

# ERCIM NEWS



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

Number 41

April 2000

## Special Theme: Web Technologies

**NEXT ISSUE:**  
July 2000  
Special Theme: Robotics

# CONTENTS

## KEYNOTE

- 3 Tim Berners Lee:  
The future Web and Europe – the Web and  
future Europe

## JOINT ERCIM ACTIONS

- 4 ERCIM launches the DELOS Network of  
Excellence for Digital Libraries *by Carol Peters*
- 5 Cor Baayen Award 2000 – Call for Candidates
- 5 Joint European-US Strategic Workshops  
*by Jean-Eric Pin*
- 6 ANFAS: Data Fusion for Flood Analysis and  
Decision Support *by Simon Lambert*

## SPECIAL THEME: WEB TECHNOLOGIES

- 7 Web Technologies – Introduction  
*by Robert Caillau*
- 8 Components for Data Intensive XML  
Applications *by Peter Fankhauser, Gerald Huck,  
and Ingo Macherius*
- 9 EXMO: Computer mediated Exchange of  
Structured Knowledge *by Jérôme Euzenat*
- 10 Escribe: Embedded Structured Content  
Representation in Repositories  
*by Jérôme Euzenat*
- 11 Workflow Demonstrator *by Damian Mac Randal*
- 12 Promoting XML in Italy *by Laura Abba*
- 13 The XMLC Browser *by Fabio Vitali*
- 14 Using XML for Data Management *by Victoria  
Marshall, Brian Matthews and Kevin O'Neill*
- 15 Corporate Knowledge Management through  
Intranet/Internet *by Rose Dieng*
- 16 Multilingual Metadata to access Social Science  
Databy *Brian Matthews and Michael Wilson*
- 17 ICS-VRP: a Tool for Parsing and Validating RDF  
Metadata & Schemas  
*by Karsten Tolle and Vassilis Christophides*
- 18 XHTML 1.0 for Device-Independent Web Access  
*by Steven Pemberton*
- 19 Amaya: a Testbed for Web Protocols and  
Formats *by Irène Vatton*
- 20 XML Accessibility Guidelines  
*by Daniel Dardailler*
- 21 The W3C in 7 Points
- 23 Web Images and Blind Users  
*by Zdeněk Mikovec, Martin Klíma,  
Dušan Pavlica and Pavel Slavík*
- 24 Adaptive Browsing: Supporting All Users on the  
Web *by Constantine Stephanidis*
- 25 Vector Graphics *by David Duce*
- 27 More Quality for Services in Internet  
*by Juha Koivisto*

- 28 The Social Web Cockpit: A Tool to Support  
Knowledge Sharing Communities  
*by Wolfgang Gräther and Wolfgang Prinz*
- 30 Eliminating the I/O bottleneck from World Wide  
Web Proxies *by Evangelos P. Markatos*
- 31 WEBPATH – a 3-D Browsing History  
Visualisation *by Emmanuel Frécon*
- 32 The sView System  
*by Fredrik Espinoza and Markus Bylund*
- 34 Building a multinational R&D proposal through  
the Web *by Andrew Hallan, Alexander M.  
Samsonov and Wolfgang Ziegler*
- 35 The Collaborative Web  
*by László Kovács and András Micsik*
- 37 Separation of Concerns in Multimedia Services  
*by Olle Olsson*
- 38 Increasing Security in Virtual Enterprise  
Communication for Mobile Environment  
*by István Mezgár, Tamás Szabó  
and Zsolt Kerecsen*
- 40 Creating Hypertextual Dossiers in XML for the  
Italian Parliament  
*by Andrea Marchetti and Patrizia Andronico*
- 41 Providing the Context for Business Decisions  
*by Brian Matthews*
- 42 Virtual Physics and Astronomy  
*by Detlef Skaley and Udo Zlender*
- 44 CommonGIS – Common Access to  
Geographically Referenced Data *by Hans Voss,  
Natalia Andrienko and Gennady Andrienko*

## RESEARCH AND DEVELOPMENT

- 46 StraDiWare: Linking Business and IT Models  
using Graph Theory *by Tony Conway*
- 47 Computer Modelling in Geosciences  
*by Stanislav Saic and Josef Jeřek*
- 48 Efficient Use of Networks through equal  
Bandwidths *by Annette Kik and Henk Nieland*
- 49 Interactions in Virtual Environments  
*by Lakshmi Sastry and David Boyd*
- 50 Alternative Processing of Verbal Data  
*by Milan Mareš*
- 51 Towards the Distributed Virtual Education and  
Training Environment *by Igor Nikitin, Valery  
Burkin and Stanislav Klimenko*
- 52 Wavelets improve Information on Earthquakes  
*by Henk Nieland*

## EVENTS

- 53 SOFSEM'99  
*by Jan Pavelka*
- 54 Announcements

## IN BRIEF

## The future Web and Europe – the Web and future Europe



Photo: A. Eberhard/NASA

**Tim Berners-Lee is director of the W3C consortium, and a Principal Research Scientist at MIT. He invented the Web in 1989, and recently wrote the book 'Weaving the Web'.**

The World Wide Web Consortium (W3C) has gained a lot from ERCIM. The strong overlap between the membership of ERCIM and the W3C offices has made it very easy to set up the European offices; it has helped W3C to be connected into Europe. Concerning Europe, a frequently asked question is: Why did the Web take off more quickly in the US, when it was invented in Switzerland? There are many reasons. The Internet was far more deployed in the US than in Europe. When I invented the Web in 1989, the Internet had just got through to CERN. Many US universities were connected on the Internet, which enhanced a rapid spread of the Web across the USA. The American entrepreneurial spirit may also have something to do with it. There are many start-ups in America. The European ethos to doing things is discussing them first and then doing them. The huge difference in diversity between the US and Europe is a third factor. Europe does not have the same model language block. There is an incentive to put up a website in America, because so many people understand English.

Diversity is very valuable, and Europe should use this. We get a much richer set of ideas with boundaries which allow people to think independently. At the same time, you need balance between diversity and homogeneity. ERCIM is part of that: it builds bridges between different cultures and allows technical ideas to move freely within the academic community and across different borders.

There are several directions in which Europe could now go. American industry seem to move ahead faster, but Europe could still very well be ahead of the next wave, because it has a lot of the know-how. There could be an excellent opportunity for Europe to address issues such as intellectual property rights, negotiation, and privacy. The Web is not yet done. The Web as you see it now, is just an information dissemination system for most people. I'd always wanted it to be a collaborative system, in which creating links, creating hypertext, is just as simple and intuitive as browsing currently is.

ERCIM provides a way of moving information and facilities between different people in different countries. It coordinates and facilitates, and the Web is a tool for doing that. Europe should think carefully about how to use the Web, in order to build a new society on top of that. Europe is going through a large change, finding out how to construct a society of many nations. The Web can give Europe more of a varied set of connections, which will make it very resilient and able to remain stable but at the same time to move very fast. This interesting general question I leave with you to think it over.

## ERCIM launches the DELOS Network of Excellence for Digital Libraries

by Carol Peters

23 participants attended the kick-off meeting of the new Network of Excellence for Digital Libraries, organised by ERCIM in Sophia Antipolis, 16-17 February 2000. The Network aims at continuing and

extending the activities of the DELOS Working Group on Digital Libraries, and intends to promote European research in the digital library domain by providing a set of 'services' for the digital library community.

The aim of the DELOS Network of Excellence for Digital Libraries (DLs) is to provide an open context in which an international research agenda for future research activities in the digital libraries domain can be developed and continuously updated. The Network will constitute a reference point for all DL projects funded by the 5th FP-IST, stimulating the exchange of experiences and know-how in this multidisciplinary domain, and will also establish close contacts with relevant application communities. It will make testbeds available, facilitate their interoperability, and provide mechanisms for the evaluation of models, techniques, and approaches, and the exchange of open-source software components. It will also contribute to the definition of relevant standards. Appropriate take-up models will be studied and tested to facilitate the exploitation of new DL technologies in the industrial/commercial environment. A dense network of links with the international DL research community will be created.

The activities of the Network will be organised under Forums and will be open to the European DL research world and the relevant application communities (electronic publishing, libraries, cultural heritage, archives, etc.).

**Digital Library Research Forum.** This activity will be aimed at stimulating the advancement of DL-related research and encouraging scientific collaboration between European teams working in this domain through the organisation of: (i) thematic workshops, (ii) the European Conferences on Digital Libraries (ECDL), (iii) an international research exchange programme.

**Digital Library Evaluation Forum.** DL evaluation implies defining new metrics for components and new combinations of components in order to measure performance aspects related to DL interoperability, metadata and resource discovery, multilingual access, etc. DELOS will promote the building and operation of publicly available reusable

evaluation infrastructures: (i) a digital library test suite and (ii) a cross-language system evaluation infrastructure.

**Standardization Forum.** Working Groups will be set up to (i) examine specific clusters of relevant standards; (ii) effect the transfer of research results into implementation; (iii) disseminate the results through workshop proceedings, reports, Web-based observatories, etc.

**Training and Technology Transfer Forum.** This forum is aimed at providing training possibilities and demonstrating the effectiveness and benefits of DL technologies within European industrial environments. Summer schools will be organised on key issues in the DL domain.

**International Cooperation Forum.** Collaborations will be established with those countries for which a scientific cooperation agreement is defined under the 5FP. In particular with the US (through NSF), Russia (through the Russian Foundation for Basic Research), Mediterranean countries, and Japan.



Participants in the kick-off meeting.

The Network of Excellence has 28 member organisations: the academic community is represented by leading European research teams working on key topics in the digital library area from eight ERCIM institutions (CNR-IEI, CWI, ETHZ, FORTH, GMD, INRIA, SICS, MTA SZTAKI) and from nine other leading European academic organizations (Universities of Athens, Dortmund, and Mannheim, Norwegian University of Science and Technology, Technical University of Crete, Bavarian Research Centre for Knowledge-Based Systems, Milan Polytechnic, Austrian Academy of Sciences, Social Sciences Information Centre - Bonn. The application community

is strongly represented by members from libraries, (National Library of Portugal, Italian National Institute for Union Catalog), electronic publishing (Springer and Elsevier), education (Arts and Humanities Data Service - UK), archives (Italian Ministry of Cultural Heritage), and broadcasting (BBC – UK). There are also two IT industrial members: Eurospider Information Technology, Switzerland, and Intracom, Greece). As one of the guiding principles of the Network will be that of “openness”, delegates from other research groups or interested communities will be invited to participate in activities of the Network as appropriate.

An international advisory board has been set up to approve and evaluate the programme of activities planned by Network. Members of the Board are:

Stephen M. Griffin, Chris Rusbridge, Makoto Nagao, Yuri Hohlov, Rudi Schmiede, representing US, UK, Japanese, Russian, and German national DL programmes, respectively. The Board will meet annually but will be kept in continual contact with the work of the Network through regular reports from the scientific coordinator: Costantino Thanos, IEI-CNR, Pisa.

**Links:**

**DELOS website:**  
<http://www.iei.pi.cnr.it/DELOS/>

**Please contact:**

**Costantino Thanos – IEI-CNR**  
**Director – DELOS Network of Excellence**  
**Tel: +39 050 315 2910**  
**E-mail: [thanos@iei.pi.cnr.it](mailto:thanos@iei.pi.cnr.it)**

## Joint European-US Strategic Workshops

by Jean-Eric Pin

**The European Commission and the US National Science Foundation will be organising a series of strategic research workshops to identify key research challenges in Information Technologies.**

The strategic workshops will be organised jointly under the auspices of the European Commission's IST-FET programme and the US National Science Foundation CISE division. Focus is on key emerging areas of mutual interest on which a world-class series of workshops would be relevant. One outcome of each workshop will be recommendations for future joint European-US activities in the domains covered by the workshop.

ERCIM has been chosen to lead the European part of this activity and to solicit proposals from the European research community. ERCIM is interested in bottom-up ideas for high-level workshops from the European scientific community.

This activity is supervised by a Strategic Workshop Review Committee (SWRC) composed of high-level independent scientists from Europe and the US. The committee will select the themes by criteria such as:

- long-term/high risk nature of the research involved, justifying risk sharing at an international level
- high potential payoffs both in the EU and the US that make up for the long-term/high risk nature of research
- existence of sufficient scientific and technological bases in both the US and the EU to entail balanced research efforts.

Scientists interested in this co-operation can submit their suggestions by a short online questionnaire available at the ERCIM website.

**Links:**

**EU-NSF Joint Strategic Workshops Initiative:** <http://www.ercim.org/EU-NSF/>

**Please contact:**

**Jean-Eric Pin – ERCIM Office**  
**Tel: +33 4 9238 5010**  
**E-mail: [jean-eric.pin@ercim.org](mailto:jean-eric.pin@ercim.org)**

### CALL FOR CANDIDATES

## Cor Baayen Award 2000

The Cor Baayen Award for the most promising researcher in computer science and applied mathematics was created in 1995 to honour the first ERCIM President. Any young researcher from one of the ‘ERCIM countries’ (Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Norway, Slovakia, Sweden, Switzerland, The Netherlands or the UK) can be nominated. The award consists of a cheque for 5000 EURO. The selected fellow will be invited to the ERCIM meetings in autumn.

### Rules for Nomination

- nominations for each country are made by the corresponding ERCIM Executive Committee member (national contact); those who wish to propose a candidate must contact the ERCIM Executive Committee member for their country (see <http://www.ercim.org/contacts/execom/execom.html>).
- nominees must have carried out their work in one of the fourteen ‘ERCIM countries’
- on the date of the nomination, candidates must have been awarded their PhD (or equivalent) not older than 2 years
- Each institute can nominate up to two persons from its country
- a person can only be nominated once
- the selection of the Cor Baayen Fellow is the responsibility of the ERCIM Executive Committee.

### Deadline

Nominations are to be received by the national contacts by 30 June 2000.

**Links:**

**Cor Baayen Award:**  
<http://www.ercim.org/activity/cor-baayen.html>

**Please contact:**

**Lubos Brim – CRCIM**  
**Tel: +420 5 4151 2323**  
**E-mail: [lubos.brim@ercim.org](mailto:lubos.brim@ercim.org)**

# ANFAS: Data Fusion for Flood Analysis and Decision Support

by Simon Lambert

Several ERCIM partners are working together in a new project under the European Information Society Technologies programme. The ANFAS project aims to

develop a decision support system for flood prevention and protection, integrating the most advanced techniques in data processing and management.

Large-scale flooding is a serious problem in many parts of the world. In Asia, countries such as China and Bangladesh regularly suffer, while in Europe many countries are at risk. There is danger of loss of life and serious damage to the economies of the areas affected.

For this reason, the ANFAS project has been set up to include partners from both the European Union, other European countries, and China. The partners bring together expertise in soil sciences, remote sensing, scientific computing, computer vision, internet technology, geographic information systems, knowledge-based systems and other areas. The ERCIM partners are:

- ERCIM itself in the project management role
- FORTH, with expertise in Geographic Information Systems
- SRCIM: Slovak Academy of Sciences, Institute of Informatics, one of the links with the end-users and also with expertise in geophysical flow modelling and high-performance computing
- CLRC, bringing expertise in knowledge-based assistance and cooperative capabilities for information systems
- INRIA, with expertise in geophysical flow modelling and high-performance computing.

The project team also includes three other European partners (Matra System & Information, the French geological survey and ore exploration - BRGM, Reading University) and three Chinese institutes (Institute of Automation (IOA), Institute of Atmospheric Physics (IAP), Institute of Remote Sensing Applications (IRSA), of the Chinese Academy of Sciences).

When there is a risk of flooding, decision makers have to decide what are the most appropriate actions to take: evacuation of the population, reinforcement of dikes, intentional breaking of dikes, etc. The aim



Floods regularly cause life loss and damage like here in France in 1999. The aim of the ANFAS project is to develop a simulation and prevention tool to limit flood damage.

of the ANFAS project is to develop a simulation and prevention tool to help decision-makers in taking decisions that will limit flood damage.

The functionality of the ANFAS system will be:

- to perform flood simulation, simulating water flow propagation based on scenarios. The user will be able to interact with the scene that models the real site in order to add/remove dikes or other human constructions
- to assess flood damage. Flood assessment can be done either using the simulation results or using the remote sensed images of a real flood event
- at a prospective level, to analyse floodplain morphology including changes in riverbed and subsurface properties due to repeated floods at a given site.

The output of the system will be:

- visualisation: consisting of two-dimensional maps of the extent of the

flood. If simulation is performed, time sequence maps tracing the propagation of the simulated flood will be provided

- direct impact assessment, evaluation of the total affected area
- regional and statistical analyses over the affected area
- simple socio-economic assessment of damage directly caused by flooding.

The final system will be an immensely valuable tool in planning national and regional responses to the threat of flooding, and will be adaptable to other areas.

#### Links:

Anfas website:  
<http://www.ercim.org/ANFAS/>

#### Please contact:

Bruno Le Dantec – ERCIM Office  
 Tel: +33 4 9238 5010  
 E-mail: [bruno.le\\_dantec@ercim.org](mailto:bruno.le_dantec@ercim.org)

# Web Technologies – Introduction

by Robert Caillau

**The current crop of articles in this issue's special theme is all about the Web. Or is it? I read the articles with interest, sometimes a frown, sometimes a smile and sometimes a sigh.**

The most interesting observation for me was to see how much XML has become a tool of choice for experimenting. This is good news. When we were working on spreading the Web, HTML was clearly a second priority, and it remained in a pitiful state from 1990 to 1993. It proved impossible to get people at large to understand that SGML with its DTDs was a better technology, so we had to wait until there had been enough corruptions and style problems. That made people look for any good proposal, and XML was one. So that was a smile.

But the other thing I observed is that Java is the unchallenged choice for doing the actual horsework: getting some semantics into it all. Somehow semantics remains an untrappable ghost. When I studied programming languages, in the days when computers would still sit up, I always got confused as to where others defined the boundary between syntax and semantics. Notation does not solve the problem. Only execution does. The prime example of how real programmers define real semantics is C, that ugliest of all ugly systems. But as with many monopolies, one of C's problems is that it works. Java should give us portable semantics. But does it? It's all in the intractable classes. Yet, there is a lot of good stuff, and certainly programming feels orders of magnitude better in Java than in C (there are of course still those curly brackets). I always believed that ultimately the Web needed a programming language in the Turing machine sense. This was sometimes frustrating to communicate, and many, including friends, went off into the blue yonder of declarative stuff, often so esoteric that no sane person would spend her time on it. But, be it implicitly, we now have a programming language, unfortunately a proprietary one. So that was a sigh.

When I first heard of the Dublin Core metadata, I got enthusiastic and asked someone to prepare a talk on the state of metadata. An interesting discussion ensued, though I was quite annoyed at how people thought it could all be very simple. So at one subsequent meeting I showed them the lid of a pot of mayonnaise as sold by the local supermarket. It had on it the date of fabrication, the sell-by date and the best-before date. How could the Dublin Core people think they could do it with a single date, if just the local grocery store had used three for decades? So that was a frown. But I am greatly relieved to see how much influence RDF has now obtained.

Another idea (there I go again...) that should have been with the Web from the start was vector graphics. Before GIF, we (I) used PostScript for Web drawings. This was certainly 2D vector graphics, but PostScript was to some a word you could not say without conjuring up the devil, lightning and sulphurous smells. Then, to my great disappointment, GIF took over, but look,

SVG is on the horizon, it is non-proprietary and we may yet see it replace PostScript. So there is another smile.

RDF, XML and SVG all at some level are re-inventions of wheels that have been turning in many places for a long time. It is necessary to re-invent once in a while, because only that way can we integrate, improve, get the youngsters enthusiastic, and become non-proprietary. Smile! (SMIL?)

I discovered a number of cats. Cats are independent agents over which you have no control. If they like you, they may spend some time sitting on your lap and purr, giving you an impression of cosy well-being. There are several articles on systems that are supposed to do things for you. But I see cats (mind you, I'm owned by a cat myself).

Finally, there are some very interesting reports on techniques for improving efficiency and for applying the Web in social interaction. The range of this collection is wide, if you are an eclectic reader as I am, you will find a lot of stimulating stuff here!

**Robert Cailliau – Web Communications & Public Education,  
Education & Technology Transfer Division  
CERN, European Organization for Nuclear Research**

# Components for Data Intensive XML Applications

by Peter Fankhauser, Gerald Huck and Ingo Macherius

XML-based applications blur the lines between Internet, publishing and database technology. While classic SGML was focused on presentation oriented publishing applications, typical XML deployments in e-commerce bias toward message oriented middleware, transactions

and data exchange scenarios. At GMD Institute for Integrated Publication and Information Systems (IPSI), a component-based framework has been designed and implemented to allow for the rapid development of data intensive, XML based applications.

The eXtensible Markup Language (XML) is the next generation data format for structured information interchange on the World Wide Web. Under the auspices of the World Wide Web Consortium (W3C), XML has grown into a family of standards integrating key technologies from three previously independent domains: documents, databases, and the Internet. This powerful mix is a strategic component of the rapidly growing 'dot com' industry.

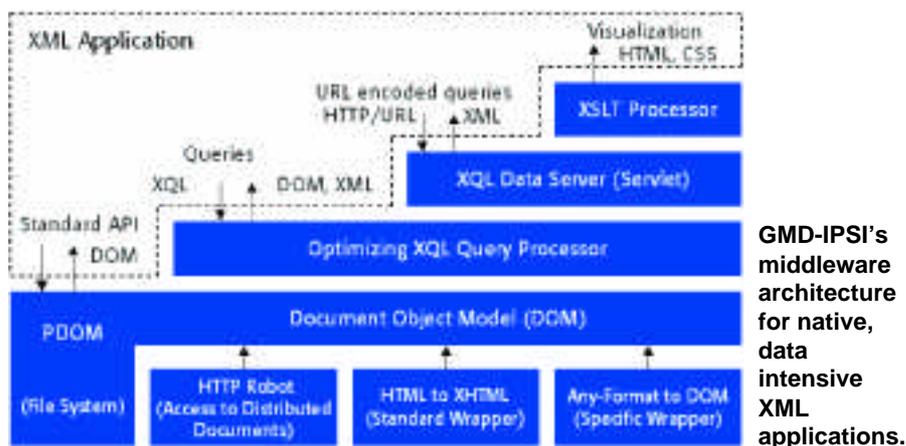
## A Native Approach to XML Processing

Many of today's XML technologies originate from document processing, where scalability was not considered important. In e-commerce, however, scalability is essential. The use of XML with database technologies overcomes this problem, but raises others, such as data model and query paradigm mismatches. XML has a semistructured nature, which is incompatible with the flattened structure of relational tables, and is not easily decomposable into objects. Queries do not capture the expressiveness of XML.

We designed middleware components, which overcome the mismatches without sacrificing the simplicity of XML. They provide two important DBMS capabilities: declarative queries and transaction safe persistence. Thus the components can cope with the weakly structured, high volume XML data typically generated by wrappers from legacy data sources such as HTML pages.

### Component 1: The Persistent DOM

The Document Object Model (DOM) is a platform- and language-neutral interface for XML standardized by the W3C and widely used throughout the industry. It provides a standard set of objects for representing XML documents, a standard model of how these objects can be combined, and a standard interface for accessing and manipulating them. The



PDOM (Persistent Document Object Model) is an object manager, which transparently maps standardized W3C-DOM API method calls into operations on binary files, enabling processing of XML documents far beyond main memory limits. Unlike traditional databases, the DOM does not require a schema and thus allows skipping the design intensive setup phase typical of traditional DBMS.

The implementation is scalable and achieves a throughput of several MB of XML data per second on a standard PC. PDOM is lightweight, in that the code size is relatively small as well as programming effort is less compared to DBMS based solutions.

### Component 2: The XQL Query Processor

The Extensible Query Language (XQL) is a declarative, path-oriented query language for XML. It includes most operations familiar to SQL, eg selection, restructuring, joins, and views, and handles the semi-structured nature of XML. Introduced first at W3C's conference on XML query languages in 1998, XQL has since been implemented by several large IT-vendors. IPSI's query processor implements the complete XQL

proposal, and augments it with extensions to cross-document joins and restructuring of results. Its robust and efficient mix of algebraic and physical query optimization techniques yields superior performance. The processor can also be used on top of any W3C compliant DOM implementation, including the PDOM.

### Component 3: The XML-Broker Data Server

Both the XQL processor and the PDOM were integrated into an XML data server, making their functionality accessible through the HTTP protocol. Our extensive experience showed the importance of diversity in interfaces. Thus we support the low-level DOM API, string-based queries, URL with embedded queries, and means to post-process query results with XSLT (XSL Transformations). Complicated information processing tasks on XML data can be concatenated into pipelines, encoded in standard URL syntax – ready to be bookmarked and reused.

### Applications: From Molecules to Markets

The first deployment of our components was the nationally funded RELIMO project, whose goal is the integration of data sources of interest to drug designers.

This still evolving technology was chosen by the IST funded OPELIX and eBroker projects as a platform for B2B e-commerce applications. The software package is also being distributed on the Web. So far, more than one thousand have been downloaded. This popularity made its commercialisation viable, and resulted in a product called the Infonbyte XQL Suite.

#### Research meets Business

The components discussed in this article are part of an ambitious project at IPSI,

the GMD's XML Competence Center. This center has already become a focal point for cooperation between industry, research partners and standardization bodies such as W3C. Our state-of-the-art know-how and proven research background in database technology, document management, publishing, information retrieval, and graphical user interfaces make GMD-IPSI a natural partner for advanced XML information management issues, and their application in the evolving digital economy.

#### Links:

XML Competence Center:  
<http://xml.darmstadt.gmd.de/>  
 XQL and PDOM download:  
<http://xml.darmstadt.gmd.de/xql/>

#### Please contact:

Peter Fankhauser – GMD-IPSI  
 Tel: +49 6151 869 939  
 E-mail: [xmlcc@darmstadt.gmd.de](mailto:xmlcc@darmstadt.gmd.de)

## EXMO: Computer mediated Exchange of Structured Knowledge

by Jérôme Euzenat

**The World Wide Web has generalised document exchange on a world-wide basis. XML promises to do the same thing for the exchange, manipulation and transformation of structured data. However, if this provides interoperability, it does not help for intelligibility.**

Exmo is an INRIA action which studies the communication of formalised knowledge among people mediated by computers. In the communication process, the computer can add value to its medium and memory functions by performing advanced operations such as formatting, filtering, categorising, consistency checking, or generalising. The goal of Exmo is the development of theoretical and software tools for helping the organisation, manipulation, presentation and combination of structured knowledge chunks though allowing communication between people. The knowledge is represented in a formally defined language (which can range from XML - structured, general but without semantics - to knowledge representation languages - structured, semantically characterised, but with a narrow scope of expression).

In order to warrant the suitability and intelligibility of knowledge for the users, this requires to develop an abstract understanding of representations and the transformations applied to them. The research of the Exmo action is focussed in two directions. The transformation topic accounts for the modification of knowledge during the communication while the communication topic accounts

for the preservation of intelligibility through transformations.

A transformation is a computational way of generating representations from other representations (not necessarily in the same language). Transformations are used for putting representations together or generating specific representations for a particular need. The goal of the transformation topic is the elaboration of a 'general theory of transformations' based on the properties that transformations satisfy (rather than on the representations or the transformation themselves). Such properties include content or structure preservation, source traceability, or conversely, confidentiality. For instance, it can be useful to define a transformation which delivers a documentation to a customer but which hides some information from the initial source (eg, if a design document includes cost study, it is not advised to communicate this information to subcontractors). In the context of collaborative work, it is useful to establish the properties of such a filter in order to know where to implement it in the organisation. Conversely, when the elaboration of a representation is a collaborative and continuous process (in

the concurrent engineering framework), it is necessary to apply treatments which do not challenge the current stage of development. It is thus relevant to warrant the preservation of content.

One outcome of this work is the capability to decide, given an organisation composing several transformations, and the properties of these transformations, the properties satisfied by the system as a whole. This should be a precious tool for the future information system architects.

More generally, given a set of transformation types and a set of transformation composition operators, one goal is to establish if a property can be decided for an instance of this. Exmo does not aim at helping the implementation of transformations or at inferring properties from implementations. However, some (generally simple) transformation languages implement only one specific interaction type (eg filter) and thus our work is connected to transformation languages in that way.

