Number 86, July 2011

ERCIM NEWS

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

Special theme:

ICT for Cultural Heritage

Also in this issue:

Keynote
ICT and Cultural Heritage: Research, Innovation and Policy
by Khalil Rouhana

Joint ERCIM Actions
University of Cyprus joins ERCIM

Research and Innovation
Three-Dimensional Reconstruction of a Nanoparticle at Atomic Resolution
by Joost Batenburg and Sandra van Aert

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 9,000 copies. The printed version of ERCIM News has a production cost of €8 per copy. Subscription is currently available free of charge.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
Tel: +33 4 9238 5010, E-mail: contact@ercim.eu
Director: Jérôme Chailloux
ISSN 0926-4981

Contributions
Contributions must be submitted to the local editor of your country

Advertising
For current advertising rates and conditions, see http://ercim-news.ercim.eu/ or contact peter.kunz@ercim.eu

ERCIM News online edition
The online edition is published at http://ercim-news.ercim.eu/

Subscription
Subscribe to ERCIM News by sending email to en-subscriptions@ercim.eu or by filling out the form at the ERCIM News website: http://ercim-news.ercim.eu/

Next issue
October 2011, Special theme: "Ambient Assited Living"

ICT and Cultural Heritage: Research, Innovation and Policy

The 'European Competitiveness Report 2010' identified the cultural and creative industries as one of Europe's most dynamic sectors, accounting for 3.3% of EU GDP and 3% of employment. Beyond this important contribution to GDP, the creative and cultural sectors are indeed a vehicle of significant lifestyle changes and progress, including the development of modern skills, adapting teaching and learning and inter-generational and intercultural dialogues.

Digital technologies continue to radically transform our approach to creativity and culture. They modify not only the way we access and preserve cultural assets but they also provide us with unique tools to better create and communicate. ICT has become a major vehicle for enriching our "creative capital" that underpins our societal vitality, our economic growth and ability to compete on a global scale.

Khalil Rouhana, Director, Digital Content & Cognitive Systems European Commission, DG Information Society and Media

This is why digital content, is one of the important policy areas of the Digital Agenda for Europe (DAE). The policy aim is to strengthen Europe's cultural creativity, develop further the internal market for digital content, preserve our cultural resources and make them widely accessible to all citizens.

The DAE approach consists of three interlinked pillars:

1. Creating the best framework conditions for the creation and diffusion of cultural assets including copyright issues and open access: A new copyright strategy has been suggested by the Commission in May 2011 that clarifies the legal framework for digitizing cultural heritage, making it accessible to all and unlocking the potential for its re-use. It includes notably a legislative proposal on orphan works in view of facilitating the digitization of millions of works that are in copyright, but for which the right-holders cannot be found.

2. Supporting the effort of digitization of all European cultural content and providing a platform for aggregating digitized cultural resources with an easy and single access point: Digitizing our cultural heritage is a huge task. A recent report has estimated the cost of digitizing our entire cultural heritage at 100 billion euro, highlighting the need for careful planning, since public money is scarce and duplication of effort should be avoided. At the same time the sheer magnitude of the effort calls for private investments in digitization, not least because the private sector will hugely benefit from a thriving content ecosystem in Europe.

   A central pillar in Europe's strategy in this area is Europeana, Europe's digital library, archive and museum. It aims at making Europe's cultural resources and scientific records easy to use for leisure, work or studies. It is also destined to act as a hub for the creative industries, facilitating the use of cultural resources in innovative products and services. Europeana currently gives direct access to more than 19 million digitized objects from more than 1,500 cultural institutions. It will reach 30 million by 2015.

3. Financing R&D and innovation: In 2011-12, the Commission is investing around 100 M€ per year in technologies for digital content and cultural heritage. Support goes to research and innovation in tools for content creation, access and preservation. It targets application areas spanning from cultural arts to entertainment and education. The underpinning technologies include computer vision, advanced graphics, simulation and visualization tools, 3D immersive environments, cognitive systems, multi sensorial interactions and semantic based content search.

The EU investment has helped to build a solid research community bridging the gap between traditional ways of dealing with cultural heritage and the opportunities offered by ICT. For example, The EU-funded project V-City - The Virtual City (http://vcity.diginext.fr) aims at an innovative system for reconstructing, visualizing and exploiting complete, large-scale and interactive urban environments. The technology (combining computer vision, 3D modelling and virtual reality) can be used by archaeologists and other cultural heritage professionals, or for business intelligence. Another example is the EU-Funded project eCUTE (http://ecute.eu/) developing a tool for cultural understanding and empathy in children and young adults. It uses a view of culture based on intercultural sensitivity model. In terms of structuring and coordination, the Network of Excellence GaLA on Serious Games for education and training aims at building a European Virtual Research Centre on Serious Games (http://www.elios.dibe.unige.it/gala/).

By combining policy, research and innovation, the DAE actions aim at making our cultural heritage a driver for innovation and creativity and to enable the wider diffusion of European culture and values worldwide. "We only live to discover beautiful things, all the rest is just waiting..." These words of the poet explain well our continuous quest to create and to access and preserve cultural resources, these millions of "beautiful things" for generations to live for.

The views expressed in the article are the sole responsibility of the author and in no way represent the view of the European Commission and its services.

2 **Editorial Information**

KEYNOTE

3 **ICT and Cultural Heritage: Research, Innovation and Policy**
by Khalil Rouhana,
Director, Digital Content & Cognitive Systems
European Commission, DG Information Society and Media

JOINT ERCIM ACTIONS

6 **University of Cyprus joins ERCIM**
by Marios Dikaiakos

7 **ERCIM Evolves with New Organisational Structure**

8 **fet11 Conference a Great Success**

9 **ERCIM Alain Bensoussan Fellowship Programme**

SPECIAL THEME

Martin Doerr, FORTH-ICS and David Arnold, University of Brighton, UK have been guest editors for this special theme section on "ICT for Cultural Heritage"

Introduction to the Special Theme

10 **ICT for Cultural Heritage**
by Martin Doerr and David Arnold

12 **Semi-automatic 3D Acquisition and Reassembly of Cultural Heritage:The SeARCH Project**
by Patrick Reuter, Nicolas Mellado, Xavier Granier, Isabelle Hairy, Robert Vergnieux and Nadine Couture

14 **Digital Acquisition and Structural Analysis of the "Rognosa" Tower in San Gimignano**
by Maria Girardi, Cristina Padovani, Marco Callieri and Massimiliano Corsini

15 **Real-World Geometry and Generative Knowledge**
by Thomas Schiffer, Christoph Schinko, Torsten Ullrich and Dieter W. Fellner

16 **Digital Documentation of World Heritage: The Scottish Ten Project**
by Douglas Pritchard and Daisy Abbott

18 **3D-COFORM: A Large-Scale Digital Production Environment**
by Katerina Tzompanaki, Martin Doerr, Maria Theodoridou and Sven Havemann

19 **Attaching Semantics to Document Images Safeguards our Cultural Heritage**
by Elena Console, Anna Tonazzini and Fabio Bruno

21 **Tangible Culture - Designing Virtual Exhibitions on Multi-Touch Devices**
by Martin Hecher, Robert Möstl, Eva Eggeling, Christian Derler and Dieter W. Fellner

22 **VISITO Tuscany: Landmark Recognition for Cultural Heritage**
by Giuseppe Amato, Fabrizio Falchi and Paolo Bolettieri

23 **National Gallery in Prague**
by Stanislav Mikeš, Michal Haindl and Radek Holub

25 **Macedonia: From Fragments to Pixels**
by Xenophon Zabulis, Dimitrios Grammenos, Antonis A. Argyros, Dimitrios Sifakis and Constantine Stephanidis

26 **Discovering Knowledge from Sumerian Economic Documents**
by Wojciech Jaworski

28 **Multimodal Computer-Assisted Transcription of Ancient Documents**
by Verónica Romero, Alejandro Héctor Toselli, Enrqiue Vidal, Elsa Cubel and Joan Andreu Sánchez

29 **Reducing OCR Errors in Gothic-Script Documents**
by Lenz Furrer and Martin Volk

30 **Deciphering the Story of the Museum**
by Eoin Kilfeather

32 **Non-linear Digital Storytelling for the Battleship "G. Averof"**
by Ralf Klamma and Yiwei Cao

33 **YUMA – Crowd Sourced Metadata Enrichment for Online Collections**
by Joachim Jung, Rainer Simon and Bernhard Haslhofer

34 **Tidy City: Learning about Cultural Heritage through a User-Created Mobile Game**
by Richard Wetzel

35 **Design, Implementation and Evaluation of a User Generated Content Service for Europeana**
by Nicola Aloia, Cesare Concordia, Anne van Gerwen, Carlo Meghini and N. Zeni

36 **Visual Semantic Browsing - A New Way to Access Digital Collections**
by Alan Payne and Peter Fry

38 **Data Quality of Public Heritage Data in Professional Applications**
by Jacco van Ossenbruggen

39 **Discovering Europeana Users' Search Behaviour**
byDiego Ceccarelli, Sergiu Gordea, Claudio Lucchese, Franco Maria Nardini, Raffaele Perego and Gabriele Tolomei

40 **Preparing the Ground for the German Digital Library**
by Kai Stalmann, Marion Borowski and Sven Becker

42 **Managing and Archiving Digital Audio: The GAMELAN Project**
by Jérôme Barthélemy

43 **The Open Kunstkammer Data Project**
by Vladimir Ivanov

44 **Pat.mapa: Linking Catalan Cultural Heritage**
by Mercè López, Oriol Almirall and Sergi Fernández

46 **GeoMemories: A Spatial-Temporal Atlas of the Italian Landscape**
by Andrea Marchetti, Elizabeth Jane Shepherd and Maurizio Tesconi

47 **Nephele – Smart Digital Memory of Artwork Conservation**
by Barbara Zitová, Miroslav Beneš and Janka Hradilová

49 **Soft Sensors: The New Frontier in Measurement for Non-Invasive Monitoring of Cultural Heritage Sites**
by Umberto Maniscalco and Giovanni Pilato

50 **Mathematical Models for Copper Sulphation**
by Fabrizio Clarelli, Barbara De Filippo, and Roberto Natalini

RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes

52 **Three-Dimensional Reconstruction of a Nanoparticle at Atomic Resolution**
by Joost Batenburg and Sandra van Aert

53 **Robust Methods for Image Processing in Anthropology and Biomedicine**
by Jan Kalina

54 **DIRECT: the First Prototype of the PROMISE Evaluation Infrastructure for Information Retrieval Experimental Evaluation**
by Nicola Ferro

55 **Potential Cell-phone Dangers: Latest Studies**
by Harry Rudin

57 **Crowdsourcing Culture with a Twist**
by Pär Hansson

58 **Building a Building for Nanotechnology**
by Harry Rudin

EVENTS

59 **4th Summer School on Network and Information Security**

59 **Facilitating an Indian Delegate's Participation at the ICT Proposers' Day 2011**
by Tom Williamson

60 **Announcements**

IN BRIEF

63 **Jos Baeten New General Director of CWI**

63 **Marie–Paule Cani Winner of the Eurographics Outstanding Technical Contributions Award 2011**

63 **Twenty-Five Years .nl Domain**

University of Cyprus joins ERCIM

by Marios Dikaiakos

The University of Cyprus (UCY) is delighted to join ERCIM as its 19th member. UCY is represented in ERCIM by both the Faculty of Pure and Applied Sciences and the Faculty of Engineering through the Department of Computer Science and the Department of Electrical and Computer Engineering respectively. Dedicated to producing and promoting high-end research, UCY strives to establish a position at the forefront of European research. In that respect, participation in ERCIM is expected to further enhance the strong links that UCY has with other European academic and research institutions.

The University of Cyprus was established in 1989, as the first public University of the country, and admitted its first students in 1992. Since then the University has developed fully fledged undergraduate and graduate programmes within all its departments and managed to host an academic community of 424 faculty members, 6240 students (4691 of whom are undergraduates and 1549 postgraduates) and 352 administrative staff members. Research is the basic strategic element for the University´s continuous growth. During the last few years the University of Cyprus has managed to acquire important research funding mainly from external sources with the major contributor being the European Union (circa 40%). The quantity and quality of the research programs present stable and high indices of growth, which shows the efficiency and effectiveness of the research activities sustained at the University.

The Department of Computer Science (http://www.cs.ucy.ac.cy) undertakes world-class, high-impact research in many areas of Computer Science, including:

Research at the Unviversity of Cyprus.

- Artificial Intelligence (knowledge representation, planning, agents, cognitive neuroscience, neural networks)
- Computer Architecture (multicore processor architecture, power and temperature aware micro-architectures, dataflow)
- Databases and Information Management (mobile and wireless data management, sensor networks, information retrieval)
- Computer Graphics (virtual and augmented reality, graphics)
- Computer Networks (fixed, wireless, ad hoc and vehicular networks)
- Distributed Computing and Web Technologies (cloud & grid computing, content distribution networks)
- Software Engineering (context-aware services, middleware, component-based software engineering)
- Theoretical Computer Science (distributed algorithms, fault-tolerance, algorithmic game theory, concurrency theory, formal methods).

Over the last 6 years, the Computer Science Department has been involved in more than 130 research projects funded by both the European Union and the Cyprus Research Promotion Foundation with a total funding of over 13 million Euros. Currently, the Department has 21 faculty mem-

Students at the Department of Computer Science.

bers and 64 research staff members. The Department offers a Bachelor degree in Computer Science, four Master degrees in Computer Science, Internet Computing, Intelligent Systems and Advanced IT (professional), and a Ph.D. degree in Computer Science. It hosts an advanced research infrastructure including cloud, grid and high-performance computing resources, advanced virtual reality and visualization equipment, and wireless sensor network and mobile computing facilities.

The Department of Electrical and Computer Engineering (ECE) was established in 2003. In a relatively short period of time it has managed to establish high quality research, teaching and outreach activities in numerous fields, namely:
Biomedical Engineering
- Computer Networks
- Digital Hardware Design and Embedded Systems
- Electronics and Microelectronics
- Robotics, Instrumentation
- Sensors, and Nanotechnology
- Intelligent Systems and Control
- Microwaves and Photonics
- Power and Renewable Energy
- Signal and Image Processing
- Telecommunication Systems
- Trustworthy System Design.

As of December 2010, it has 15 full-time academic staff, and this number is projected to rise over the coming years. In the first few years of the department´s existence, the faculty was awarded more than 75 research projects and grants in the aforementioned fields, with a combined budget of approximately €15 mil., 60% of which coming from the EC. This funding success continues, and is considered essential to the future aspirations to be a leading centre of ECE in Europe.

UCY considers research as a basic strategic element for its continuous growth. In joining ERCIM, we anticipate that UCY will advance its research prospects and capabilities. Evidently, ERCIM can act as the medium to strengthen partner ties and expand strategic collaborations. Moreover, it can provide great assistance in the attraction of, and participation in EC funded projects. Indirectly, via increased research activity, both departments can shape their research and teaching curricula in a way that reflects modern academic and industrial thinking and state of the art technology. Consequently, the University of Cyprus is delighted to join ERCIM and looks forward to being an active participant. Three members of staff will represent UCY in ERCIM, professor Marios Dikaiakos as the representative in the ERCIM general member assembly with Dr. Theocharis Theocharides as his substitute. Professor George Papadopoulos represents UCY on the ERCIM News editirial board.

Link:
http://www.cs.ucy.ac.cy/

Please contact:
Marios Dikaiakos
University of Cyprus
E-mail: mdd@cs.ucy.ac.cy

ERCIM Evolves with New Organisational Structure

With the first general assembly of the ERCIM association (AISBL - association internationale sans but lucratif; international not-for-profit association) in Trondheim, Norway, 10 June 2011, ERCIM has accomplished a major organisational restructuring. An important change in ERCIM's membership policy is coinciding with the creation of ERCIM AISBL. ERCIM is now open to multiple members per country while retaining its membership criteria of excellence and active participation. The aim is to build a strong European community of research organisations to promote science, research and innovation.

