technology
Tel: +358 40 839 8405
E-mail: santtu.toivonen@vtt.fi

Integrating Multimedia Components into a Semantic Web

by Benjamin Jung and Lyndon JB Nixon

The concept of separating content, transformation and presentation data is one of the key requirements for constructing a Semantic Web. It has been successfully applied in textual publishing for decades (using SGML and DSSSL) and could finally revolutionise composition and access of multimedia components such as graphics, audio and video. What role does XML play in integrating multimedia seamlessly into a Semantic Web?

During the electronic publishing process content passes through a number of conversion layers. Each layer modifies the content to adjust and customise it in more detail to create user-oriented valuable information. The reverse generalisation process is used to break up documents into basic and original components such as content, transformation and presentation data. Traditional Web formats (ie HTML) no longer allow reverse engineering of information, as all components are unidentifiably blended together.

The Table shows the four basic layers and exemplarily some of the vocabularies available. Each layer corresponds to one distinct stage in the publishing process. Vocabulary-specific processing engines are use to handle the data (eg XSLT engines to process XML documents with XSLT style sheets).

The physical storage and content datasource is represented by the Content Layer; the Transformation Layer physically changes the original content by adding, deleting (filter), modifying (process) and converting the content into a data format supported by the end user application. The Presentation Layer defines and adds style and rendering information, used by the application to sense (read, listen, watch) the final publication. The Application Layer finally represents all applications, capable of presenting information according to users' preferences.

To cater for seamless and automatic integration of local resources into a Semantic Web, content has to be available independently from its transformation and presentation data. XML in combination with XML-related vocabulary specifications (see Table) offers a solid framework to organise and keep content, transformation and presentation data separately and in the most accessible way.

The definition of content, transformation and presentation data might be obvious in the context of textual publishing, but how are these components defined in graphical, audio and video publications, to become part of a Semantic Web?

Graphics

Images using raster graphics formats such as JPEG and GIF are widely used on the Internet. They are composed of a two-dimensional grid of pixels, the basic unit of programmable colour. While this makes information retrieval in terms of colour feasible, accessing other image information such as objects, shapes and relations is ambiguous if not impossible. For purposes of decomposition, images using raster technology are comparable with badly designed HTML pages, where content and presentation data is welded together, and it is impossible to extract pure content for further processing.

One XML-based solution to describing vector graphics is the SVG format, a sequence of mathematical statements that places lines and shapes into a two-dimensional space. It clearly defines the contents of the image (mathematical objects as well as their grouping) and its transformation in terms of location, size, applicable filters and deformations. An optional CSS file keeps presentation data such as colours and margins.

Audio

Separating content from presentation data in audio files appears to be more complex than it is for textual or graphical data. Traditionally, audio was distributed to the user as a single source (eg cassette, CD). Nevertheless, professional recording studios use single tracks for each instrument/voice and compile them together before distribution. One approach with an emphasis on audio delivery over the Internet is MusicXML, which uses XML to describe sheet music. Ideally, presentation data such as speed, volume, key and instruments will be kept separate from the MusicXML content file. Possible 'presentation engines' for MusicXML would include a sheet music renderer (for print) as well as an audio player.

	Text	Graphics	Audio	Video
Application Layer	XHTML, PDF, …	SVG, CGM, …	MIDI, PDF, …	MPEG-4 (BIFS)
Presentation Layer	CSS, FO	CSS	Finale templates, ...	MPEG-4 (OD)
Transformation Layer	XSL, DSSSL, DOM, fxp			
Content Layer	Various XML vocabularies, (XML-) databases			

Publication architecture.

Video

Video is the greatest challenge for content and presentation separation. 'Shaped video' is already used by TV studios to blend video objects together ; eg in weather forecasts, the weather forecaster and the background map are generated separately. However, object extraction from the result video stream is difficult. MPEG-4 is an ISO/IEC standard for the composition of audiovisual scenes consisting of both natural and synthetic media. While the format is binary, there is an XML representation for content interchange. With MPEG-4's Binary Format for Scenes (BIFS), multiple arbitrarily shaped video objects may be composed in an audiovisual scene together with video-quality synthetic images. The presentation of each video object is defined in an 'object descriptor' (OD) which includes decoding information, synchronisation and support for scalability. As objects in a BIFS scene can be manipulated individually, MPEG-4 offers the possibility of extending media adaptation into the realm of video. Furthermore, MPEG-4 has an XML representation for content interchange, which allows it to be tightly integrated with the Semantic Web. Lyndon's research is using XML-encoded metadata to dynamically generate MPEG-4 scenes, which, using shaped video techniques, would make adaptive video possible.

XLink

XLink, the next generation linking and part of the XML family of specifications, plays an important role for seamlessly integrating multimedia components into a Semantic Web. Having used the aforementioned concepts of separation, XLink allows link sources and targets to be defined in any granularity, eg ranging from a single letter to a paragraph (text), from a single line to the grouping of objects (graphics), from a single note to a set of measures (audio) and from a single frame to an entire scene (video).

Links:
http://www.cs.tcd.ie/Benjamin.Jung/
http://www.fokus.gmd.de/research/cc/magic/employees/lyndon.nixon/

Please contact:
Benjamin Jung, Trinity College Dublin
Tel: +353 608 1321
E-mail: benjamin.jung@cs.tcd.ie

Lyndon JB Nixon, Fraunhofer FOKUS
Tel: +49 30 3463 7103
E-mail: nixon@fokus.gmd.de

Ontologies for Semantic Web Components

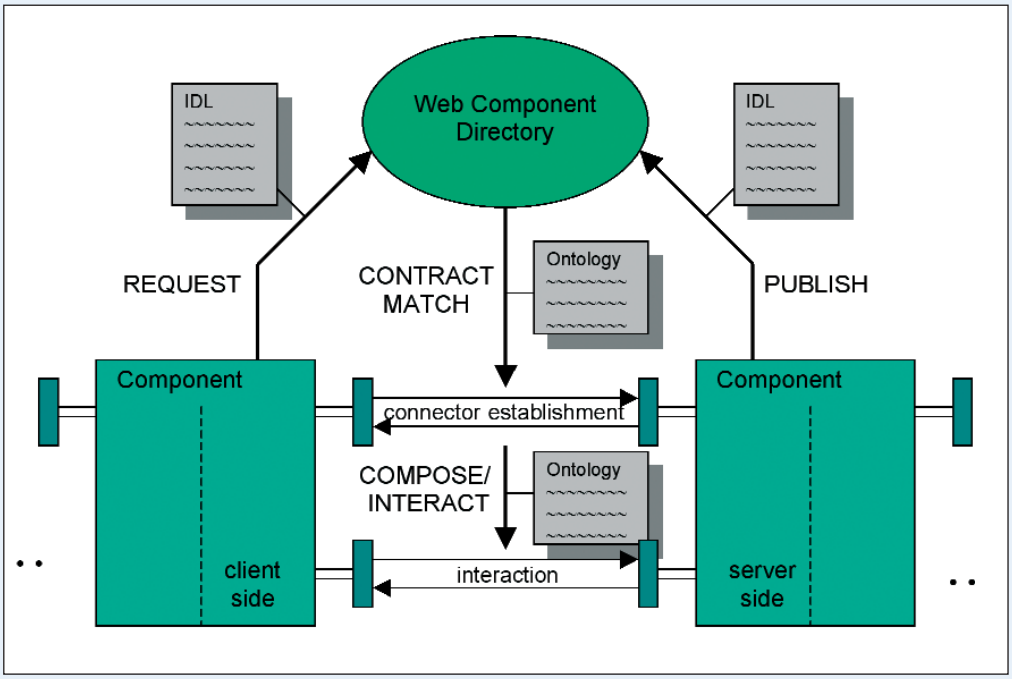
by Claus Pahl

Enabling the Web for software components can be ideally supported through the use of Semantic Web technologies. The Semantic Web activity aims to introduce meaning to the Web using ontologies. Ontologies can create a shared understanding of application domain and development process knowledge that is crucial for component development activities such as matching and connecting requestors and providers of component services.

Making the Web more meaningful and open to manipulation by software applications is the objective of the Semantic Web activity. Research at the School of Computer Applications at Dublin City University investigates Web components, which utilise Semantic Web technologies to enable component-based software engineering (CBSE) for the Web platform. Efforts in taking the Web from a document-oriented to a services-oriented environment — bundled in the Web Services Framework (WSF) — are a first step towards the Web as a software development and deployment platform. The purpose of the WSF is to define a distributed computing model for the Web and to provide languages and protocols for service description, discovery and invocation. However, adding semantics to this framework is essential if Web components instead of Web services are considered.

Providing reusable software components and plug-and-play style software deployment is the central objective of CBSE. Components are reusable, internally coherent software entities that are specified using interface description languages (IDL). Traditional IDLs focus on the syntactical description of services that are provided by the component. However, semantic annotations of components in form of semantic IDLs are essential for Web components. In order to facilitate precise retrieval and matching and to guarantee the reliable and correct execution of components, semantic service descriptions and component interaction processes need to be captured in semantic IDLs. To achieve a high degree of reusability, not only which services a component provides, but also which services a component requires needs to be explicitly specified in an interface. Our research aims to solve problems identified in relation to semantics in the specification of behaviour of component services and the component's interaction processes. Here, representing knowledge about components is as important as knowledge about the process of component development.

The Semantic Web objective to open the Web to manipulation by software applications in our case means to enable component development activities that are based on the semantical representation of software components. Using the Web as an infrastructure for components development requires support for two different activities – see Figure 1. Firstly,



Web Component Life-Cycle. A component request might by matched by a server component published in a Web component directory. Successful matching results in direct communication between client and server component to establish a connector for service invocation and reply.

discovery and retrieval functionality – possibly in form of a Web component directory or marketplace – allows providers of services and components to publish descriptions of their software entities and requestors to discover and retrieve them. A notion of contracts ties matching components together. However, unlike for a document-oriented Web, a second form of support is needed. Interaction support allows components to be composed through connections for interaction. Essentially, knowledge about all life cycle processes of components is required.

Providing meaning and a shared understanding are crucial for Web-based CBSE. Ontologies are means to define the concepts and properties of a domain and to provide facilities to reason about them. A central objective of our research is the formulation of ontologies for Web component development. An application domain ontology describes the domain of the software application under development. A software development ontology describes the component development entities and processes. Ontology technology integrates notions to describe components and their behaviour and an inference system to reason about composition. These ontologies create a shared understanding between software developers and software agents that allow a developer to describe provided or required components and agents to reason about matching of suitable components and to control the composition and interaction of components.

Classical formalisms that we have used to represent component-related knowledge are process calculi and modal logics. Ontologies for Web-based component development, however, have to adopt Web standards and technologies. DAML+OIL is a Web ontology language. DAML+OIL can be seen as a very expressive description logic. Knowledge representation techniques suitable for the classical use of the Web are, however, not necessarily directly suitable for software components and software development, for example if a process- or transition-oriented view on components is taken. Our approach to Web components is based on a description logic geared towards the representation of component properties and the support of reasoning for matching. To enable modal reasoning about components and their behaviour in a description logic framework is a central objective of our research. A correspondence between description logics and dynamic logic – a modal logic of programs – that relates roles and transitions allows us to integrate classical software engineering techniques into a Semantic Web environment.

Link:
http://www.computing.dcu.ie/~cpahl/CompTech.html

Please contact:
Claus Pahl, Dublin City University
Tel: +353 1 700 5620
E-mail: Claus.Pahl@computing.dcu.ie

Concepts on Enriching, Understanding and Retrieving the Semantics on the Web

by Arne Sølvberg, Sari Hakkarainen, Terje Brasethvik, Xiaomeng Su, Mihhail Matskin and Darijus Strasunskas

The COEUR-SW program at the Department of Computer and Information Science at NTNU in Norway seeks to capture the heart of the Semantic Web.

The Information Systems on the Web comply with the semiotic triangle. In the COEUR-SW triangle, a concept in a Universe of Thought is related to an uttered symbol in a Universe of Language, the symbol is related to a referent in a Universe of Structure, and the referent is related back to the concept. The concept, the symbol and the referents are related to a context in the Universe of Discourse (UoD), as embedded in the Web. The COEUR-SW program approaches the UoD from three interrelated angles and thus seeks to capture the heart of the Semantic Web.

The Web has three major roles in Information Systems; one role as the dominant medium for information dissemination, one as the tool for information compilation, and one as the evolving information repository. The roles comprise enterprise and user interface issues (UoT), information categorisation and interpretation issues (UoS), and information storage and information access issues (UoL).

In order to support dissemination and compilation of information, the repositories may be organised according to the principles of integration or interoperability. In the former, the relevant information repositories are viewed as one large distributed database, organised according to a common schema. In the latter, every information system is considered to consist of autonomous subsystems.

The integration principle is dominant in the classical approaches to information systems design. The alternative principle of semantic interoperability is more compliant with the Semantic Web. For interoperable systems to communicate there must be some agreement on coding as well as meaning of data, but only for the data that is interchanged. The mutual understanding must be as wide as it is necessary for a particular data interchange, but need not be wider.

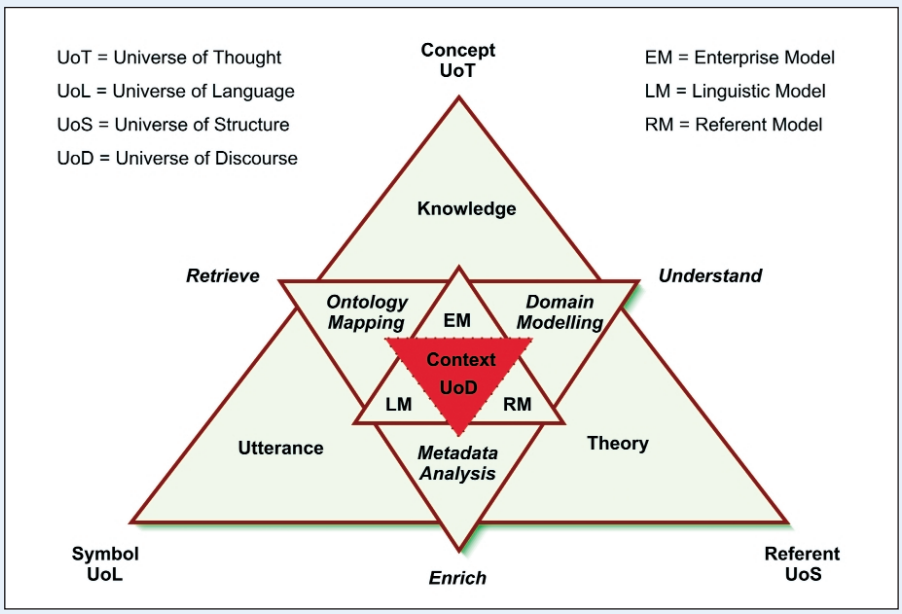
Semantic Interoperability

Data semantics is the relationship between data and what the data stand for. In order to obtain mutual understanding of interchanged data, the actors have to share a model of what the data represent. Semantic interoperability is about how to achieve such mutual understanding.

In order to achieve this we use concepts and symbols. The concept is the unit of thought. The symbol is the unit of language. Theories, such as Newton's theory of motion, are structures of thoughts, and referents like force, mass and acceleration are the units of these structures. Conceptual knowledge comes wrapped in symbols, eg mathematical notation, words or diagrams, which are the linguistic expressions of knowledge. In order to access the ideas of other people, ie the concepts in their UoT, we must understand the conceptual structures that are employed in the common UoS. Furthermore, we need to understand the relationship between the symbols and the ideas they stand for.

Agent Systems

Agents in a multi-agent system are characterised by abstraction, interoperability, modularity and dynamism. These qualities are particularly useful in that they can help to promote open systems, which are typically dynamic, unpredictable, and highly heterogeneous, as is the Internet.



The COEUR-SW program at IDI, NTNU in Norway, approaches the UoD from three interrelated angles. Domain modelling is used in order to capture the knowledge within a UoT into a man/machine understandable theory in a UoS. Ontology Mapping is used in order to capture the utterances in a UoL into man/machine-retrievable knowledge. Metadata analysis is used in order to capture the theories in a UoS into enriched man/machine generative utterances in a UoL. In all the approaches inter-related conceptualisation of the UoD is used in order to capture the heart of the Semantic Web.

Within a multi-agent system, agents represent their 'view of the world' by explicitly defined ontologies. The interoperability of such a multi-agent system is achieved through the reconciliation of these views by a commitment to common ontologies that permit agents to interoperate and cooperate. However, between different agent systems, ontologies are developed and maintained independently of each other. Thus two agent systems may use different ontologies to represent their views of the domain. This is often referred to as an ontology mismatch. In such a situation, interoperability between agents is based on the reconciliation of their heterogeneous views.

Web Resources

Recently, several general-purpose models for describing Web resources have emerged. The intention of the industry-driven initiatives is to provide a metadata description framework for interconnected resources. A resource can be viewed as a bibliographic document (DC, MARC), electronic Web resource (RDF, HTML) multimedia object (SMIL), ontology (OIL, DAML), database concept (MDIS, OIM), or case tool structure (XMI). Many of the models result from standardisation work as carried out in industry-driven coalitions like W3C, MDC, and OMG.

Emerging academic research work has adapted and applied some of the models provided by the industry coalitions. Metadata is recognised as equally important for describing the components from which we build services and systems, regardless of whether they are on the Internet or not. The research so far has focused on technical aspects of the frameworks, semantic coverage of the models, and reasoning and querying mechanisms. The area is becoming mature enough for comparative studies and prescriptive theories on the proposed modelling frameworks.

Activities

The overall intention of the COEUR Semantic Web research program at IDI, NTNU is to explore and exploit existing research results in the areas of conceptual modelling, information systems analysis and design, agent technology, and information systems architectures. Our attacking point is semantic interoperability. Our main armoury consists of enriched descriptions and representations for each of the Universes of the Semantic Web triangle. Our ultimate target is to enable information systems development for Semantic Web applications in areas such as information services, e-commerce, knowledge management, and cooperative systems.

Currently, several activities are under way within this program:

- the Referent Modelling Language project, RML
- the Ontology Mapping Using Text Categorisation project, OMUT
- the Semantic Modelling of Documents project, SeMDoc
- Conceptual Metadata Analysis and Design in Information Service Development, COMAD
- the Adaptive Distributed Information Service project, ADIS
- the Traceability in Cooperative Systems Development project, CoTracSyDev.

Links:
http://www.idi.ntnu.no/coeur-sw/
http://www.idi.ntnu.no/coeur-sw/rml/
http://www.idi.ntnu.no/coeur-sw/omut/
http://www.idi.ntnu.no/coeur-sw/semdoc/
http://www.idi.ntnu.no/coeur-sw/comad/
http://www.idi.ntnu.no/coeur-sw/adis/
http://www.idi.ntnu.no/coeur-sw/cosy/

Please contact:
Arne Sølvberg, NTNU
Tel: +47 7359 3438
E-mail: asolvber@idi.ntnu.no

Negotiating on Your Own Terms!

by Declan O'Sullivan and Vincent Wade

In Multimedia and Hypermedia Systems the Knowledge and Data Engineering Group at Trinity College Dublin has pioneered research in virtual Web-based environments, adaptive hypermedia systems for customised or personalised user experience, usability of Web-based learning, metadata-driven (XML based) approaches for developing re-usable digital resources and XML-based digital resource discovery services in a global search context.

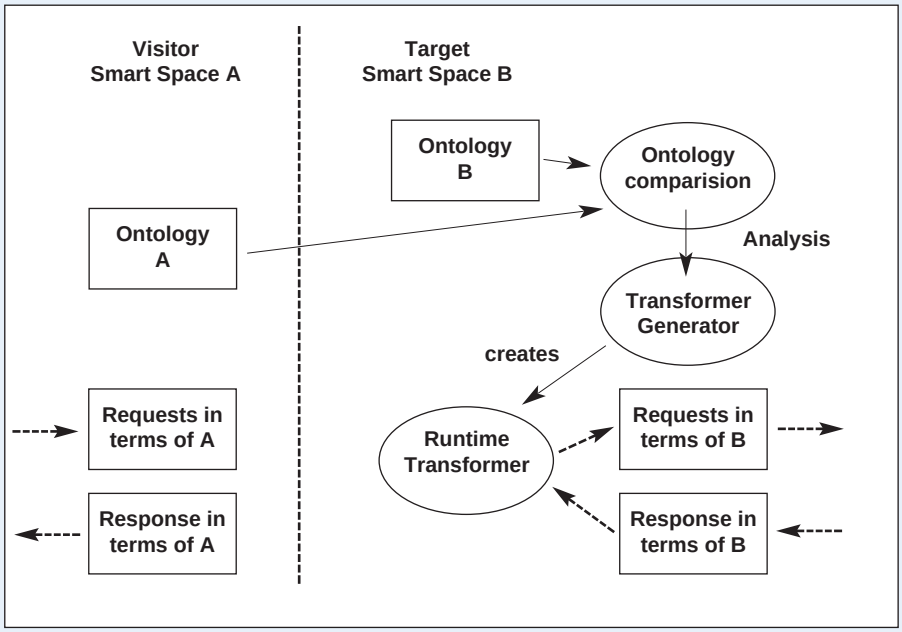
Achieving semantic interoperability when interconnecting information systems of different parties continues to be a difficult problem. Moreover, a solution that allows for semantic interoperability to be achieved at runtime through dynamic discovery of meaning and dynamic translation between parties is increasingly needed. In particular, such a solution could have a profound impact on how e-business is conducted. For example, each party in a B2B relationship would conduct business using the natural terminology of the business, leaving the translation between vocabularies to be automatically undertaken by the system. For example, one party might call the date for delivery 'Due Date' with the format 'Day/Month/Year' and the other party calls it 'Delivery Date' with format 'Month/Day/Year'. Such interoperability issues have been dealt with in the past at schema design time (with distributed database solutions), through the hand-crafting of interoperability gateways (with system integrator solutions) or by forcing each business to conform to a standard vocabulary (eg ebXML). Of course, the difficulty comes as the nature of each business and their respective ontologies evolve over time. More and more what is required is the runtime comparison of ontologies

and the dynamic construction of gateways.

Research that is being undertaken within the Knowledge and Data Engineering Group of Trinity College Dublin looks at finding such a solution, and draws upon ongoing research into the Semantic Web, ontologies, electronic negotiation, and dynamic adaptation of software systems. The chosen application area, which provides a rich source of requirements, is the management of ubiquitous computing/smart spaces environments. A Smart Space is a physical space rich in devices and software services that is capable of interacting with people, the physical environment and external networked services. The aim of the Smart Space is to bring tangible benefits to people in support of their tasks and activities. Take for example, a cardiac surgeon who is woken by her alarm clock. Upon sensing that she has got up and is having a shower, the home smart space sends messages to the coffee machine, toaster and fruit press appliances to have breakfast ready. Her PDA is activated and downloads any emergency case patient files that may have arrived overnight, and alters the day's appointment and work schedule if necessary. Upon leaving the house, the home management system configures the intruder alarm, the call forwarding/answering service and utility management services accordingly. The primary objective of smart space management is the dynamic runtime adaptation of smart space devices and software services to provide the necessary support for people's tasks and activities. In addition it involves the lifecycle and operational management of the devices and services offered by the smart space.

Today's computers interact with users in terms of low level abstractions – applications and individual appliances. A smart space will take over most low level management of applications and appliances so that the user interacts in terms of 'activities' (that represent user intent). In this scenario, it cannot be assumed that ordinary users will limit their expressions to a set of standardised activities. Indeed this would run contrary to the intent of smart spaces to support users to perform their activities seamlessly and efficiently. Where a user is confined to a single operator domain, it can be foreseen how the smart space can learn which user-specified activities can be mapped to which smart space services, and which of those services can be executed simultaneously without conflict at runtime, in accordance with policies set for that space. However, in order to enable the free movement of users between smart spaces, information must be made available to the visited smart space regarding which equivalent or adapted services the user will want to avail themselves of, and what quality of service will be expected/agreed to. Each smart space will want to negotiate using the activity/service and contracting terminology that is most familiar. This results in the need to 'bridge' between the terminology of each party during the negotiation of the necessary service level agreements. However, given the wide diversity of users, smart space capabilities and smart space operators envisaged in the future, the negotiation of such service level agreements needs to be automatic, with human intervention being a very exceptional case.

One solution being researched in the group involves a platform that uses Semantic Web technologies to dynamically bridge the terminologies of the two parties at runtime. First the ontologies of the two parties (represented in XML Topic Maps) are compared. The comparison of the ontologies is then used to derive XML stylesheets that transform the negotiation from one party's terminology to the other party's terminology and vice versa. This comparison and dynamic bridging is undertaken at runtime and requires no a priori knowledge of the party requesting the service level agreement. Key to our approach is that each ontology needs only to adhere to a small subset of common concepts, with the rest of the ontology being defined to reflect the terminology most natural for the particular smart space. The approach is illustrated in the Figure.



Overview of Smart Space Mobility Negotiation tool.

This work is undertaken as part of the Irish government HEA funded M-Zones project. The M-Zones Research Programme is a multi-disciplinary, inter-institutional research programme that engages in fundamental research in Management and Control systems for integrating Multiple Smart Spaces.

Link:
http://kdeg.cs.tcd.ie

Please contact:
Declan O'Sullivan or Vincent Wade
Trinity College Dublin
Tel: +353 1 608 1765
E-mail: Declan.osullivan@cs.tcd.ie,
Vincent.wade@cs.tcd.ie

Articles in this Section

gViz: Visualisation Middleware for E-Science: A New Project in the UK E-Science Programme

by Julian Gallop

E-science is intended to enable global collaborations over the Internet and provide access to very large, distributed resources, including data collections, computation and networks of specialised experimental facilities. A new project, funded by the UK E-science Core Programme aims to research and develop visualisation middleware for e-science.

Computer-assisted visualisation is already a key technology for understanding large-scale simulations and observations and, when appropriately configured, will enable collaborations of researchers to analyse the results of evolving e-scientific resources. To be successful at this, visualisation technology needs to be adapted to the evolving Grid, which underpins e-science. gViz, funded by the UK E-science Core Programme started on August 2002, aims to research and develop visualisation middleware for e-science.

Situations

Using the Grid presents new situations for anyone developing visualisation software:

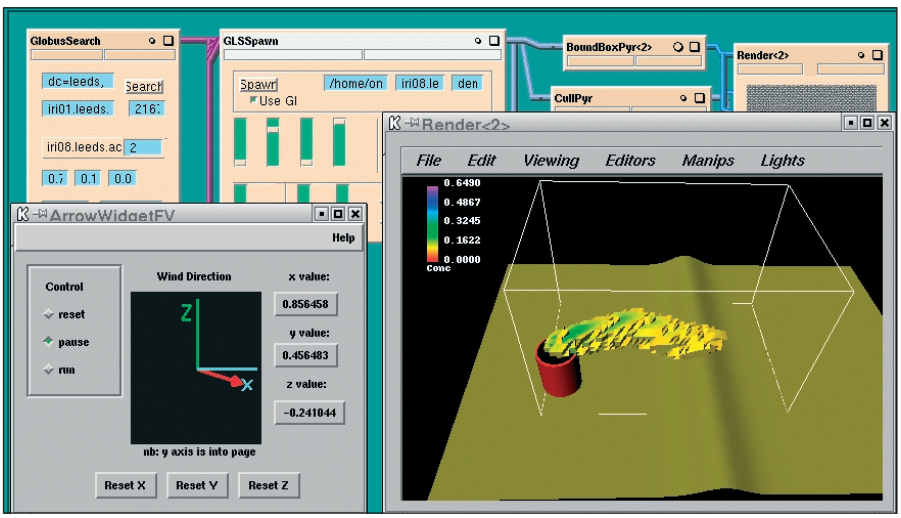
- An intensive computation on the Grid may generate large results datasets. To be viable, some visualisation algorithms would also need to be run on the Grid.
- Visualizing a large distributed dataset or multiple heterogeneous ones requires access to Grid data management interfaces.
- The existing models for collaboration need to be extended. A virtual organisation may involve dynamic collections of people who may alternate between virtual meetings and asynchronous communication. Also the virtual organisation may not be able to enforce one specific visualisation system.
- There will be a need to integrate visualisation software with emerging Grid technologies, such as GSI to handle authentication, Grid replica management to handle large data files and other Globus facilities.
- Visualization frameworks — whether dataflow networks such as Iris Explorer or interpretative programming languages such as PV-Wave or Java programming systems such as VisAD — can include computational modules when the whole framework is on a single processor or a fixed, limited distributed system. This leads us to ask whether there is a possibility that the framework can control computation modules executing on the Grid.

Work Packages

There are three main subdivisions of work:

- two workpackages aim to grid-enable two very different visualisation systems, Iris Explorer and pV3, which are both exemplars of useful system classes
- another workpackage aims to develop compression techniques suitable for transmitting very large streams of data
- the remaining workpackage aims to build up appropriate XML languages for describing the structure and format of data to be analysed and for processing tasks. In addition to the technical definitions, this workpackage will, early in the project, run a workshop on XML for visualisation.

The Figure shows some earlier work carried out at the University of Leeds (the lead partner in the new project) as part of a previous UK e-science demonstrator project (see http://www.visualization.leeds.ac.uk/CovisaG/). This shows at a basic level that it is possible to communicate with the Grid using the framework provided within an off the shelf visualisation system. The top of the figure shows some modules connected together, two of which are wrappers for Grid services. The bottom of the figure shows a rendering window which shows a case study of hazardous pollutants. This work was demonstrated at the opening of the UK National e-Science Centre in April. The new project will need to go on to exploit Grid capabilities including security, flexibility, symmetry between collaborators and extensibility.



Screenshot of a session using the Iris Explorer visualization system accessing a Grid application.

Also relevant to the new project is previous work on component-oriented visualisation carried out by CLRC RAL and the University of Lancaster (project Visual Beans (http://www.acu.rl.ac.uk/VisualBeans/).

Applications
The intention is to develop techniques and tools which are usable across a wide range of applications. However within the project, certain applications will be used as testbeds and these include heart modelling (in computational biology) and climate prediction.

Partners
The academic partners are the University of Leeds (lead partner), the University of Oxford, Oxford Brookes University and CLRC Rutherford Appleton Laboratory.

In accord with the funding rules, the industrial partners make, in kind or directly, a total contribution which matches the funding received from the e-science programme. These partners are IBM UK, Nag, and Streamline Computing.

Followup
This new project may be of interest to those who are developing Grid applications which require visualisation or those who are developing more advanced Grid tools and believe that visualisation software may make use of them. It would be useful to hear from European researchers who are interested in this project.

Links:
http://www.bitd.clrc.ac.uk/
and follow links to R&D
then Visualization Middleware.

Please contact:
Julian Gallop, CLRC
Tel: +44 1235 44 5569
E-mail: Julian.Gallop@rl.ac.uk

Dealing with Massive Volumetric Visualisation: Progressive Algorithms and Data Structures

by Rita Borgo, Valerio Pascucci and Roberto Scopigno

This article summarises recent results obtained in the area of visualization of massive datasets exploiting the strength of subdivision techniques and distribution algorithms. The focus is on trying to meet storage requirements designing visualisation algorithms suitable for distributed environments.

Projects dealing with massive amounts of data need to carefully consider all aspects of data acquisition, storage, retrieval and navigation. The recent growth in size of large simulation datasets still surpasses the combined advances in hardware infrastructure and processing algorithms for scientific visualisation. As a consequence interactive visualisation of results is going to become increasingly difficult, especially as a daily routine from a desktop. The visualisation stage of the modelling-simulation-analysis activity, still the ideal effective way for scientists to gain qualitative understanding of the results of simulations, then becomes a bottleneck in the entire process. This problem poses fundamentally new challenges, both to the development of visualisation algorithms and to the design of visualisation systems.There is a need at system level to design the visualisation process as a pipeline of modules processing the data in stages. In this way a flow of data is created that needs to be optimised globally with respect to magnitude and location of available resources. To address these issues we have been working on a new progressive visualisation algorithm that subdivides the input grid following regular criteria and then traverses the regular structure produced extracting the cells containing isovalues. These cells are then organised in a hierarchical structure (from coarse to fine levels) and subsequent levels of detail are constructed and displayed to improve the output image.

We separate the data extraction from its display. The hierarchy is built by a single process that traverses the input 3D mesh. A second process performs the traversal and display. The schema allows us to render partial results at any given time while the computation of the complete hierarchy is progressing. The regularity of the hierarchy makes it possible to create a good data-partitioning schema that allows us to balance processing time and data migration time. The subdivision criteria is itself suitable for parallelisation. The algorithm works in the following way:

Subdivision: we generate a partition of the input data. Figure 1b shows the 2D edge-bisection refinement sequence that we adopt to sub-sample a 2D regular grid. The coarse level is basically triangular mesh. Each refinement step inserts a new vertex on an edge and splits the adjacent triangles along this edge into two. Instead of 'triangles' we reason in terms of '2D diamonds' considering each cell as the composition of the two triangular halves that will be split by the bisection. For the 2D case we can subdivide the diamonds into two main classes: first class square-shaped diamonds (or type 0 diamonds, Figure 2a), and second class rhombus-shaped diamonds (or type 1 diamonds, Figure 2b). Each diamond is characterised by a centre (the centre of the bisection edge).

In 3D, the bisection edge refinement becomes a schema for the subdivision of tetrahedral meshes. It applies to tetrahedral meshes in the same way that it applies to 2D triangular meshes. With respect to the 2D case it maintains all the properties. Each cell is still subdivided

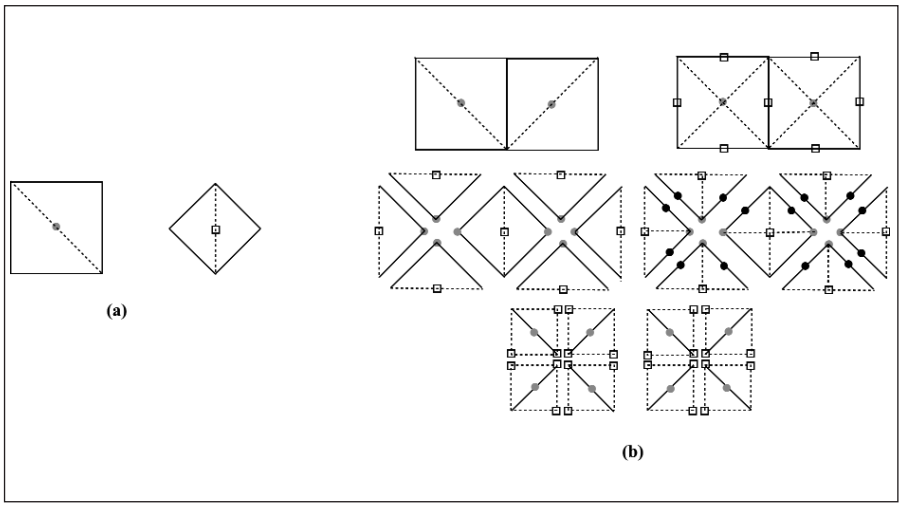


Figure 1: a) 2D First and second level diamonds; b) 2D subdivision.

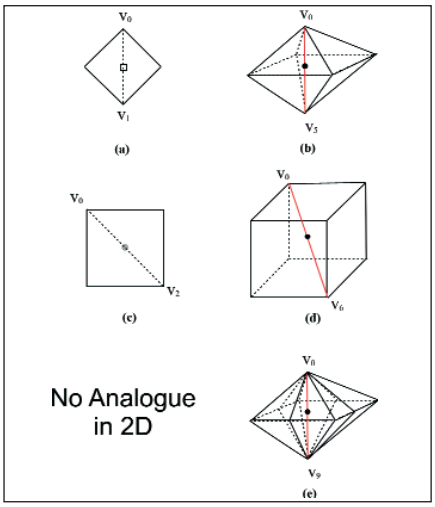


Figure 2: 2D diamonds vs. 3D diamonds.

along 'the longest edge', which corresponds to the main diagonal of the cell itself. Each bisection gives birth to an arbitrary number of diamond-like new cells. For regular grids, the starting cell is a first class cube-shaped diamond cell (see Figure 3a). The bisection of a first class diamond cell gives birth to second class octahedral-shaped diamond cells (Figure 3b-c). Each new second class diamond cell (Figure 3d) is bisected along its main diagonal and third class esahedral-shaped diamond cells are created (see Figure 3e-f). Each esahedral cell (Figure 3g) is bisected along its longest diagonal and generates first class diamond cells (Figure 3h-i). In 3D, we need to repeat the bisection procedure three times to reach a stage in which a new bisection gives birth to first class cube-shaped cells, returning to the subdivision step for the starting cell (see the third level diamond in Figure 2e) .

Hierarchical Organisation: The subdivision schema implies a hierarchical organisation of the dataset. It is easy to organise such a hierarchy in a tree-like data structure and to extract a sort of 'seed set', made up of cells whose internal range of isovalues includes the isovalue target. The seed set generated corresponds to an adaptive traversal of the mesh at different levels. The hierarchy always starts with a first class diamond cell and proceeds through each level of refinement with an alternation of second, third (for the 3D case) and first class diamond cells.

Isosurface Extraction: Each generated diamond internally holds a 'piece' of the total isocontour that needs to be extracted. To perform the extraction, for the 3D case, we subdivide each diamond cell into tetrahedra and apply a marching tetrahedra algorithm. Each isocontour is updated within a single tetrahedron and then composed to update the global isosurface within the set T of all tetrahedra around the bisection edge.

The key novelty of the present schema is that, by providing a set of local rules for continuous geometric transitions (geomorphs) of one level of resolution into the next, we bridge the gap between adaptive techniques and multiresolution decimation-based techniques. Moreover, the regularity of the schema permits the design of an efficient run-time data partitioning and distribution algorithm which will reduce local memory requirements and will make use of the potential of the currently under-exploited distributed environment.

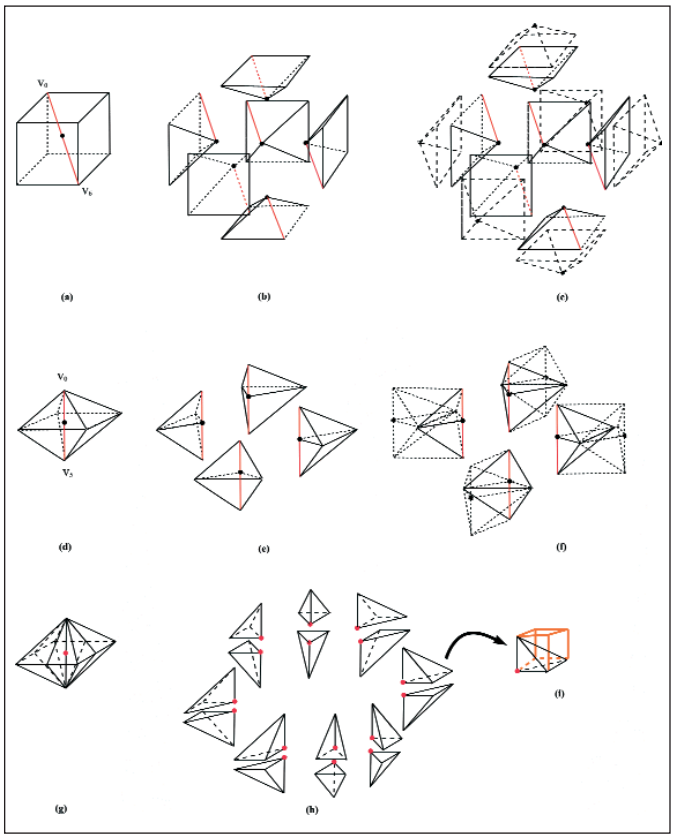


Figure 3: First, second and third level diamonds and respective subdivision and fusion with adjacent cells.

Please contact:
Rita Borgo
Pisa University and ISTI-CNR
Tel: +39 050 315 3471
E-mail: borgo@di.unipi.it

Roberto Scopigno, ISTI-CNR
E-mail: r.scopigno@cnuce.cnr.it

Valerio Pascucci
Lawrence Livermore
National Lab, USA
E-mail: pascucci@llnl.gov

RADIUS: Looking for Robots' Help in Computer Science Research and Education

by Jozef Kelemen and Aleš Kubík

RADIUS is a project at the Institute of Computer Science at the Silesian University in Opava, Czech Republic, which reflects the shift from viewing computation as an execution of a sequence of basic computational steps, towards viewing it as the activity of a community of interacting basic computational entities. The acronym RADIUS is inspired by the robot of the same name from the play 'R.U.R' (Rossum's Universal Robots) by the Czech writer Karel Capek.

RADIUS, based on the experience of the project participants in such varying scientific areas as formal language theory, artificial intelligence, artificial life, natural computing, multi-agent systems and agent-based economies, reflects this paradigmatic change in two ways; research in robotics and decentralised intelligence underlying systems, and the incorporation of ideas from studies in this field into the computer science and applications curricula.

Our results and experiences in the theory of grammar systems (GSs), a recent area of formal language theory, have proven that formal languages can also be described and studied in terms of (finite) collections of grammars rather than individual grammars. This theory provides us with an effective framework for research in fields such as multi-processing, decentralisation, cooperation and emergence of collective phenomena, and can thus be viewed through the lens of the above new computational paradigm.

In addition to research, grammar systems as a part of Masters and doctoral curricula have been presented by Alica Kelemenová and Jozef Kelemen at Czech and other European universities (Silesian University, Opava, University at Hradec Králové, both Czech Republic, University of Economics at Bratislava, Slovakia, Vienna University of Technology, Austria, and Rovira i Virgili University, Tarragona, Spain). A tutorial was delivered by Jozef Kelemen on the connection of a new, non-functional modularisation of systems and grammar systems at the 9th ECCAI Advanced Course on AI in Prague, 2-13 July 2001.

Figure 1: The Khepera robots.

We have supervised Masters and PhD theses in GSs; for example, a formal framework and computer simulation as a tool for experimentation have been used in order to prove emergence of some phenomena in agent-based computational economies in the PhD dissertation of the second author of this article.

Besides investigating multi-agent systems as grammar systems, we have experience with teaching agent-based software engineering and the theory and practice of multi-agent systems as a regular Masters-level university course. A collection of lecture notes prepared by Ales Kubík and published by the Silesian University for this purpose was the first comprehensive introductory textbook on the topic published in the Czech language.

In spring 2002, we set up a laboratory for experimentation in information-processing aspects of collective robotics. It is equipped with K-team Khepera robots. We study societies of robots as open information systems of physically embodied agents executing tasks in a shared dynamic environment. For programming the robots we use C and Java, with a Lisp interface for controlling them remotely.

RADIUS is focused on studying the emergence of complex behaviours from the interactions of simple agents, in order to make emergence more tractable through the design and implementation of multi-agent architectures on the theoretical ground provided by grammar systems. Proven capabilities of formal tools can thus be experimentally tested in real environments. In conjunction, we study how the behaviour of a group of robots can evolve from an initial specification, and how to design robot controllers using neural nets, interactive Turing machines, biologically inspired computing etc.

Their modularity and easy-to-use interface make Kheperas suitable for teaching distributed programming and the information-processing base of robotics, and for teaching basic computer programming in a non-traditional manner, stressing the ideas of Stein's new modularity rather than the traditional functional modular decompo-

Figure 2:
The authors playing with the robots.

sition. By programming robots, students must interact with autonomous units that are driven by the controller, with several processes with 'endless' loops that compete for the robot control. Students face the problem of programming real devices equipped with sensors and effectors which interact with the dynamic environment. This helps them to realise not only the particular characteristics of programming open information systems, but the limitations – which from our perspective is the most important point – of the object-oriented analysis and design of information systems in comparison with agent-based analysis and design of distributed applications or even robotic societies.

It is intended that RADIUS will continue to be used in the formulation and solution of further problems of new modularity, collective robotics, artificial life, computational economies and other fields within the area of grammar systems. It will also find employment in the experimental verification of emerging hypotheses (we plan to build our own peripherals in order to enhance the sensory/effectory capabilities of robots), and in incorporating new ideas into our programming courses, making them more effective, attractive and interesting for our students.

Looking forward to further common work in order to achieve these goals, we acknowledge the long-term, intensive and fruitful cooperation with our partners, especially with the Research Group on Modelling Multi-Agent Systems, SZTAKI, and the above-mentioned universities.

Link:
http://ui.fpf.slu.cz

Please contact:
Jozef Kelemen
Silesian University, Opava, Czech Republic
Tel: +420 653 684 361
E-mail: kelemen@fpf.slu.cz

Ales Kubik
Silesian University, Opava, Czech Republic
Tel: +420 653 684 367
E-mail: kubik@fpf.slu.cz

Consensus Creation on Universal Access in Health Telematics

by Constantine Stephanidis and Demosthenes Akoumianakis

IS4ALL (IST-1999-14101) 'Information Society for All' is an EC-funded Thematic Network seeking to establish a wide, interdisciplinary and closely collaborating network of experts (Working Group) to provide the European Health Telematics industry with a code of practice detailing how to appropriate the benefits of Universal Design.

The objectives and technical approach of the IS4ALL project have been presented in ERCIM News Nos. 46 and 43. This article focuses and elaborates on the project's activities that aim to facilitate transfer of know-how and know-why regarding Universal Access to potential recipients in the Health Telematics industry.

To this end, the perspective of the IS4ALL project is three-fold. Firstly, there is a commitment to consolidate and codify currently available knowledge and best practice on Universal Access by means of a continuous data collection effort. In a second stage, methods and techniques for Universal Access are validated by reference to designated scenarios and use cases from the domain of Health Telematics. Finally, a broad range of outreach activities are planned to facilitate awareness raising and consensus creation on the issue of Universal Access amongst Health Telematics practitioners. The prime target of these outreach activities is to facilitate inter- and intra-sectoral transfer of knowledge with the aim to:

- increase Health Telematics practitioners' propensity to innovate as a result of process improvements resulting from the adoption of Universal Access practices
- guide them towards the appropriation of the benefits of innovative applications and services following the principles of Universal Access.

In devising suitable mechanisms to facilitate transfer of know-how and know-why, the commitment is towards support measures intended to promote in the Health Telematics sector a favourable environment for the exchange of knowledge, experience and practice on the issues pertaining to Universal Access. The prime focus is on mechanisms that foster networking and learning by doing. Two such clusters of mechanisms have been defined to facilitate collaboration and dissemination of information, respectively.

Collaboration-Oriented Mechanisms

IS4ALL promotes active collaboration with players in the Health Telematics sector by:

- Formally inviting key representatives from the Health Telematics sector to become involved as partners in the IS4ALL consortium. In this way, the IS4ALL project has acquired strong links with user communities, such as MS-HUGe, as well as industrial networks such as EHTEL. Additionally, these representatives provide the project with the collective wisdom, requirements and needs of the Health Telematics sector, thus focusing technical work and 'screening' the results

so as to ensure their relevance, comprehensiveness and applicability.
- Initiating and sustaining direct contacts with target user communities, which may be national Health Telematics providers and networks, industries, or consortia, in an effort to elicit requirements and formulate reference scenarios, so that Universal Access methods and techniques are validated. At the same time, the target user communities gain feedback on their technical developments in the form of usability reports, recommendations for design improvements, participation in seminars and access to information and project deliverables.
- Organising thematic meetings in the form of Focus Groups or workshops. Experts are invited to review progress and technical achievements, and to help the IS4ALL project consolidate and refine its findings and fine-tune the presentation of its technical achievements. Thus far, the project has organised one Focus Group seeking to collect information on Universal Access methods, techniques and tools, while three workshops are planned to take place by September 2003 to refine and finalise the project's final deliverable, which is a comprehensive code of Universal Access practice in Health Telematics.

Information Dissemination Mechanisms

The outreach activities of the IS4ALL project amount to organising and extending invitations for participation to seminars, tutorials and workshops, preparing and periodically updating project presentation materials, and maintaining the project's website. Seminars constitute a prime instrument for disseminating information in a timely fashion and are delivered, free of charge, in the context of important events organised by members of the Health Telematics sector (ie European and international conferences, workshops, etc).

IS4ALL has already delivered three seminars, the content of which varied to reflect progress and technical achievements, focusing respectively on:
- presenting Universal Access methods as adopted by practitioners in other sectors of the industry, such as industrial engineers, architects and developers of interactive applications and services
- elaborating Health Telematics reference scenarios to ensure their validity and relevance
- illustrating the methods in practice through designated validation exercises.

IS4ALL plans to organise three more seminars by September 2003. For more information on the IS4ALL project and scheduled activities, please visit the project's website.

Link:
http://is4all.ics.forth.gr

Please contact:
Constantine Stephanidis, ICS-FORTH
Tel: +30 810 391741
E-mail: cs@ics.forth.gr

A Body-Monitoring System with EEG and EOG Sensors

by Mária Bieliková

A team of four students from the Slovak University of Technology in Bratislava has developed a prototype mobile sleep laboratory — a body-monitoring system with EEG and EOG sensors. The students explored the concept of intelligent data collection from human body sensors, and with this system, won third prize at the annual IEEE Computer Society International Design Competition (CSIDC 2002) World Finals.

The Body-Monitoring System (BMS) is designed as a mobile device that is able to collect measured data and to act according to instructions set by a supervisor. The system consists of a body-monitoring network (see Figure 1). In order to recognise the monitored person's state, the monitor unit connects to various body sensors and i/o devices using either wired or wireless communication technologies. Data from all sensors is collected, stored and analysed in real time and, according to the analysis, actions may then be performed. A computer is used as an interface to the body-monitoring network, and developed software allows a supervisor to configure the monitor unit for the monitored person, to connect sensors and i/o devices, define and upload instructions for monitoring and download collected data.

The monitor unit software consists of a communication module (responsible for connecting and controlling sensors, and for gathering and pre-processing measured data), a storage module (for storage of collected data), and a policy interpretation module (responsible for controlling the behaviour of the monitor unit according to instructions defined by a supervisor).

Two types of drivers are introduced. The role of a communication driver is to hide the way in which data is transmitted. There is one driver for every type of communication interface, eg a Bluetooth driver or an IEEE 802.11b driver. The communication driver does not care about the data itself; this is the role of device drivers. Each type of sensor has its own device driver. When a device driver receives a message from one of its sensors it decodes the message and informs the policy engine about the state

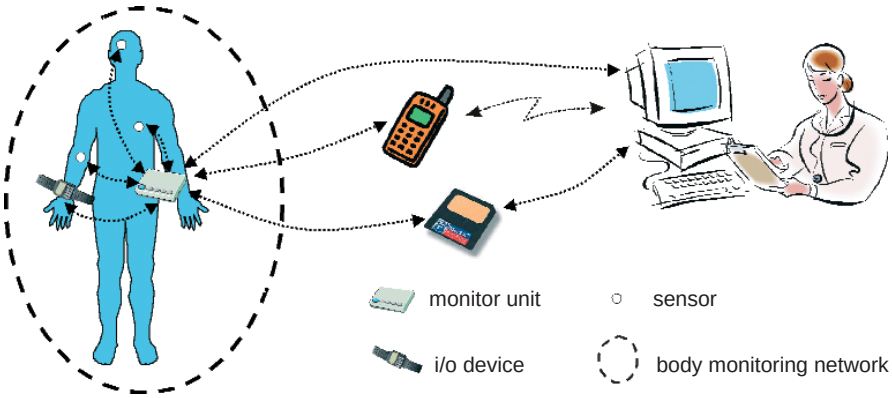


Figure 1: Overview of the Body Monitoring System.

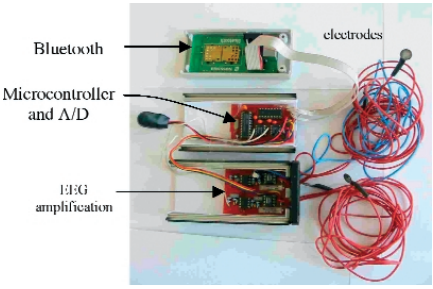


Figure 2: EEG sensor protoype.

of the sensor. To send/receive a message to/from a sensor, the device driver uses the corresponding communication driver.

The behaviour of the monitor unit is controlled by interpretation instructions defined by a supervisor (called policies). A policy describes the monitor unit's response to events reported by sensors. Policies are written in the Policy Markup Language (based on XML). It enables general policies to be written using 'virtual' objects that are replaced by physical devices in actual runtime. The monitor unit prototype is implemented in Microsoft-embedded Visual C++ 3.0 on Compaq iPaq Pocket PC HR3870.

To validate the system design the students tested it in a specific field of medicine – sleep research. To cope with the problem of sleep disorders, sleep laboratories in hospitals are used to monitor patients overnight. However, patients are influenced by the hospital environment, and usually show different sleep patterns to patients at home. As a solution to this problem a prototype mobile sleep laboratory was developed for home use. The prototype employs an electroencephalograph (EEG, which monitors brain waves), an electrooculograph (EOG, which monitors eye movement) and a thermometer. Analysis of EEG and EOG data allows identification of all sleep stages.

Figure 3: The EOG sensor in action.

Sensor implementation goes out from a common sensor platform designed in the course of this project. The common platform contains a detecting element, amplifiers and filters, an AD converter, a microprocessor and a Bluetooth module. Figure 2 depicts the EEG sensor prototype. The developed EEG sensor could also be used for continuous EEG examination. Such an examination is necessary for patients suffering from epilepsy.

The most distinctive features of the described project are:

- it provides wireless communication between various sensors attached to the human body
- it provides a unified way of controlling several sensors by a single monitor device. Sensors may also influence each other, eg it is possible to start monitoring blood pressure using a manometer as soon as an ECG (electrocardiograph) sensor reports a heart problem
- it is possible to adapt monitoring to the patient's state, ie measurements and alert messages are controlled and provided according to the current context
- it allows complex schemas for sensor control to be defined and represented in a specialised language called PML (Policy Markup Language), which is powerful enough to describe mutual dependencies and cooperation among several sensors, and even to define complex schemas for taking medicine
- it provides a secure way to transmit and store measured data
- it can be used for monitoring almost any interesting body parameters, since the design supports different sensors using practically any communication technology.

Although the students concentrated on monitoring the human body for medical purposes, the design of the Body-Monitoring System could also be used in many other fields (eg pulse rate monitoring in sports science, prevention of Sudden Infant Death Syndrome and monitoring of people working in dangerous environments).

Acknowledgement

I would like to thank our students Peter Blstak, Matus Horvath, Peter Lacko and Marian Lekavy for their excellent project work and presentation at the CSIDC World Finals in Washington, DC. The work reported here was partially supported by the Slovak Science Grant Agency, grant No. G1/7611/20, and by the Computer Society IEEE via the Computer Society Design International Competition 2002.

Links:
SUT CSIDC web page:
http://www.dcs.elf.stuba.sk/csidc

CS-IEEE CSIDC web page:
http://www.computer.org/csidc

Please contact:
Mária Bieliková
Slovak University of Technology (SRCIM)
Tel: +421 2 602 91 473
E-mail: bielik@elf.stuba.sk

WebRemUSINE: A Tool for Usability Evaluation of Web Applications

by Laila Paganelli and Fabio Paternò

Creating a Web site allows millions of potential users with various goals and knowledge levels to access the information that it contains. For this reason, interest in usability evaluation of Web sites is rapidly increasing.

We have developed a method and an associated tool to detect usability problems in Web interfaces through a remote evaluation where users and evaluators can be separated in time and space. Our approach combines two techniques that are usually applied separately: empirical testing and model-based evaluation. The reason for this integration is that models can be useful to detect usability problems but their use can be much more effective if they can be related to the actual use of a system. Our tool is able to analyse the possible inconsistency between the actual user interactions and the task model of the Web site that describes how its concrete design assumes that activities should be performed. To support remote evaluation, we have developed a technique that makes it possible to record user actions during a site visit. The analysis of the log data is based on a comparison of the traces of actions performed with the structure of the task model. This analysis provides evaluators with a number of results that are related to the tasks that users intend to perform and the Web pages and their mutual relationships.

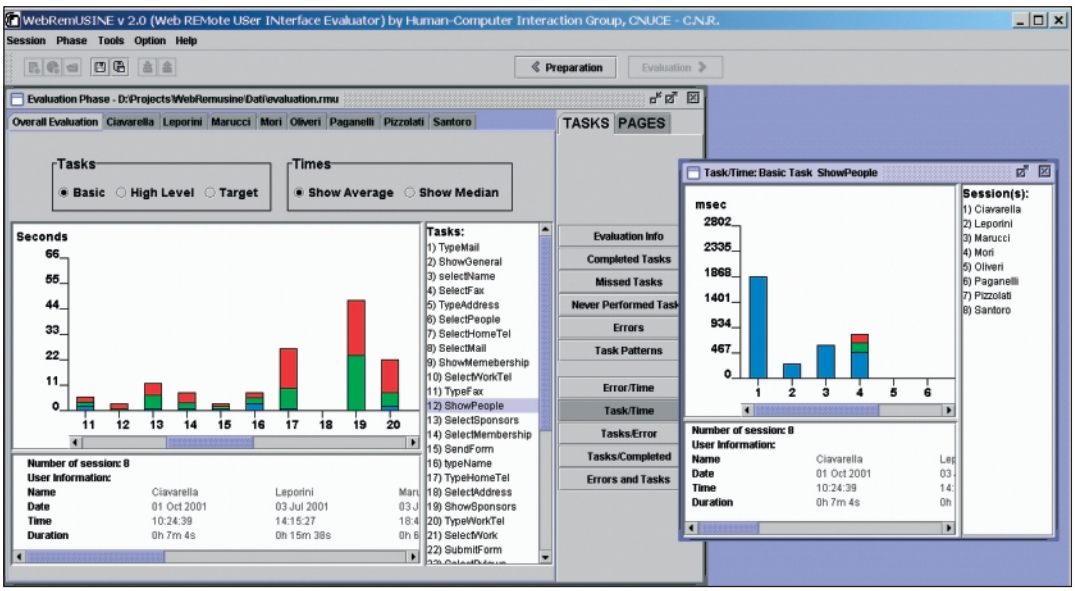
In our case we follow a hybrid approach because our environment is able to analyse data relative to user interactions and then compare them to the task model corresponding to the design of the Web site. The solution that we have adopted in WebRemUSINE to identify user intentions is to display the high-level tasks that are supported by the Web site asking the user to indicate explicitly what task they want to perform. WebRemUSINE compares the logs with the task model and provides results regarding both the tasks and the Web pages supporting an analysis from both viewpoints.

The method is composed of three phases: *preparation*, which consists of creating the task model of the Web site, collecting the logged data and defining the association between logged actions and basic tasks; *automatic analysis*, where WebRemUSINE examines the logged data with the support of the task model and provides a number of results concerning the performed tasks, errors, loading time; and *evaluation*, the information generated is analysed by the evaluators to identify usability problems and possible improvements in the interface design.

The environment is mainly composed of three modules: the ConcurTaskTrees editor CTTE) (publicly available at http://giove.cnuce.cnr.it/ctte.html); the logging tool that has been implemented by a combination of Javascript and Java applet to record user interactions; and WebRemUSINE, a Java tool able to perform an analysis of the files generated by the logging tool using the task model created with CTTE.

Link:
http://giove.cnuce.cnr.it/webremusine.html

Please contact:
Fabio Paternò, ISTI-CNR
Tel: +39 050 315 3066
E-mail: fabio.paterno@cnuce.cnr.it



WebRemUSINE interface.

Digital Factories, Production Networks

by László Monostori, Géza Haidegger, József Váncza and Zsolt János Viharos

Today's complex manufacturing systems operate in a changing environment rife with uncertainty. The performance of manufacturing companies ultimately hinges on their ability to rapidly adapt their production to current internal and external circumstances. The Hungarian national R&D project, led by SZTAKI, combines the concepts of intelligent manufacturing systems and digital enterprises in addressing these problems.

Manufacturing systems of our epoch work in a rapidly changing environment full of uncertainties. Besides internal factors such as malfunctions and breakdowns, the main external reasons for uncertainty are:

- fast-increasing and diversified customer demands
- the growing role of one-of-a-kind production, and rapid sequences of new tasks
- an increase in the number and speed of communication channels
- the appearance of new technologies
- changes made by partners (suppliers, distributors, customers, purchasers)
- market instability (eg the sudden changes of raw material prices).

Increasing complexity is another characteristic that is evident both in production processes and systems and in enterprise structures.

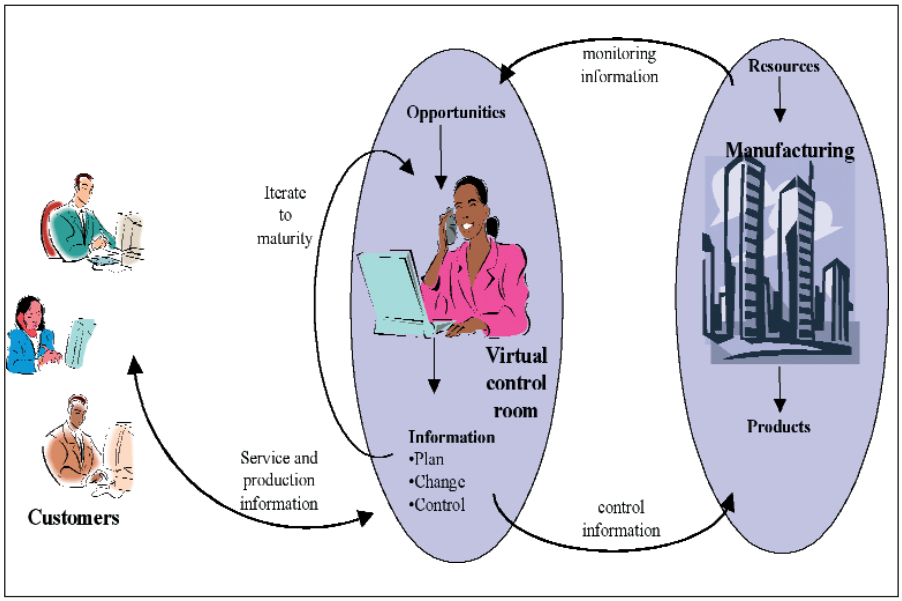


Illustration of the concept of the NRDP project on digital enterprises.

The concept of the digital enterprise, that is, the mapping of the important elements of enterprise processes by means of information technology tools, gives a unique way of managing the above problems. However, the optimal or nearly optimal exploitation of the huge amount of information available is impossible without the effective application of the methods and tools of artificial intelligence. In this way, systems can be developed that are expected to solve, within certain limits, unprecedented and unforeseen problems on the basis of even incomplete and imprecise information. This is the early concept of Intelligent Manufacturing Systems (IMS) by the late József Hatvany, from 1983.

The project partners for 'Digital Enterprises, Production Networks', which is supported by the National Research and Development Program (NRDP) in Hungary, form a well-balanced 'academia-industry' cluster. They include GE Hungary Rt, a big manufacturing enterprise, and MT-System Ltd, an information technology SME, both of which represent industry, and the Budapest University of Technology and Economics, the Miskolc University and SZTAKI, representing academia.

The Figure illustrates the concept of a digital, distributed enterprise representing the framework for the research. The vision of Digital Enterprises is to provide the capability to 'Manufacture in the Computer'. It incorporates a modelling and simulation environment that enables the fabrication and/or assembly of products including the associated manufacturing processes to be simulated by the computer. This capability takes into account all the variables in the production environment, from shop-floor processes to enterprise transactions. Effectively, it accommodates the visualisation of interacting production processes, process planning, scheduling, assembly planning, and logistics from the resource to the enterprise level.

The main characteristics of the project correspond to the above main tendencies:

- unified management of technical and economical problems (including concurrent engineering of products and production systems and the vertical integration of technical design and technological levels)
- purposeful, effective application of up-to-date computerised methods to optimise and control rapidly changing complex production structures in an environment full of uncertainties

- integrated application of tele-presence and interactive multimedia techniques towards the establishment and industrial application of virtual control rooms (VCRs), and the involvement of customers in the different phases of design and production.

The project has the following — partly overlapping — main directions, which will be treated in a comprehensive way:

- management and scheduling of large-scale projects
- tele-presence and interactive multimedia
- monitoring of complex production structures.

This national R&D project aims to make all the important production-related information available and manageable in a controlled, user-dependent way by the efficient application of information and communication technologies. The development and application of intelligent decision support systems will help enterprises to cope with the problems of uncertainty and complexity, increase their efficiency, join in production networks and to improve the scope and quality of their customer relationship management.

Please contact:
László Monostori, SZTAKI
Tel: +36 1 279 6159
E-mail: laszlo.monostori@sztaki.hu

GenoStar: A Bioinformatics Platform for Exploratory Genomics

by François Rechenmann

As a direct consequence of the spectacular progress of experimental methods and devices, the diversity and volume of biological data are steadily increasing. In this context, the GenoStar consortium has developed an integrated, interactive and easy-to-use computer platform which helps the biologists to turn these data into knowledge.

The GenoStar consortium was created at the end of 1999. Partly supported by the French ministry of research, it brings together four partners: two biotech companies, Hybrigenics (Paris) and GENOME express (Grenoble), and two research institutes, in biology, the Pasteur Institute (Paris), and in computer science, INRIA Rhône-Alpes (Grenoble). At the end of May, the consortium presented the first version of its platform, which will be made available to academic users by the end of this year.

The GenoStar platform has been designed according to a specific view of the genomic world as a complex network of biological entities and their relationships. As a very simple example, a gene and a protein can be seen as entities linked together through the 'is-coding-for' relationship. The gene is related to its chromosome, the chromosome to its species, and so on. More abstract entities can be described in the same way, such as a set of homologue genes of different species or the components of a eucaryotic gene, ie the (coding) exons and the (non-coding) introns.

GenoStar is structured into four modules: GenoCore, GenoAnnot, GenoLink and GenoBool. As indicated by its name, GenoCore is the central module that provides the various services needed by the other three application modules. It is thus in charge of the management of the entities and their relationships within a knowledge base relying on an advanced entity-relationship model; it offers persistence, query, and editing facilities.

Data, described as values of the attributes of these entities, must however be analysed with the help of adequate methods, which are provided by the application modules. Each of the three existing application modules, GenoAnnot, GenoLink and GenoBool, plays a different role in accordance with the general view of the genomic world supported by GenoStar.

GenoAnnot allows the biologist to identify regions of interest in a genomic

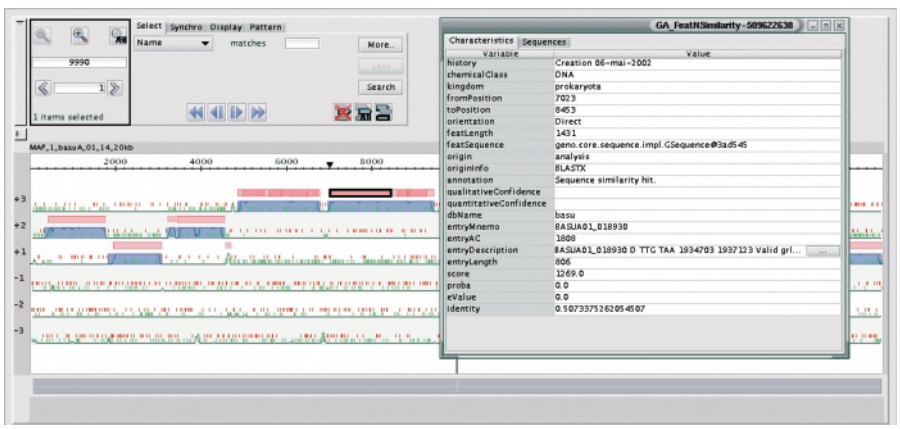


Figure 1: The results produced by several sequence analysis methods are simultaneously displayed along the genomic sequence.

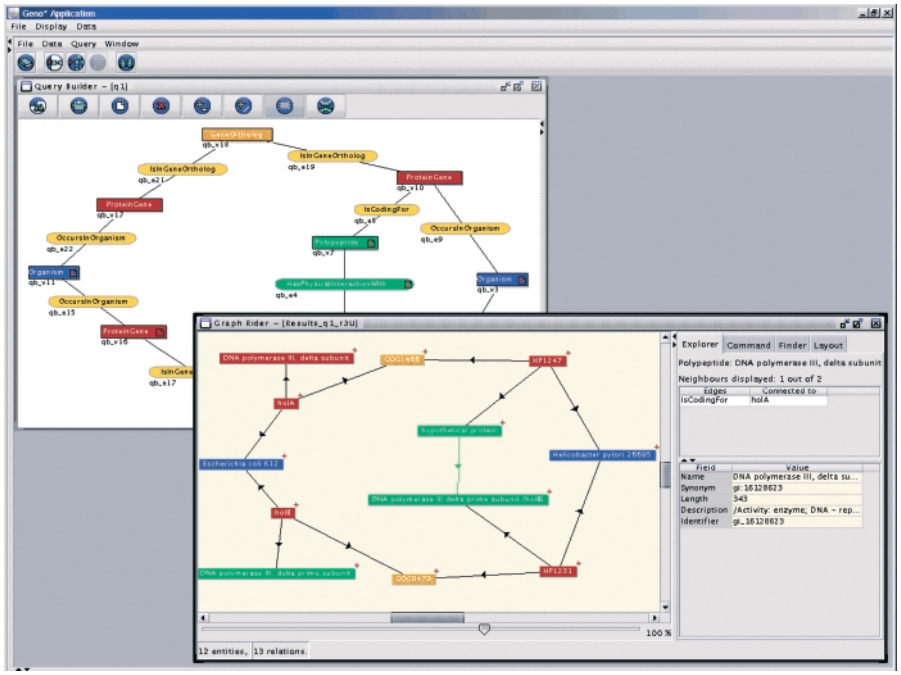


Figure 2: A GenoLink query is a partial network of entities and relations which is matched against the whole knowledge base.

sequence. An important example of such a region is a gene, which can be seen as a portion of the genome which contains the information required by the cell machinery to make a protein. In the present version of GenoStar, GenoAnnot provides several methods for identifying genes in procaryotic (bacterial) genomes. Clearly, each time a gene is identified, a corresponding instance of the gene class is created and adequately related to the other entities. The results of the identification methods can be displayed on one-dimensional maps of the genome (see Figure 1). The biologist can then easily compare the predictions of concurrent methods and eventually make a decision on their validation.

In GenoStar, a knowledge base is populated by entities and relations, which are either imported from external databases or produced by GenoAnnot. The GenoLink application module allows the biologist to explore this network of entities. Complex queries can be expressed as partial networks that are searched for in the network of the knowledge base (see Figure 2). Through this exploration process, the biologist may be able to infer new relations between previously unrelated entities. A typical example of such an inference is the prediction of the function of a gene (ie the function of the protein for which the gene codes), using information of other genes to which it is related by some pertinent relations.

The use of GenoAnnot and GenoLink results in the addition of instances to existing classes: GenoAnnot creates instances of entity classes, GenoLink creates instances of relationship classes. To complete the exploratory mission of the GenoStar platform, the GenoBool module provides the user with data analysis methods that can enable the identification of new classes of pertinent entities or relationships (Figure 3). GenoBool offers various ways to transform data before applying classical data analysis methods.

As stated in the introduction, progress in experimental methods and tools is leading to the emergence of new types of data, which obviously require sets of adapted methods. GenoStar has been designed so that it is easy not only to add methods and strategies to the existing modules, but also to add new modules.

At the present time, the next important step is to deliver a version of GenoAnnot dedicated to the analysis of eucaryotic genomes. This is still an open problem in bioinformatics due to the existence of very large intergenic regions and to the structure of the genes, which are made up of exons and introns. This version, together with important improvements and extensions in the other modules, will allow the consortium to deliver next year a commercial version of the GenoStar platform.

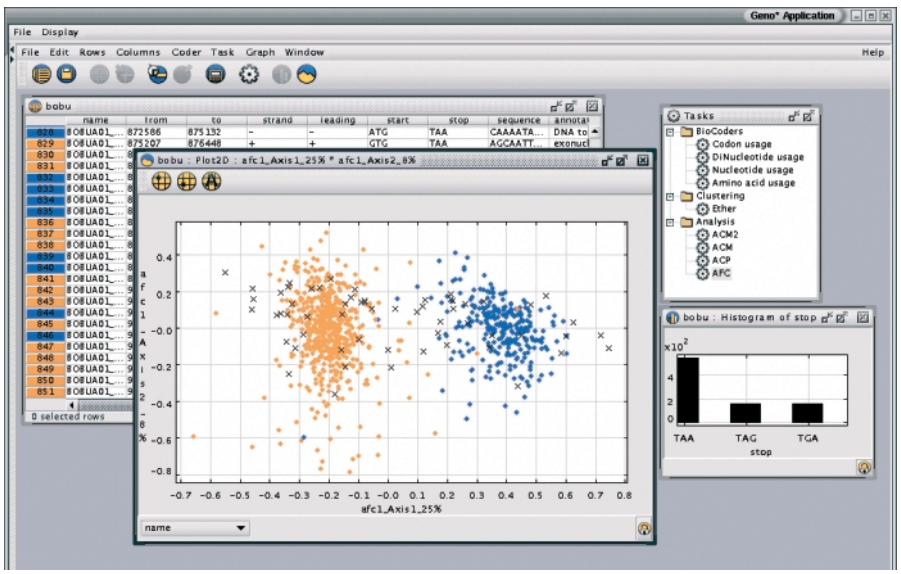


Figure 3: GenoBool is the 'data-mining' module of GenoStar. Properly encoded data are submitted to data analysis methods.

Link:
http://www.genostar.org

Please contact:
François Rechenmann, INRIA
Tel: +33 4 76 61 53 65
E-mail: Francois.Rechenmann@inria.fr

sponsored by ERCIM

SEKE'02 Conference Report

by Filomena Ferrucci

The fourteenth International Conference on Software Engineering and Knowledge Engineering (SEKE'02 conference), sponsored by ERCIM, was held in Sant'Angelo d'Ischia, Italy, from 15 to 19 July 2002. Ten years after the 1992 conference, which was held in Capri, the enchanted location offered by the isle of Ischia has been chosen to allow participants to enjoy a unique scenario for scientific discussion.

The general chairs of this edition were Genny Tortora and Shi-Kuo Chang, while the program chairs were Filomena Ferrucci and Giuliana Vitiello. The international program committee of the conference comprised 55 representatives from 15 countries. The conference was organised by the Dipartimento di Matematica e Informatica of the University of Salerno in co-operation with ACM/SIGSOFT.

The SEKE conference is a premier forum for academic and industrial researchers and practitioners in the fields of software and knowledge engineering. Continuing its long-term tradition, the aim of the conference is to provide a fruitful forum of discussion about present and future issues for researchers and practitioners of Software Engineering and Knowledge Engineering.

The main conference topics are exemplified by the talks given by the invited speakers: Lionel C. Briand illustrated why and how the two worlds of Software Engineering and Knowledge Engineering can both gain by better communication, Witold Pedrycz discussed the synergies emerging between Software Engineering and Computational Intelligence, Xindong Wu's survey covered the Data Mining techniques that can be used to distill knowledge from data, and Paul Dantzig described recent evolutions in the architecture and design of high volume Web sites.

A surprising number of researchers answered to the Call for Papers. 171 papers were submitted from all over the world touching different topics of the Conference. A peer review process was applied, each paper being reviewed by three independent and qualified referees. As a result, the technical program was organised in 26 sessions and included 71 full papers, 19 short papers, and 6 posters. There were sessions concerned preeminently with the field of software engineering. The corresponding papers touched classic software engineering issues, from requirements engineering to validation/verification, but also newer, emerging topics that saw a large number of contributions.

Indeed, many authors proposed solutions to some of the new challenges posed by hypermedia and the World Wide Web while others addressed applications in fields such as education/training or computer-supported cooperative work. The papers in the session on artificial intelligence approaches to software engineering dealt with several issues, from problem understanding by knowledge modeling through the use of case-based reasoning for domain analysis, to support systems for the generation of integration tests. There were sessions which provided contributions on knowledge engineering tools and techniques, on knowledge representation and retrieval, and on soft computing. Most of the applied studies were featured in two sessions, one on system applications and experience, the other on industrial applications. These studies reported on the application of Software Engineering and Knowledge Engineering techniques, tools and processes to a specific domain or a significant software project.

Along with the regular sessions, the program of this edition included two thematic workshops. Both workshops addressed topics of current and relevant interest — one, chaired by Athula Ginige, on Web Engineering; the other, conducted by Günther Ruhe, on Software Engineering Decision Support. Moreover, a half-day tutorial was delivered by Colin Atkinson. This seminar presented a model-driven (UML-based) approach to component engineering. The conference also hosted a European session where representatives of the EU illustrated the guidelines for the preparation of projects for the 6th Framework Programme of the European Commission.

The conference attracted over 160 participants from 23 countries and all continents. There were many students and young researchers among the participants. A relaxed and friendly atmosphere fostered additional 'off the record' informal discussions among the participants, and the social program included a Pizza Welcome Reception, a sightseeing boat trip and gala dinner, and a trip to Capri.

The proceedings of SEKE'02 were published by ACM Press and contain papers from the invited and contributed talks. A special issue of the International Journal of Software Engineering and Knowledge Engineering (IJSEKE) will be dedicated to a selection of the best papers from the Conference. The purpose of the special issue, to be printed in 2003, is to offer to the authors of such papers an opportunity to present a more detailed description of their work. Guest editors of this IJSEKE special issue are Filomena Ferrucci and Giuliana Vitiello.

Information about the Conference and the program is available at http://www.dmi.unisa.it/seke02

Please contact:
Filomena Ferrucci,
Università di Salerno, Italy
Tel: +39 089 965 273
E-mail: fferrucci@unisa.it

The Trading Agent Competition – TAC 2002

by Joakim Eriksson and Sverker Janson

The finals of the third annual Trading Agent Competition were held on 28 July 2002, co-located with AAAI-02 in Edmonton, Canada. The actual games took place on the Internet, with the game and auction servers running at SICS in Kista, Sweden. The agents resided at the home locations of the participating research groups.

Trading Agents

A trading agent is a computer program that acts on online markets on behalf of one or more clients, trying to satisfy their preferences. In the trading agent competition (TAC), agents are travel agents with the goal of assembling travel packages in a five-day period, from flight tickets, to hotel rooms, and event tickets. In a game (Figure 1), eight agents, each representing eight clients, compete for a limited amount of goods on a total of 28 different auctions, for 12 minutes. The agents receive a score based on how well client preferences are satisfied (sum of client utilities) and the average score over a number of games determines the winner.

TAC combines a fairly realistic model of the Internet commerce of the future with challenging problems in automated reasoning and decision making:

- Relative to their clients, TAC agents have to solve a combinatorial assignment problem, where goods available to an agent are packaged into bundles, and delivered as such to the clients. This problem is related to that of an auctioneer determining winning bids in a combinatorial auction. Successful trading agents must therefore make use of similar integer programming or constraint programming techniques.
- Relative to other agents, TAC agents instead have a strategy problem. To determine which strategy to use in a competitive situation is not easy. Essentially, there is no known way to compute the best course of action. One important strategic component in TAC is estimating auction closing prices.

Good software trading agents can potentially handle more complex combinations of goods, larger numbers of goods and markets, a wider range of market types, and faster markets with more fine-grained goods than their human counterparts.

TAC Software

For TAC-02, SICS developed open-source TAC server software, freely available to all TAC competitors as well as to other researchers, students, and educators in the field. The previous two competitions, TAC-00 and TAC-01, were played on the Michigan Internet AuctionBot platform, a configurable auction server available as an Internet service but not available as downloadable software.

The SICS TAC server (Figure 2) is written using a combination of SICStus Prolog and Java and consists of two main subsystems, the Game Server, running the actual games and communicating

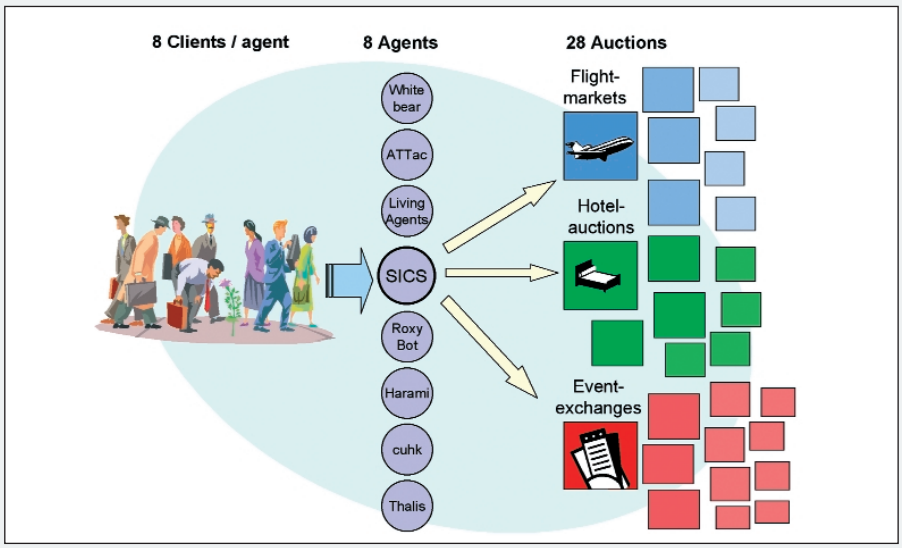


Figure 1: A TAC game has 8 agents, each with 8 clients, and 28 simultaneous auctions.

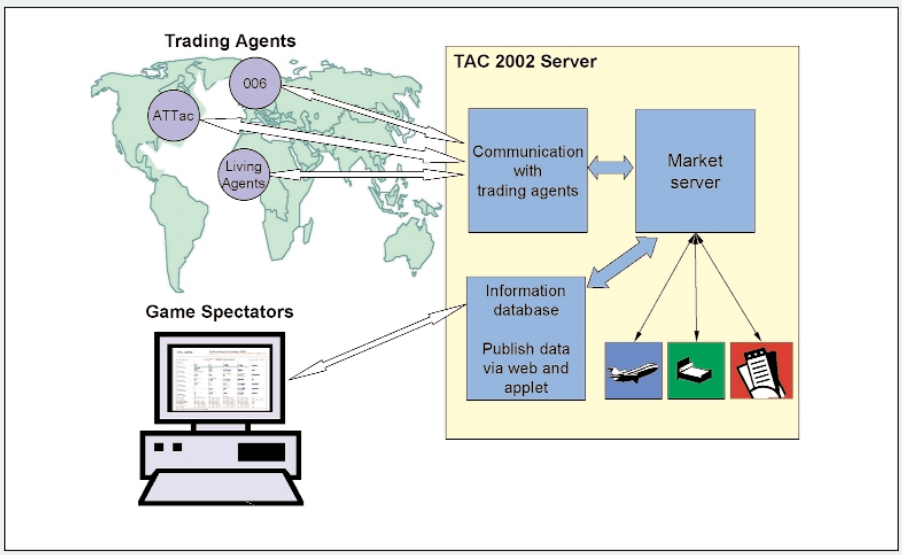


Figure 2: SICS TAC server and clients.

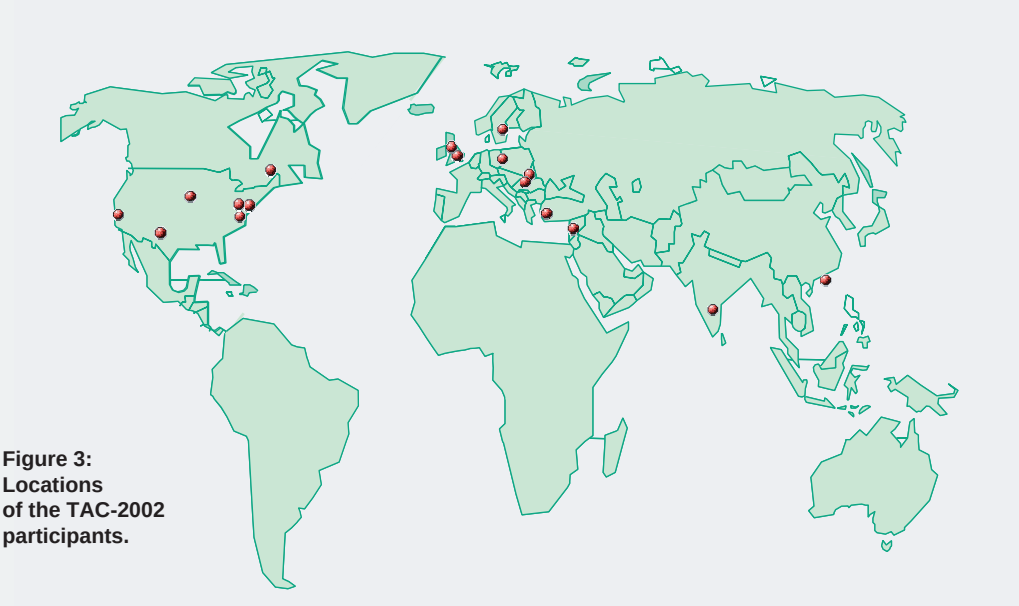


Figure 3: Locations of the TAC-2002 participants.

with trading agents, and the Information Server, for managing participants, games scheduling, and collecting and distributing game information and statistics via the Web and game monitoring applets.

SICS also developed new TAC agentware, a Java software toolkit for developing TAC agents, making it easy to get started. Most of the new TAC entrants based their agents on the SICS TAC agentware.

TAC-02 Games and Agents

The winner in this year's tight field of competitors was 'whitebear', developed by Ioannis A. Vetsikas at Cornell University with the the University of Southampton as a close runner-up.

The TAC-02 finals were the conclusions of several weeks of qualification and seeding rounds involving thousands of games, with different combinations of agents and different random client preferences:

- Qualification: select top 16 out of 26 agents for the semi-finals (120 games/agent)
- Seeding: seed agents into two semi-final groups (440 games/agent)
- Semi-finals: select top 4 agents of each group for the finals (14 games/agent)
- Finals: determine winner (32 games/agent).

The number of games in the finals is quite small, due to the desire to run them in a single day. This increases the element of chance in the game and is likely to be changed in future competitions.

Each year most teams build their agents on the ideas and solutions of the previous year's top-performing agents. Methods, models, and technologies used in TAC-02 agents include:

- Linear Programming – for finding optimal set of travel packages, many inspired by the optimiser used by TAC-00 winner, ATTac
- Machine Learning – automatic learning of closing prices on hotels, trading policies for entertainment, etc
- POMDP, Partially Observable Markov Decision Processes
- Genetic Algorithms – for finding an optimal set of travel packages
- Price Prediction – various ways of predicting price including 'Walrasian competitive equilibrium'.

Many of the agent teams did not have time to fully implement their ideas and approaches to the TAC problem. There are still many untested ideas, which will guarantee interesting results for coming TAC events.

The future of TAC

Following the 2002 finals, discussions about the future of TAC have been intense. The competition has increased in popularity, and is widely recognised as a both fruitful and entertaining means of addressing important issues in the field of agents. While participant performance has increased considerably in the first three competitions, the TAC game is by no means a fully solved problem and will remain as the 'TAC classic' in future competitions.

There is consensus that the TAC community is ready for the introduction of an additional trading game, introducing new aspects and corresponding research issues. But there are several different candidate proposals, ranging from minor modifications of 'TAC classic' to quite different games involving spectrum auctions, stock markets, and supply chain management. There is also consensus that TAC will be proposed as an official workshop in a major AI or agents conference such as IJCAI or AAMAS, potentially attracting more interest and participants.

We look forward to seeing even more European participants in future competitions. You can get a head start in 'TAC classic' by downloading and trying out the SICS TAC server and agentware.

Link:
http://www.sics.se/tac/

Please contact:
Joakim Eriksson, SICS
Tel: +46 8 633 1541
E-mail: joakime@sics.se
http://www.sics.se/~joakime/

sponsored by ERCIM

CONCUR 2002 Conference Report

by Luboš Brim

The 13th International Conference on Concurrency Theory, sponsored by ERCIM, was hosted by the Faculty of Informatics of Masaryk University Brno, Czech Republic from 19-24 August 2002. There were 174 registered participants to CONCUR 2002 and its satellite events from 18 countries around the world.

The purpose of the CONCUR conferences is to bring together researchers, developers and students in order to advance the theory of concurrency and promote its applications. Interest in this topic is continuously growing as a consequence of the importance and ubiquity of concurrent systems and their applications, and of the scientific relevance of their foundations. The scope covers all areas of semantics, logics and verification techniques for concurrent systems. Topics include concurrency-related aspects of models of computation and semantic domains, process algebras, Petri nets, event structures, real-time systems, hybrid systems, probabilistic systems, model checking, verification techniques, refinement techniques, term and graph rewriting, distributed programming, logic constraint programming, object-oriented programming, typing systems and algorithms, security, case studies, tools and environments for programming and verification.

The first two CONCUR conferences were held in Amsterdam (NL) in 1990 and 1991, with the following locations being Stony Brook (USA), Hildesheim (D), Uppsala (S), Philadelphia (USA), Pisa (I), Warsaw (PL), Nice (F), Eindhoven (NL), University Park (Pennsylvania, USA), and Aalborg (DK). The proceedings of CONCUR 2002 have appeared in Springer LNCS as volume 2421.

The technical programme of the conference and its satellite events delivered 116 presentations in total, from which 18 were invited talks and invited tutorials. Out of 101 regular papers submitted to CONCUR 2002, 32 were accepted for presentation at the conference. The conference also included five talks by invited speakers. Both Amir Pnueli (Weizmann Institute of Science, Rehovot, IL) and Vaughan Pratt (Stanford University, USA) reappeared in the role of CONCUR invited speakers after ten years, and further invited talks were given by Wan Fokkink (CWI Amsterdam, NL), Alexander Rabinovich (Tel Aviv University, IL) and Davide Sangiorgi (INRIA Sophia Antipolis, F). Additionally, there were two invited tutorials, given by Julian Bradfield (University of Edinburgh, UK) and Andrew D. Gordon (Microsoft Research, Cambridge, UK). The best student paper award went to 'Local Model-Checking Games for Fixed Point Logic with Chop' by Martin Lange.

CONCUR 2002 was accompanied by a special event called Tools Day organised by Ivana Černá, and by the following satellite workshops:

- CMCIM'02 (Categorical Methods for Concurrency, Interaction, and Mobility), organised by Alexander Kurz, held on 24 August 2002
- EXPRESS'02 (9th International Workshop on Expressiveness in Concurrency), organised by Uwe Nestmann and Prakash Panangaden, held on 19 August 2002
- FATES'02 (Formal Approaches to Testing of Software), organised by Robert M. Hierons and Thierry Jéron, held on 24 August 2002
- FOCLASA'02 (1st International Workshop on Foundations of Coordination), organised by Antonio Brogi and Jean-Marie Jacquet, held on 24 August 2002
- INFINITY'02 (4th International Workshop on Verification of Infinite State Systems), organised by Antonín Kučera and Richard Mayr, held on 24 August 2002
- MTCS'02 (3rd International Workshop on Models for Time-Critical Systems), organised by Kim Larsen and Walter Vogler, held on 24 August 2002
- PDMC'02 (1st International Workshop on Parallel and Distributed Model Checking), organised by Luboš Brim and Orna Grumberg, held on 19 August 2002.

Photos: Jan Holeček.

Impressions from CONCUR 2002.

The conference took place in the Czech Republic for the first time. The organisers attempted to give the conference additional stimuli like moderate conference fees and cheap accommodation on the University campus to make it attractive to participants from Central and Eastern Europe and to young researchers all over the world.

Link:
Conference web page:
http://www.fi.muni.cz/concur02

Please contact:
Lubos Brim, CRCIM
Tel: +420 54151 2323
E-mail: brim@fi.muni.cz

sponsored by ERCIM

CEOI 2002 — Central European Olympiad in Informatics 2002

by Gabriela Andrejková and Jan Vinař

CEOI 2002 took place from 30 June to 6 July 2002 in the Slovakian city of Kosice, under the auspices of the Ministry of Education of the Slovak Republic.

ERCIM played an important role both in organising this event (the three organising bodies — the Faculty of Science of P. J. Šafárik University, the Faculty of Mathematics, Physics and Informatics of Comenius University, and the Slovak Society for Computer Science are connected with ERCIM as members of SRCIM) and in providing necessary funds by means of a financial grant.

The participating teams from Croatia, the Czech Republic, Germany, Hungary, Iran, Poland, Romania, Slovakia and Slovenia (with 60 contestants and team leaders in all) demonstrated their skills in writing and testing programs by solving six problems ranging from graph theory to theory of games. The solutions were evaluated automatically, taking into account accuracy (including borderline cases) as well as the time and space complexity of the algorithms. Three gold medals were awarded: the winner was Peter Bella from Slovakia, and the next two places were taken by Victor Costan and Daniel Dumitran from Romania.

The CEOI, which is open to secondary school students in Central Europe, aims to:

- encourage students' interest in informatics and information technology in general
- test and prove their competence in solving problems with the help of computers
- promote the exchange of knowledge and experience with other students of similar interests and qualifications
- help establish personal contacts between young people of the Central European region.

Additionally, its purpose is to provide training for the students participating in the International Olympiad in Informatics (IOI), and to initiate discussion and cooperation in informatics education in the secondary schools of Central European countries. In our view this, the ninth CEOI, was successful in fulfilling these aims. Next year's CEOI will be held in Münster, Germany, and Poland will host CEOI 2004.

Link:
http://cs.science.upjs.sk/ceoi/

Please contact:
Gabriela Andrejková
P. J. Šafárik University, Košice (SRCIM)
Tel: +421 95 62 21 128
E-mail: andrejk@kosice.upjs.sk

Forthcoming Events sponsored by ERCIM

- **SOFSEM 2002 — 29th Annual Conference on Current Trends in Theory and Practice of Informatics**, Milovy, Czech Republic, 24-29 November 2002 http://www.sofsem.cz/
- **Euroweb 2002 Conference**, Oxford, UK, 17-18 December 2002 http://www.w3c.rl.ac.uk/Euroweb/
- **Coding and Cryptography WCC2001**, Versailles, France, March 24-24, 2003 http://www-rocq.inria.fr/codes/WCC2003/

For more information about ERCIM's conferences, workshops and summer school sponsorship programme, see http://www.ercim.org/activity/sponsored.htm

CALL FOR PAPERS

ECIR'03 — 25th European Conference on Information Retrieval Research

Pisa, Italy, 14-16 April 2003

The annual European Conference on Information Retrieval Research, organised by the Information Retrieval Specialist Group of the British Computer Society (BCS-IRSG), is the main European forum for the presentation of research results in the field of information retrieval.

We encourage the submission of papers reporting original, unpublished research results in information retrieval, broadly intended to cover all aspects of computer-based access to data with poorly specified semantics. This includes (but is not limited to) the searching, browsing, filtering, classifying, clustering, mining, tracking and summarising of written text, hypertext, spoken text, music, images, video, and Web resources. More information on the list of topics that will be covered by the Conference can be found on the ECIR'03 Web site.

Important Dates

- Paper submission: 20 November 2002
- Acceptance/rejection notification: 21 December 2002
- Submission of camera-ready copy: 20 January 2003.

Accepted papers will appear in the proceedings published by Springer-Verlag in the Lecture Notes in Computer Science series. Awards will be presented to the author of the Best Paper and to the author of the Best Student Paper.

More Information:
http://ecir03.isti.cnr.it/

CALL FOR PAPERS

Fourth International Workshop on Energy Minimisation Methods in Computer Vision and Pattern Recognition – EMMCVPR'2003

Lisbon, Portugal, 7-9 July 2003

Many problems in computer vision and pattern recognition (CVPR) are couched in the framework of optimisation. The minimisation of a global quantity, often referred to as the energy, forms the bulwark of most approaches in CVPR. Disparate approaches such as discrete and probabilistic formulations on the one hand, and continuous, deterministic strategies on the other, often have optimisation or energy minimisation as a common theme. Instances of energy minimisation arise in Gibbs/Markov modelling, Bayesian theory, geometric and variational approaches and in areas in CVPR such as object recognition/retrieval, image segmentation, registration, reconstruction, classification and data mining.

The aim of this workshop, the fourth in a series, is to bring together researchers with interests in these disparate areas, but with an underlying commitment to some form of optimisation. Although the subject is traditionally well represented in major conferences on CVPR, this workshop provides a forum wherein researchers can report their recent work and engage in more informal discussions. As with the previous editions the proceedings will be published by Springer Verlag in the Lecture Notes on Computer Science (LNCS) series.

Topics include (but are not restricted to):
- Gibbs/Markov modeling
- Probabilistic networks and graphical models
- Variational formulations, level sets, and PDEs
- Deformable models and registration
- Graph matching
- Statistical pattern recognition
- Supervised learning
- Unsupervised learning
- VC-theory and support vector machines
- Information theoretic methods
- Model selection
- Combinatorial optimisation
- Interior point methods
- Image reconstruction and coding
- Neural networks for classification and regression
- Markov-Chain Monte Carlo methods
- Relaxation labeling
- Advanced mean-field methods
- Self-organising networks
- Evolutionary / genetic approaches
- Applications.

Important Dates
- Paper submission deadline: 6 January 2003
- Notification of acceptance: 1 March 2003
- Camera-ready paper due: 1 April 2003.

More information:
http://www.emmcvpr.org/

CALL FOR PAPERS

3rd ESA Workshop on Millimetre Wave Technology and Applications: Circuits, Systems, and Measurement Techniques

Espoo, Finland, 21-23 May 2003

The third European Space Agency (ESA) workshop on the technology and applications of millimetre waves will be organised by MilliLab jointly with ESA/ESTEC, VTT and Helsinki University of Technology. The aim of this international workshop is to bring together people involved in research and industrial development of millimetre wave components and systems, and to explore common areas and synergies in the development of millimetre wave techniques for commercial and scientific ground-based and space-borne applications.

More information:
http://www.vtt.fi/millilab/docs/FirstCall2003.pdf

CALL FOR PAPERS

WSCG 2003 — 11th Conference in Central Europe on Computer Graphics, Visualisation and Computer Vision

Plzen, Czech Republic, 3-7 February 2003

the WSCG conference has been held annually since 1992. The conference is devoted to:
- Algorithms, Parallel & Distributed Graphics, Computer Aided Geometric Design
- Graphics Architecture & Visualization HW, Image Based Rendering, Mathematical Aspects of Graphics
- Global Illumination, Ray Tracing, Radiosity
- Computational Geometry, Surface Meshing, Modeling, Constraint Motion, Simulation
- Virtual Reality & VR Interaction, Animation, Viewing Dynamic World Morphing & Warping
- Visualization, Computer Vision & Image Processing, Pattern Recognition
- WWW Technologies, Hypermedia, Human Computer Interface (Graphical&Vision,Haptic)
- CAD/CAM & GIS Systems, Education Aspects, Applications.

The conferences are interdisciplinarily oriented, presenting the state-of-the-art, as well as the latest research results and applications. The WSCG conferences are intended to promote the exchange of ideas over a broad range of topics.

Keynote speakers:
- Gabor T. Herman: 'Blobs for Representing and Visualizing 3D Objects', City University of New York, USA
- Dirk Bartz: 'Large Model Visualization: Techniques and Applications', University of Tübingen, Germany.

More information:
http://wscg.zcu.cz

ERCIM NEWS

ERCIM News is the magazine of ERCIM. Published quarterly, the newsletter reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of over 9000 copies.

ERCIM News online edition is available at http://www.ercim.org/publication/ERCIM_News/

ERCIM News is published by ERCIM EEIG, BP 93, F-06902 Sophia-Antipolis Cedex
Tel: +33 4 9238 5010, E-mail: office@ercim.org
ISSN 0926-4981

Director: Bernard Larrouturou
Central Editor:
Peter Kunz
peter.kunz@ercim.org
Local Editors:
AARIT: Wolfgang Hawlik
w.hawlik@ocg.at
CLRC: Martin Prime
M.J.Prime@rl.ac.uk
CRCIM: Michal Haindl
haindl@utia.cas.cz
CWI: Henk Nieland
Henk.Nieland@cwi.nl
CNR: Carol Peters
carol@iei.pi.cnr.it
FORTH: Leta Karefilaki
karef@ics.forth.gr
FhG: Michael Krapp
michael.krapp@iuk.fhg.de
FLN: Patrik Hitzelberger
hitzelbe@crpgl.lu
INRIA: Bernard Hidoine
bernard.hidoine@inria.fr
NTNU: Truls Gjestland
truls.gjestland@ime.ntnu.no
SARIT: Harry Rudin
hrudin@smile.ch
SICS: Kersti Hedman
kersti@sics.se
SRCIM: Gabriela Andrejkova
andrejk@kosice.upjs.sk
SZTAKI: Erzsébet Csuhaj-Varjú
csuhaj@sztaki.hu
TCD: Ann McNamara
Ann.McNamara@cs.tcd.ie
VTT: Pia-Maria Linden-Linna
pia-maria.linden-linna@vtt.fi

Free subscription

You can subscribe to ERCIM News free of charge by:

- sending e-mail to your local editor
- contacting the ERCIM office (see address above)
- filling out the form at the ERCIM website at http://www.ercim.org/

sponsored by ERCIM

Euroweb 2002 Conference
The Web and the GRID: from E-Science to E-Business

Oxford, UK, 17-18 December 2002

The conference theme is intended to prompt debate on a convergence of developments pioneered for e-science on the GRID and as Web services to provide business applications. The communities who identify with the Web, the GRID, Web Services, Grid Services and Semantic Web should not be isolated from each other, but need to come together to unify their approaches to meet the real needs of information, data and knowledge technology users:

- how will frameworks like Web services, GRID Services, and .NET address the issue of Internet-aware programs using the services that are offered by other programs?
- what is XML's role in managing and routing data and services?
- what kinds of problems lend themselves to resource sharing?
- should Web services be inspectable whilst running? Are there lessons to learn from GRID services about quality of service management ?
- what contribution can Semantic Web technologies make to service directories, quality of service policies etc?
- where are the practical business applications requiring the terabits/sec of speed, terabytes of storage and petaflops of processing used in GRID computing?
- what are the barriers to adoption and assimilation of these technologies for business, and how should they be overcome?

EuroWeb 2002 will be a major international forum at which research on GRIDs and Web Services is presented. EuroWeb 2002 is endorsed by the UK and Ireland office of W3C, and by IW3C2.

More information:
http://www.w3c.rl.ac.uk/Euroweb/

Call for Participation

'iEX' — Internet Expo 2003

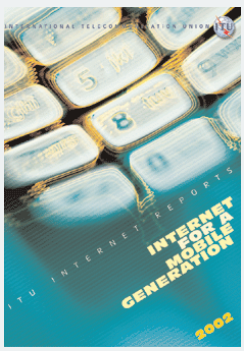
Zurich, Switzerland, 5-7 February 2003

The target attendee is the Internet professional. The meeting has two parts: a tradeshow and a conference. Attracting some 30,000 visitors last year, the iEX tradeshow this year will have three main sections: Internet/intranet solutions, networking products, and Web services. The conference portion attracted some 3000 attendees last year. This year the topics are e-commerce, intranet, Web-mastering, software development (particularly Web related), and network planning strategies. Workshops and case studies are included in the program.

More information:
http://www.iex.ch/

Book Review

The ITU released a new publication, 'Internet for a Mobile Generation', on 17 September, 2002. If you are concerned

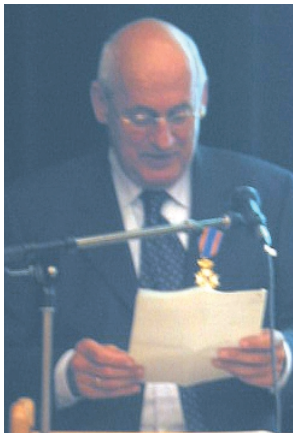


about the mixture of two explosive growth areas, the Internet and mobile communications, you should have this 240-page book at your side. You will find almost all the information you need in this book: a summary of the competing technologies and their applications, case studies from various countries, how 200 economies are performing in terms of mobile and Internet technologies, regulatory and policy aspects, and sample scenarios for the future. Extracts are given on the ITU's Web site: http://itu.int/mobileinternet/. The book may also be ordered through that site. This is the fourth in a series of studies written under the supervision of Tim Kelly who writes with a clear and delightful style.

Harry Rudin, Consultant and Swiss Local Editor for ERCIM News

CNR — **Piero Maestrini** has been appointed Director of the newly constituted 'Istituto di Scienza e Tecnologie dell'Informazione' (ISTI), located in the CNR Research Area, Pisa. ISTI-CNR is the result of a union between the CNR

Institutes of CNUCE and IEI and will conduct R&D activities in the fields of computer science and IT. Currently, the main areas of interest are currently: software engineering, formal methods for specification and verification, networked multimedia information systems, Web-based information systems, knowledge discovery and data mining, image and signal processing, visual computing and computer graphics, human-computer interaction, high performance computing, wireless and mobile networks, dependable computing, structural mechanics and spaceflight dynamics. Further information can be found at http://www.isti.cnr.it/.



CWI — **Jaco W. de Bakker** received a high Royal Decoration (Ridder in de Orde van de Nederlandse Leeuw), during his 'retirement symposium' at CWI on 30 August 2002. Jaco de Bakker is one of the pioneers of computer science in The Netherlands, especially in the field of mathematical semantics of programming languages and reasoning on program correctness. He worked for 38 years at CWI, and put the word 'Informatics' in the name of the former Mathematical Centre. From 1973 on, De Bakker was Professor at the Vrije Universiteit in Amsterdam, and in 1989 he became a Member of the Royal Netherlands Academy of Arts and Sciences. He wrote more than 150 scientific articles and books. De Bakker is proud that 32 scientists who worked in his group were appointed professor. Also at managerial level, De Bakker played an important role. CWI will miss him and wishes him all the best for the future.

CWI — **Professor Edsger Wybe Dijkstra**, world-famous pioneer of the science and industry of computing, died after a long struggle with cancer on 6 August 2002 at his home in Nuenen, the Netherlands, at the age of 72. Dijkstra worked at the Mathematisch Centrum (MC/CWI) between 1952 and 1962. During his Amsterdam years, he worked as the first 'programmer' (his own modest description) in the Netherlands, participating in the design of the MC's first commercially marketed computer,

Photo: CWI

the X1. Perhaps his greatest achievement during these years was the writing with Jaap Zonneveld of the world's first ALGOL60 compiler. From 1962 until 1984 he held a chair at the Eindhoven University of Technology. His fame was augmented through his fundamental studies of parallel programming and his insights into the construction of correct programs, and he was an eloquent advocate of the methodology of structured programming. Dijkstra was a Burroughs Corporation research fellow from 1973-1984. From 1984 until his retirement in 1999 he worked at the University of Texas in Austin. He wrote over 1300 books and papers, all of which are digitally accessible. Dijkstra was the 1972 recipient of the ACM Turing Award, often viewed as the Nobel Prize for computing. He was also a recipient of numerous honorary degrees and awards. Dijkstra is universally acclaimed as the most famous scientist in the history of the Mathematical Centre, and was arguably one of the most influential computer scientists of the twentieth century. Dijkstra is survived by his wife, Ria, and three children. For more information, see: http://www.cs.utexas.edu/users/EWD/

ERCIM PhD Fellowship Programme

ERCIM offers 18-month fellowships in leading European information technology research centres. Fellowships are available in areas covered by the ERCIM Working Groups and other fields. This offer is available for PhD holders from all over the world.

Deadline for application: 31 October 2002

Conditions

Applicants must:
- have a PhD degree (or equivalent) or be in the last year of the thesis work
- be fluent in English
- be discharged or get deferment from military service.

Fellowships are of 18 months duration, generally spent in two institutes. The fellow will receive a monthly allowance which may vary depending on the country. In order to encourage the mobility, a member institution will not be eligible to host a candidate of the same nationality.

Detailed information and the online application form is available at:
http://www.ercim.org/fellowship/

ERCIM – The European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development, in information technology and applied mathematics. Its national member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

Austrian Association for Research in IT
c/o Österreichische Computer Gesellschaft
Wollzeile 1-3, A-1010 Wien, Austria
Tel: +43 1 512 02 35 0, Fax: +43 1 512 02 35 9
http://www.aarit.at/

Central Laboratory of the Research Councils
Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
Tel: +44 1235 82 1900, Fax: +44 1235 44 5385
http://www.cclrc.ac.uk/

Consiglio Nazionale delle Ricerche, ISTI-CNR
Area della Ricerca CNR di Pisa,
Via G. Moruzzi 1, 56124 Pisa, Italy
Tel: +39 050 315 2878, Fax: +39 050 315 2810
http://www.isti.cnr.it/

Czech Research Consortium
for Informatics and Mathematics
FI MU, Botanicka 68a, CZ-602 00 Brno, Czech Republic
Tel: +420 2 688 4669, Fax: +420 2 688 4903
http://www.utia.cas.cz/CRCIM/home.html

Centrum voor Wiskunde en Informatica
Kruislaan 413, NL-1098 SJ Amsterdam,
The Netherlands
Tel: +31 20 592 9333, Fax: +31 20 592 4199
http://www.cwi.nl/

Fonds National de la Recherche
20, Montee de la Petrusse
L-2912 Luxembourg
Tel. +352 26 1925 31, Fax +352 26 1925 35
http:// www.fnr.lu

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
Tel: +30 81 39 16 00, Fax: +30 81 39 16 01
http://www.ics.forth.gr/

Fraunhofer-Gesellschaft,
Information and Communication Technology Alliance
Schloß Birlinghoven, D-53754 Sankt Augustin, Germany
Tel: +49 2241 14 0, Fax: +49 2241 14 2889
http://www.fraunhofer.de/

Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
Tel: +33 1 3963 5511, Fax: +33 1 3963 5330
http://www.inria.fr/

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and
Electrical Engineering, N 7491 Trondheim, Norway
Tel: +47 73 59 80 35, Fax: +47 73 59 36 28
http://www.ntnu.no/

Swedish Institute of Computer Science
Box 1263,
SE-164 29 Kista, Sweden
Tel: +46 8 633 1500, Fax: +46 8 751 72 30
http://www.sics.se/

Swiss Association for Research in Information Technology
Dept. Informatik, ETH-Zentrum, CH-8092 Zürich,
Switzerland
Tel: +41 1 632 72 41, Fax: +41 1 632 11 72
http://www.sarit.ch/

Slovak Research Consortium for Informatics and
Mathematics, Comenius University, Dept.of Computer
Science, Mlynska Dolina M, SK-84248 Bratislava,
Slovakia, Tel: +421 7 726635, Fax: +421 7 727041
http://www.srcim.sk

Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézete
P.O. Box 63, H-1518 Budapest, Hungary
Tel: +36 1 279 6000, Fax: + 36 1 466 7503
http://www.sztaki.hu/

Trinity College
Department of Computer Science,
Dublin 2, Ireland
Tel: +353 1 608 1765, Fax: 353 1 677 2204
http://www.cs.tcd.ie/ERCIM

Technical Research Centre of Finland
P.O. Box 1200
FIN-02044 VTT, Finland
Tel:+358 9 456 6041, Fax :+358 9 456 6027
http://www.vtt.fi/tte

Order Form

If you wish to subscribe to ERCIM News ***free of charge*** *or if you know of a colleague who would like to receive regular copies of ERCIM News, please fill in this form and we will add you/them to the mailing list.*

Send, fax or email this form to:

ERCIM NEWS
Domaine de Voluceau
Rocquencourt
BP 105
F-78153 Le Chesnay Cedex
Fax: +33 1 3963 5052
E-mail: office@ercim.org

Name:

Organisation/Company:

Address:

Post Code:

City:

Country

E-mail:

Data from this form will be held on a computer database. By giving your email address, you allow ERCIM to send you email

You can also subscribe to ERCIM News and order back copies by filling out the form at the ERCIM website at http://www.ercim.org/publication/Ercim_News/