The communication topic aims at contributing to the proper understanding of the knowledge by people. Formal semantics developed for knowledge

representation systems is suited to the use of a system by one user because it is expected that (s)he interprets in a coherent way the terms (identifiers) used. However, when several users communicate, this understanding becomes problematic. Work developed in the field of consensual ontology construction could help to solve the problem of term interpretation. However, other problems are involved. For instance, a user can express knowledge under the form of class hierarchies and first order clauses and then communicate it by using an interoperability language. But if this last language expresses the whole knowledge with clauses (though preserving the semantics of the assertions), the initial user will hardly recognise (and hardly understand) the semantically equivalent

result. Hence, when a transformation operates a translation between formal languages, good understanding cannot be ensured by meaning preservation. A semiotic treatment must be applied for ensuring the meaning reconstruction. This treatment comes as a complement to the sheer semantic treatment used in the knowledge representation area.

In order to contribute to the solution of this last problem, we investigate the embedding in XML of the semantics of the language used. Thus, the form can be preserved as much as possible (form is relevant to human understanding) and interoperability is ensured through the availability of the semantics (correct computer treatment mainly depends on semantics).

One step further on this line leads to the consideration, together with structure and semantics, of interpretation policies.

These concepts are applied to technical memories and representation of web document content. However, it will spread to all the aspects of information systems in the future (especially in the context of the generalisation of computer supported collaborative work and the development of the so-called 'semantic web').

**Links:**

<http://www.inrialpes.fr/exmo/>

**Please contact:**

Jérôme Euzenat – INRIA  
Tel: +33 4 7661 5366  
E-mail: [Jerome.Euzenat@inrialpes.fr](mailto:Jerome.Euzenat@inrialpes.fr)

## Escrire: Embedded Structured Content Representation in Repositories

by Jérôme Euzenat, Rose Dieng and Amedeo Napoli

**Content representation seems unavoidable in some areas of the future web. Although there are many candidates as content representation languages, their respective merits in this context are not yet known.**

An Intranet or, more generally, the use of the Internet technology, is an opportunity for companies, to publish and share knowledge often difficult to reach in documentary form. The numerical and digitised documents can be made available in a standard and transparent way to all the users concerned. The ambition, in the long term, is to produce knowledge servers allowing the search and the handling of the corporate resources. However, the limits of this approach appear quickly: the organisation and maintenance of the sites appears an expensive task and full-text search is not very effective.

The representation of the content of documents is becoming a necessity. Content representation allows to manipulate content, to make search by analogy, by specialisation, similarity, etc. XML enables to insert content representation (through RDF) within the documents (in XHTML or other XML format) and knowledge representation formalisms are good candidates for representing content.

ESCRIRE is a coordinated action of three INRIA teams (Acacia, Exmo, Orpailleur). One of its first objectives is to propose an implementation model for this. But there are various knowledge representation formalisms and their respective qualities (compared to each other) are not exactly known. The main goal of ESCRIRE thus consists of comparing three types of knowledge representation formalisms (conceptual graphs, object-based knowledge representations and description logics) from the standpoint of the representation and the handling of document content. That will enable to highlight the desired properties and to evaluate qualitatively and quantitatively the performances of the implementation formalisms. Beyond the better knowledge of the techniques implied by the various projects, this work will advance the state of the art.

A set of documents and a set of queries have been selected in a coordinated way. The efficiency of the respective formalisms will be assessed with regard to these

queries. Each team specifies the integration of the formalism in XML and develops a query evaluation strategy depending on the considered formalism. The formalisms together with full-text search will be evaluated along a predefined protocol for assessing qualitative criteria (eg query expressivity, legibility) and quantitative ones (eg precision, recall). This evaluation will provide a precise analysis of the strengths and weaknesses of each knowledge representation formalism of content representation.

**Links:**

<http://www.inrialpes.fr/exmo/cooperation/escrrire/>

**Please contact:**

Jérôme Euzenat – INRIA  
Tel: +33 4 7661 5366  
E-mail: [Jerome.Euzenat@inrialpes.fr](mailto:Jerome.Euzenat@inrialpes.fr)

# Workflow Demonstrator

by Damian Mac Randal

As part of the World Wide Web Consortium's Leverage Action project (W3C-LA), several small demonstrators were developed to promote European adoption of the Web and W3C protocols. At RAL, a pilot workflow

system was developed to explore XML/RDF and their interaction with other client side protocols. The demonstrator focuses on delivery of the workflow steps from multiple servers to remote actors.

Business processes are now recognized as being key to thinking about organizations, their effectiveness, and their ability to cope with change. IT support for business process is provided by workflow management technology: typically, a workflow system will provide support for each of the activities in a process by offering the necessary work (and information needed to complete it) to those staff with the correct attributes to undertake that work. A server retains the information on the process specifications and the state of each process which is currently active. Major issues in workflow management include the portability, coping with the geographical distribution of users, relationships with organizational Intranets, and consistency of access.

All workflow management issues are addressable by Web server / browser

technologies, by building on the W3C developments in HTML, DOM, XML and RDF (Metadata) in particular, together with the portability features offered by Java and the servlet API. Workflow is seen as a challenging application area that will explore many of the emerging Web technologies, such as push technology, security, client-side processing, etc.

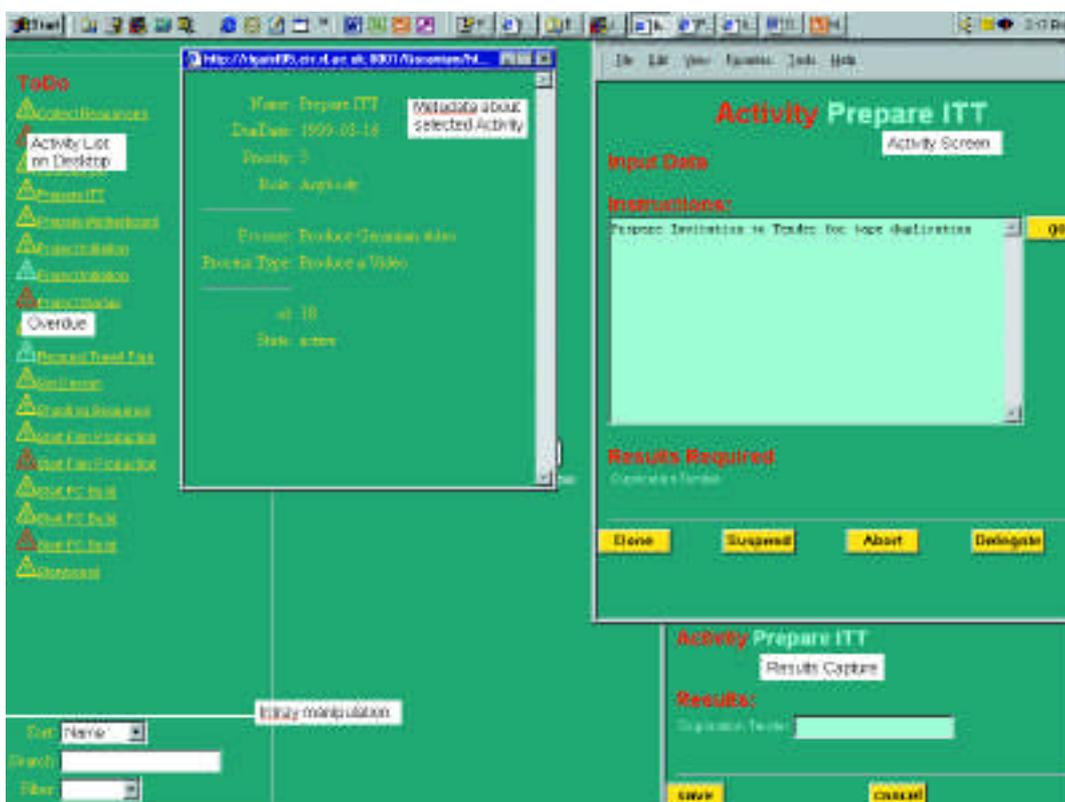
The W3C Workflow demonstrator focuses on the delivery of activity information to the appropriate actors as required by the process, in a timely, secure and useful manner while minimizing network and server loads.

Activity information consists of instructions as to what has to be done, and 'parameters', that is data items required to carry out the activity and the data generated by the activity. Some of this

can be sensitive, restricted to specific actors identified by the process. There is also a significant amount of content about the activity, such as the DueDate, containing process, status, etc. This metadata is useful to the actor independently of the activity itself. Since most actors will have many activities allocated to them, some sort of Intray is required, listing outstanding activities and allowing selection of the next. Access to metadata, such as DueDate, is clearly essential to prioritize selections. Equally clearly, the Intray has to be kept up-to-date.

The demonstrator uses a 3-tier architecture:

- a database holding the process and activity data, state information and actor information.
- Java servlets (on a Web server) providing the process 'engine'



Screenshot of the demonstrator Intraday and Activity pages.

managing the delivery of activities to actors and the process flow

- scripts on the actor's Web browser providing local processing of the Inray and Activity pages.

### Intrays

An Inray is a dedicated page (or desktop) displaying an up-to-date list of the activities allocated to the actor. The list provided by the Web server uses RDF to record the metadata about each of the activities. The Inray page, or more precisely, scripts in the page then use this RDF to display a scrollable list of the activities, highlighting overdue items. Further scripts allow the actor to sort and filter the list without going back to the

Web server. By stopping the mouse pointer over a list item, the full metadata for that activity can be displayed in a popup window. Finally, so-called push technology (actually a scheduled check/refresh) is used to keep the list up-to-date.

### Activity Pages

Clicking on an activity in the inray indicates to the workflow engine that the actor intends to work on that activity. A new Activity page is created using an XML description of the activity sent by the Web server. (The Activities metadata is also sent as RDF embedded in the XML). The Activity page displays the Activity instructions and data, and if

appropriate connects the 'Go' button to the required application. When the actor is finished, clicking the 'done' button causes the results form to popup. Clicking 'send' on this sends the results back in XML to the workflow engine and closes the popup and activity window. The workflow engine saves the results, marks the activity as completed, identifies the next activity and 'pushes' it to the appropriate Intrays.

#### Please contact:

**Damian Mac Randal – CLRC**  
Tel: +44 1235 44 5403  
E-mail: damian@inf.rl.ac.uk

## Promoting XML in Italy

by Laura Abba

**During 1999, the Institute for Telematic Applications of the Italian National Research Council (IAT-CNR) promoted the creation of XML-Italy, an open working group similar in constitution to the Internet technical**

**groups. One of the objectives of XML-Italy is to promote R&D collaborative activities between Italian research and industry.**

XML-Italy's primary goal is to coordinate existing national expertise in order to spread XML know-how in an organised manner. Technicians, researchers, operators and vendors interested in XML have joined the group, and it has become a point of reference for the Italian scientific and industrial community. The group facilitates the approach to the new technologies, and thus contributes to a more efficient and safer management of information on the net.

The Italian XML Group has four main goals:

- to disseminate know-how on a national basis through the organisation of technical transfer activities
- to facilitate the exchange of ideas via e-mail and meetings
- to create an XML information bank or reference point both for developers and for users
- to demonstrate the new technologies through the development of prototypes and simple applications.

Participation in the activities of the group is open. The work is carried out during

the meetings and through the 'group communication' system via electronic mail at the address: xml@xml.it. More than 230 public and private Italian organisations subscribe to this list and the average traffic is about 4 e-mails per day. Technical problems – such as XML parsers, CR/LF treaty in XML, applicative co-operation using XML instead of CORBA, parameter transfer in xslt, WAP and XML files, MS-WORD document representation in XML, etc - are debated. Requests for assistance and information on markup languages, particularly XML, for publishing, data exchange, and the visualisation of complex and structured documents on the net are frequently received.

Many of the most relevant research activities and applications of XML technologies in Italy were presented last June in Bologna, at the first XML-Italy workshop (<http://www.xml.it/bologna.html>). More than 200 researchers attended this event. The next XML-Italy meeting will be held in Pisa, at the Research Area of the National Research Council, 9-10 May 2000.

In addition to XML-Italy, IAT-CNR is also currently responsible for several research programmes on technologies for the processing, exchange and protection of information. Languages and models for the representation of the information content of data files are being studied for use by the next generation browsers and web servers, in order to facilitate access to distributed data and surfing on the net. The considerable changes that have been made in the architectures of information systems, following the rapid evolution of the de facto standards, are also being investigated with particular attention. The technologies used for this line of research are WEB, JAVA, and XML.

#### Links:

Activities of XML-Italy: <http://www.xml.it/>

#### Please contact:

**Laura Abba – IAT-CNR**  
Tel: +39 050 315 2633  
E-mail: Laura.Abba@iat.cnr.it

# The XMLC Browser

by Fabio Vitali

**XMLC is a modular browser for XML documents. It is the result of a project born initially to display generic extensions of HTML on Web browsers, and later became a browser for arbitrary XML documents. This activity, which began as a simple student project at**

**the New Jersey Institute of Technology (USA) in 1996, has now become an important enterprise, involving faculty and graduate students of the University of Bologna with additional support from the DaimlerChrysler research centre in Berlin.**

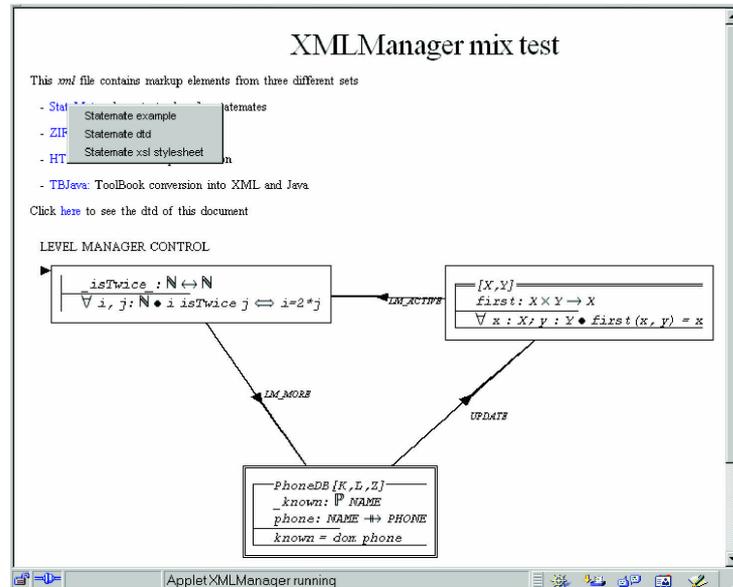
The World Wide Web has developed in an idiosyncratic, difficult-to-predict fashion. For instance, the Web community has valued the development of standards and protocols over that of functionalities. This has led to the creation of some dozens of different languages and protocols that are necessary to master the task of creating satisfactory websites.

In our opinion this richness of languages is clear evidence that while, on the one hand, there exists the possibility of implementing a large number of interesting functionalities, on the other hand, unfortunately the Web does not enforce or even encourage their use. This is left up to the willingness and awareness of individual authors of Web pages and sites.

The XML family represents a considerable advancement over previous languages and standards. The possibility given by XML to define a syntax (ie, a Document Type Definition, or DTD) tailored for one's own document classes, and to use standard XML tools to create, verify and exchange data is a real bonus. Additionally, XSL adds much to XML in terms of reach and flexibility, including a mapping language that can be used to transform one XML document into another.

Our long-term aim is to create an environment that, while relying on several existing Web languages and protocols, can provide fundamental and substantial hypertext functionalities in a streamlined and easy way. Currently we are concentrating on browsing and displaying hypertext data.

In order to provide flexible support for the special rendering needs that authors often have, we propose the use of 'displets'. Displets are software modules (currently Java classes) that are associated with each element of an XML document



The XMLC environment showing a complex XML document.

and provide the rendering behaviour for that element. Support for the most common element types is provided (for instance, text elements and paragraphs) but, at any time, it is possible to add new modules enabling specialised rendering semantics for specific needs.

Our implementation of displets is based on the XMLC prototype. XMLC can be considered as a very general architecture to provide very sophisticated functionalities to documents created in the XML format. However, the architecture of XMLC can be fruitfully used for more than visualisation, as it provides an extremely general way to associate behaviours with XML elements, and thus to produce active documents that perform computations, enact goals, produce results. In fact, the XMLC prototype uses XSL styles to associate XML elements of the source document with Java classes that can be displayed within a generic applet, the XML Manager. Furthermore, these classes can be made to respond to events, and to be activated in any manner decided by their implementers.

We have applied the displet approach to hypertext functionalities, implementing full support for XPath, XPointer and XLink, the tripartite hypertextual architecture of XML, and also in software engineering applications, providing displets for Z specifications, UML schema, finite state machines in the StateMate form, and Petri nets.

In our opinion, XMLC is the most customizable and expandable architecture for displaying XML documents. Full information on the XMLC browser and references to additional literature can be found on our website.

#### Links:

**XMLC website:**  
<http://www.cs.unibo.it/projects/displets/>

#### Please contact:

**Fabio Vitali – University of Bologna**  
 Tel: +39 051 2094 872  
 E-mail: [fabio@cs.unibo.it](mailto:fabio@cs.unibo.it)

# Using XML for Data Management

by Victoria Marshall, Brian Matthews and Kevin O'Neill

Recognising the limitations of HTML for the representation, discovery and exchange of structured data, the World Wide Web Consortium (W3C) has developed an alternative, XML, together with a set of related standards. This allows user communities to

define their own formats for structured data whilst allowing easy and open integration with other XML formats. XML has been widely taken up as a potential enabler of a wide variety of applications.

XML is an open and simple format for representing structured data tailored to the Web, and as each user community can provide its own format, there is no loss in semantics. Many user communities have been developing XML formats for the exchange of information including MathML for mathematics, Chemical Markup Language (CML) for chemistry,

any other XML format. Thus graphics can be described in a flexible yet compact format compared with, say, binary formats such as GIF or JPEG.

## Common Interfaces to Data Sources

XML can define schema formats for classes of documents. These XML schemas can be used to define common

## Different User Views

XML is a purely structural format; there is no semantics provided for defining its appearance. This separation of structure and presentation is an important feature of XML: the data may not be displayed at all, but passed straight to some program, or target database, so presentation information may be superfluous. Also, by providing presentational information by means of a stylesheet, different views can be presented on the same data, for example, data in trees, tables, graphs, or charts. These views can be tailored to the user requirements. Using the new W3C recommendation XSL, a high degree of interactivity with the data can be provided on the client, without requiring further interaction with the server.

ISE has been experimenting with providing different views on the laboratory personnel database (see figure 1). This allows the user to explore the structure of the organisation without downloading further data from the server. The same data set also can be used to generate alternative views by simply changing the colours and interface, or by combining the data with SVG to, say, provide an annotated map of the laboratory offices (see figure 2).

Figure 1: A view on a personnel database.

Figure 2: Alternative view on the personnel database.

Bioinformatic Sequence Markup Language (BSML) for the human genome project, Extensible Scientific Interchange Language (XSIL), DDI for Social Science Data, and Schematic Vector Graphics (SVG) for diagrams.

The latter has been developed by ISE at RAL working with W3C (see article on page 24). SVG uses graphical primitives such as boxes, lines, and circles which can be integrated with other XML formats. An SVG graphic fragment can be embedded within an HTML document, and can itself have labels which are of

interfaces to live data sources, independently of the concrete representation of the repository. This raises issues in using the hierarchical structure of XML, essentially derived from the structured document community, to represent relational database semantics. XML is a relatively poor language for expressing concepts such as primary keys or constraints. Nevertheless, the common interface layer it offers means that it can be a powerful mechanism for exchanging data between heterogeneous data sources.

## Links:

<http://www.itd.clrc.ac.uk/>

## Please contact:

Victoria Marshall – CLRC  
Tel: +44 1235 44 6799  
E-mail: [V.A.Marshall@rl.ac.uk](mailto:V.A.Marshall@rl.ac.uk)

# Corporate Knowledge Management through Intranet and Internet

by Rose Dieng

**The ACACIA research team at INRIA-Sophia-Antipolis, aims at offering methodological and software support in form of models, methods and tools for knowledge**

**management ie, for building, managing and distributing a corporate memory.**

The ACACIA team studies the problems raised by the dissemination of knowledge through a knowledge server via the company's Intranet or via the Web: we consider the Web as a privileged way for supporting management of knowledge distributed either inside a company or between several companies (for example in the framework of a project memory or of technological monitoring). We aim at building knowledge servers enabling search for information in a heterogeneous corporate memory, this search being 'intelligently' guided by ontologies and ontological annotations. We focus on the case of a corporate memory materialised in the form of XML documents. The ontologies can then be expressed either in RDF Schema, or in Sowa's Conceptual Graphs or in CML (CommonKADS Modelling Language).

## Research on Knowledge Servers

We developed WebCOKACE, a knowledge server for distributing CommonKADS expertise models on the Web. It enables to explore expertise models and to perform hypertextual navigation in expertise models and towards associated electronic documents thanks to a Web browser .

## Research on XML Galaxy

Considering that information search in the memory can be guided by knowledge models and semantic metadata, Acacia studies information search in XML documents semantically annotated by RDF metadata or by conceptual graphs. We built a translator transforming an ontology expressed in CommonKADS, into an RDF schema that may serve as an annotation model of the corporate XML documents. We built an ontological filtering engine to answer simple requests of the user. We proposed a model of interpretation of RDF and RDF Schema in terms of conceptual graphs. We developed a translator of RDF Schema/RDF towards conceptual graphs.

It enables to exploit projection on conceptual graphs in order to retrieve the adequate XML documents semantically annotated by RDF metadata thus translated into conceptual graphs. This work is implemented in JAVA in the OSIRIX tool. We are studying exploitation of complex inference mechanisms on conceptual graphs in order to enable different kinds of approximate search of information.

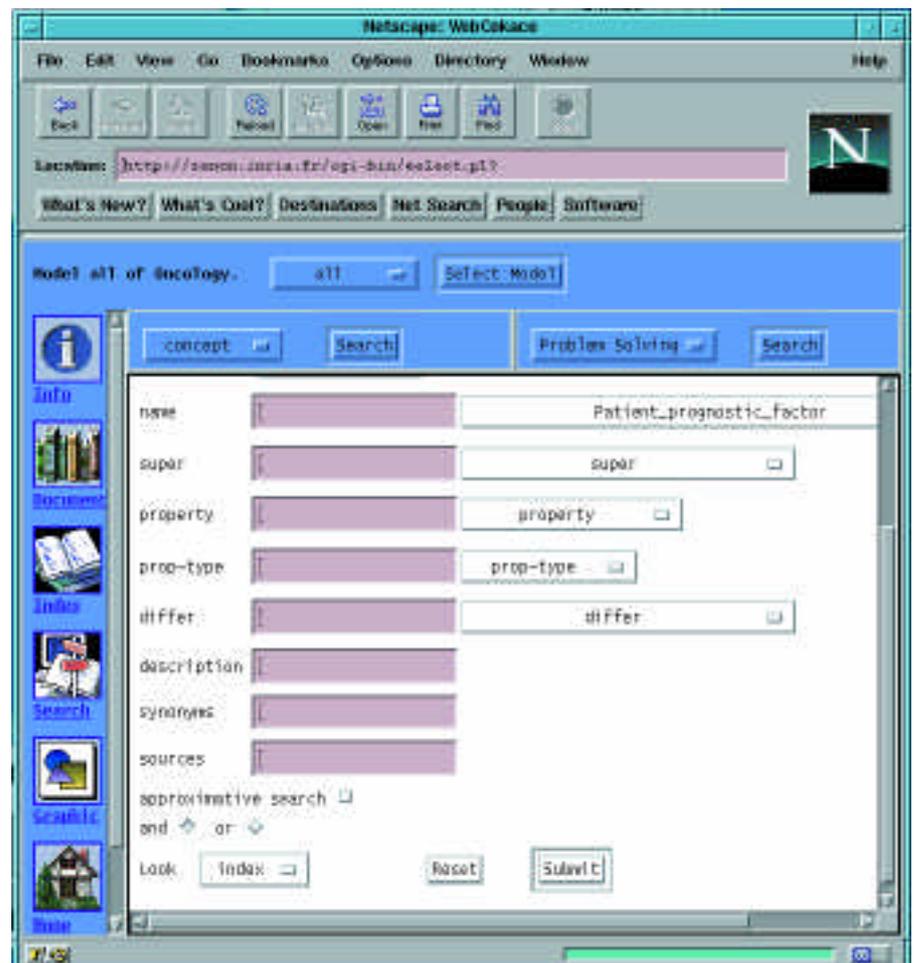
We also developed an inference engine able to perform elementary, event-guided, processings, on XML documents: consistency management, elementary

deductions. This engine was applied to make deductions using a XML-expressed knowledge base in road accident analysis.

Moreover we are studying an exploitation of RDF and RDF Schema for taking into account the notion of viewpoint. The purpose is to enable to represent an ontology with multiple viewpoints, corresponding for example to multiple professions in a company.

## Applications

We developed RESEDA, a system aimed at supporting road accident analysis. Installed on the intranet of INRETS, it



**WebCokace, a CommonKADS Expertise Model Web Server: example of search.**

enables the investigators of this institute to fill the accident dossiers in a computational form, with an intelligent support. The check-lists constituting an accident dossier are reconstituted from a generic XML document and from data stored in a database. According to the data entered about the current accident, RESEDA exploits a knowledge base (written in XML format) in order to provide the user with suggestions useful for his task of accident analysis. These suggestions stem from a knowledge base that was built thanks to acquisition and modelling of knowledge of experts of INRETS. RESEDA also exploits a base of generic scenarios (written in XML format) in order to suggest the most plausible generic scenarios that may be applicable for the reconstitution of the current accident analysed by the user. The scenarios are visualized thanks to XSLT stylesheets. RESEDA illustrates an application combining knowledge engineering and XML-based technologies.

### Cooperative projects

The ACACIA project takes part in two IST projects that started officially in February 2000:

- COMMA (Corporate Knowledge Management through Agents) aims at building a multi-agent system, made of several agents cooperating for information search in a corporate memory materialised by XML documents, these agents having capabilities to adapt to the user. This project involves ATOS, CSELT, Deutsche Telekom, INRIA, LIRMM and University of Parma
- C-WEB (Community Webs), that aims at designing a generic platform to support community-Webs. This project involves EDW-International, ERCIM, ICS-FORTH, and INRIA.

The ACACIA team also takes part in the cooperative project ESCRIRE (Embedded Structured Content Representation In REpositories) with EXMO and ORPAILLEUR, two other

teams of INRIA (See articles on pages 9 and 10). This project aims at comparing three types of knowledge representation formalisms (conceptual graphs, object-based knowledge representation and description logics) from the standpoint of the representation and the handling of document content.

Last, the ACACIA project takes part in the Working Group on Agent-Based Mediation Among Information Systems, Data, and Knowledge Management, in the framework of the AgentLink SIG on Intelligent Information Agents.

### Links:

ACACIA website: <http://www.inria.fr/Equipes/ACACIA-eng.html>  
WebCokace: <http://inria.fr/acacia/Cokace>

### Please contact:

Olivier Corby and Rose Dieng – INRIA  
Tel: +33 4 9238 7871/7810  
E-mail: [Olivier.Corby@sophia.inria.fr](mailto:Olivier.Corby@sophia.inria.fr),  
[Rose.Dieng@sophia.inria.fr](mailto:Rose.Dieng@sophia.inria.fr)

## Multilingual Metadata to access Social Science Data

by Brian Matthews and Michael Wilson

**Most countries have a national archive for social science data. This is data about social attitudes, and financial and environmental details, which describe the current and past state of each nation. This information derives from government statistics, and through academic studies, often commissioned by**

The European funded LIMBER project, which began in January 2000, brings together CLRC-RAL, Intrasoft, the UK national archive at Essex University, the Norwegian national archive at NSD, and three other national archives within Europe. The vision behind LIMBER is the interoperability of data, for example, from WHO Health archives, with social science datasets on behaviour, and genetic datasets from the Human Genome Project. These can be integrated together to show the localisation of potential genetically abnormal populations. This would then be presented through a multilingual interface so that it can be used for policy making and planning.

LIMBER will achieve this by providing a uniform metadata description. Metadata allows the explicit specification of the semantics of data, relationships between data, and its quality (recency, accuracy etc). We describe LIMBER's key features.

### Multilingual Thesaurus

Using a controlled vocabulary to index metadata increases the relevance to retrieval and when this is structured in a thesaurus it further helps to refine searches. LIMBER will extend existing thesauri by using an XML representation and adding 'equivalent terms' in other languages. This provides the terms to catalogue datasets, and, using equivalent

**government. These archives are usually a few hundred gigabytes in size, and accessible partially via the Internet and the Web. However, it is becoming increasingly important to allow access across national and linguistic boundaries so that decision makers have a comparative picture of European society.**

terms in a multilingual search, relevant data across the archives can be discovered.

### Indexing

Creating metadata usually takes considerable effort, and whilst its benefits are evident to users, data contributors have little incentive to spend this effort. To ameliorate this, an automatic tool will index the metadata, by scanning for relevant terms and converting these to the controlled vocabulary, in the language of the metadata.

### Multilingual User Interface

The multilingual thesaurus will allow users to perform searches using their own

language. To refine their search, the thesaurus will be presented in the user's natural language. The relevance of data returned, can be shown by displaying terms from the controlled vocabulary in the user's own language.

#### Metadata using RDF

Metadata is a very active area with many proposed schemes. The problem with standardising metadata definitions, is to achieve the correct layering, so that domain specific elements are defined on a generic base technology. Recognising this, the World Wide Web Consortium (W3C) has defined a recommendation for web-based metadata, the Resource

Description Framework (RDF). This defines a model, syntax, and a representation of schema to capture semantics in terms defined by authoritative bodies. RDF is defined in XML, so as browsers are supporting XML, they are also supporting RDF, providing economic and easy to use tools.

LIMBER will develop a metadata model (in XML and RDF) for social science datasets to allow their integration within and across archives, building on the existing work of the Data Documentation Initiative (DDI) from University of Michigan, the most advanced metadata structure currently proposed for social

science. RDF also offers a uniform vehicle to manage multilingual thesaurus as full ontologies. Further, common RDF format allows the seamless integration with metadata from related fields, such as geographical, environmental, and health data.

#### Links:

<http://www.linglink.lu/hlt/projects/limber/>

#### Please contact:

Michael Wilson – CLRC  
Tel: +44 1235 44 6619  
E-mail: M.D.Wilson@rl.ac.uk

## ICS-VRP: a Tool for Parsing and Validating RDF Metadata & Schemas

by Karsten Tolle and Vassilis Christophides

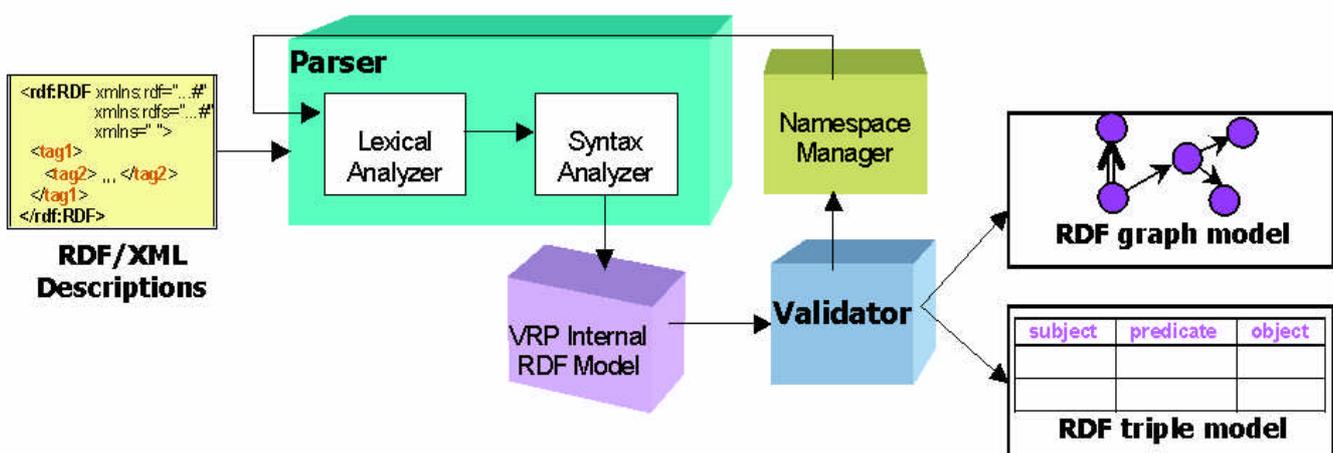
The Web provides a simple and universal infrastructure to exchange various kinds of information. In order to share, interpret, and manipulate information worldwide, the role of metadata is widely recognized. Indeed, metadata allow us to easily locate information available in the Web, by providing descriptions about the structure and the

content of the various Web resources (eg data, documents, images, etc.) and for different purposes. The emergence of the Resource Description Framework (RDF) is expected to enable metadata interoperability across different communities or applications by supporting common conventions about metadata syntax, structure, and semantics.

More precisely, it provides a) a Standard Representation Language for Web metadata; and b) a Schema Definition Language (RDFS) to interpret (meta)data using specific class and property hierarchies (ie vocabularies). Moreover, RDF/RDFS offer a syntax for representing metadata and schemas in XML, enabling the creation and exchange of RDF

descriptions in a both human readable and machine understandable form. Many information providers like ABC News, CNN and Time Inc., Web portals like Open Directory as well as Web browsers like Netscape, and search engines like Altavista, Yahoo and Webcrawler already support the RDF proposal. Unfortunately, existing RDF parsers (eg, SiRPAC) check

only the well-formedness of RDF resource descriptions according the W3C RDF M&S specifications. For this reason, we have developed the ICS-FORTH Validating RDF Parser (VRP) allowing the validation of RDF resource descriptions against the associated RDFS schemas, as well as of the schemas themselves.



ICS-FORTH Validating RDF Parser.

RDF is based on a directed graph model that alludes to the semantics of resource description. The basic idea is that a Resource (identified by a URI) can be described through a collection of Statements forming a so-called RDF Description. A specific resource together with a named property and its value is an RDF statement. RDFS schemas are then used to declare vocabularies, ie collections of classes and properties, that can be used in resource descriptions for a specific purpose or domain. VRP is a tool to analyze, validate and process RDF descriptions based on standard compiler generator tools for Java, namely CUP/JFlex (similar to YACC/LEX). As a result, users do not need to install additional programs (eg, XML Parsers) in order to run VRP while they can easily update or extend the VRP BNF grammar in case of changes in the RDF/RDFS specifications. VRP is a 100% Java(tm) development understanding embedded RDF in HTML or XML and providing full Unicode support. The quick LALR grammar parser (ie CUP) as well as the stream-based parsing support (ie JFlex) ensure good performance during the processing of large volumes of RDF descriptions.

The most distinctive feature of VRP is its ability to verify the constraints specified in the RDF Schema specification. This allows the validation of both the RDF descriptions against one or more RDFS schemas, and the schemas themselves. The VRP validation module relies on (a) a complete and sound algorithm to translate descriptions from an RDF/XML form (using both the Basic and Compact serialization syntax) into the RDF core model (ie triples) (b) an implementation of this model in Java to efficiently verify the RDFS constraints.

To favour metadata reusability, RDF supports a) the sharability of RDFS schemas using the XML namespace mechanism (ie provide only incremental modifications to a base schema in order to create a new variant); and b) the creation of RDF (meta)data using multiple schemas at the same time (ie merging different types of metadata). This implies to maintain for real scale applications, several interconnected RDF schemas than can be potentially used to describe Web resources. To meet these requirements VRP supports validation across several namespaces: we can connect to remote namespaces in order to

import the external statements we need to validate our RDF descriptions. Note that the RDF and RDFS namespaces are necessary for every RDF description and therefore their statements are by default included into VRP.

Currently VRP provides a command line interface with various options to generate a textual representation of the internal model (either graph or triple based). In the future, we plan to implement a graphical user interface to visualize the analyzed RDF statements as well as to interact with VRP during parsing/ validation. Finally, there is an ongoing effort to develop VRP APIs in order to facilitate the integration of the program into other systems (eg on-line loaders to DBMS).

#### Links:

ICS-VRP website:  
<http://www.ics-forth.gr/proj/isst/RDF/>

#### Please contact:

Vassilis Christophides – FORTH-ICS  
 Tel: +30 81 391628  
 E-mail: [christop@ics.forth.gr](mailto:christop@ics.forth.gr)  
 Karsten Tolle – Universität Hannover  
 E-mail: [storr@t-online.de](mailto:storr@t-online.de)

## XHTML 1.0 for Device-Independent Web Access

by Steven Pemberton

**The World Wide Web Consortium (W3C) has released the XHTML 1.0 specification as a W3C Recommendation. This new specification represents cross-industry and expert community agreement on the importance of XHTML 1.0 as a bridge to the Web of the future. A W3C Recommendation indicates**

**that a specification is stable, contributes to Web interoperability, and has been reviewed by the W3C membership, who favour its adoption by the industry. XHTML 1.0 was developed in a W3C working group chaired by CWI researcher Steven Pemberton.**

HTML currently serves as the lingua franca for millions of people publishing hypertext on the Web. While that is the case today, the future of the Web is written in W3C's Extensible Markup Language (XML). XML is bringing the Web forward as an environment that better meets the needs of all its participants, allowing content creators to make structured data that can be easily processed and transformed to meet the varied needs of users and their devices.

In designing XHTML 1.0, the challenge was how to design the next generation language for Web documents without making obsolete what is already on the Web. The answer was to take HTML 4.0, and rewrite it as an XML application. In simple terms that means making the HTML a well-formed XML document where start and end tags are always there and match precisely. Empty elements have to use the correct syntax. XHTML 1.0 allows authors to create Web documents that work with current HTML

browsers and that may be processed by XML-enabled software as well. Authors writing XHTML use the well-known elements of HTML 4 (to mark up paragraphs, links, tables, lists, etc.), but with XML syntax, which promotes markup conformance.

The benefits of XML syntax include extensibility and modularity. With HTML, authors had a fixed set of elements to use, with no variation. With XHTML 1.0, authors can mix and match

known HTML 4 elements with elements from other XML languages, including those developed by W3C for multimedia (Synchronized Multimedia Integration Language – SMIL), mathematical expressions (MathML), two dimensional vector graphics (Scalable Vector Graphics – SVG), and metadata (Resource Description Framework – RDF).

In addition to its extensibility, moving from HTML to XML via XHTML 1.0 lays the foundation for making Web content available to millions more users. People browsing the Web with cell phones or other mobile devices want Web content tailored to their needs. People with disabilities need ways to transform content into accessible formats.

XML documents can already be transformed using Extensible Stylesheet Language Transformations (XSLT), and

rendered using independent style sheets such as CSS style sheets. XHTML 1.1, already under development, coupled with device-specific style sheets and Composite Capability/Preference Profiles (CC/PP) - a protocol which allows a user to describe both user preferences and device capabilities - will bring mobile and other devices to the Web as full participants.

The XHTML 1.0 Recommendation was written by members of the HTML working group, which includes key industry players such as Ask Jeeves, CNET, Gateway 2000, CWI, GMD, Hewlett-Packard, HTML Writers Guild, IBM, JetForm, Microsoft, MITRE, Philips Electronics, Phone.com, Quark, Stack Overflow, Sun Microsystems, and WebTV Networks.

CWI always has been at the forefront of developments on the Internet and the

World Wide Web. In 1988 CWI was Europe's first non-military Internet site, and became a Dutch and European backbone site. Since then it was involved with the development of the Web programming language Python, HTML, CSS (Cascading Style Sheets), and SMIL, the Web's multimedia language, as well as spinning off companies such as NLNet that set up the initial Dutch Internet infrastructure, and General Design, the first Dutch Web Design company.

#### Links:

<http://www.w3c.org/>

#### Please contact:

Steven Pemberton – CWI  
Tel: +31 624 671 668  
E-mail: [steven.pemberton@cwi.nl](mailto:steven.pemberton@cwi.nl)

## Amaya: a Testbed for Web Protocols and Formats

by Irène Vatton

**Amaya is the World Wide Web Consortium's ( W3C) client which acts as both a browser and an authoring tool. It has been designed with the primary purpose of being a testbed for experimenting with, testing and demonstrating new specifications and extensions of Web protocols and formats.**

Amaya is the World Wide Web Consortium's ( W3C) client. It provides a complete Web browsing and authoring environment and has a WYSIWYG style of interface, similar to that of the more popular commercial browsers. With the extremely fast moving nature of Web technology, Amaya plays a central role at the Consortium. Easily extended to integrate new ideas into its design, Amaya provides users with many specialized features including multiple views, where the internal structural model of the document can be displayed alongside the browser's view of how it should be presented on the screen. Amaya has a counterpart called Jigsaw[1] which plays a similar role on the server side.

An important benefit of using Amaya as an editor is that it implements the W3C specifications very carefully, thus guaranteeing the interoperability and legacy of the markup it produces.

Integration is the word that best characterizes Amaya. Since the beginning, Amaya adopted the spirit of the Web as a distributed collaborative medium and it integrated seamlessly the two functions: authoring and browsing. At any time, a user is able to edit any document that s/he is browsing, without having to resort to another tool. As Amaya is able to work on several open documents, it's very easy to copy/paste information from one page to another page or to link pages by a simple click. Moreover, the user is able to publish documents on Web servers using the standard HTTP PUT method.

#### Nested Structures in Amaya

With the introduction of XML, it's now possible to create very rich documents on the Web, mixing text, graphics and mathematical expressions. Amaya allows users not only to browse these rich documents including HTML[2],

MathML[3], and SVG[4] elements, but also provides editing functions for creating them. Figure 1 shows an Amaya window where the three precedent elements are combined. XML namespaces are used to distinguish between each set of elements.

#### Styling

Amaya has extensive support for the W3C's Cascading Style Sheets (CSS[5]) and offers a simple-to-use interface enabling users to write style sheets for Web documents without having to know the details of the CSS syntax. You can create colored text, change the font, set the background color and control other presentational effects. CSS style sheets may also be used to alter the look of graphics or mathematical expressions. Amaya provides several features for testing, reusing, and publishing CSS style sheets. You can also temporarily apply or remove a CSS file to/from a document in order to test its effect.



which may fit more closely the needs of their information system.

Within W3C, the HTML language is now migrating from SGML to XML – this is called XHTML – including a modularization of HTML to suit the needs of a larger community (mobile users, Web TV, etc). XML is therefore not to be seen as a replacement of HTML, but as a new building layer on top of which HTML is to be placed, next to other languages designed by W3C, such as MathML (for representing mathematical formula), SMIL (for synchronizing multi media), SVG (for scalable graphics), etc., and other new languages designed by other organizations (such as OpenEBook, XML-EDI, etc).

### Problem statement

The Web Accessibility Initiative (WAI) has done extensive work in the HTML area, resulting in lots of new functionalities being added to the version 4.0 of the language. One area of concern with the advent of XML is that the freedom of design it brings will result in a loss of accessibility features, present today because of HTML pervasive presence and widely available specification. For instance, one can design a set of XML tags that would prevent the creation of accessible document, such as:

```
<MENU>New England Restaurant </MENU>
<APPETIZER>Clam Chowder
<PHOTO url="clam.jpg"/>
</APPETIZER>
```

with no way to include an alternate textual description of the photo, for instance.

But let's start by defining what we mean by accessible DTD and documents. An XML DTD is accessible if it enables and promotes the creation of accessible documents. A document is accessible if it can be equally understood by its targeted audience regardless of the device used to access it. For User-centric XML based languages, the message is simple: be device independent and export your semantics as much as you can. While the priority is stronger on the first aspect (multi-modality), both aspects are important, as without the knowledge of the meaning of the XML elements and attributes, there is little chance that

alternative user agents can do something intelligent with just the document bits.

This semantic knowledge can be provided through human readable documentation of course, but having machine readable assertions of some semantics that can then be used to present the document in various media is paramount for seamless access (ie, you don't need a programmer, you just need a program).

### Guidelines for Designers of User Interface-oriented XML Tagset

This section provides a list of proposed abstract guidelines. Some examples of checkpoints are provided, and detailed checkpoints and techniques that DTD designers can follow to achieve accessibility when designing new XML DTDs still have to be defined by WAI and W3C.

1. Ensure that authors can associate a text description with any non-text content (graphics, sound, multimedia, scripts, etc.):
  - make sure this is done in the most natural way possible
  - point at XHTML modules for OBJECT, MAP element, SMIL switch, etc.
2. Create semantically-rich languages. Do not define presentation elements or attributes:
  - Enable the use of style sheets (linked or internal).
  - use the standard linking mechanism
  - define element types that allow classification and grouping
  - define element types that identify important text content.
3. Export semantics:
  - provide accessible documentation
  - provide schemas and a mean to access it
  - reuse accessible DTDs, schemas, etc. (xlink, xml:lang).
4. design an accessible user interface: provide default style sheets for multiple output modes:
  - describe navigable structures
  - use CSS or XSLT to describe a basic outline view
  - document navigable structures

- describe how discrete, sequential, structured, and search navigation mechanisms should work.

An additional advice we give to DTD designers is that in their specification itself (the documentation) they always emphasize the accessibility features of their new language and try to include accessibility as part of any conformance statement that they introduce (be it for the document themselves, or for readers/editors of the language). See the Scalable Vector Graphics (SVG) specification for an example of both practices (<http://www.w3.org/Graphics/SVG/>).

### Conclusion

Our primary message is simple: be device independent, and export your semantics as much as you can. We believe abstract guidelines and verifiable checkpoints/techniques (using implementation mechanisms associated with abstract guidelines) are the best way to address this problem and we are in process of defining them in the framework of the WAI at W3C.

#### Links:

WAI: <http://www.w3.org/WAI>  
 WAI and XML: <http://www.w3.org/WAI/PF/xmlgl>

#### Please contact:

Daniel Dardailier – W3C  
 Tel: +33 4 9238 7983  
 E-mail: [danield@w3.org](mailto:danield@w3.org)

## The W3C in seven points



The W3C was created to lead the Web to its full potential by developing common protocols that promote its evolution and ensure its interoperability. It is an inter-national industry consortium jointly run by the MIT Laboratory for Computer Science (MIT LCS) in the USA, INRIA and Keio University in Japan. To date, over 390 organizations are members of the Consortium. For more information, see <http://www.w3c.org/>



### 1. Universal Access

W3C defines the Web as the universe of network-accessible information (available through your computer, phone, television, or networked refrigerator...). Today this universe benefits society by enabling new forms of human communication and opportunities to share knowledge. One of W3C's primary goals is to make these benefits available to all people, whatever their hardware, software, network infrastructure, native language, culture, geographical location, or physical or mental ability. W3C's Internationalization Activity, Mobile Activity, TVWeb Activity, Voice Browser Activity, and Web Accessibility Initiative all illustrate our commitment to universal access.



### 2. Semantic Web

People currently share their knowledge on the Web in language intended for other people. On the Semantic Web ("semantic" means "having to do with meaning"), we will be able to express ourselves in terms that our computers can interpret and exchange. By doing so, we will enable them to solve problems that we find tedious, to help us find quickly what we're looking for: medical information, a movie review, a book purchase order, etc. The W3C languages RDF, XML, and digital signatures are the building blocks of the Semantic Web.



### 3. Trust

The Web is a collaborative medium, not read-only like a magazine. In fact, the first Web browser was also an editor, though most people today think of browsing as primarily viewing, not interacting. To promote a more collaborative environment, we must build a 'Web of Trust' that offers confidentiality, instills confidence, and makes it possible for people to take responsibility for (or be accountable for) what they publish on the Web. These goals drive much of W3C's work around digital signatures, annotation mechanisms, group authoring, versioning, etc.



### 4. Interoperability

Twenty years ago, people bought software that only worked with other software from the same vendor. Today, people have more freedom to choose, and they rightly expect software components to be

interchangeable. They also expect to be able to view Web content with their preferred software (graphical desktop browser, speech synthesizer, braille display, car phone...). W3C, a vendor-neutral organization, promotes interoperability by designing and promoting open (non-proprietary) computer languages and protocols that avoid the market fragmentation of the past. This is achieved through industry consensus and encouraging an open forum for discussion.



### 5. Evolvability

W3C aims for technical excellence but is well aware that what we know and need today may be insufficient to solve tomorrow's problems. We therefore strive to build a Web that can easily evolve into an even better Web, without disrupting what already works. The principles of simplicity, modularity, compatibility, and extensibility guide all of our designs.



### 6. Decentralization

Decentralization is a principle of modern distributed systems, including societies. In a centralized system, every message or action has to pass through a central authority, causing bottlenecks when the traffic increases. In design, we therefore limit the number of central Web facilities to reduce the vulnerability of the Web as a whole. Tolerance of errors is the necessary companion of distributed systems, and the life and breath of the Internet, not just the Web.



### 7. Cooler Multimedia!

Who wouldn't like more interactivity and richer media on the Web, including resizable images, quality sound, video, 3D effects, and animation? W3C's consensus process does not limit content provider creativity or mean boring browsing. Through its membership, W3C listens to end-users and works toward providing a solid framework for the development of the Cooler Web through languages such as the Scalable Vector Graphics (SVG) language and the Synchronized Multimedia Integration Language (SMIL).

# Web Images and Blind Users

by Zdeněk Mikovec, Martin Klíma, Dušan Pavlica and Pavel Slavík

**The Blind Information System (BIS) is a system for interpretation of graphical information for blind users designed and developed at the Czech Technical University. The project is developing a picture**

**description methodology based on XML which characterizes pictures and charts by structure. It also looks at objects and the relations among them and then makes this information accessible to blind users.**

User-friendly communication with computers is a very important issue. Development of user interfaces is at the centre of attention of the software industry in the few last years and due to technological development new methods of communication are brought into everyday use. The past 15 years were characterised by intensive development of graphical user interfaces that increased the user-friendliness and thus the efficiency of communication with computers. This allowed a wider audience to use computers. Nevertheless there is one user group that has been seriously affected by the introduction of graphical user interfaces – blind users.

The amount of graphically presented information is steadily increasing and very often the textual information can not be understood without perception of the attached image. It is obvious that blind users have to get access to graphical information to be able to understand the information presented. The solution is to create an alternative textual description the blind users can work with. The easiest way is to create the list of objects that the picture consists of which the user can search or perform other operations on.. This approach does not provide the user with any structural information about relations among objects in the picture investigated.

The aim of our BIS (Blind Information System) was to create a web tool that would allow the user to browse graphical information. The basic idea is to create an alternative picture description in textual form that would be appended to optional section of some image formats like GIF or PNG.

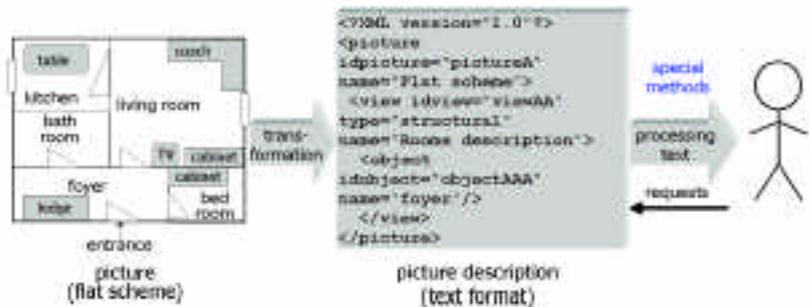
This approach would then allow the picture to be displayed by means of standard tools that are an integral part of common browsers like Netscape, Internet

Explorer, etc. Using a browser module that allows the user to browse the textual description of a picture it is possible to obtain information about picture both the structural one and the list of objects in the picture.

Besides the 'pure' pictorial information the picture also contains information of a semantic nature. This type of information covers relations like 'a person talks to another person', 'a girl waves to a boy', etc. Our system allows the creation of both types of descriptions and to browse them. In both cases a hierarchical description of a picture is created. The

implemented and tested (for testing purposes the system is implemented as a stand-alone Java application). The response from our blind volunteers has been very positive. The approach used allowed them to get relatively easy orientation in pictures. It is important to note that the use of the system is limited to pictures that can be characterised by structure, objects and relations among them. These can only be described by the formalism implemented.

Future work will be devoted to the development of the plug-in version for Internet browsers (eg Netscape) and the



## Picture perception process.

appropriate formalism by which it is possible to create such a description is a grammar that can describe both structures (describing structural and semantic view of the picture). The description grammar was implemented in XML. The use of XML has several advantages. Because XML is a standard the system is easily portable into different environments. Another advantage is that XML can be processed in web environments (XML was specially designed for use on Internet, and all newly developed web applications are able to work with XML documents). In the figure it is possible to see our approach to the picture perception process. The system that allows the creation of picture descriptions and picture browsing has been successfully

extension of the approach described into 3D. This means that scenes described by means of VRML could be both displayed in a normal way and investigated by means of textual queries concerning the scenes internal structure. In such a way the blind users will get access to information on the web currently not available to them.

### Links:

<http://sgi.felk.cvut.cz/cgi-bin/toASCII/Research/bis/>

### Please contact:

Pavel Slavík – Czech Technical University  
Tel: +420 2 2435 7480  
E-mail: [slavik@felk.cvut.cz](mailto:slavik@felk.cvut.cz)

# Adaptive Browsing: Supporting All Users on the Web

by Constantine Stephanidis

The AVANTI Web browser aims to facilitate access to Web-based information sources, applications and services, by potentially all categories of end users (including people with special needs), in different contexts of use.

The design space for the AVANTI Web browser was large and necessitated a move away from traditional user interface development approaches, which cannot cater for the foreseen levels of diversity and variability in the target user population and contexts of use. The novel Unified User Interface Development method was adopted in order to support

the practical application of the concepts and principles of 'User Interfaces for All' during the user interface development life cycle. Unified User Interfaces address the individual requirements of potentially all users in different contexts of use, by employing self-adaptation, to cater for their diverse, and often dynamically changing, characteristics. Self-adaptation

in the AVANTI browser is based on user (dis)abilities, skills and knowledge, requirements and preferences, as well as the characteristics of the context of use. Adaptations in the browser take place during the initiation of interaction (adaptability), as well as at run-time (adaptivity).

Two instances of the interface that demonstrate adaptation based on the characteristics of the user and the usage context are presented in Figure 1(a) and Figure 1(b). Specifically, Figure 1(a), presents a typical instance intended for use by a user familiar with Web browsing. Note the additional functionality that is available to the user (eg, a pane where the user can access an overview of the document itself, or of the links contained therein, an edit field for entering the URLs of local or remote HTML documents), as well as the facilitation of the link selection (links are presented as buttons, arguably increasing their affordance).

The second instance, Figure 1(b), presents the interface in 'kiosk mode', ie, making use of the kiosk-based information system metaphor. Note that the typical windowing paradigm is abandoned and replaced by a simpler, content oriented interface (for example, scrollbars have been replaced by 'scroll buttons', which perform the same function as scrollbars, but have different presentation and behavioral attributes – eg, they 'hide' themselves if they are not selectable). Furthermore, the interface 'toolbar' contains 'content-sensitive' buttons, ie, buttons which differ based on the actual information content of the pages being viewed.

Figure 1(c), presents disability-oriented adaptations in the browser's interface. The interface is presented with automatic scanning activated. The additional toolbar that was automatically added in the user interface is a 'window manipulation'

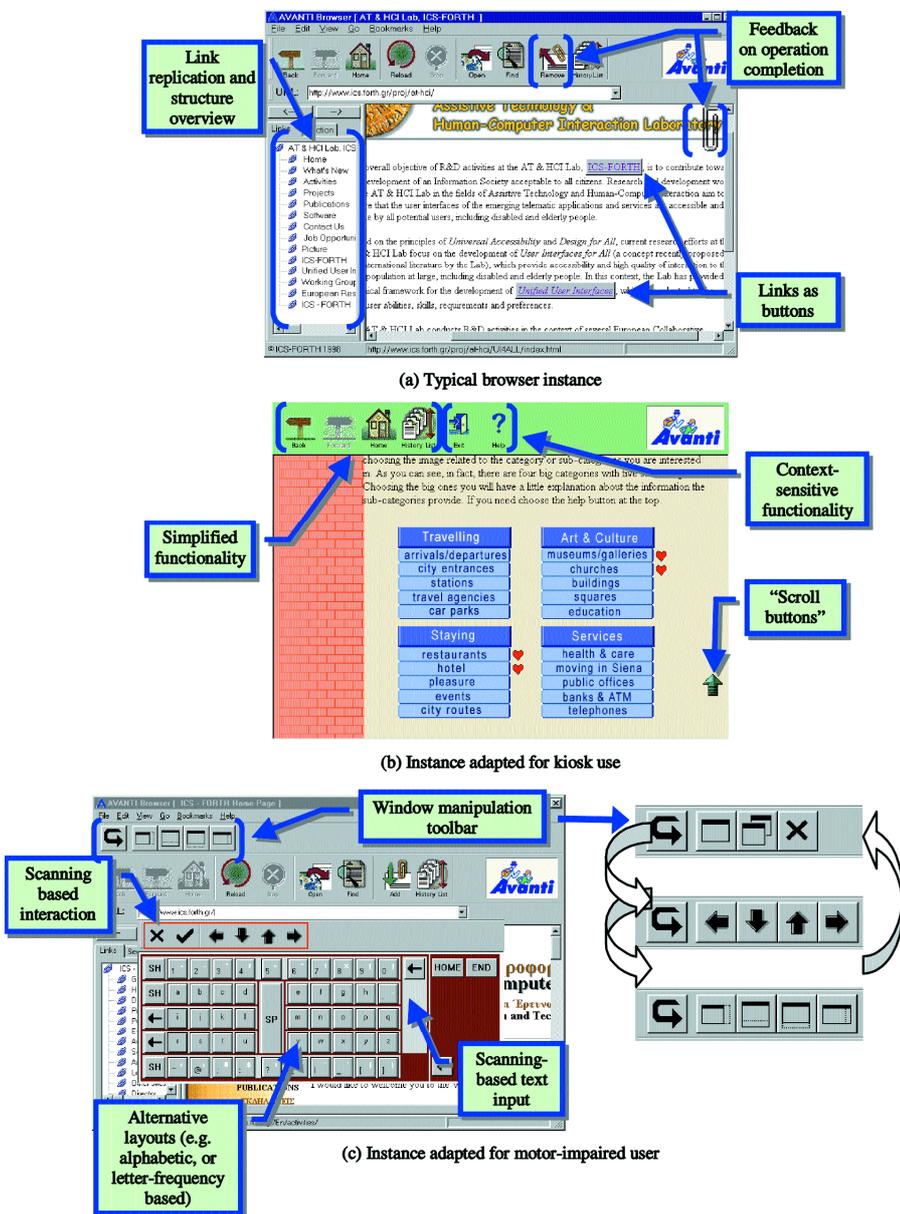


Figure 1: Instances of the browser interface.

toolbar, containing three sets of controls enabling the user to perform typical actions on the browser's window (eg, resizing and moving). The three sets of controls, as well as the 'rotation' sequence between the sets are also depicted on Figure 1(c). These sets occupy the same space on the toolbar, to better utilize screen real estate, and to speed up interaction; the user can switch between them by selecting the first of the controls. In the same figure one can also see the on-screen, 'virtual' keyboard activated for text input (interaction with the keyboard is also scanning-based), as well as the scanning highlighter over the toolbar on the screen keyboard that provides feedback on the current selectable item.

Run-time adaptations in the AVANTI Web browser aim to enhance the quality

of interaction by providing real-time support and guidance to the user as he/she interacts with the system. The system monitors the interaction session, identifies possible interaction gaps and triggers the activation of appropriate dialog patterns or interface adaptations. The different categories of run-time adaptations include the provision of awareness prompting, adaptive guidance and help facilities, and the activation of alternative interaction patterns for each browser task.

The AVANTI Web browser prototype, developed by ICS-FORTH, is currently the only system known to provide access to the Web for a range of user categories and different contexts of use. Ongoing work involves the support of different platforms and different terminal devices, as well as the compliance to new

guidelines and recommendations for User Agents and Web Content Accessibility. The AVANTI consortium, partly funded by the ACTS Programme of the European Commission, comprised: ALCATEL Siette (Italy) - Prime contractor; CNR-IROE (Italy); ICS-FORTH (Greece); GMD (Germany); University of Sienna (Italy); MA Systems (UK); MATHEMA (Italy); VTT (Finland); ECG (Italy); University of Linz (Austria); TELECOM ITALIA (Italy); EUROGICIEL (France).

**Please contact:**

**Constantine Stephanidis - ICS-FORTH**  
Tel: +30 81 391741  
E-mail: [cs@ics.forth.gr](mailto:cs@ics.forth.gr)

## Vector Graphics

by David Duce

**This article describes two recent developments in 2D web graphics, WebCGM and Scalable Vector Graphics. Work in progress at CLRC and Oxford Brookes**

**University exploring the translation of higher level languages into Scalable Vector Graphics is also described.**

Ever since the introduction of images via GIFs in the Mosaic browser, the only way of including simple vector graphics (flow diagrams, floor plans, simple layouts etc) in a web document has been to compose the vector graphics using some other system and translate the result into a pixel-based image. This leads to an inefficient representation, inability to search for the content in the diagram (particularly the text), and an inability to make changes. Furthermore, the resulting image is completely inaccessible to people with disabilities or users working in environments where visual representations are inappropriate, unless alternative descriptions are included. Two recent developments are changing this picture: WebCGM and Scalable Vector Graphics. Both address 2D graphics. Some of the requirements for the inclusion of 3D graphics in web pages have been addressed by the VRML (Virtual Reality Modelling Language) standard; 3D graphics is outside the scope of this article,

other than to note that VRML is being further developed within the Web3D consortium, see <http://www.web3d.org/> for details.

Vector graphics offers a number of advantages over image-based graphics: vector graphics are fully scalable without loss of resolution, file sizes are independent of resolution (leaving aside considerations of level of detail), a rich set of primitives is typically provided, text is represented as text strings, hotspots can be defined by drawing elements. There are both proprietary and open systems around, for example the proprietary systems PDF (Adobe), WMF (Microsoft) and DXF (Autocad). Open systems include the ISO/IEC standard CGM (Computer Graphics Metafile) and the recommendation under development in the World Wide Web Consortium SVG (Scalable Vector Graphics). CGM and SVG are briefly described below.

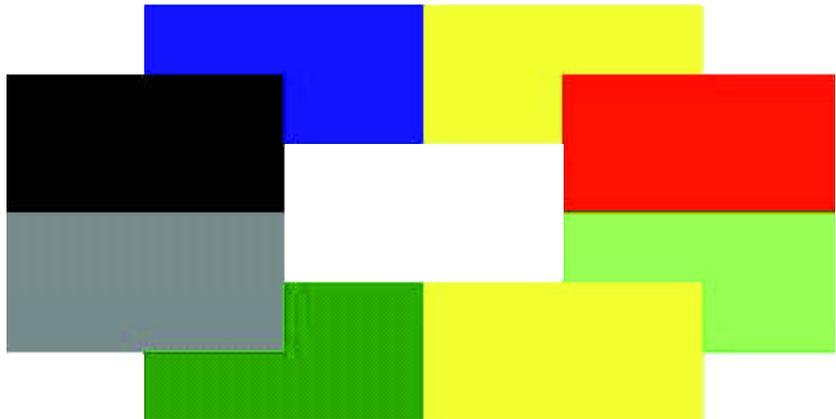
### WebCGM

CGM, the Computer Graphics Metafile, is an ISO/IEC standard for the storage and transfer of picture description information. The first edition of CGM appeared in 1987 and a number of extensions have been standardized since then. CGM is widely used to describe technical illustrations and has a strong following in fields of engineering including the aeronautical and automobile industry and the defence industry. CGM provides a rich set of primitives and attributes for describing pictures. Abstract and concrete syntax are separated in a nice way, the functionality of the drawing elements are defined in Part 1 of the standard and three concrete representations (binary, clear text and character encodings) are defined in Parts 2 to 4. Another important concept in CGM is the idea of a profile. Profiles enable the standard to be tailored to the needs of specific application areas in well-controlled ways, for example by

```

<pic>
  <right>
    <box fill color = "blue" fill = "true" />
    <box fill color = "yellow" fill = "true" />
  </right>
  <down>
    <box fill color = "red" fill = "true" />
    <box fill color = "lime" fill = "true" />
  </down>
  <left>
    <box fill color = "yellow" fill = "true" />
    <box fill color = "green" fill = "true" />
  </left>
  <up>
    <box fill color = "gray" fill = "true" />
    <box fill color = "white" fill = "true" />
  </up>
</pic>

```



### Web Schematics markup of a diagram and the corresponding visible representation generated by translation to Scalable Vector Graphics.

restricting the range of primitives and attribute values that need to be supported in order to meet the specific needs of a particular area.

WebCGM is a profile of CGM designed to meet the needs for scalable 2D vector graphics in electronic documents on the World Wide Web. CGM was developed by the CGM Open Consortium in collaboration with W3C staff and with the support of the W3C-LA leveraging action funded by the EU. Staff at a number of ERCIM institutes have been engaged in this activity. The publication of WebCGM as a W3C Recommendation represents a significant interoperability agreement between major players in this field. As well as providing a rich set of primitives, WebCGM also includes elements to group graphical objects for linking, layering and searching.

There are two points to note about WebCGM are that there isn't an XML representation of WebCGM and also it does not use the web style sheet mechanisms. For some applications, these are significant limitations.

#### Scalable Vector Graphics (SVG)

SVG is a Recommendation under development in the World Wide Web Consortium. The SVG activity started around 1998 with the submission of a proposal called Web Schematics by Bob Hopgood and David Duce at RAL and Vincent Quint at INRIA. Shortly afterwards a consortium led by Adobe

submitted a proposal called PGML and a few weeks later a consortium led by Microsoft submitted a proposal entitled VML (Vector Markup Language). W3C responded by establishing a working group on scalable vector graphics and after an initial requirements phase, the first working draft of SVG appeared in February 1999.

Since that time SVG has matured considerably and prototype implementations have started to appear. Unlike WebCGM, SVG is represented in XML and fully utilizes the style sheet and DOM mechanisms provided by other W3C Recommendations. In terms of application areas, it is fair to say that WebCGM is slanted more towards the CAD and engineering areas, whereas SVG is slanted more towards business graphics and graphics arts. SVG is based on the painter's drawing model. Shapes defined by paths may be filled and stroked in a variety of styles. Fill styles include rich gradient and radial fills and filter effects such as Gaussian blur, specular and diffuse lighting may be applied to shapes or groups of shapes. Animation facilities are also under development in conjunction with other W3C Working Groups.

#### Web Schematics

This is work in progress at RAL and now Oxford Brookes University to explore how higher level (in a sense modeling) languages, can be translated into SVG. So far an XML language based on the

Unix pic language for picture description has been designed and its translation into SVG using the DOM and a scripting language has been demonstrated. The language provides a more declarative way to describe diagrams than SVG.

#### Links:

WebCGM:  
<http://www.w3.org/Graphics/WebCGM/>  
 SVG:  
<http://www.w3.org/Graphics/SVG/>  
 SVG plug-in viewer:  
<http://www.adobe.com/SVG>

#### Please contact:

David Duce – Oxford Brookes University  
 and CLRC  
 Tel: +44 1865 48 4379  
 E-mail: [daduce@brookes.ac.uk](mailto:daduce@brookes.ac.uk)

# More Quality for Services in Internet

by Juha Koivisto

Internet traffic consists of packets. How well the packets flow can be characterized by Quality of Service (QoS) parameters. Several approaches to provide better service for selected applications have been defined. A research team at VTT Information

Technology is building test networks to evaluate QoS technologies in various configurations and combinations. Particularly interesting are configurations where QoS support is available only in very limited parts of the network.

Internet traffic consists of packets. These Internet Protocol (IP) packets carry all the data of, for instance, Web, e-mail and IP phone applications. How well the packets flow can be characterized by Quality of Service (QoS) parameters. QoS parameters include bandwidth (bits/second), average delay (transmission time from sender to receiver), delay variation, packet loss rate, and error probability. Which parameters are the most important depends on the applications. WWW users want as broad bandwidth and low packet loss probability as possible in order to shorten download times. Users of interactive voice applications cannot take advantage of the bandwidth exceeding a certain limit, but they find short delays important. In general, real-time applications are more sensitive to QoS than other applications.

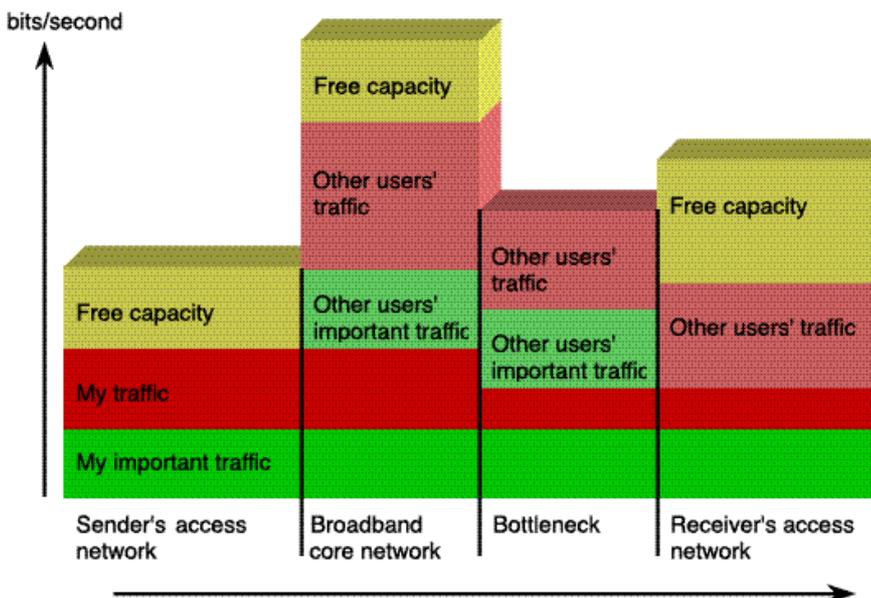
When no special actions are taken, IP packets are transmitted as fast as the bandwidth and load on the network permit. However, this best-effort service

is not always sufficient. Users want better quality for important applications. They are even ready to pay higher prices to guarantee the smooth operation of critical applications. Therefore, several approaches to provide better service for selected applications have been defined:

- Network managers may configure routers to favour certain applications, workstations or sub-networks. The routers set priorities based on the addresses and port numbers of the senders and receivers. Such priorities tend to be fairly static.
- Applications may categorize IP packets into different classes. Then the network knows which packets are, for example, less important and drops them in overload cases. This kind of differentiation of services does not require end-to-end signalling. However, there is no general agreement about the exact meaning of the traffic classes.
- Resources may be reserved in advance through, for instance, the Resource Reservation Protocol. This way, some

applications can get guaranteed QoS, which is independent of the load other users put on the network. However, establishing and maintaining such reservations is complex, and classifying the packets is heavy for network components. This is why resource reservations are usually not supported. Before resources can be reserved, it must be decided who is allowed to make the reservations, and who is going to pay for them, etc.

- Smart applications that adapt to available QoS can improve perceived service quality. For instance, an adaptive video encoder may decrease resolution and send fewer bits per second when the network is heavily loaded. Adaptive applications need to get feedback about the actual QoS.
- The network may be enhanced to provide more bandwidth, or separate networks may be used for critical applications. This may be the simplest but not necessarily the cheapest approach.



**Important applications can maintain high quality by having high priority in the network bottleneck.**

Internet traffic between two users in different organizations goes through several administrative domains. A combination of various approaches can be used to improve QoS. For example, resource reservation in an access network can be combined with priorities in the core network. It is not necessary to control QoS on all the nodes on the path from sender to receiver, but it is important to control the bottleneck of the connection. We are building test networks to evaluate QoS technologies in various configurations and combinations. The goal is to find out when it is worthwhile to use QoS mechanisms.

One problem with many QoS mechanisms is that they require extra network infrastructure, modifications to protocols, or co-operation between the sending and receiving applications. Often

such an infrastructure is not available, or end-users are not allowed to dynamically configure it. However, there are also simple mechanisms that set very little requirements on the environment. One example is the process-priority-based method that was developed in the Moses project. The method slows down selected applications by reducing the processor time they get. Through the properties of the Transmission Control Protocol (TCP), this will eventually slow down the sending rate on the other end and leave more bandwidth for critical applications. This method does not require any extra support from the network or the sending end or any modifications to existing

applications and it gives a larger share of the limited bandwidth to important applications. The solution is useful when no other differentiation mechanism is available and the network bottleneck is not loaded by other users. This situation is quite common with slow wireless and modem access links. For example, a user may want to give a larger share of the access network capacity to an interactive voice application, and less capacity to WWW downloading.

This work is part of the Moses project. Other research topics in Moses include mobility support technologies on network and application layers, and communi-

cation software development methods. Moses is a co-operation project between VTT, Lappeenranta University of Technology, HPY Research, Nokia Research Center, and Tekes (the National Technology Agency).

**Please contact:**

**Juha Koivisto – VTT Information  
Technology**  
Tel: +358 9 456 5664  
E-mail: juha.koivisto@vtt.fi

## The Social Web Cockpit: A Tool to Support Knowledge Sharing Communities

by Wolfgang Gräther and Wolfgang Prinz

**The Social Web Cockpit integrates different services that have been developed in the Social Web research programme at GMD Institute for Applied Information Technology to support a knowledge sharing**

**community. This includes a web-search based on recommendations, a concept index to develop a community vocabulary, shared workspaces for knowledge sharing, and an awareness service.**

Computer supported co-operative Work (CSCW) research yielded a number of groupware systems for the support of cooperative work such a shared document-spaces or workspaces, shared applications, and workflow systems. These applications support primarily work or knowledge sharing processes that are pre-planned or that involve well-defined, closed, or small groups of people. This paper describes systems that aim at the support of knowledge sharing within larger and more loosely coupled groups of people that share a common interest, ie knowledge sharing communities. The presented systems originate from our 'Social Web Research Programme' at GMD-FIT which aims to explore and demonstrate how we can turn information environments into rich communication and interaction environments. These are:

- BSCW, a Web-based shared workspace system
- Nessie, an application-independent awareness infrastructure
- Concept Index, a system for concept-based indexing of Web documents

- LiveMarks, a system of information processing software agents running on our agent platform SoaP.

We have developed a service, called the 'Social Web Cockpit', or 'cockpit' for short, that integrates these service and that turns websites into meeting places where visitors may become aware of each other, open communication channels, and exchange information and knowledge with each other or with experts. Usually people visit a website in search of some information. So there is some chance that its visitors may have overlapping interests, if the site matches their expectations, backgrounds, or motivations. These persons might well profit from each other's knowledge by sharing opinions or experiences or offering advice. Some visitors might enter long term relationships and eventually evolve a community - if only they became aware of each other. Similar functions are provided by systems such as CoBrow or ThirdVoice. The Social Web Cockpit goes beyond that. Visitors may enrich the website with their

knowledge by recommending pages and related links, by annotating and discussing them. People may add their own documents in a shared workspace, they may highlight key phrases and develop their special terminology. These concepts and terms are automatically cross-referenced with the site and related pages to support content-based navigation. Search facilities take into account concepts and collaborative ratings. Visitors may select pages from the site to be monitored for changes that will be indicated upon their next visit.

We expect that web-sites that are augmented with the cockpit and the Social Web Services may become very attractive for knowledge sharing communities because they offer the chance to meet again, to see what is new, and to inspect the growing number of contributions by its visitors. Eventually, the visitors may consider themselves as a virtual community that is held together by their common interests and this special place for meeting and trading information and knowledge.

Many websites would benefit from added community support: commercial sites that want to attract visitors, sites of professional organizations, self-help groups, clearing houses, digital libraries, project portals. Nowadays, however, community support is mostly restricted to chat, annotation and rating facilities, discussion forums, or forms to submit new links. The combination of a powerful shared workspace system, a collaborative recommendation service, a community vocabulary for knowledge organization, and a notification service, as offered in our Social Web Cockpit, is unique.

### Cockpit Services

The cockpit is a small window that occupies only little space on screen; all actions and controls can be dragged out in separate tiny windows. Therefore, most users don't need to reorganize their desktop when using the cockpit (see figure 1). After the cockpit has been launched and when a web-page is visited, a user becomes immediately aware of other visitors and the communities at the site. A cockpit is a place for social encounters so users can get in contact with each other using different communication tools. A chronology of community visits supplies further details about the distribution of visits over the last hours, days, or month. This helps users to decide on the best time for visiting a community.

The Social Web Cockpit uses BSCW to enable cooperation over the Web. BSCW (Basic Support for Cooperative Work) is a shared workspace system which supports document upload, event notification, group management, discussion forums, meetings, and much more. In the context of the Social Web Cockpit a community is a group of people sharing similar interests and tasks. In our approach, the documents which the community collects and produces are the main carrier of knowledge. Community knowledge is stored in BSCW shared workspaces. Creation of communities and community membership is implemented by BSCW functions. Users can rate the current page on a scale of 5 (poor, passable, fair, good, excellent). This personal rating is sent to the BSCW system, where it is added to other ratings of the page within the present community. The (changed) median of all ratings is

returned to the cockpit and displayed using the corresponding icon. Ratings are used in LiveMarks to help users to retrieve useful information from the Web.

The cockpit uses the NESSIE awareness infrastructure to turn websites into meeting places. Members of a community become aware of each other and get in touch. While visiting web pages, users get informed about the community to which the web page belongs. The median of all ratings of a web page – made by members of the community – is also shown. Users can easily contact experts of this community. Awareness information is visualized by different indicators; for example, changes of icons, detailed listings of present members in a specific community, etc. Members of a community may easily start an agent which monitors



Cockpit User Interface.

web pages. The monitoring agent adjusts to the users' preferences: the user may specify what web page should be monitored, for what keywords and at what dates and in what frequency this should be done. Users are notified of changes in a web page by a little yellow asterisk in the monitor icon.

The Social Web Cockpit uses the ConceptIndex system to enable members of a community to construct their own shared vocabulary. This vocabulary captures characteristic words and phrases; an index based on this vocabulary cross-references all web pages of this community. Web pages may be viewed with the community vocabulary; phrases in the web page are highlighted and links to the index are provided. A word or phrase can easily be added to the vocabulary. The index with all the cross-references is automatically updated.

The cockpit uses a recommender system, which enables members of a community to share queries and to collaboratively retrieve and collect useful web pages. While a member of a community submits

a query to a search engine, software agents in the background look for similar queries of the community with related and recommended web pages. The result of the search is a mixture of web pages originating from the recommender database and pages coming from the search engine. When the member recommends such a web page (by rating it good or excellent), the page is put into the link collection of this community.

### Summary

The different services developed by the Social Web Research Programme are combined to provide a comprehensive set of services for a knowledge sharing community. This includes a web-search based on recommendations, a concept index to develop a community vocabulary, shared workspaces for information

exchange, and an awareness service. We believe that the extension of a web-site with these services will foster the community building process because it enables a web-site to become a meeting place for a knowledge sharing community.

### Acknowledgement

We would like to thank the teams of BSCW, Concept Index, LiveMarks and Nessie of the FIT.CSCW research group who all cooperated to make the Social Web Cockpit operational and A. Voss who contributed to an earlier version of this article.

### Links:

**Social Web Cockpit:**  
<http://orgwis.gmd.de/cockpit/>  
**Social Web Research Programme:**  
<http://orgwis.gmd.de/projects/SocialWeb/>

### Please contact:

**Wolfgang Prinz – GMD**  
 Tel: +49 2241 14 2730  
 E-mail: [wolfgang.prinz@gmd.de](mailto:wolfgang.prinz@gmd.de)

# Eliminating the I/O bottleneck from World Wide Web Proxies

by Evangelos P. Markatos

**World Wide Web proxies are being widely used today, in order to reduce network traffic, web server load, telecommunication costs, and to provide Internet access to users behind a firewall. Web proxies are typically required to handle a heavy load which reaches up to millions of requests per day. To make**

**matters worse, this load is expected to increase sharply in the near future when users will be able to connect to the Internet via fast modems. A research team at ICS-FORTH investigates methods to enhance the performance of Web proxies thus enhancing web server performances.**

Although the network bandwidth has been traditionally thought to be the main bottleneck in a proxy's performance this is not true anymore. Recent research results suggest that a proxy server may spend a significant percentage of its time on file-system related activities.

Experimental evaluation of the web proxy at Digital Palo Alto firewall revealed that the disk I/O overhead of caching turns out to be much higher than the latency improvement from cache hits. Thus, to save the disk I/O overhead the Digital Palo Alto proxy server is typically run in its non-caching mode. This disk I/O overhead can be partially attributed to the fact that file systems were designed and optimized for an every-day pattern of use that is completely different from the needs of a web proxy:

- Contrary to read-dominated traditional file systems, disk accesses in web proxies are dominated by write requests: Each URL request that does not hit in the proxy's cache must be fetched from the network and stored into the proxy's local disk. Thus, each URL miss usually results in a disk write operation. Disk read operations are invoked when a URL request misses the proxy's main memory cache, but hits the proxy's disk cache. Our measurements suggest that for a web proxy file-system write operations are usually 3-5 times more than file-system read operations.
- Although file-update operations are frequent in traditional file systems, Web proxies rarely update their files: Most of the data on the web change rather infrequently. Published results suggest that the average life of a web document can easily be several days up to several weeks (or even months) long. This fact implies that most URLs stored on a proxy's disk may never be updated;

they will usually only be replaced by other URLs when the proxy runs out of space.

Thus, traditional file system technology can not keep up with the disk I/O needs of busy web proxies. Our current work focuses on identifying and eliminating the sources that create the disk I/O bottleneck. We investigate the effects of storage management policies for web proxies based on the following principles:

- Avoid file management operations: current proxies store a single URL in each file. Given that each file creation/deletion operation requires 1-2 disk operations, and that most disks can usually perform no more than 50 operations per second, then an average proxy can not deliver more than 20-30 URLs per second, or roughly 100 KBytes/sec, a rate of execution that is orders of magnitude lower than the data transfer rates that current disks can sustain. This data rate is even lower than most Internet connections. We advocate that proxies should not store one URL per file, but use a small number of large files to store all URLs. This approach reduces the number of operations associated with creating, opening, closing, and deleting files.
- Avoid disk operations: proxies should write data in large chunks and reduce the aggressive prefetching being employed by modern operating systems: prefetching is likely to hurt performance since proxy data are not accessed sequentially.
- Preserve locality of reference: each stream of URL requests that arrives in a web proxy from a single client has a significant amount of temporal locality (eg every time a client requests an HTML document (s)he will request the document's embedded images as well). Current proxies destroy this locality of

reference by serving several clients at a time, and effectively storing documents requested by different clients in nearby disk locations. The locality of reference that apparently exists in web client requests would be preserved and exploited to achieve a better disk data layout.

We have started testing our approach using a combination of experimental evaluation and simulation. Our initial results are very encouraging: they suggest that our method breaks through the I/O bottleneck and effectively improves performance over traditional proxies (like SQUID) by more than an order of magnitude.

#### Links:

<http://archvlsci.ics.forth.gr/OS/www.html>

#### Please contact:

Evangelos P. Markatos – ICS-FORTH  
Tel: +30 81 391655  
E-mail: [markatos@ics.forth.gr](mailto:markatos@ics.forth.gr)

# WEBPATH — a 3-D Browsing History Visualisation

by Emmanuel Frécon

**WEBPATH is a virtual reality based application that provides an easily recognisable presentation of your browsing history. The different dimensions of a 3-D**

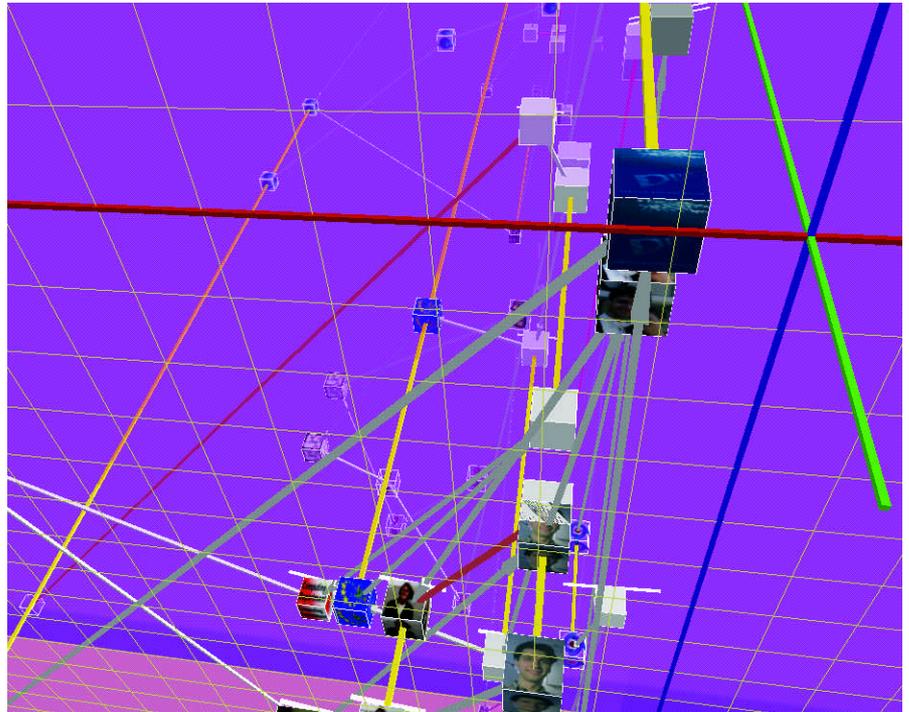
**space are used to position representations of documents using different metrics, together with their relationships.**

Only a few years ago, Web browsing was seen as an interesting possibility. It is now part of our working and leisure activities, and the time we spend looking for information ‘out there’ increases. As it becomes easier to publish documents, as the number of users, and thus publishers, increases and as the number of documents grows, searching for information is turning into a cumbersome and time-consuming operation. Moreover, because of the loose inter-connection among documents, people have difficulty returning to previously visited pages and remembering where they have been.

SICS and Lancaster University have prototyped WEBPATH, a virtual reality (VR) based application to be used alongside traditional Web browsers. A driving motivation for this work is specifically to assist users in scenarios where they are searching for pages previously visited. These scenarios include requests such as: “I want to find a page I visited this morning, all I know is that it had no images in it”; “I remember the page was from Asia somewhere”; “It had many links and the latest news in it”; and “It took a very long time to load”.

WEBPATH provides a 3D structure to the network of documents that have been visited during a number of browsing sessions. A three-dimensional presentation was chosen for maximum flexibility in data visualisation. The different dimensions of a 3-D space are used to position representations of documents using different metrics, together with their relationships. The result is an easily recognisable presentation of your browsing history.

The presentation can be tailored in terms of its overall structure and the appearance of individual pages. When a user loads a new document (using any of the methods offered by existing browsers: bookmarks, history facilities, manual location entry,



**A 3D network of documents that have been visited during a number of browsing sessions. The vertical axis portrays the time.**

etc.), WEBPATH will create a representation for that document and the user will be smoothly transported in front of it. Each Web document is represented as a cube that is labelled with the document’s title and appears as any one of: the page’s background image, an image in-lined in the HTML document or simply the background colour of that page. In order to indicate the structure of the user’s browse path, WEBPATH connects the cubes with arrows. The colour of the arrow indicates the nature of the link, for example, does the destination document originate from the same site as the previous document?

WEBPATH uses three orthogonal dimensions to visualise the browsing history. A number of metrics are available, such as page size or number of images contained in a document. Any of these metrics can be applied to the two horizontal axes (x and z). The vertical axis

(y) is always reserved to portray the time at which the documents were visited, the last visited document is always the top-most one. If a document happens to be revisited, it will generate a new cube, but all representations for the same document are vertically linked to ease recognition.

Clicking on a cube within the visualisation will cause the user’s browser to re-load the associated Web page. When the user selects a new document location using the different direct methods (ie bookmarks, history facilities or manual entry), he/she starts a new browsing session and a semitransparent horizontal plane, covering the whole visualisation area is generated just below the new representation. When many planes have been created (ie when a user has been engaged in several Web sessions) older documents, towards the bottom of the visualisation space, will appear through an increasingly dense fog, thus, reducing the user’s awareness of their

previous results. When looking back for previous documents, all the cubes resulting from a browsing sub-session are “isolated” in-between two semitransparent layers, which visually separates them from the rest of the data.

WEBPATH is a tool that unobtrusively visualises the users trail of URLs as they browse the Web, which can be tailored to best suit the users working needs.

WEBPATH provides users with a graphical view of their recent activities, enabling them to for example, see how a given document was reached and visually perform complex searches for previous documents visited. WEBPATH has raised substantial interest, ranging from casual users to scientific researchers. The authors are currently looking into commercial ways to distribute an improved version of the software.

**Links:**  
<http://www.comp.lancs.ac.uk/computing/users/gbs/webpath/>  
**Please contact:**  
 Emmanuel Frécon – SICS  
 Tel: +46 8 633 15 34  
 E-mail: emmanuel@sics.se

## The sView System

by Fredrik Espinoza and Markus Bylund

**The sView system developed at SICS makes networked services available to users in an environment that is close to the user. It allows**

**services to collaborate and further research will be done in the area of user assisted service collaboration using sView as a platform.**

The role of the Internet is about to change again. After the hype of web portals we look forward to an Internet in which networked services dominate. However, problems with the currently prevailing infrastructure for networked services, the World Wide Web, slow down this process. We argue that there are three main problem areas: limited support for service collaboration, producer dominance, and the poor support for service interaction in today’s web browsers.

A Personal Service Environment (PSE) enables a rich interaction between networked services and their users. A PSE, which is private to an individual

user, enables user interaction with networked services. The PSE can store and execute service logic and data, and it can move from computer to computer as the user moves between access nodes within the network. Preferably, the PSE is run on a computer under the user’s immediate control (such as his/her workstation at work or personal computer at home), but if the user is not directly represented on the network, the PSE can run on a server that is dedicated to running PSEs.

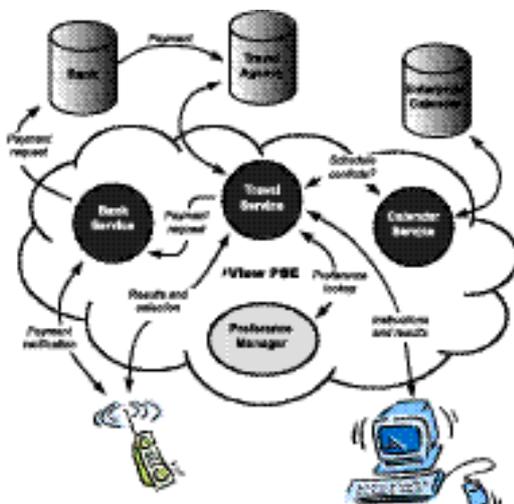
The choice of method for interaction between service logic within a PSE and its user is open. The PSE can provide a range of channels for services to interact

with their users (whichever channel is available for the moment), eg HTML over HTTP, WML over WAP, ASCII over SMS, GUIs via SWING, etc.

### Implementation

The sView system is a complete system for user-service interaction that implements PSEs. The system is authored exclusively in Java 1.2, which brings several advantages. The choice of implementation language makes it easier to use the system on different platforms. The popularity of Java makes it easy to reach a broad user group, and finally, Jini technology can be used for naming, searching and distribution of service components.

The sView system constitutes a very thin layer of infrastructure between the service providers and their users. Nothing is assumed about communication protocols between base services and service components within a PSE, or about interaction protocols between service components and the devices that render service presentations. Similarly, no languages, interpreters, or interaction styles are forced upon service components that wish to collaborate within the PSE. The thin properties of the sView system are a deliberate design choice. We want the system to be specific in its overall goal (to enable user-service interaction), but open when it comes to methods of achieving the goal.



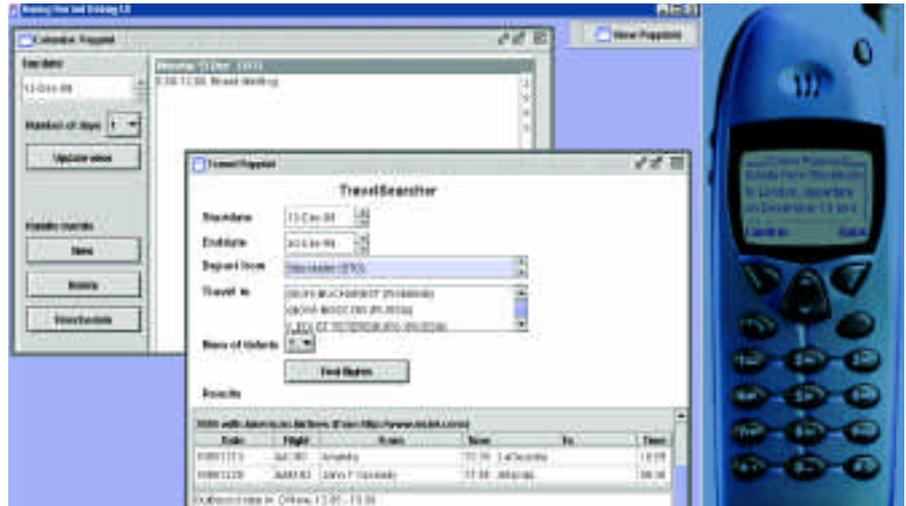
**Figure 1:**  
 A schematic overview of the described usage scenario.

### Sample Scenario

Below we describe a usage scenario that illustrates some features of the PSE. A man is about to make a business trip to a foreign city. Using his Personal Service Environment, he locates an electronically mediated travel agency and initiates a dialog with it. The travel agency uploads a travel service component to the user's PSE. Once in the PSE the travel service receives the man's instructions, via a standard graphical user interface (GUI), to make a flight and hotel reservation for his planned trip. Then the man turns his attention to something else and leaves the office bringing his cellular phone.

The travel service now makes use of a number of information sources in order to accomplish its task. It searches the PSE for a preference manager and asks it about its client's complete name and address, as well as his seating and smoking preferences. It also locates a calendar within the PSE and checks when the man must be back and if the trip conflicts with any of his other appointments.

Having collected all background information, the travel service turns to its base service trying to find an appropriate flight and hotel. The service finds three alternatives that all match the man's request, preferences, and schedule. The travel agency is now ready to get back to the client with the result of the search. However, since the man is no longer available via the desktop computer (as can be concluded since the screen saver has been on for quite a while), the service contacts him via his cellular phone. The man, now on the train on his way home, selects one of the alternatives and instructs the travel agency to go ahead with the reservation. The travel service accepts the request and starts searching the client's PSE again, this time for a service that provides payment. One of the man's services, a bank service, is willing to provide payment, but only after a confirmation by the client (this is also done through the interface of the cellular telephone). Having everything that is needed, including payment, the travel service now executes the man's request by instructing its base service to buy the flight tickets and make the hotel reservation.



**Figure 2: Screenshots from interactions with service components in the sView system. The calendar service and the travel service from the scenario above interact with the user through a SWING GUI. The payment service interacts via a cellular telephone. The snapshot of the cellular phone is taken from the Nokia WAP Toolkit 1.2 that is used to simulate WAP enabled cellular phones in absence of real ones (see <http://www.forum.nokia.com/developers/wap/wap.html>). WAP phones are available for European 900MHz/1800MHz digital networks, but not yet for American 800MHz/1900MHz digital networks.**

The above scenario illustrates three important aspects of the PSE:

- the PSE allows, and actively supports, service collaboration to take place within the environment
- since the infrastructure does not require HTML/HTTP, interaction through a number of different devices without loss of interaction state is possible
- the user shares personal information (such as preferences) with a special purpose component within the PSE (the preference manager). This makes it easy for the user to add, change, inspect, and retract information without having to contact every service that is used. At the same time, services have a central source for such information for every user.

### Features

In addition to what the above scenario illustrates, the sView system demonstrates the following noteworthy features:

- service components can be downloaded to the PSE for local execution.
- the use of three different interfaces for interaction with the same service (HTML/HTTP, WML/WAP, GUI/SWING).
- the interaction state can be preserved between usage sessions and throughout the transfer of interaction between devices

- sView allows services to collaborate. In itself it only provides the locale for this collaboration, a locale that is very intimately tied to the user. However, it is our intent to do further research in the area of human assisted service collaboration using sView as a platform.

A Personal Service Environment (PSE) enables a rich interaction between networked services and their users. We have introduced the concept of a Personal Service Environment as a solution to many of the problems with using the World Wide Web for mediation of networked services. The main contribution of the PSE is that it increases its user's control of networked services. Services are offered support for spontaneous collaboration between peers. The sView system will also serve as a platform for further research in the area of human assisted service collaboration.

### Links:

sView home page: <http://www.sics.se/~espinoza/sview/index.htm>

### Please contact:

Fredrik Espinoza and Markus Bylund  
Tel: +46 8 633 15 15, +46 18 471 70 53  
E-mail: [espinoza@sics.se](mailto:espinoza@sics.se), [bylund@sics.se](mailto:bylund@sics.se)

# Building a Multinational R&D Proposal through the Web

by Andrew Hallan, Alexander M. Samsonov and Wolfgang Ziegler

Last December, 20 scientific institutions and industrial companies led by GMD set up a Europe-wide consortium to write a proposal for the 2nd call of the EU's Information Society Technologies R&D programme. The preparation of a successful proposal required close collaboration between consortium members, together with strict adherence to the

European Commission's formats and procedures. The various parts of the proposal were prepared at the widely dispersed partner sites and distributed, discussed and combined in a shared workspace on the Web. This workspace was created and managed using BSCW, a tool developed by GMD. The consortium never held a meeting.

Scientific research repeatedly exposes new problems. Some of these, such as the investigation of non-linear waves in solids, require exceptional computing resources, beyond the capacity of any one computing centre. The aim of our proposal was to make it easier for scientists working on these problems to combine resources from a variety of high-performance computing centres.

Our consortium was a large one, spread across Europe from Spain to Russia and from England to Armenia. By the time that we had assembled the right partners, we had only two months to write the proposal. Writing an EU R&D proposal is a complex process, where scientific objectives and approach must be reconciled with the distribution of tasks between partners and an overall budget. The geographical spread of our partners and the lack of time effectively prevented us from holding the meetings typical of R&D consortia.

The solution was to organise all the collaboration through the Internet. But from the beginning it was obvious that just setting up e-mail exploders and an ftp-site would be insufficient for discussion and management of documents. There would be confusion and wasted work because of missing functionality, such as:

- no reasonable synchronisation and versioning
- no document description available along with each document
- no history function, showing who changed what and when
- no easy, platform independent facility for navigating the document store
- no efficient and consistent way to manage the access rights for the documents

- no simple way to add comments or ranking to documents.

We decided to set up a shared workspace for the partners using the BSCW tool developed by GMD. BSCW allowed partners to use any Web browser to check the status of all documents, to upload and download files, and to add their own comments. It also provided smooth integration with word processing, database and scheduling applications on the partners' own workstations. The security features and versioning enabled the editors to maintain a complete and consistent view of the whole proposal, accessible to all the partners. With the underlying clarity and security provided by BSCW, ad-hoc e-mails and telephone calls were used just to resolve difficult points.

In less than two months, we succeeded in building a complex, high-quality proposal of 100 pages plus an administrative and financial database, all through the virtual workspace provided by BSCW. Although almost all partners were using BSCW for the first time, it was accepted and used from the beginning thanks to the intuitive GUI and the comprehensive on-line help. As a result, the consortium as a whole never had a meeting. The four partners managing the proposal held just four working meetings face-to-face. At the end of the process all the documents required by the Commission were available online in the workspace, allowing electronic submission of the proposal to Brussels within a couple of minutes.

The advanced Web technology of BSCW allowed successful and cost-effective collaboration on a pan-European scale. It reduced dramatically total travel expenses for partners and saved considerable time

for creative work. Negotiation and preparation of the proposal was done almost entirely via the Internet, and fax and mail were negligible. Moreover, the positive experience with long distance on-line document preparation and collaboration encouraged us to reduce the budget for travel and other expenses in the project itself.

Once the proposal is accepted by the European Commission, we will continue to use BSCW Web technology. This will allow a smooth and efficient transition through contract negotiations and into the management of the project itself. Furthermore, part of our project will be devoted to the further development of Web technology to manage user accounts and to ensure security in geographically dispersed high-performance computing. Results of the project will also be available in the European Grid Forum EGRID.

#### Links:

BSCW website: <http://bscw.gmd.de/>

#### Please contact:

Wolfgang Ziegler – GMD  
Tel: +49 2241 14 2258  
E-mail: [Wolfgang.Ziegler@gmd.de](mailto:Wolfgang.Ziegler@gmd.de)

# The Collaborative Web

by László Kovács and András Micsik

The omnipotent presence of World Wide Web information sources and technologies made it feasible to build groupware applications using Web technologies. At the Department of Distributed Systems of SZTAKI, three projects are dealing with

the development of collaborative Internet services integrating ideas from Computer Supported Cooperative Work: rating within a digital library, a collaborative information filtering project and an open voting Web service.

## AQUA-2, User Interface for Rating within Digital Library

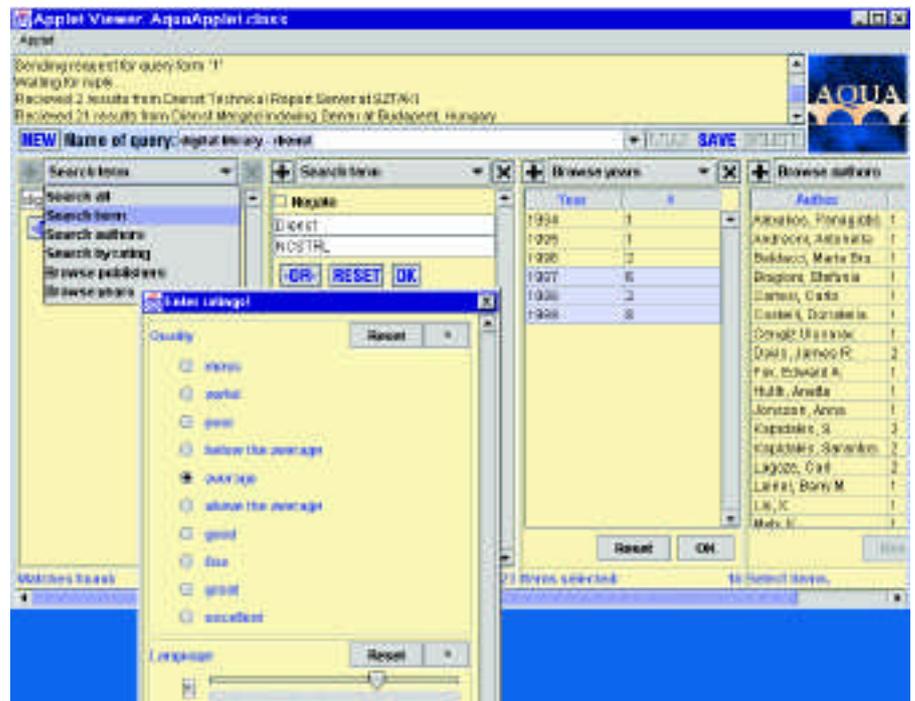
Communities Techniques for querying digital library collections ranging from searching and browsing to relevance feedback, result ranking etc. are usually combined and users construct several queries until they get the desired results or the final query expression. A query interface that eases the building of the query expression is developed within the AQUA project (Advanced Query User Interface Architecture). The first version of AQUA was integrated with the Networked Computer Science Technical Reports Library (NCSTRL) using Dienst distributed digital library system.

The system builds upon the query chain paradigm: a sequence of alternate search and browse operations to filter the result sets of the preceding operations. Individual query or search operations are displayed via a GUI panel, the query chain is given by a series of panels. A great variety of query methods is offered. During an iterative query refinement, users build query chains that can be modified/refined at any point.

The AQUA framework and environment can ease future research as well as user experiments towards the development of advanced digital library query techniques and visualization methods.

AQUA is further developed towards a collaborative tool for digital library community use. AQUA-2 helps sharing knowledge and opinions about documents while AQUA (the first version) helped only sharing queries as well as exploration strategies among named or anonymous community members.

AQUA-2 supports digital document rating. Peer rating groups (reviewers) can be organized. They can attach their own



A query in AQUA, and the panel for rating a document.

ratings to documents they have visited. Associated ratings can help other users to find the most relevant (high-)quality documents. Ratings given by the users are collected in a SELECT server (SELECT is an EC funded project for social and collaborative filtering described below). The development of AQUA-2 is partly funded by KNIXMAS INCO project supported by EC.

## SELECT, Social and Collaborative Filtering within WWW and Usenet News

Internet users need tools to find the information most valuable to them within the limited time they may have available. Information filters that make use of recommendations derived from individual users' past choices and/or recommendations derived from the behavior of other

users, especially those who have displayed similar tastes and interests in the past, are built within the SELECT project supported by EC Telematics Applications Programme.

Automatic methods where the system derives the filtering conditions from user actions or evaluations of documents using NLP (natural language processing) techniques are also being developed. Two different user interface techniques were developed: a bookmarklet-based thin SELECT client and a local proxy-based fat client. The thin SELECT client has the advantage of requiring virtually no installation and being accessible from any standard WWW browser. The proxy-based approach allows the content of a document (WWW page or Usenet News article) to be modified in different ways

presenting the actual rating information of the document as well as other kinds of quality indicators. Thus, the SELECT architecture may be used as a gateway for the associated quality metainformation. The proxy-based client may also perceive/interpret the actual user behavior and translate it into implicit rating information.

The server consists of several modules for passive filtering (recommendations are based on the submitted ratings of documents), for active filtering (recommendations are based upon the ratings of documents submitted by people with interests and/or rating histories that match those of the user), a rating database containing ratings of registered and/or anonymous users, as well as a profile database containing information about users and their interests.

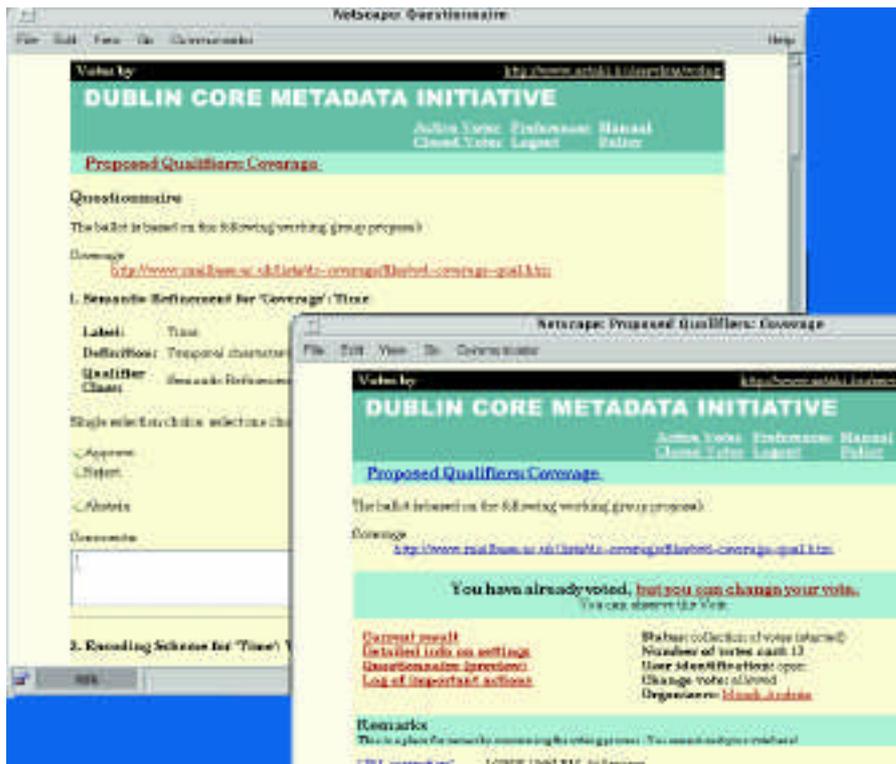
The prototype implementation of SELECT has recently been completed, user groups evaluations as well as the development of a new open SELECT framework for public testing of new collaborative filtering ideas and techniques are being initiated.

**WWW-based Open Voting and Survey Service**

The importance of computerized votes and related activities (polls, surveys, etc.) have been increasing since the globalization of the world's information societies. Computer supported decision making processes can have beneficial effects on society; these tools may be able to open and clarify decision making processes, to help the evolution of people's democratic attitudes.

A Reference Model for voting activities was defined as the first step. The voting process is split into the following subprocesses: initiation, collection, analysis, distribution, and decision making. A high-level vote control language was created for the definition of vote subprocesses.

A prototype voting and survey subsystem was developed at the Department of Distributed Systems of SZTAKI under the umbrella of Web4Groups of EC Telematics Application Programme. As a follow-up activity an Internet voting



**A vote process for the standardization of Dublin Core element modifiers.**

service was developed which has been operating as a public service of SZTAKI for more than one year allowing any user or closed group to create vote processes of any type.

Besides the basic functions of the service it has some features not found in other similar systems:

- user access via WWW as well as E-mail
- many built-in question types and evaluation algorithms
- support of multilingual votes
- user identification, user and group definition for different access permissions,
- flexible control scripts to control the vote process events,
- replaceable graphical layout.

A version of the voting system is also used in a spin-off company of Kapsch AG., which promotes the Web4Groups WWW-based groupware application.

The public voting service at SZTAKI has many applications in various types of voting activities. It served as a public opinion collection tool in polls such as a questionnaire on user requirements by European R&D projects, a survey on the

new graphical design of one of the WWW sites of the Hungarian government, surveys of a newspaper and a public library. It was used for decision polls within closed groups (selecting a logo for a digital community, a survey about a university lecture, deciding on the new name for a company, the approval of draft recommendations by the Dublin Core international standardization board, etc.). The system allows also to investigate the beneficial effect of multi-round votes in reaching the best consensus (two phase poll for the new state slogan for Victoria state (Australia) number plates).

**Links:**  
**AQUA project:**  
<http://www.sztaki.hu/sztaki/aszi/dsd/aqua>  
**SELECT project:**  
<http://www.omega.it/research/select>  
**SZTAKI vote service:**  
<http://www.sztaki.hu/services/voting>  
**NCSTR:** <http://www.ncstrl.org/>  
**ETRD (ERICM Technical Reference Digital Library) user interface:**  
<http://www.sztaki.hu:8000/>

**Please contact:**  
**László Kovács – SZTAKI**  
**Tel: +36 1 209 5286**  
**E-mail: laszlo.kovacs@sztaki.hu**

# Separation of Concerns in Multimedia Services

by Olle Olsson

**What can you do to make your customer feel good when (s)he accesses some service you offer via the web? A good approach is to personalize the service to his/her interests, and adapt the form of presentation and interaction to the situation at hand.**

The ESPRIT project MAPPA (Multimedia Access through Persistent Personal Agents) is targeted on supporting multimedia-based services adapted to the individual user. The two-year project (running 1998 - 2000) is performed by a consortium of six partners, from Sweden, UK and Ireland.

The goal of the project is to develop technology and methods that enables adaptable systems to be built. Such systems offer personalized service access, based on the paradigm of a automated personal agent interacting with the user via multimedia assets. The personal agent is able to build and evolve a model of user preferences based on the user's interactions. As the personal agent is persistent, it is able to track changes in user preferences over time.

The application goals are to provide services to registered customers to a vendor, support their explorations of available products and provide announcements of new products that match the interests of the customer.

The technical goals are to develop generic models of multimedia product presentations, of user preferences, and of the tasks that the personal agent can perform.

Part of the agenda of the project is to explore how customers can access the offered services via a range of devices and network technologies.

## Individualized Services

The broadened acceptance of the World Wide Web as a medium for access to information and services has recently been accompanied by a growth in the spectrum of commercially available devices through which the web can be accessed. Service providers are becoming aware of the need to offer services that provide satisfaction to their users, and

packaging these services in a way that makes them accessible on a wide range of access devices. The technologies range from public kiosks, desktop PCs and laptops, to palmtops and mobile phones; connected through broadband WANs, called-up connections, or wireless connections.

The emergence of XML offers relief in certain aspects; the markup languages used for information delivery can be handled within a common framework. But having a common basis for representation languages does not solve the real problem; how to optimally provide individualized services.

One of the aims of the MAPPA project is to explore ways of dynamically adapting user interaction to the context in which interaction occurs. The approach is to provide a clear separation between the semantics of the services and the way the services are presented when they are communicated across the web.

A second way of decomposing the problem is the separation of Multimedia asset representation of information from the tasks that the user performs on the basis of this information. The tasks correspond to what happens in the system when the user interacts with the multimedia assets. Information representation and presentation is basically what XML and other markup languages support. Support for tasks are typically provided through scripts embedded in the document itself, or via cgi-bin processing.

Individualised multimedia access depends heavily on dynamic adaptation at the user interface. Ideally, the full service implementation could be sent to the user's access device, eg in the form of Java applets managing/controlling both the multimedia presentation and the actual service tasks. But this is in conflict with

the desire to be able to support access from WAP-phones, as Java applets are not (yet) supported on such phones.

A goal pursued in this project is to enable service tasks to be automatically configured and distributed to the user's access device. There are a number of constraining factors influencing this, eg what representation, presentation and execution resources are available in the access device? What communication resources are available? What behavioral patterns are expected of the user? 'How do the tasks map to response-time priorities'? What are the sizes of the different kinds of data-sets that support the tasks and to what extent should data and task implementations be protected from exposure in the access device?

We have developed a set of models corresponding to different perspectives on an interactive service. These include domain models, view models, interaction models, presentation models, and action models. A representation of interactive access to a service is constructed by instantiating these models, adapting them to the constraints of the delivery environment, combining them to a complete models, and then producing a representation that can be transported to the client's access device for execution.

Service tasks are classified into a set of categories, ranging from 'interaction' tasks (tasks that the user performs while exploring the information space and preparing his fundamental service request) to 'service-effect' tasks (tasks that involve effects/actions that take place in the service-provider's environment). As an example from the application domain we are working in, typical interaction tasks are finding, inspecting, and evaluating available products, and composing the content of a 'shopping cart'. All such interactions have no semantical persistent side effects in the

vendors environment. The prototypical example of a service-effect task is submitting a request expressing that a specific set of products are to be shipped to the customer and payment drawn from his account. Obviously, service-effect tasks are extremely domain dependent, and are handled as wrapped components with a declarative description of their software interfaces. On the other hand, interaction tasks typically exhibit certain kinds of patterns, which makes it possible to model these in a semantically more meaningful way rather than programming them.

Interaction tasks are modeled declaratively, at a level of granularity that enables the desired service behavior to be achieved on the user's access device. To each task there is an associated description of how this task relates to and has effects on the information model underlying the service. This description is expressed in a representation that can be transformed into a form that can be executed or interpreted in the presentation environment on the device used.

The structure of the information model maps to/from XML, which enables domain data and multimedia assets to be generated, operated on by generated/configured tasks, and presented using standard browsers. The delivery environments experimented with include standard web-browsers with HTML or XML capabilities and JavaScript or Java applet support, WAP-phones with WML and WMLScript, as well as an experimental Java-based presentation system using a flexible protocol for interacting with server-based functional components. An XSL-based transformation system has also been evaluated (Cocoon), to determine to what extent this type of standard support tools can be used for producing the actual representation.

The rapid evolution in commercially available access technologies that we see today, introduces problems that will persist for a long time. New technologies enable new patterns of behavior, which will change the expectations of users, and this will change the ways in which people

actually will take advantage of the fact that they will be 'continuously on-line'. Technologies and services should adapt to the user, not the other way around, and the user will expect the services to be more 'intelligently interactive'. This will take us away from the old paradigm of the Web as a document presentation technology to the new paradigm of the network (whether wireless or not) being a platform for intelligent interactive distributed services. This means that we will need techniques and methods for smart dynamic distribution of a set of adapted tasks that together provide a set of services to the user.

#### Links:

Mappa website:  
<http://www.sics.se/mappa>

#### Please contact:

Olle Olsson – SICS  
Tel: +46 8 633 15 19  
E-mail: [olleo@sics.se](mailto:olleo@sics.se)

## Increasing Security in Virtual Enterprise Communication for Mobile Environment

by István Mezgár, Tamás Szabó and Zsolt Kerecsen

**A research project to develop a secure communication architecture for mobile communication in virtual enterprises started at SZTAKI in 1999, supported by the Hungarian Scientific Research Fund (OTKA). The operation of virtual enterprises is based on the Internet,**

**thus a large amount of valuable technical-related data moves on the network. Since the management of virtual enterprises will be controlled also from mobile devices in the near future, security problems of mobile communication are of vital importance.**

Mobile devices (mainly mobile phones) can be used today for more purposes than just as making phone calls. Additional functions like mobile Internet browsing, ATMs for banking transactions, etc. will be available for daily use and virtual enterprises are candidates for using these types of new functions extensively.

A virtual enterprise (VE) is a co-ordinated network of autonomous production units (factories, firms) with the goal of producing a product while exchanging all needed information via a computer network. The lifecycle of a VE - formation,

communication, operation - is strongly connected with the Internet. During communication in a virtual enterprise, a huge amount of extremely valuable technical data and information (development-, product-, process data beside business information) moves through the network, making security a vital concern. Although there are already many applications available, the lack of proper co-operative, effective and secure software tools in this field delays the massive spread of real VEs.

Today, e-commerce represents the best way in efforts of convergence and globalisation. From an information point of view, the future of VEs moves in the same direction. E-commerce is the 'fore-runner' of VE, there are many similarities in their functions and in the techniques they use.

Increasing the security level of virtual enterprise communication in a mobile environment is a vivid demand. Enhanced security can contribute to the growth of the number of realized VE solutions. In the future, management activities in VE



Figure 1. The WAP protocol Stack.

will be accessed and controlled from mobile devices as well. This access has to be very secure, and a general tool for safety could be the smart card (SC). Smart cards can be used for a variety of purposes, such as cryptography, identification (combined with biometry), authentication (digital signature).

The aim of our project is to define and describe a software architecture that properly addresses mobility and security related needs of VEs. This includes collecting a protocol set where security is based on smart-card technology and which provides mobility in accessing virtual enterprise services.

#### The Applied Techniques

In order to fulfil the compatibility requirements, in developing the new software architecture, standardized communication technologies have to be used. The most widely used protocol to deliver the Internet content is HTTP, however, it does not behave in accordance with the mobile environment. Therefore, a new protocol - a de facto standard - has been defined for the wireless world, called WAP (Wireless Application Protocol), which makes reliable and fast access of data and services possible.

Wireless carriers are known for their low bandwidth, high latency, and unpredictable reliability. Wireless devices represent limited processing power, memory and simple user interface. All these limitations are addressed by the WAP, making it the most suitable platform for supporting sophisticated telephony and information services on hand-held devices. The most important parts of the WAP protocol stack are outlined in Figure 1.

The security layer in the WAP is WTLS (Wireless Transport Layer Security). This protocol is based on the TLS protocol, which makes client-server authentication possible and performs cryptographic operations. In addition, the application level security can be accessed using the WMLScript (Wireless Markup Language Script). However, in order to address most of the security requirements of the customers, some of the security functionalities have to be performed in a tamper-resistant device. For this purpose WAP uses WIM (WAP Identity Module), usually implemented by a smart card, possibly together with the SIM (Subscriber Identity Module) card.

According to WAP, not only a mobile client is able to access the services, but also the server side is no more localised, so needs of the mobility-based new VE concept are fulfilled. In the case of smart card applications a new technology is smartX that defines a complete framework to develop smart card applications. By separating the application process (the logic of the application) from the application protocol (card-specific layer), smartX makes quick and efficient migration to a new smart card possible (see Figure 2).

SmartX relies on SML (smartX Markup Language), a description language that describes the smart card application data and processes. SML leverages XML technology with a DTD (Document Type Definition), which describes the semantics and the grammar to define application protocols. SML is an implementation of XML for the smart card industry.

Our activities in the close future will include the description of VE functions

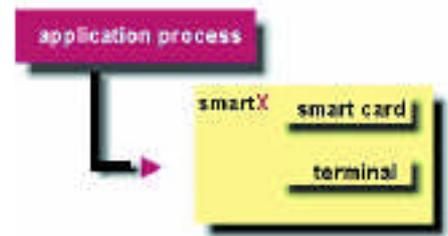


Figure 2. SmartX encapsulates both the smart card and the terminal application.

in XML, since HTML based Internet is moving towards XML. The description of VE in XML including the enterprise resources, data transactions and product flows is a new approach, it will make far more easy the building of the connection between the content and the communication channels. By combining WAP and smartX technology we obtain a very flexible, secure mobile architecture which together with VE represented in XML will provide a safe business communication environment.

#### Conclusions

HTML and HTTP/TCP standards are not appropriate for mobile communication devices. The WAP and WML have been developed to bridge this problem in such way that wireless devices can be connected to traditional WEB servers, tools and techniques. Secure communication is very important in wireless communication, and as the management of virtual enterprises moves towards the mobile world, the development of a secure mobile communication architecture is a vivid demand. In order to reach a flexible and adaptable communication infrastructure the representations will have to be written in different XML-based languages.

#### Please contact:

István Mezőgar – SZTAKI  
Tel: +36 1 209 6143  
E-mail: mezzgar@sztaki.hu

Tamás Szabó – Nokia Hungary Ltd.  
Tel: +36 20 984 9074  
E-mail: tamas.szabo@nokia.com

Zsolt Kerecsen – BablakSoft Ltd.  
Tel: +36 1 327 5821  
E-mail: zsolt.kerecsen@scala-rnd.com

# Creating Hypertextual Dossiers in XML for the Italian Parliament

by Andrea Marchetti and Patrizia Andronico

During the last two years, a working group at the Institute for Telematic Applications of the Italian National Research Council (IAT-CNR) has been following the evolution of XML-related technologies. This group was requested by the Italian Chamber of

Deputies to collaborate on the creation of a website to be presented during the course of the Annual European Parliamentary Technology Assessment (EPTA) Conference held in Rome, December 1999.

The EPTA Conference signalled the introduction of new high-tech equipment into the Italian Chamber of Deputies. A meeting-room for members of the Italian parliament was equipped with desks, each with their own built-in PC and Internet access. The Conference website provided access to the texts of the speakers' presentations and documentation on the themes of the Conference. Links were built between the various documents on the basis of a thesaurus, consisting of topics identified by experts in the fields of interest.

For each term of the thesaurus, documentation was created containing:

- a definition
- associated terms in the thesaurus
- one or more quotations
- pointers to related websites
- bibliography.

This documentation had to be collected, processed and rendered homogeneous in a very short period (approximately 10 days) by people working in different geographical areas. A requirement on the application was that it should allow real-time processing on the documentation, in

order to be able to extract phrases from the texts, at the moment of their presentation, and insert them as enrichments of the thesaurus entries.

For each of the topics identified, the logical schemata used to structure the material made available (presentations, articles, thesaurus with its associated documentation) were defined in the form of XML DTDs. All the material was edited in XML and experts then inserted further markups to signal thematic associations between different items. The material was then sent to IAT for validation and for an initial processing aimed at transferring information extracted from the thesaurus and inserting it in the documentation associated to the terms of the thesaurus. Selected images were also associated with this documentation.

At present, very few browsers allow the visualisation of XML pages. For this reason, all of the material produced was translated into HTML and inserted into the website.

## Layout of the Website

The information present on the browser had to include:

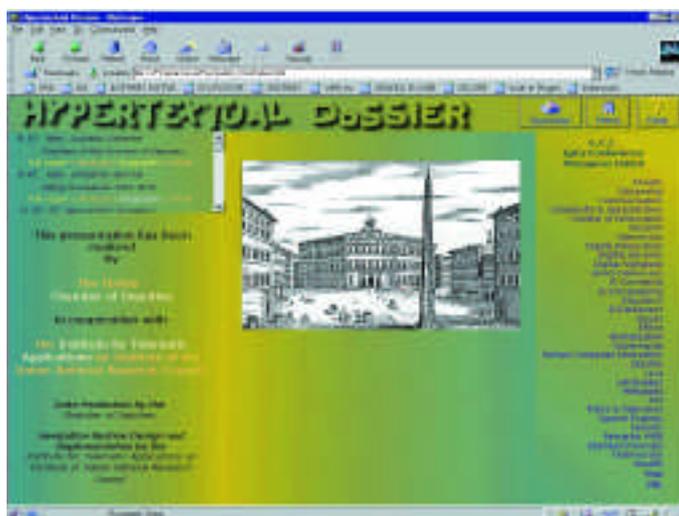
- the programme of the day;
- the documents available for each speaker (full paper, abstract, bibliography, others), directly accessible from the programme;
- the index of terms from the EPTA Conference Thesaurus;
- the associated documentation.

This requirement led us to adopt a frame architecture. Each frame contained a type of document. The frames were organised so that the page was read in a circular direction, following the order given in the above list. At the centre of the page, the thesaurus documentation frame formed the central point of navigation.

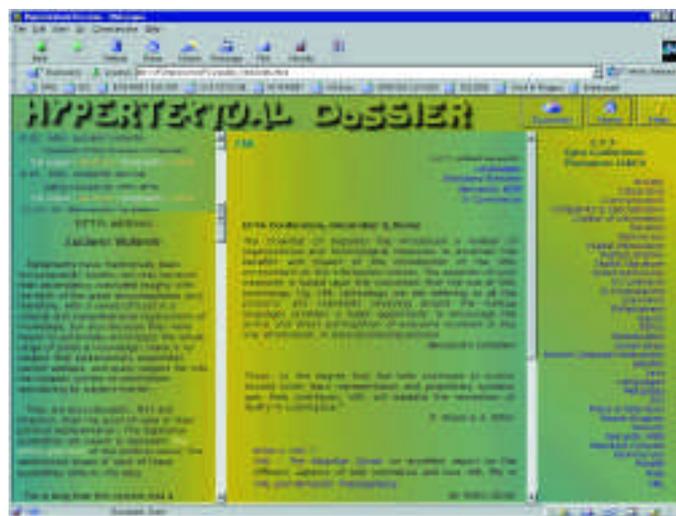
For each document type, a specific CSS stylesheet was defined.

## Conclusions

This application involved defining a complex model to organise the information. The IT and document experts worked together on the development of this model: the IT operators were involved in the choice of



The Home Page of the Hypertextual Dossier.



The Hypertextual Dossier during navigation.

documentation, and the documentation experts were similarly consulted with respect to the choice of technologies. This integration contributed greatly to the success of this experimental application; the website was set up much more quickly than could have been expected.

The site has functions (such as the consultation of the documents in various languages) which have not yet been activated, partly as a result of lack of time and partly because available browsers do

not yet fully support XML and associated standards. Currently, the conversion into HTML is performed off-line. In the near future it is expected that it will be possible to develop a version of the site in which the documents remain in XML format, being translated into HTML only when requested by a browser that does not support XML. The results of the XML project Apache may be used for this. The site can be consulted via Internet at: <http://epta.camera.it>, selecting the link 'Hypertextual Dossier'

**Links:**

**EPTA:**

<http://www.atkinsoft.ndirect.co.uk/epta/>

**Apache project:**

<http://xml.apache.org>

**Please contact:**

**Andrea Marchetti or Patrizia Andronico – IAT-CNR**

**Tel: +39 050 315 2649 or +39 050 315 2090**

**E-mail: [andrea.marchetti@iat.cnr.it](mailto:andrea.marchetti@iat.cnr.it) or**

**[patrizia.andronico@iat.cnr.it](mailto:patrizia.andronico@iat.cnr.it)**

## Providing the Context for Business Decisions

by Brian Matthews

**Information is the major asset of business today. However, if there is insufficient knowledge about the business context in which information is supplied, then the user may not be in a position to judge its**

**relevance and as a consequence, may not be able to make correct business decisions. The problem is to provide the correct information to the correct user at the correct time.**

Information Systems Engineering at CLRC is working with a major international communications company to explore ways to enhance the dissemination of the corporate knowledge assets. As the Web has become a universal medium for both external and internal communication, it is natural to use it for this project. The emerging Web technologies, especially the Extensible Markup Language (XML) and the Resource Description Framework (RDF) offer a powerful means for managing and distributing data as well as the context in which it is presented.

The company maintains a single data model for corporate information. This forms the basis for generating particular data models to be used for particular situations, such as a product management system, marketing system, or customer handling system. By conforming to a common basis, the exchange of information between these systems becomes easier, allowing rapid and consistent information to be transmitted. However, maintaining and using this model is a real problem. If the model is hard to understand and explore, then it is hard to persuade systems developers to use it.

The project is using an XML representation of this UML model to transmit it via the corporate Intranet. This gives the opportunity to present it to the user within widely available XML enabled browsers, providing a universal and economic interface onto the data model. The use of an XML format combined with the scripting languages such as JavaScript, and the Extensible Stylesheet Language (XSL) allows the user to move between different views of the data within the browser, without reference to the server. This allows the possibility of user profiles providing customisable views. Further, the data model can also be modified in XML and transmitted back to the central repository. Thus an easy to use editor can be distributed to the user via their desktop browser.

As the model is developed, decisions are taken on the design of the model. These decisions and the reasons behind them are vital to the understanding of any future user. RDF is being used to provide such contextual information, which is supplied to the end user who can then discover the context in which the model has been designed. Thus, links to explanatory text, constraints, and company standards can be provided. Such explanations can also

be in audio or video formats. The RDF description can also be used to apply constraints on the data model, or constraints on the contextual information itself. As the data model becomes subject to extension and restriction according to the circumstances in use, additional context can be provided to track the changes and record the additional reasoning behind its evolution. Thus an audit trail can be established through the variants on the model. Again all this can be presented and edited across the WWW using standard browsers.

**Please contact:**

**Brian Matthews – CLRC**

**Tel: +44 1235 44 6648**

**E-mail: [B.M.Matthews@rl.ac.uk](mailto:B.M.Matthews@rl.ac.uk)**

# Virtual Physics and Astronomy

by Detlef Skaley and Udo Zlender

**The Virtual Physics and Astronomy (VPA) project is a Web-based innovative learning platform, with multimedia capabilities and intuitive user interfaces, developed at GMD. It is intended for schools, universities, or self-study to learn physics and astronomy.**

Today's life requires accessing information and knowledge in relatively short time scales, which means new methods of knowledge transfer must be established. The Internet has brought forward enormous technical possibilities which not only can compete in many ways with classical text books, but also provides new grounds through its multimedia capabilities. Visualisation has always been regarded as an important element to aid in understanding complex scientific theories and experiments which up to recently required rigorous special soft- and hardware requirements. Nowadays, however, graphics, videos, animations, together with appropriate description texts can be accessed directly on the Internet using an ordinary personal computer. Moreover, new learning tools can be established in combination with exercises. In this way the student can learn by experimenting in the virtual worlds. For instance measurements with a simple stop watch could be sufficient to carry out some of these exercises such as computing the mass of a planet.

The project Virtual Physics and Astronomy (VPA) is an innovative learning platform, primarily considered for schools, universities, or self-study to learn physics and astronomy. Through its multimedia capabilities and intuitive user interfaces, VPA enhances the learning effects. The latter is further amplified when visualisation is combined with interactivity which allows the learner to explore the virtual environments individually. Next to 3D representation accurate programming of scientific simulations are essential for such a learning platform. VPA simulations are carried out in Java, hence a standard Web browser, and some additional and freely available plug-ins are the only requirements for using the program on the Internet.

## Examples

The virtual optical bench provides examining the laws of geometrical optics. The student may place lenses, mirrors, screens, and light sources on a holding bar to study the light rays in detail.

Interesting experiments can be performed by modifying the optical properties of complete instruments by changing the properties of its elements eg, changing the refraction indices or radii of curvature of the lenses (see figure 1).

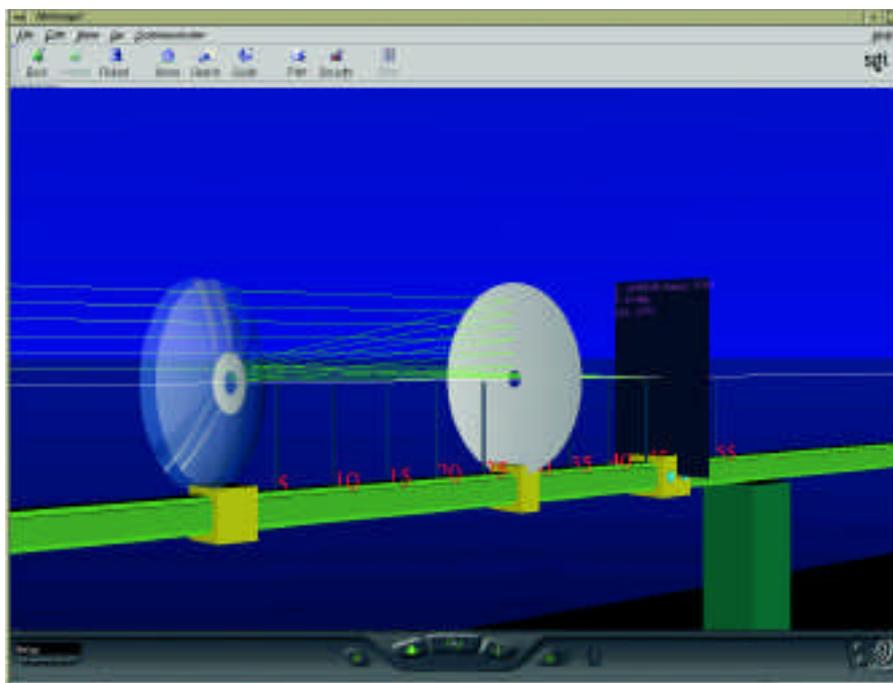


Figure 1: Maksutov telescope on virtual optical bench.

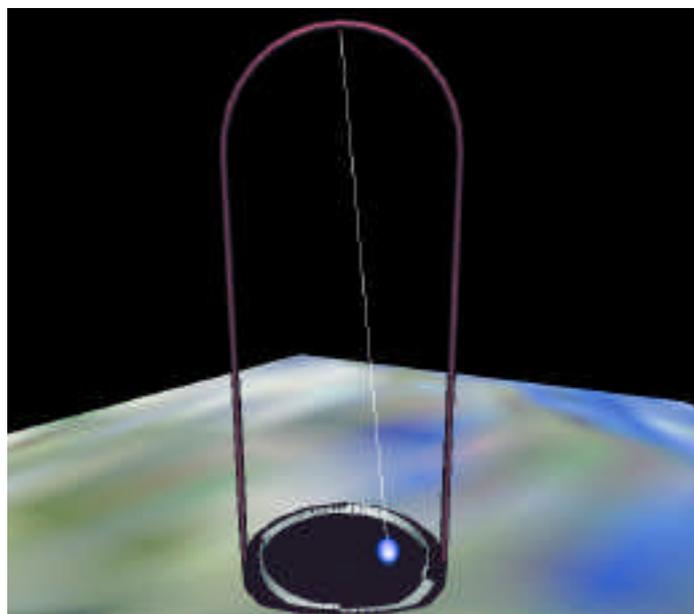


Figure 2:  
The Foucault  
Pendulum.

Physical experiments like the one performed by Jean Leon Foucault in 1851, by which he proved the earth's rotation can be understood using the Virtual Reality Modelling Language (VRML). Virtual Foucault's Pendulum may aid the user to determine the geographical latitude by measuring the rotation velocity of the pendulums plane of oscillation (see figure 2).

The Virtual Planetarium allows facts of our solar system to be explored. By navigating freely in 3-dimensional space, the student can study the planets' orbits

relative to the ecliptic plane from different viewpoints. In addition calculations of physical laws in real-time enable the realistic representations of the planets' motions. The latest development of this software provides Java3D stereoscopic 3D-presentation using recent-of-the-shelf 3D-graphics hardware (see figure 3).

Discovering supernovae explosions, neutron stars, black holes, or quasars are some of the outcomes of modern astronomy. Virtual astronomy can help to illustrate and explain some of the exciting cosmic facts. The combination

of real images, videos and VRML models of those examples clarifies the important results from that recent field of research. In here a student might be instructed to observe and measure the distance of a galaxy from the oscillation period of a cepheid star. Or perhaps consider two penetrating galaxies, where the mass of the inner galaxy can be determined from the rotational period of the outer galaxies' debris. The examples mentioned here can be found among many others under the following URL: <http://imk.gmd.de/delta>.

### Techniques

Although the project consists of a wide range of different topics in Physics and Astronomy, the underlying structure of each individual application is very similar. Most of them consist of three major parts:

- **Computer Simulations and Calculations:** Are required to give insight into complex dynamical processes. They might be written in any desired programming language like C, C++, Java etc.
- **Network Connection:** Interfacing between client (visualisation front-end) and server applications (simulation providing site) can be realised by Common Gateway Interface (CGI, Perl) or Remote Method Invocation (RMI, Java).
- **Visualisation and Animations:** Are done local on the users machine in Java, Java3D or with help of the Virtual Reality Modelling Language (VRML). Static examples are just downloaded and displayed (VRML), animation of parts of the scene require Java, or Java in combination with extensions to VRML eg External Authoring Interface (EAI).

Promising future developments like MPEG-4 will be used as they become generally available.

### Links:

<http://imk.gmd.de/delta>

### Please contact:

Detlef Skaley – GMD  
Tel: +49 2241 14 2987  
E-mail: [Detlef.Skaley@gmd.de](mailto:Detlef.Skaley@gmd.de)

Udo Zlender – GMD  
Tel: +49 2241 14 2485  
E-mail: [Udo.Zlender@gmd.de](mailto:Udo.Zlender@gmd.de)

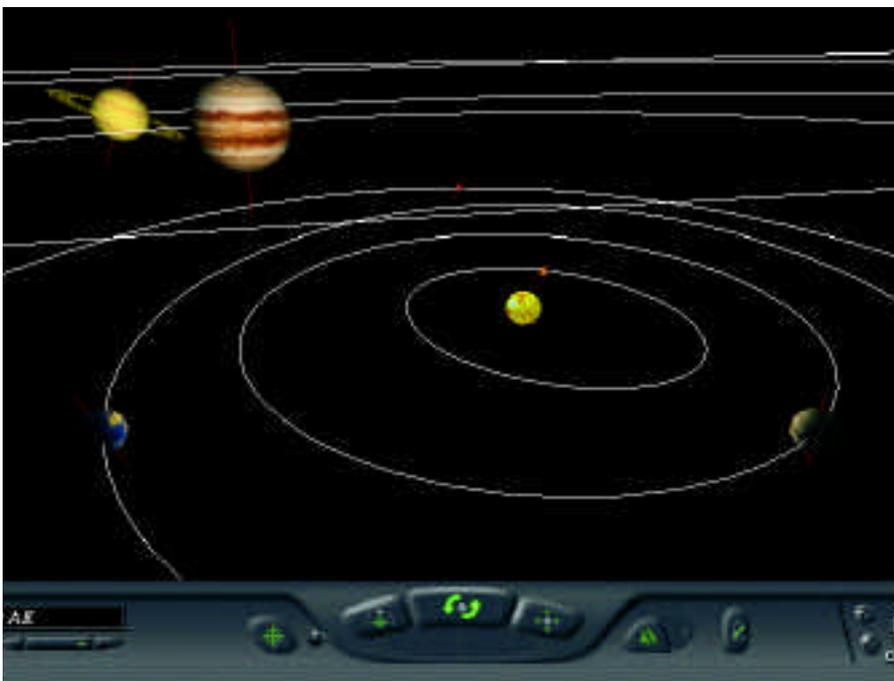


Figure 3: Virtual Planetarium.



Figure 4:  
Virtual  
Astronomy.

# CommonGIS – Common Access to Geographically Referenced Data

by Hans Voss, Natalia Andrienko and Gennady Andrienko

**Public access to the immense volume of existing geo- and geo-referenced thematic data and their exploitation is of significant value for the development of an open and democratic ‘information society’ and a true global market. The widespread use of such data and GIS (geographical information system) will promote general public awareness and further social cohesion. Publicly**

**available geo-data is, however, of little use unless people can easily access and easily exploit it. CommonGIS, a research project coordinated by the Knowledge Discovery Team of GMD Institute for Autonomous Intelligent Systems, is developing Web-based tools for access and analysis of geo-data that can be utilized likewise by skilled and casual, non-expert users.**

Geo-data encompass various thematic or statistical data on demography, economy, education, culture, history, etc., which are associated with objects and locations in space. Probably the best way to explore such data is to visualize them on maps. The key-thought of CommonGIS (<http://commongis.jrc.it>) is thus to make geo-data commonly accessible and usable for everyone, from everywhere, by providing a Web-based Geographical Information System (GIS) with knowledge-based functions for the automatic generation of thematic maps. To a very large extent, the user should not at all need to worry about visualization issues but rather focus on the problem solving process and analytical task.

Ordinary GIS are not well suited for this because they suffer from at least one of three intrinsic pitfalls: they tend to be architecturally closed, monolithic, and costly environments; their use requires specific technical skills; they require the user to think more about obtaining visually nice presentations than about the selection of appropriate data and solving problems with their use. The CommonGIS project is now (as of February 2000) in its 16th month, and current achievements include:

- A data characterization schema (DCS), and a data characterization language (DCL) as an instantiation of the schema has been defined. The DCL is used to describe properties of and relationships between attributes of a set of thematic data set that shall be published. The semantics, in particular, is utilized to further automated support of data visualization and analysis.
- A prototype for the support of building new applications (application builder) has been realized, and the resulting application modules comprising a new

Web-based, intelligent GIS for automated thematic mapping have been implemented.

- Several demonstrator applications showing the potential of the new technology to data providers and users were put in place.

In the following we will address these achievements in some more detail. A basic requirement was that CommonGIS should easily connect to existing data sets, which could be stored in large databases or kept in spreadsheets, just to name extreme options. For this purpose, CommonGIS provides a set of converters for various table data formats as well as data base connectivity. Internally, all geographic and thematic data are stored in a data base (ODBC/JDBC interface). Aside from this base functionality, the proper goal of providing smart support also to non-expert users can only be satisfied if the application program itself incorporates knowledge: knowledge about the data of the specific application, and knowledge about principles of cartographic visualization and analysis.

As basis for the latter we utilize the knowledge base for cartographic visualization from Descartes (<http://borneo.gmd.de/and/java/iris/>), which was developed at GMD's AiS institute (see ERCIM News, July 1998). In addition, some ideas are being incorporated from the cartographic visualization system VIZARD from the Fraunhofer Institute FhG-IGD. Regarding knowledge about the application data, the original Descartes system also included some means to describe the semantics of the data. However, the language used for defining characteristics of and relationships between attributes was somewhat ad-hoc, not documented, and its usage was not

supported by suitable editors and checkers. So one specific task of CommonGIS was to develop a language that is applicable to any domain of spatially related thematic data, and to provide tools that make it easy for data providers to build their own applications.

As a result, a so-called data characterization scheme (DCS) was developed. In contrast to other work known from literature, the DCS provides a rich arsenal of concepts, data structures and operators for defining semantic domain models of the given data. The reason for this particular emphasis is that the focus of previous work was mostly on the visual presentation of data on then static maps, while in CommonGIS we are dealing with very interactive maps. In addition, we want to support users in handling more complex analytical tasks. For example, in the context of exploring some demographic data the CommonGIS system should become able to automatically identify, formulate, and support analytical tasks like “Compare gender structure of population in different age groups” or “Look at the distribution of a specific age group (0-14 years, 15-64 years, or older than 64) across the countries”. In the latter example, it could also propose that using relative values (age group in percentages of the total population) would probably make more sense than using absolute numbers.

The DCS was defined using UML, and was studiedly kept on a conceptual level. It can thus be used as a schema that can be instantiated by different syntactical variants. For use in the CommonGIS software one such language was XML, and corresponding parsers were generated. This instantiated schema is called DCL (data characterization language).

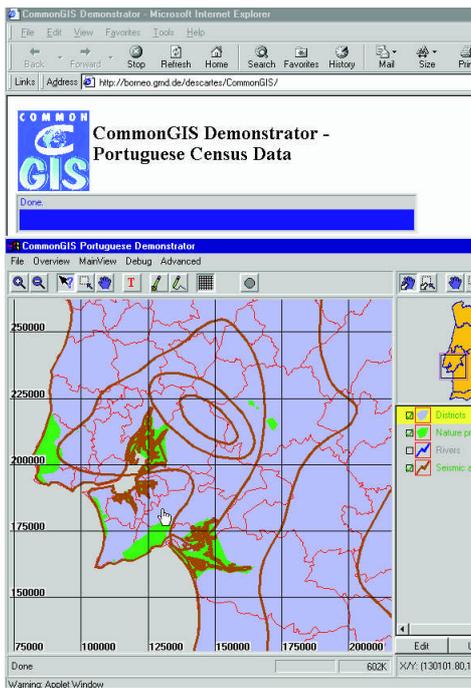


Figure 1: User interface of CommonGIS in the Internet.

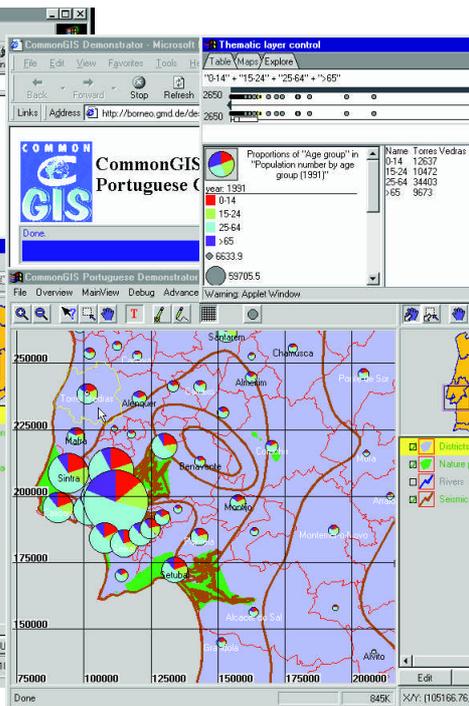


Figure 2: Pie charts are generated for showing the age structure of the population.

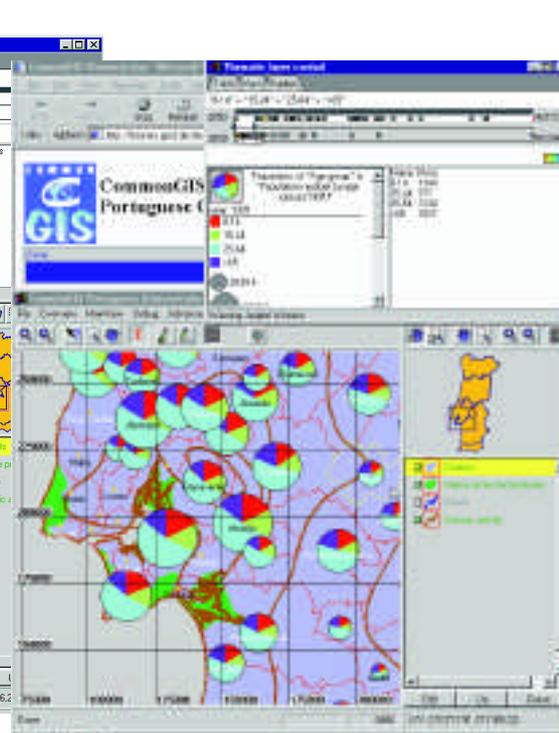


Figure 3: Pie charts: focus on areas with relatively small population.

The CommonGIS software is developed as an open, object-oriented, distributed system with a client-server architecture. The user interface is realized as Java-applets, thus providing comfortable access and interactivity while only requiring a standard (Java-enabled) Web-browser. Currently, what the user sees after starting CommonGIS in the Internet looks like something that one would expect to see when running a typical GIS on a local PC (see figure 1). In fact, what was taken as the base technology is the Java-based Lava/Magma GIS from PGS (see. <http://www.pgs.nl>). A salient feature of Lava/Magma is that it performs sophisticated caching of data for optimization of performance. This is particularly needed when using larger maps that would take too long to download as a whole. One achievement of CommonGIS was to make the Descartes system running from within Lava/Magma. The user can thus select and define a desired map with ordinary GIS operations, and then select certain thematic variables for visual display on this map. Thereby the full interactivity of Descartes stays available as the systems were redesigned so that Descartes can use the Lava/Magma display methods for presenting the results of interactive manipulations.

A demonstrator of CommonGIS using Portuguese Census Data is available in the Internet. It may be run from either web site, [http://commongis.jrc.it/commongis/sw/commongis\\_first/CommonGIS/](http://commongis.jrc.it/commongis/sw/commongis_first/CommonGIS/) or <http://borneo.gmd.de/descartes/CommonGIS/>. After the applet is loaded, it connects to the server, retrieves information, and displays a map. The HTTP protocol is used for communication between the client and the server. This allows the system to run in the Internet or in Intranets without being disabled by firewalls. If a geographic layer is associated with thematic information, this information can be visualized on a thematic map. For this purpose one pushes the 'T' (Thematic data) button. In response the system will list all available thematic variables. After the selection of one or several variables CommonGIS will propose one or more suitable visualization methods that are selected on the basis of the available semantic information about the data.

Using the Portuguese example, one may select four variables with population numbers of different age groups. CommonGIS will take into account relationships among the variables: it knows that they are comparable and all together make the total population. On

this basis, one of the proposed visualizations is a pie-chart presentation (see figure 2). The sizes of circles are proportional to the total population in the municipalities of Portugal, and the sizes of the segments show proportions of different age groups.

From the sizes of the pie charts one can observe that population numbers in the capital, Lisboa, and in its surroundings are much larger than in other districts. This makes the signs in those districts too small for seeing proportions of different age groups. As a remedy, all visualization techniques are supplied with specially designed controls for the interactive manipulation of the display. In the example one may decide to not show signs for the largest districts just by moving a slider, and thereby implicitly enlarge the small signs so that their structure becomes better visible (see figure 3).

Currently, two pilot applications are running. The above mentioned one was developed by CNIG (Centro Nacional de Informacao Geografica, Portugal), which is also responsible for the overall usability analysis of CommonGIS. Another application is under development by Dialogis GmbH together with the City of

Bonn. Other applications for customers of PGS and InGeoForum (which is affiliated to ZGDV – sister institution to FhG-IGD – and to the Hessian State Surveying office) will soon follow. The two industrial partners of the consortium (PGS and Dialogis) are developing and will commonly market a commercial version of the software. They will also incorporate results of CommonGIS in their current commercial products, Lava/Magma and DialoGIS, respectively, where DialoGIS is a commercial version

of Descartes. Furthermore, the consortium is taking efforts towards the standardization of the data-characterization schema. It is expected that a standardization in this field will promote ‘intelligent’ GIS software interoperability, thus broadening a potential market. The Joint Research Center of the European Commission EC-JRC, Institute ISIS, is co-managing the whole project effort with GMD, and together with its subcontractor GISIG, will be instrumental in the dissemination and standardization process.

**Links:**

**Homepage of the project:**  
<http://commongis.jrc.it/>  
**Demonstrators:**  
[http://commongis.jrc.it/commongis/sw/commongis\\_first/CommonGIS/](http://commongis.jrc.it/commongis/sw/commongis_first/CommonGIS/)  
<http://borneo.gmd.de/descartes/CommonGIS/>  
**Further publications about the project:**  
<http://borneo.gmd.de/descartes/>

**Please contact:**

**Hans Voss – GMD**  
**Tel: +49 2241 14 2532**  
**E-mail: [hans.voss@gmd.de](mailto:hans.voss@gmd.de)**

## StraDiWare: Linking Business and IT Models using Graph Theory

by Tony Conway

**The aim of StraDiWare is to formalise the linkage between quantified business objectives, decisions to implement IT systems and the IT systems themselves. The delivered system will also provide formal linkage descriptions between dependencies within IT projects.**

**This will be achieved by a heterogeneity-reconciling linkage mechanism between the various tools and information representations. The use of graph language formalism is seen as the essential feature of this mapping.**

At present IT systems rarely satisfy the aim of supporting business objectives: there is thus a high level of business dissatisfaction. The StraDiWare project aims to reduce this by coupling business objectives to IT support. There is a large and growing market for analysis and recommendation of remedial actions for pre-existing (‘legacy’) IT systems and for the development of new systems strictly controlled by quantified business objectives.

Thus the StraDiWare project complements Action 18 "Euromethod pilot projects in the CEE countries" in the PECO Action Plan by providing the essential linkage from IT projects (managed using Euromethod) to higher level business objectives and processes (outside the scope of Euromethod). Nevertheless, StraDiWare is not simply an extension of Euromethod: it is equally applicable to other IT development methodologies, such as SSADM.

StraDiWare is concerned with the business process in which enterprise objects interact with the implementation of IT projects: this is referred to as Strategic Planning (SP). SP is a rather complex task that has several basic stages and which produces various products. It

is typically a long duration process that is at present partially supported by several discrete methodologies. The benefits of using the StraDiWare tool set and methodology will be that business and IT managers can see at a glance:

- the impact of a change in business objectives on existing IT systems and planned future developing IT systems
- the effect of change in the design / implementation of an IT system on the achievement of business objectives through a multi-media representation supported by a formal model. This will provide monitoring and control of the developing IT project in terms of quantified business objectives.

The proposed delivery mechanism for the methodology as realised in the StraDiWare software product range will be via Web technology. The universality of browser availability and advances in technology recommendations from W3C (notably XML and RDF) combine to make this an appropriate and cost-effective delivery mechanism. StraDiWare will use this advanced multi-media user interface to produce outputs that are easily understood by non-expert users. This is necessary so that business managers, who may not be familiar with the technical aspects of IT projects and

their development methodologies, can nevertheless use StraDiWare to understand the linkages between business objectives and IT projects. A major goal of the project is to produce an implementation of the CASSP (Computer Aided Support for Strategic Planning) software tool.

StraDiWare brings together appropriate end-users in realistic business situations in the IT support development cycle with IT SMEs in both CEE and EU countries supported by one CEE and one EU research Institute each with appropriate (commercially tested and developing) expertise to transfer to the IT SMEs.

StraDiWare is a Research & Technological Project supported by the Commission of the European Union under the Copernicus scheme.

**Links:**

**StraDiWarehome page:**  
<http://www.itd.clrc.ac.uk/Activity/StraDiWare>

**Please contact:**

**Keith G Jeffery – CLRC**  
**Tel: +44 1235 44 6103**  
**E-mail: [K.G.Jeffery@rl.ac.uk](mailto:K.G.Jeffery@rl.ac.uk)**

# Computer Modelling in Geosciences

by Stanislav Saic and Josef Jeřek

The progress made in computer technology during the recent years allows to solve more difficult tasks also in the field of geoscience. Many new results can be obtained by reconsidering old mathematical/

numerical solutions and using powerful computers. A typical example is the three dimensional dynamical modelling of rigid objects moving in viscous fluid and the flow around them.

In ductile shear zones and in metamorphic rocks in general, the geometry of fabrics of rocks is a valuable source of information on the kinematics and dynamics of flow. In order to understand the fabric development, structural geologists use numerical or analogue modelling of the behaviour of rigid objects embedded in a viscous matrix. Numerical models are based on continuum mechanics and, with more or less accuracy, describe the motion of a rigid body and/or the flow around it.

A pioneering work in this domain was carried out by G.B. Jeffery in 1922, who developed equations of the rotation of an isolated rigid ellipsoidal particle in a slowly moving Newtonian viscous fluid.

Analytical or numerical integration of these equations give the trajectory of a particle in a flow described by a velocity gradient tensor. Jeffery's work has been broadly used and discussed by those who studied the development of shape-preferred orientation in rocks. Jeffery also developed equations that can be used for numerical evaluation of the surrounding velocity field. It is surprising that the equations have not been used for this purpose up to nowadays. The possible reasons are a rather complicated procedure of development, the need for an efficient algorithm for the numerical evaluation of elliptic integrals in Jeffery's formulas and a possibility of computer implementation of the whole problem. Moreover, the original manuscript of

Jeffery contained some errors (most probably misprints) that did not affect the equations of rotation of the inclusion, but significantly affected those for the surrounding velocity field. We present a procedure based on the equations of Jeffery that were re-developed and corrected. We found an efficient way to compute the elliptic integrals that are needed for the solution and produced software for computer modelling and visualisation of the problem. In the set of three figures we demonstrate a result of modelling with an application to structural geology.

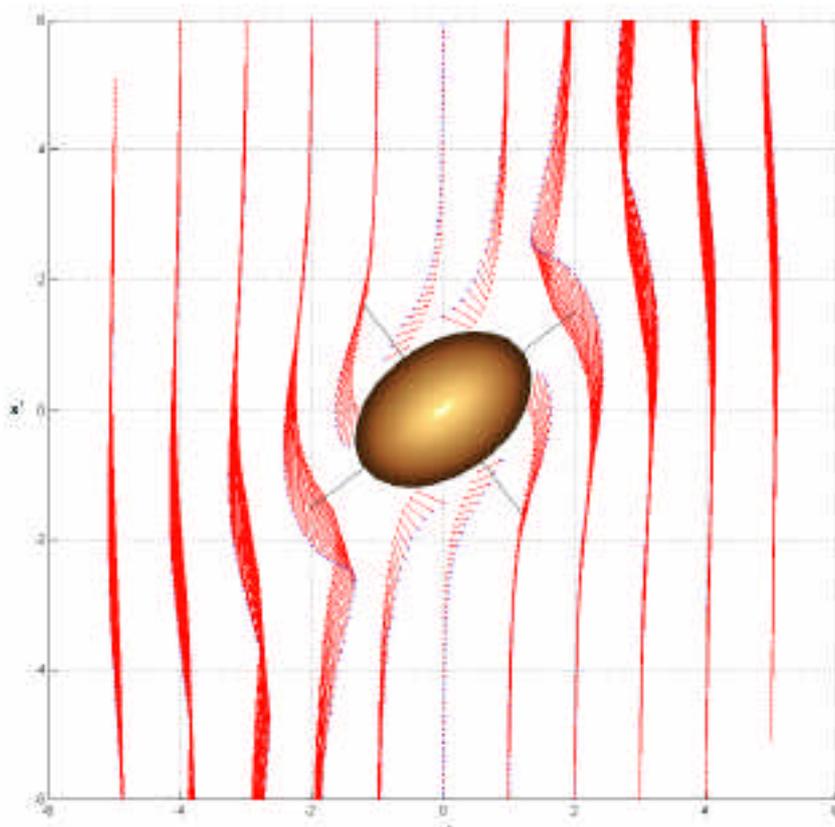
The porphyroblast is approximated by a growing ellipsoidal inclusion rotating around its longest axis parallel to the rotation axis of simple shear. Markers representing external foliations were initially arranged in lines and moved with the viscous fluid until they were captured by the growing porphyroblast.

We conclude that, in comparison to recently used procedures, the procedure and computer program we have developed represent a fully three-dimensional hydrodynamical model. It permits considering ellipsoidal inclusions and a general type of flow, and therefore to study how the asymmetry of the inclusion and different kinematic conditions influence flow patterns and structures around rigid objects in a ductile matrix.

The research was supported by the Grant Agency of the Czech Republic under Grant No. 102/96/0419.

Please contact:

Stanislav Saic – CRCIM  
Tel: +420 2 6605 2211  
E-mail: ssaic@utia.cas.cz



A two-dimensional cut through a rigid ellipsoidal particle rotating in a viscous flow in a regime of simple shear. Arrows indicate the velocity field in a set of chosen markers (point) of the moving fluid.

# Efficient Use of Networks through equal Bandwidths

by Annette Kik and Henk Nieland

**Network providers offering integrated services aim at optimal use of the available capacity. The best method to satisfy the very diverse user requirements also when the network is almost full, is to allot to each data**

**transmission an equal portion of the bandwidth. This result follows from research at CWI on queueing models for performance analysis of telecommunication systems with integrated services.**

In integrated services like (mobile) telephone with Internet, Web browsing or Internet with video communication, the various traffic flows need different treatment. For example, speech should not suffer much delay, lest it becomes incomprehensible, whereas for e-mail this is less problematic. In dealing with these problems it makes sense to distinguish between 'stream' traffic and 'elastic' traffic. The first mainly consists of real-time connections (telephone, video), which are prone to delay in the transmission. This requires some guarantee in capacity. Elastic traffic (bulk data, e-mail) allows fluctuations in transmission speed as long as the total transmission time remains 'acceptable'. The present Internet is focused on handling elastic traffic: transmission time increases with the number of users.

The network can be used more efficiently by integrating stream and elastic traffic. For example, pauses in a telephone conversation can be used for sending e-mail. Since stream traffic varies (eg, the

number of phone calls), so does the capacity for elastic traffic. Efficiency is improved by applying the method of 'processor sharing', in which every user is immediately allotted an equal portion of the bandwidth. Here the average waiting time is determined by 'first order' factors like the mean size of the data rather than the variance (in First-Come-First-Served models the dependence on variance can be detrimental to performance). CWI research showed that this favourable feature of processor sharing is preserved under varying network capacity.

In collaboration with Dutch Telecom (KPN Research) the optimal integration strategy for stream and elastic traffic was found to be a division of the bandwidth into two distinct parts. One part is reserved for elastic (data) traffic only, in the other preference is given to stream traffic. If the latter needs less capacity, the unused capacity is given to the elastic traffic. This strategy provides some protection for both types of traffic, as

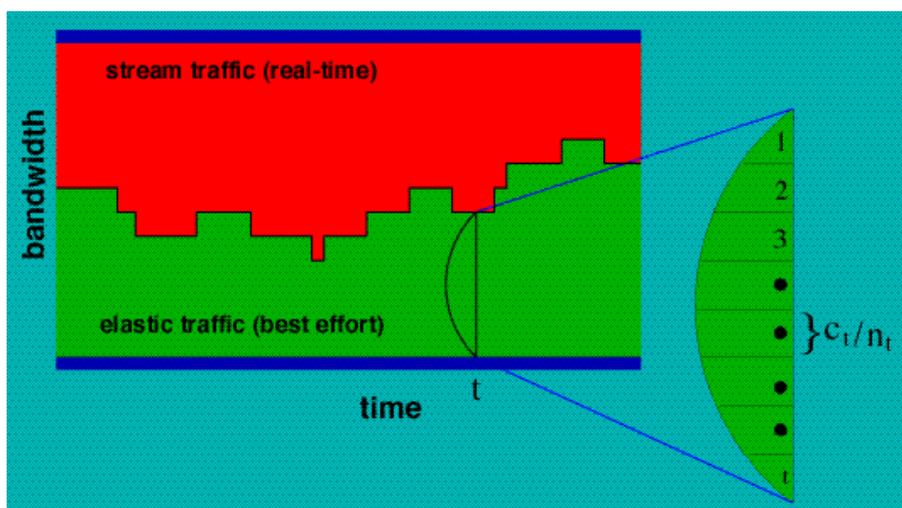
opposed to using the 'law of the jungle', which gives every user instant access if capacity is available. Implementation in the switches of a network both guarantees quality-of-service and achieves a high network efficiency.

Integrated services in mobile communication is at present a focal point of interest, as is demonstrated, for example, by considerable media attention for the Wireless Application Protocol (WAP). Current research at CWI concerns comparison of several integration strategies of data and real-time in a mobile network. Here performance analysis is even more important than in fixed networks, because the capacity is much more limited.

#### Please contact:

Sem Borst – CWI  
Tel: +31 20 592 4092  
E-mail: Sem.Borst@cw.nl

Sindo Núñez-Queija – CWI / INRIA  
Tel: +31 20 592 4168 / +33 4 9238 7637  
E-mail: sindo@cw.nl / rnunez@sophia.inria.fr



**Network capacity is efficiently used by simultaneous transmission of stream traffic (speech, video) and elastic traffic (data). At time  $t$  the available bandwidth for data ( $c_t$ ) is divided among  $n_t$  packets to be transmitted.**

# Interactions in Virtual Environments

by Lakshmi Sastry and David Boyd

At CLRC Virtual Reality Centre, located at Rutherford Appleton Laboratory, work is underway to implement a 3D interaction toolkit as part of the INQUISITIVE project. The toolkit architecture has been derived from a task-based analysis of a range of immersive VR-based applications for large-scale engineering design

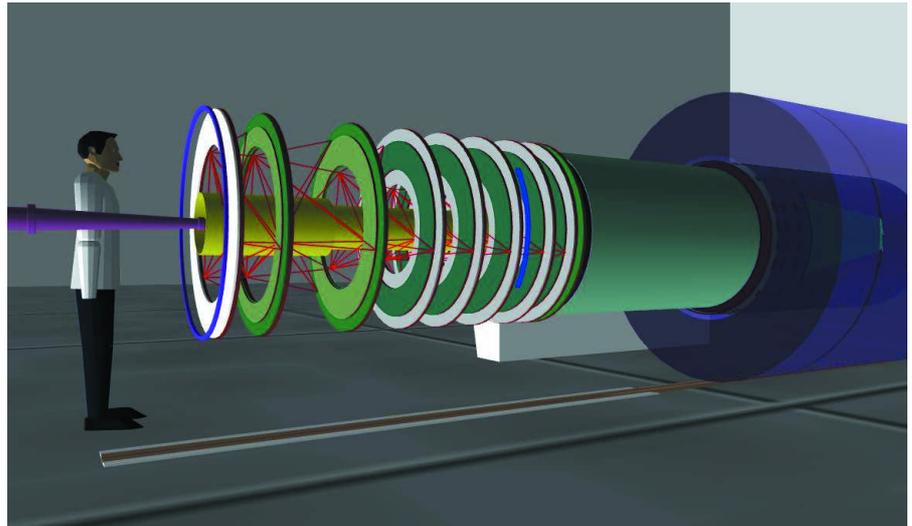
reviews, operational training and scientific visualisation over the period of the last four years. A VR system specific subset of the toolkit is an Engineer's Toolbox (ET) which is being developed as part of the smaller project Delivering Advanced Virtual Interactions to the Desktop (DAVID).

Virtual Reality (VR) technology and techniques have the potential to deliver intuitive user interfaces with more natural styles of interaction between the user and his application which will enhance productivity and performance. Facilities for user interface and interaction development within today's VR systems are rudimentary, limited and limiting. The availability of a generic interaction toolkit with portable modules makes a significant contribution to the rapid development and successful application of 3D VR interaction techniques to a wide range of virtual environments.

The aim of the RAL component of the INQUISITIVE project is the development of a portable interaction toolkit for VR applications which will improve support for developing user interaction within task-oriented virtual environment applications. The design of the interaction toolkit is based on our extensive experience in developing and successfully deploying immersive end user applications for large scale engineering design reviews, operational training and visualization within virtual environments.

## Toolkit Architecture

All application tasks, however complex, can be implemented in terms of a combination of tasks from the four basic classes of user interaction – navigation, selection, manipulation and data input, in virtual environments. Each basic interaction task can be realised using a number of possible interaction techniques. For example movement can be implemented using the magic carpet or point-fly techniques. Each application will identify one or more interaction techniques appropriate for carrying out the tasks required in that application. This in turn will guide the definition of the interaction processes needed to realise



A virtual environment for Atlas Inner Detector alignment checking.

those techniques. The application tasks are then achieved by suitable combinations of these interaction processes. Based on this analysis, the main functional components which the toolkit must provide are:

- a set of interaction processes for the four classes of basic interaction tasks
- a set of generic virtual interaction objects such as toolbox
- a run-time interaction framework.

The components of the runtime interaction framework are:

- A Contextual Interpreter which interfaces to the device drivers to obtain the measures and triggers from the input devices and convert these into the VE co-ordinate system. It then takes the dynamic constraints and current state of the interaction objects provided by the Interaction Manager and interprets the measures and triggers in that context. Finally the Contextual Interpreter calls the appropriate interaction processes to generate the event tokens which it passes to the Interaction Manager. The Contextual

Interpreter is independent of the host VR system.

- An Interaction Manager which monitors the changing state of user interaction within the VE. It receives the event tokens from the Contextual Interpreter. It finds out the current state within the VR system's runtime object database through the VR run-time manager. The operation of the Interaction Manager must be customised for each host VR system.

This architecture can be seen to map on to existing VR system kernels such as Maverik from the University of Manchester, UK and dvMockup from Parametric Technology.

## Requirements Capture, Analysis and Modelling Methodology

We are using UML, the Universal Modelling Language, to capture and transform high level user interaction techniques and tasks into detailed interaction processes between the user, the system and the underlying application leading to an implementable design. The

use case methodology is being used to identify the core classes that each basic interaction task will require and the derived classes and class variables and methods for each interaction technique. We have realised that even the simplest interaction process in VEs requires several agents co-operating together and monitoring various system states concurrently for task achievement. This places significant requirements on communication between the components of the interaction toolkit and between the interaction toolkit and the VR system manager. The interaction, activity and class diagrams are helping to clearly identify the appropriate classes and entities and how they collaborate within each process. Its ease of use and its ability to describe and analyse interaction processes from both the perspectives of user requirement and of communication between the user, the interaction toolkit, the VR system and the user's application is promising.

### Proof of Concept – An Engineer's Toolbox

Over the past four years we have developed many VR-based applications for engineering design review, operational training and computational data visualisation in collaboration with other departments as part of supporting RAL's engineering and scientific activities. The figure shows the VR model of the Atlas Inner Detector with lines-of-sight laid out to check clearance and clashes. Such applications have demonstrated the potential of virtual reality technology and interaction techniques to deliver cost effective productivity benefits and greater insight into engineering design and computational data. The utility of the INQUISITIVE Interaction Toolkit is to be realised via the DAVID project to deliver an application toolkit to the engineers and scientists so that use of VR becomes part of their design and analysis methodology. Thus the aim of the DAVID project is to develop a portable

'virtual engineering toolbox' which will include intuitive handling operations on model parts for simulation of assembly and maintenance tasks as well as sets of tools for measurements and alignment checking tasks in immersive and desktop environments. In addition, it can include facilities for superimposing and visualising analytical results from computational simulations such as metal deformation, fluid flow, thermal and stress analysis simulations. The implementation is currently under progress using the commercial VR system dvMockup which links closely with the ProEngineer CAD software.

#### Links:

The INQUISITIVEProject:  
<http://www.itd.clrc.ac.uk/Activity/INQUISITIVE/>

#### Please contact:

Lakshmi Sastry – CLRC  
Tel: +44 1235 44 6892  
E-mail: m.sastry@rl.ac.uk

## Alternative Processing of Verbal Data

by Milan Mare

**The vague quantitative values which we call here verbal data are usually processed by means of the theory of fuzzy numbers or fuzzy quantities. This theory is well elaborated and the algebraical properties**

**of fuzzy numbers are well known. It is worth mentioning that these properties are not completely identical with the properties of deterministic numbers  $\tilde{n}$  but they are well known and analytically managed.**

The attempts to apply the theory of fuzzy quantities to the formal analysis of practical situations like, eg, the CPM (critical path method) optimization techniques show another serious discrepancy connected with this, today already classical, processing of fuzzy quantities based on so called extension principle. It is used for the summation and multiplication of fuzzy quantities and it has a very logical formal structure consisting in re-combination of possible values and their possibilities of the input fuzzy quantities. Anyhow, the extension principle enlarges the range of possible values (namely in the case of summation) in a significant measure and, namely if it is applied repeatedly in processing a large number of fuzzy quantities, the extent of possible values of the resulting fuzzy quantity becomes enormously (and unacceptable) large.

The everyday experience with practical manipulation with vague data (and many traditional techniques do manipulate with them without admitting that they process fuzzy quantities – let us remember quantitative units like 'pinch', 'handful', 'step', 'two-days marching distance', 'proportional weight', 'acceptable profit', etc.) shows that the results of it are much less dispersed and their realistic values are concentrated near some periodic achieved 'expected' values. As such practical vagueness is usually connected with verbal description of quantitative values, we use the term 'verbal data' for their characterization. The above practical experience shows that there exists a contradiction between the fuzzy set theoretical model and real state regarding the processing of such verbal data. Its roots may be quite deep and worth for

discussion, which was already done. Let us mention, at least briefly, its essence.

It appears to be the crucial point that the operations based on the extension principle do not reflect the specific semantic structure of the fuzzy quantities modelling the verbal data. Namely, these fuzzy quantities are, in fact, combinations of at least two components. Verbal datum like, eg, 'approximately 8' consists of a 'crisp core'  $\tilde{n}$  its proper quantitative value (in our case 8) and a semantic 'shape' mathematically describing the structure of vagueness connected with the verbal expression (in our example with 'approximately'). Meanwhile the crisp cores are real numbers in the deterministic sense and they can be processed by means of classical deterministic operations, the shapes are mathematically real-valued

functions mapping the real line into the unit interval with modal value 1 achieved for the argument 0. They can be interpreted as some kind of normalized membership functions of the fuzzy quantities modelling the verbal data. As the shapes are connected with the semantic component of the fuzzy datum, they would be rather processed by fuzzy logical methods using operators like maximum, minimum, or product, etc., even if the formula of the extension principle is also one of the applicable methods.

Each algebraical operation realized over such verbal data is to be decomposed, then, into operation with crisp cores

leading to a deterministic real value, and logical or semantical operation over shapes leading to a resulting shape. This output crisp core and shape define the fuzzy quantity which reflects the result verbal datum. Due to the applied operation over shapes the possible values of the result are more or less closely concentrated near their crisp core (it is worth mentioning that usually much more closely than if the extension principle is applied).

The above method can be very easily completed also for the manipulation with verbal data having 'anonymous' crisp core (like 'several', 'few', etc.) or combined

structure of relation between vague and crisp component (eg 'something between 20 and 25', 'a bit more than 10', etc.). In all these cases it offers the possibility to reflect the logical relation between the processed verbal data much better than the traditional approaches limited to the extension principle.

**Please contact:**

Milan Mare – CRCIM  
Tel: +420 2 688 4669  
E-mail: mares@utia.cas.cz

## Towards the Distributed Virtual Education and Training Environment

by Igor Nikitin, Valery Burkin and Stanislav Klimenko

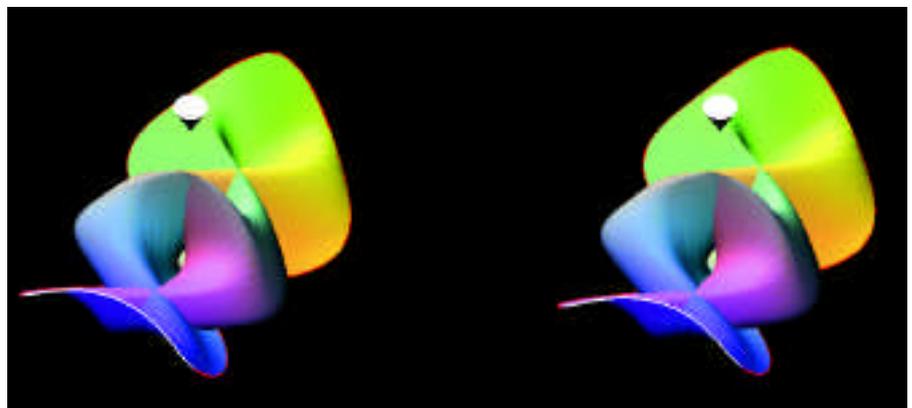
**Interactive installations in a virtual environment, supporting the educational process, are created at GMD Institute for Media Communication. Each installation is reflected in the Internet as an**

**introductory course to correspondent discipline. Related video-films, recorded in a virtual environment, are offered as illustration material.**

The modern educational processes actively uses the methods of computer graphics and scientific visualization. Wide opportunities are opened by the emerging technology of virtual environments, which can be used for the creation of highly interactive virtual laboratories intended for teaching different disciplines. GMD's Institute for Media Communication participates in the INTAS project 'Visualization of Complex Physical Phenomena and Mathematical Objects in Virtual Environment', supporting the development of virtual laboratories in natural sciences. Three installations are currently created in GMD's immersive virtual environment system CyberStage, serving education in higher mathematics, theoretical physics, astronomy and astrophysics.

### Topological Zoo

This installation demonstrates the main properties of elementary topological objects (Moebius band, Klein bottle, projective plane etc.) and gives the possibility to explore interactively these complex constructions arising in higher mathematics. The installation is



**This cross-eyed stereopair shows a sample object from relativistic string theory: world sheet. Arrow points to physical singularity (4D pinch point).**

accompanied by a web-tutorial on the topology of low dimensional manifolds, introducing all necessary concepts, and illustrated by 3 video-films.

### In the Microworld

This is an elementary introduction to a theory of relativistic strings, a branch of theoretical physics, exploring the properties of elementary particles. String theory is selected due to its increasing popularity among the theorists and

because it accumulates almost all mathematical systems used in modern theoretical physics. To attend this Internet-course, no special pre-requisites are needed, except 1st year University lectures on higher mathematics. The course includes a lot of graphical material, particularly, stereoscopic movies in the form of stereopairs and anaglyphs. Several applications on visualization of string dynamics are available, playing the role of interactive practicum. The course

is also supported by an installation in CyberStage. It is a prototype for a virtual lecture room of the future, where complex animated 3D forms, created by theoretical physics, can be directly observed and interactively studied. The related course material can be displayed on virtual screens using virtual slide projectors.

### Virtual Planetarium

Experimental course on basics of astronomy, delivered inside the immersive virtual environment system, and describing interactively all surrounding astronomical objects (planets, constellations, stars, nebulae etc). Currently the installation includes 3D models of 30 objects in the Solar System, 3200 nearby stars, a large database, containing textual descriptions

of all the objects in a scene, interactive map of constellations and tools for search and navigation.

### Applications

The described courses and videos were used in the educational process at Moscow Institute for Physics and Technology, University of Helsinki, University of Kaiserslautern and other educational institutions in Europe and USA. Virtual laboratories in GMD's high-end immersive virtual environment systems were presented at several local and remote exhibitions, students and scholar excursions.

The content of virtual laboratories can be integrated to Internet-accessible library of components, where the teacher can get

the objects necessary for presentation. All applications can be adapted to simpler projection systems for demonstrations on the screen of the University's lecture hall or the wall of real planetarium.

### Links:

Introductory course on string theory:  
<http://viswiz.gmd.de/~nikitin/course>  
 Tutorial on topology of low dimensional manifolds:  
[http://viswiz.gmd.de/~nikitin/vismat\\_html](http://viswiz.gmd.de/~nikitin/vismat_html)  
 Virtual Planetarium in CyberStage:  
<http://viswiz.gmd.de/~nikitin/stars/starsCB.html>  
 Related research projects:  
<http://viswiz.gmd.de/~nikitin/viz.html>

### Please contact:

Igor Nikitin – GMD  
 Tel: +49 2241 14 2137  
 E-mail: [Igor.Nikitin@gmd.de](mailto:Igor.Nikitin@gmd.de)

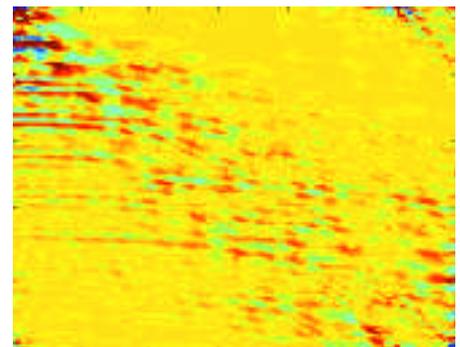
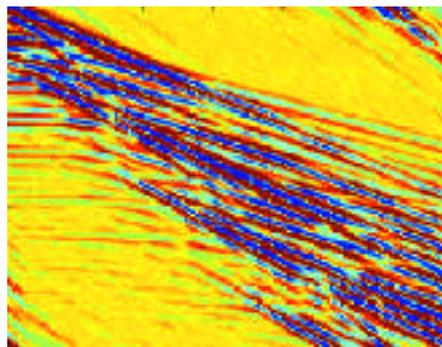
## Wavelets improve Information on Earthquakes

by Henk Nieland

**Tremors caused by an earthquake contain valuable information on its nature. Hence, close examination of these often very fanciful vibration patterns is highly important. CWI has analyzed seismic signals with the wavelet method. It turns out that in this way structures**

**in the signal are revealed that otherwise would remain hidden. The research was carried out jointly with the Dutch National Meteorological Institute KNMI in a project financed by the Dutch Technology Foundation STW.**

Apart from natural earthquakes, which remain unpredictable so far, there are also those caused by man, in order to explore the structure of the earth's crust (for example to trace possible oil sources). Vibrations reflected on the higher earth layers are recorded at an observation station and analyzed. Such a seismic signal consists of several waves (phases). Precise determination of the arrival times of two important waves, the P-phase (primary) and the S-phase (shear), is a crucial step in fixing the quake's type and location. The real art is to extract these two waves from the total signal, in which other waves appear alongside several types of noise. An algorithm based on wavelets was developed that automatically deducts the S-phase from the signal, given the precise arrival time of the P-wave (the S-phase always arrives later). The algorithm was tested at KNMI on more than threehundred 'local' quakes (vibrations of gas fields in Groningen,



**Observed reflections on earth layers of seismic vibrations. On the left the original distorted data, on the right the same data after cleaning up with the wavelet method.**

mini-quakes in Limburg and the Ardennes, etc.). The difference with manual determination of the arrival time remained far below the accepted norm and was clearly smaller than what the best method so far could achieve. Before implementation at KNMI for automatic S-wave detection, more tests on a large

set of 'non-local' quakes (Japan, California, etc.) will follow.

A wavelet is a wave packet of small extension. By shifting it as a mould along the signal, and measuring at every time point how well the fit is, certain features may be revealed which are missed otherwise. The wavelet method,

developed since the mid 1980s, is particularly good at detecting sharp transitions or multi-scale (fractal) structures in a signal. It was proposed by the French geophysicist Jean Morlet as an alternative to the traditional Fourier method, which describes a signal (function) as a sum of sine and cosine waves of a certain frequency and amplitude. Because of their infinite extension the fit of such waves with the signal only gives global information

(average over the whole signal). In several cases this works well, for example to detect the influence of ocean waves in the wave pattern of the North Sea, but local features are usually missed. Wavelets also have a certain frequency and amplitude, but because of their small extension they are able to describe local features in a signal. Mathematically they are more complicated than the simple sine or cosine functions, but by now the theory is well-established. Wavelets are successfully

applied not only in the analysis, but also the processing of signals (efficient storage and transmission of sounds and images). Present research at CWI concerns so-called nonlinear wavelets, with application to information retrieval from multimedia databases.

**Please contact:**

Patrick Oonincx – CWI  
Tel: +31 20 592 4177  
E-mail: [patricko@cwi.nl](mailto:patricko@cwi.nl)

## SOFSEM'99

by Jan Pavelka

**After celebrating its silver anniversary in Slovakia, SOFSEM came back to Milovy in Moravia, Czech Republic. SOFSEM'99, held from 27 November to 3 December 1999, was true to its new character and mission, established five years ago: a full-fledged**

**international conference with international invited speakers and international refereed submitted papers (known as contributed talks). As in several preceding years, SOFSEM received support from ERCIM and several other sponsors from the IT industry.**

From its previous history SOFSEM has inherited its generalist and multidisciplinary character. This is attested to by the themes of '99: Trends in Theory, Core Technologies, Software and Information Engineering, From Data to Knowledge, and Applications. These were covered by 19 invited talks and reinforced by selected submitted papers.

The PC (program committee) chair and co-chair as well as the OC (organising committee) chair were approved by the SOFSEM Steering Committee in 1997. As usual, the Steering Committee recruited the invited speakers, leaving the rest (the composition and co-ordination of the PC, the preparation of the time table, and the supervision during the conference) to the PC chair and co-chair.

The 23 PC members – 10 Czech, 3 Slovaks, and 10 researchers from 6 other countries – reviewed 46 submitted papers coming from 19 countries and selected 18 contributions to be presented at SOFSEM.

SOFSEM'99 was attended by 138 persons with ERCIM institutes well represented. Although the majority of participants were Czech (82) there were 21 from Slovakia, 8 from Germany, 5

from the Netherlands. The UK, France, Spain, and Sweden had 3 participants each. The remaining 7 countries were represented by smaller numbers.

Serving as a meeting ground for professionals from both the theory and the practice of computing has always been an important ingredient of the SOFSEM mission. I am glad I can report that also SOFSEM'99 has served this purpose well. Closing the gap between theory and practice was also the topic of a panel discussion held on the first day of the conference.

The busy time schedule of SOFSEM'99, in which invited and contributed talks were supplemented by commercial presentations and a cultural program, still left time for lively discussions, which often focused on sketching research projects with newly discovered partners.

For the fifth time, the proceedings were published by Springer-Verlag in the Lecture Notes in Computer Science (LNCS) Series (1725) and contain papers from the invited and contributed talks.

I am greatly indebted to my colleagues in the Steering Committee, the PC co-chair Gerard Tel, all PC members, the OC chair

Jan Staudek and his experienced organising team. Together, they have produced another successful SOFSEM.

The history of SOFSEM, which now includes fresh information from SOFSEM'99, can be found under the general SOFSEM Web page at <http://www.ics.muni.cz/sofsem/sofsem.html>.

SOFSEM 2000 will be held probably at the same location, again with ERCIM heavily present in its organization. The new PC Chair is Vaclav Hlavac from the Czech Technical University, Prague. Keith G Jeffery (CLRC RAL) and Jiri Wiedermann (CRCIM) will serve as co-chairs. Jan Staudek will again be OC chair. For the on-line information on SOFSEM 2000 please see the above mentioned web page.

**Links:**

<http://www.ics.muni.cz/sofsem/sofsem.html>

**Please contact:**

Jan Pavelka – DCIT, Ltd.  
PC Chair SOFSEM'99  
Tel: +420 2 3536 3342  
E-mail: [pav@dcit.cz](mailto:pav@dcit.cz)

# ERCIM NEWS

ERCIM News is the in-house magazine of ERCIM. Published quarterly, the newsletter reports on joint actions of the ercim partners, and aims to reflect the contribution made by ercim to the European Community in Information Technology. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. ERCIM News has a circulation of 7000 copies.

ERCIM News online edition is available at [http://www.ercim.org/publication/ERCIM\\_News/](http://www.ercim.org/publication/ERCIM_News/)

ERCIM News is published by ERCIM EEIG, BP 93, F-06902 Sophia-Antipolis Cedex  
Tel: +33 1 3963 5303, E-mail: [office@ercim.org](mailto:office@ercim.org)  
ISSN 0926-4981

Director: Jean-Eric Pin  
Central Editor:

Peter Kunz  
E-mail: [peter.kunz@ercim.org](mailto:peter.kunz@ercim.org)  
Tel: +33 1 3963 5040

Local Editors:

CLRC: Martin Prime  
E-mail: [martin@inf.rl.ac.uk](mailto:martin@inf.rl.ac.uk)  
Tel: +44 1235 44 6555

CRCIM: Michal Haidl  
E-mail: [haidl@utia.cas.cz](mailto:haidl@utia.cas.cz)  
Tel: +420 2 6605 2350

CWI: Henk Nieland  
E-mail: [henkn@cw.nl](mailto:henkn@cw.nl)  
Tel: +31 20 592 4092

CNR: Carol Peters  
E-mail: [carol@iei.pi.cnr.it](mailto:carol@iei.pi.cnr.it)  
Tel: +39 050 315 2897

FORTH: Constantine Stephanidis  
E-mail: [cs@csi.forth.gr](mailto:cs@csi.forth.gr)  
Tel: +30 81 39 17 41

GMD: Dietrich Stobik  
E-mail: [stobik@gmd.de](mailto:stobik@gmd.de)  
Tel: +49 2241 14 2509

INRIA: Bernard Hidoine  
E-mail: [bernard.hidoine@inria.fr](mailto:bernard.hidoine@inria.fr)  
Tel: +33 1 3963 5484

SICS: Kersti Hedman  
E-mail: [kersti@sics.se](mailto:kersti@sics.se)  
Tel: +46 8633 1508

SINTEF: Truls Gjestland  
E-mail: [truls.gjestland@informatics.sintef.no](mailto:truls.gjestland@informatics.sintef.no)  
Tel: +47 73 59 26 45

SRCIM: Gabriela Andrejkova  
E-mail: [andrejk@kosice.upjs.sk](mailto:andrejk@kosice.upjs.sk)  
Tel: +421 95 622 1128

SZTAKI: Erzsébet Csehaj-Varjú  
E-mail: [csehaj@sztaki.hu](mailto:csehaj@sztaki.hu)  
Tel: +36 1 209 6990

VTT: Pia-Maria Linden-Linna  
E-mail: [pia-maria.linden-linna@vtt.fi](mailto:pia-maria.linden-linna@vtt.fi)  
Tel: +358 0 456 4501

## Free subscription

You can subscribe to ERCIM News free of charge by:

- sending e-mail to your local editor; posting paper mail to the address
- filling out the form at the ERCIM website at <http://www.ercim.org/>

## CALL FOR PARTICIPATION

### QofIS'2000 – Quality of Future Internet Services

Berlin, 25-26 September 2000

The workshop focuses on end-to-end services over QoS assured Internet, on service creation, configuration and deployment.

Sponsors: T-Nova (Deutsche Telekom), Cisco Systems, Nokia, IST Programme of the European Union, GMD FOKUS

For more information:

<http://www.fokus.gmd.de/events/qofis2000/>  
E-mail: [qofis2000@fokus.gmd.de](mailto:qofis2000@fokus.gmd.de)

## CALL FOR PARTICIPATION

### Cross-Language Evaluation Forum

A Cross-Language System Evaluation activity is now being launched in Europe. The activity is sponsored by the DELOS Network of Excellence for Digital Libraries (see page x of this number) in collaboration with the US National Institute of Standards and Technology (NIST) and the TREC Conferences. The Cross-Language Evaluation Forum (CLEF) will run three main evaluation tracks in 2000, testing multilingual, bilingual and monolingual (non-English) information retrieval systems. There will also be a special sub-task for domain-specific cross-language evaluation.

The results of the activity will be presented during a two-day Workshop on Multilingual Information Access, 21-22 September in Lisbon, Portugal, immediately after the fourth European Conference on Digital Libraries (ECDL 2000: see <http://www.bn.pt/org/agenda/ecdl2000>).

Those intending to participate in CLEF 2000 are requested to send an e-mail to Carol Peters ([carol@iei.pi.cnr.it](mailto:carol@iei.pi.cnr.it)), as soon as possible, indicating in which task(s) they intend to participate.

For further information and deadlines, see: <http://www.iei.pi.cnr.it/DELOS/CLEF/>

## SPONSORED BY ERCIM

### The Ninth International World Wide Web Conference

Amsterdam, 15-19 May 2000

Join five world leaders in the 21st Century Web community as they deliver keynote speeches at The Ninth International World Wide Web Conference (WWW9), the only Web conference where leaders from academia, research, government and industry meet on a global stage. The conference will be held May 15-19, 2000, at the Amsterdam RAI, in Amsterdam, The Netherlands, with Web-related hardware, software, and services expositions to open May 16-18. Registration for the conference is available online at <http://www9.org>.

The keynote speakers include:

- Egbert-Jan Sol, Vice President Technology at Ericsson: 'Making the Internet Mobile for Everybody'
- Paul de Bot, Vice President Business Development at Philips Digital Networks: 'The Convergence of Internet and Television Technologies: How will it Change the Living Room Experience?'
- Charles W. Davies, Group Development Director at Psion PLC: 'Wireless Information Devices and the Mobile Internet'
- M. Graeber Jordan, CEO of GJordan.com LLC: 'The Spirit of the WWW in the Corporate Intranet'
- Lawrence Lessig, Berkman Professor of Law at Harvard Law School: 'Cyberspace's Constitution'

The conference will last five full days with Tutorials and Workshops on day one, Keynotes and tracks for Refereed Papers, Poster Sessions, Web and Industry track, Web/Internet and Society track, Culture track, Panels and Invited Speakers on days two through four, and a Developers' Day on day five. The conference is also a unique opportunity to learn about the latest developments on Web standards, due to the extensive W3C track, organized by the W3C itself.

For more information: <http://www9.org/>

**SINTEF Telecom and Informatics has signed a contract with BAE-Systems** to develop a system for natural communication in rough environments (NACRE). The concept for this system was presented on the ERCIM 10th anniversary CD-Rom. BAE-Systems is the second largest company in Europe within aviation, space technology and defence systems with more than 90.000 employees. The Norwegian Air Force is currently negotiating a contract for renewal of fighter aircraft. The NACRE-system will be implemented in the Eurofighter which is a possible candidate for this contract.

**SINTEF has celebrated its 50th anniversary.** The research foundation SINTEF was founded in 1950 by The Norwegian Institute of Technology in order to strengthen the ties between the university and Norwegian industry and to promote applied research. SINTEF operates in close collaboration with the Norwegian University of Science and Technology (NTNU). The two institutions cooperate in projects and share laboratories and equipment. Together, they form a center of expertise of high international standard. SINTEF also cooperates with the University of Oslo. With about 1800 employees and an annual turnover of NOK 1,5 billion, SINTEF is one of Europe's largest independent research organizations. Contracts from the public and private sectors provide 92% of the operating revenue.

**The work of four members of CLRC's staff were recognised in the 2000 New Year Honours list,** amongst them Dr Gordon Walker and Professor Bob Hopgood, both of whom were awarded the Order of the British Empire (OBE) by the Queen. Dr Gordon Walker has worked at Rutherford Appleton Laboratory for 37 years and was recently appointed Chief Operating Officer for CLRC, responsible for the day-to-day operations of RAL and Daresbury Laboratory, its sister laboratory in Cheshire. Professor Bob Hopgood has been involved in computer graphics since 1959 including pioneering work

on computer animation. He currently chairs the BSI Committee responsible for computer graphics and image processing standardisation. More recently he has been involved with the World Wide Web Consortium (W3C) helping to develop standards for the Web and setting up W3C European regional offices including the W3C UK Office at RAL. Bob is Professor of Computer Science at Brunel University where he has taught since 1968. He also has an Honorary Doctorate from the Technical University of Darmstadt, awarded for his trend setting activities in the field of computer graphics. Bob Hopgood is retiring from CLRC in April 2000, he is taking up the position of Head of Offices for W3C and will be based at CLRC – so he won't escape entirely! Bob will also continue with his work and positions at UK universities.

**Professor Brian Eyre has been appointed to the new post of Chairman of CLRC (CCLRC).** Eyre who has been a member of CCLRC's



**Brian Eyre, new Chairman of CLRC.**

council since 1998, is appointed for one year from 1 April 2000. Dr Gordon Walker, the present Director for Research & Development at CCLRC, has been appointed Chief Operating Officer and will be Acting Chief Executive from 1 April 2000 until a new Chief Executive is appointed through open competition. Dr Albert Westwood completes his term of office as Chairman and Chief Executive of CCLRC on 31 March 2000. In future CCLRC will have an independent part time Chairman and a full time Chief Executive, bringing it into line with the other six Research Councils in the UK.

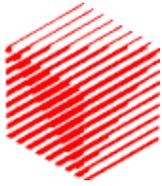
**The CNR Informatics Institutes in Pisa have moved** from their separate sites in the historical city centre to the new CNR Research Area



**The new Pisa research campus.**

on the outskirts of town. CNUCE, IAT, IEI, ILC and IMC, whose research activities encompass practically all sectors of IT and Computer Science, are now located on the same site. The new research campus will provide them with a number of important facilities, such as a Conference centre and printing services, which previously were lacking. When the Pisan Area of Research is fully operational, it will be constituted by 15 CNR research institutes working in various fields of the applied sciences (including Biology and Medicine, Physics, Chemistry, Environmental and Earth Sciences) with approximately 600 research, technical and administrative staff.

**ERCIM organizes Information Days on the European Information Society Technologies Programme in Latin America** with support from the European Commission. The objective is to raise awareness among the Latin American S&T actors about the opportunities related to the calls for proposals issued by this programme. Detailed information is given on the strategy of the programme, the content of its annual workplan, the available modalities for participation, and the coming 'roadmap' for its activities. ERCIM will produce a report to identify subjects of common interest and recommend actions or mechanisms for funding bodies in order to increase the R&D co-operation between Europe and Latin America in Information Technologies. After a very successful Information Day 16 March in Mexico City, two further Info Days will be held 4 April in Rio de Janeiro and later on in Buenos Aires. For details, see <http://www.ercim.org/EU-LA/>.



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.



Central Laboratory of the Research Councils  
Rutherford Appleton Laboratory, Chilton, Didcot, GB-Oxon OX11 0QX, United Kingdom  
Tel: +44 123582 1900, Fax: +44 1235 44 5385  
<http://www.clrc.ac.uk/>



Consiglio Nazionale delle Ricerche, IEI-CNR  
Via Alfieri, 1, I-56010 Ghezzano (Pisa) Italy  
Tel: +39 050 315 2878, Fax: +39 050 315 2810  
<http://www.iei.pi.cnr.it/>



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



Centrum voor Wiskunde en Informatica  
Kruislaan 413, NL-1098 SJ Amsterdam, The Netherlands  
Tel: +31205929333, Fax: +31 20 592 4199  
<http://www.cwi.nl/>



Danish Consortium for Information Technology  
DANIT co/CIT, Aabogade 34, DK - 8200 Aarhus N, Denmark  
Tel: +45 8942 2440, Fax: +45 8942 2443  
<http://www.cit.dk/ERCIM/>



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



GMD – Forschungszentrum Informationstechnik GmbH  
Schloß Birlinghoven, D-53754 Sankt Augustin, Germany  
Tel: +49 2241 14 0, Fax: +49 2241 14 2889  
<http://www.gmd.de/>



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



Swedish Institute, of Computer Science  
Box 1263, S-164 29 Kista, Sweden  
Tel: +46 8 633 1500, Fax: +46 8 751 7230  
<http://www.sics.se/>



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



Stiftelsen for Industriell og Teknisk Forskning ved Norges Tekniske Høgskole  
SINTEF Telecom & Informatics, N-7034 Trondheim, Norway  
Tel :+47 73 59 30 00, Fax :+47 73 59 43 02  
<http://www.informatics.sintef.no/>



Slovak Research Consortium for Informatics and Mathematics  
Dept. of Computer Science, Comenius University, Mlynska Dolina M SK-84215 Bratislava, Slovakia  
Tel: +421 7 726635, Fax: +421 7 727041  
<http://www.srcim.sk>



Magyar Tudományos Akadémia – Számítástechnikai és Automatizálási Kutató Intézet  
P.O. Box 63, H-1518 Budapest, Hungary  
Tel: +36 1 4665644, Fax: + 36 1 466 7503  
<http://www.sztaki.hu/>



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