For more than 20 years, ERCIM has operated as a grouping of members and associates, one per country, working cooperatively. During recent years, the number of European research organisations in information and communication technologies and applied mathematics has increased considerably. ERCIM is evolving to maintain its position as the main recognized expert grouping in Europe and to increase its impact on European strategy and policy. The ERCIM community, continuing its mission 'Cooperating for Excellence in Research', is now supported by two organisations: The ERCIM association (AISBL) and the ERCIM European Economic Interest Group (EEIG).

ERCIM AISBL

The ERCIM not-for-profit international association is carrying out and supervising all scientific activities of ERCIM.

ERCIM EEIG

The ERCIM European Economic Interest Group is composed of a subset of members from the association, hosting the ERCIM office and the European branch of the World Wide Web Consortium (W3C).

During the first general assembly, the association elected the board of the association composed of the following members:
- President: Keith Jeffery, STFC, UK
- Vice-President: Domenico Laforenza, IIT-CNR, Italy
- Treasurer: Dick Broekhuis, CWI,The Netherlands
- Secretary: Jerzy Tiuryn, University of Warsaw, Poland

and in addition:
- Dimitris Plexousakis, Institute of Computer Science, Foundation for Research and Technology (FORTH) – Hellas, Greece and Patrick Furrer, Swiss Association for Research in Information Technology (SARIT), Switzerland, jointly responsible for scientific aspects (projects, working groups)
- Claude Kirchner, INRIA, France, responsible for human capital
- Andreas Rauber, AARIT, Austria, responsible for outreach.

For more information, please contact the ERCIM office at contact@ercim.eu

fet[11] Conference a Great Success

***fet*[11], the European Future Technologies Conference and Exhibition 2011, held in Budapest 4-6 May 2011, was the second edition of a new forum dedicated to frontier research in information and communication technologies. *fet*[11] was a unique conference on visionary, high-risk and long-term research in information science and technology. Featuring an exceptionally broad range of scientific fields, the event seeded new ideas across disciplines that will reshape the future.**

The conference, jointly organised by ERCIM, SZTAKI and the Future and Emerging Technologies (FET) Units (Open and Proactive) of the European Commission, exceeded all expectations. It attracted more than 1000 participants and received huge media coverage. It involved key policy makers, and featured a mix of seven keynotes, a panel dis-

ERCIM President Michel Cosnard awarding a prize winner.

Neelie Kroes opens the exhibions.

cussion, 30 scientific sessions, 100 poster presentations and a science cafe with energy-packed ignite-style presentations. Hands-on exhibitions with 30 booths ran throughout, in parallel to the conference, showcasing the latest research developments in future and emerging information technologies.

fet[11] also marked the official launch of the FET Flagship Pilots by Neelie Kroes, Vice President of the European Commission and Commissioner for the European Digital Agenda. FET Flagships are large-scale initiatives that aim to achieve a visionary technological goal over 10 years with a budget of up to one billion Euro for each Flagship. To prepare the launch of the FET Flagships, six Pilot Actions were funded for one year starting in May 2011. In the second half of 2012 two of the pilots will be selected and launched as full FET Flagship Initiatives in 2013 (see http://cordis.europa.eu/fp7/ict/programme/fet/flagship/).

More than a traditional scientific conference, *fet*[11] was a unique intellectual event which created an atmosphere of excitement for the opportunities presented by FET-type research in Europe.

Recorded sessions

All plenary sessions (keynotes, panel discussion, political adresses, closing ceremony) were recorded and can be viewed via the fet11 website.

Awards

The participants were invited to vote for the best exhibition and poster. The following winners were awarded by the European Commission with a beautiful glass sculpture made by the Hungarian artist Laszlo Lukacsi and sponsored by ERCIM:

- Exhibit first prize: "The future of biomimetic machines" by Anna Mura et al.
- Exhibit second prize: "A new kind of robot: ECCEROBOT" by Owen Holland et al.
- Exhibit third prize: "Future technologies to support collaborative solutions for grand challenges (Biological water safety, Augmented collaboration, e-Infrastructures for science)" by Szabolcs Tokes, Peter Galambos, Robert Lovas et al
- Poster first prize: "Simple formations in large societies of tiny mobile artifacts" by Bastian Degener et al.
- Poster second prize: "From sensorimotor knowledge to abstract symbolic representations by Marek Rucinski et al.
- Poster third prize: "Methodological bridges for complex systems" by Emanuela Merelli et al.

Zoltan Csefalvay, Hungarian Minister of State for Economic Strategy and Parliamentary Affairs, Ministry for National Economy.

Proceedings

The scientific conclusions and perspectives from the keynotes, topical sessions, and posters are being published by Elsevier as proceedings entitled "fet11essence". The proceedings will be freely available from the fet11 website.

More information: http://www.fet11.eu

ERCIM Alain Bensoussan Fellowship Programme

The ERCIM Alain Bensoussan Fellowship Programme enters the third round of fellowships supported by the COFUND programme of the European Commission with the next application deadline 30 October. More than 50 fellowhips have already been granted under the COFUND scheme and ERCIM plans to co-fund some more 100 more fellowships during the next two years.

Who can apply?

The fellowships are available for PhD holders from all over the world.

What is the duration?

Fellowships are generally of 24 months duration, spent in two of the ERCIM institutes. A fellowship of 12 months duration spent in one institute might also be offered.

Application deadlines:

Twice per year: 30 April and 30 September.

How to apply?

Only online applications are accepted. The application form will be online one month prior to the application deadline.

Which topics/disciplines?

Topics cover most disciplines in computer science, information technology, and applied mathematics.

Where are the fellows hosted?

Fellows can be hosted at ERCIM member institutes only (the current ERCIM member institutes are listed on the back page of this issue). When an ERCIM member is a consortium (AARIT, CRCIM, IUA, PEG, PLERCIM, SARIT, SpaRCIM) the hosting institute might be any of the consortium's members. When an ERCIM Member is a funding organisation (FNR, FWO/FNRS), the hosting institute might be any of their affiliates.

What are the conditions?

- have obtained a PhD degree during the last eight years (prior to the application deadline) or be in the last year of the thesis work
- be fluent in English
- be discharged or get deferment from military service
- the fellowship is restricted to two terms (one reselection possible)
- have completed the PhD before starting the grant.
- a member institute cannot host a candidate of the same nationality
- a candidate cannot be hosted by a member institute, if by the start of the fellowship, he or she has already worked in this institute for a total of six months or more, during the last three years.

How are the fellows selected?

Each application is reviewed by scientists, and the criteria for selection are:

- scientific expertise of the applicant
- quality of scientific publications
- relevance of the fellow's research agenda
- interest/added-value for the ERCIM consortium
- previous mobility / professional experiences.

The number of available positions depends on the needs of the member institutes and their available funding.

More information: http://fellowship.ercim.eu/

Introduction to the Special Theme

ICT for Cultural Heritage

by Martin Derr and David Arnold

ICTs' use for Cultural Heritage has been growing very fast as part of the explosion in digital arts and humanities research driven by both public interest in heritage and the opportunity to enhance intellectual enquiry for Arts and Humanities researchers. Societal interest is at many levels from heritage themes in entertainment, including tourism and computer games, to political interest in heritage and its affect on our sense of identity and citizenship. Like education and health, heritage is personal, though influenced by the collective through commonalities in personal experiences.

The field is also maturing intellectually in a similar way to ICT adoption in other fields, progressing through stages of: replicating pre-existing manual processes; increasing inter-disciplinary collaboration and co-development and; co-development of tools which could not have been conceived or developed without the trust and understanding of previous stages. Leading research in digital humanities has reached this stage, which not only leads to novel techniques in arts and humanities but in computer science. Stokes characterizes this type of result as "use-inspired basic research" which he places in Pasteur's Quadrant in his scheme[1].

Cultural heritage may be considered as tangible (artefacts, the written word and documentation) or intangible (oral tradition, stories, performance, etc.), but the distinction may be illusory - written documents require interpretation; the significance of an object depends on the individual's cultural context etc. ICTs are currently most commonly used with the most tangible aspects, but the longer term challenges, both for the humanities and for ICTs, lie with representing and interrogating meaning, which inevitably spans both tangible and intangible.

Research inspired by:		Considerations of Use?	
		No	Yes
Quest for fundamental understanding	Yes	Pure Basic Research (Bohr)	Use-inspired basic research (Pasteur)
	No		Pure Applied Research (Edison)

Figure 1: Stokes' Quadrant Model of Scientific Research (after [Stokes, 1997], p73)

Increasingly, 3D digital model representations are providing new challenges and opportunities[2]. Apart from data volume, there are specific challenges concerning: metadata; the semantics of shape; digital provenance; and long-term preservation. Some of these areas are also challenge digital texts, images and videos but become highlighted when working with 3D. Even more challenging is the integration of knowledge across multiple sources and their metadata.

Libraries and Digital Libraries have survived perfectly well without information integration, providing good, homogeneous finding aids for scholars who know which collection to access. However a cultural-historical research space (museum, library or archive) provides access to primary knowledge about objects in very different organizations and archival material. The information in a museum or archive catalogue will not match the subject classification structure in a library context for example. The museum object is more like an illustration or witness of the past, than the raw information in its own right, some of which may be held as curatorial interpretations within the museum catalogue.

Cultural historical research requires knowledge of "possible pasts" –facts, events, material, social and psychological influences and motivations. It comes to life from understanding contexts by pulling together bits and pieces of related facts from disparate resources, which can typically not be classified under subjects in an obvious way. It lives from taking into account all known facts.

With the advent of interconnected sources, researchers are expected to search data from geographically distributed sources. This is impossible to achieve without ICT support

[1] [Stokes, 1997] Stokes, Donald, E. (1997) Pasteur's Quadrant: Basic Science and Technological Innovation, The Brookings Institution, Washington (ISBN 0-8157-8177-6), pp180.

[2] David Arnold and Guntram Geser, The EPOCH Research Agenda for the Applications of ICTs to Cultural Heritage. Full Report (May 2008) ISBN 978-963-9911-03-1, available on-line at http://www.epoch.eu

and sophisticated search tools. Video, image and 3D resources exacerbate these challenges with the embedded semantics of image and shape that may not be explicitly recorded in the digitized data and metadata, but require content-based analysis to understand the artefact's "meaning".

All digitization generates new metadata, but with 3D additional factors are involved. Models are produced by post-processing scanned or image data and models may be produced using a variety of technologies and algorithms. Understanding the data capture methods and modelling operations may influence future humanities research. For example one study might compare the shape of two objects (eg archaeological fragments) to determine whether they have common origins or analyse two models from different dates to determine whether they represent the same object (for stolen artefacts) or have changed (for condition monitoring over time). The complexities of acquiring 3D mean that different approaches are being taken to cater for the objects' physical properties. Techniques for acquiring building models are very different from those used for jewellery or costumes for instance, which complicate the design of search tools.

3D highlights other metadata challenges which are always there in principle but not often addressed. Digitizing a book may result in replicating the original – including the original pagination, headers and footnotes, and perhaps the handwritten notations of various owners. Where appearance is preserved by scanned images the challenge of relating image and text data remains.

3D involves complex relationships because many artefacts are composed of sub-parts. Even apparently single items – such as statues – may have several pieces. A complete building has many elements with their own interest – for example the decorative style permeated through dispersed and recurring features.

Detecting segmentation and the inter-relation of parts challenges both the digital representation and search operations. The re-use of a sub-part may also complicate the relationships it represents – a jewel remounted, or the re-use of stone from one historic building in a later building, for example.

At a deeper level, shape embodies semantics that are easily recognisable to the human eye but very difficult to analyse automatically. An archaeologist may recognise part of a cup but automated detection from a collection of shards may be unsolved. An expert may tag a shard as "cup" to simplify the search, but automated search involves feature detection. To detect things produced by the same artist, we would have to define the artist's "stylistic signatures" from features, as opposed to extracting from text metadata tags. New challenges also arise in detecting co-referencing from digital models of shape and (say) texts that describe an artist's life. Finally, shape may represent conventional meaning for which there is no linguistic equivalent – gesture or analogy for example. There is no agreed vocabulary of shape elements which would underpin the detection of such meanings from surface characteristics, although application such as face recognition and computer games interfaces may be leading the way.

There is a long way to go before the full potential of ICTs for Cultural Heritage is achieved. In the meantime, the articles in this special issue illustrate the range of research in the field and we look forward to seeing future developments which address these big challenges.

The authors' work which underpins this article has been supported by 3D-COFORM ("Tools and Expertise for 3D Collection Formation" – http://www.3d-coform.eu) – a large scale integrating research project co-funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 231809; 2008-2012. The objective is to improve 3D digitization tools and processes to make 3D documentation a practical alternative for tangible cultural heritage. It builds on many of the challenges recorded in the EPOCH Research Agenda for the use of ICTs with tangible cultural heritage.

Links:

The EPOCH Research Agenda for the Applications of ICTs to Cultural Heritage:
http://public-repository.epoch-net.org/publications/RES_AGENDA/research_agenda.pdf

3D COFORM project:
http://www.3d-coform.eu

Please contact:

Martin Doerr
FORTH-ICS, Greece
E-mail: martin@ics.forth.gr

David Arnold
University of Brighton, UK
E-mail: D.Arnold@brighton.ac.uk

Semi-automatic 3D Acquisition and Reassembly of Cultural Heritage: The SeARCH Project

by Patrick Reuter, Nicolas Mellado, Xavier Granier, Isabelle Hairy, Robert Vergnieux and Nadine Couture

The SeARCH research project (Semi-automatic 3D Acquisition and Reassembly of Cultural Heritage) brings together archeologists and computer scientists with one unique objective: the virtual reassembly of broken artifacts, and especially the broken statues that were surrounding the lighthouse of Alexandria. The gathering of researchers of these two fundamentally different scientific origins makes it possible to combine expert archeological high-level knowledge with the power of computer graphic visualization techniques and geometry p rocessing algorithms. The integration of both disciplines is achieved by designing efficient human-computer interaction techniques that use semi-automatic geometry-driven acquisition, visualization, and reassembly techniques.

Archeological artifacts are often broken into a large number of fragments, and, when reassembling the fractured objects, cultural heritage professionals are confronted by huge three-dimensional puzzles. The reassembly of the fragments is generally done manually, but this task can be very tedious, and in some cases impossible.

This is particularly true for the broken colossal statues that were surrounding the Lighthouse in Alexandria built in the 3rd century BC. The fragments were deteriorated by erosion, weathering, and the earthquakes in the 10th and 14th century AC. Most of the fragments are still underwater in the Mediterranean Sea, in a spectacular submarine archeological site of about one hectare in area next to Alexandria's Eastern Harbor, and only a few fragments have been brought to the surface. Due to the availability of fragments at different locations with varying access policies, digitally acquiring the fragments and reassembling the corresponding 3D models virtually is probably the only feasible solution to reason about the past.

With this common objective in mind, in 2009, the SeARCH project was born, a three year project funded by the French Agence Nationale de la Recherche (ANR). The SeARCH project strives to develop semi-automatic techniques for the acquisition, visualization, and reassembly of the 3D models. This project gathers the experience of four research partners, specialized in Archeology (Centre d'Études Alexandrines in Alexandria), in operational 3D scanning for Archeology (Ausonius Archéovision at Bordeaux University), in data processing for Computer Graphics (INRIA Bordeaux Sud-Ouest), and in Human Computer Interaction as well as simulation of physical phenomena (ESTIA Research).

In this first half of the project, we identified three major achievements: first, the digital acquisition of the fragments of the surrounding statues, even those that are still underwater. Second, the expressive visualization that highlights details of the acquired fragments that are sometimes invisible to the human eye when inspecting the real fragments. Third, a semi-automatic reassembly technique that allows an archeologist to drive sophisticated geometry processing algorithms in order to find the most probable reassemblies.

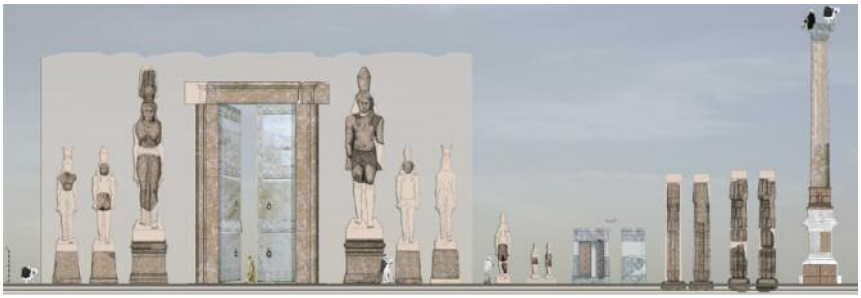


Figure 1: Restitution of parts of the lighthouse and monuments raised beside it (© Isabelle Hairy - CEAlex).

Fragment acquisition

The first step in the virtual reassembly is the on-site digital acquisition of the fragments in order to obtain 3D models. From the 3000-plus available fragments on the archeological sites in Alexandria, the cultural heritage professionals identify those with the highest reassembly potential. Instead of using traditional 3D laser range scanning that cannot be used underwater, our acquisition process is based on photogrammetry and is quite simple: we take about a hundred photographs per fragment covering the entire surface, with much less infrastructural effort than traditional techniques. Then, salient feature points are identified automatically in every photograph, and the correspondences of the same feature in different photographs are detected. This makes it possible to infer the 3D positions of the features, and hence to reconstruct the entire 3D model with submillimeter accuracy. With this acquisition protocol, we have even been able to reconstruct the immersed fragments during an underwater acquisition campaign. So far, we have reconstructed 15 fragments from the submarine site, and 34 fragments from three different museums.

Expressive visualization

Exploring the physical fragments on-site is sometimes difficult due to their size and the lighting conditions, especially for underwater fragments. Conversely, the 3D models can be inspected virtually in laboratory conditions. We have developed novel expres-

Figure 2: An underwater photographer during the 3D acquisition, and the reconstructed 3D model.

sive visualization techniques that use differential geometry to find the best lighting conditions to highlight all the features from the fragments. In our context, this is particularly important since we have to distinguish between the features of the fragment and the undesired effects from the century long suffered erosions. Our visual enhancement techniques stress features at different scales, ranging from vast ridge and valley lines to precise hand-made stone engravings. Since all our techniques operate in real-time on the graphical processing unit after a short preprocessing step, the 3D fragments can be comfortably explored by cultural heritage experts, sometimes revealing more detail than inspecting the physical fragments on-site.

Semi-automatic reassembly

The availability of accurate 3D models allows the archeologists for the first time to study the fragments coming from different locations in one coordinate frame. After visual exploration of the fragments, potential reassembly candidates are identified. For the pairwise virtual reassembly, we designed ArcheoTUI, an easy-to-use tangible user interface that makes it possible to relatively position the two fragments as if they were in the user's hands: in each hand, the user manipulates an electromagnetically tracked prop, and the translations and rotations are directly mapped to the corresponding virtual fragments on the computer screen when a corresponding foot pedal is pressed down. During the manipulation of the fragments, we provide feedback to the user in real-time by showing the geometric error of the reassembly. Furthermore, we provide a visual representation of the locally best match that is computed by optimizing the geometric compatibility of the two fragments with respect to the fragments' positions and orientations. Hence, the virtual fragments virtually "snap" to the locally optimal best match.

Future work

For the remainder of the project, we will continue to study the fragments with the challenging goal of finding new archeologically plausible reassemblies. We are also trying to physically model the phenomena that happened to the fragments during the deterioration. Furthermore, we are striving to develop a tangible interaction tool so that archeologists can easily make hypotheses about deformations.

Most of the members of the SeARCH project were involved in establishing V-MusT.net, a European Network of Excellence dedicated to Virtual Museums (2009-2013). Motivated by our attractive application case, we are currently planning an exhibition at a museum to show the results of the SeARCH project.

This work was supported by the ANR SeARCH project, grant ANR-09-CORD-019 of the French Agence Nationale de la Recherche.

Links:
http://anr-search.labri.fr/
http://v-must.net

Please contact:
Patrick Reuter
Bordeaux University, INRIA Bordeaux Sud-Ouest, LaBRI Bordeaux, France
E-mail: preuter@labri.fr

Figure 3: Enhancing details with expressive visualization.

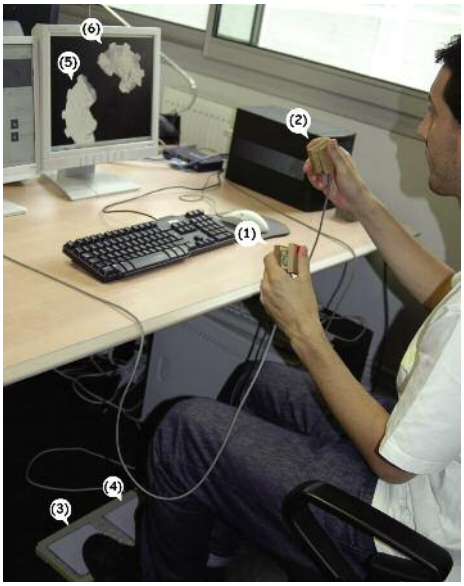


Figure 4: ArcheoTUI, a tangible user interface that makes it possible to relatively position two fragments as if they were in the user?s hands. In each hand,the user manipulates a prop (items 1 and 2). For each prop, there is a corresponding foot pedal (items 3 and 4). Only when the corresponding foot pedal is pressed down, the translations and rotations are directly mapped to the corresponding virtual fragment on the display (items 5 and 6).

Digital Acquisition and Structural Analysis of the "Rognosa" Tower in San Gimignano

by Maria Girardi, Cristina Padovani, Marco Callieri and Massimiliano Corsini

We present a set of integrated tools designed and developed for the digital acquisition and structural analysis of historical monuments and buildings.

The assessment of monuments requires various types of expertise and technologies. The St@rt project (Sciences and Technologies for the Artistic, Architectural and Archaeological Tuscan heritage), funded by the region of Tuscany (Italy) for the three year period 2007-2010, aimed at integrating different research and administrative expertise in the field of cultural heritage. The project's activities focused on the 13th century "Rognosa" tower in San Gimignano, included in the Unesco World Heritage List in 1990, as a case study.

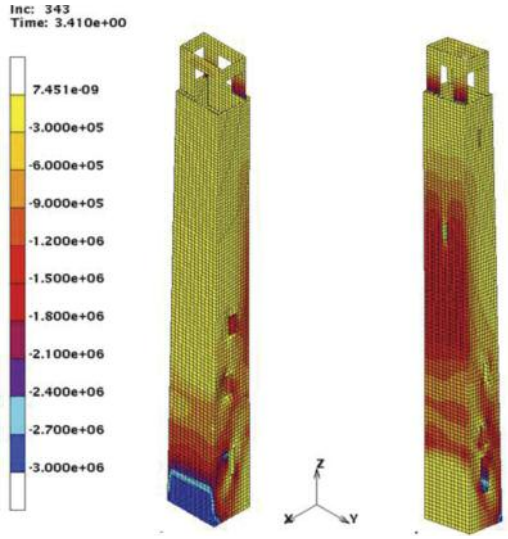
Figure 1: Acquisition of the tower's geometry by means of a TOF laser scanner (on the left), 3D digital model of the Rognosa tower and the Cathedral Square (on the right).

We generated a finite element mesh of the Rognosa tower using a three-dimensional digital model of the tower and the surrounding buildings, realized by a RIEGL time-of-flight (TOF) laser scanner. The tower's surface was acquired from four different positions, each selected to provide the best possible coverage of one side and, at the same time, a good view of an adjacent side, in order to facilitate the alignment process and provide coverage of all details from multiple angles.

Additional scans were carried out to acquire the facade of the building encompassing the tower and the front arcade overlooking the square, where a portion of the tower's base is visible. The model has been created by processing the laser scanner data using MeshLab, an open source software developed at the Visual Computing Lab and other geometric processing tools specifically designed to build a 3D model starting from a set of acquired range maps. Acquisition of the tower's inside geometry is impossible because it is only accessible via an internal staircase and the view from below of its highest portion is partially hidden by wooden floors. Thus, information about the geometry of the inner part of the tower has been obtained via classical measuring methods.

Although masonry towers represent an important part of the world's ancient architectonic heritage, their mechanical behaviour is, as yet, not well known. In fact, their response to dynamic actions is nonlinear and heavily dependent on the construction techniques utilized. Moreover, they are characterized, even in the static case, by high compressive stresses that can cause additional damage in correspondence to openings or other geometrical irregularities.

The mechanical behaviour of the tower was studied using the finite element code NOSA, developed by the Mechanics of Materials and Structures Lab in which masonry is assumed to be a nonlinear elastic material with



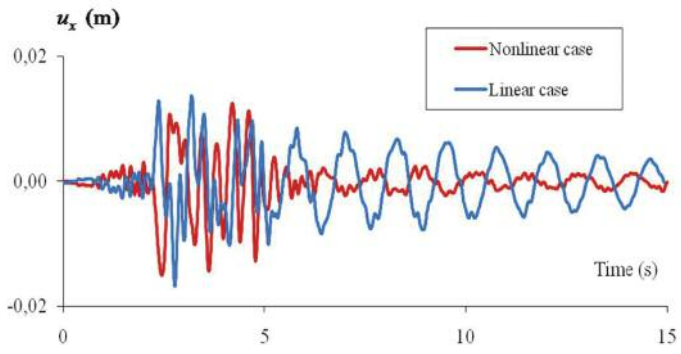


Figure 2 (left): Dynamic analysis: distribution of the stress Tzz (Pa) on the external surface of the tower at time 3.41 s.

Figure 3 (above): Dynamic analysis: relative x-displacement of the top of the bell chamber with respect to the top of the tower vs. time.

zero tensile strength and bounded compressive strength. The behaviour of the tower, subjected to its own weight and to time-dependent loads that model seismic excitations, was investigated.

After the static analysis, the tower was subjected to the horizontal component of the Nocera Umbra earthquake of 1997. The Nocera Umbra earthquake lasted 41.30 s and generated a maximum acceleration (PGA) of 4.3192 m/s^2.

Figure 2 shows the distribution of stress Tzz on the external surface of the tower at time t = 3.41s. The maximum crashing damage is recognized at the tower's base, while the maximum values of the fracture strain tensor are reached in the highest part of the tower, focused near the openings and in correspondence with the bell chamber.

Bell towers typically exhibit a certain vulnerability in correspondence with the bell chamber, which is the structure's highest element and is usually built separately from the rest of tower's structure. Figure 3 shows the behaviour of the bell chamber pillars of the Rognosa tower during the earthquake. The relative x-displacement of the top of the bell chamber with respect to the top of the tower is plotted vs. time. For the sake of comparison, the displacement corresponding to the analysis performed assuming the structure made of a linear elastic material, is reported.

The study highlights the importance of developing integrated procedures involving technologies of both visual computing and computational mechanics to obtain realistic models of monumental structures in complex architectural contests.

Links:
[1] The Mechanics of Materials and Structures Laboratory (MMS Lab) website: http://www.isti.cnr.it/research/unit.php?unit=MMS
[2] The Visual Computing Laboratory (VC Lab) website: http://www.isti.cnr.it/research/unit.php?unit=VC
[3] The MeshLab website: http://meshlab.sourceforge.net
[4] http://www.springer.com/materials/mechanics/book/978-3-540-79110-24

Please contact:
Marco Callieri, Maria Girardi
ISTI-CNR, Italy
E-mail: marco.callieri@isti.cnr.it, maria.girardi@isti.cnr.it

Real-World Geometry and Generative Knowledge

by Thomas Schiffer, Christoph Schinko, Torsten Ullrich and Dieter W. Fellner

The current methods of describing the shape of three-dimensional objects can be classified into two groups: composition of primitives and procedural description. As a 3D acquisition device simply returns an agglomeration of elementary objects (eg a laser scanner returns points) a real-world data set is always a – more or less noisy – composition of primitives. A generative model, on the other hand, describes an ideal object rather than a real one. Owing to this abstract view of an object, generative techniques are often used to describe objects semantically. Consequently, generative models, rather than being a replacement for established geometry descriptions (based on points, triangles, etc.), offer a sensible, semantic enrichment.

In combination with variance analysis techniques, generative descriptions can be used to enrich digitized artifacts. Detailed mesh comparisons can reveal the smallest changes and damage. These analysis and documentation tasks are valuable not only in the context of cultural heritage but also in engineering and manufacturing.

The Institut für ComputerGraphik und WissensVisualisierung (CGV) at Technische Universität Graz and Fraunhofer Austria tackled this problem and created a workflow, which automatically combines generative descriptions with reconstructed artifacts and performs a nominal/actual value comparison. The bridge between both the generative and the explicit geometry description is very important: it combines the accuracy and systematics of generative models with the realism and irregularities of real-world data. In this way, digitized artifacts can benefit from procedural modelling techniques: procedural models include expert knowledge within its object description; eg classification schemes used in architecture, archaeology and civil engineering can be mapped to procedures and algorithms, which realize a generative shape.

A generative description is like a shape template (eg to construct a cup) with some free input parameters (eg height and radius). For a specific object only its type and its instantiation parameters have to be identified. Identification is required by digital library services for markup, indexing and retrieval. The classification and semantic meta data in general are obviously important in the context of electronic product data management, product life cycle management, data exchange and storage.

The approach implemented by Fraunhofer Austria and CGV performs this semantic enrichment in three steps: The first step registers a generative model (including its free parameters) to the real-world data set. A generative model can be regarded as a function, which generates geometry, when called with some parameters. The registration step analyzes such a function, identifies whether it can describe the given artifact, and if so it determines the best-fit parameter set; ie no other parameter set can describe the digital artifact any better (in combination with the generative function). In a second step, the dif-

Figure 1: Comparison between a reconstructed vase and the results of a generative enrichment with real-world geometry. The light red parts do not have a counterpart in the input data set, whereas the light brown parts correspond to the input data set.

ference between the best-fit generative model and the input scan is computed. The results of this variance analysis are stored in a simple texture. Finally, the obtained variance can be visualized using X3D technology. The last step generates an X3D file containing the best-fit generative model, the texture of distance values and shader code capable of applying the difference as displacements. The X3D solution offers an integrated, standard compliant approach for visualization and documentation purposes. Furthermore, the standardized X3D format is self-contained and does not depend on an external viewer application.

The generated X3D file incorporates the following components:

- The geometry of the generative, best-fit reference mesh as indexed face set
- A texture storing distance values to the digitized artifact
- Vertex, geometry and fragment shaders for displacement and lighting purposes

To demonstrate our workflow we use an example data set of the Museum Eggenberg collection. The vase (see Figure 1 – left) has been reconstructed using the photogrammetric service ARC3D (http://www.arc3d.be), which takes a set of photographs and returns a reconstructed triangle mesh. Furthermore, an existing generative shape template of a vase with 13 degrees of freedom (ie 13 real-valued input parameters) has been passed to our workflow. The resulting X3D file describes the real-world geometry as well as the generative structure. Its included shader code is capable of applying the difference as displacements and allows a human viewer to switch selectively between the two model descriptions or displays both of them simultaneously (Figure 1 – right illustrates these results). The rendering shows the clean generative model (light red) for parts, which do not have a counterpart in the input data set, and the geometric offset (light brown) for parts, which do have a corresponding part in the input data set.

There are some drawbacks when displacing vertices on a curved surface (holes or overlaps appear) or when the input data has no correspondence to the generative representation. These issues will be the focus of future work.

Links:
http://www.cgv.tugraz.at/euclides
http://www.fraunhofer.at/vc
http://www.arc3d.be

Please contact:
Thomas Schiffer
TU Graz, Austria
E-mail: t.schiffer@cgv.tugraz.at

Digital Documentation of World Heritage: The Scottish Ten Project

by Douglas Pritchard and Daisy Abbott

The Scottish Ten project uses cutting edge technology to create exceptionally accurate digital models of ten UNESCO World Heritage Sites in order to better conserve and manage them.

Since 2009, the Digital Design Studio at the Glasgow School of Art, Historic Scotland and the CyArk Foundation have been involved in a five-year project to create precise 3-dimensional documentation of Scotland's five UNESCO World Heritage Sites and five international heritage sites. The project incorporates an extensive use of aerial lidar, terrestrial laser scanning, HDR photography, 2D and 3D visualization software and other technologies.

The objectives of the Scottish Ten project (http://www.scottishten.org) are to record important historical sites for the benefit of future generations in Scotland and overseas, test emerging survey and documentation technologies to see how they can be employed in the heritage sector, and provide precise digital media to site managers to better care for the heritage resource.

Four of the overseas sites have been selected based upon Scottish Government international objectives in North America, Japan, India and China. Internationally, the projects encourage the exchange of techniques and best

practices between the various agencies and build cultural connections. At the end of each international project, all of the data are given over to the site managers.

The acquired data from the Scottish Ten sites will be used for a number of purposes - from conservation practice to education programme development. Measured drawings and intensity values generated by the laser scanners will enable the Scottish and international site managers to better identify problems and quantify any decay of the monuments. At the same time, the data will be used to enhance interpretative and education programmes through the use of dimensionally accurate and photorealistic 3D modelling. This will provide unprecedented virtual access to the sites that may be restricted or closed to the public.

Of the Scottish UNESCO sites, New Lanark and Neolithic Orkney have been thoroughly documented and modelled. Both sites were extensively terrestrially laser scanned and digitally photographed. The documentation of the world heritage zone within the city of Edinburgh is currently under development. A lidar-based 3D model has been recently completed and the terrestrial scanning will begin later this summer. Given its size and the number of buildings within the World Heritage Site, this is an ambitious undertaking and the project is expected to be completed by the end of 2012.

The planning for the fourth site, the islands of St. Kilda, is well underway.

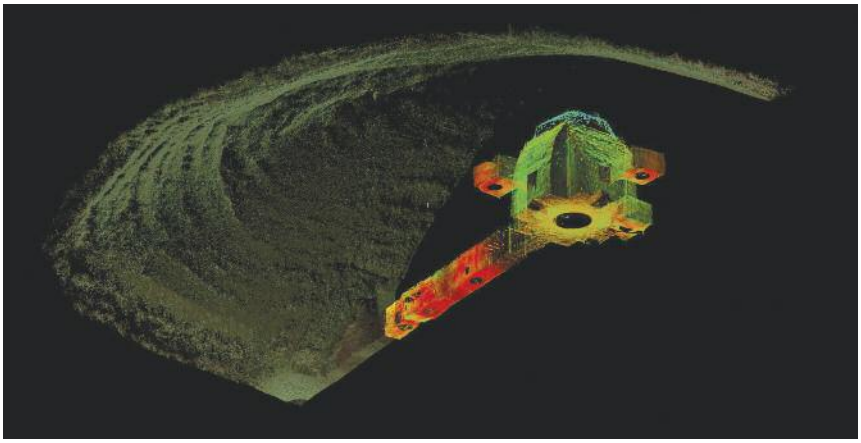


Figure 1: 3D point cloud rendering of Maes Howe from the Neolithic Orkney Scottish Ten project.

The combined 6 member Glasgow School of Art/Historic Scotland team will be documenting the Village Bay settlement area and other parts of the archipelago during the early summer. Weather permitting; the site documentation work will last for two weeks. The Antoine Wall project is scheduled to begin in 2012.

Internationally, the Mount Rushmore National Memorial in South Dakota was thoroughly documented in May 2010. The project was challenging, in part due to the terrain but also because of erratic weather conditions. To capture the entire surface area of the sculpted monument, a custom-manufactured rig was used to position the laser scanner on each of the four Presidents' faces. The procedure was physically demanding and relied heavily on the expertise of the US Parks Technical Ropes Team.

The Rani ki Vav (Queen's Stepwell) in Gujarat, India is intended as the next international heritage site. The well was built in 1050AD and consists of six stepped terraces into the ground. The walls are adorned with over 400 Hindu sculptures. The team will begin documentation work at the end of September 2011.

On the 21st and 22nd September the Scottish Ten project will feature prominently at the 2011 Digital Documentation Conference in Glasgow (http://www.digitaldocumentation.co.uk/). In November 2011, the results from St. Kilda project will be showcased at the UNESCO Remote Access Conference in Edinburgh.

Link:
http://www.scottishten.org

Please contact:
Douglas Pritchard
Glasgow School of Art, UK
Tel: +44 141 566 1452
E-mail: d.pritchard@gsa.ac.uk

Figure 2: The Rani ki vav / the Queens Step Well, Patan India.

Figure 3: 3D point cloud rendering of the four sculpted presidents from the Mount Rushmore Scottish Ten project.

3D-COFORM: A Large-Scale Digital Production Environment

by Katerina Tzompanaki, Martin Doerr, Maria Theodoridou and Sven Havemann

The systematic large-scale production of digital scientific objects, such as 3D models, requires much more infrastructure than a classical digital archive connected to a workflow manager. The size of the data to be handled, the distribution of expertise, acquisition and production sites, and the complexity of the processes involved require an innovative integrated environment that combines content management and information retrieval (IR) services with a centralized knowledge management in order to monitor, manage and document processes and products in a flexible manner.

The 3D-COFORM project aims to advance the state-of-the-art in 3D-digitization and to make 3D-documentation of cultural and other material objects an everyday practice by providing an integrated environment of services. The implementation and research period started in December 2008 with a duration of 48 months (European Community's FP7 IP n° 231809; 2008-2012).

As part of this project, the 3D-COFORM Repository Infrastructure (RI) was designed and is being implemented as a sustainable, distributed repository for a massive quantity of large digital objects and their metadata capturing scientific descriptions and monitoring all processes of data generation and processing. The RI is designed to allow different kinds of user (researchers, academics, web users, etc) to store, manage, query and retrieve whatever temporary and permanent digital objects are created in the course of digitizing physical objects, post-processing, and down to the final 3D products. Furthermore, it is designed to support their use and reuse in presentations and the scientific discourse about the modelled objects and their features, for art conservation, virtual reconstruction or hypothesis building about historical provenance and use. The distinctive innovation of the RI is that it can store the complete processing history of the digital artifact, its digital provenance. It is this feature that allows assessment of the authenticity of the resulting high-quality model, and reasoning about its properties, fidelity and fitness for purpose.

Data acquisition for 3D models may take place in a studio, but more commonly one has to go to the physical objects themselves, which may involve mobile equipment, open-air conditions, lack of network connections and reduced local processing capabilities. Particular processing services may be offered by specialized companies only in their premises, which results in data distribution. The acquisition is often costly, or even unique, requiring data replication. Data may be sensitive, or in the terabytes scale, making access rights, location and transfer an issue. Selection of data, manual processes, refinements, re-processing with improved methods, taking a series of measurements etc. contribute to the need of a centralized, integrated knowledge management in order to cope with defining the ultimate products, reasoning about their properties, and garbage collection.

The RI offers the aforementioned features through a central entry point, the RI-API webservice, following a highly compatible SOA approach (SOAP). Behind it there is a webservice-extensible Cloud Computing System (CCS) that can interface to other CCSs and Linked Open Data (LoD) based serv-

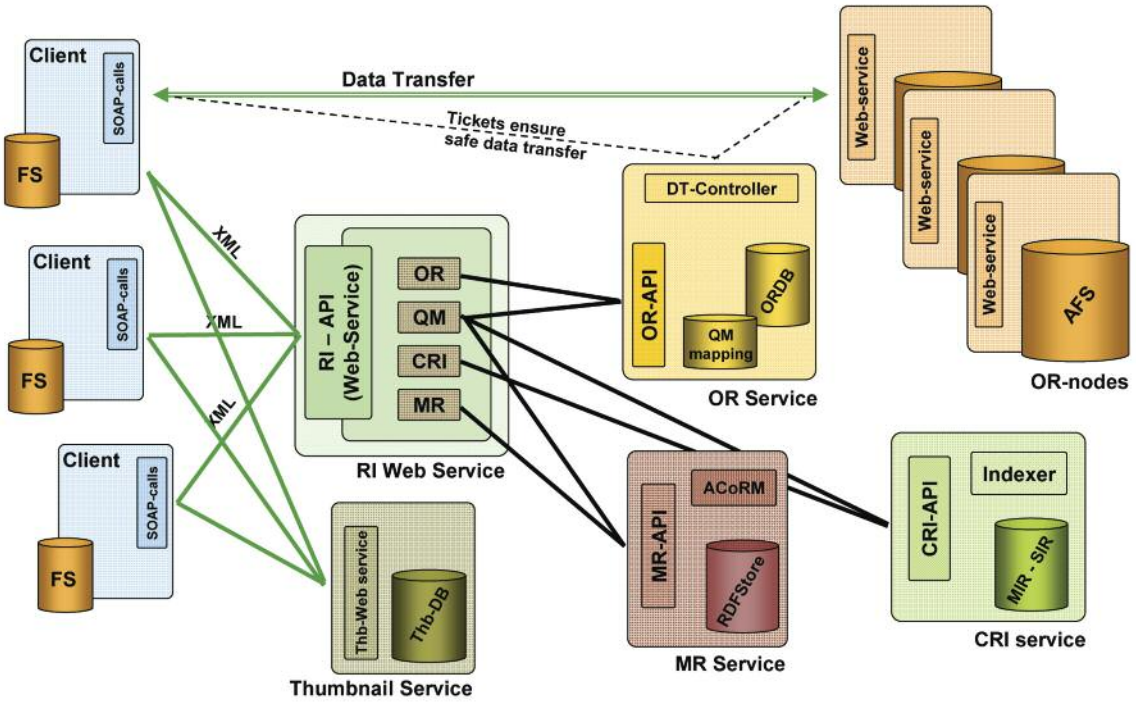


Figure 1: Overview of the Repository Infrastructure architecture: The clients communicate with the repository via SOAP through a central webservice. The RI dispatches the requests to its components (the MR, OR, and CRI Services). Data transfer is performed directly between clients and OR nodes, as initiated by the OR Service which controls the distributed OR nodes.

ices. The essential internal components of the RI are: (a) the Object Repository (OR), a distributed mass data storage layer, (b) the Metadata Repository (MR), an integrated semantic network layer, (c) the Content Retrieval Indices (CRI) for different search modalities by several content modules, and (d) the Query Manager (QM), which provides a single homogeneous access point to query the three components. The RI offers a central thumbnail database and provides a webservice for handling http thumbnail management requests.

The Object Repository (OR) connects to a (potentially large) number of distributed OR nodes for the physical data storage. The central OR Service provides access security, data integrity and risk-of-loss control, including metadata backup. It consists of a relational database (ORDB) for data file management, the query manager (QM) for mapping the relational database to the RDF format, the DT-Controller module for controlling the data transfer between client computers and OR nodes, and between OR nodes (replica management). The distinctive feature of the OR component is that all data transfers are logged for legal reasons: Not only the acquisition and post-processing of digital 3D assets are expensive, but high-quality 3D models can even be used for creating a high-quality physical replica, eg, through 3D-printing. So it is becoming ever more widely understood that 3D datasets are valuable assets that must be treated carefully, and that their proliferation needs to be faithfully recorded and controlled.

The Metadata Repository (MR) is an RDF triple store that aims at providing a common place to reason on, query, manipulate and export provenance metadata concerning any temporary or permanent object stored in the OR and related metadata about the modelled reality. Metadata are recorded in files in the units of creation, physically backed-up in the OR together with their related content files, while in the MR a semantic network is built with the integrated metadata information. The MR is based on a homogeneous global schema - an extension of the CIDOC CRM (ISO21127) that models provenance metadata (CRMdig). It comprises physical object descriptions, annotations and co-reference information, format and compatibility information of 3D models, historical events, and real world objects. All this information is stored in a coherent semantic network that enables useful and complex inferences to support content management and research, comprising even diverse content indexing and retrieval mechanisms for 3D objects. An integral part of the MR is the Annotation and Co-reference Manager (ACoRM). The Annotation Manager connects links into content segments of any kind and dimension (areas with no modifications on the original object), with semantic information capturing the related scientific discourse. The Co-reference Manager allows for collapsing duplicate URIs in the semantic network without loosing their provenance. This feature acts as a "mending" mechanism of the semantic network and will contribute significantly to the reasoning performance and the future connection into a Linked Open Data (LoD) world.

The responsibility for the RI design and implementation is shared between FORTH-ICS (Greece) and CGV, TU-Graz (Austria), while other partners in the Project deal with the creation of a rich Integrated Visual Browser interface to the RI and an immense spectrum of 3D tools, all integrated via the RI.

Links:
3D-COFORM Project: http://www.3d-coform.eu/
CIDOC CRM: http://www.cidoc-crm.org/
CIDOC CRM v5.0.2 Encoded in RDFS: http://www.cidoc-crm.org/rdfs/cidoc-crm
CRMdig 2.5 Encoded in RDFS: http://www.ics.forth.gr/isl/rdfs/3D-COFORM_CRMdig.rdfs

Please contact:
Katerina Tzompanaki, Martin Doerr, Maria Theodoridou
FORTH-ICS, Greece
E-mail: {katetzob, martin, maria}@ics.forth.gr

Sven Havemann
CGV, TU-Graz, Austria
E-mail: s.havemann@cgv.tugraz.at

Attaching Semantics to Document Images Safeguards our Cultural Heritage

by Elena Console, Anna Tonazzini and Fabio Bruno

Extracting and archiving information from digital images of documents is one of the goals of the project AMMIRA (multispectral acquisition, enhancing, indexing and retrieval of artifacts), led by Tea-Sas, a service firm based in Catanzaro, Italy, with the collaboration of two Italian research teams, the Institute of Information Science and Technologies of CNR in Pisa, and the Department of Mechanical Engineering of the University of Calabria in Cosenza. AMMIRA is supported by European funding, through the Italian regional program for integrated support to enterprises.

Gathering as much information as possible from the documents of our past is essential to preserve our cultural heritage for future generations, especially in an era when the documents are migrating from traditional supports towards the new electronic devices. For many historical documents, this involves capturing their appearance, as well as mitigating distortions and degradations, in order to help human or automatic readers to extract their content. All the information extracted must then be organized to facilitate storage, access and retrieval.

Depending on the type of document, the first step is to choose the most accurate

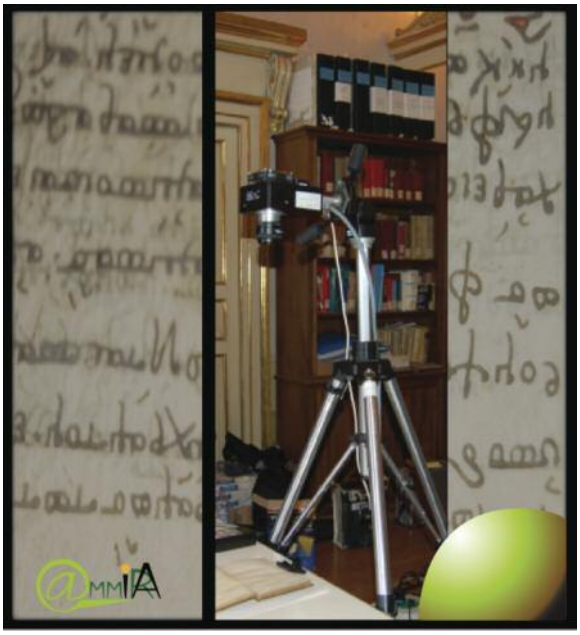


Figure 1: Capturing pages from an ancient book: the AMMIRA DTA Chroma multispectral camera.

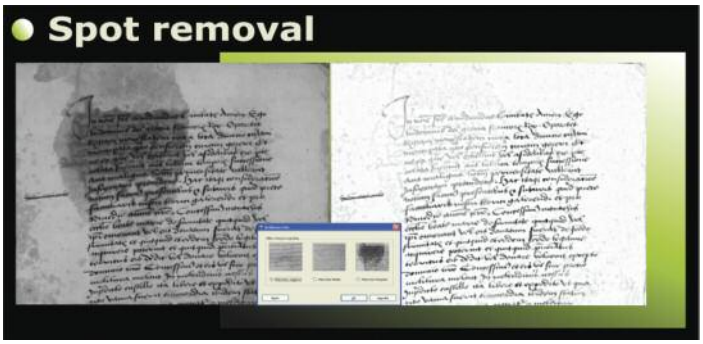


Figure 2: A screenshot from the AMMIRA image manipulation user interface.

and detailed data capture modalities. These modalities include the acquisition geometry and the spatial and spectral ranges and resolutions. The imaging system that we have adopted consists of a high-resolution, multispectral, computer-controlled camera with three visible and two infrared channels, coupled with a visible-ultraviolet lighting system and a structured-light projector for 3D acquisition. Thus, depending on the importance and the state of preservation of a document, we can represent it through panchromatic or colour images in the visible range, infrared reflectance or transmission images, and ultraviolet fluorescence images and, in some cases, a precise 3D shape. This flexibility of representation can help us detect and isolate many kinds of hidden features, and also account for geometrical distortions caused by deformations of the document support. The 3D acquisition augments the level of information that we are able to preserve, because deformations and degradations often describe the history of a manuscript. So, the fruition of the digital replica of a manuscript or an entire book acquires a new dimension as these can be represented as a 3D object in the space.

However, all these features are still raw data, ie a more or less detailed representation of the document appearance with no semantics relating to it. The images must next be freed from interferences and distortions, extracting all the features to which, at some level, semantics can be attached (eg "main text", "footnotes", "images-graphics", etc.). The software system that supports our imaging system is capable of performing many of these tasks. First, we can spatially co-register the available images, after correcting 2D or 3D geometric distortions if necessary. This provides us with a set of data maps with precisely located pixels. Further processing includes virtual restoration, ie removal of distortions and interference (stains, blur, bleed-through), and extraction of features and patterns. The latter task is accomplished through several fully or partially automatic approaches, based mainly on linear data models.

The key to incorporating semantics into the processed data is to adopt an efficient metadata schema that traces all the processing steps applied to any piece of data and all its relationships to other stored material, including all the administrative and descriptive information needed and, eventually, enabling content-based retrieval from a large data repository. The parts classified as text can be translated into machine-readable form through automatic or semi-automatic character recognition systems. The AMMIRA project began in September 2009 and will be completed in September 2011. The partners all have a consolidated experience in this field. The research is expected to continue in the future with the aim of implementing a fully integrated hardware-software system.

Link:
AMMIRA homepage:
http://www.ammira.eu

Please contact:
Elena Console
TEA-SAS, Italy
E-mail : elena@teacz.com

Anna Tonazzini
ISTI-CNR, Italy
E-mail : anna.tonazzini@isti.cnr.it

Fabio Bruno
University of Calabria, Italy
E-mail: f.bruno@unical.it

Tangible Culture – Designing Virtual Exhibitions on Multi-Touch Devices

by Martin Hecher, Robert Möstl, Eva Eggeling, Christian Derler and Dieter W. Fellner

Cultural heritage institutions such as galleries, museums and libraries increasingly use digital media to present artifacts to their audience and enable them to immerse themselves in a cultural virtual world. With the application eXhibition:editor3D, museum curators and editors have a software tool at hand to interactively plan and visualize exhibitions. The software is running on standard PCs as well as multi-touch devices, which allow a user to utilize intuitive gestures for positioning exhibition objects. Furthermore, multi-touch technology offers the integration of collaborative work into a decision making process.

In cultural heritage institutions, digital content has gained increasing importance over recent years. The availability of scanned 3D objects, images or videos of historical artifacts offers many advantages. Exhibits, which would be secured in an archive, can now be presented to a wide audience. Furthermore, virtual artifacts cannot only be used to replace existing objects, but also to complete fragments or to show them in a state never seen before. In digital form the content of an exhibition can be visualized with highlights and enriched with information. These enriched digital artifacts reach an audience all over the world without replacing the real museum. As a consequence, it is a reasonable and sensible extension to a museum visit and to the whole museum scene.

To bring the digital content into a presentable form, an authoring tool is needed. eXhibition:editor3D allows the composition of content into a virtual museum. The eXhibition:editor3D application, developed by the Austrian Institute of Information and Communication Technologies of JOANNEUM RESEARCH, is tailored entirely to a curator's needs and requirements. Museums, galleries and other cultural heritage institutions can utilize it for multiple purposes such as:

- the enrichment of existing exhibitions
- the design and preview of upcoming exhibitions
- archiving of temporary exhibitions
- the creation of interactive 3D-presentations for multi-media terminals
- interactive catalogues or websites
- advertising.

The primary target for the development of the eXhibition:editor3D was user friendliness and usability. As technical details are hidden behind an intuitive interface, curators can fully concentrate on realizing their ideas. The software uses two-dimensional graphical user interfaces and planning concepts based on floor plans. Within these plans the author places exhibition objects virtually by drag'n'drop techniques. Having placed all 3D models, images and video exhibits, the two-dimensional representation of the virtual exhibition can be exported any time to a three-dimensional world, which offers an immersive impression. The integrated preview comprises a realtime visualization of the virtual exhibition in 3D. Additionally placed sound elements – such as a narrator or ambient music – complete the virtual world.

The traditional way of interaction with virtual exhibits uses the mouse pointer as input device. This paradigm of user interaction has changed. Multi-touch technology is the interactive and intuitive means of providing a new user experience. In 2011, JOANNEUM RESEARCH and Fraunhofer Austria joined forces to extend the capabilities of eXhibition:editor3D. Fraunhofer Austria contributed its know-how on multi-touch devices. Now modern gesture-based user interfaces help to design exhibitions as easily as handling a mobile device or smart-phone: grabbing, rotating and scaling of objects can be done with a fingertip. Another ben-

Figure 1: Moving objects with a fingertip (left), collaborative work on a single exhibition (middle), floor based GUI interface of the application eXhibition:editor3D (right).

efit of the multi-touch technology is the possibility to collaboratively work on a single exhibition. Around a multi-touch table, as shown in Figure 1, authors can jointly discuss ideas and simultaneously position and adjust exhibition artifacts. The project cooperation between JOANNEUM RESEARCH and Fraunhofer Austria got off to a good start with two major enhancements of eXhibition:editor3D, namely the possibility to design virtual exhibitions with a modern gesture input interface and instant visual feedback of the resulting 3D scene. The combination of these two features provides exhibition editors a fast and flexible workflow for efficiently designing virtual exhibitions appropriate for a wide range of cultural heritage themes.

Links:
http://www.exhibition3d.at
http://www.fraunhofer.at/vc
http://www.joanneum.at

Please contact:
Martin Hecher
Fraunhofer Austria Research GmbH
Tel: +43 316 873 5410
E-mail: office.graz@fraunhofer.at

VISITO Tuscany: Landmark Recognition for Cultural Heritage

by Giuseppe Amato, Fabrizio Falchi and Paolo Bolettieri

VISITO Tuscany (VIsual Support to Interactive TOurism in Tuscany) is a research project which investigates techniques for producing an interactive guide, accessible via smartphone, for tourists visiting cities of art. The system applies image analysis and content recognition techniques to recognize photographed monuments. The user just has to take a picture of a tourist landmark to obtain pertinent information on his or her smartphone.

In the last few years, the problem of recognizing landmarks has received growing attention from the research community. As an example, Google presented its approach to building a web-scale landmark recognition engine that was also used to implement the Google Goggles service [3]. VISITO Tuscany [1] also addresses this issue, investigating and developing technologies in order to produce an interactive and customized advanced tour guide service to visit Tuscan cities of art. More specifically, it focuses on offering services to be used:

- During the tour – through the use of new generation mobile devices in order to improve the quality of the experience. The mobile device enables users to get detailed information about the artistic objects they are looking at or about their location. While taking pictures of monuments, places and other close-up objects, the users indicate what appears to them to be most interesting. When a picture is taken it is processed by the system to infer the user's interests and to provide relevant and customized information. For example, if a user takes a picture of the bell tower of Giotto, they can get detailed information (historical, artistic, structural techniques, etc) on this monument.
- Before the tour – to plan the visit in a better way. Information from other users of the system and their experiences , together with information already included in the database system and, more generally, on the web, can be employed by users to better plan their own visit. Interaction will take place through 3D graphic-based techniques.
- After the tour – to keep the memory alive and share it with others. The user can access the pictures taken and the itinerary followed through an interaction based on 3D graphics. Information and experiences can be shared with other users by creating social networks.

Figure 1: Tourist information on a smartphone.

One of the main novelties of VISITO Tuscany, is the possibility of obtaining information about monuments by taking a picture of the tourist landmark with a smartphone. The acquired image is analyzed and the landmark recognized so that the user can obtain relevant information related to the monument.

The landmark recognition system is composed of three main components: a client application that runs on a mobile phone, an image classifier that recognizes landmarks contained in pictures, and a digital library containing descriptions of various monuments. At the moment of writing, we have created recognizers for hundreds of monuments in three cities in Tuscany: Florence, Pisa, and San Gimignano. The mobile application is already available for the Android platform [2] and will be soon available also for iPhone. When the user takes a picture of a monument, the picture is first sent to the classifier that checks if one of the available monuments is recognized. When a monument is recognized, the description is retrieved from the digital library and sent back to the mobile device.

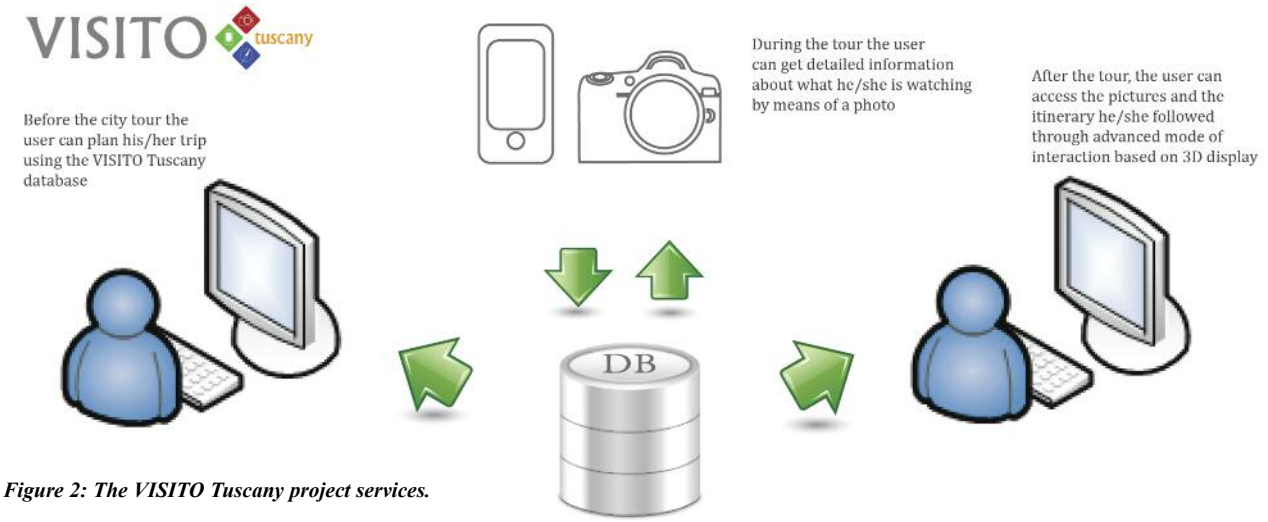


Figure 2: The VISITO Tuscany project services.

Landmark recognition is performed using local features and kNN based classification algorithms. We defined a new approach that relies on a revision of the single label kNN classification algorithm. More specifically, we propose an algorithm that first assigns a label to each local feature of an image query. The label of the image is then assigned on the basis of the labels and confidences assigned to its local features. In other words, our kNN approach is based on the similarity among the local features of the query image and the ones in the training set rather than similarity among whole images.

The VISITO Tuscany project is funded by the Tuscan Region and is coordinated by ISTI-CNR. The consortium includes three ISTI-CNR laboratories (Networked Multimedia Information Systems, Visual Computing and High Performance Computing), the security laboratory of IIT-CNR, and three private companies: Alinari24Ore, Hyperborea, and 3Logic MK. We thank the municipalities of Florence, Pisa, and San Gimignano for providing us with all authorizations to build the demonstrator.

Links:
[1] http://www.visitotuscany.it
[2] https://market.android.com/details?id=it.visitotuscany
[3] http://www.google.com/mobile/goggles

Please contact:
Giuseppe Amato
ISTI-CNR, Italy
Tel: +39 050 3152810
E-mail: giuseppe.amato@isti.cnr.it

National Gallery in Prague

by Stanislav Mikeš, Michal Haindl and Radek Holub

Scientists at Institute of Information Theory and Automation (UTIA), Academy of Sciences of the Czech Republic have developed a detailed virtual model of the Department of Modern Art of the National Gallery in Prague. This impressive seven-storey building holds the gallery's collection of contemporary images, drawings and statues as well as several temporary exhibitions. This virtual model serves as a comprehensive 3D information system with navigation support for visitors and as an interactive tool for exhibition designers and curators. Within the comfort of their own home, visitors can experience an animated online thematic visit to their selected works of art and also print a map with a proposed personalized route.

While navigation in real world, ie travelling to a specific target location, often poses a challenging and only partially-understood problem, especially in unknown environment, navigation in a virtual reality (VR) environment is even more difficult due to many missing real world cues. A major problem for users of virtual environments is maintaining knowledge of their location and orientation while they move through the space because perceptual judgements are biased within a virtual environment.

The proposed solution for navigation in this huge gallery building, which has seven exhibition floors and two large exhibition halls in the ground floor, is based on the graph structure. One's actual position in a complex virtual scene is depicted as a highlighted point in an overlaid transparent map of the building floor plan. This basic navigation graph structure is constructed semi-automatically and it is subsequently locally changed by the exhibition editor which places new exhibition panels into the building interior and thus locally changes the navigation route structure. The optimal navigation route is automatically generated using graph algorithms and user defined constraints. For example, we assume that a visitor will never walk closer than half a metre to the walls, will pass through each exit in its centre and larger spaces are covered with walking loops with a minimum diameter of one metre, etc. Each floor plan is than supplemented with a preset route graph structure based on the basic

Figure 1: National Gallery in Prague virtual model.

building structure. Single corridors, lifts and staircases are represented as graph edges, while doors, branching or turning points are graph vertices. This initial graph structure which represents initialization of navigation routes can be generated semi-automatically based on the floor plans. Narrow corridors have single graph edges while wider corridors or halls can have several graph loops. This automatically proposed graph structure (primary graph) can be interactively edited using the exhibition editor. Superfluous edges or vertices can be removed while new edges and vertices can be added. Single edges or vertices can also be shifted to other positions and vertices can be supplemented with additional attributes such as emergency exits, lifts, staircases and doors.

The virtual National Gallery model allows virtual exhibitions to be built interactively using our exhibition editor. The editor was devised for exhibition architects to support and speed up their exhibitions proposals. The editor loads a requested floor plan and allows insertion of single exhibition panels and specification of their parameters such as single dimensions, colour and covering material. Single paintings from the gallery database are subsequently set out on these exhibition panels and other supplementary data can be attached, for example information about a painter in the corresponding pop-up window. When the exhibition editing is ready, it is exported into the VRML building model and can be immediately checked in the browser either walking around or in the generated walk-through movies.

Generation of a path based on parameters provided by the user is performed automatically in a module that considers the ground plan of a 3D scene as a labelled graph. The labels represent various kinds of information such as accessibility of a particular location for disabled people. Single edge attribute also encompasses physical length, thus it is possible to estimate the real time needed to walk a specified route in the real National Gallery of Prague as well as the time needed for an exhibition sightseeing tour. This navigation route is subsequently used for generation of a movie that represents the virtual walkthrough. This walkthrough can be demonstrated using avatars or simulating the viewpoint of a visitor. Visitors can watch not only an animated thematic visit to their selected works of art from home over internet, but also print a map with a proposed personalized route. It is also possible to print the floor plan with the suggested route highlighted. If disabled or wheel chair access is required, the generated route will avoid staircases, instead taking a slightly longer route via the lift.

Please contact:
Stanislav Mikeš
UTIA ASCR - CRCIM, Czech Republic
Tel: +420 266052262
E-mail: xaos@utia.cas.cz

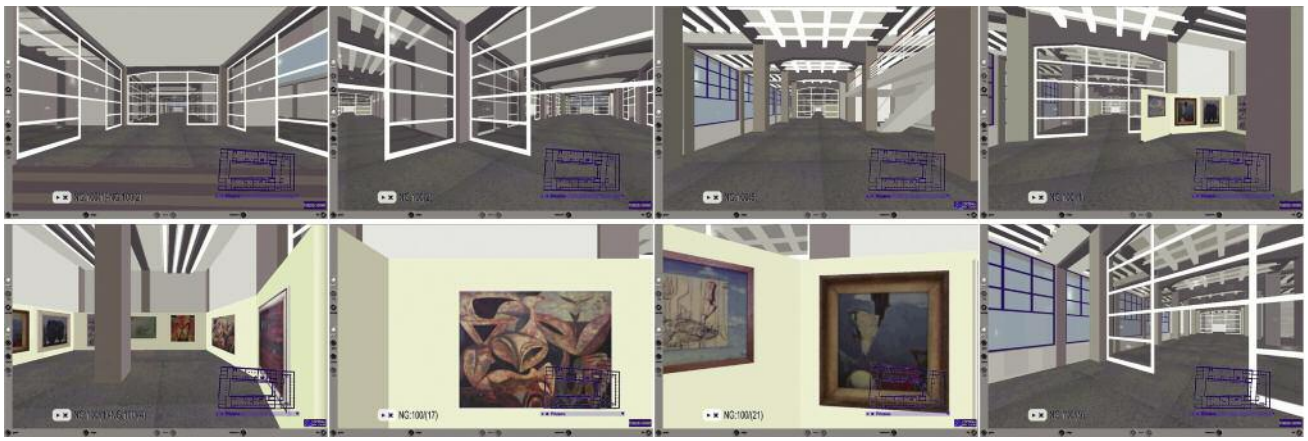


Figure 2: Selected navigation frames (2,5,11,25 top, 27,32,38,42 bottom) around a virtual exhibition.

Macedonia: From Fragments to Pixels

by Xenophon Zabulis, Dimitrios Grammenos, Antonis A. Argyros, Dimitrios Sifakis and Constantine Stephanidis

Forget all about the traditional "Do Not Touch" museum rule! In 2010, the Institute of Computer Science of the Foundation for Research and Technology – Hellas (ICS-FORTH) and the Archaeological Museum of Thessaloniki (AMTh) collaborated in the creation of a special exhibition of prototypical interactive systems with subjects drawn from ancient Macedonia, named "Macedonia from fragments to pixels" (see link below). The exhibition is hosted by the AMTh at its premises and is open to the general public.

The exhibition comprises seven interactive systems which are based on some of the research outcomes of the Ambient Intelligence Programme of ICS-FORTH. The digital content of these systems includes objects from the Museum's permanent collection and other objects related to Macedonia. In brief, the interactive systems are:

1. *Panoptes*, a digital exhibition-catalogue system which allows visitors to browse its content and examine details of images, read accompanying texts and follow threads of information. A children's windmill is located between two touch screens. Blowing on the windmill, the largest collection of gold wreaths in the world unfolds before the visitor. By pausing on any image, additional information is provided on about the wreath and its symbolisms.
2. *Cryptolex*on, the hidden crossword, is a game loved by all ages, which combines entertainment with education. The names of ancient gods and heroes are hidden within a matrix of letters for the visitors to discover.
3. *Peridexion* offers museum visitors the possibility of exploring an object and/or a subject in depth. In this particular case, the subject is a masterpiece of 6th c. BC Athenian black-figured pottery, known as the Crater of Lydos, which portrays the legend of the hunt of a monstrous boar in Calydonia, Aetolia. The system has a touch screen that presents a view of an object at a time. Visitors can select the object or the view they wish to see and then discover points of interest and relative multimedia information, or zoom in on any detail at will.
4. *Polyapton*, a large interactive screen that can be used by several visitors who all wish to explore simultaneously multifaceted information on a subject. It presents one of the best-preserved ancient Greek paintings, from the Macedonian tomb of Agios Athanasios near Thessaloniki depicting a symposium. The tomb is not open to visitors, so Polyapton provides a rare opportunity to enjoy it. The multi-touch screen recognizes the touch of many fingers or hands, but also specific objects, at the same time. Visitors can 'scroll' across the painting and focus on points of interest with multimedia information by touching them, zoom in on details with the use of a paper magnifying lens, while an infrared torch displays a modern rendition of the painting.
5. *Multimodal Diverse Travel* comprises a table, whose surface is covered by a printed map on which the location of various cities and other notable sites is projected. White paper tablets with a coloured frame are at the visitors' disposal. When the magnifying glass is placed over a city, related images, videos and texts appear on the tablet. For every city there are multiple information "pages", which can be viewed by touching virtual buttons at the bottom of the tablet.

Figure 1: Two young visitors exploring multiple layers of information in the electronic reproduction of a wall painting with Polyapton. Besides a finger touch, the system is sensitive to the touch of particular objects with different functionalities.

Figure 2: The Minister of Culture and Tourism of Greece explores locations of archaeological interest in Multimodal Diverse Travel. As the users move rectangular pieces of paper on the map, location-based, interactive multimedia content appears on them.

Figure 3: Macrographia presents a large artifact in actual size and updates visual content based on the location and walkthrough trajectories of visitors.

6. *One day in a farmstead* allows the public to visit an ancient farmstead that has been excavated at Asprovalta, near Thessaloniki, by enriching with multimedia information a scale model realized by the AMTh. By moving a white paper tablet over the areas of the farmstead, the visitor can view and learn more about the excavation findings and gain details on ancient rural life.
7. *Macrographia* is a system that presents very large images, which visitors can explore by walking around in a room. One of the walls of this room comprises the projection screen. The projection of content on the screen depends on the location of each visitor in the room. Visitors enter the room and the system follows the movement of each one separately. The painting is divided into five sections, which correspond to the prey of the hunters. When someone stands in front of a section, depending on the distance from the screen, the image she/he views and the caption underneath change. There are four levels of information: the present state of the painting, an artist's sketch, an artist's modern rendition, and notable details.

The exhibition is very different from an archaeological exhibition, as the public can have a novel relationship with exceptional artifacts of the past through the use of new technologies, enjoying an interactive experience that combines information and learning with entertainment. In this way, visitors have the opportunity to approach classical antiquity in a novel manner: by exploring digital reproductions of ancient masterpieces. Rare and fragile artifacts are now at the public's fingertips with the use of modern and user friendly technology. All systems are multilingual and support interaction by one or multiple visitors.

This work was supported by the "Ambient Intelligence Programme" of ICS-FORTH.

Link:
[1] Macedonia: from fragments to pixels, 2010: http://www.makedonopixels.org

Please contact:
Constantine Stephanidis
ICS-FORTH, Greece
Tel. : +30 2810 391741
E-mail: cs@ics.forth.gr

Discovering Knowledge from Sumerian Economic Documents

by Wojciech Jaworski

A vast amount of knowledge is contained in large collections of unstructured or weakly structured text documents, which started to emerge soon after the discovery of writing. We develop a methodology, which allows users to seek not only for information localized in specific documents but also knowledge spread across an entire document collection.

Our recent research has focused on the Sumerian Economic Text Corpus from the Ur III period. Sumerians lived from prehistoric times until late 3rd millennium BC in lower Mesopotamia (modern Iraq). Sumer was the first highly developed urban civilization, which used cuneiform script. During the reign of the 3rd dynasty of Ur (2100 BC-2000 BC), whose power extended as far as present Iraq and western Iran, the state introduced a centrally planned economy with an extensive bureaucratic structure.

Civil servants used clay tablets to record data about agriculture and factory production, worker salaries, summaries of executed work, distribution of commodities, goods and animals, lists of sacrificed animals, travel diets and other economical information.

Archaeologists have excavated about 100 000 tablets from this period. A corpus of over 45 000 tablets is available electronically, stored in the form of Latin transliteration (ie documents are represented using Latin alphabet and each cuneiform sign is replaced by its reading). For our studies, we have selected a subcorpus of 11 891 documents concerning distribution of domestic animals. This subcorpus consists of circa 850 000 Sumerian signs, each representing either a word or a syllable.

Figure 1 presents the contents of a typical Sumerian document. This document reports the transfer of lambs from three people to ab-ba-sa6-ga, an official of the Ur III state. The transfer took place on the 23rd day of the month sze-kin-ku5 in the year when the high priest of goddess Innana was elevated to office. The third verse of the document is ambiguous. We cannot determine whether ga is a part of a name or a part of an animal description.

Economic documents are an essential source of information about ancient Sumer. The corpus contains crucial information about economic, social and political history of the state, as well as its political system and administration structure. Sources of this type provide the most complete information about the daily life of those days.

Owing to the large number of documents, the task of finding relevant ones is intractable for human readers. On the other hand classical information retrieval techniques fail when confronted with the Sumerian language. Our search engine, dedicated to Sumerian Economic documents, offers a solution to these problems. However, vital information is spread across a vast number of simple documents. In order to extract it we must process document contents into computer understandable format.

We take advantage of the fact that the Ur III Economic Text Corpus has restricted subject-matter which allows

```
&P123831 = OIP 121, 101
tablet
obverse
1.  1(disz) sila4 ur-mes ensi2          1 lamb ur-mes governor
2.  1(disz)# sila4 da-da dumu lugal     1 lamb da-da son of king
3.  1(disz)# sila4 ga-ga-mu             1 milky lamb ga-mu or 1 lamb ga-ga-mu
reverse
1.  u4 2(u) 3(asz@t)-kam                Day 23
$ 1 line blank
3.  mu-DU                               delivery
4.  ab-ba-sa6-ga i3-dab5                ab-ba-sa6-ga received
5.  iti sze-KIN-ku5                     month sze-kin-ku5
6.  mu en {d}inanna ba-hun              Year when high priest of goddess Innana
                                                          was elevated to office
left
1.  3(disz)                             3
```

Figure 1: An example of transliterated cuneiform tablet from UR III.

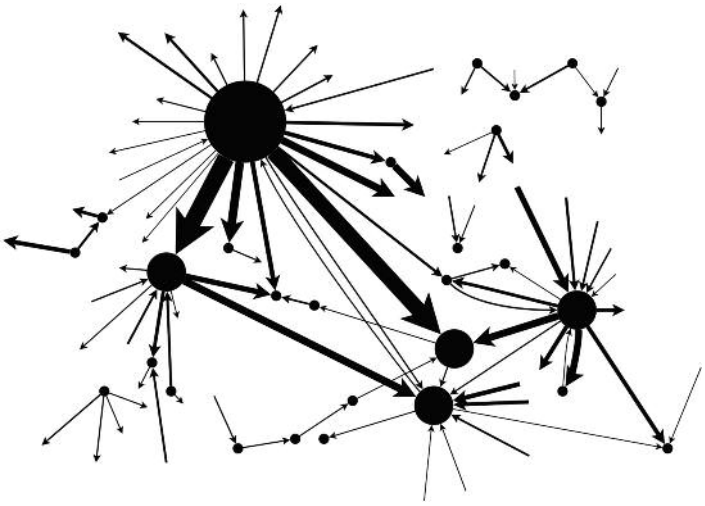


Figure 2: Animal flows between the officials in UR III Kingdom.

us to represent the structure of information enclosed in documents by means of ontology. This ontology represents the domain knowledge: it splits the set of objects described in documents into categories (types) and it determines relationships between these categories.

We process documents by means of a grammar which describes the way in which phrases are constructed from words and other simpler phrases. We assume that syntactic operations constructing compound phrases on the basis of simpler ones correspond to the presentation of complex objects by means of their components. As a result of parsing we obtain a logical formula which represents the contents of the document.

The content of all documents constitutes a vast knowledge base of Sumerian economy. We use it to determine relationships between Sumerian officials in terms of number of animals that were transferred between them. We represent this information in terms of the graph of animal flow. Vertices of the graph represent officials. Graph edge width is proportional to the number of animals transferred between individuals. In figure 2, we present a fragment of the animal flow graph. We selected edges labelled with animal quantities greater than 900. The complete graph that encloses all extracted transactions has 2754 vertices and 5275 edges.

Among other possible applications of our knowledge base we explored:

- observation of seasonal economic fluctuations as well as macro-economic changes that happened during the Ur III period;
- detection documents that describe the same object or event;
- reconstruction of broken documents and determining contents of missing ones.

We also develop a query language that allows us to retrieve information according to semantic patterns.

In future we plan to pursue the ultimate goal of creating a model of the Sumerian economy.

Links:
http://www.ur3.historia.uw.edu.pl
http://www.mimuw.edu.pl/~wjaworski

Please contact:
Wojciech Jaworski
Institute of Informatics, University of Warsaw, Poland
E-mail: wjaworski@mimuw.edu.pl

Multimodal Computer-Assisted Transcription of Ancient Documents

by Verónica Romero, Alejandro Héctor Toselli, Enrqiue Vidal, Elsa Cubel and Joan Andreu Sánchez

A multimodal interactive approach for transcription of ancient documents is proposed. In this approach, user's feedback directly facilitates improvements to system accuracy while multimodality increases system ergonomy and user acceptability.

Huge historical document collections residing in libraries, museums and archives are currently being digitalized for preservation purposes and to make them available worldwide through large on-line digital libraries.The number of these on-line digital libraries is dramatically increasing. This is in part due to the ever-decreasing costs of digital storage devices and to recent advances in image digitalization and processing technologies. Thanks to these advances, hundreds of terabytes worth of ancient document digital images have been collected and can be easly made available to the historians and the public alike. However, such huge amounts of data are only of very limited use in their present raw digital image form and current efforts focus on technologies aimed at reducing the human effort required for the annotation of the raw images with informative content. In the case of text images, which are among the most numerous and interesting, the most informative annotation level is their (palaeographic) transcription into an adequate textual electronic format that would provide new ways of indexing, consulting and querying these documents.

Figure 1: Using the MM-CATTI system with a touch-screen.

Given the kind of (typically handwritten) text images involved in ancient (or even more recent historic) documents, state-of-the-art Handwritten Text Recognition (HTR) technology is very far from offering useful solutions to the transcription problem and heavy human intervention is often required to check and correct the results. This post-editing process is both inefficient and inconvenient to the user.

As an alternative to fully manual transcription and post-editing, a multimodal interactive approach is proposed here where user feedback is provided by means of touch-screen pen strokes and/or more traditional keyboard and mouse operation. User's feedback directly facilitates improvements in system accuracy, while multimodality increases system ergonomy and user acceptability. Multimodal interaction is approached in such a way that both the main and the feedback data streams help each other to optimize overall performance and usability.

In this new multimodal interactive approach for transcription of text images (Multimodal Computer-Assisted Transcription of Text Images - MM-CATTI), the user is directly involved in the transcription process, in which following a preset protocol, he/she validates and/or corrects the HTR output during the process. The protocol that rules this interaction process is formulated in the following steps: The HTR system proposes a full transcription of the input handwritten text line image. Then, the user validates the longest prefix of the transcription which is error-free and enters some on-line touch-screen pen-strokes and/or some amendment keystrokes to correct the first error in the suffix. An on-line HTR feedback subsystem (or HFR) is used to decode this input. In this way, a new extended consolidated prefix is produced based on the previous validated prefix, the on-line decoding word and the keystroke amendments. Using this new prefix, the HTR suggests a suitable continuation of it. These previous steps are iterated until a final, perfect transcription is produced.

MM-CATTI is shown to work quite well by an implemented Web-based Demo (http://cat.iti.upv.es/iht/). Figure 1 shows a user interacting with the MM-CATTI system by means of a touch-screen. The on-line form of such MM-CATTI system allows collaborative tasks with thousands of users across the globe to be carried out, thus reducing notably the overall image recognition process. Since the users operate within a web browser window, the system also provides cross-platform compatibility and requires no disk space on the client machine.

To test the effectiveness of the MM-CATTI approach, experiments were carried out on several corpora corresponding to different handwritten text transcription tasks. From the reported results on these corpora and assuming for simplicity that the cost of correcting an on-line decoding error is just similar to that of another on-line touch-screen

interaction, the estimated human effort to produce error-free transcription using MM-CATTI is reduced by as much as 15% on average, regarding to the classical HTR system. In other words, from every 100 words misrecognized by a conventional HTR system, a human post-editor will have to correct all the 100 erroneous words, while a MM-CATTI user would correct only 85 – the other 15 are corrected automatically by the MM-CATTI system.

This multimodal interactive paradigm has been proposed by the Pattern Recognition and Human Language Technology (PRHLT) group from the Universidad Politécnica de Valencia (UPV). PRHLT is currently involved in the use of interactive techniques for machine translation and transcription through the large-budget MIPRCV Spanish project (http://miprcv.iti.es). The MIPRCV project is leaded by the PRHLT research group. MIPRCV establishes a five-year research programme to develop pattern recognition approaches that explicitly deal with the challenges and opportunities entailed by the human-interaction paradigm.

Please contact:
Joan Andreu Sánchez
Universidad Politécnica de Valencia, Spain
Tel: +34 963877253
E-mail: jandreu@dsic.upv.es

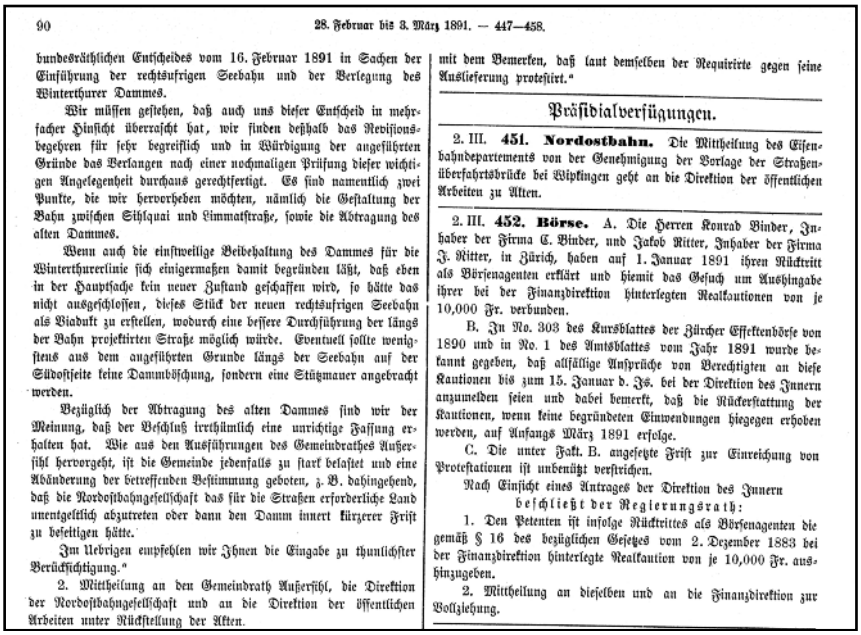
Reducing OCR Errors in Gothic-Script Documents

by Lenz Furrer and Martin Volk

In order to improve optical character recognition (OCR) quality in texts originally typeset in Gothic script, an automatic correction system can be built to be highly specialized for the given text. The approach includes external dictionary resources as well as information derived from the text itself.

The resolutions by the Zurich Cantonal Government from 1887 to 1902 are archived as printed volumes in the State Archive of Zurich. The documents are typeset in Gothic script, also known as blackletter or Fraktur. As part of a collaborative project between the State Archive and the Institute of Computational Linguistics at the University of Zurich, these texts are being digitized for online publication.

The aims of the project are automatic image-to-text conversion (OCR) of the approximately 11,000 pages, the segmentation of the bulk of text into separate resolutions, the annotation of metadata (such as the date or the archival signature) as well as improving the text quality by automatically correcting OCR errors. From an OCR perspective, the data are most challenging, as the texts contain not only Gothic type letters – which lead to a lower accuracy compared to antiqua texts – but also particular words, phrases and even whole paragraphs printed in antiqua font. Although we were lucky to have available an OCR engine capable of processing mixed Gothic and antiqua texts, the alternation of the two fonts still has an impairing effect on the text quality. Since the interspersed antiqua tokens can be very short (eg the abbreviation 'Dr.'), their diverting character is sometimes not recognized. This leads to badly misrecognized words due to the quite different shapes of the typefaces; for example antiqua 'Landrecht' (Engl.: citizenship) is rendered as completely illegible 'I>aii<lreclitt', which is clearly the result of using the inappropriate recognition algorithm for Gothic script.



90 28. Februar bis 3. März 1891. — 447—458.

bundesräthlichen Entscheides vom 16. Februar 1891 in Sachen der Einführung der rechtsufrigen Seebahn und der Verlegung des Winterthurer Dammes.

Wir müssen gestehen, daß auch uns dieser Entscheid in mehrfacher Hinsicht überrascht hat, wir finden deßhalb das Revisionsbegehren für sehr begreiflich und in Würdigung der angeführten Gründe das Verlangen nach einer nochmaligen Prüfung dieser wichtigen Angelegenheit durchaus gerechtfertigt. Es sind namentlich zwei Punkte, die wir hervorheben möchten, nämlich die Gestaltung der Bahn zwischen Sihlquai und Limmatstraße, sowie die Abtragung des alten Dammes.

Wenn auch die einstweilige Beibehaltung des Dammes für die Winterthurerlinie sich einigermaßen damit begründen läßt, daß eben in der Hauptsache kein neuer Zustand geschaffen wird, so hätte das nicht ausgeschlossen, dieses Stück der neuen rechtsufrigen Seebahn als Viadukt zu erstellen, wodurch eine bessere Durchführung der längs der Bahn projektirten Straße möglich würde. Eventuell sollte wenigstens aus dem angeführten Grunde längs der Seebahn auf der Südostseite keine Dammböschung, sondern eine Stützmauer angebracht werden.

Bezüglich der Abtragung des alten Dammes sind wir der Meinung, daß der Beschluß irrthümlich eine unrichtige Fassung erhalten hat. Wie aus den Ausführungen des Gemeindrathes Außersihl hervorgeht, ist die Gemeinde jedenfalls zu stark belastet und eine Abänderung der betreffenden Bestimmung geboten, z. B. dahingehend, daß die Nordostbahngesellschaft das für die Straßen erforderliche Land unentgeltlich abzutreten oder dann den Damm innert kürzerer Frist zu beseitigen hätte.

Im Uebrigen empfehlen wir Ihnen die Eingabe zu thunlichster Berücksichtigung."

2. Mittheilung an den Gemeindrath Außersihl, die Direktion der Nordostbahngesellschaft und an die Direktion der öffentlichen Arbeiten unter Rückstellung der Akten.

mit dem Bemerken, daß laut demselben der Requirirte gegen seine Auslieferung protestirt."

Präsidialverfügungen.

2. III. **451. Nordostbahn.** Die Mittheilung des Eisenbahndepartements von der Genehmigung der Vorlage der Straßenüberfahrtsbrücke bei Wipkingen geht an die Direktion der öffentlichen Arbeiten zu Akten.

2. III. **452. Börse.** A. Die Herren Konrad Binder, Inhaber der Firma C. Binder, und Jakob Ritter, Inhaber der Firma J. Ritter, in Zürich, haben auf 1. Januar 1891 ihren Rücktritt als Börsenagenten erklärt und hiemit das Gesuch um Aushingabe ihrer bei der Finanzdirektion hinterlegten Realkautionen von je 10,000 Fr. verbunden.

B. In No. 303 des Kursblattes der Zürcher Effektenbörse von 1890 und in No. 1 des Amtsblattes vom Jahr 1891 wurde bekannt gegeben, daß allfällige Ansprüche von Berechtigten an diese Kautionen bis zum 15. Januar d. Js. bei der Direktion des Innern anzumelden seien und dabei bemerkt, daß die Rückerstattung der Kautionen, wenn keine begründeten Einwendungen hiegegen erhoben werden, auf Anfangs März 1891 erfolge.

C. Die unter Fakt. B. angesetzte Frist zur Einreichung von Protestationen ist unbenützt verstrichen.

Nach Einsicht eines Antrages der Direktion des Innern

beschließt der Regierungsrath:

1. Den Petenten ist infolge Rücktrittes als Börsenagenten die gemäß § 16 des bezüglichen Gesetzes vom 2. Dezember 1883 bei der Finanzdirektion hinterlegte Realkaution von je 10,000 Fr. auszuhingeben.

2. Mittheilung an dieselben und an die Finanzdirektion zur Vollziehung.

Figure 1: Sample page of the processed texts in Gothic script.

The main emphasis of the RRB-Fraktur project is on the post-correction of recognition errors. The evaluation of a limited number of text samples yielded a word accuracy of 96.96 %, which means that one word out of 33 contains an error (for example, 'Regieruug' instead of correct 'Regierung', Engl.: government). We aim to significantly reduce the proportion of misrecognized word tokens by identifying them in the OCR output and determining the originally intended word with its correct spelling. The task resembles that of a spell-checker as found in modern text processing applications, with two major differences. First, the scale of the text data (with an estimated 11 million words more than 300,000 erroneous word forms are to be expected) does not allow for an interactive approach, asking a human user for feedback on every occurrence of a suspicious word form. Therefore we need an automatic system that can account for corrections with high reliability. Second, dating from the late 19th century, the texts show historic orthography, which differs in many ways from the spelling

encountered in modern dictionaries (eg historic 'Mittheilung' versus modern 'Mitteilung', Engl.: message). This means that using modern dictionary resources directly, cannot lead to satisfactory results. Additionally, the governmental resolutions typically contain many toponymical references, which are not covered by general-vocabulary dictionaries. Regional peculiarities are also evident, for instance pronunciation variants or words not common elsewhere in the German speaking areas (eg 'Hülfstrupp' versus the standard German 'Hilfstruppe', Engl.: rescue team), and of course there is a great amount of genre-specific vocabulary, ie administrative and legal jargon. We are hence challenged to build a fully-automatic correction system with high precision that is aware of historic spelling and regional variants and contains geographical and governmental language.

The core task of the desired correction system with respect to its precision is the categorization routine that determines the correctness of every word. For example, 'saumselig' (Engl.: dilatory) is a correct word, whereas 'Gefundheit' is not (in fact, it is misrecognized for 'Gesundheit', Engl.: health). We use a combination of various resources for this task, such as a large dictionary system for German that covers morphological variation and compounding (such as 'ging', a past form of 'gehen', Engl.: to go, or 'Niederdruckdampfsystem', a compound of four segments meaning low-pressure steam system), a list of local toponyms, the recognition confidence values of the OCR engine and more. Every word is either judged as correct or erroneous according to the information we gather from the various resources. The historic and regional spelling deviations are modelled with a set of handwritten rules describing the regular differences. For example, with a rule stating that the sequence 'th' corresponds to 't' in modern spelling, the standard form 'Mitteilung' can be derived from old 'Mittheilung'. While the latter word is not found in the dictionary, the derived one is, which allows for the assumption that 'Mittheilung' is a correctly spelled word.

The set of all words recognized as correct words and their frequency can now serve as a lexicon for correcting erroneous word tokens. This corpus-derived lexicon is naturally highly genre-specific, which is desirable. On the other hand, rare words are likely to occur only in a misspelled version, in which case there will be no correction candidate in the lexicon. Due to the repetitive character of the text's topic there is also a lot of repetition in the vocabulary across the corpus. This increases the chance that a word misrecognized in one place will have a correct occurrence in another.

In an evaluation the approach has shown promising results. The text quality in terms of word accuracy could be improved from 96.96 % to 97.73 %, which is equivalent to a 25 % reduction in the rate of word misrecognition. As the system is far from complete, considerably improved results can be expected.

With this project we demonstrate that it is possible to build a highly specialized correction system for a specific text collection. We are using both general and specific resources, while the approach as a whole is widely generalizable. We see our work as an auspicious method for improving text quality of historic OCR-converted text.

Link:
http://www.cl.uzh.ch/research/digitalisierungunderschliessung/digitalisierungUndAuszeichnungDerZuercherRegierungsratsbeschluesse_en.html

Please conact:
Lenz Furrer
University of Zurich, Switzerland
Tel: +41 44 635 67 38
E-mail: Lenz.Furrer@uzh.ch

Martin Volk
University of Zurich, Switzerland
Tel: +41 44 635 43 17
E-mail: volk@cl.uzh.ch

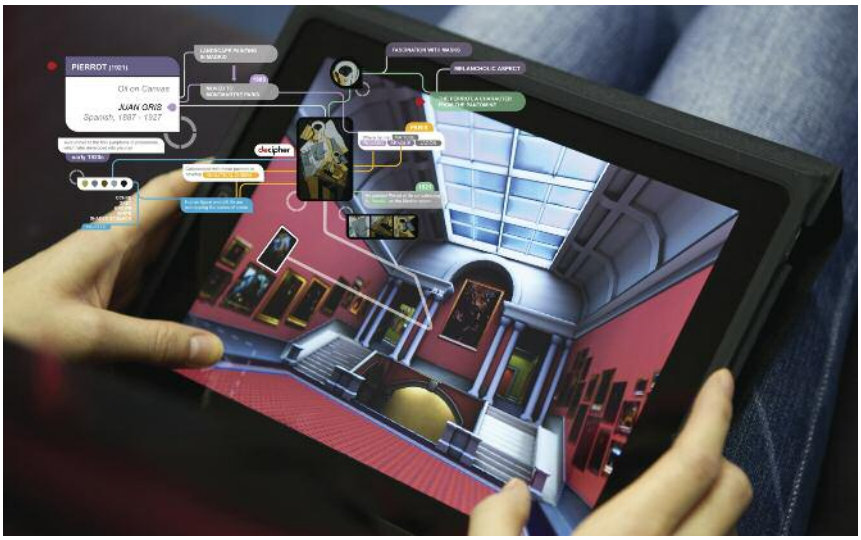
Deciphering the Story of the Museum

by Eoin Kilfeather

Cultural meaning does not reside in individual objects but in the patterns of knowledge and events, belief and thought that link them to each other and to the observer. This is why narrative is so important to the communication of, and meaningful understanding, of culture.

Digital heritage and semantic web technologies have held out the promise of nearly unlimited access to cultural knowledge. However, the vision is as beguiling – and currently as unobtainable – as the invocation in John Donne's Song (see Figure 1). The problem is that cultural meaning does not reside in individual objects (however beautiful and marvellous in themselves) but in the patterns of knowledge and events, belief and thought that link them to each other and to the observer. It is well understood that cultural narratives are not static, but change in time and in relation to the viewpoints from which they are both created and explored. Curators have developed ways of presenting their collections that guide visitors and make sense of the experience. Although we have new ways of accessing cultural objects, the narrative structures and arguments that could be found in a handcrafted presentation are often lost. There is a tension between providing user control over what is shown and the coherence of the overall experience.

Although there is a vast and growing store of digital objects, most of them do not provide the information for explicit interpretation. On the other hand many cultural websites are extensive and elegant, but their narratives are 'burnt-in' and fixed by their authors: they can never be complete, and there is a good chance that the searcher's interests will not be fully met. Also, current systems are tied to the technology platforms for which they are authored. Web-based approaches offer a two dimensional sensory experience that is limited in com-



SONG

GOE, and catche a falling starre,
Get with child a mandrake roote,
Tell me, where all past yeares are,
Or who cleft the Divels foot,
Teach me to heare Mermaides singing,
Or to keep off envies stinging,
And finde
What winde
Serves to advance an honest minde.

Figure 1 (top): A Donne' poem.

Figure 2 (left): A mock-up of a visualization.

parison with the real world; immersive and haptic systems still require expensive special-purpose hardware, whose use is restricted to small numbers of people at the heritage sites themselves.

The DECIPHER project proposes new solutions to the whole range of narrative construction, knowledge visualization and display problems. It will combine much richer, event-based metadata with causal reasoning models. This will result in a reasoning engine, virtual environment and interfaces that can present digital heritage objects as part of a coherent narrative, directly related to individual searches and user contexts. This will allow the user to interactively assemble, visualize and explore, not just collections of objects, but the knowledge structures that connect and give them meaning.

The partners bring together the skills and experience in the technical fields required by the project, with the authorial and curatorial authority of national institutions, the innovative impetus of a technology-based SME, and the drive to market of a large company that combines heritage and media interests. They are: 1) Two national heritage institutions, the National Gallery of Ireland and the Irish Museum of Modern Art, with major collections that span the range of physical and born digital art works. 2) A large company, Alinari 24 Ore (IT), which is both the world's oldest photographic archive and a part of a large, modern media group. 3) A research-based SME System Simulation Ltd. (UK), specialising in software for professional media and content management. 4) Three leading academic research centres, Dublin Institute of Technology (IE), the Open University's Knowledge Media Institute (UK), and Brno University of Technology (CZ) expert in knowledge management, narrative construction, data mining, semantic web technologies, language technologies, multimodal interfaces and digital heritage applications.

Background work carried out by KMi has succeeded in uncovering interesting patterns from data by using a combination of machine reasoning, knowledge visualization and user tagging. DECIPHER will produce an advance in the process by combining much richer, event-based metadata with causal and dynamic reasoning models. We will do this by associating an object with a set of events, making it possible to apply machine reasoning in a much more interesting way, not only to identify clusters of events or objects but also to construct narratives.

To do this we need to generate the event-based metadata. In this project, colleagues in Brno University of Technology are producing the event-based semantic metadata automatically from networked sources and the traces of social interaction. Artificial intelligence techniques, using Kripke structures (a type of nondeterministic finite state machine proposed by Saul Kripke in 1963 to represent the behaviour of a system)and Kripke semantics (a formal semantics for non-classical logic systems, also known as relational semantics or frame semantics) then make it possible to identify and formally represent regions and tipping points in sets of content and related events. By maintaining and accruing the metadata associated with the objects and transitions, the machine representation becomes richer and more powerful. This will make it possible to see if there are relevant transitions or regions in other people's narratives, thereby enabling social interpretations and recommendations. It will also support counterfactual or 'what if' reasoning by showing the effect of moving or removing a transition.

Clearly, it is not enough to develop a narrative: we must present it in a way that allows the visitor to explore the knowledge space and objects in the richest way possible, by taking full advantage of whatever technology platform is available (fixed or mobile, 2D or 3D, window or immersive). The visual formalisms, interfaces and interaction techniques draw on DIT's previous work in mobile interaction and virtual museum spaces. Customised rich interfaces support digital curation in other ways. For example, some installation art consists of a digital file and the curatorial instructions for "how to show it." Such a system that links media resources, curatorial metadata and multi-modal interfaces would allow a much deeper exploration of such works.

The DECIPHER project (FP7-270001-Decipher) began in January 2011 and runs for three years. It is supported by the EU's 7th Framework Programme.

Link:
http://decipher-research.eu/

Please contact:
Eoin Kilfeather
Digital Media Centre, Dublin, Ireland
E-mail: ekilfeather@dmc.dit.ie

Non-linear Digital Storytelling for the Battleship "G. Averof"

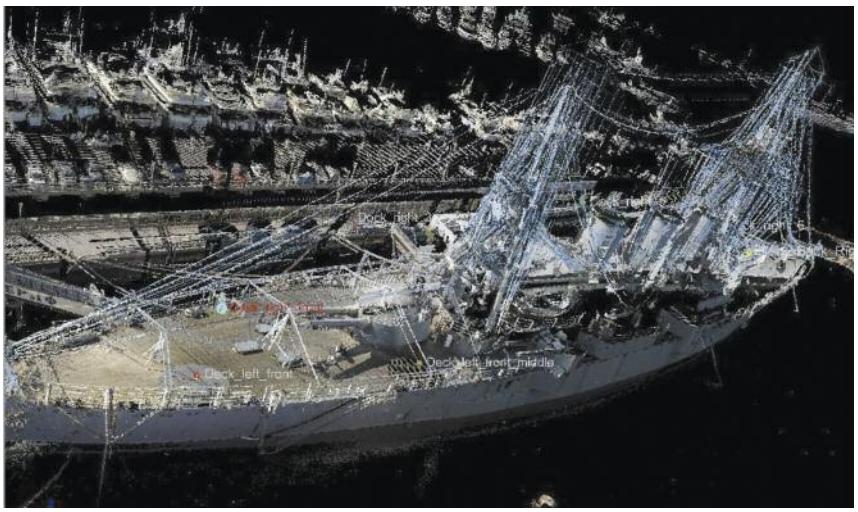
by Ralf Klamma and Yiwei Cao

Digital multimedia, web and mobile technologies bring new experiences to museums and cultural heritage management. Within the German-Greek project of non-linear digital storytelling for the battleship "G. Averof", advanced storytelling approaches were applied to the museum with a large multimedia archive on the historic battleship.

The battleship "G. Averof" is the world's only surviving heavily armoured cruiser of the early 20th century and serves as a museum operated by the Greek Navy in Faliron today. Non-linear digital storytelling for the battleship "G. Averof" was an interdisciplinary research project between the Chair of Information Systems and Databases at RWTH Aachen University, Aachen, Germany and Harokopio University, Athens, Greece. Starting from 2009, this project was supported by the IKYDA program, an integrated action program between the German Academic Exchange Service (DAAD) and the Greek State Scholarship Foundation (I.K.Y). The Greek partners contributed their expertise in geographic information systems and cultural heritage while the German partners provided their experience in multimedia information systems design. Matthias Jarke, Ralf Klamma, Yiwei Cao as well as a number of research assistants and students visited Harokopio University, while Emmanuel Stefanakis, Eleni Gadolou and Haroula Papadaki visited RWTH Aachen University to exchange knowledge in a couple of workshops.

The research project aimed at the promotion and enrichment of the museum archives for a new generation of museum visitors used to web and mobile technologies. With our advanced web-based storytelling environment "Virtual Campfire" we were able to reuse the large multimedia collection about the battleship "G. Averof" as well as to trace the journeys of the flagship through time and space by the utilization of digital map technologies. Moreover, we created a Web 2.0 community around the traditional media where visitors can add their own media and stories, thus creating new perspectives on the daily life on a battleship. To stimulate this, we created multimedia stories documenting the coal loading process for the ship as well as the battle of Lemnos. The stories could be consumed and refined in every web browser but also on mobile smart phones.

Mobile Campfire, the mobile version of Virtual Campfire has been launched on iTunes App Store. Mobile Campfire enables user communities to create, annotate, search, and share photos and multimedia stories on iPhone, iPod and iPad. Mobile Campfire extends the semantic-enhanced video annotation service for Virtual Campfire to enable collaborative tagging of videos and photos.



"G. Averof" Battleship.

A group of research assistants and from Aachen visited Harokopio University in November 2009 to document the cultural heritage for future projects. We did a 3D laser scanning campaign with the Riegl LMS-Z390i 3D scanner with over 20 scans from different viewpoints and created the only 3D digital model of the "G. Averof".

The story of the battleship "G. Averof" continues. Interdisciplinary cooperation enhances the applicability of research methods in cultural heritage management. Upon request of the Culture Section of the German Foreign Office an evaluation of the historic city walls of Ghazni in Afghanistan, the future Islamic Culture Capital in 2013, has recently been conducted by cultural heritage experts from the RWTH Aachen Center of Documentation and Conservation. All these activities have addressed new community requirements for Virtual Campfire and are now under consideration, such as cloud based video processing for cultural sites. In such environments, the documentation often has to be performed with cheap and available hardware. Processing costs and time of multimedia documents can be dramatically improved by using modern processing paradigm like cloud computing.

Links:
http://dbis.rwth-aachen.de/cms/projects/IKYDA
http://dbis.rwth-aachen.de/cms/projects/virtualCampfire
http://itunes.apple.com/de/app/mobile-campfire/id365323266?mt=8

Please contact:
Ralf Klamma
RWTH Aachen University, Germany
E-mail:
ralf.klamma@dbis.rwth-aachen.de

YUMA – Crowd Sourced Metadata Enrichment for Online Collections

by Joachim Jung, Rainer Simon and Bernhard Haslhofer

Annotations have been a means of scholarly communication and an invaluable research tool for centuries. The Open Source YUMA Universal Media Annotator takes this paradigm and makes it available for the collaborative annotation of online media resources. The YUMA suite of tools can currently be used for image, map, audio and video annotation. It is designed to be integrated into any host environment (like a digital library portal or an online media collection) and enables mash-ups across such environments by exposing annotation data according to the Linked Data principles. Linked Data is also the basis for one of YUMA's unique features: Semantic Tagging. A semi-automatic mechanism provides users with tag suggestions that can be used to effortlessly augment annotations with structured context information, eg about places, persons of interest or historical periods.

Major memory institutions (libraries, archives, museums, or audio and video collections) have, over the last 2-3 decades, made substantial efforts to bring their collections closer to their users. These efforts have had two strands: (i) the digitization of original collections (like books or administrative records) and (ii) the start of new, "born-digital" collections. While these collections are now often accessible to the public via the World Wide Web, tools that support actual research – scholarly analysis, annotation, and communication – have been scarce. The YUMA Universal Media Annotator (YUMA) aims to provide some of these tools.

Figure 1: YUMA map annotation screenshot.

YUMA is an Open Source suite of browser-based applications that allow users to annotate different types of media content. It is being developed by the Digital Memory Engineering (DME) research group of the AIT Austrian Institute of Technology in cooperation with the Research Group Multimedia Information Systems at the University of Vienna. The system has seen a number of iterations since 2004 when a first proof of concept was developed as part of the BRICKS (Building Resources for Integrated Cultural Knowledge Services) EU project. It has seen further work in TELplus, a project of 'The European Library', the common portal of Europe's national libraries. The current system represents a complete overhaul both in terms of the technology and the user interface. It is currently being developed further as part of the EuropeanaConnect best practice network, which will run until October 2011. EuropeanaConnect is one of the projects set up to build Europeana, the portal that aims to give access to Europe's museums, libraries, archives and audio-visual collections.

YUMA is based on a distributed architecture. It is designed to be integrated into a host environment – eg an online library portal – and lacks typical portal features like user management. Instead it foresees appropriate APIs and authentication mechanisms which allow the host environment to use YUMA as an external, loosely-coupled service. The system consists of two core elements: the Annotation Suite, an extendable set of browser-based end-user tools for annotating content of specific media types (currently digital image, audio and video files, as well as digitized maps); and the Annotation Server, a common "backend" service used by all of those tools.

The Annotation Suite offers similar functionality across all supported media types: users can create new annotations, view or reply to existing ones and keep track of discussions around individual items or particular annotations via RSS feeds. Each tool provides appropriate selection features for annotating specific parts of an item: shape drawing for images or maps, or time range selection for audio and video material. The map annotation tool (Figure 1) includes a special interface with panning and zooming functionality (think Google Maps), and a set of geographical features such as map geo-referencing and overlay.

YUMA introduces a novel semi-automatic Semantic Tagging approach that lets users make their annotation semantically more expressive by adding links to relevant 'Linked Data' resources. To support users in this task, the tool automatically generates suggestions based on an analysis of the annotation text (and the selected geographic area of a map) and pre-configured Linked Data sets (eg DBpedia and Geonames). Suggestions are presented in the form of a tag cloud (Figure1) from which the user can add proposed links to their annotation. Relevant properties of added resources are stored as part of the annotation metadata: eg alternative language

labels, spelling variants or geo-coordinates. The thus enriched metadata can later be exploited in the portal to enable advanced search functionality, eg search in multiple languages, search by synonymous names, or geographical search.

The Annotation Server is the storage and administration backend of the YUMA Annotation Framework. It can be deployed with different relational database systems (such as MySQL or PostgreSQL). The different applications in the Suite access, store, update, and delete annotations through a REST API. The Server also offers search (through a GUI as well as through an API) and basic administration features, and provides the infrastructure for the RSS feed syndication. The server, in turn, can also act as a Linked Data resource itself: it exposes annotations with unique URIs, which return an RDF representation when resolved. To provide data interoperability, the tool relies on the OAC model, an emerging ontology for describing scholarly annotations of Web-accessible information.

Besides further development of YUMA's feature set and user interface, future work primarily addresses system evaluation. We currently investigate the effect that structured metadata generated collaboratively by users through Semantic Tagging has on search & retrieval. For that purpose, we have created the COMPASS Map Labeller, an online portal used for the study of a map annotation use case. A first outcome of this effort will be a 'ground truth' for the evaluation of map search engines. In a second phase, this ground truth will be used to carry out precision and recall analyses on an annotated map collection. The analysis will help to quantify the improvement that is gained in terms of the quality of obtained search results, when search is performed on metadata enriched by Semantic Tags vs. on traditional metadata only.

Links:
YUMA: Demonstration: http://dme.ait.ac.at/annotation, source YUMA code: http://github.com/yuma-annotation
BRICKS: http://www.brickscommunity.org/
TELplus: http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/
The European Library: http://www.theeuropeanlibrary.org/
EuropeanaConnect: http://europeanaconnect.eu
Europeana: http://europeana.eu/
Linked Data: http://linkeddata.org/
OAC: http://www.openannotation.org/spec
COMPASS Map Labeler: http://compass.cs.univie.ac.at/

Please contact:
Joachim Jung, Rainer Simon
AIT Austrian Institute of Technology / AARIT
E-mail: joachim.jung@ait.ac.at, rainer.simon@ait.ac.at

Bernhard Haslhofer
Cornell University, USA
E-mail: bernhard.haslhofer@cornell.edu

Tidy City: Learning about Cultural Heritage through a User-Created Mobile Game

by Richard Wetzel

Tidy City is a location-based scavenger hunt game for Android phones which allows everybody to create and publish their own content. During the game, players have to solve riddles that deal with historic or current topics and elements of the city. This enables players to learn about the past and present while exploring the city.

The game starts with the riddles scattered around in the environment which the players can see on a map view on the mobile phone. They now need to walk over to the riddle in order to collect it. This enables them to read a short description and see a photo. From this information the players have to figure out two things:

1) What is the object in question?
2) Where is this object located in reality?

The subjects of the riddles might include statues and monuments, architecturally interesting buildings, historical sites or anything else conceived of by the mission creator. Sometimes the solution to a riddle is very simple; often however, the players need to think hard and walk around the city with their eyes open to come up with the solution.

When the players think they are standing at the correct location (evaluated by GPS), they can try to solve the riddle, and if they do so they earn points. When players solve a riddle a new description and image for the riddle become available which give the players further information about the riddle that they just solved. When the players have collected and solved all riddles of a given mission, they have won the scenario.

Creating a new mission for Tidy City is an easy process, consisting of two steps. At first, the mission author uses the Tidy City Scout app for Android phones and walks around the city in order to get inspiration, take photos and record GPS positions. This data can then be uploaded to the server where it can be accessed from a web interface. