# ERCIM NEWS

Special:

## The Future
# WEB

# The Web of Things

Progress in communications technology has been characterized by a movement from lower to higher levels of abstraction. When, first, computers were connected by telephone wires, you would have to run a special program to make one connect to another. Then you could make the second connect to a third, but you had to know how to use the second one too. Mail and news was passed around by computers calling each other late at night, and, for a while, email addresses contained a list of computers to pass the message through, such as: timbl@mcvax!cernvax!cernvms.

### It's not the wires – it's the computers

The ability to use this communication power between computers wasn't truly useful until the Internet. The Internet allowed one to forget about the individual connections. It was thought of as the 'Internet Cloud'. Messages went in from one computer, and appeared in another, without one having to worry about how they were broken into packets, and how the packets were routed from computer to computer.

### It's not the computers – it's the documents

This power of communication between computers wasn't easily usable by the majority of people until the Web came along. The realization of the Web was: "It's not the computers which are interesting, it's the documents!". The WWW protocols (URI, HTTP, HTML) defined how documents could be sent between Web servers and Web browsers. With these innovations, the user entered a Web of interconnected documents. She didn't have to worry about how the protocols work underneath, with two exceptions.

When things went wrong, she had to be able to figure out whether it was a problem with her connection to the Internet, with the URI in the link she was following, or an error on the server end. This involved looking under the hood, as it were. But that was, and still is, the exception.

There is another reason to be aware of what is happening. The information one browses comes from a particular server, whose name has been registered to a particular person or organization. The trust you put in that information relates to who that organization is. The serving organization is responsible for keeping the URIs you bookmark today alive tomorrow. So the Web is just a web of documents, except one has to be aware of the social aspects implied by the underlying level.

### It's not the documents – it's the things

The power of the Web hadn't reached its full potential until the semantic Web came along. The semantic Web's realization is, "It isn't the documents which are actually interesting, it is the things they are about!". A person who is interested in a Web page on something is usually primarily interested in the thing rather than the document. There are exceptions, of course – documents are certainly interesting in their own right. However, when it comes to the business and science, the customers, the products, or the proteins and the genes, are the things of interest. A good semantic Web browser, then, shows a user information about the thing, which may have been merged from many sources. Primarily, the user is aware of the abstract Web of connections between the things



*Tim Berners-Lee, director of the World Wide Web Consortium and inventor of the Web.*

Photo: Donna Coveney.

– eg: this person is a customer who made this order which includes this item which is manufactured by this facility ... and so on.

There are again the same two exceptions. When things go wrong, the user must be able to look under the hood to see whether the document was fetched OK but had missing data, or the document was not fetched OK, in which case what was the underlying Web problem. And again, when the user is looking at a bit of data in the data view, perhaps a point on a map or a cell in a table, then she must be able to see easily which document that information came from.

One thing, or person, may have many URIs on the semantic Web. Often such a URI has within it the URI of the document which has some information about the thing: a gene as defined in the Gene Ontology; a protein as defined in this taxonomy; a citizen as defined by the Immigration and Naturalization Service glossary, and so on. Similarly, the URI of the document has within it the DNS name of the computer. So the social structure can be seen within the URI.

### The World of Things ... on the Web

The Web of things is built on the Web of documents, which is built on the Web of computers controlled by DNS owners, which itself is build on a set of interconnected cables. This is an architecture which provides a social backing to the names for things. It allows people to find out the social aspects of the things they are dealing with, such as provenance, trust, persistence, licensing and appropriate use as well as the raw data. It allows people to figure out what has gone wrong when things don't work, by making the responsibility clear.

The last level of abstraction is the Web of real things, built on top of the Web of documents, which is in turn built on the network of computers.

*Tim Berners-Lee*

*Next issue: April 2008*
*Next special theme: Maths for Everyday Life*

# Denmark joins ERCIM

***DANAIM (Danish Research Association for Informatics and Mathematics), a research consortium established by seven major Danish universities and university collaborations, has become a member of ERCIM. Denmark is the nineteenth country to join ERCIM.***

The aim of ERCIM is to foster collaborative work within the European research community in information and communication technology (ICT) and applied mathematics, and to increase cooperation with European industry. With the DANAIM membership, ERCIM continues to expand its geographical reach and increase its impact.

### DANAIM Represents almost all Danish Universities

DAIAIM is a consortium of seven universities: University of Southern Denmark, University of Aarhus, Roskilde University, Technical University of Denmark, IT Vest – networking Universities, University of Copenhagen and Aalborg University. We hope that the remaining two Danish universities become members in a few years. In Denmark there is a long tradition of international collaboration with other universities and collaboration with industry. DANAIM will strengthen the cooperation between Danish and other European researchers and industry.

DANAIM represents 440 researchers within the field of computer science and related fields and an annual turnover of more than €30 million. More than 75% of the students enrolled for computer science or related areas are enrolled at the institutions represented by DANAIM.

### Reasons for Joining ERCIM

DANAIM chair Chrstian S. Jensen finds many good reasons for joining ERCIM: "Research is fundamentally international. Membership in ERCIM represents an opportunity for the partners to strengthen their international ties with research organisations across Europe. The DANAIM partners view ERCIM's many working groups as attractive fora for scientific exchange and networking. Collaborative research projects often grow out of networks. By joining ERCIM, the partners hope to expand and invigorate their European networks, thereby increasing their participation in European projects."

On the more specific level, one of the major reasons for joining ERCIM was the fellowship program. As Christian S. Jensen notes, "The partners are particularly excited about the prospects for becoming part of ERCIM's fellowship program, as this highly successful program attracts very large numbers of applicants on an annual basis."

The research spirit in Denmark is collaborative. Consequently, it is in the interest of DANAIM to influence the strategic directions for ICT research at a European scale – an objective the DANAIM consortium hopes to be able to contribute to via ERCIM by supporting ERCIM's objective of influencing the research directions.

The board of directors for DANAIM consists of Kim Skak Larsen, Department of Mathematics and Computer Science, University of Southern Denmark; Kurt Jensen, Department of Computer Science, University of Aarhus; Hans Vestman Jensen, Institute of Information and Media Studies, University of Aarhus; Henning Christiansen, Department of Communication, Business and Information Technologies, Roskilde University; Flemming Nielson, Department of Informatics and Mathematical Modelling, Technical University of Denmark; Jens Bennedsen, it-vest – Networking Universities; Jørgen P. Bansler, Department of Computer Science, University of Copenhagen; and Christian S. Jensen (chair), Department of Computer Science, Aalborg University.
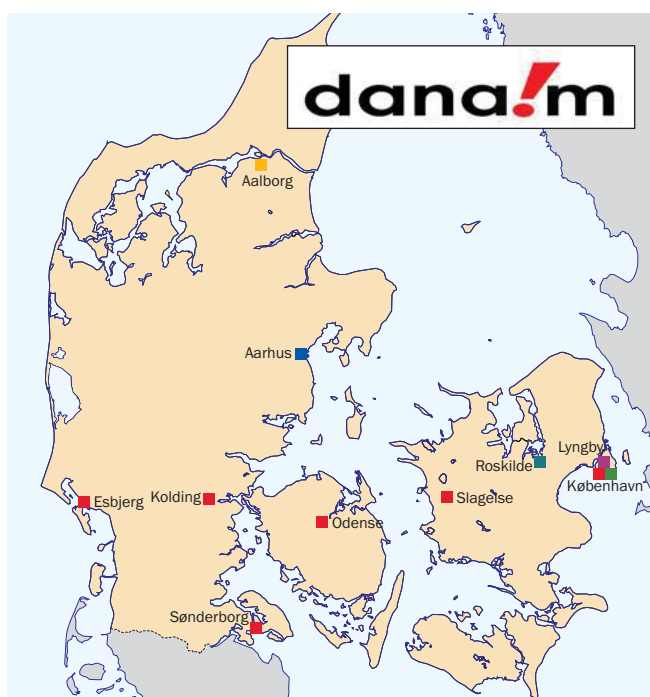


*The DANAIM consortium.*

**Link:**
http://www.danaim.dk/

**Please contact:**
DANAIM
Att: Henriette Frahm
c/o Aalborg Universitet, Institut for Datalogi
Selma Lagerlöfs Vej 300
9220 Aalborg Ø
Denmark
E-mail: frahm@cs.aau.dk

- ■ Aalborg University
- ■ University of Aarhus
- ■ University of Copenhagen
- ■ University of Southern Denmark
- ■ Technical University of Denmark
- ■ Roskilde University
- ◎ IT University West (an educational network between the four university institutions in the Western part of Denmark; Aarhus School of Business, University of Southern Denmark, Aalborg University and University of Aarhus.)

# The Future of the Web:
# An ERCIM View

by Keith G. Jeffery

It is impossible to imagine modern life without the Web. Tim Berners-Lee's original vision, so simple and elegant, has become part of the fabric of our lives. It is hard to realize that in less than two decades a line-mode 'techie' tool from a research laboratory has become indispensable to all for information, communication, business and research. Within the first few years of the Web, teams worked on aspects such as Web-database integration – leading to an explosion of e-business, especially B2C (Business to Customer) – and stylesheets to assist in consistent presentation. As time progressed and the Web became increasingly widespread, accessibility issues were addressed. Graphics recommendations such as PNG and SVG were developed. Synchronized multimedia (SMIL) was developed. Teams were already working on trust and security and on metadata to describe Web pages when 'Weaving the Web' was published, emphasising the Web of trust and semantic Web. The development of Java allowed the production of applets and servlets, leading to the client-server model being replaced by a new three-level architecture. Web services with XML provided data transfer and workflow capabilities for business. Web access from mobile devices became possible and the Mobile Web Initiative was launched by W3C, which has been in existence since the early years, providing a forum for consulting members and pushing through recommendations as effective standards.

## Web of Trust

With some background in PICS (Platform for Internet Content Selection) various researchers have been developing systems for trust and security in a Web environment. The definition of standard document formats (using XML) to exchange information about companies, products and services, and the provision of SLAs (service level agreements), has underpinned the trust effort. Cryptography with page-content encoding/decoding has been a major part of security. Privacy has also been addressed, from techniques for authenticating users to methods of authorizing (and therefore also restricting) access.

## Semantic Web

The development of the Resource Description Framework (RDF), essentially a binary relational representation, allowed more expressiveness to be encoded in Web applications. DAML and OIL – early ontology languages – led to the recommendation OWL for constructing domain ontologies to be used in assisting interoperation (or for input validation, query improvement and answer explanation). There is still a long way to go, but semantics are gradually being managed in the Web environment.

## Mobile Web

The increasing capacity of handheld devices and networks has led to demand for access to the Web on mobile devices. The possibilities include data input from attached detectors (eg video), full-quality video and sound output, and push technology to make the device intelligent with respect to the user's objectives and the current spatio-temporal environment.

## Web 2.0

The potential of the Web as a platform for interaction and cooperation is now emerging with the so-called Web 2.0 ideas. These include not only the use of services but also participation by users individually or collectively (collective intelligence). This can occur, for example, through blogs and Wikis with tagging, faster page loads and interactions using AJAX technology, access to multiple information sources and their integration through 'mashups', peer-to-peer connectivity and increasingly intelligent push technology.

## Web Services

The development of Web services provided a sort of application programming interface (API) to allow programming of the Web, rather than simply authoring or generating content and then making it available. With XML encoding of the information, cooperative working including workflow is made possible. It also made it possible to consider SOA (service-oriented architecture), which leads to shorter development times for applications and easier maintenance. However this also entailed metadata descriptions of services and business processes to allow discovery and use. There has been an explosion of metadata offerings in this domain – many application-domain specific – and some standardisation is needed.

## GRIDs

Starting as metacomputing (linked supercomputers) in North America in the mid- to late nineties, the Europeans proposed that GRID services be based on Web services so that GRIDs could become a ubiquitous facility. In this vein, work by the Next Generation Grid Expert Group of the European Commission has led to much R&D activity in services and their specification with metadata.

## Convergence

The future of the Web lies first with the convergence of the semantic Web, Web of trust and Web 2.0, followed by twinning with GRIDs. This combined e-infrastructure will allow more or less unlimited access to information, business processing, entertainment, education, research and so forth.

## Standards

The outstanding barriers to realizing this vision concern the development and acceptance of standards. This is partly because some standards (recommendations in W3C terms) are inadequate, and partly because there are competing commercial interest groups. However, to provide ubiquitous interoperability these standards must be developed, accepted and evolved in a structured way.

As they were ten years ago, the key standards required now are in the areas of semantics (particularly of metadata to allow management of services and management of information resources) and trust/security (another metadata-controlled area). Resolution of these areas will allow confident progression to the vision. ERCIM researchers are involved in projects addressing these very issues and are working together with the W3C Europe, located within ERCIM.

**Please contact:**
Keith G. Jeffery, ERCIM President
Science and Technology Facilities Council, UK
E-mail: kgj@rl.ac.uk

# Second DELOS Conference on Digital Libraries

*The second DELOS Conference, which was held on 5-7 December 2007 in Tirrenia, Pisa, Italy, not only provided a forum to present and discuss recent research advances in the field of Digital Libraries, but also provided the opportunity to summarise some of the main achievements of DELOS during the past four years.*

After four years of activity, the DELOS Network of Excellence on Digital Libraries, supported by the European Union under the Sixth Framework Programme and managed by ERCIM, has reached the end of its funded lifetime. However, the activities of DELOS actually started many years ago, first with the DELOS Working Group at the end of the nineties, and then through the DELOS Thematic Network, under the Fifth Framework Programme, from 2001 to 2003. Over these years, the objective of DELOS has been to fill the gap between current Digital Library practice and the needs of modern information provision. The goal has been to foster the development of technology that will eventually overcome the current restrictions of today's information systems.

In these years, it is widely recognised that DELOS has played a significant role in
• the formation of an active European Digital Library research community
• the formulation of a vision for the future of the field
• fostering collaborative research in the direction of this vision.

The DELOS community has provided significant contributions to many key components of digital libraries. This is evidenced by the vast volume of scientific papers (more than 500) produced by the DELOS community in the last four years and by the number of projects and applications in which members of DELOS have provided or are providing a contribution. As important examples, we can cite the Digital Library Reference Model and the DELOS Digital Library Management System, which were both presented at the Conference.



*Yannis Ioannidis presenting the DELOS achievements.*



*DELOS fields of research.*

The conference was attended by more than 80 participants. The 33 presentations reflected the main fields of interest of the DELOS community (DL Architectures, Interoperability, Information Access, Image and 3D Objects Retrieval, User Interfaces, Personalization, Preservation, Evaluation). One session was dedicated to the DELOS Reference Model, which constitutes one of the major legacies of DELOS to the digital library community. The final morning session consisted of presentations and demonstrations of the Delos-DLMS, a globally integrated prototype of a Digital Library Management System, offering DL services and specialized functionality provided by members of the DELOS community. The DelosDLMS currently offers text and audio/visual searching, relevance feedback tools, novel interfaces and information visualization, annotation of retrieved information, and support for sensor data streams.

There were also three invited talks from key players in the field: Fabrizio Gagliardi, Microsoft, who described "Technical Computing at Microsoft Research", Maurizio Lenzerini, University of Roma La Sapienza, on "Ontology-based information access", and Stefan Gradmann, University of Hamburg, talking about "Interoperable Information Space – moving towards the European Digital Library".

The results achieved by DELOS over the years were summarised in a concluding session both by Yannis Ioannidis in a presentation entitled "DELOS Achievements" and also by

# Launch of the DELOS Association for Digital Libraries

*Launch of the DELOS Association for Digital Libraries - In order to keep the DELOS spirit alive after the end of the DELOS Network of Excellence sponsored by the Sixth Framework Programme, a non-profit association has been established. The main objective of this initiative will be to continue the work of DELOS by promoting research activities in the field of digital libraries through the organisation of workshops, working groups, schools, etc. Membership is open both to individuals and to organisations, and the membership fee for the first year has been set at €50 for individuals and €2000 for organisations. Complete information, including the Statute of the Association and a membership application form can be found on the DELOS Web site at http://www.delos.info/Association/.*

a number of participants who presented various activities that have seen the light as a result of work which begun in or was sponsored by DELOS. These activities include a number of independent EC funded projects, eg MultiMatch, TrebleCLEF, LOGOS, DILIGENT, DPE, etc. as well as national initiatives such as Rinet in Greece, and the series of IRCDL conferences in Italy. The conference proceedings and copies of all presentations can be found on the DELOS Web site at http://www.delos.info/.

Although, sadly, this was the last official event of the DELOS Network of Excellence, DELOS will live on in through the newly formed "DELOS Association" and it is hoped that many of the DELOS events, such as the summer schools and the scientific workshops, will continue in the future. For more information about the DELOS Association, see the announcement in the box or visit the DELOS Web site.

**Link:**
http://www.delos.info/

**Please contact:**
Costantino Thanos
DELOS scientific coordinator
ISTI-CNR, Italy
E-mail: thanos@isti.cnr.it

# 'The infinity Initiative' - ICT Research meets Standardization

*ERCIM and ETSI, the European Telecommunications Standards Institute, have co-signed a landmark Memorandum of Understanding (MoU) to launch their future close collaboration.*

Collaborative research and development in ICT (Information and Communication Technologies) plays a major role in determining the prosperity of Europe. The MoU encapsulates the organisations' joint aims to:
• exploit areas of synergy between Research and Standardisation in the innovation food chain
• accelerate R&D innovation cycles
• reduce time-to-market.

The organisations' commitment to their future collaboration was initiated by an inaugural high level seminar at ETSI headquarters on 29 November 2007, the first in a series of envisaged joint events. The seminar took place in Sophia Antipolis, France, where both ERCIM and ETSI have their operations.

The seminar included presentations from the European Commission DG Information Society & Media and the Joint Research Council (JRC) as well as from top level representatives from Alcatel-Lucent, the Austrian Academy of

Science, BT, Ericsson, Fraunhofer-Gesellschaft, University College Dublin and the World Wide Web Consortium.

According to ETSI Director-General, Walter Weigel: "ICT markets are shaped by [ETSI] Standards which enable the widespread use of new technologies developed after extensive research activities led by organisations such as ERCIM and so cooperation between our two communities can only be beneficial. This is an exciting development for all innovation food chain stakeholders and our corridors are already buzzing with ideas for the future".



*Walter Weigel, ETSI Director-General (left) and Keith Jeffery, ERCIM President*

ERCIM President, Keith Jeffery said: "The greatest satisfaction for researchers is to see their ideas carried forward to products and services that are used widely. There are many examples where ERCIM researchers have achieved this. However, this cooperation with ETSI allows us to accelerate the innovation chain by overlapping research, testing, standardisation and exploitation, thus creating a leading position for the ICT industry within Europe, the results of which will then spread -like the World Wide Web- to the rest of the world."

"The infinity Initiative" envisages a number of joint events including advanced workshops, seminars and meetings. It aims to strengthen the links between standardisation and research and has the support of the European Commission. The name "The infinity Initiative" was chosen in allusion to the enduring and futuristic nature of research.

**Link:**
Seminar programme, agenda, presentations and speakers' biographies:
http://www.etsi.org/WebSite/NewsandEvents/infinity_home.aspx

**Please contact:**
Catherine Marchand
ERCIM office
E-mail: catherine.marchand@ercim.org

# IV Grids@Work Conference in China

by Bruno Le Dantec

*Further to the success of the previous three events, ERCIM, ETSI, and INRIA teamed again this year to organise the 4th GRID Plugtests. The event, hosted by CNIC, the Computer Network Information Center of the Chinese Academy of Sciences, took place in Beijing from 28 October to 2 November 2007.*

The events brought together more than 130 Grid experts and users from academia and industry from many European countries and China. It was co-located with a series of conferences and tutorials at which several Grid research projects met and networked. Participants learned about Grid perspectives for research activities and industrial applications, compared some running European and Chinese Grid middleware and were informed on the efforts made by the Grid community toward interoperability.

## Workshops, Tutorials and 'All EC Projects' Meetings

EchoGrid, the project on European and Chinese Cooperation on the Grid, held its second public workshop. Over two days, presenters from Europe and China exchanged and compared



*Plugtest participants.*

past work and exciting ideas for the coming years on four topics, namely: new programming paradigms, trust and security, Grid workflow and Grid middleware for industrial applications. The outcomes will be used to draft the EchoGrid roadmap during the next consortium meeting in Roma in January 2008.

The GridComp project (Grid Programming with Components) organised its internal management meeting together with an open Proactive/GCM tutorial; this was used to present interactively the ProActive middleware for parallel, distributed and multi-threaded computing. This middleware provides the reference implementation of the Grid Component Models (GCM). At the end of the day, participants were able to develop, deploy and test ProActive/GCM applications.

Discussions relating to the Bridge project (Cooperation between Europe and China to Develop Grid Applications) were centred on successfully obtaining interoperability requirements from the application partners, and finalizing the design of the joint interoperability architecture needed by the project. The first Bridge prototype will be delivered before the end of the year.

All three European projects are partially funded by the European Commission. EchoGrid and Gridcomp are managed by ERCIM, while Bridge is managed by Fraunhofer-Gesellschaft. In the middle of the week, EchoGrid organised an all-EC Projects meeting to exchange information between the above projects and to organise a dedicated session on the description and comparison of the different middleware in use by Chinese and European projects. Representatives of five types of middleware participated in this exercise (Proactive/GCM, EGEE gLite, GRIA, CNGrid GOS and CROWNE), using a dedicated questionnaire defined prior to the meeting. A very interesting and exciting question and answer session followed the presentations.

## IV Grids@Work Plugtest

The goal of this year's challenge was to write the best Grid distributed program that could count a maximum number of solutions within a limited amount of time to the well-known N-Queens problem. The finalist teams competing had to write a program able to work out solutions to the N-Queens problem using a Grid. All jobs were submitted from Beijing using the ProActive middleware on a worldwide Grid made available by Grid5000 (FR), In Trigger (JP), PoweRcost (IT) and DAS (NL). These jobs were then executed on the Grid during a fixed time slot and results evaluated according to their performance (number of solutions found, ability to use a maximum of grid nodes, CPU time etc). The winner of this 2007 contest is KAAPI-MOAIS, France represented by Xavier Besseron, with ~ 40 379 billion solutions found. Deployed on 3654 workers, it computed for the first time N=23 in 35min 7s and N=22 in 3min 21s.

This highly oriented Grid week was a clear success. It facilitated the exchange of good experiences and best practices on the use of Grid open standards for Grid middleware and applications interoperability, and helped to reinforce existing links between the two Grid communities and to promote long and lasting partnerships.

**Links:**
IV Grids@Work conference:
http://www.etsi.org/plugtests/grid/IVGRID_PLUGTESTS.htm
EchoGrid: http://echogrid.ercim.org/
GridComp: http://gridcomp.ercim.org/
Bridge: http://www.bridge-grid.eu/

**Please contact:**
Florence Pesce
ERCIM office
Tel: +33 49238 5073
E-mail: florence.pesce@ercim.org

# EVOL 2007 - Third International ERCIM Symposium on Software Evolution

by Tom Mens and Kim Mens

*The third edition of the International Symposium on Software Evolution took place in Paris, 5 October 2007, under the auspices of the ERCIM Working Group on Software Evolution. The event gathered people from academia and industry to identify and discuss recent advancements and emerging trends in state-of-the-art research and practice in software evolution.*

The symposium was organised in 'La Maison Internationale' in Paris, in co-location with the 23rd IEEE International Conference on Software Maintenance (ICSM 2007). The organisers were Tom Mens (coordinator of the ERCIM Working Group on Software Evolution), Kim Mens (lecturer at UCL, Belgium), Maja D'hondt (Chargé de Recherches at INRIA) and Ellen Van Paesschen (ERCIM postdoctoral research fellow). Sponsoring was obtained from ERCIM, INRIA and Université des Sciences et Technologies de Lille.

As in previous years, the event was very successful. 37 people from twelve different countries attended, and in total we received 29 regular submissions and three tool demonstra-

tions. An international programme committee of experts in the field took charge of ensuring the quality of accepted papers. Fifteen submissions and two tool demonstrations were accepted for presentation, and eleven of those were selected for inclusion in the post-proceedings, which will be published as part of the scientific open-access journal Electronic Communications of the EASST in 2008. This corresponds to an overall acceptance ratio of 34.4 %.

The presentations covered a wide variety of research topics, including research advances in software refactoring, open-source software evolution, reengineering of legacy code, software product lines, and evolution problems pertaining to component-based, aspect-oriented or model-driven software. The two tool demonstrations focused on the support of feature modelling and maintaining design regularities in evolving code.

**Links:**
WG Web site:
http://w3.umh.ac.be/evol/
Workshop Web site:
http://w3.umh.ac.be/evol/events/evol2007.html
ICSM Web site:
http://conferences.computer.org/icsm/

**Please contact:**
Tom Mens
Université de Mons-Hainaut, Belgium
Tel: +32 65 37 3453
E-mail: tom.mens@umh.ac.be

## Software Evolution

As a joint effort of the ERCIM Working Group on Software Evolution, Tom Mens and Serge Demeyer have edited together a book entitled 'Software Evolution' with contributions by international authorities in the field of software evolution.

Software has become omnipresent and vital in our information-based society, so all software producers should assume responsibility for its reliability. While 'reliable' originally assumed implementations that were effective and mainly error-free, additional issues like adaptability and maintainability have gained equal importance recently. For example, the 2004 ACM/IEEE Software Engineering Curriculum Guidelines list software evolution as one of  ten key areas of software engineering education.

Tom Mens and Serge Demeyer, both international authorities in the field of software evolution, together with the invited contributors, focus on novel

trends in software evolution research and its relations with other emerging disciplines such as model-driven software engineering, service-oriented software development, and aspect-oriented software development. They do not restrict themselves to the evolution of source code but also address the evolution of other, equally important software artifacts such as databases and database schemas, design models, software architectures, and process management. The contributing authors provide broad overviews of related work, and they also contribute to a comprehensive glossary, a list of acronyms, and a list of books, journals, Web sites, standards and conferences that together represent the community's body of knowledge. Combining all these features, this book is the indispensable source for researchers and professionals looking for an introduction and comprehensive overview of the state of  the art. In addition, it is an ideal basis for an advanced course on software evolution.



*Software Evolution*
*Tom Mens, University of Mons-Hainaut, Belgium;*
*Serge Demeyer, University of Antwerp, Belgium (Eds.)*
*Springer, ISBN 978-3-540-76439-7*

# 12th International ERCIM Workshop on Formal Methods for Industrial Critical Systems

by Stefan Leue and Pedro Merino

*Last year's ERCIM Workshop on Formal Methods for Industrial Critical Systems (FMICS) was affiliated with the Computer Aided Verification (CAV) conference, held in Berlin on 1-2 July 2007.*

The aim of the FMICS workshop series is to provide a forum for researchers who are interested in the development and application of formal methods in industry. In particular, these workshops are intended to bring together scientists and practitioners who are active in the area of formal methods and interested in exchanging their experience in the industrial usage of these methods. These workshops also strive to promote research and development for the improvement of formal methods and tools for industrial applications.

The topics for which contributions to FMICS 2007 were solicited included, but were not restricted to, the following:
• design, specification, code generation and testing with formal methods
• verification and validation of complex, distributed real-time systems and embedded systems
• verification and validation methods that aim at circumventing shortcomings of existing methods with respect to their industrial applicability
• tools for the design and development of formal descriptions



*From left: Pedro Merino ERCIM FMICS Working Group coordinator and workshop co-organiser; Ganesh Gopalakrishnan winner of the FMICS best paper award; and Stefan Leue, University of Konstanz, workshop co-organiser.*

• case studies and project reports on formal methods related projects with industrial participation (eg safety critical systems, mobile systems, object-based distributed systems)
• application of formal methods in standardization and industrial forums.

The workshop included five sessions of regular contributions and three invited presentations, given by Charles Pecheur, Thomas Henzinger and Gérard Berry. In addition to the participant's proceedings, a volume in the Lecture Notes in Computer Science (LNCS) series published by Springer Verlag will be devoted to FMICS 2007. The papers included in the LNCS volume were selected from a second round of peer reviewing.

FMICS 2007 attracted 33 participants, some of them are members of the FMICS Working Group, from fourteen different countries.

We wish to thank the members of the programme committee and the additional reviewers for their careful evaluation of the submitted papers (fifteen papers have been selected out of 31 submitted). We are very grateful to the local organisers of the CAV conference and to University of Dortmund for allowing us to use their online conference service.

Following a tradition established over the past years, the European Association of Software Science and Technology (EASST) offered an award to the best FMICS paper. This year's award was given to Robert Palmer, Michael DeLisi, Ganesh Gopalakrishnan and Robert M. Kirby for their paper 'An Approach to Formalization and Analysis of Message Passing Libraries'.

**Link**:
FMICS Working Group and the next FMICS workshop:
http://www.inrialpes.fr/vasy/fmics

**Please contact:**
Pedro Merino
University of Malaga / SpaRCIM, Spain
E-mail: pedro@lcc.uma.es
http://www.lcc.uma.es/~pedro

## ERCIM-Sponsored Events

*ERCIM sponsors up to ten high-quality events per year (conferences, workshops and summer schools).*

*Conditions:*
• *ERCIM sponshorship must be acknowledged in all official printed publicity material relating to the event*
• *promotional and informational material received prior to event commencement must be distributed to the conference participants as a part of the participant's conference package*
• *event organisers must provide free-of-charge attendance to one ERCIM representative, should ERCIM decide to send a PR person. They must also provide a booth in a reasonably prominent place, where this person can present and distribute further materials or provide further information to participants*
• *event organisers must include the ERCIM logo with a link to ERCIM Web pages on the title (reference) page of the conference for the period starting with successful PR sponsorship negotiation, for the remaining existence of these Web pages.*

*If you would like to associate ERCIM to an event, please contact danny.deschreye@ercim.org.*

# ERCIM Working Group Workshops on Dependable Embedded Systems

*The ERCIM Working Group (WG) 'Dependable Embedded Systems' organised two successful workshops in cooperation with DECOS, the Integrated Project 'Dependable Embedded Components and Systems'. Both workshops were chaired by the WG coordinators Erwin Schoitsch from the Austrian Research Centers (ARC) and Amund Skavhaug, NTNU.*

In cooperation with Euromicro, the ERCIM/DECOS workshop 'Dependable Smart Systems: Research, Industrial Applications, Cooperative Systems, Standardization and Certification' took place on 28 August 2007 in Lübeck, Germany. Thirteen participants enjoyed eleven presentations on dependable embedded systems theory and research, validation and certification, industrial (product line development for SMEs) and transport applications. A very engaged discussion – especially on automotive embedded systems visions and futures (road safety) – closed the full day sessions.

The 2nd DECOS/ERCIM/COOPERS workshop 'Dependable Embedded Systems - Challenges, Impact, Solutions, Examples, Professional and Academic Education and Training' took place during the SAFECOMP 2007 conference in Nuremberg, Germany on 18 September 2007. 23 attendees from Europe (including two Russians) discussed dependable embedded systems research, embedded systems trust and security issues, industrial applications such as very large sensor networks, and road safety. A panel of industrial suppliers for automotive and road traffic advanced driver assistance and traffic management embedded systems concluded the event. The papers and discussion results have been collected, and will be reviewed and published as ERCIM workshop proceedings.

Additionally, a special session on dependable embedded systems was organised at INDIN 2007, the Industrial Informatics conference in Vienna, 23-26 July 2007. The papers are part of the conference proceedings published by IEEE.

The WG presented ERCIM at a joint DECOS/ERCIM booth at the workshops described above. ERCIM was also present at the Vehicle Electronics conference in Baden-Baden, Germany, 10-11 October 2007, which was attended by 1500 participants, and the ITEA2 Symposium, held in Berlin from 18-19 October, 2007. ITEA2, a unique opportunity to gain an insight into one of Europe 's most important and strategic ICT initiatives, attracted almost 500 participants.

**Link:**
http://its.arcs.ac.at/ercim/

**Please contact:**
Erwin Schoitsch
Dependable Embedded Systems WG coordinator
Austrian Research Centres GmbH – ARC, Austria
Tel: +43 50550  4117
E-mail: erwin.schoitsch@arcs.ac.at

# ERCIM STM Working Group Award for the best PhD Thesis on 'Security and Trust Management'

*The ERCIM Working Group (WG) on Security and Trust Management (STM) establish an award for the best Ph.D. thesis in this area of security and trust management in order to increase the visibility of the young researchers inside ERCIM as well as in the larger European scientific community.*

The topics of interest include, but are not limited to:
• rigorous semantics and computational models for security and trust
• security and trust management architectures, mechanisms and policies
• networked systems security
• privacy and anonymity
• identity management
• ICT for securing digital as well as physical assets
• cryptography.

Applications for the 2008 award are open to all PhD holders who defended their thesis during 2007 in any European university. Applications consisting of the PhD thesis (in pdf format), a short summary of the thesis in English, and letters of support from researchers from at least two different ERCIM STM WG member institutions must be submitted by email not later than 20 February 2008 to the STM Working Group coordinator Fabio Martinelli (Fabio.Martinelli@ercim.org).

Applications will be evaluated by a committee and judged on the basis of scientific quality, originality, clarity of presentation, as well as potential impact of the results. The committee members for the 2008 competition are:
• Gilles Barthe, INRIA
• Theo Dimitrakos, BT
• Dieter Gollmann, Hamburg University of Technology
• Günter Karjoth, IBM
• Javier Lopez, University of Malaga
• Fabio Martinelli, CNR
• Mogens Nielsen, University of Aarhus
• Jean-Jacques Quisquater, Université catholique de Louvain
• Peter Ryan, University of Newcastle
• Pierangela Samarati, University of Milan

The award ceremony will be held during the 4th International Workshop on Security and Trust Management (STM08) on 16-17 June in Trondheim, Norway, collocated with EUROPKI08 and IFIP TM08. The winner will be invited for giving a lecture at STM08.

**Link:**
STM WG: http://www.iit.cnr.it/STM-WG

**Please contact:**
Fabio Martinelli, IIT-CNR, Italy
E-mail: Fabio.Martinelli@ercim.org

# The European Commission is launching 24 new Research Projects on ICT Security for €90 million under the FP7-ICT Theme

by Jacques Bus

*Twenty-four new research projects are being launched by DG INFSO following the result of the first call for proposals in the area of 'secure, dependable and trusted infrastructures', which is part of the 7th Framework Programme (FP7) Information and Communication Technologies (ICT) theme. Most of the projects started on 1st January 2008 and all will be operational within the next few months.*

Among the 24 projects, there are five Integrated Projects (IPs), one Network of Excellence (NoE), fourteen Focused Research Actions (STREPs) and four Coordination Actions (CAs). The duration of individual projects varies from two to four years and the funding is ranging from less than €1 million (CAs) to more than €9 million (some IPs). The projects cover an extensive range of hot security-related research topics in ICT. The overall focus of the research is on developing knowledge and technologies needed for building a trustworthy ubiquitous information society, based on secure and resilient network and service infrastructures. A central issue of the new research is the empowerment of end-users through proper handling of their digital identities and respecting their privacy when they interact in the digital world. The research topics can be broadly grouped into four closely inter-related thematic areas, with each project touching upon several areas (see figure).

## 1. Dynamic, Reconfigurable Service Architectures
Modern IT systems tend to replace individual components and applications with compositions of services distributed over a network. It is vital to define and ensure trust, security and dependability properties when composing services, either at design time or dynamically at run-time. Four research projects investigate the implications of this paradigm shift in service-oriented architectures in terms of trust and security.

## 2. Network Infrastructures
Heterogeneous network infrastructures that are secure, resilient and able to cope with different attacks and threats are needed in order to deliver trustworthy and always available applications and services. Research funded in this area covers, on the one hand the need to design, build and maintain security in such network infrastructures. On the other hand, it covers threat prevention mechanisms with a focus on the early identification and modelling of emerging threats. This will enhance the ability of operators to detect anomalies and to counter possible cyber attacks.

## 3. Identity Management, Privacy, Biometrics, and Trust Policies
As users are bringing more and more of their corporate and private data within the realm of the Internet, they realise the potential, and also the risks that technology has created for the protection and management of sensitive interactions, information and data. Several projects investigate new technological solutions for the secure handling of digital identities and sensitive data by adopting a user-centric, privacy-enhancing federated identity approach. Specific research areas addressed include: new privacy and identity management schemes for social networking and Web 2.0 environments and for complex real-life applications like healthcare; novel solutions combining biometrics and crypto-hashing methods for fighting digital identity theft; new concepts in multi-modal biometric authentication; and innovative biometric authentication means for new mobile services.

## 4. Underpinning Technologies for Trustworthy Infrastructures
Four new projects, one of which is a Network of Excellence, will develop a number of technology building blocks underpinning trustworthy network and service infrastructures.
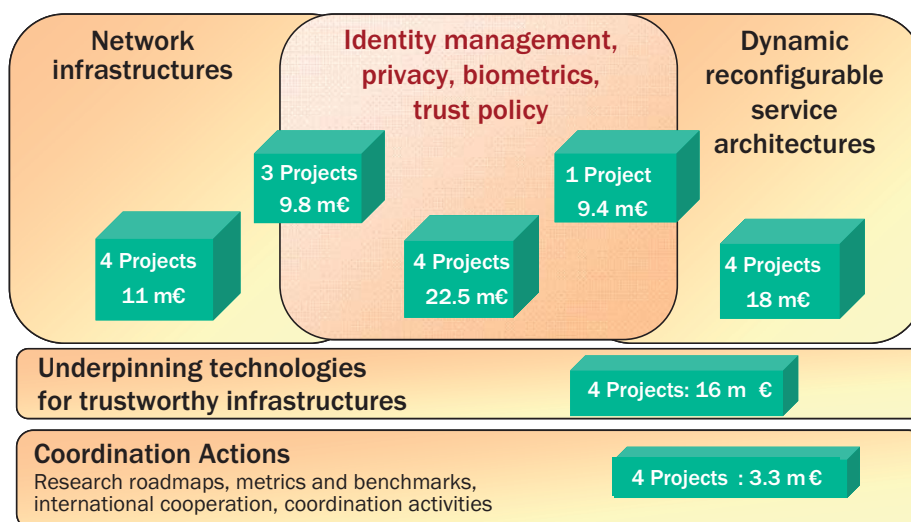


*ICT Research in 'secure, dependable and trusted infrastructures': research areas covered by the 24 new FP7 projects.*

These projects will mainly focus on: achieving further advances in cryptographic algorithms and protocols and their efficient hardware and software implementations; developing trusted computing architectures for embedded computing platforms; and, building tools to help programmers in designing and developing more secure and robust software.

## Coordination Actions

Finally, four new Coordination Actions (CAs) will address sharing of knowledge and collaboration between researchers in various areas of ICT security. Two of them cover the coordination of research, respectively, in cyber-threat defence and in resilience measuring and benchmarking in computer systems and components. Another CA supports and coordinates international collaboration with a number of developed countries. A fourth one addresses the definition of new areas of work in ICT Security Research, to further EU's strategic thinking and positioning in the field.

*Disclaimer: The content of this paper is the sole responsibility of the author and in no way represents the view of the European Commission or its services.*

**Link:**
Synopsis of the new projects funded under FP7 in the field of ICT security:
http://cordis.europa.eu/fp7/ict/security/projects_en.html

**Please contact:**
Jacques Bus, Head of Unit "Security", DG Information Society and Media, European Commission
E-mail: Jacques.Bus@ec.europa.eu

# European Digital Library Foundation welcomed by the Commissioner

*The European Commission endorsed the work of the European Digital Library Foundation when high level EC officials met members for the formal handover of their statutes on 27 November 2007.*

"Europe's citizens should all be able to enjoy our rich cultural heritage. This Foundation is a significant step towards making that ambition come true," Commissioner Viviane Reding, responsible for Information Society and Media commented. "It shows the commitment of Europe's cultural institutions to work together to make their collections available and searchable by the public through a common and multilingual access point online."

Foundation members include the key European heritage and information associations. Their statutes commit members to work in partnership to:
• provide access to Europe's cultural and scientific heritage though a cross-domain portal
• co-operate in the delivery and sustainability of the joint portal

• Stimulate initiatives to bring together existing digital content
• Support digitisation of Europe's cultural and scientific heritage

The European digital library is developing its prototype site for launch next year. To coincide with the formal handover to the Commissioner, the Foundation announced the CITY as the first of the site's themes.

The CITY is a broad theme that will enable the prototype site to show the European urban experience from several perspectives. Emerging ideas include:
• cities of the future/cities of the past
• migration and diaspora
• trade and industry
• design, shopping and urban cool
• pox, cholera and the plague: the route to urban health
• archaeology and architecture
• utopias and cities of the imagination
• riot and disorder
• palaces and politics

The digital library project is gathering digitised content from European archives, museums, audio-visual collections and libraries. It will use maps, artefacts, photos, sound, film material, books, archival records and artworks to explore two millennia of connectivity between Europe's cities.

Dr Wim van Drimmelen, foundation member and Director of the Koninklijke Bibliotheek, the Netherlands, which hosts the European digital library initiative, said, "archives, museums, audio-visual collections and libraries are collaborating in order to guarantee that their resources can be brought together in a virtual world, regardless of where the original is held. It's what users expect. However, the benefits are not just for the users. The development work is providing an excellent forum for knowledge transfer between the domains, which increases our collective ability to respond to changing user needs and remain relevant in the fast-moving digital environment."

The European digital library is hosted by the Koninklijke Bibliotheek, the national library of the Netherlands, and led by the Conference of European National Librarians. The initiative is one of the Commission's flagship i2010 initiatives.

In August 2006, the Commission adopted a Recommendation on digitisation and digital preservation which urged European Union member states to set up large-scale digitisation facilities, so as to accelerate the process of getting Europe's cultural heritage online via the European digital library. In November 2006 the idea of a European digital library was strongly endorsed by the Culture Ministers of all EU Member States and was recently backed by the European Parliament in its resolution of 27 September 2007.

**Link:**
http://www.europeandigitallibrary.eu/edlnet/

**Please contact:**
Jonathan Purday
Digital Library Foundation
E-mail: Jonathan.purday@kb.nl

Introduction to the Special Theme

# The Path to Web n+1

by Lynda Hardman and Steven Pemberton

The Sapir-Whorf Hypothesis postulates a link between thought and language: if you haven't got a word for a concept, you can't think about it; if you don't think about it, you won't invent a word for it. The term "Web 2.0" is a case in point. It was invented by a book publisher as a term to build a series of conferences around, and conceptualises the idea of Web sites that gain value by their users adding data to them. But the concept existed before the term: Ebay was already Web 2.0 in the era of Web 1.0. But now we have the term we can talk about it, and it becomes a structure in our minds, and in this case a movement has built up around it.

There are inherent dangers for users of Web 2.0. For a start, by putting a lot of work into a Web site, you commit yourself to it, and lock yourself into their data formats. This is similar to data lock-in when you use a proprietary program. You commit yourself and lock yourself in. Moving comes at great cost. This was one of the justifications for creating the eXtended Markup Language (XML): it reduces the possibility of data lock-in - having a standard representation for data helps using the same data in different ways too.

As an example, if you commit to a particular photo-sharing Web site, you upload thousands of photos, tagging extensively, and then a better site comes along. What do you do? How about if the site you have chosen closes down (as has

happened with some Web 2.0 music sites): all your work is lost. How do you decide which social networking site to join? Do you join several and repeat the work? How about geneology sites, and school-friend sites? These are all examples of Metcalf's law, which postulates that the value of a network is proportional to the square of the number of nodes in the network. Simple maths shows that if you split a network into two, its value is halved. This is why it is good that there is a single email network, and bad that there are many instant messenger networks. It is why it is good that there is only one World Wide Web.

Web 2.0 partitions the Web into a number of topical sub-Webs, and locks you in, thereby reducing the value of the network as a whole.

The advantage of the semantic Web, however, is that it allows your data to be distributed over the Web and not centralised. By adding metadata to your data it allows aggregation services to offer the same functionalities as Web 2.0 without the lock-in and potential loss of data should the service die. To enable the inclusion of non-textual media in the semantic Web we need to extend existing languages, and be aware that existing interfaces are based on assumptions that may no longer apply. Articles in this special issue describe emerging technologies that will contribute to enabling making use of modalities beyond the visual and audible, for the whole, mobile, Web population.

The Web 2.0 phenomenon could not have been predicted when the first Hypertext Transfer Protocol (http) was being developed. Similarly, as we are going beyond human-created links to machine integrated data, we cannot forsee the applications and social phenonena which may emerge, even in the near future.

In this special issue we are not seeking to predict where the Web will go along its apparently accelerating path, but give insights into emerging core technologies that will enable the continuing revolution in our species' relationship to creating, using and maintaining information.

**Please contact:**
Lynda Hardman, CWI
and Eindhoven University of Technology, The Netherlands
E-mail: Lynda.Hardman@cwi.nl

Steven Pemberton, CWI, The Netherlands and W3C
E-mail: Steven.Pemberton@cwi.nl

# Which Future Web?

An interview with Frank van Harmelen

**The semantic Web will be a considerable part of the future Web. What is the difference between the semantic Web and artificial intelligence? And what about Web 2.0? Frank van Harmelen, computer scientist in the Netherlands and a specialist in the semantic Web, answers some questions.**

*The semantic Web initiative is often said to address the same issues that have already been approached for thirty years in Artificial Intelligence. What makes the semantic Web, with its focus on ontologies and reasoning so different?*

There is indeed a widespread misconception that the semantic Web is 'AI all over again'. Even though the two may have tools in common (eg ontologies, reasoning, logic), the goals of the two programmes are entirely different. In fact, the goals of the semantic Web are much more technical and modest. Rather than seeking to build a general-purpose all-encompassing global Internet-based intelligence, the goal is instead to achieve interoperability between data sets that are exposed to the Web, whether they consist of structured, unstructured or semi-structured data.

Tim Berners-Lee (W3C) devoted an entire presentation to the confusion between AI and semantic Web in July 2006 (see the link below). This presentation also does a very good job of busting some of the other myths surrounding the semantic Web, such as that the semantic Web is mainly concerned with hand-annotated text documents, or that the semantic Web requires a single universal ontology to be adopted by all.

*Web 2.0 appears to be the new kid on the block - everybody's darling, loved both by academia and industry. The semantic Web, on the other hand, has fallen from grace, owing to numerous unmet promises. How do you regard the coexistence of these two Webs and what role will Web 2.0 assume in the semantic Web's story?*

My feeling is that this question is based on a false premise, namely that "the semantic Web has fallen from grace, owing to numerous unmet promises". The SemTech conference, an annual industry-oriented event organised in the past three years in San Jose, California, attracted 300 attendants in 2005, 500

*Frank van Harmelen.*

Picture: Irma Bulkens.

attendants in 2006, and 700+ attendants in 2007. Its European counterpart, the European Semantic Technologies Conference, attracted 200+ attendants to its first event in Vienna in May 2007, of whom 75% were from companies. This would suggest that research and interest in the semantic Web are alive and well.

In fact, semantic technology is in the process of an industrial breakthrough. Here is a quote from a recent (May 2007) Gartner report, the industry watcher not known for its love of short-lived hypes:

"Key finding: During the next ten years, Web-based technologies will improve the ability to embed semantic structures in documents, and create structured vocabularies and ontologies to define terms, concepts and relationships. This will offer extraordinary advances in the visibility and exploitation of information - especially in the ability of systems to interpret documents and infer meaning without human intervention." And: "The grand vision of the semantic Web will occur in multiple evolutionary steps, and small-scale initiatives are often the best starting points."

Turning to the substance of your question: There is widespread agreement in the research world that Web 2.0 and semantic Web (or Web 3.0) are complementary rather than in competition. For example, a science panel at the WWW07 conference in May 2006 in Edinburgh came to the following consensus: Web 2.0 has a low threshold (it's easy to start using it), but also has a low ceiling (folksonomies only get you so far), while Web 3.0 has a higher threshold (higher startup investments), but has a much higher ceiling (more is possible).

The aforementioned Gartner report has useful things to say here as well. It advises the combination of semantic Web with Web 2.0 techniques, and predicts a gradual growth path from the current Web via semantically light-weight but easy to use Web 2.0 techniques to higher-cost/higher-yield Web 3.0 techniques.

*And what about automated means of learning ontologies, relationships between entities, and so forth - that is, resorting to natural language processing, text mining, and statistical means of knowledge extraction and inference. Do you regard these techniques as complementary to the manual composition of ontologies or rather inhibitory?*

My attitude towards the acquisition of ontologies and the classification of data objects in these ontologies is: if it works, it's fine. Clearly relying solely on the manual construction of ontologies puts a high cost and a low ceiling on the volume of knowledge that can be coded and classified. Hence, I expect that the techniques you mention will play an ever-bigger role in the range of semantic technologies. I see no reason why such techniques are 'bound to fail' ? instead I am rather optimistic about their increasingly valuable contribution.

*All great technological inventions and milestones are marked by the advent of a killer application. What could/will be*

*the semantic Web's killer app? Will there be one at all?*

I find the perennial question for the 'killer app' always a bit naive. For example: we can surely agree that the widespread adoption of XML was an important technical innovation. But what was XML's 'killer app'? Was there a single one? No. Instead there are many places where XML facilitates progress 'under the hood'. Semantic Web technology is primarily infrastructure technology. And infrastructure technology is under the hood, or in other words, not directly visible to users. One simply notices a number of improvements. Web sites become more personalized, because under the hood semantic Web technology is allowing your personal interest profile to be interoperable with the data sources of the Web site. Search engines provide a better clustering of results, because under the hood they have classified search results in a meaningful ontology. Desktop search tools become able to link the author names on documents with email addresses in your address book, because under the hood, these data formats have been made to interoperate by exposing their semantics. However, none of these applications will have 'semantic Web technology' written on their interface. Semantic Web technology is like Nikasil coating in the cylinders of a car: very few car drivers are aware of it, but they are aware of reduced fuel consumption, higher top speeds and the extended lifetime of the engine. Semantic Web technology is the Nikasil of the next generation of human-friendly computer applications that are being developed right now.

**Links:**
http://www.w3.org/2006/Talks/
0718-aaai-tbl/Overview.html
http://www.cs.vu.nl/~frankh/

**Please contact:**
Frank van Harmelen
Vrije Universiteit Amsterdam,
The Netherlands
Tel: +31 20 598 7731/7483 (secr)
E-mail: Frank.van.Harmelen@cs.vu.nl

# KAON2 – Scalable Reasoning over Ontologies with Large Data Sets

by Boris Motik

*Scalability of ontology reasoning is a key factor in the practical adoption of ontology technologies. The KAON2 ontology reasoner has been designed to improve scalability in the case of reasoning over large data sets. It is based on a novel reasoning algorithm that builds upon extensive research in relational and deductive databases.*

Ontologies – vocabularies of terms often shared by a community of users – are being applied in science and engineering disciplines as diverse as biology, geography, astronomy, agriculture and defence. Nowadays, ontologies are usually expressed in the W3C standard language called the Web Ontology Language (OWL). OWL ontologies consist of a schema part, called a TBox, which describes the concepts and relationships in the domain of discourse, and a data part, called an ABox, which describes the actual data in the application. An efficient reasoner is the cornerstone of most OWL-based applications. It implements the formal semantics of OWL and thus provides the application with query answering capabilities.

While reasoning over OWL ontologies is a provably intractable computational problem, it has been observed that the ontologies encountered in practice rarely involve a combination of constructs that leads to intractability. By relying on sophisticated optimizations, reasoners were developed that can handle ontologies with large Tboxes, yet these still do not provide adequate performance on ontologies containing large ABoxes. This has so far prevented the usage of OWL in applications that depend on large data sets, such as metadata management and information integration.

Parallel to the development of reasoning techniques for OWL, significant effort has been invested into improving the scalability of relational and deductive databases. In particular, numerous optimizations of query answering in (disjunctive) datalog (a widely used deductive database language) are known and have proven themselves effective in practice. It is therefore natural to try to improve the scalability of ABox reasoning in OWL by building on this large body of existing work.

This idea has been realized in a new reasoner called KAON2. The architecture of the reasoner is shown in Figure 1. The central component of the system is the reasoning engine, which implements a completely new reasoning algorithm. A query Q over an ontology consisting of a TBox T and an ABox A is answered by first reformulating T as a set of clauses in first-order logic, and then transforming the result into a disjunctive datalog program DD(T). The latter step is the key part of KAON2: based on certain novel results in resolution-based theorem proving, it ensures that all answers of Q over T and A can be equally computed by evaluating Q over DD(T) and A. The main benefit of such a transformation is that, to evaluate Q w.r.t DD(T) and A, the disjunctive datalog engine can use the optimization techniques known from deductive databases, such as (disjunctive) magic sets or join-order optimizations.

This approach to query answering has shown itself to be practical and effective in cases where the TBox is rather simple but the ABox contains large amounts of data. On such ontologies, KAON2 has shown performance improvements over the state of the art of one or more orders of magnitude. KAON2 has thus become the platform of choice for numerous research proj-
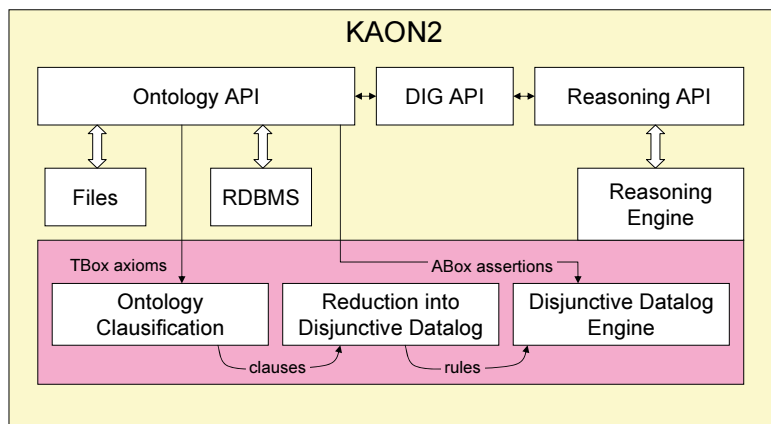
*Figure 1: KAON 2 architecture.*

ects, such as FIT (EU IST 27090), OntoGov (EU IST 507237), NeOn (EU IST-2005-027595), X-Media (EU FP6-26978), and KnowledgeWeb (EU FP6-507482). Furthermore, ontoprise GmbH, the vendor of ontology-based software infrastructure based in Karl-sruhe, Germany, is integrating KAON2 into its product suite and is using the tool in a commercial setting.

KAON2 is written in Java and can be used free of charge for non-commercial purposes. The tool has emerged as a result of the author's PhD work at the University of Karlsruhe, Germany. Its development was continued at the University of Manchester, UK, and is currently taking place at the University of Oxford, UK.

*Boris Motik has been awarded the 2007 Cor Baayen Award for a most promising young researcher in computer science and applied mathematics by ERCIM.*

**Links:**
http://kaon2.semanticweb.org/
http://www.ontoprise.de/

**Please contact:**
Boris Motik
Computing Laboratory,
Oxford University, UK
E-mail: boris.motik@comlab.ox.ac.uk

# The WIKINGER Project – Knowledge-Capturing Tools for Domain Experts

by Lars Bröcker

*The semantic Web offers exciting opportunities for scientific communities: knowledge bases with underlying ontologies promote the process of collaborative knowledge creation and facilitate research based on the published body of work. However, since most scientific communities do not possess the knowledge required to build an ontology, these opportunities tend not to be taken up. The WIKINGER project aims to provide tools that largely automate the creation of an ontology from a domain-specific document collection.*

There is a multitude of interesting content on the Web that would greatly benefit from inclusion into the semantic Web, eg cultural heritage collections, educational resources and scientific digital libraries. These collections offer a plethora of possibilities for knowledge applications if only they could be made accessible through ontologies. However, as long as domain experts require the assistance of ontology engineers to even begin the process of building a shallow ontology for their own domains, this is unlikely to happen.

Although there are tools available that guide the creation of ontologies, they are pitched to the level of an experienced ontology engineer. Domain experts wanting to create an ontology must therefore become ontology-creation experts before they can make use of these tools. A more sensible situation would be to provide tools to guide the ontological layperson through the process of creating an ontology for their domain, by automating the creation as much as possible.

WIKINGER is a joint project involving the department for Computer Linguistics at the University of Duisburg-Essen, the Fraunhofer-Institut für Intelligente Analyse und Informationssystem (IAIS), and the Commission for Contemporary History (KfZG). The goal is to create ontologies from collections of scientific documents; these are accessed using a Wiki containing the entities and concepts relevant to the domain, as well as qualified associations connecting them.

Named entity recognition (NER) is employed to automatically extract from the text collection entities of different concept classes. Domain experts provide examples for these concept classes using WALU, the WIKINGER Annotation and Learning Environment. Figure 1 shows a typical screenshot of the tool in action. Entities are marked up in different colours in the text, and are thus easily discerned at a glance. They are then assigned to a concept class by pressing the corresponding button in the lower part of the editor. WALU contains helper functions that mark up entities automatically, eg date/time information, previous annotations or entities provided in list form. The results of this step are sent back to the WIKINGER server, which then trains the extraction algorithms to extract entities on the larger scale of the document collection.

Whereas this step provides the nodes of the semantic network, the following step is concerned with the provision of the edges connecting these nodes. These edges represent the relations between the different concept classes. They are determined in a relation discovery module that first discerns all statistically relevant relations between classes in a training part of the corpus. The association rule mining and clustering algorithms are applied to this task. The association rules highlight interesting concept combinations in the text collection. The relations present in the occurrences of different rules are separated using clustering algorithms.

The relation discovery process can be governed by domain experts using a Java application. It generates the association rules, performs the clustering and allows the selection and labelling of relations that are to be included in the ontology. The different clusters can be reviewed on screen, meaning they can be merged with other similar clusters, discarded in the event of mistakes, and finally labelled if they fit into the ontology that is to be created. Results are sent back to the server that applies them on the automatically extracted entities in order to populate the ontology. The server then provides access to the ontology for Web applications or via a SPARQL Protocol and RDF Query Language.

The pilot application for our system focuses on the domain of contemporary history, in particular the social and political history of German Catholicism. Historians from the KfZG are already using the tools described above for the creation of annotation examples and the subsequent selection of appropriate relations for the ontology describing their domain. They are reporting no difficulties in adapting to these new tools. The metaphors used in WALU imitate known behaviours used in marking up text passages, and thus integrate easily into a familiar pattern of workflow. Slightly more effort is required to become acquainted with the relation discovery application, since the actions performed here are new to the experts.

The prototype of the system is in its final development stages, and its release to the scientific community is planned for early 2008. Work conducted on different data sets shows the transferability of the approach: WALU for example achieved second place in the EVALITA shared task on NER in Italian newspapers in 2007.

The project is funded by the German Federal Ministry of Research and Education in the programme 'e-Science und Wissensvernetzung' (e-Science and Knowledge Networking). Work on the project began in October 2005, and will be completed in September 2008.

**Link:**
http://www.wikinger-escience.de

**Please contact:**
Lars Broecker
Fraunhofer-Institut Intelligente
Analyse- und Informationssysteme,
Germany
Tel: +49 2241 14 1993
E-mail:
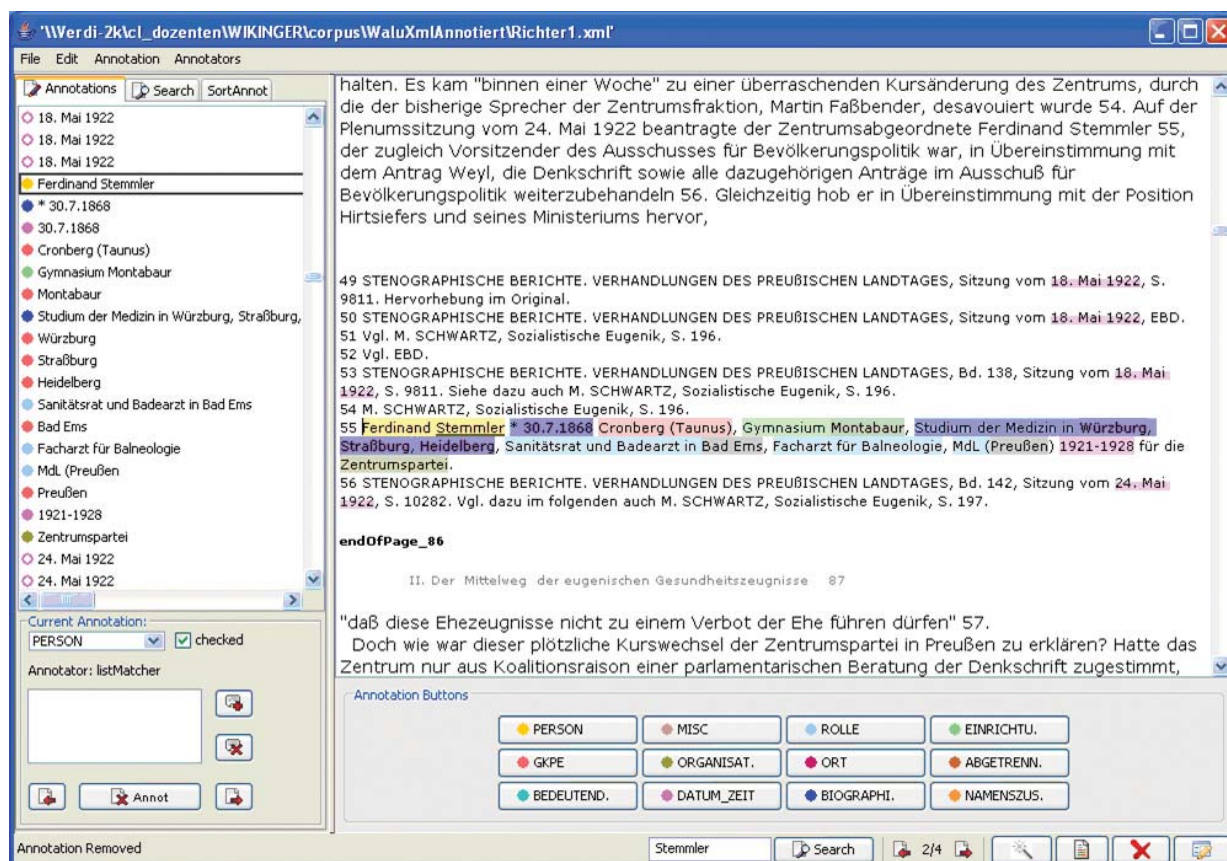Lars.Broecker@iais.fraunhofer.de



*Figure 1: Screenshot of WALU.*

# Tagging and the Semantic Web in Cultural Heritage

by Kees van der Sluijs and Geert-Jan Houben

*While it is generally desirable that huge cultural collections should be opened up to the public, the paucity of available metadata makes this a difficult task. Researchers from the Eindhoven University of Technology in the Netherlands have built a Web application framework for opening up digital versions of these multimedia documents with help of the public. This leads to a win-win situation for users and content providers.*

The Regional Historic Centre Eindhoven (RHCe) governs all historical information related to the cities in the region around Eindhoven in the Netherlands. The information is gathered from local government agencies or private persons and groups. This includes not only enormous collections of birth, marriage and death certificates, but also posters, drawings, pictures, videos and city council minutes. One of its goals is to make these collections available to the general public. However, for the videos and pictures very little metadata is available, which makes indexing this data for navigation or searching very hard. Moreover, most of the fragile material is stored in vaults and is thus physically inaccessible to the public.

Through the use of simple tagging techniques, users could collaboratively help to provide metadata for previously uncharted collections of multimedia documents. By applying semantic and linguistic techniques these user tags can be enriched with well-structured ontological information provided by professionals. This gives content providers access to high-quality metadata with little effort, eg, for archival purposes. In return, the rich metadata in combination with ontological sources offers users more powerful navigation and search facilities, allowing them to get a grip on the 'sometimes very large' multimedia collections and find what they were really looking for.

Web engineering researchers from the Eindhoven University of Technology took on the challenge of building a Web application framework that would open up digital versions of these multimedia documents to the public in a meaningful way. A prototype framework, called Chi, provides a faceted browser view over the different data sets. This means that the data can be navigated or searched via a number of different dimensions.

Photo and video collections share three dimensions within scenes: time, location and a set of keywords that are associated with a scene. In Chi these dimensions are described by a specialist ontology. Professionals adapt the metadata schema of every collection by aligning notions of these dimensions with those in the ontologies. Because of these aligned ontologies users can homogeneously navigate heterogeneous collections of multimedia.

For every dimension Chi has a separate visualization in the user interface. For time there is a timeline, for location Google Maps is used, and for the keywords we present a graph which represents the relatedness of terms. In this way, results sharing a characteristic in one dimension can be grouped in these interfaces. This representation means that for a given picture or video, other pictures and videos that share a characteristic can be found.

The RHCe institute is aiming for good metadata descriptions of all its archived data, but its few domain specialists have only limited time and the collections are huge. For them, the
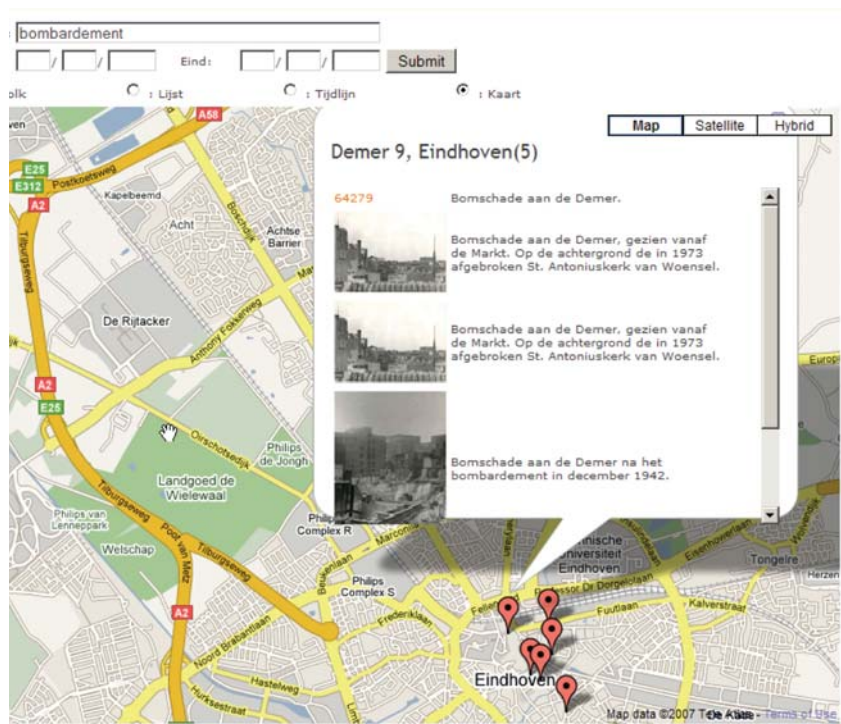


*Figure 1: Visualization of location using Google Maps (in Dutch).*

biggest benefit from Chi is getting metadata from the users. However, users do not want to fill in large forms to provide well-structured data about the scenes: many will find this too complicated or too time-consuming. Therefore, Chi uses several tricks to keep things simple while still obtaining this information. First it uses an ordinary tagging mechanism by which users can enter keywords or small sentence fragments, called tags, to describe a scene. Then via several matching mechanisms that take syntax, semantics and user reinforcement into

account, users are presented with alternatives to their input tags. These can be clicked on; while users are not obliged to click on an alternative, if they do, this is registered. If users often select a specific alternative for some tag, we assume there is some kind of relationship between the original and the selected tags, and we use this relationship to improve tagging suggestions and user queries.

Users can see the tags of others and can agree or disagree with a tag by clicking a thumbs-up or a thumbs-down icon. If many people agree on a certain tag for a certain scene we assume the probability is high that this tag is a correct description for a scene. Finally these tags are presented to the RHCe domain specialist who provides a final judgment. In effect, this information is reused to identify those users who often present good tags and those who don't. The good taggers are given more weight to calculate if a tag is useful.

Chi is currently being finalized in a second version and prepared for more extensive user testing. The first version of Chi has already been tested by end-users, and their responses and feedback were very positive and encouraging. This has motivated RHCe to continue with the researchers along this path of applying semantic technology.

**Links:**
http://wwwis.win.tue.nl/~ksluijs/
http://wwwis.win.tue.nl/~hera/
http://www.icwe2007.org

**Please contact:**
Kees van der Sluijs
and Geert-Jan Houben
Eindhoven University of Technology,
The Netherlands
Tel: +31 40 247 5154
E-mail: k.a.m.sluijs@tue.nl,
G.J.Houben@tue.nl

# Infrastructure for Semantic Applications - NeOn Toolkit Goes Open Source

by Peter Haase, Enrico Motta and Rudi Studer

*The NeOn project is investigating the entire life cycle of networked ontologies that enable complex, semantic applications. As the amount of semantic information available online grows, semantic applications are becoming increasingly ubiquitous, Web-centric and complex. The NeOn Toolkit and the NeOn methodology lie at the core of the NeOn vision, defining the standard reference infrastructure and the standard development process for creating and maintaining large-scale semantic applications. The first version of the NeOn Toolkit for ontology engineering has just been released in open source by the Neon Consortium.*

NeOn is a 14.7 million euro project involving fourteen European partners. It started in March 2006 and will have a duration of four years. NeOn addresses the complete R&D cycle of the emerging generation of semantically enriched applications, which exist and operate in an open environment of highly contextualized, evolving and networked ontologies. It aims to achieve and facilitate the move from feasibility in principle, to concrete cost-effective solutions that can support the design, development and maintenance of large-scale, semantic-based applications. In particular, the project is investigating methods and tools for managing the evolution of networked ontologies, for supporting the collaborative development of ontologies, and for the contextual adaptation of semantic resources.

NeOn has developed an open service-centred reference architecture for managing the complete life cycle of networked ontologies and metadata. This architecture is realized through the NeOn Toolkit and complemented by the NeOn methodology for system development using networked ontologies.

## NeOn Toolkit Launched
The NeOn Toolkit provides a next-generation ontology engineering environment for semantic applications. It is designed around an open and modular architecture, which includes infrastructure services such as registry and repository, and supports distributed components for ontology management, reasoning and collaboration in networked ontologies. Building on the Eclipse platform, the Toolkit provides an open framework for plug-in developers. The NeOn Toolkit is freely available in open source as the reference implementation of the NeOn architecture. The first open-source release of the Toolkit was just recently launched at the latest NeOn Glowfest at ISWC07 – the International Semantic Web Conference – in Busan, Korea on the 13th November. NeOn Glowfest is a series of seminars and discussions intended to gather together the members – both users and developers – of the NeOn community.

## NeOn in Practice
NeOn pays special attention to integrating research with work practices. The methodology, toolkit and infrastructure are therefore intertwined with their deployment and testing from the early phases of the project. NeOn uses a case-centred methodology, which means that our research results are applied to real-world case studies involving partners from industry and public bodies.

Three use cases provide testbeds for NeOn technology that show how the research innovations can be applied in practice and potentially exploited. We apply NeOn technology in two sectors with early-adopter businesses, where we can achieve rapid and significant impact. The two case studies are carried out as follows:
• Ontology-driven stock over-fishing alert system is a study led by the Food and Agriculture Organisation of the UN, which focuses on the agricul-

*Figure 1: NeOn helps ensuring sustainable fishery with semantic technologies.*

tural sector and information management for hunger prevention. As the title suggests, it involves the management of alerts to avoid over-fishing in already stretched global waters.

- Supporting collaboration in pharmaceutical industry is a case study concerned with infrastructure and APIs to bridge the currently used proprietary systems for managing financial and product knowledge interoperability in the networks/clusters of pharmaceutical labs, companies and distributors in Spain.

The NeOn project is co-funded by the European Commission's Sixth Framework Programme under grant number IST-2005-027595. The consortium includes European universities, businesses and user organisations. Among the research institutions are world-leading groups in the fields of ontology, collaborative technology, context management and human-computer interaction, such as the Open University and the University of Sheffield from the UK, the Universities of Karslruhe and Koblenz-Landau from Germany, University Polytechnic Madrid (UPM) from Spain, INRIA, the Josef Stefan Institute from Slovenia and the National Research Council of Italy. The participating companies are Software AG, iSOCO, Ontoprise and Atos Origin.

**Links:**
http://www.neon-project.org/
http://www.neon-toolkit.org/

**Please contact:**
Peter Haase
Institute AIFB
University of Karlsruhe, Germany
Tel: +49 721 608 3705
E-mail: haase@aifb.uni-karlsruhe.de

# Bridging the Clickable and Semantic Webs with RDFa

by Ben Adida

*RDFa, a W3C proposal about to enter Last Call, will help bridge the clickable and semantic Webs. Publishers of current HTML Web pages can augment their output with interoperable, structured data. Web users can extract and process this data straight from its pleasant visual rendering within their browsers. RDFa will enable semantic Web data to progressively emerge from the existing, visual Web.*

Semantic Web technologies, and in general the components of the 'data Web', are maturing at a rapid pace. A number of commercial products support the design of RDF (Resource Description Framework) schemas and the publication of RDF datasets. Key semantic-Web specifications, SPARQL (SPARQL Protocol and RDF Query Language) and GRDDL (Gleaning Resource Descriptions from Dialects of Languages), have recently reached Recommendation status with the W3C. One of the important questions remaining is how to connect the Semantic Web to the existing 'clickable Web', the Web just about everyone knows and loves. RDFa, a W3C proposal on the verge of entering Last Call, is one promising answer, a potential bridge between the clickable and semantic Webs. HTML already contains some constructs to provide the beginnings of data structure, for example with the *rel* attribute. Publishers are anxious to achieve this kind of functionality: Google recommends the use of this *rel* attribute with the keyword '*nofollow*' to indicate that search engines should 'not follow' this link in their search relevance algorithms, eg when a publisher attempts to increase their URL's standing by posting it in multiple blog comments.

With RDFa, an HTML author can augment presentational markup with RDF structure. This begins very simply, by providing extra information about clickable links using the existing HTML *rel* attribute. In the following example, we declare a copyright license for the current document:

```
This document is licensed under a
<a rel="license"
href="http://creativecommons.org/
licenses/by/3.0/">
   Creative Commons license</a>
```

Note how the clickable link's *href* is used both for rendering purposes and for machine readability. This idea, often called DRY (Don't Repeat Yourself), is a key principle of the RDFa design.

A Web publisher may want to make statements about more than the current

page, of course, and RDFa fully supports this. An image may be Creative-Commons licensed as follows:

```
<div about="photo1.jpg">
  This photo is licensed under a
<a rel="license"
href="http://creativecommons.org/
licenses/by/3.0/">
  Creative Commons license
</a>
</div>
```

With RDFa, an author is not limited to the existing reserved keywords in the baseline HTML vocabulary (next, prev, license, etc.). It is trivial to import the Dublin Core vocabulary, for instance, using the xmlns declaration mechanism. Once this is done, an author can use the vocabulary in her document. The following example does just that, using, this time, the *property* RDFa attribute to indicate a literal field, rather than a URL:

```
This document was created by:
<span property="dc:creator">
  Ben Adida
</span>.
```

RDFa supports most of the power of RDF, including the ability to create more complex graphs of data. For example, if we want to provide more information about the creator of this page, we can introduce a blank node to which we attach a number of additional data fields. The *rel* attribute is used without a corresponding *href*, which indicates the creation of a blank node, and the contained RDFa statements are automatically applied to this blank node.

```
This document was created by:
<div rel="dc:creator">
<span property="foaf:name">
  Ben Adida</a>,
  [<a rel="foaf:mbox"
href="mailto:ben@adida.net">
email</a>]
</span>.
```

In the above example, the creator of the current document is a an entity named Ben Adida, with email address ben@adida.net. Additional fields, including a phone number, a physical address, an organizational affiliation, etc. can be added just as easily, and further data depth, such as the physical address of the author's organizational affiliation, can be provided.

A number of RDFa tools are emerging, with parsers in a number of programming languages including Python, Java, PHP, Ruby, and JavaScript [1]. Within the browser, the RDFa Buttons [2] can be used in any Web browser to highlight RDFa structured data and convert it to RDF/XML or RDF/N3. Developers can easily extend these code snippets to provide specific actions based on this data, eg preparing a blog post about a Creative-Commons resource using appropriate attribution name and URL. Even more interesting is the Operator Firefox Add-On [3], which notifies users of the presence of RDFa as they browse, with content-specific action buttons automatically enabled based on the data found on the page. For example, Operator might allow one-click adding of an event to a calendar, of a phone number to an address book, or of a song to a playlist.

More information about RDFa markup can be found in the RDFa Primer [4], and the RDFa community and ecosystem is reachable on the RDFa Information Web site [5].

**Links:**
[1] http://rdfa.info/rdfa-implementations/
[2] http://www.w3.org/2006/07/SWD/RDFa/impl/js/
[3] https://addons.mozilla.org/firefox/addon/4106
[4] http://www.w3.org/TR/xhtml-rdfa-primer/
[5] http://rdfa.info/

**Please contact:**
Ben Adida
Creative Commons
E-mail: ben@adida.net

# A Framework for Video Interaction with Web Browsers

by Pablo Cesar, Dick Bulterman and Jack Jansen

*In order to make multimedia a first-class citizen on the Web, there is a need for major efforts across the community. European projects such as Passepartout (ITEA) and SPICE (IST IP) show that there is a need for a standardized mechanism to provide rich interaction for continuous media content. CWI is helping to build a framework that adds a temporal dimension to existing a-temporal Web browsers.*

Web 2.0 is not so much a technological revolution as an evolution in the attitude of end-users towards the Web. What started as a global library is becoming a social meeting place in which users can share views and content. Where the initial focus was on a repository of static documents, the future focus will be on the provision of dynamic document services, such as the one shown in Figure 1.

Some of the future scenarios that motivate our work are:
• *E-tourism:* an online guide to a city that includes videos of the place of interest. The presentation can dynamically export the coordinates of the locations presented in the videos, which can be used to represent the guided tour in an external engine such as Google maps, as shown in Figure 1. In addition, a GPS-enabled phone can

request the current coordinates and import them into a presentation that shows the location of the user.
• *E-learning:* an e-learning portal that includes synchronized videos and slides. An interactive test controlled by an external calculation engine can provide results to the media player, allowing the learning material to be adapted to the knowledge of the student.
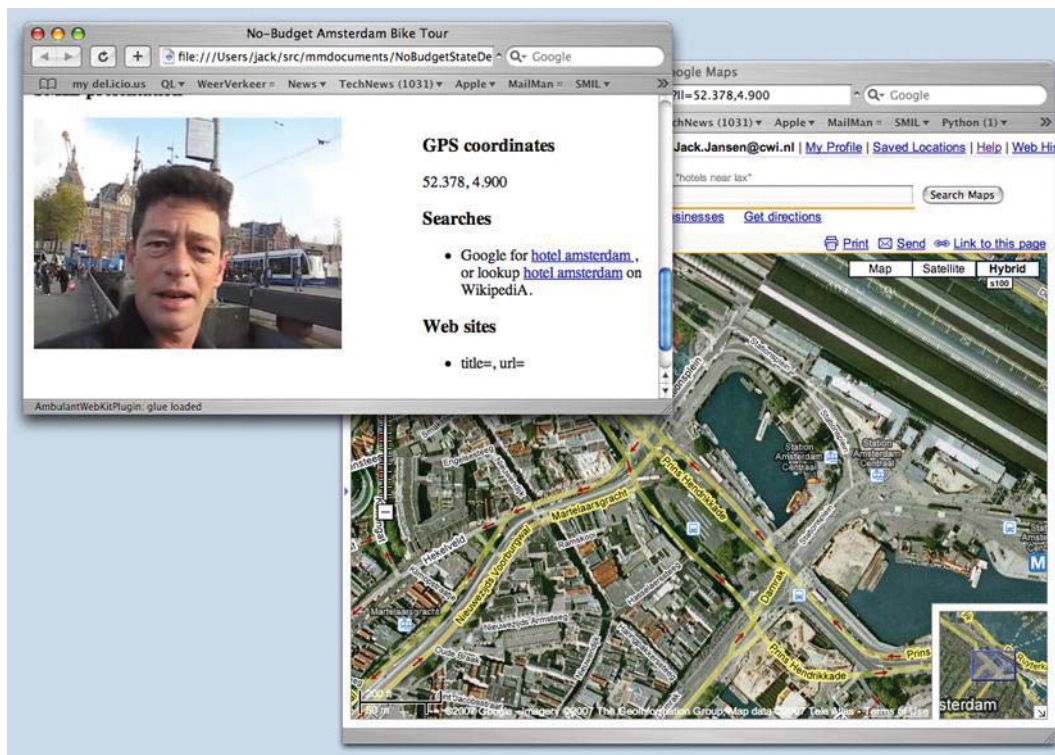
- *E-commercials:* media-based commercials can be customized to a specific user by, for example, displaying the name of the user and adapting the media presentation based on user preferences.
- Persistent segmentation: for example, by allowing the user to explicitly pause a presentation and then restart it at some later point – possibly days or weeks later.

In order to realize the services-oriented vision with video, an interaction model needs to be defined that transcends the traditional control set of start, stop and pause. The content within the video element will need to be triggered from external, peer-level content, as in the e-commercial scenario. That content in turn needs to trigger related content within the context of a higher-level embedding, as in the e-tourism scenario.

The scenarios show that there is a clear need for richer temporal semantics when integrating a conventional (X)HTML browser interface with multimedia documents. To this end, we wrap videos with an external data model, to extend content-related (not content-based) interaction. The data model – rather than the video encoding – is the focal point for sharing, mashing and reusing individual objects.

Following the lead of XForms, our data model is defined as a small XML docu-

ment. This data model is language-independent and can be shared between different XML-based documents such as (X)HTML, SMIL, or SVG. In addition, the framework provides support for defining and manipulating the value of variables in the data model. Moreover, the framework provides the mechanism by which variables can be evaluated at runtime and the state variable values saved for the next time the media document is played.

By exporting the data model to the outside world, it becomes possible for the media document to affect other contexts, eg the (X)HTML presentation. At the same time, external engines can affect the media presentation. So, unlike embedded video players, in our scenarios the video plays an active role in the Web page.

At the moment, the framework is implemented in the Ambulant open-source SMIL player. The work sketched in this article has been submitted to the W3C's SYMM working group under the name of smilState. It is expected to be integrated in the SMIL 3.0 release in early 2008. We are also actively participating in the W3C Backplane work to use the results from this and other Web groups to integrate a broadly consistent framework for sharing the data model across XML-based languages.

**Links:**
http://www.cwi.nl/projects/Ambulant/distPlayer.html
http://www.w3.org/TR/NOTE-HTMLplusTIME
http://www.w3.org/TR/SMIL/

**Please contact:**
Pablo Cesar
CWI, The Netherlands
Tel: +31 20 592 4332
E-mail: Pablo.Cesar@cwi.nl

# Creating, Organising and Publishing Media on the Web

by Raphaël Troncy

*Creating, organising and publishing findable media on the Web raises outstanding problems. While most applications that process multimedia assets make use of some form of metadata to describe the multimedia content, they are often based on diverse metadata standards. In the W3C Multimedia Semantics Incubator Group (XG), we have demonstrated, for various use cases, the added value of combining several of these into a single application. We have also shown how semantic Web technology can help in making these standards interoperable.*

The W3C Multimedia Semantics XG has been chartered to show how metadata interoperability can be achieved by using semantic Web technology to integrate existing multimedia metadata standards. Between May 2006 and August 2007, 34 participants from fifteen different W3C organisations and three invited experts have produced several documents: these describe relevant multimedia vocabularies for developers of semantic Web applications, and the use of RDF and OWL to create, store, exchange and process information about image, video and audio content. The Semantic Media Interfaces group at CWI has sponsored and co-chaired this W3C activity.

The main result is a Multimedia Annotation Interoperability Framework that presents several key use cases and applications showing the need for combining multiple multimedia metadata formats, often XML-based. We use semantic Web languages to explicitly represent the semantics of these standards and show that they can be better integrated on the Web. We describe two of the use cases here.

### Managing Personal Digital Photos

The advent of digital cameras and camera phones has made it easier than ever before to capture photos, and many people have collections of thousands of digital photos from friends and family, or taken during vacations, while travelling and or parties. Typically, these photos are stored on personal computer hard drives in a simple directory structure without any metadata. The number of tools, either for desktop machines or Web-based, that perform automatic and/or manual annotation of content is increasing. For example, a large number of personal photo management tools extract information from the so-called EXIF header and add this information to the photo description. Popular Web 2.0 applications, such as Flickr, allow users to upload, tag and share their photos within a Web community, while Riya provides specific services such as face detection and face recognition of personal photo collections.

What remains difficult however, is finding, sharing, and reusing photo collections across the borders of tools and Web applications that come with different capabilities. The metadata can be represented in many different standards ranging from EXIF (for the camera description), MPEG-7 (for the low-level characteristics of the photo) to XMP and DIG35 (for the semantics of what is depicted). To address this problem, we have designed OWL (Web Ontology Language) ontologies to formalize the semantics of these standards, and we provide converters from these formats to RDF in order to integrate the metadata. A typical use case is a Web application that can generate digital photo albums on the fly based on both contextual and semantic descriptions of the photos, and aesthetic and visual information tailored to the user.

### Faceting Music

Many users own digital music files that they have listened to only once, if ever. Artist, title, album and genre metadata represented by ID3 tags stored within the audio files are generally used by applications to help music consumers manage and find the music they like. Recommendations from social networks, such as lastFM, combined with individual FOAF (Friend of a Friend) profiles comprise additional metadata that can help users to purchase missing items from their collections. Finally, acoustic metadata, a broad term which encompasses low-level features such as tempo, rhythm and tonality (key), are MPEG-7 descriptors that can be automatically extracted from audio files, and which can then be used to generate personalized playlists.

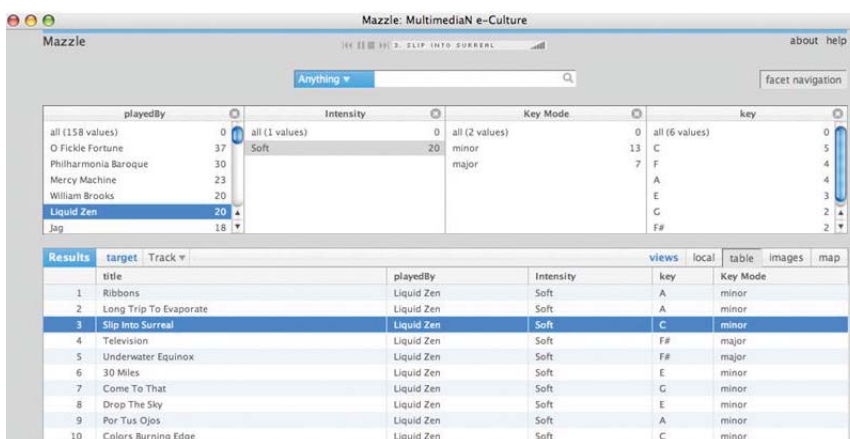Using the Music Ontology, a converter generates RDF descriptions of music tracks from the ID3 tags. The descrip-



*Figure 1: Mazzle provides a faceted browser interface on your personal music collection and can automatically build personalized playlists.*

tions are enhanced with low-level descriptors extracted from the signal and can be linked to recommendations from social community Web sites. The user can then navigate a personal music collection using an iTunes-like faceted interface that aggregates all this metadata (see Figure 1).

### COMM: From MPEG-7 to the Semantic Web

MPEG-7 is the de facto standard for multimedia annotation that appears in every use case developed by the Multimedia Semantics XG. Media content descriptions often explain how a media object is composed of its parts and what the various parts represent. Media context descriptions need to link to some background knowledge and to the creation and use of the media object. To

make MPEG-7 interoperable with semantic Web technology, we have designed COMM, a core ontology for multimedia that has been built by reengineering the ISO standard, and using DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) as its underlying foundational ontology to support conceptual clarity and soundness as well as extensibility towards new annotation requirements.

Current semantic Web technology is sufficiently generic to support annotation of a wide variety of Web resources, including image, audio and video resources. The Multimedia Semantics XG has developed tools and provided use cases showing how the interoperability of heterogeneous metadata can be achieved. Still, many things need to

be improved! The results from the Multimedia Semantics XG and the report from the W3C 'Video on the Web' workshop will provide valuable input to identifying the next steps towards making multimedia content a truly integrated part of the Web.

**Links:**
http://www.w3.org/2005/Incubator/mmsem/
http://multimedia.semanticweb.org/COMM/
http://www.w3.org/2007/08/video/

**Please contact:**
Raphaël Troncy
CWI, The Netherlands
Tel: +31 20 592 4093
E-mail: Raphael.Troncy@cwi.nl

# Empowering Online Communities to Manage Change - How to Build Viable Organisations Online

by David Lewis and Kevin Feeney

*The World Wide Web is witnessing an explosion in new forms of online community. These are built on advances in Web content posting, social networks and the wisdom of crowds. However, we are still largely ignorant about the factors that enable online communities to react successfully to change. Change can involve a number of things: swings in membership and levels of engagement; the emergence of internal conflicts; changing relationships with other online communities and offline organisations; and increasingly, changes caused by new computer-mediated communication technology.*

This is significant because there are real structural problems with how online communities manage change. Though communication channels in online communities are inherently open, the channels through which management decisions are made and enforced are not. The management technology available for online communities reflects the centralized, specialist mindset of professional administrators of information and communication technology.

Typically in online communities, the authority to manage communication channels also imparts authority over the social behaviour of the community. Management decision-making in online communities is therefore highly centralized, which restricts the participation of the broader community in reacting to change. The viability of a community is therefore precariously dependent on the

competence and good grace of the individuals concerned.

Researchers at the Knowledge and Data Engineering Group at Trinity College Dublin have recently completed the development and evaluation of a ground-breaking mechanism for collective management decision-making in online communities. Their approach uses progressive grounding of a rule-based model. This advances a collective approach to managing change as a means to building viable online communities. The system developed is a novel specialisation of policy-based management, termed Community-Based Policy Management (CBPM).

With its novel use of self-defining groups as the fundamental structural abstraction, CBPM differs from previous policy-based management approaches

that use the centrally defined roles typical of existing access control systems. A group is simply established by a set of people engaged in a shared activity. By defining subgroups and federated groups through explicit mandates for exercising decision-making authority, an organisation of self-managing groups can be formed around the evolving needs and experiences of an online community. These mandates can be progressively grounded as patterns of authority change, or as new models of resources or context emerge, and their impact on the distribution of authority is captured. By controlling the interconnection of different social and IT management rules between groups, the portion of an organisation's current rule set that collaborating decision-makers must understand is restricted, thereby making collective decision-making more scaleable. This also delivers fast runtime policy rule
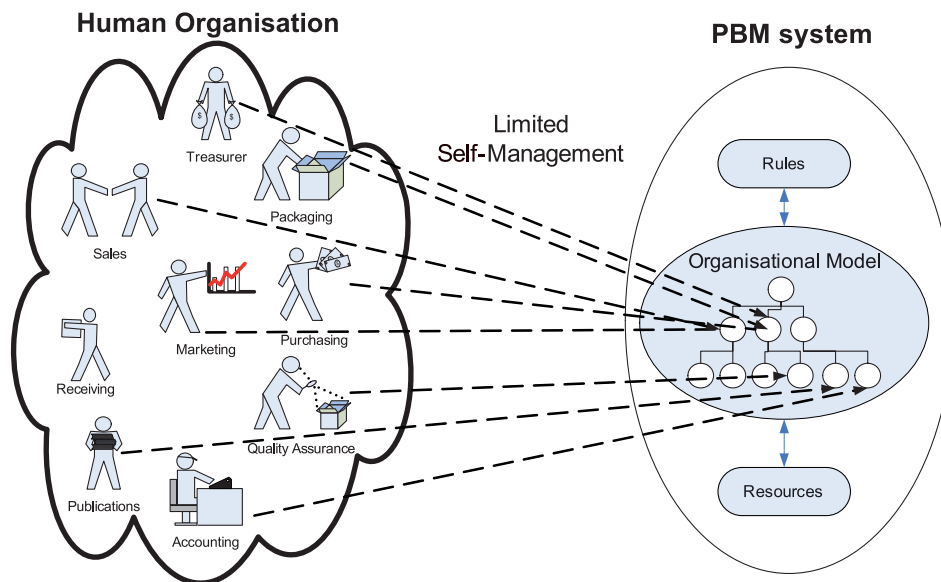
**Human Organisation**

Treasurer

Packaging

Sales

Marketing

Purchasing

Receiving

Quality Assurance

Publications

Accounting

Limited **Self-**Management

**PBM system**

Rules

Organisational Model

Resources

*Figure 1:*
*Policy-Based Management*
*(PBM) using a group-oriented*
*organisational model.*

checking. Clear provenance of rules ensures that when policies authored in one part of a community clash with policies or goals from another, the causes of the conflict are immediately identified, thereby quickening their resolution. Group structure and policy rules are explicitly modelled, and as a result, policy conflicts and the parties necessarily involved in their resolution can be identified. This enables online communities to actively reflect on their management processes.

The Indymedia network (an online network for global news reporting that is independent of large news corporations) was selected as a case study for CBPM. The Oscailt Content Management System, which is the Indymedia community server, was modified to enforce all access control decisions according to policy rules. The organisational form of the community was tracked over a year, during which time CBPM proved able to accurately model structural changes. This illustrated that CBPM possesses the power to manage changes in the distribution of authority and to resolve problems in organisational structure through detection and analysis of policy conflicts. The explicit, declarative nature of the policy rules, when aligned using CBPM to the structure of community, was therefore shown to support reflection upon that structure by its members.

Future work will develop further visual, interactive Web-based tools for reflect-

ing on organisational structure, resource semantics and the existing policy-rule set. This will empower online communities to move from mere discussion fora to viable, value-generating organisations. Such empowerment has the potential to radically enhance the role of online voluntary associations: in non-profit activities, in public policy deliberation, and in broadening the inclusive management of public bodies. Similarly, the approach can support devolved decision-making and an agile team working in corporations or commercial virtual organisations.

Such use of CBPM will bring into sharp focus the dynamics and distribution of power in online organisations, and will highlight how strategies can evolve for balancing control and trust in achieving shared and competing goals. Ultimately, the explicit modelling of successful strategies will facilitate the definition and reuse of online community management patterns.

Further CBPM applications are being developed using a Web service implementation. It is ideal for applications where resource ownership, and thus decision-making authority, are necessarily diffuse. Integrated applications include Ubiquitous Computing under the M-Zones project (www.m-zones.org), and now the NEMBES (www.nembes.org) project and Dynamic Spectrum Management under the Irish Centre for Telecommunication Value Chain Research (www.ctvr.ie).

# Owela: Open Web Laboratory for Innovation and Design

by Pirjo Näkki

*For most users, the Web is about communication rather than technology. The rise of so-called social media shows that when people are provided with simple tools and easy access to online content, they find new ways to utilize the Internet. The VTT Open Web Lab (Owela) is studying how social media tools can also be utilized in innovation and product design processes.*

As increasing numbers of people spend their time in online communities, companies are also trying to find their way into these networks. One of the new ways of utilizing the Internet is the open innovation platform, where companies develop new technology, products and services openly with customers and end users. For companies, the Web serves as a direct channel for communication, feedback and sharing ideas with users.

Owela is an online laboratory that utilizes social media features for participatory design and open innovation. It has been developed in the VTT research project 'Social media in the crossroads of physical, digital and virtual worlds' (SOMED). This project started in 2006 and will continue until August 2008. It is developing technology and applications that support the creation of user-friendly and value-adding applications with socially created input for everyday use.

Owela serves as a virtual working environment, in which VTT is studying social media as a phenomenon of digital culture, and is testing and developing new social software for computer-mediated communication. In Owela we can improve our understanding of future user needs as well as gathering feedback from a wider audience. Furthermore, Owela is used to study how new social media tools can be utilized to enrich interactive Web-based research methods. Owela was launched in April 2007 and is under continuous development based on experience and user feedback.

Traditional user-centred design begins with user-need acquisition, use-context analysis, and an examination of users' social and physical environments. In different phases of the product design, users participate in interviews, observations and testing. The user research process requires a lot of time and resources from the researchers.
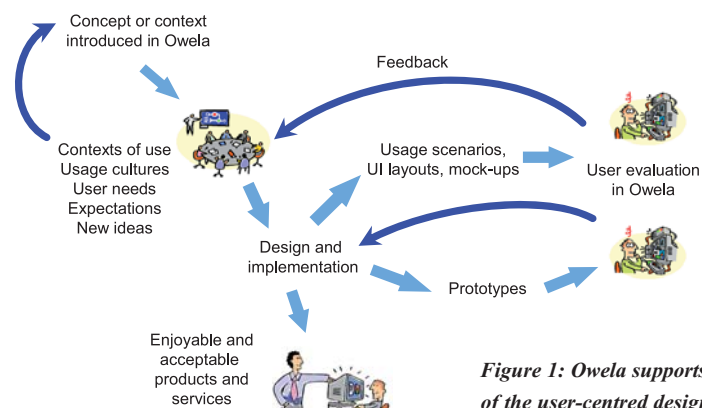


*Figure 1: Owela supports all phases of the user-centred design process.*

In the new era of the Internet, when applications and services are produced with more agility, the traditional thorough user research process seems too time-consuming. New, easier and faster ways to perform user-centred design are therefore needed. As both a means of communication and a meeting place, the Web serves as a good medium for participatory design, since that is where the users are already. Our goal is to define and implement the processes and tools for user-driven product development so that the whole user-centric research process can be carried out reliably and efficiently.

At the moment, Owela is centred around a blog-based tool called Idea-Tube, with which users may browse, comment on, and rate ideas, concepts and scenarios of new products and services. Other tools for collaboration between users, developers and researchers are chat and test lab, in which users may test new prototypes and give feedback. In addition to these qualitative research methods, quantitative online questionnaires can be used.

Owela is not only a collection of different tools but an active community of users interested in new product and service development. In this kind of design community, optimum results are achieved if users' motivation comes from the community and participation, rather than from external rewards. Direct contact with users is needed, along with communication via Web site and emails. Furthermore, the design process must be clear to the users: they must see how their participation affects the design.

Owela can be combined with other user-centred design methods and utilized as a communication channel between face-to-face studies. User research can be done either publicly or, for confidential user studies, in restricted environments. It is also possible to link Owela tools to other existing Web communities. In the future, mobile features will also be developed.

As a follow-up project, we have been planning the project 'Information Technologies supporting the Execution of Innovation Projects' (ITEI) in cooperation with seven European universities and 22 companies. This will involve further development of methods for user-driven innovation and the online tools that support it.

**Links:**
Owela: http://owela.vtt.fi
SOMED: http://www.vtt.fi/proj/somed/

**Please contact:**
Pirjo Näkki, VTT Technical Research Centre of Finland
Tel: +358 20 7225897
E-mail: pirjo.nakki@vtt.fi

# The Principle of Identity Cultivation on the Web

by Pär J. Ågerfalk and Jonas Sjöström

*Web 2.0 and the commercial interest in open-source software both reflect a current trend towards increased user involvement in product and service development. To stay competitive in this era of open innovation, companies must learn to trust users as codevelopers and to make use of the Web as an instrument for identity cultivation.*

A fundamental question in relation to the success of these emerging gift cultures is what motivates people to contribute time and knowledge without any apparent return, at least not in the immediate monetary sense. According to economists Josh Lerner and Jean Tirole, the two major motivations are career concerns and ego gratification, which they collectively refer to as the signalling incentive. By contributing to a Web community people gain reputation and status within that community. Creating and maintaining one's identity and

others in order for them to recognize one's contributions. In a similar vein, individuals' personal identity yis important to corporations in order for them to recognize the characteristics and needs of their customers, and to tailor their own Web presence, thus building their corporate identity.

However, while identity and recognition are important on the Web, the flipside of the identity coin is that of personal integrity. Consider the following story for example, which was reported

trail led to Thelma Arnold, a 62-year-old widow in Lilburn, GA, who confirmed the searches were indeed hers. Shortly after this report, AOL removed the search data from its site and apologized for its release, but the detailed records continue to circulate online. The story does not reveal whether or not Ms. Arnold benefited from her strengthened identity in this particular community. However, the example clearly illustrates that some of the traces we leave on the Web are less intentional and probably less ego gratifying than we would wish.

The aim of this research is to increase our understanding of how the Web is used as an identity creation system, both for individuals and corporations. We have analysed the use of Amazon.com from the perspective of identity creation based on how and with whom we communicate when using such a system. This includes the importance of considering both the intentional actions of users and the traces left behind by those actions. The traces we leave behind, both knowingly (such as when reviewing a book) and unknowingly (by merely using the underlying technology of the Web infrastructure) are the foundation for our Web identity. In order to take advantage of the signalling incentive, a Web site therefore needs to provide users with instruments that allow them to develop a proper understanding of their 'online conversations' and their contribution to the development of their own identity. We refer to this as the 'principle of identity cultivation'.



*Figure 1: Main concepts and relationships relevant to identity cultivation on the Web.*

social status thus appears to be the main driving force of these emerging phenomena. A successful company must show commitment and build a strong corporate identity to attract people (ie visitors), and an individual must commit fully to a community in order to build a strong personal identity: both are motivated by the signalling incentive. Hence, while personal identity is important to oneself, it is also important to

in The New York Times. Last year, a team within AOL publicly released search data of more than 650,000 users. Although actual user names were replaced with random numbers, it was possible to track all the search terms used by a given individual, and thus to physically locate that individual. Apparently, No. 4417749 conducted hundreds of searches over a three-month period and eventually the data

Striving for recognition in some community thus motivates users seemingly altruistically to contribute to the development of the value of Internet services. In the quest for strong corporate identity, Web 2.0-aware companies, such as Amazon.com, are building their services around user engagement. This attracts people to their Web site and engages them in the codevelopment of
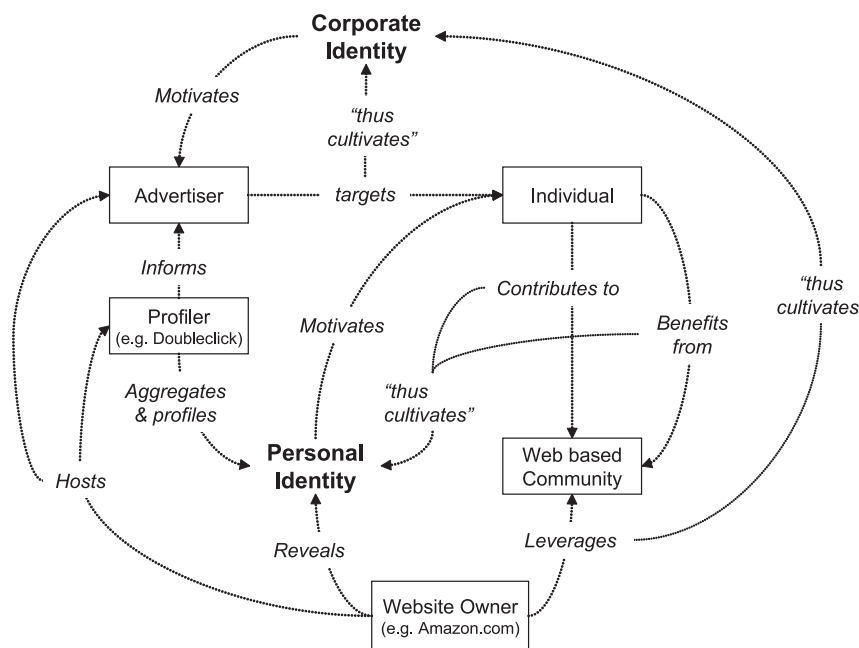
the site, which in turn gives Amazon.com a strong corporate identity. This strong identity and the large number of committed visitors also makes Amazon.com an attractive platform on which third parties can build their identity through personalized advertisements. This has an impact on our view of the Web from a design perspective: we need to acknowledge multiple interests and the implications for gathering, storing and utilizing information about both intentional user actions and the possibly unintended traces left behind.

For more details on this study, see the paper "Sowing the Seeds of Self: A Socio-Pragmatic Penetration of the Web Artefact", which recently won the Best Paper Award at the 2007 International Conference on the Pragmatic Web (http://www.pragmaticweb.info). The paper is available from the authors on request and from the ACM digital library (http://www.acm.org/dl/).

**Please contact:**
Pär J. Ågerfalk
Lero – The Irish Software Engineering Research Centre, University
of Limerick and Uppsala University
Tel: +353 61 213543
E-mail: par.agerfalk@lero.ie

Jonas Sjöström
Jönköping International Business School and Linköping University
Tel: +46 36 101784
E-mail: jonas.sjostrom@jibs.hj.se

# Understanding the Hidden Web

by Pierre Senellart, Serge Abiteboul and Rémi Gilleron

*A large part of the Web is hidden to present-day search engines, because it lies behind forms. Here we present current research (centred around the PhD thesis of the first author) on the fully automatic understanding and use of the services of the so-called hidden Web.*

Access to Web information relies primarily on keyword-based search engines. These search engines deal with the 'surface Web', the set of Web pages directly accessible through hyperlinks, and mostly ignore the vast amount of highly structured information hidden behind forms that composes the 'hidden Web' (also known as the 'deep Web' or 'invisible Web'). This includes, for instance, Yellow Pages directories, research publication databases and weather information services. The purpose of the work presented here, a collaboration between researchers from INRIA project-teams Gemo and Mostrare and the University of Oxford (Georg Gottlob), is the automated exploitation of hidden-Web resources, and more precisely the discovery, analysis, understanding and querying of such resources.

An original aspect of this approach is the avoidance of any kind of human supervision, which makes the problem quite broad and difficult. To cope with this difficulty, we make the assumption that we are working with services relevant to a specific domain of interest (eg research publications) that is described by some domain knowledge base in a predefined format (an ontology). The approach is content-centric in the sense that the core of the system consists in a content warehouse of hidden-Web services, with independent modules enriching the knowledge of these services so they can be better exploited. For instance, one module may be responsible for discovering relevant new services (eg URLs of forms or Web services), another for analysing the structure of forms and so forth.

Typically the data to be managed is rather irregular and often tree-structured, which suggests the use of a semi-structured data model. We use eXtended Markup Language (XML) since this is a standard for the Web. Furthermore, the different agents that cooperate to build the content warehouse are inherently imprecise since they typically involve either machine-learning techniques or heuristics, both of which are prone to imprecision. We have thus developed a probabilistic tree (or XML) model that is appropriate to this context. Conceptually, probabilistic trees are regular trees annotated with conjunctions of independent random variables (and their negation). They enjoy nice theoretical properties that allow queries and updates to be efficiently evaluated.

Consider now a service of the hidden Web, say an HTML form, that is relevant to the particular application domain. In order that it can be automatically reused, an understanding of various aspects of the service is necessary; that is, the structure of its input and its
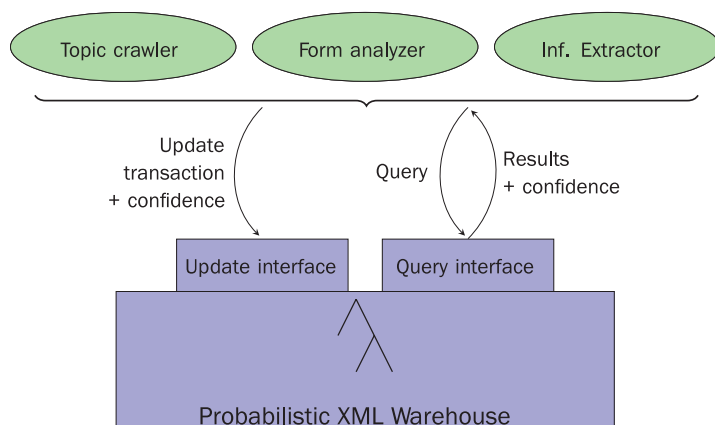


*Figure 1: Probabilistic content warehouse, updated and queried by various modules that analyse the hidden Web.*

output, and the semantics of the function that it supports. The system first tries to understand the structure of the form and relate its fields to concepts from the domain of interest. It then attempts to understand where and how the resulting records are represented in an HTML result page. For the former problem (the service input), we use a combination of heuristics (to associate domain concepts to form fields) and probing of the fields with domain instances (to confirm or invalidate these guesses). For the latter (the service output), we use a supervised machine-learning technique adapted to tree-like information (namely, conditional random fields for XML) to correct and generalize an automatic, imperfect and imprecise annotation using the domain knowledge. As a consequence of these two steps, the structure of the inputs and outputs of the form are understood (with some degree of imprecision of course), and thus a signature for the

service is obtained. It is then easy to wrap the form as a standard Web service described in Web Service Definition Language (WSDL).

Finally, it is necessary to understand the semantic relations that exist between the inputs and outputs of a service. We have elaborated a theoretical framework for discovering relationships between two database instances over distinct and unknown schemata. The problem of understanding the relationship between two instances is formalized as that of obtaining a schema mapping (a set of sentences in some logical language) so that a minimum repair of this mapping provides a perfect description of the target instance given the source instance. We are currently working on the practical application of this theoretical framework to the discovery of such mappings between data found in different sources of the hidden Web.

At the end of the analysis phase we have obtained a number of Web services, the semantics for which is explained in terms of a global schema specific to the application. These services are described in a logical Datalog-like notation that takes into account the types of input and output, and the semantic relations between them. The services can now be seen as views over the domain data, and the system can use these services to answer user queries with almost standard database techniques.

**Please contact:**
Pierre Senellart
INRIA Futurs, France
Tel: +33 1 74 85 42 25
E-mail: pierre@senellart.com

# Static Analysis of XML Programs

by Pierre Genevès and Nabil Layaïda

*Static analysers for programs that manipulate Extensible Markup Language (XML) data have been successfully designed and implemented based on a new tree logic by the WAM (Web, Adaptation and Multimedia) research team, a joint lab of INRIA and Laboratoire d'Informatique de Grenoble. This is capable of handling XML Path Language (XPath) and XML types such as Document Type Definitions (DTDs) and XML Schemas.*

Since its introduction a decade ago, Extensible Markup Language (XML) has gained considerable interest from industry and now plays a central role in modern information system infrastructures. In particular, XML is the key technology for describing and exchanging a wide variety of data on the Web. The essence of XML consists in organising information in tree-tagged structures conforming to some constraints, which are expressed using standard type languages such as DTDs, XML schemas and Relax NG. XML processing can be seen as transforming these structures using tree-oriented query languages such as XPath expressions and XQuery within full-blown transformation languages such as XSLT.

With the ever-increasing information flow in the current Web infrastructure, XML programming is becoming a key factor in realizing the trend in Web serv-

ices aimed at enhancing machine-to-machine communication. There still exist important obstacles along this path: performance and reliability. Programmers are given two options, Domain-Specific Languages such as XSLT, or general-purpose languages augmented with XML application programming interfaces such as the Document Object Model (DOM). Neither of these alternatives is a satisfactory answer to performance and reliability, nor is there even a trade-off between the two. As a consequence, new paradigms are being proposed and all have the aim of incorporating XML data as first-class constructs in programming languages. The hope is to build a new generation of tools that are capable of taking reliability and performance into account at compile time.

One of the biggest challenges in this line of research is to develop automated and tractable techniques for ensuring static-

type safety and optimization of programs. To this end, there is a need to solve some basic reasoning tasks that involve very complex constructions such as XML types (regular tree types) and powerful navigational primitives (XPath queries). In particular, every future compiler of XML programs will have to routinely solve problems such as:
- XPath query emptiness in the presence of a schema: if one can decide at compile time that a query is not satisfiable then subsequent bound computations can be avoided
- query equivalence, which is important for query reformulation and optimization
- path type-checking, for ensuring at compile time that invalid documents can never arise as the output of XML processing code.

All of these problems are known to be computationally heavy (when decid-

able), and the related algorithms are often tricky.

In the WAM research team we have developed an XML/XPath static analyser based on a new logic of finite trees. This analyser consists in compilers that allow XML types and XPath queries to be translated into this logic, and an optimized logical solver for testing satisfiability of a formula of this logic.

The benefit of these compilers is that they allow one to reduce all the problems listed above, and many others, to logical satisfiability. This approach has a couple of important practical advantages. First of all, one can use the satisfiability algorithm to solve all of these problems. More importantly, one could easily explore new variants of these problems, generated for example by the presence of different kinds of type or

schema information, with no need to devise a new algorithm for each variant.

The system has been implemented in Java and uses symbolic techniques (binary decision diagrams) in order to enhance its performance. It is capable of comparing path expressions in the presence of real-world DTDs (such as the W3C SMIL and XHTML language recommendations). The cost ranges from several milliseconds for comparison of XPath queries without tree types, to several seconds for queries under very large, heavily recursive, type constraints, such as the XHTML DTD. These preliminary measurements shed light for the first time on the cost of solving static analysis problems in practice. Furthermore, the analyser generates XML counter-examples that allow program defects to be reproduced independently from the analyser.

The development of these analysers was initiated by the Web Adaptation and Multimedia research team at INRIA, Grenoble - Rhône-Alpes Research Centre, France. The project commenced in October 2005 and the full system will soon be released publicly.

# Seaside – Advanced Composition and Control Flow for Dynamic Web Applications

by Alexandre Bergel, Stéphane Ducasse and Lukas Renggli

*Page-centric Web application frameworks fail to offer adequate solutions to model composition and control flow. Seaside allows Web applications to be developed in the same way as desktop applications. Control flow is modelled as a continuous piece of code, and components may be composed, configured and nested as one would expect from traditional user interface frameworks.*

Seaside is a framework for building dynamic Web applications, uniquely combining object-oriented and continuation-based approaches. Seaside applications are built by composing stateful components, each encapsulating a small portion of the page. Programmers are freed from the concern of providing unique names, since Seaside automates this by associating callback functions with links and form fields. Control flow is expressed as a continuous piece of code and in addition, Seaside offers a rich application programming interface (API) to integrate with the latest Web 2.0 technology, such as AJAX (Asynchronous JavaScript) and Comet (Server Push).

Seaside is implemented in Smalltalk, a dynamically typed programming language. Seaside inherits powerful reflective capabilities from the underlying language. Web applications may be debugged while the application is running. Inspection and modification may

occur on objects on the fly. Source code is changeable and recompilable without interrupting the running application, and there is no need to restart a session.

Whereas most other Web application frameworks work in a page-centric fashion, Seaside makes use of stateful components that encapsulate a small portion of a page. Developers can compose the user interface as a tree of individual components, and often these components are reused over and over again, within and between applications. A basic set of ready-made widgets is also provided to handle user interactions.

Seaside offers a mechanism by which objects may be registered to be backtracked. With every response sent to the client, Seaside takes a snapshot of and caches registered objects. This allows previous application states to be restored in a controlled fashion, for

example when the user is using the 'back' button in the Web browser.

Seaside uses programmatic XHTML generation. Instead of repeatedly pasting the same sequence of tags into templates, it provides a rich API to generate XHTML. This approach not only avoids common problems with invalid markup, but also allows markup patterns to be easily abstracted into convenient reusable methods. CSS (Cascading Style Sheets) are used to give the application a professional look.

Seaside also provides callback-based request handling. This allows developers to associate a piece of code with anchors and form fields, which are then automatically performed when the link is clicked or the form is submitted. This feature makes it almost trivial to connect the view with its model, as Seaside abstracts all serialization and parsing of query parameters away.

Callbacks are used in a natural way to define a control flow, for example to temporarily delegate control to a sequence of other components. The flow is defined by writing plain source code. Control statements, loops and method calls are mixed with messages to display components. Whenever a new view is generated, the control flow is suspended and the response is sent back to the client. Upon a new user interaction the flow is resumed.

In order that the definition of control flows as part of a Web application be as seamless as possible, Seaside internally stores a 'continuation' whenever a new component is displayed. This suspends the current control flow and allows one to resume it later on. Since continuations may be resumed multiple times, Seaside supports the use of Web browsers' 'back' and 'forward' buttons at any time. The execution state is automatically restored to the requested point and everything behaves as the developer expects.

To complement the expressiveness of the Smalltalk programming language, a set of tools including a memory analyser, a speed profiler, a code editor and an object inspector are included. The debugger supports incremental code recompilation, and enables highly interactive Web applications to be built quickly and in a reusable and maintainable fashion.

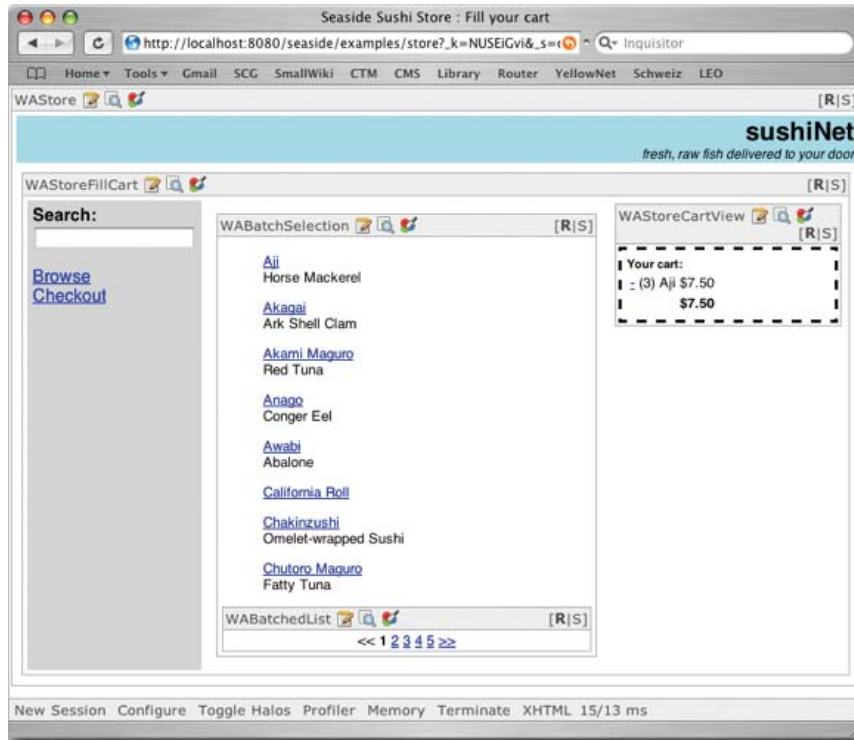Seaside is open-source software distributed under the MIT licence. It is under



*Figure 1: A Web application built with Seaside.*

constant development by an international community using Squeak Smalltalk. Due to the platform independence of Smalltalk, Seaside can be run on almost any platform. The Seaside application server can be bridged with industrial scale servers, such as Apache. Seaside is supported by major commercial Smalltalk vendors, GemStone Smalltalk and Cincom VisualWorks, as part of their product strategy, and is widely used in an industrial context.

The official Seaside Web site uses Pier, an open-source content management system written on top of Seaside.

**Link:**
http://www.seaside.st

**Please contact:**
Alexandre Bergel
INRIA Futurs, Lille, France
E-mail: alexandre.bergel@inria.fr

# EAGER: A Novel Development Toolkit for Universally Accessible Web-Based User Interfaces

by Constantina Doulgeraki, Alexandros Mourouzis and Constantine Stephanidis

*EAGER is an advanced toolkit that helps Web developers to embed in their artefacts accessibility and usability for all. Web applications developed by means of EAGER have the ability to adapt to the interaction modalities, metaphors and user interface elements most appropriate to each individual user and context of use.*

The constantly evolving Web is an unprecedented and continuously growing source of knowledge, information and services, potentially accessible by anyone, at any time and from anywhere. Despite the universality of the Web and the predominant role of Web-based user interfaces in the evolving Information

Society, current approaches to Web design do not embrace the notion of adaptation and the principles of 'design for all'. Consequently, they fail to satisfy the individual interaction needs of target users with different characteristics. A common practice in contemporary Web development is to deliver a single user-

interface design that meets the requirements of an 'average' user. However, this 'average' user is in fact an imaginary entity, and differs radically from the profiles of a large portion of the population. This is particularly the case for people with a disability, elderly people, novice IT users and users on the move.

| Target population | User tasks | Context of use | Means of access |
|---|---|---|---|
| Language Age Background Skills Preferences Disability (blind, motor impaired, deaf, cognitive impaired) … | Work Socialisation Entertainment Education Surfing Commerce Government … | Home Office School Car Internet café … (light, noise, privacy, security, etc.) | Platform (Public terminals, PCs and Laptops, PDA and smart phones) Assistive Technology (screen readers, scanning, Browser (Explorer, Netscape, Firefox) … |

*Figure 1: Dimensions of diversity - some examples.*

The development of the EAGER toolkit was motivated by the emerging need to support individuals in achieving full participation in the knowledge society (especially individuals and groups at risk of exclusion), and to offer the means to address challenges associated with population aging and high disability rates all over the world. In these terms, the main goal of this work is to contribute to the collective vision of mainstreaming and radically improving the accessibility, usability and general user experience and acceptance of computer-based products and services. Thus, it is perceived as a contribution in reducing the 30% of the European population currently not using ICT.

EAGER supports, and provides the means for, the development of inclusive Web-based interfaces that are capable of adapting to multiple and significantly different user profiles and contexts of use (see Figure 1). It is an advanced toolkit that aids Web developers in following the Unified Web Interfaces (UWIs) method that builds on well-established 'design for all' principles. In partic-ular, EAGER allows Microsoft.NET developers to create or revise existing interfaces such that they are able to adapt to the interaction modalities, metaphors and user interface elements most appropriate to each individual user, according to profile information based on user and context-specific parameters..

In this context, a number of user interface elements were designed in various forms (polymorphic task hierarchies) according to specific user and context-related parameter values (see Figure 2). These include user expertise and characteristics (eg disability) and interaction devices and platforms. This phase provided feedback during the development of EAGER, which involved the design of alternative interaction elements, and mechanisms for facilitating the dynamic activation-deactivation (decision making) of the interaction elements and modalities based on individual user interaction and accessibility preferences.

In brief, EAGER is an advanced library of: (a) the core UWI architectural com-ponents; (b) primitive UI elements with enriched attributes, eg buttons, links and radios; (c) structural page elements, eg page templates, headers, footers and containers; and (d) fundamental abstract interaction dialogues in multiple alternative styles, eg navigation, file up-loaders, paging styles and text entry.

As a means to validating the proposed toolkit, a prototypical, fully-functional Web portal, namely the new portal for the European Design for All and e-Accessibility Network (EDeAN), was implemented with the EAGER toolkit and evaluated in terms of its accessibility and usability. The development of the new EDeAN portal is carried out in the framework of the EC-funded Coordination Action DfA@eInclusion (Design for All for eInclusion, contract no. 0033838). Through this process, the produced interfaces were also assessed against the W3C accessibility guidelines and thereby improved. This ensures that interfaces developed by means of EAGER conform to the Web Content Accessibility Guidelines. In summary, this development confirmed that EAGER offers significant benefits to developers and ensures the delivery of accessible, usable and satisfying Web-based interfaces. In fact, the process of employing EAGER supports the development of elegant interfaces and is significantly less demanding in terms of the time, experience and skills required from the developer than the typical process of developing Web interfaces for the 'average' user.

Overall, EAGER constitutes a significant contribution towards embedding accessibility, graceful transformation and ease of use for all in future and existing Web-based applications. Ultimately, this will help individuals – particularly people at risk of exclusion – to fully participate in the knowledge society.

**Link:**
Towards Unified Web-based User Interfaces, Technical Report 394, ICS-FORTH:
http://www.ics.forth.gr/ftp/tech-reports/2007/2007.TR394_Towards_Unified_Web-based_UI.pdf

**Please contact:**
Constantine Stephanidis
ICS-FORTH, Greece
Tel: +30 2810 391 741
E-mail: cs@ics.forth.gr
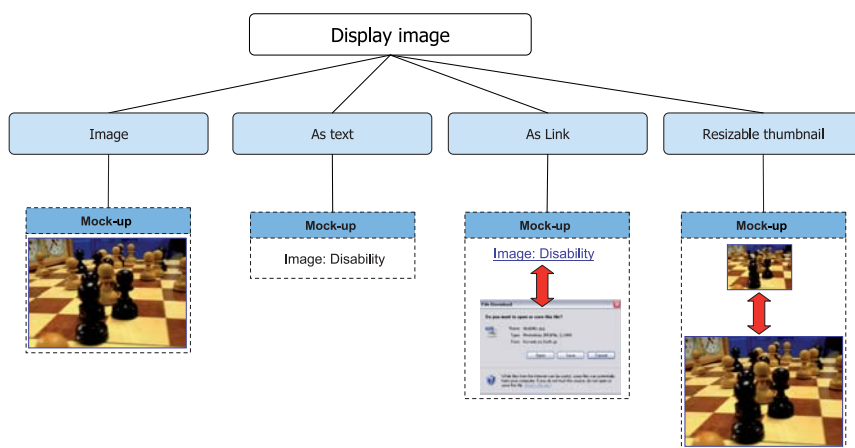


*Figure 2: Examples of alternative representations for images.*

# Migratory Web User Interfaces

by Fabio Paternò, Carmen Santoro and Antonio Scorcia

*An environment developed at the Human Interfaces in Information Systems (HIIS) Laboratory of ISTI-CNR supports Web user-interface migration through different devices. The goal is to provide user interfaces that are able to move across different devices, even offering different interaction modalities, in such a way as to support task continuity for the mobile user. This is obtained through a number of transformations that exploit logical descriptions of the relevant user interfaces.*

One important aspect of pervasive environments is giving users the ability to freely move about while continuing to interact with services available through a variety of interactive devices (cell phones, PDAs, desktop computers, digital television sets, intelligent watches and so on). Many interesting issues arise from such new environments, even for Web applications. Indeed, one significant potential source of frustration is the necessity of continuously restarting a

and that it can be developed with any authoring environment. This is reasonable given that the desktop version of a Web application is usually the only version existing. In addition, the user interface and the content for the desktop version are usually the most extended, thus providing a good basis for the transformations that aim to obtain versions adapted for different types of platform. Our migration platform is composed of a proxy service and a number of spe-

duct their regular access to the Web application and then ask for a migration to any device that has already been discovered by the migration server. Two factors make possible the support of migration across devices that enable various interaction modalities: first, the use of a logical language for user interface descriptions that is independent of the modalities involved (TERESA XML), and second, a number of associated transformations that incorporate



*Figure 1: An Example of Web User Interface Migrating across Multiple Interactive Platforms.*

session after each interaction device change. Migratory Web Interfaces can overcome this limitation and support continuous task performance. This requires that the interactive part of Web applications be able to follow users and adapt to the changing context of use while preserving their state.

We have developed a new solution for supporting migration of Web application interfaces between different types of device. Our migration environment is based on a service-oriented architecture involving multiple clients and servers, and supports Web interfaces with different platforms (fixed and mobile) and modalities (graphical, vocal, and their combination). Our solution aims to be very general. We assume that a desktop version of the relevant applications already exists in the application servers,

cific services and can be hosted by any system. In order to subscribe to our migration environment, a device must run a client. The purpose of this client is to notify the server about the availability of the device to the migration services, and to provide the server with information regarding the device. The migration trigger can be activated only from the source device, either by a user action or a system event. In addition, the source interface is closed after migration; this prevents someone else picking up the source device and corrupting the input.

Our solution is able to detect any user interaction performed at the client level. The state resulting from the different user interactions can then be retrieved and associated with a new user interface version that is activated in the migration target device. Users can therefore con-

design rules and take into account the specific aspects of the target platforms.

When a migration has to be processed, and regardless of which agent (the user or the system) starts the process, the Migration Manager, acting as the main server module, retrieves the original Web page(s) desktop version that is to be migrated. Once it is retrieved, the Migration Manager builds the corresponding logical descriptions at a different abstraction level. The purpose of the logical description is to identify the basic tasks supported by the original version of the application (eg editing values, selecting elements, activating functionalities etc). Once the logical description is obtained, it is used as input for the Semantic Redesigner service, which performs a redesign of the user interface for the target platform.

This involves adapting the logical description to the interaction resources of the target device.

When the target user interfaces have been generated, a specific service is used to identify the page that should be first uploaded to the target device (which is that supporting the last basic task performed on the source device). Indeed, it can happen that one desktop page corresponds to multiple pages (for example when considering a mobile user interface). In order to support task continuity throughout the migration process, the state resulting from the user interactions with the source device (filled data fields, selected items, cookies etc) is gathered through a dynamic access to the Document Object Model (DOM) of the pages in the source device. In a completely transparent fashion, this information is then associated to the corresponding elements of the newly generated pages and adapted to the interaction features of the target device by one module of our architecture. We have developed generators for various Web implementation languages (XHTML, XHTML MP, VoiceXML, X+V) and even for non-Web languages (such as Java for the digital TV).

# Paving the Way of Designers - Towards the Development of Multimodal Web User Interfaces

by Adrian Stanciulescu, Jean Vanderdonckt, Benoit Macq

*Multimodal Web applications provide end-users with a flexible user interface (UI) that allows graphical, vocal and tactile interaction. As experience in developing such multimodal applications grows, the need arises to identify and define major design options of these applications in order to pave the way of designers to a structured development life cycle.*

At the Belgian Laboratory of Computer-Human Interaction (BCHI), Université catholique Louvain, we argue that developing consistent and usable multimodal UIs (involving a combination of graphics, audio and haptics) is an activity that would benefit from the application of a methodology which is typically composed of: (1) a set of models gathered in an ontology, (2) a method which manipulates the involved models, and (3) tools that implement the defined method.

## The Models

Our methodology considers a model-based approach where the specification of the UI is shared between a set of implementation-independent models, each model representing a facet of the interface characteristics. In order to encourage user-centred design, we take the task and domain models into consideration right from the initial design stage. The approach involves three steps towards a final UI: deriving the abstract UI from the task and domain models, deriving the concrete UI from the abstract one, and producing the code of the corresponding final UI. All the information pertaining to the models describing the future UI is specified in the same UI Description Language: UsiXML.

## The Method

Our method considers all stages of the software development life cycle to be covered according to design principles, from early requirements through to prototyping and coding. This approach benefits from a design space that explicitly guides the designer to choose values of design options appropriate to the multimodal UIs, depending on parameters. In order to support these aspects our approach is based on a catalogue of transformation rules implemented as graphs.

## The Tool

We developed MultiXML, an assembly of five software modules, to explicitly support the design space. In order to reconcile computer support and human control, we adopt a semi-automatic approach in which certain repetitive tasks are executed, partially or totally, while still offering some level of control to the designer. This involves:
- manual selection of design option values by the designer, with the possibility of modifying this configuration at any time
- automatic picking of transformation rules from the transformation catalogue corresponding to the selected options, followed by their application by TransformiXML module (see Figure 1) to reduce the design effort.

## Case Studies

The described methodology was tested on a couple of case studies using the Opera browser embedded with IBM Multimodal Runtime Environment. One of these applications considers a large-scale image application (see Figure 2) that enables users to browse a 3x3 grid map (ie translate, zoom in or zoom out) using an instruction pattern composed of three elements: Action, Parameter Y and Parameter X. The X and Y parameters identify the coordinates of the grid unit over which the action is applied (eg Action = <<translate>>, Parameter Y = <<top>>, Parameter X = <<left>>). The application enables users to specify each element of the pattern by employing one of the graphical, vocal and tactile modalities or a combination of them.
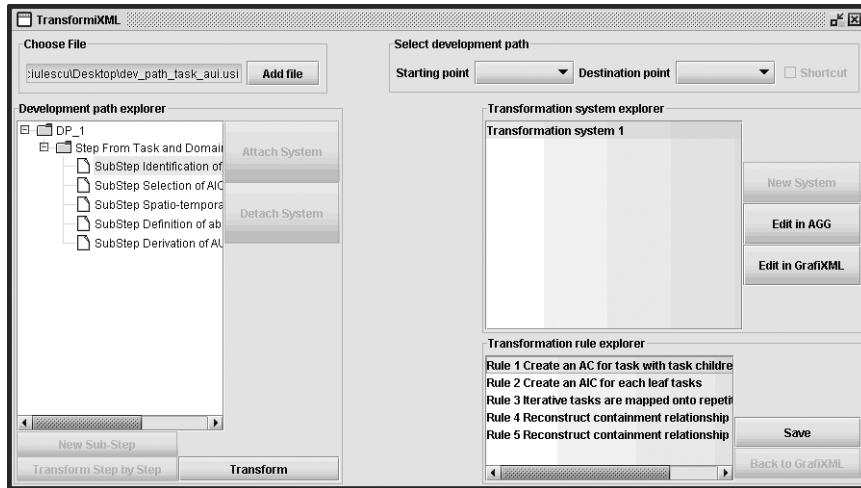
*Figure 1: The TransformiXML module.*

## Motivations

One of the main advantages of our design space is the independence of any existing method or tool, making it useful for any developer of multimodal UIs. The design options clarify the development process by providing options in a structured way. This method strives for consistent results if similar values are assigned to design options in similar circumstances, and also means that less design effort is required. Every piece of development is reflected in a concept or notion that represents some abstraction with respect to the code level, as in a design option. Conversely, each design option is defined clearly enough to drive the implementation without requiring any further interpretation effort. Moreover, the adoption of a design space supports the tractability of more complex design problems or a class of related problems.

## The Project

The current work was conducted in the context of the Special Interest Group on Context-Aware Adaptation of the SIMILAR project. The group began in December 2003 and was coordinated by BCHI. Five other partners were involved: ICI, ISTI-CNR, Universiy Joseph Fourier, University of Castilla-La Mancha and University of Odense. The goal of the group was to define, explore, implement and assess the use and the switching of modalities to produce interactive applications that are aware of the context of use in which they are executed. They can then be adapted according to this context so as to create UIs that remain the most usable. In the future, BCHI will continue working on the design, implementation and evaluation of Ambient Intelligence systems as a member of the ERCIM Working Group Smart Environments and Systems for Ambient Intelligence (SESAMI).

**Links:**
http://www.isys.ucl.ac.be/bchi
http://www.usixml.org
http://www.similar.cc
http://www.ics.forth.gr/sesami/index.html

**Please contact:**
Adrian Stanciulescu,
Jean Vanderdonckt
and Benoit Macq
Université catholique de Louvain,
Belgium
Tel: +32 10 47 83 49, +32 10 47 85 25,
+32 10.47.22.71
E-mail:
stanciulescu@isys.ucl.ac.be
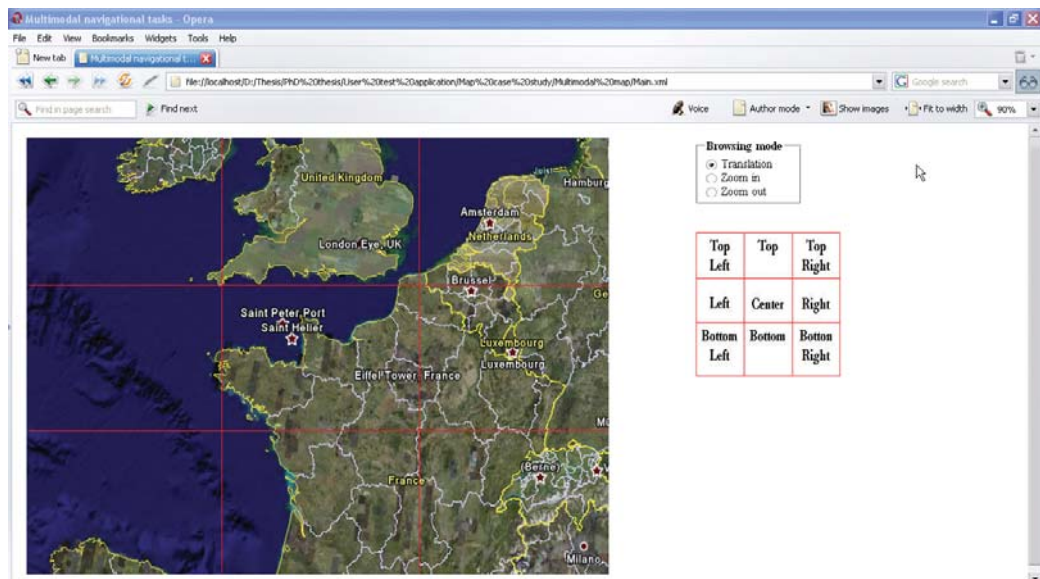vanderdonckt@isys.ucl.ac.be,
benoit.macq@uclouvain.be

*Figure 2:
Multimodal navigation in
large-scale images.*

# Distributed Personalization: Bridging Digital Islands in Museum and Interactive TV

by Lora Aroyo

*Personalization and user experience are key challenges for effectively using current consumer electronics. The VU University Amsterdam demonstrated in two projects the use of Semantic Web technology for personalization: CHIP – combining the experience in a physical museum with mobile devices and the Web, and iFanzy – a personalized selection of the digital TV content in a cross-media environment.*

With the miniaturizing of the computer and its massive deployment in end-user electronics the 'disappearing computer' becomes an inseparable part of the everyday life. The ICT landscape is developing into a highly-interactive distributed environment where people more than ever become early adopters of technology. This rapid progressing of the 'technology embraced by public' creates a risk of information overload if the content is not adapted to the preferences and current user context. The experiences become more personal and engaging across multiple media.

However, the integration of the different user profiles is limited, lacking transparency in the use of personal data between applications. For example, usage data collected by Amazon.com is only valid within the Amazon domain; it

cannot be used for other information services, nor can Amazon cater for different 'modes' of a user. Such personalization is still local and the user experiences are still 'digital islands' – within the boundaries of single applications, not interconnected to allow users capitalizing on the full potential of distributed multi-device environments. Methods and techniques that embrace the complexity of the distributed digital space of devices, information and users are needed to collect the existing fragments of personalization components and provide the essential glue to support personalization across the boundaries of particular applications and devices.

CHIP (Figure 1), an interactive application of semantic Web technology, illustrates how the physical museum, the mobile devices and the Web can be

combined in order to maximize the museum visitors' experience anytime and anywhere. Users' interests and art preferences expressed as semantic metadata of cultural heritage, offer a personalized way to explore digital and physical museum collections through multiple devices (eg PC, mobile phone, PDA) and sensor technology, eg RFID. The CHIP (Cultural Heritage Information Personalization) project is guided by a use case provided by the Rijksmuseum Amsterdam, and is funded by the NWO CATCH program.

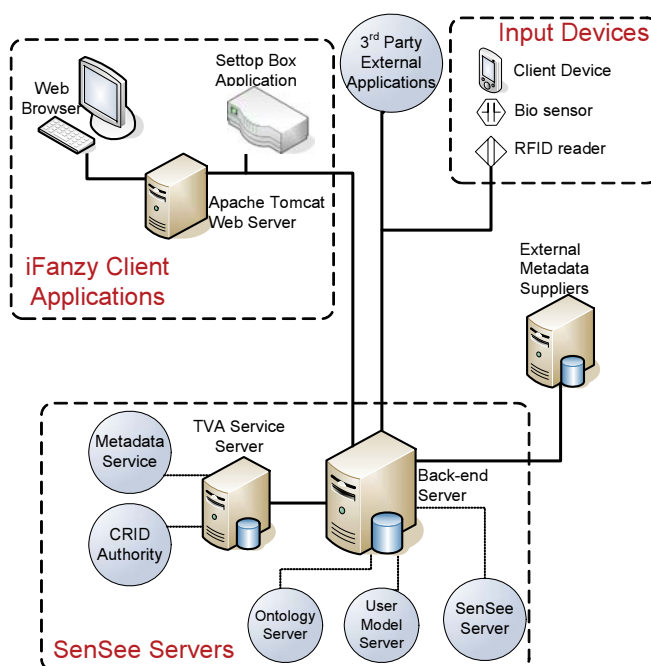In iFanzy personalized cross-media TV guide background knowledge is used for personalized access to TV content in a cross-media environment with TV settop box, PC and mobile. The continuous observation of the user in all spaces (eg online and with various sen-



*(a) iFanzy set-top box interface.*



*(b) iFanzy Web interface.*



*(c) iFanzy space architecture.*

*Figure 1: CHIP application.*

sors at home) allows for a context-sensitive adaptation. This work has been realized as a collaborative effort between Stoneroos Interactive Television, Eindhoven University of Technology and VU University Amsterdam with a Eureka ITEA funding.

To realize distributed personalization in these interactive environments three aspects have been identified:
- interaction spaces, ie virtual (Web) space, physical space and the mobile space
- nformation challenges, ie data integration, modeling and presentation
- space-integration principles, ie user experience optimization strategies.

The physical interaction space, eg, at home or in the museum encompasses task-specific devices such as TV and information displays supporting user's prime goal to consume content. In the virtual interaction space, referring to the Web, users typically work with multiple applications and perform information-intensive tasks. The mobile interaction space regards portable and hand-held devices each containing a fragmented portion of user's data. Further, space-integration principles for utilizing the three spaces for distributed personalization are postulated:
- *Complementarity ('use the right space'):* Use the applications and devices in the three spaces in a com-

plementary way, so that the overall user's experience is improved. For example, the iFanzy Web application is handy for searching, browsing and text-typing tasks. Where a mobile device is used for user identification and one-click-away actions.
- *Minimization ('do not bother the user'):* Minimize the explicit user input, while maximizing the background collection of user and context data and its integration from different applications. This allows in CHIP to start instantly the personalization process, and prevent a 'cold start' in the TV recommendation in iFanzy.
- *Context-awareness ('know where you are serving'):* Consider the specific context for presenting information to the user in order to achieve good confluence of the content and the context. Support user awareness about the current context both while navigating in the museum and interacting with the iFanzy recommender.
- *Vocabulary bridging ('talk the user's language'):* Provide ways of bridging the gap of terminology between expert-created applications and the end-user ways of naming items, searching and browsing.

The information integration, modeling and presentation have been realized with semantic Web technology. Museum and TV metadata have been provided in RDF/OWL format, as well

as relevant vocabularies, eg TV-Anytime, XMLTV Genres, IMDB Genres and Locations, WordNet, IconClass, Getty AAT, TGN and ULAN.

The RDF/OWL graphs resulting from the aligning and enriching of these vocabularies and metadata allowed for building a combined (virtual and physical) user model and further apply it (1) for (semi-)automatic generation of virtual and physical museum tours, and (2) semantic recommendations of TV programs. In CHIP AAT Style metadata had been added and further mapped to the artists in the Rijksmuseum collection, to allow for serendipitous recommendations based on relationships between artist and artworks in the same style or relationships between artists, eg collaborator and student_of. The use of RFID technologies allows for an instant interaction with artworks in the museum and for its logging for the user profile.

**Please contact:**
Lora Aroyo, VU University Amsterdam, The Netherlands
Tel: +31 20 598 2868/7718
E-mail: l.m.aroyo@cs.vu.nl

# Discovery and Selection of Services on the Semantic Web

by Dimitrios Skoutas, Alkis Simitsis and Timos Sellis

*Scientists from the National Technical University of Athens and the IBM Almaden Research Centre are proposing a novel infrastructure for ranking and selecting Web services using semantic Web technology. This approach uses the measures of recall and precision to evaluate the similarity between requested and provided services and expresses that similarity as a continuous value in the range of 0 to 1.*

Service-oriented architectures constitute a key technology for providing interoperability among heterogeneous systems, and integrating inter-organisation applications in a loosely coupled fashion. Due to their increasing popularity and adoption and the widening availability of Web services, the problem of effectively and efficiently discovering and selecting appropriate serv-

ices to meet specific user or application requirements or to compose complex workflow processes becomes a critical issue. On the other hand, the semantic Web is now enhancing the current Web with machine-processable metadata, giving formal and explicit meaning to information and thus making it processable not only by humans but also by software agents.

Our approach aims at the discovery and selection of appropriate advertised services that match a requested service by combining techniques from both the aforementioned areas. Figure 1 depicts a high-level overview of the individual steps – sequentially numbered in the figure – that are required in such a process. In the remainder of the article, we elaborate on the semantic matching.

## Service Descriptions

UDDI provides syntactic interoperability and a classification scheme to describe service functionality. However, its support for semantics is limited, affecting the recall and precision of the obtained results. Several initiatives exist for the purpose of semantically enriching the descriptions of Web services (OWL-S, WSDL-S, WSMO), so that a high degree of automation can be achieved in performing fundamental tasks such as service discovery and composition. Hence, services are described in terms of their input and output parameters, preconditions and effects. Each parameter is semantically annotated using terms in a corresponding OWL (Web Ontology Language) domain ontology. A service description may also contain non-functional parameters regarding quality-of-service (QoS) aspects.

## Semantic Matchmaking

The process of matching semantic Web services is essentially based on the use of logic inference to check for equivalence or subsumption relationships between the ontology classes that correspond to the parameters in the service descriptions. Typically, five types of match are identified: exact, if the request is equivalent to the advertisement; plug-in, if the request is subsumed by the advertisement; subsume, if the request subsumes the advertisement; intersection, if the intersection of the request and the advertisement is satisfiable; and disjoint, otherwise.

## Ranking Discovered Services

The ability of a search engine to rank the returned results is essential for two reasons. First, given that exact matches are often rare, and there is potentially a huge number of partial matches, ranking facilitates the location of the most suitable candidates. Second, users often do not have sufficiently clear information needs. Thus, consequent searches may be performed, and the ranking of intermediate results provides useful feedback for refining the search criteria until converging to a search that retrieves the final item(s).

We consider the following desiderata for a ranking mechanism for Web services:

- the notion of similarity should be based on assessing the semantic similarity between the request and service
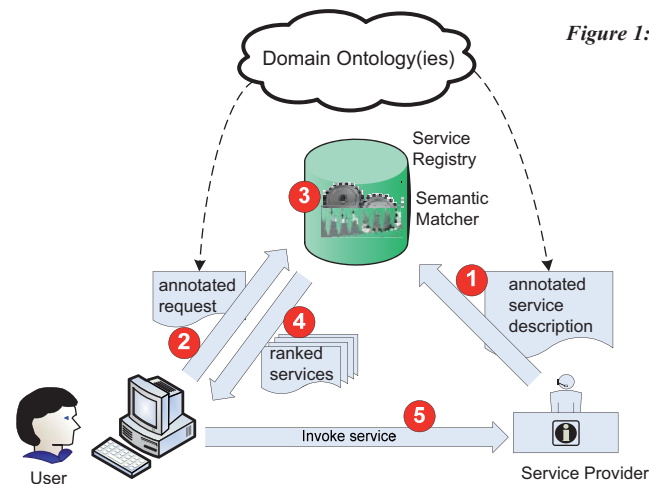


*Figure 1: System overview.*

parameters, including inputs, outputs, preconditions, and effects
- the similarity function should be asymmetric, so as to distinguish whether the service capabilities are a superset or a subset of those requested
- the similarity function should consider the semantic information encoded in the domain ontology, including both the class hierarchy and the properties of classes, as well as restrictions specified on these properties, so that the higher expressiveness of OWL ontologies can be exploited
- the ranking mechanism should be flexible and customizable, and be able to adapt to specific application needs or user requirements and preferences.

We use the notions of recall and precision to assess the similarity between requested and provided services. Similar to existing works, the proposed discovery mechanism uses the available semantic information provided by the domain ontology, and performs logic inference to estimate the degree of match between the requested and offered capabilities. However, the degree of match is not expressed in a discrete scale, such as the five types of match discussed above, but as a continuous value in the range [0..1]. This allows the handling of cases where a large number of candidate services provide the same type of match, ie partial matches can be appropriately ranked. Moreover, using these two measures provides an intuitive way of capturing asymmetry in the matching. Essentially, this encourages advertisers to be honest with their descriptions: the service provider is obliged to strike a balance between these two factors in order to achieve a high rank. Ranking exploits the semantic information encoded both in the class hierarchy and the properties

of the classes, including hierarchy of properties and value or cardinality restrictions. It is thus useful for both applications relying on taxonomies and those employing more expressive ontologies.

Finally, the proposed ranking mechanism is flexible and customizable, allowing the consideration of user preferences. The user may determine the relative importance of each search parameter, and apart from presenting a single rank for each candidate service, more detailed results may also be provided (eg separate values for recall and precision or the degree of match for specific parameters). This aids the user in identifying the most suitable service or refining the search criteria.

## Prototype Implementation and Future Work

Our approach is complemented by a prototype implementation based on Jena, a Java framework for building semantic Web applications, and Pellet, an open-source OWL-DL reasoner. While currently only functional service parameters are considered, in the near future the ranking mechanism will also support QoS parameters. Another scheduled extension involves the facilitation of service composition and the ranking of composite services.

Please contact:
Dimitrios Skoutas
National Technical University of Athens, Greece
Tel: +30 210 7721436
E-mail: dskoutas@dblab.ece.ntua.gr

# QoS-Based Web Service Description and Discovery

by Kyriakos Kritikos and Dimitris Plexousakis

*The success of the Web service paradigm has led to a proliferation of available services. While sophisticated semantic (functional) discovery mechanisms have been invented to overcome UDDI's syntactic solution, the number of functionally equivalent Web services returned is still large. The solution to this problem is the description of the non-functional aspect of Web services, in particular quality of service (QoS), which is directly related to their performance. We are currently developing a semantic framework, which includes ontologies and matchmaking algorithms, for the semantic QoS-based description and discovery of Web services.*

The aim of this work is to develop a complete framework for the semantic QoS-based description and discovery of Web services. This framework will consist of ontological specifications for the semantic description of QoS of Web services, QoS metrics and specification matchmaking algorithms for the QoS-based Web service discovery and a prototype system implementation for managing these specifications and algorithms.

Associated with each Web service is a set of non-functional attributes/characteristics that may impact its quality of service. Each QoS attribute is measured by one or more QoS metrics, which specify the measurement method, schedule, unit, value range, authority and other measurement details. A QoS specification (offer or demand) of a Web service is materialized as a set of constraints on certain set of QoS metrics. These constraints restrict the metrics to have values in a certain range or in a certain enumeration of values, or to have just one value. Current modelling efforts of QoS specifications in fact differ only in the expressiveness of these constraints. However, when it comes to the modelling of QoS metrics, these efforts fail. The main reason is that their QoS metric model is syntactic, poor (ie not capturing all aspects of QoS metric definition) and not extensible. Based on the above deficiencies, we have conducted extensive research that resulted in the identification of a set of requirements for a complete QoS-based Web service description. By taking into account these requirements, we have developed an upper ontology of QoS-based Web service description called OWL-Q. This is comprised of many sub-ontologies (facets), each of which can be extended independently of the others. This ontology also complements

OWL-S, the W3C submission for the functional specification of Web services. The sub-ontology specifying a QoS metric is depicted in Figure 1.

The most prominent QoS-based Web service discovery algorithms fail to produce accurate results because they rely either on syntactic descriptions of QoS metrics or on semantically poor QoS metric descriptions. Hence, in both cases, they cannot infer the equivalence



*Figure 1: The QoS metric sub-ontology of OWL-Q.*

of two QoS metrics based on descriptions provided by different parties (eg Web service provider and requester). Different specifications can occur for two reasons: either because of a different perception of the same concept, or a different type of system reading or measurement (low-level or high-level) for the same metric. Provided that two QoS metric descriptions are expressed in OWL-Q, we have developed a rule-based QoS metric-matching algorithm that infers the equivalence of the two metrics. This algorithm is composed of three main rules, each corresponding to a different case in a comparison of two metrics. The last rule is recursive and

reaches the final point of checking the equivalence of two mathematical formulae in order to infer the equivalence of the metrics. Equivalency of mathematical expressions is generally undecidable, but powerful mathematical engines like Mathematica and Maple can successfully deal with many hard cases.

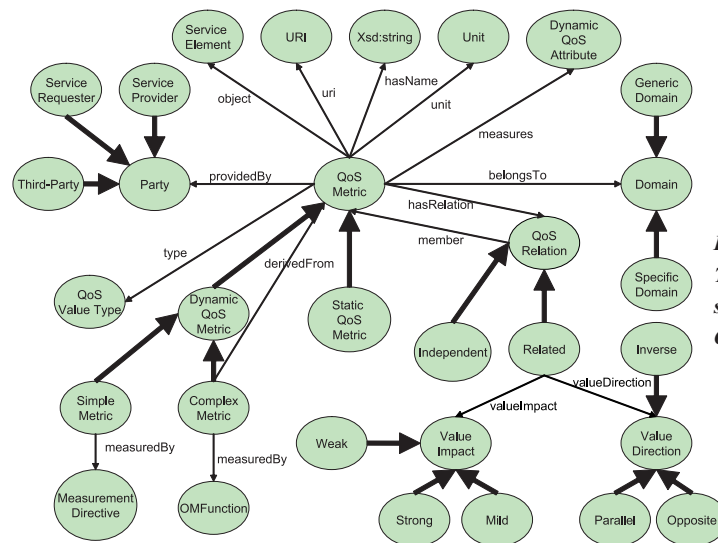One of the most prominent approaches for QoS-based Web service discovery transforms QoS-based Web service specifications into Constraint Satisfaction Problems (CSPs). Each CSP is checked for consistency (ie whether it admits any solution at all). Matchmaking is then performed according to the concept of conformance (ie if every solution of the offer is a solution to the demand). We have extended this approach in three ways: a) using our QoS metric-matching algorithm, we assign the same CSP variable to equivalent metrics; b) if equivalent metrics use different units, then a unit transformation procedure is carried out on the appropriate parts of the provider's CSP; and c) the types of discovery results were expanded.

We are currently in the development phase of our QoS-based Web service discovery process by using the Pellet reasoner for inferencing and the ECLiPSe Constraint Logic Programming (CLP) system for solving constraints. The Java programming language is used as a bridge between them. The architecture of the system under development is shown in Figure 2. Upon completion, the QoS metric-matching algorithm will be evaluated using synthetic data and simulation methods. In addition, we are currently investigating the relaxation of hard constraints into soft ones for solving the CSPs in cases where the QoS-based Web service matchmaking procedure fails to produce any useful result. Regarding QoS modelling, we plan to devise mid-level ontologies describing domain-independent QoS metrics for helping Web service providers and requesters in expressing QoS-based Web service specifications. Moreover, we intend to extend OWL-Q by modelling other types of non-functional attrib-
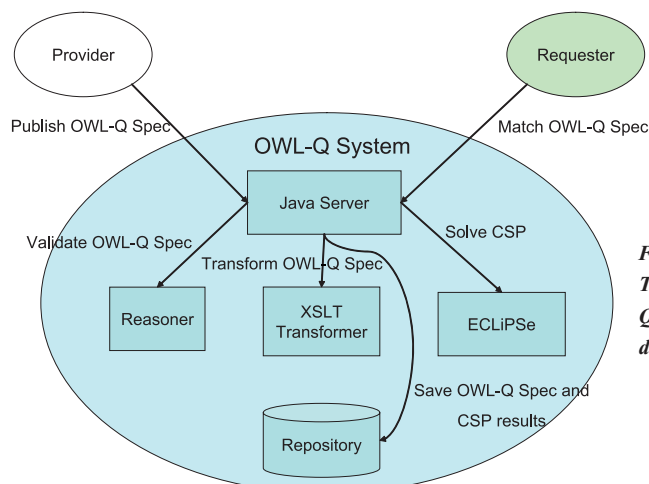
utes like context (location, time, user preferences, service execution environment etc). Last but not least, we envisage QoS-aware, dynamic Web service composition: this will be done by enforcing global QoS constraints, and for every metric, using metric evaluation functions imposed on any possible workflow (Web service) composition construct to produce global QoS metric values through construct reduction.



*Figure 2:*
*The architecture of the QoS-based Web service discovery process.*

**Links:**
http://www.ics.forth.gr/isl/index.html
http://www.w3.org/Submission/OWL-S/
http://pellet.owldl.com/
http://eclipse.crosscoreop.com/

**Please contact:**
Kyriakos Kritikos
ICS-FORTH, Greece
Tel: +30-2810-391637
E-mail: kritikos@ics.forth.gr

# Automating the Creation of Compound Web Applications

by Walter Binder, Ion Constantinescu, Boi Faltings and Radu Jurca

*The creation of compound, service-oriented Web applications is a tedious, manual task. The software designer must search for relevant services, study their application programming interfaces (APIs) and integrate them into the desired application, while also taking into account non-functional aspects such as service cost or reliability. We have been investigating models and algorithms to automate the service integration process, resulting in novel service composition algorithms that combine artificial intelligence (AI) planning techniques with advanced dynamic matchmaking in large-scale Web service directories.*

Building compound Web applications by integrating standardized Web services promises many benefits, such as reduced development effort and cost, ease of maintenance, extensibility and reuse of services. Service-oriented computing promotes the construction of applications by composing services. Service-oriented architectures maximize decoupling between services and create well-defined interoperation semantics based on standard protocols.

Service-oriented Web applications integrate the distributed Web services (WS) offered by various service providers. Web services are advertised in directories, which publish their programming

interfaces, grounding details (protocol, address etc) and eventually service-level agreements (SLAs) specifying the conditions and cost of service usage. Current technology includes Universal Description Discovery and Integration (UDDI) for Web service directories, Web Service Definition Language (WSDL) for Web service advertisements, WS-Agreement for SLAs, and Business Process Execution Language for Web Services (BPEL4WS) for representing compound services. In order to leverage available Web services, software designers must be aware of the service interfaces and structure their applications so that they can be integrated. Software designers searching

for suitable services (also using traditional search engines) are confronted with a large and dynamically changing search space, which in general must be explored manually. Such a manual approach to creating service-oriented systems is tedious and can hardly take advantage of the rapidly growing market of available Web services.

As a joint project between the University of Lugano and the Artificial Intelligence Lab of the Ecole Polytechnique Fédérale de Lausanne (EPFL), we have been exploring new techniques for automating the creation of compound service-oriented Web applications based on user requirements, while taking into

account large-scale directories of Web service advertisements. As an essential tool for the design of service-oriented Web applications, we promote the ServiceComposer. This takes a specification of user requirements as input and interacts with Web service directories, as well as with the software designer, in order to create a compound service that meets those requirements. Our approach supports traditional settings in which Web service advertisements merely convey interface information at the level of method signatures, as well as advanced settings in which semantically enriched service advertisements provide specifications of service behaviour. In the latter case, languages such as OWL-S or Web Service Modelling Ontology (WSMO) may be used to specify service semantics; the ServiceComposer uses the additional information to improve composition quality.

At the core of the ServiceComposer are planning and optimization algorithms that combine available services. On the one hand, the generated compound service must fulfil the functional requirements specified by the user. On the other hand, the ServiceComposer should choose individual Web services so as to minimize the cost of service usage and maximize reliability and availability of the compound service. Ideally, the ServiceComposer discovers a set of available Web services that can be assembled to meet all requirements. However, often there are specific aspects of the desired application that cannot be covered by off-the-shelf services, or the software designer may want to avoid using certain Web services, eg because of high cost or because of lack of confidence in the service reliability. Thus, the ServiceComposer cannot be a fully automated tool, since it needs to interact with the software designer. Rather, it offers design alternatives, suggests Web services to be integrated, and helps to minimize development effort by making use of available Web services wherever possible.

So far, as original scientific contributions, we have developed novel algorithms for service matchmaking and automated service composition, allowing to efficiently assemble services advertised in large-scale directories in order to meet given user requirements. We have succeeded in interleaving and integrating AI planning algorithms with remote directory access, for which we designed a dedicated query language and new matchmaking algorithms. Moreover, our work on automated Web service composition takes partial matches into account, increasing the success rate of our algorithms. In contrast, prior work on automated composition required a fixed and limited set of service advertisements (in the order of a few hundred services) that was hard-coded in a reasoning engine and did not support partial matches.

Our ongoing research concentrates on optimizing compound services with respect to non-functional properties such as service cost and service reputation. Moreover, we are developing a new middleware for dependable service-oriented applications. This will support failure recovery through dynamic replanning of service compositions, as well as proactive improvement of compound services, taking service reliability and performance into account. The latter features require a sophisticated monitoring infrastructure to be integrated into our middleware. Regarding concrete applications, we have demonstrated the feasibility of our approach with an evening planner that automatically integrates Web services for route planning, public transport schedules, restaurant recommendations, movie schedules etc. Currently, we are exploring ways to partly automate the creation of upcoming, social-based Web structures such as Web mashups.

**Links:**
A Flexible Directory Query Language for the Efficient Processing of Service Composition Queries:
http://www.igi-pub.com/articles/details.asp?id=6573

A Multi-agent System for the Reliable Execution of Automatically Composed Ad-hoc Processes: http://dx.doi.org/10.1007/s10458-006-5836-0

Flexible and Efficient Matchmaking and Ranking in Service Directories: http://doi.ieeecomputersociety.org/10.1109/ICWS.2005.62

Large Scale, Type-Compatible Service Composition:
http://doi.ieeecomputersociety.org/10.1109/ICWS.2004.1314776

**Please contact:**
Walter Binder
University of Lugano, Switzerland
E- mail: walter.binder@unisi.ch

Ion Constantinescu
Digital Optim LLC
E-mail: ic@opt.im

Boi Faltings, Radu Jurca
Artificial Intelligence Laboratory,
Ecole Polytechnique Fédérale
de Lausanne (EPFL), Switzerland
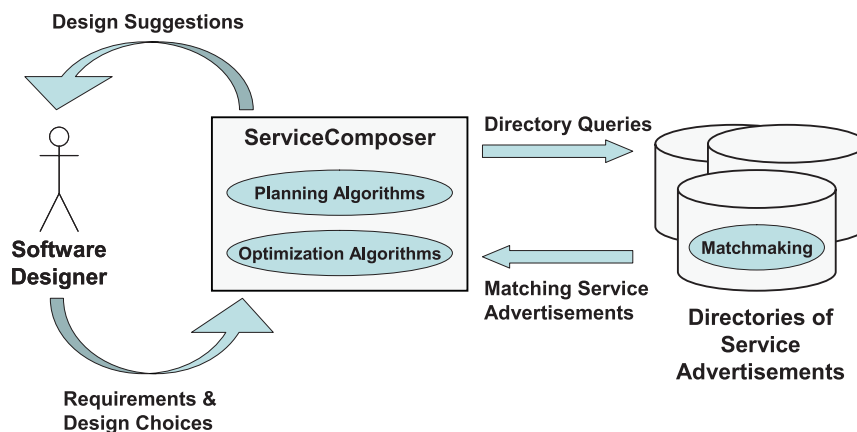E-mail: {boi.faltings,
radu.jurca}@epfl.ch

*Figure 1: The ServiceComposer helps the software designer in selecting and integrating available Web services that are advertised in distributed directories. The ServiceComposer uses AI planning and optimization algorithms and dynamically interacts with remote Web service directories that provide advanced matchmaking functionality.*

# Epigenetic Modelling

by Dimitri Perrin, Heather J. Ruskin, Martin Crane and Ray Walshe

*The field of epigenetics looks at changes in the chromosomal structure that affect gene expression without altering DNA sequence. A large-scale modelling project to better understand these mechanisms is gaining momentum.*

Early advances in genetics led to the all-genetic paradigm: phenotype (an organism's characteristics/behaviour) is determined by genotype (its genetic make-up). This was later amended and expressed by the well-known formula $P = G + E$, encompassing the notion that the visible characteristics of a living organism (the phenotype, $P$) is a combination of hereditary genetic factors (the genotype, $G$) and environmental factors ($E$). However, this method fails to explain why in diseases such as schizophrenia we still observe differences between identical twins. Furthermore, the identification of environmental factors (such as smoking and air quality for lung cancer) is relatively rare. The formula also fails to explain cell differentiation from a single fertilized cell.

In the wake of early work by Waddington, more recent results have emphasized that the expression of the genotype can be altered without any change in the DNA sequence. This phenomenon has been tagged as epigenetics. To form the chromosome, DNA strands roll over nucleosomes, which are a cluster of nine proteins (histones), as detailed in Figure 1. Epigenetic mechanisms involve inherited alterations in these two structures, eg through attachment of a functional group to the amino acids (methyl, acetyl and phosphate). These 'stable alterations' arise during development and cell proliferation and persist through cell division. While information within the genetic material is not changed, instructions for its assembly and interpretation may be. Modelling this new paradigm, $P = G + E + EpiG$, is the object of our study.

To our knowledge, no previous efforts have sought to model directly the mechanisms that affect epigenetic changes. Biological research on epigenetic phenomena is ongoing, but while some very promising articles are being published, most still contain only qualitative descriptions of epigenetic changes. This is not ideal when trying to develop computer-based models, but it is also not unusual. Over a decade ago the basics of HIV infection were understood, but quantitative data were sparse. Yet as early as 1992, differential equation models were proposed, while cell-mediated micro-models date from the 1990s. As more data have become available, these models have improved in sophistication, incorporating features such as shape-space formalism and massively multi-agent, parallel systems.
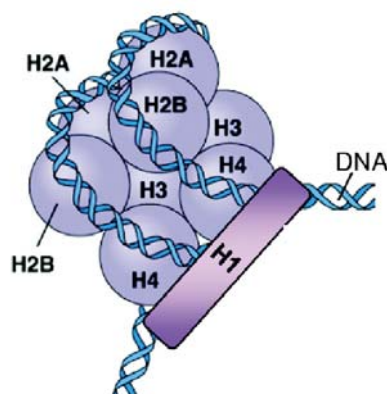


*Figure 1: A nucleosome, the fundamental subunit of the chromosome (adapted from C. Brenner, PhD thesis, Université Libre de Bruxelles, 2005).*

As a first step, we propose a microscopic model for chromatin structures. From the current biological results, it clearly appears that each unit (eg histone, DNA strand or amino acid) has a distinct role in epigenetic changes, and this role can alter depending on the type or location of the unit (eg which particular amino acid, what part of the DNA strand etc). For efficiency, this is best modelled using an object-oriented approach and a C++ implementation. The main objective of this early model is to provide a description and hierarchy for epigenetic changes at the cell level, as well as an investigation into the dynamics and time scales of the changes. These results will then be used to 'feed into' other models. Already in development are approaches such as agent-based modelling of cell differentiation and complex recurrent networks of cancer initiation by epigenetic changes.

Another early model uses Probabilistic Bayesian Networks. These represent a set of variables and their probabilistic dependencies and are constructed as directed acyclic graphs, for which nodes represent variables and arcs encode conditional dependencies between the variables. The variables can be of any type, ie a measured parameter, a latent variable or even a hypothesis. These networks can be used for inference, parameter estimation and refinement, and structure learning. This approach has been successfully used in medicine (eg breast cancer diagnosis) and biology (eg protein structure prediction), and epigenetic mechanisms appear amenable to such techniques.

Though still in its infancy, the project is gaining momentum and early work on the different approaches looks very promising. Active involvement from biologists and medical researchers is currently being sought in order to secure access to data and guarantee model realism (as highlighted by a presentation at the International Agency for Research on Cancer in early December 2007). Previous modelling experience from the group promises sensible integration of the various approaches and efficient implementations. Several publications and presentations are expected in the coming year, all of which will appear on the group's Web site (link below).

**Links:**
http://www.computing.dcu.ie/~dperrin/
http://www.computing.dcu.ie/~msc/publications.shtml

**Please contact:**
Dimitri Perrin
School of Computing, Dublin City University, Ireland
Tel: +353 1 700 8449
E-mail:
dimitri.perrin@computing.dcu.ie

# Contract-Oriented Software Development for Internet Services

by Pablo Giambiagi, Olaf Owe, Anders P. Ravn and Gerardo Schneider

*The 'COSoDIS' project - Contract-Oriented Software Development for Internet Services - is developing novel approaches to implementing and reasoning about contracts in service-oriented architectures (SOA). The rationale is that system developers benefit from abstraction mechanisms in working with these architectures. Therefore the goal is to design and test system modelling and programming language tools to empower SOA developers to deploy highly dynamic, negotiable and monitorable Internet services.*

As recently as several years ago, technology gurus predicted that the next big trend in software system development would be service-oriented architecture: that is, a successful integration of loosely coupled services belonging to different organisations, sometimes competing but on specific tasks collaborating, would storm the world. This would create a myriad of new business opportunities, enabling the formation of virtual organisations in which small and medium-sized enterprises would join forces to thrive in an increasingly competitive global market. The dream lives on, and the industry is pouring resources into developing and deploying Web services. Yet the degree of integration achieved between different organisations remains low. Collaboration presumes mutual trust, and wherever trust is not considered sufficient, business people turn to contracts as a mechanism to reduce risk. In other words, for the SOA to deliver its promised advantages, developers need cost-effective contract management solutions.

The main problems and open issues identified for supporting Web services development are the following:

1. Formal definition of generic contracts: currently, no unified system of syntax and semantics exists for contracts (in particular for quality of service (QoS) and security).
2. Negotiable and monitorable contracts: a contract must be negotiated until both parties agree on its final form, and it must be monitorable in the sense that there must be a way to detect violations. Current programming languages provide little support for negotiable and monitorable contracts.
3. Combination of object-orientation and concurrency models based on asynchronous message-passing. The shared-state concurrency model is not suitable for Web service development.
4. Integration of XML into a host language, reducing the distance between XML and object data models.
5. Harmonious coexistence at the language level of real-time and inheritance mechanisms.
6. Verification of contract properties: the integration of contracts in a programming language should be accompanied by the ability to guarantee essential properties. Guaranteeing the non-violation of contracts can be done in (at least) four different ways: (a) with runtime enforcement (eg through monitors); (b) by construction (eg through low-level language mechanisms); (c) with static program analysis techniques; or (d) through model checking.

None of the above can be used as a universal tool; they must be combined.

In addressing these issues and problems, we need to develop a model of contracts in an SOA that is broad enough to cater for at least QoS and security contracts. A basic requirement is the ability to seamlessly combine real-time models (for QoS specifications) and behavioural models (essential to constrain protocol implementation and enforce security). Contract models should also address discovery and negotiation. However, the formal definition of contracts should only be a first step towards the more ambitious task of providing language-based support for the programming and effective usage of such contracts.

Some contracts may be seen as wrappers that 'envelop' the code/object under the scope of the contract. Firewalls, for instance, may be seen as a kind of monitor of the contract between a machine
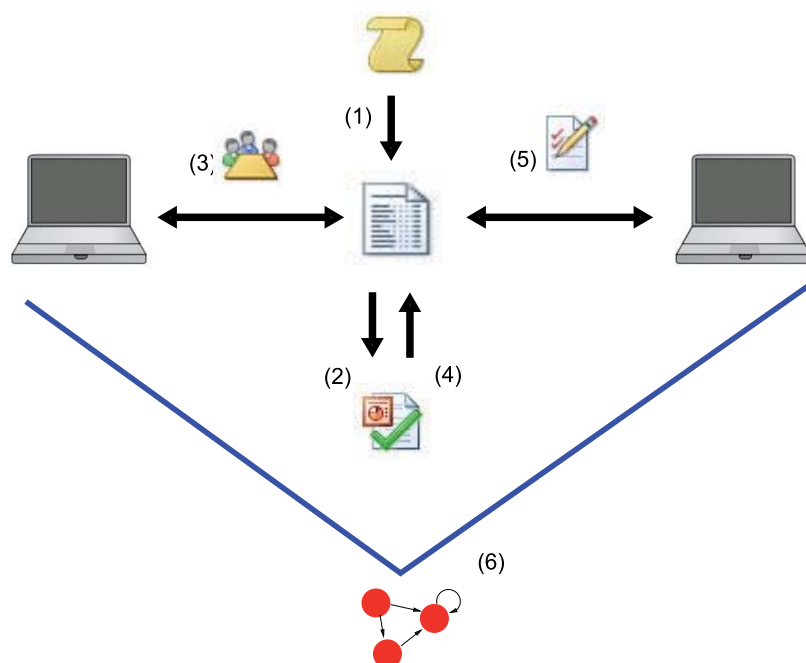


*Figure 1: A contract (template) is generated in an electronic version (1); the contract is checked to be free of contradictions (2); a negotiation starts (3); different versions of the contract are checked (4); a final contract is signed (5); a runtime monitor guarantees the contract fulfilment (6) .*

and the external applications wanting to run on that machine. It would be interesting to investigate a language primitive to create wrapped objects that are correct by construction.

Contracts for QoS and security could also be modelled as first-class entities using a 'behavioural' approach, through interfaces. In order to tackle time constraints (related to QoS), such interfaces need also to incorporate time.

Finding languages or notations for describing timing behaviour and requirements is easy: the real challenges are in analysis. Besides syntactic extensions, the language needs to have time semantic extensions in order to allow extraction of a timed model, eg a timed automaton. This model may be checked with existing tools such as Kronos or Uppaal, while other properties may instead be proven correct by construction (eg wrappers).

In practice, many properties can only be validated through runtime approaches. A promising direction is to develop techniques for constructing a runtime monitor from a contract, which will be used to enforce its non-violation (cf ongoing work with Java Modeling Language and Spec#).

In summary, our aim is to develop language-based solutions to address the above problems through the formalization of contracts as enriched behavioural interfaces, and the design of appropriate abstraction mechanisms that guide the developer in the production of contract-aware applications.

The COSoDIS project is a Nordic project funded by the Nordunet3 committee. The partners involved are the University of Oslo (Norway), Aalborg University (Denmark) and SICS (Sweden).

**Link:**
http://www.ifi.uio.no/cosodis/

**Please contact:**
Gerardo Schneider
University of Oslo, Norway
Tel: +47 22 85 29 71
E-mail: gerardo@ifi.uio.no
http://folk.uio.no/gerardo/

# gCube: A Service-Oriented Application Framework on the Grid

by Leonardo Candela, Donatella Castelli and Pasquale Pagano

*gCube is a new service-oriented application framework that supports the on-demand sharing of resources for computation, content and application services. gCube enables the realization of e-infrastructures that support the notion of Virtual Research Environments (VREs), ie collaborative digital environments through which scientists, addressing common research challenges, exchange information and produce new knowledge. gCube is currently used to govern the e-infrastructure set up by the European integrated project DILIGENT (A Digital Library Infrastructure on Grid-Enabled Technology).*

The long-term journey towards the e-science vision demands e-infrastructures that allow scientists to collaborate on common research challenges. Such infrastructures provide seamless access to necessary resources regardless of their physical location. These shared resources can be of very different natures and vary across application domains. Usually they include content resources, application services that manipulate these content resources to produce new knowledge, and computational resources, which physically store the content and support the processing of the services. Many e-infrastructures already exist, at different levels of maturity, and support the sharing of and transparent access to resources of a single type, eg SURFNet (information), GriPhyN (services), EGEE (computing and storage). However, they are still too primitive to support a feasible realization of VREs. The DILIGENT infrastructure, released recently by the homonymous project, will overcome this limitation by supporting in a single common framework the sharing of all three types of resource.

The core technology supporting such e-infrastructure is a service-oriented application framework named gCube. gCube enables scientists to declaratively and dynamically build transient VREs by aggregating and deploying on-demand content resources, application services and computing resources. It also monitors the shared resources during the lifetime of the VRE, guaranteeing their optimal allocation and exploitation. Finally, it provides mechanisms to easily create dedicated Web portals through which scientists can access their content and services.

From the technological point of view, gCube provides: (i) runtime and design frameworks for the development of services that can be outsourced to a Grid-enabled infrastructure; (ii) a service-oriented Grid middleware for exploiting the Grid and hosting Web Services on it; (iii) a set of application services for distributed information management and retrieval of structured and unstructured data.

Runtime frameworks are distinguished workflows that are partially pre-defined within the system; they include Grid-enabled services and application services, where the former coordinate in a pure distributed way the action of the latter, while relying on a high-level characterization of their semantics. Design frameworks consist of patterned blueprints, software libraries and partial implementations of state-of-the-art application functionality, which can be configured, extended and instantiated into bespoke application Grid services.

The service-oriented Grid middleware provides all the required capabilities necessary to manage Grid infrastructures. It eliminates manual service deployment overheads, guarantees opti-

mal placement of services within the infrastructure and opens unique opportunities for outsourcing state-of-the-art implementations. Rather than interfacing with the infrastructure, the software which implements the application services is literally handed over to it, so as to be transparently deployed across its constituent nodes according to functional constraints and quality-of-service (QoS) requirements. By integrating the gLite system released by the Enabling Grid for E-sciencE (EGEE) project for batch processing and management of unstructured data, gCube also allows the large computing and storage facilities provided by the EGEE infrastructure to be properly exploited. With over 20,000 CPUs and 5 million Gigabytes of storage, EGEE is the largest operational Grid infrastructure ever built.

gCube application services offer a full platform for distributed hosting, management and retrieval of data and information, and a framework for extending state-of-the-art indexing, selection, fusion, extraction, description, annotation, transformation and presentation of content. In particular, gCube is equipped with services for manipulating information objects, importing external objects, managing their metadata in multiple formats, securing the information objects to prevent unauthorized access, and transparently managing replication and partition on the Grid.
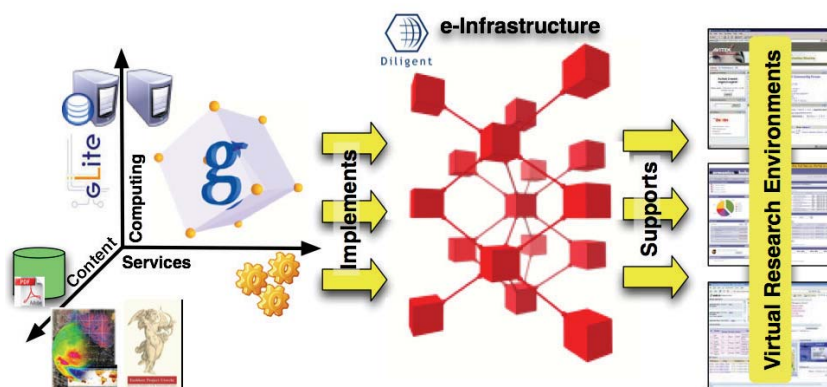


*Figure 1: gCube in action.*

In order to promote interoperability, gCube services are implemented in accordance with second-generation Web service standards, most noticeably SOAP, BPEL, WSRF, WS-Notification, WS-Security, WS-Addressing, and JSR168 Portal and Portlets specifications.

gCube is the result of the collaborative efforts of more than 48 researchers and developers in twelve different academic and industrial research centres: Institute of Information Science and Technologies ISTI-CNR (IT), University of Athens (GR), University of Basel (CH), Engineering Ingegneria Informatica SpA (IT), University of Strathclyde (UK), FAST Search & Transfer (NO), CERN European Organisation for Nuclear Research (CH), 4D SOFT Software Development Ltd (HU), European Space Agency ESA (FR), Scuola Nor-

male Superiore (IT), and RAI Radio Televisione Italiana (IT), and ERCIM.

Information about the project and detailed information about the software, currently available in its beta version, can be found at the gCube Web site.

**Links:**
gCube:
http://www.gcube-system.org/
DILIGENT:
http://www.diligentproject.org/
EGEE:
http://www.eu-egee.org/

**Please contact:**
Donatella Castelli, Pasquale Pagano
ISTI-CNR, Italy
E-mail: donatella.castelli@isti.cnr.it,
pasquale.pagano@isti.cnr.it

# The Decision-Deck Project – Developing a Multiple Criteria Decision Analysis Software Platform

by Raymond Bisdorff and Patrick Meyer

*The Decision-Deck project is developing an open-source generic Multiple Criteria Decision Analysis (MCDA) software platform composed of modular components. Its purpose is to provide effective tools for decision-aid consultants, for researchers in the field of MCDA, and for operations research teachers.*

The Decision-Deck project's objective is to provide an open-source software composed of various modular components, which pertains to the field of Multiple Criteria Decision Analysis (MCDA). It should give a user the ability to add, modify or simply use existing plugged-in functionalities ('plugins'). These con-

stituents can either be complete MCDA methods, or elements common to a large range of procedures. The typical end-user of the Decision-Deck platform is an MCDA researcher, an MCDA consultant or a teacher in an academic institution. The following MCDA methods have already been implemented: sorting

of alternatives into ordered classes based on an outranking relation (IRIS), progressive best-choice method based on an outranking relation (Rubis), best-choice method based on an additive aggregation model accepting imprecise information on the scaling coefficients (VIP), and ranking of alternatives with
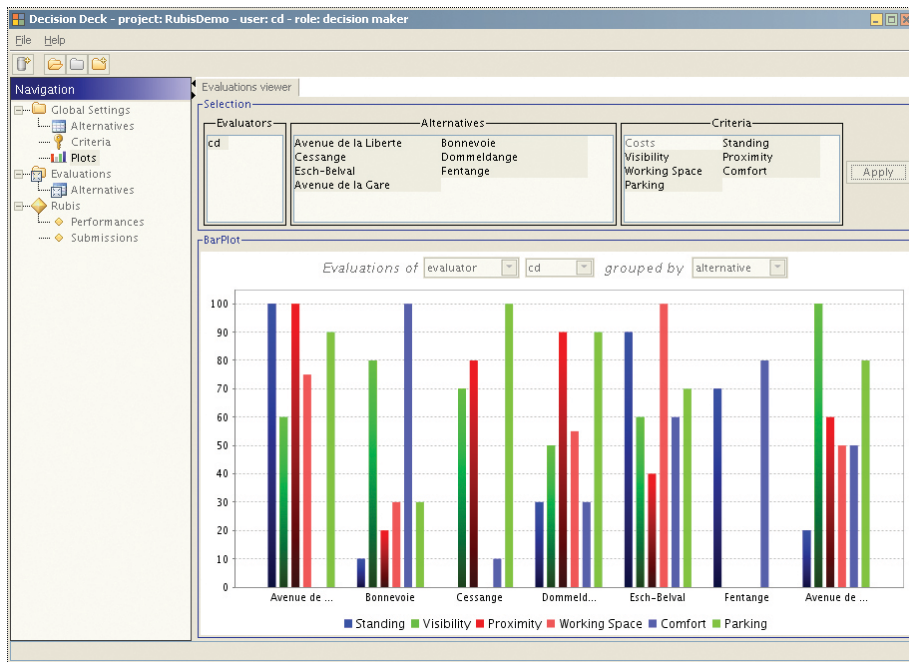
*Figure 1: One of the interesting features offered by the Decision-Deck software is the common availability of visualization resources as illustrated in the picture above. The snapshot, taken from a D3 Rubis plugin session, shows the performances of the alternatives on a subset of criteria in a column chart style.*

a set of value functions (UTA-GMS/GRIP).

In order to make these various software plugins reusable by the whole community, their implementation must be compliant with a set of predefined conventions and formats. Quite obviously, the success of the Decision-Deck project relies on a collaborative development effort from the MCDA community.

The Decision-Deck software is written in the Java programming language and is therefore platform-independent. Its latest version can be downloaded from the collaborative software development management system Sourceforge. Two kinds of implementation designs are available: a Java client which locally implements the MCDA methods (D2), and a distributed Web service and AJAX based architecture, serving the MCDA methods from a distributed Web server (D3). For example, the Rubis choice method is implemented as such a Web service on a server at the University of Luxembourg.

One of the most valuable features of the Decision-Deck software is the effective consideration of specific roles such as decision maker, evaluator, coordinator or facilitator in a given decision analysis project. For instance, evaluators from a variety of distant locations may communicate their evaluations via their local D2 clients to the common decision analysis project under the supervision of the project coordinator, whereas the decision maker may input personal

preferences via method-specific criteria tuning facilities offered in his/her local client (see Figure 1).

One of the tasks of the Decision-Deck project is to develop and maintain the Universal MCDA Modelling Language (UMCDA-ML), an XML standard which describes in a generic way the inputs and outputs of MCDA methods, as well as the different steps of a decision analysis workflow. The purpose of UMCDA-ML (see Figure 2) is to allow easy integration of MCDA Web services such as the Rubis server above, while also facilitating communication and data exchange between core components of the software platform.

The founding partners of the Decision-Deck project, which started in early 2006, are the MathRO laboratory of the

Faculty of Engineering of Mons, the Lamsade laboratory of the University Paris-Dauphine, the ILIAS laboratory of the University of Luxembourg and Karmic Software Research.

**Links:**
http://www.decision-deck.org/
http://decision-deck.sourceforge.net/
http://sourceforge.net/projects/decision-deck/
http://ernst-schroeder.uni.lu/rubisServer
http://charles-sanders-peirce.uni.lu/bisdorff/

**Please contact:**
Raymond Bisdorff
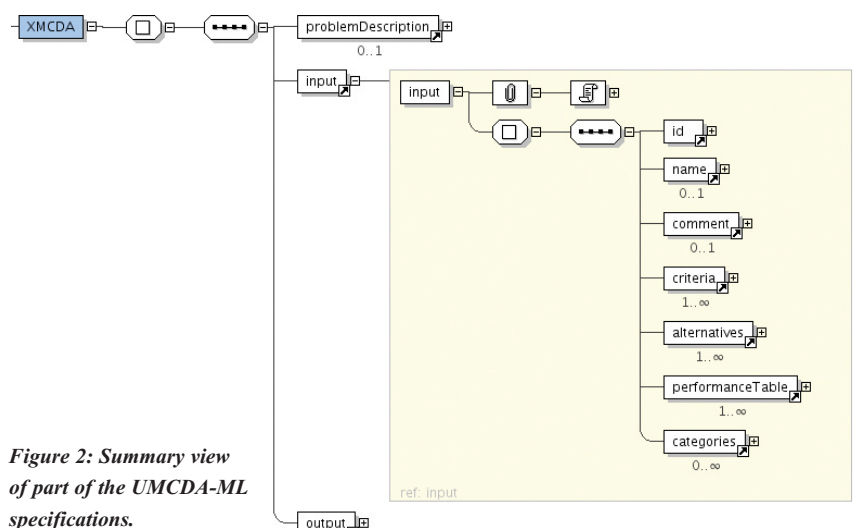University of Luxembourg
E-mail: raymond.bisdorff@uni.lu

*Figure 2: Summary view of part of the UMCDA-ML specifications.*

# Lightspeed Communications in Supercomputers

by Cyriel Minkenberg, Ronald Luijten and Francois Abel

*Optics offer significant potential as the communication infrastructure of choice for future supercomputers. The exploratory OSMOSIS project has resulted in a prototype of an electronically controlled optical packet switch with an aggregate capacity of 2.5 Tb/s. This device is designed specifically to meet high-performance computing (HPC) requirements.*

Recently, the rate of progress in CPU performance has slowed down considerably, mainly because of power constraints. Nowadays, the preferred approach to increasing computer performance is to use many processors in parallel, each relatively simple and power-efficient, thus giving rise to massively parallel computer architectures with thousands or even tens of thousands of processors making up a single machine. To collectively solve a problem, these processors need to communicate among each other with low latency and high bandwidth. In such a machine, the network used to interconnect the processors is a crucial factor in determining the system's overall performance.

Optical technology holds substantial promise for interconnection networks in HPC systems for a number of reasons. First, an optical switch typically consumes significantly less power than a comparable electronic switch because its power consumption is proportional to the packet rate rather than to the bit rate; second, fibre is much better suited than copper to the transmission of high data rates over long distances; and third, optics neither generate nor are sensitive to electromagnetic interference.

In the Optical Shared MemOry Supercomputer Interconnect System (OSMOSIS) project, Corning Inc and IBM Research have jointly explored this promise by developing an optical network switch prototype. The project was sponsored by the US Department of Energy. The result (see Figure 1) of this four-year endeavour is the highest-capacity optical packet switch in the world, with 64 ports running at 40 Gb/s, switching fixed-size packets (cells) of 256 bytes at an aggregate rate of 1.25 billion packets per second. A state-of-the-art electronic controller, which computes an optimal switch configuration once in every packet slot of 51.2 ns, ensures that throughput and reliability are maximized and latency minimized.



*Figure 1: The OSMOSIS optical switch prototype, partially populated.*

Optical switching has three main drawbacks, namely the absence of a practical, dense, fast-access optical memory technology, the complexity of its optical control and arbitration, and the cost of fast optical switching components. The OSMOSIS project addressed these drawbacks by adopting a hybrid electro-optical approach, using electronics to implement buffering and scheduling and optics for transmission and switching.
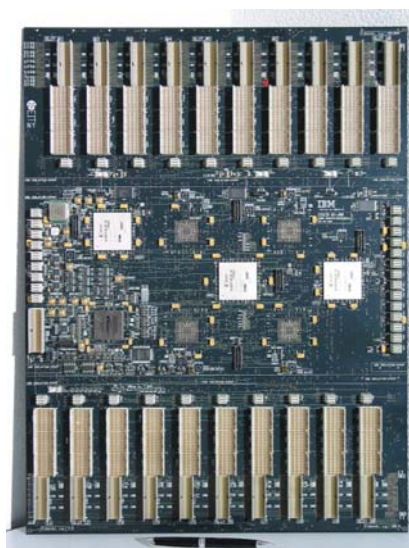


*Figure 2: The midplane of the OSMOSIS controller (57×43 cm$^2$).*

HPC switches must deliver very low latency, an extremely low bit-error rate, a high data rate, and extreme scalability. Moreover, they must be efficient for bursty traffic, with demands ranging from very small messages (eg collective operations and syncs) to bulk dataset transfers (eg memory pages). The main challenge for OSMOSIS was the short packet duration, which resulted in a small overhead budget of 12.8 ns for switching time, synchronization time, line coding, forward error correction and header information. This was achieved by using Corning's high-speed semiconductor optical amplifier (SOA) technology, which provided sufficiently fast switching times (< 3 ns).

Moreover, unlike most other optical switches, which are designed to be reconfigured at a relatively slow rate using a circuit switching or container-switching approach, OSMOSIS was designed to perform just-in-time packet-by-packet switching, reconfiguring the data path every 51.2 ns. This put a heavy burden on the controller, which was addressed by IBM's novel highly distributed controller architecture. The challenge was magnified by being limited to using only field-programmable gate arrays (FPGAs) for

reasons of cost and flexibility. The controller's final implementation uses a 36-layer midplane (see Figure 2) holding 40 daughter cards and a total of 48 high-end FPGAs.

The key requirement in computing systems is low latency, as latency typically translates directly into idle processor time. To meet the requirement of less than 1 µs latency measured application-to-application, we designed novel control algorithms to reduce latency by up to 50%. Further requirements for our prototype system included a bandwidth efficiency of 75% user payload, a maximum link utilization of at least 95%, a bit-error ratio less than 10-21, and scaling capability to support more nodes (>= 2048) and higher line rates (>= 160 Gb/s).

The OSMOSIS project has demonstrated the viability of SOA-based optical packet switching. It is the first system-level optical packet switch with more than two ports operating at 40 Gb/s that does not resort to container switching.

OSMOSIS also sports a number of architectural and algorithmic innovations. For instance, it uses a combination of a speculative transmission scheme with dual receivers per output to significantly reduce latency. Its deeply pipelined controller architecture uses multiple independent scheduling engines in parallel to achieve high utilization. Its highly distributed nature allows physical distribution across multiple FPGA devices to manage complexity. At the heart of the controller is one of the world's most complex board designs (36 layers, 13,000 wires, 45,000 connections), which was awarded a Mentor Graphics PCB Technology Leadership Award. It also features one of the highest con-

centrations of high-end FPGAs in a single system.

To make the OSMOSIS switch architecture commercially viable and fit for practical deployment, drastic cost reductions through dense integration of optical components and ASIC (application-specific integrated circuit) integration of controller logic will be needed. The critical technology to make this level of integration cost-effective is already under development.

**Link:**
http://www.zurich.ibm.com/~fab/Osmosis/

**Please contact:**
Cyriel Minkenberg
IBM Research, Switzerland
E-mail: sil@zurich.ibm.com

# Mobile-Based Wireless Sensor-Actuator Distributed Platform in Pervasive Gaming

by Pär Hansson and Olov Ståhl

*Wide-area wireless sensor/actuator networks can be used across a broad range of application domains. Within the Integrated Project on Pervasive Gaming (IPerG), we use these types of systems to establish a computer-supported citywide role-playing game. SICS develops one such platform, with research spanning such areas as wireless sensor/actuator hardware platforms, distribution software platforms, configuration software, and tools for user-created content.*

IPerG is an EU-FP6 project on Pervasive Gaming - games that integrate the technical approaches of computer gaming with emerging interface, wireless and positioning technology to create game experiences that combine both virtual and physical game elements. IPerG researches technical, design and business aspects of pervasive games. Whereas the player of a 'regular game' can participate in a game only at certain places and times, pervasive games expand the boundaries spatially (eg playing on the streets, across the globe, use of positioning), temporally (eg interlacing gaming with everyday life), and socially (eg non-players can be invited, players can produce content).

The project's infrastructure research theme is focused on identifying the requirements that pervasive gaming puts on supporting technology such as

devices, software platforms and networks. One aspect that makes pervasive games different from traditional PC- or console-based games is that the players are often required to move around in the physical environment. This has several implications for the design of software platforms that support such games, and often means that techniques and principles used to achieve today's computer-based multiplayer games need to be modified or cannot be used at all. PC- and console-based multiplayer games typically assume a networking environment where connections are stable and latencies are low or moderate. This is different from many pervasive games, in which players are mobile and the use of radio-based communication means that the risk of high latencies and frequent disconnections cannot be neglected. One of the aims of the IPerG project is to develop platforms that are

adapted to the specific characteristics of pervasive games, including:
• frequent disconnection: devices will often move in and out of connection
• mobile players: players move about forming ad-hoc arrangements of devices
• lightweight: must be able to run on low powered devices, especially phones
• heterogeneous: needs to run over multiple networks including WLAN, GSM/GPRS/3G and Bluetooth
• sensors/actuators: use of sensor or actuator hardware, for instance as part of purpose-built game props.

One of the platforms developed by SICS within IPerG is called PART (Pervasive Applications RunTime). A typical PART game session consists of a number of handheld or embedded devices, eg mobile phones or microcon-

*Figure 1: A small PART session consisting of processes running on a PC, a phone and a microcontroller.*

trollers, which are connected to some sensor or actuator hardware. On each such device runs a PART process, which is an application that has been built using the PART middleware. These processes are in turn connected to each other, network-wise, via point-to-point connections, which allow them to communicate and exchange data using the PART protocol. Such network connections can be set up in an ad-hoc fashion, since PART does not dictate any particular connection topology (eg a star shaped client-server topology). Furthermore, it supports several communication protocols such as tcp, http, and btspp. In addition, PART also provides support for distributed objects, which can be used to represent a game state that needs to be shared and synchronized among a set of game processes.

A PART process that connects to some sensor hardware will typically create a distributed object representing this sensor. The process will also dynamically add properties to this object, representing different types of sensor readings (eg temperature and humidity). In a similar way, hardware that controls actuators such as, for example, a light switch, can be represented by distributed objects and properties. One of the key features of the PART object model is that objects created by one process can be discovered and accessed by other processes. This allows remote processes to read or set the values of properties of such objects, providing access to the hardware they represent.

IPerG has implemented and demonstrated its results in a number of showcase games. These include Live-Action Role Playing games and Location-based games, both of which have been focused around ubiquitous computing concepts for interacting with the physical world. Game technology setup is centred around players roaming around a city, carrying a mobile phone and any number of extra devices, all highly connected to players of the same game.

Typically these 'ubicomp' games are not played on the mobile screen; rather, they emphasize real-world interaction through other means and employ the phone as a communication node running the PART process. Extra devices, carried or placed in the environment, are automatically discovered using Bluetooth and then become part of the gaming experience. Support for devices is easily extensible and ranges from custom-built modular hardware that can be fitted with arbitrary sensors/actuators (eg touch, light, temperature, vibration, audio, RFID etc) to off-the-shelf units such as GPS.

The partners in the IPerG project are SICS, the Interactive Institute and Gotland University in Sweden, Nokia Research and the University of Tampere in Finland, the University of Nottingham and Blast Theory in Great Britain, and the Fraunhofer Institute and Sony Europe in Germany. The project has been running since September 2004 and will continue until February 2008.



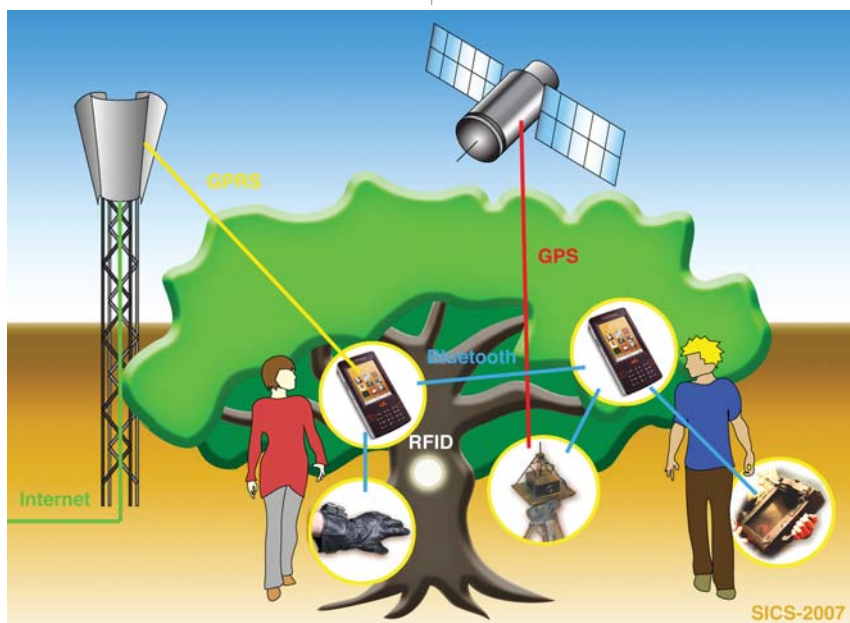*Figure 2: Typical mobile technology setup for the IPerG ubicomp demo game, with GPS, RFID reader and graphics tablet connected to players.*

**Links:**
http://part.sourceforge.net/
http://pimp.sourceforge.net/
http://www.pervasive-gaming.org/
http://prosopopeia.se/

**Please contact:**
Pär Hansson
SICS, Sweden
Tel: +46 8 633 1575
E-mail: par@sics.se

# An Informatics Research Contribution to the Domotic Take-Off

by Vittorio Miori, Dario Russo and Massimo Aliberti

*While many household appliances and hi-tech devices populate our lives, each one is independent of the others. Informatics research on domotics has the target of achieving interoperability between all the devices available in a home, automating some particular sequences of tasks and integrating the home with the outside world and especially with the Internet. The lack of an open and universal domotic standard is the principal reason for the poor use of domotics in real environments, so a new software architecture based on the SOA (service-oriented architecture) model and a domotic XML language has been created in order to realize a homogenous and independent vision of the home environment.*

Domotics is the discipline that investigates how to realize an intelligent home environment. Benefits of the use of domotic technology include improved energy efficiency, on-command entertainment, increased safety, greater independence and access to external services. This is particularly important in addressing the needs of elderly and disabled people. These services are generally based on hardware and software infrastructures commonly called domotic systems (X10, Konnex, Lon-Works, UPnP, HAVi, Jini etc), which support the main available communication standards (Ethernet, FireWire, Bluetooth, ZigBee, IrDA, proprietary buses etc). However, they are often developed by private companies and therefore distributed as proprietary and closed technology.

At present, too many domotic systems are crowding the market, and these systems are rarely interoperable (Figure 1). This represents a serious obstacle to solid market growth. Consumers should be able to choose devices according to their requirements and to criteria such as cost, performance, trends and confidence, without having to worry about issues of compatibility with their existing systems. Unfortunately, common market logic currently binds consumers, forcing them to purchase only devices that conform to a single system in order to obtain a suitable level of interoperability. Further, in our opinion none of the domotic protocols that are currently available will ever be able to achieve a truly leading role.

We thus propose the introduction of a new communication infrastructure and a new XML language to enable exchange of information between nodes belonging to distinct domotic systems. This new infrastructure, called Domotics Network (DomoNet), is based on a service-oriented architecture (SOA) model, in which services coincide with the functionalities offered by the devices. DomoNet uses Web services technology, a standard communication paradigm in distributed SOA contexts, and uses XML and XML-schemas to define a language that describes datatypes, messages, devices and their services. This universal language, called DomoML, allows a standard communication between the devices inside DomoNet.

As shown in Figure 2, DomoNet provides a network infrastructure for ambient devices in which each domotic system represents a sub-network connected to it through a dedicated gateway.

From a functional point of view, DomoNet is constituted by a Web service, called DomoNetWS, and by a set of



*Figure 1: At home, various devices belonging different brands and distinct domotic systems don't speak each others.*
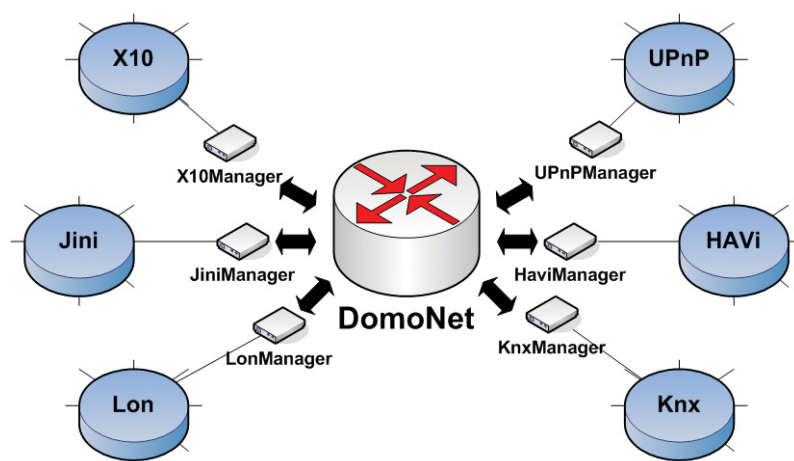


*Figure 2: domoNet architecture.*

modules that work as gateways, called techManagers, to deal with specific domotic systems. When an event generated by a device is captured by its techManager (eg a button is pressed), domoNetWS identifies it and finds the configured device and service (eg the switching on of a light) in order to execute the action.

The main task of domoNetWS is to correctly route messages between techManagers using a table containing pairs of addresses: a DomoNETAddress, assigned to every device that is connected to DomoNet, and the techManager's IP address (and a TCP port) which handles the specific domotic system.

DomoNetWS considers the house to be an Internet node that can share environments and device functionalities with any other DomoNetWS placed on the Internet. For example, functionalities and states can be shared between a user's main residence and the holiday home.

The system prototype developed at the ISTI-CNR Domotics Laboratory in Pisa is now open-source software released under GPL licence. It was written using Java and open-source libraries and tools. The prototype has been tested and is currently under further development. In its present state, it can correctly import and interact with all available devices such as televisions, washing machines and lights, regardless of the specific domotic system to which they belong.

The software has been tested using two domotic systems with different functionalities: Konnex and UPnP. These systems have two fundamental differences: first, in the ease with which devices can be configured (in Konnex this must be done by specialized staff, systems and software, whereas UPnP has an automatic plug-and-play auto-configuration system); and second, in the functions provided (Konnex is oriented towards lighting, energy effi-

ciency, safety and comfort, while UPnP focuses on multimedia applications).

Many activities based on the results of this project have been or will soon be launched. Currently, the most important effort concerns the definition of an XML language that is able to add semantics to domoNet in order to increase its usability. An initial part of this study is the development of an autoconfigurable universal home device controller, designed to be user-friendly for everyone, especially for elderly and disabled people.

**Links:**
Domonet home page:
http://sourceforge.net/projects/domonet
ISTI - CNR Domotic Laboratory:
http://www.domoticslab.it

**Please contact:**
Vittorio Miori
ISTI-CNR, Italy
E-mail: vittorio.miori@isti.cnr.it

# Natural Interaction -
# A Remedy for the Technology Paradox

by Lea Landucci, Stefano Baraldi and Nicola Torpei

*The advent of new technology means we are continuously offered high-tech devices that should make our life more agreeable, safe and pleasant. Do they succeed? In his book "The Psychology of Everyday Things", Donald Norman, a leading cognitive psychologist, warns us about the technology paradox: innovation technology risks making our life more complex every day. The intention of this warning is to make us conscious of the importance of 'human-centred design', particularly when we talk about human-computer interaction.*

The aim of Natural Human-Computer Interaction (NHCI) research is to create new interactive frameworks that integrate human language and behaviour into tech applications, focusing on the way we live, work, play and interact with each other. Such frameworks must be easy to use, intuitive, entertaining and non-intrusive. They must support interaction with computerized systems without the need for special external input-output equipment like mice, keyboards, remote control or data gloves. Instead, these will be replaced by hand gestures, speech, context awareness and body movement.

An interesting challenge for NHCI is to make such systems self-explanatory by working on their 'affordance' and intro-

ducing simple and intuitive interaction languages. Ultimately, according to M. Cohen, "Instead of making computer interfaces for people, it is of more fundamental value to make people interfaces for computers".

Our research team at the Media Integration and Communication Center (MICC) is working on natural interactive systems that exploit Computer Vision (CV) techniques. The main advantages of using visual input in this context are that visual information allows users to communicate with computerized equipment at a distance, without the need for physical contact with the equipment to be controlled. Compared to speech commands, hand gestures are advantageous in noisy

environments, in situations where speech commands would be disturbed, and for communicating quantitative information and spatial relationships.

We have developed a framework called VIDIFACE to obtain hand gestures and perform analysis of smart objects (detection and tracking) in order to understand their position and movement in the interactive area. VIDIFACE exploits a monochrome camera equipped with a near-infrared band-pass filter that captures the interaction scene at thirty frames per second. By using infrared light, the computer vision is less sensitive to visible light and is therefore robust enough to be used in public spaces. A chain of image-processing operations is applied

in order to remove noise, adapt to the background and find feature dimensions, optimized to obtain a real-time flow of execution.

## What Kind of Interfaces?

Our early results concerned vertical screens showing multimedia contents. Users point at a content display in order to access information: one significant installation is the PointAt system at the Palazzo Medici Riccardi museum in Florence. It is considered to be a good vanguard experiment in museum didactics, and has been functioning successfully since 2004. With the PointAt system, users can view a replica of the Benozzo Gozzoli fresco 'Cavalcata dei Magi' in the palace chapel, and request information on the displayed figures simply by pointing at them. Information is provided to the user in audio.

More recently, we have focused our research on TableTop multi-user frameworks that exploit table-like horizontal screens. This decision was effected by the concept of 'affordance'. According to D. Norman, this is the ability of an object to suggest the correct way in which it should be used. A table is a natural place to share with other people and can also be used to hold other objects.

Multi-user systems often involve interfaces that are rich in interactive actions and multimedia content. Our early solution used richer interaction languages to develop CV algorithms able to recognize different gestures and hand postures. However, we observed that such a solution causes an increasing cognitive load for users, who were forced to learn different gestures before using the sys-
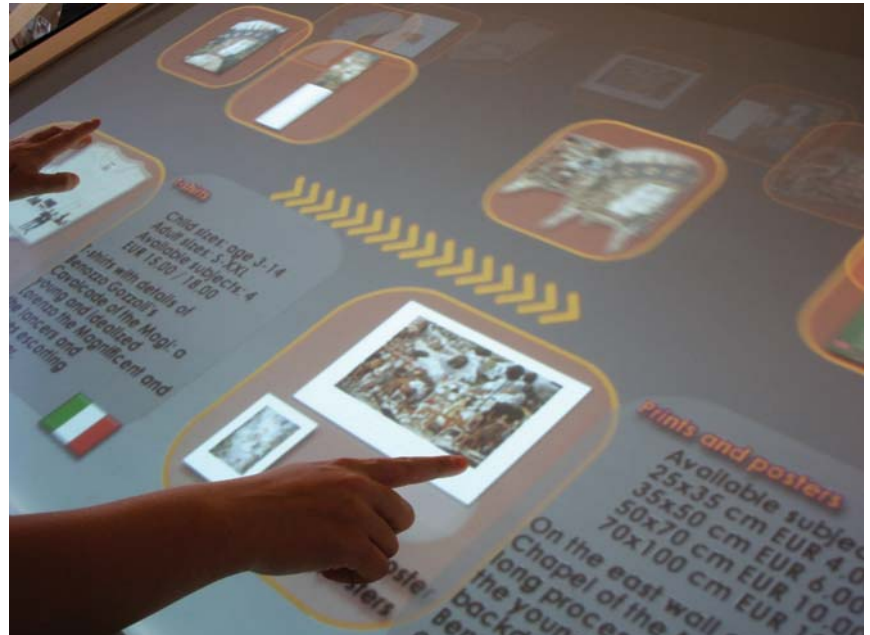


*Figure 1: VIDIFACE interactive bookshop, a tabletop multi-user natural interaction system.*

tem. This went against the concept of a self-explanatory system.

During the last year we have worked on different solutions for developing new natural interaction frameworks that exploit a minimal set of natural object-related operations (eg zoom-in for 2D images, zoom and rotate for 3D objects, open and turn page for books, unroll for foulards and so on). An interesting result is our Interactive Bookshop, which was recently prototyped and will soon be installed in the bookshop of Palazzo Medici Riccardi in Florence to display artistic objects as digital replicas.

Another solution exploits Tangible User Interfaces (TUIs) to enrich interaction

in the case of hinged interfaces. In our implementation of TUIs, users employ 'smart' physical objects in order to interact with Tabletop interfaces. Obviously the employment of such objects must be as intuitive as possible. The TANGerINE project, developed in collaboration with the Micrel Lab of the University of Bologna, exploits a 'smart cube': a wireless Bluetooth wooden object equipped with sensor node (tri-axial accelerometer), a vibra motor, an infrared LED matrix on every face and a microcontroller. We chose a cube once again because of its clear affordance: users intuitively consider the uppermost face 'active' (as if reading the face of a dice), thus conceiving the object as able to embody six different actions or roles.

**Links:**
http://www.micc.unifi.it
http://www.micc.unifi.it/vidiface
http://www.palazzo-medici.it/ita/
sperimenta.htm
http://www.micc.unifi.it/interactive-bookshop
http://www.tangerineproject.org

**Please contact:**
Lea Landucci, Stefano Baraldi and Nicola Torpei
MICC Media Integration and Communication Center,
University of Florence, Italy
E-mail: lea.landucci@gmail.com,
stefano.baraldi@gmail.com,
nicolatorpei@gmail.com
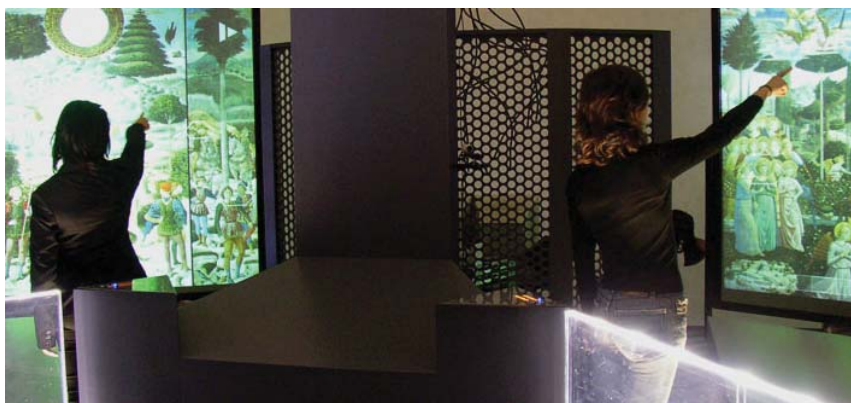


*PointAt, an interactive replica of Benozzo Gozzoli fresco 'Cavalcata dei Magi', installed in Palazzo Medici Riccardi Museum in Florence, Italy.*

SPONSORED BY ERCIM

# ISWC07 - IEEE International Symposium on Wireless Communication Systems

by Yuming Jiang

*The IEEE International Symposium on Wireless Communication Systems 2007 (ISWCS'07) was held in Trondheim, Norway, 16-19 October 2007. It was the fourth conference in the series and provided a platform for wireless communication researchers and technologists to identify and discuss technical challenges and business opportunities in wireless communication systems.*

In 2007, ISWCS became an IEEE-sponsored conference. Specifically, ISWCS'07 was sponsored by the IEEE and the IEEE Vehicular Technology Society. In addition, it was technically co-sponsored by the IEEE Communications Society and the Nordic Radio Association. ISWCS'07 also received strong institutional support from the Norwegian University of Science and Technology (NTNU), the University of Agder, and the Simula Research Laboratory. Besides, the conference received sponsorships from the City of Trondheim, the Research Council of Norway, Telenor Research and Innovation, The MathWorks and ERCIM.

The technical program of ISWCS'07 was very strong, which, in addition to paper sessions, included keynote speeches, panel discussion and tutorials. There were four keynote speeches given by internationally recognized experts from both academia and industry. Professor David Gesbert (Eurecom Institute) gave a talk on adaptation, coordination and distributed resource allocation in wireless networks. Dr. Hans Martin Ritt (The MathWorks) presented his view on the role of new development technologies in developing complex wireless systems. Professor Andreas F. Molisch (Mitsubishi Electric Research Labs and Lund University) highlighted how the MIMO antenna and channel properties impact wireless systems. Professor Mario Gerla (UCLA) talked about vehicle to vehicle communications and networking.

One panel discussion was organised. The panel members were Richard Savage (Qualcomm Europe), Stein Hansen (Telenor), Boon Sain Yeo (SensiMesh), Ralf Müller (NTNU) , Jarle Boe (Texas Instrument), and Victor Bahl (Microsoft Research). There were lively discussions during the panel session, and the panel provided a good forum for both the panelists and the audiences to debate the future of wireless communications.

Six tutorials were provided on 16th October 2007. These tutorials were carefully chosen from the proposals received by the conference. Topics of them covered a wide range of active areas in wireless communications, networking and systems. They included "Multipath Characterization of Antennas and Mobile Terminals for Diversity and MIMO Systems" by Per-Simon Kildal and Jan Carlsson (Chalmers University of Technology), "Ultra-Wideband Technology and its Standardization for Wireless Personal Area Networks" by Huan-Bang Li (NICT), "From Cognitive Radios to Cognitive Wireless Networks" by Petri Mahonen and Marina Petrova (RWTH Achen University), "Multiuser Communications" by Ralf Muller (NTNU), "Introduction to ITS Communication Solutions" by Runnar Søråsen and Per Jarle Furnes (Q-free), and "RFID Systems" by Boon Sain Yeo (SensiMesh). The conference gave each registered participant free access to one tutorial. All six tutorials were well attended.

The conference received 369 submissions and about 2/3 of the submissions were full papers. This figure exceeded the number of submissions to previous ISWCS conferences and showed that ISWCS was on the right path to greater recognition and international popularity. The review process involved all members of the Technical Program Committee and many other reviewers, all were recognized experts in their respective fields. It was taken care that the topic of each submitted paper fell within the reviewers' area of expertise. The review process was in accordance with standard single blind peer-reviewing practices, i.e., the reviewers know who the authors of the manuscript are, but the authors do not have access to the information of who the reviewers are. For most papers, at least 3 independent reviews were received. In some very few cases, we were satisfied with 2 reviews provided the two reviews were consistent in their conclusions. The result of the review process was that 170 papers were accepted for presentation at the conference and inclusion in the conference proceedings, among which 107 were oral presentation papers and 63 papers were poster presentations.

The conference received more than 210 registrations, and approximately, 180 people attended the conference each day. A Conference Dinner was arranged. About 120 people registered and attended the Dinner where the best student paper awards were given. ISWCS'07 awarded two Best Student Papers. They were chosen by a committee consisting of the General Chair, TPC Chair, Publication Chair and Speaker Chair. The decision was made based on the paper's originality, depth, and potential impact as well as the author's presentation at the conference.

I thank all people who have contributed to and made the conference a great success in Trondheim. Particularly, I am grateful to the conference organising committee, technical programme committee, reviewers, authors, presenters, sponsors, and the technical societies. My special thanks go to the students of NTNU and University of Agder for their great effort with the on-site arrangement.

ISWCS'08 will be held from 21-24 October 2008 in Reykjavik, Iceland. I wish them all the best for another successful conference.

**Link:**
http://www.iswcs.org/iswcs2007/

**Please contact:**
Yuming Jiang, ISWCS General Chair
NTNU, Norway
E-mail: ymjiang@ieee.org

# INRIA Celebrated 40 years of Activity

*INRIA celebrated its 40th anniversay with a large forum on information and communication science and technologies, and their past, present and future socio-economic impact at the Grand Palais in Lille on 10-11 December 2007. The aim was to analyse and present the 'information revolution' at the turn of the millennium, the genuine upheaval in economic and research activities, in employment organisation and social practices.*

Over the past 40 years, information and communication science and technology (ICST) has entered every area of social-economic life. It has revolutionised social practices, changed our working lives, sped up economic cycles and imposed new cultural habits and leisure activities. Information technology and software, the Internet and Web, plus a multitude of spin-offs, have developed on a scale never before wit-



*Michel Cosnard, CEO and chairman of INRIA adresses the audience.*

nessed in human history. And yet, the public at large has little awareness of the science behind these innovations.

The objective of the forum was to launch a debate on society, measuring how far we have come, analysing the changes made and attempting to explain how and why some types of technology have been so successful. Once we have that knowledge, we can look ahead to the near future, anticipate forthcoming changes, define good practices and set into motion new innovations to guarantee economic development in the years to come. The conferences brought together scientists from all disciplines and key guest speakers. It was composed of four plenary sessions and three theme sessions.

The plenary session featured the topics "Computer Science, history in the making", "Digital networks and the development of a new industrial model", "Information and Communication Technologies: change of scale and new challenges", and "New technologies: a cultural and cognitive revolution". Chaired by experts, historians, sociologists and philosophers (Michel Serres closed the forum speaking about cultural and cognitive revolution), the theme sessions focused on various aspects of ICST: historical perspectives, economic issues, societal impact and scientific challenges. The scientific theme

session discussed the questions "Is computer science really a science?" with Robin Milner, Cambridge University and Gilles Dowek, INRIA and Ecole Polytechnique; "Computer Science and Biology: a creative convergence?" with Luca Cardelli, Microsoft Research, Cambridge, Jean-Frédéric Gerbeau, INRIA and Jacques Haiech, University of Strasbourg 2; and "What scientific challenges does the future hold in store?" with Olivier Faugeras, INRIA, Bruno Sportisse, Cerea-ENPC/INRIA, and Vincent Quint, INRIA and W3C.

The economic sessions focused on the three topics "Has computing transformed engineering?", "Labels and consumers: the stakes involved in the shopping of the future?" and "From research to business: process dynamics". Speakers included Jean-François Abramatic, Ilog; Patrick Johnson, Dassault Systèmes; Armand Hatchuel, CSG, Ecole des Mines de Paris; Olivier Iteanu, specialist lawyer in information technology; Henri de Maublanc, Aquarelle.com, Acsel; Arnaud Mulliez, Auchan France; David Simplot-Ryl, INRIA; Jérôme Chailloux, ERCIM, Jean-Marie Hullot, Thierry Pollet, Alcatel-Lucent.



*Exhibitions.*

The third session discussed society-based topics such as "Surveillance society – the reality of tomorrow? (Barbara Cassin, CNRS, University Paris 4, author of 'Google-moi', Sébastien Canevet, specialist in Internet law, Alex Türk, French Senator, CNIL); "The individual in network: new cultural practices?" (Eric Bruillard, University Paris12, ENS Cachan, Jean-Pierre Dalbera, French Ministry of Culture; Patrice Flichy, University Marne la Vallée, Jean-Michel Salaün, University of Montréal, EBSI); and "Will the future of art go digital?" (Norbert Corsino, choregraph, n+n corsino, Philippe Morel, architect, EZCT; Steve Sullivan, Industrial Light & Magic).

As part of the celebrations, INRIA also organised an exhibition devoted to the history of information and communication science and technology. More than thirty demonstrations provided an overview of results today. The exhibition centred on an 'innovation village' offering scientific demonstrations, films and activities on a wide variety of subjects: health and biology, art and science, the virtual world, sustainable development, robotics, search engines, networks and sensors, languages, human-machine interaction, etc.

**More information:** http://www.inria.fr/40ans/

# ECSA 2007 -
# First European Conference on Software Architecture

by Carlos E. Cuesta and Paloma Cáceres

*The First European Conference on Software Architecture (ECSA 2007) was held in Aranjuez, near Madrid, on 24-26 September 2007. Conceived as the premier meeting place for European researchers in the area of software architecture, it has now become firmly established as one of the most prominent worldwide events in the field.*

Software architecture has just begun its second decade as a discipline within software engineering. During this period, its role in the engineering and design of complex software-intensive applications has become more and more important and widespread. The field is now experiencing the coexistence of component-based and service-oriented architectures, and because of that, this is a time of changes, which could even lead to a paradigm shift. From a general perspective, architecture presents itself as a key aspect in the design, development and evolution of software systems; the latter is perhaps the most challenging issue in the field today.

Due to the relevance of these issues, the existing successful series of European workshops on software architecture has evolved into a number of fully-fledged European conferences. Thus, the follow-up to the meetings held in the United Kingdom in 2004 (LNCS 3047), Italy in 2005 (LNCS 3527), and France in 2006 (LNCS 4344) has been the first edition of the European Conference on Software Architecture (ECSA 2007). This took place in Aranjuez (Madrid), Spain, on 24-26 September 2007 (LNCS 4758).

The conference attracted the interest of researchers in the field, increasing the number of submissions with respect to previous workshops. An initial profile of 89 abstracts provided a total of 62 contributions from 25 countries. From these, eighteen were selected for presentation, and only five were accepted as full papers. Moreover, sixteen additional submissions were selected as posters; these were exposed in a very intense poster session, held also as part of the main program. The selected papers covered many areas and promoted fruitful discussions among the participants. The structure of the program grouped them as follows:
- architecture description, with an emphasis on evolution
- architecture analysis
- architecture-based approaches and applications
- challenges in software architecture
- service-oriented architectures.

Furthermore, there were three keynote talks during the conference, given by the following outstanding researchers:
- Professor David Garlan from Carnegie Mellon University (USA), who presented the invited lecture "Software Architectures for Task-Oriented Computing"
- Professor Ron Morrison from the University of St Andrews (UK), who in his lecture described "An Active Architecture Approach to Dynamic Systems Co-Evolution"

- Professor Michael Papazoglou from Tilburg University (The Netherlands), who asked "What's in a service?" and discussed in detail the consequences and ideas around this concept.

Due to the success of ECSA 2007 and with the collaboration of all the participants, ECSA has already become one of the worldwide reference forums for software architecture. In addition, plans for future editions already include a joint event with WICSA, the premier meeting in the field, to be held in 2009.

The ECSA promoters would like to thank the Spanish Ministry of Education and Science, the the General Director of Universities and Research (DGUI) of the autonomous Government of Madrid, and Rey Juan Carlos University for their financial support; Kotasoft SI Ltd, Kybele Consulting, Open Canarias and InterSystems for their sponsorship; and the Spanish Office of the W3C for their kind collaboration.

**Link:**
http://www.kybele.es/ecsa/

**Please contact:**
Esperanza Marcos, SpaRCIM, Spain
E-mail: esperanza.marcos@urjc.es

# WWIC 2008, the 6th International Conference on Wired/Wireless Internet Communications

Tampere, Finland, 28-30 May 2008

Next generation mobile networks will be based on Internet core networks and wireless access networks. The need for efficient merging of the wired and wireless infrastructure as well as the new multimedia services and applications of next generation networks call for novel network architectures, protocols and traffic-related mechanisms.

WWIC addresses research topics such as the design and evaluation of protocols, the dynamics of the integration, the performance tradeoffs, the need for new performance metrics, and cross-layer interactions. The goal of the conference is to present high-quality results in the field, and to provide a framework for research collaboration through focused discussions that will designate future research efforts and directions. In this context, the program committee will accept only a limited number of papers that meet the criteria of originality, presentation quality and topic relevance.

WWIC is a single-track conference which has reached, within 5 years, the highest level of quality, which is reflected both in the level of participation as well as the acceptance ratio and the amount and quality of submitted papers.

### Conference Topics
The conference objectives will be pursued through highly technical sessions organised thematically and keynote talks offered by recognized experts. Topics of interest to WWIC 2008 include (but are not limited to) the following:

AAA in mobile environments, ambient networks, ad-hoc mobile networks, blended network configurations, beyond 3G networks technologies, cross layer interactions, economical issues of wireless networks, end-to-end Quality of Service support, handover techniques, heterogeneous wireless access networks, hybrid wired / wireless environments, integration of wired and wireless networks, mobile service level agreements / specification, mobility management, network design and network planning, network mobility, network coding in mobile networks, network security in mobile environments, performance evaluation of wireless systems, pricing, charging and accounting in wireless networks, QoS routing in mobile networks, QoS signalling in mobile environments, resource management and admission control, service creation and management for wireless, simulation for next generation mobile networks, traffic characterisation and modelling, traffic engineering, transport protocols and congestion control, wireless mesh networks, wireless multi-hop networks, wireless multimedia systems, wireless network monitoring, wireless sensor networks.

**More information:**
http://wwic2008.cs.tut.fi/



CALL FOR PARTICIPATION

# 5ECM, the Fifth European Congress of Mathematics

Amsterdam, 14-18 July 2008

5ECM is the fifth in a row of highly successful congresses that cover the broad range of the mathematical sciences, from very applied to very pure, and which are organised under the auspices of the European Mathematical Society (EMS). This year's congress will be organised in Amsterdam, under the special patronage of the Koninklijk Wiskundig Genootschap (Royal Dutch Mathematical Society – the KWG), by the mathematical departments of the VU University, the University of Amsterdam, and CWI. The preceding ECMs were held in Paris (1992), Budapest (1996), Barcelona (2000) and Stockholm (2004).

The scientific committee of 5ECM, headed by Lex Schrijver (CWI and University of Amsterdam) has composed a wonderful programme with ten plenary talks by outstanding mathematicians, three science lectures by top scientists and thirty-two invited talks by top specialists. A prize committee chaired by Rob Tijdeman (Leiden University) is busy selecting ten of the most promising young European mathematicians who will each be awarded a prize of €5000 and who will be invited to present a talk. In addition, there are 22 mini symposia to learn about the newest developments in the various areas of specialization of mathematics.

As a special activity we mention the presentation of the Brouwer medal which is awarded once every three years by the KWG. The Brouwer medal ceremony consists of a laudatio for the golden Brouwer medal winner and a lecture by the medal winner, followed by a welcome reception for the participants and accompanying persons of 5ECM. This ceremony will be organised on the evening of the first day of the congress.

The ECM congresses do not know submitted talks or papers, but registered participants have the occasion to present their mathematical work in the form of a poster.

The local organisation is in the hands of André Ran (VU University), Herman te Riele (CWI Amsterdam), and Jan Wiegerinck (University of Amsterdam).

### Important Dates
31 January 2008: poster submissions due
15 April 2008: early registration deadline

**More information:**
http://www.5ecm.nl/

## Alcatel-Lucent Bell Labs and INRIA Establish a Joint Research Lab

Alcatel-Lucent and INRIA have signed a joint research agreement that commits the two organisations to partner in developing advanced technologies for the next generation Internet. As announced by Valérie Pécresse, French Minister for Higher Education and Research in her opening speech on the occasion of INRIA's 40th anniversary, Bell Labs and INRIA have committed researchers and resources to jointly research and develop critical technologies to realize the challenge of self organising networks – considered by many to be key for the future of networking.

This joint research will cover four critical areas - semantic networking, high manageability, self-optimizing mobile cellular networks, and ad-hoc extended mobile networks - and will focus on generating results through new technologies, intellectual property, and milestone publications that pave the way for the telecommunications.

"The number of sensors and communication devices is bringing Internet complexity to a new dimension" said Jeong Kim, president of Alcatel-Lucent's Bell Labs. "This Joint Research Lab is a natural next step from our collaboration with INRIA which started 10 years ago." "We have developed a theory based on this approach, giving rise to a genuine paradigm of the virtuous circle of research and innovation" said Michel Cosnard, chairman and CEO of INRIA. "It is the people in industry who raise the questions that lead to new and exciting challenges for the scientists, who dream the innovations that the people in industry go on to develop." This partnership will bringing together a team of about 50 researchers from both organisations dedicated to this task.

## Dublin City Univesity to Lead Research in High-Tech Automatic Language Translation

Dublin City University (DCU) is to lead a multi-million euro research partnership funded by Science Foundation Ireland (SFI) that will develop the next generation of high tech automatic language translation. This five-year research programme will transform an important sector of Ireland's global software business – localisation - as well as a key driver of the global content distribution industry. DCU is collaborating in the project with academic partners, University College Dublin, University of Limerick and Trinity College Dublin, and with the companies IBM, Microsoft, Symantec, Dai Nippon Printing, and Idiom Technologies as well as Irish SMEs, Alchemy, VistaTech, SpeechStorm and Traslan.

The project is awarded €16.8m by SFI, and the industry partners are contributing €13.6m in materials, research services and additional funding. Ireland already has a substantial global footprint in the localisation industry – the process of adapting digital content, download manuals, software and other materials, to different languages and cultures.

The Irish project will tackle three critical problems for the localisation industry:
- *Volume:* The amount of content to be translated and localised to the destination culture and environment is growing rapidly and massively outstrips the supply of human translators.
- *Access:* Powerful, small devices such as mobile phones and PDAs require novel technologies integrating speech and text to support "on the move" delivery of, and access to multilingual information.
- *Personalisation:* A new demand has rapidly emerged for the adaptation of a huge amount of multilingual content now available on the Web, for individual needs. It needs "instant" localisation and personalisation to meet the demands of the users.

Professor Joseph van Genabith, Director of the new Centre said: "Localisation as an industrial process was developed in Ireland. We have a unique concentration of university- and industry-based research and development expertise in language technologies, machine translation, speech processing, digital content management and localisation. The research centre is going to pool that expertise and develop the next generation of language and content management technologies to support and develop the localisation industry."

**Please contact:**
Prof. Josef van Genabith
Director
National Centre for Language Technology NCLT
School of Computing, Dublin City University
Tel: +353 1 7005074
E-mail: josef@computing.dcu.ie

## Successful Start for CWI's New Building



Photo: CWI

The building activities for the new wing of CWI in Amsterdam formally started on November 20, 2007. Peter Nijkamp, chairman of the Governing Board of the Netherlands Organisation for Scientific Research - NWO), Cees de Visser, general director of NWO, Pieter Adriaans, chairman of the Governing Board of CWI, and Jan Karel Lenstra, general director of CWI, personally hammered the 'first pole' of the new wing into the ground, encouraged by more than 120 cheering visitors and a brass band. This building event completed "CWI in Bedrijf", CWI's annual day for relations (http://www.cwi.nl/cib/). The new wing and the renovation of the old building will be ready in 2009.

# Leibniz Prize for Holger Boche

Holger Boche, director of the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut HHI, has been awarded the prestigious €2.5 million Leibniz Prize 2008 for his work on the mobile radio networks of the future.



*© Fraunhofer HHI.*

*Professor Holger Boche, director of the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut HHI.*

Several key developments in the expansion of the mobile phone networks in recent years have been due to Holger Boche. On the basis of his theoretical work, Boche has advanced the understanding of complex mobile communication systems and at the same time has put his insights into practice in the technology used to standardise new mobile telephone systems. His research is of particular relevance to cross-layer optimisation, which increases the effectiveness and reliability of mobile phone networks. Boche has thus contributed significantly towards making it possible to use the existing mobile network frequencies with as few permanently installed transmitters and receivers to provide full coverage - a task which is not only a considerable scientific challenge but also has great economic potential.

Holger Boche studied information technology at the Technical University Dresden, supplementing his studies with mathematics from 1990 to 1992. Having obtained the respective diplomas, Boche went on to earn doctorates in both subjects, graduating with "summa cum laude". In 2002 – at the age of 32 – Boche accepted a professorship in mobile communications at the Technical University of Berlin. He has headed the Fraunhofer German-Sino Lab for Mobile Communications since 2003, and took on the joint directorship of the Fraunhofer Heinrich Hertz Institute together with Hans-Joachim Grallert in 2005.

The "Gottfried Wilhelm Leibniz Prize" is awarded by the Deutsche Forschungsgemeinschaft (German Research Foundation) every year since 1985 to scientists working in Germany. The DFG named eleven researchers, eight men and three women, as this year's winners of Germany's most prestigious research prize from amongst 158 candidates. Leibniz prize winners receive a grant of up to €2.5 million, which they can use flexibly according to the requirements of their research projects, over a period of up to seven years. Other winners include: Prof. Dr. Susanne Albers, Theoretical Computer Science, University of Freiburg; Prof. Dr. Martin Beneke, Theoretical Particle Physics, RWTH Aachen University; Dr. Elena Conti, Structural Biology, Max Planck Institute of Biochemistry, Martinsried, with Dr. Elisa Izaurralde, Cell Biology, Max Planck Institute for Developmental Biology, Tübingen; Prof. Dr. Stefan W. Hell, Biophysics, Max Planck Institute for Biophysical Chemistry, Göttingen; Prof. Dr. Klaus Kern, Physical Chemistry of the Materials, Max Planck Institute for Solid State Research, Stuttgart; Prof. Dr. Wolfgang Lück, Algebraic Topology, University of Münster; Prof. Dr. Jochen Mannhart, Experimental Solid State Physics, University of Augsburg.

**More information:**
http://www.dfg.de/en/news/press_releases/2007/press_release_2007_81.html

# Nokia Foundation Award to Pekka Abrahamsson

Nokia Foundation has granted its 2007 award to Pekka Abrahamsson for his merits in researching and developing global software development and processes. The €10,000 award was presented at the foundation's scholarship awards ceremony on 28 November 2007, when the foundation celebrated its 13th year of operation.

Pekka Abrahamsson is a research professor at VTT Technical Research Centre of Finland and currently head of the FLEXI-ITEA2 project, which aims to improve the competitiveness of Europe's software development by accelerating and adapting global product development and innovations. There are 37 organisations from eight European countries involved in the project. Abrahamsson is a pioneer of agile software development in Finland, and his team designed Mobile-D, the agile approach for mobile application development.

"Software is one of the key elements driving the Information, Communication and Telecommunications' role in the economy. At Nokia, the company is branching out into software and services in a unique way, compared with the mobility industry at large. Nokia is renewing and transforming business models, taking different approaches than in the past, as software increases in value," Jorma Ollila, Chairman of Nokia's Board of Directors, noted in his address at the Foundation's ceremony.

Nokia Foundation was formed in 1995. The Foundation supports the scientific development of information and telecommunications technologies and promotes education of the sector.

# ERCIM

**European Research Consortium
for Informatics and Mathematics**

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.

**W3C**  ERCIM is the European Host of the World Wide Web Consortium.

**AARIT**
Austrian Association for Research in IT
c/o Österreichische Computer Gesellschaft
Wollzeile 1-3, A-1010 Wien, Austria
http://www.aarit.at/

**CNR**
Consiglio Nazionale delle Ricerche, ISTI-CNR
Area della Ricerca CNR di Pisa,
Via G. Moruzzi 1, 56124 Pisa, Italy
http://www.isti.cnr.it/

**Crcim**
Czech Research Consortium
for Informatics and Mathematics
FI MU, Botanicka 68a, CZ-602 00 Brno, Czech Republic
http://www.utia.cas.cz/CRCIM/home.html

**CWI**
Centrum voor Wiskunde en Informatica
Kruislaan 413, NL-1098 SJ Amsterdam,
The Netherlands
http://www.cwi.nl/

**dana.m**
Danish Research Association for Informatics and Mathematics
c/o Aalborg University,
Selma Lagerlöfs Vej 300, 9220 Aalborg East, Denmark
http://www.danaim.dk/

**fonds national de la recherche**
Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
http://www.fnr.lu/

**FWO**
FWO
Egmontstraat 5
B-1000 Brussels, Belgium
http://www.fwo.be/

**FNRS**
FNRS
rue d'Egmont 5
B-1000 Brussels, Belgium
http://www.fnrs.be/

**FORTH**
Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
http://www.ics.forth.gr/

**Fraunhofer** Gruppe Informations- und Kommunikationstechnik
Fraunhofer ICT Group
Friedrichstr. 60
10117 Berlin, Germany
http://www.iuk.fraunhofer.de/

**INRIA**
Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
http://www.inria.fr/

**IUA** Irish Universities Association
Irish Universities Association
c/o School of Computing, Dublin City University
Glasnevin, Dublin 9, Ireland
http://ercim.computing.dcu.ie/

**NTNU**
Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and
Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

**PLERCIM**
Universitas Varsoviensis / Universitas Wratislaviensis
Polish Research Consortium for Informatics and Mathematics
Wydział Matematyki, Informatyki i Mechaniki
Uniwersytetu Warszawskiego
ul. Banacha 2, 02-097 Warszawa, Poland
http://www.plercim.pl/

**Science & Technology Facilities Council**
Science and Technology Facilities Council,
Rutherford Appleton Laboratory
Harwell Science and Innovation Campus
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
http://www.scitech.ac.uk/

**SpaRCIM**
Spanish Research Consortium for Informatics
and Mathematics c/o Esperanza Marcos, Rey Juan Carlos University,
C/ Tulipan s/n, 28933-Móstoles, Madrid, Spain,
http://www.sparcim.org/

**SICS** SWEDISH INSTITUTE OF COMPUTER SCIENCE
Swedish Institute of Computer Science
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

**SARIT**
Swiss Association for Research in Information Technology
c/o Professor Daniel Thalmann, EPFL-VRlab,
CH-1015 Lausanne, Switzerland
http://www.sarit.ch/

**MTA SZTAKI**
Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
http://www.sztaki.hu/

**VTT**
Technical Research Centre of Finland
PO Box 1000
FIN-02044 VTT, Finland
http://www.vtt.fi/

---

## 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
2004 route des Lucioles
BP 93
F-06902 Sophia Antipolis Cedex
Fax: +33 4 9238 5011
E-mail: office@ercim.org**

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

**I wish to subscribe to the**

☐ *printed edition*     ☐ *online edition (email required)*

*Name:*

*Organisation/Company:*

*Address:*

*Postal Code:*

*City:*

*Country*

*E-mail:*

---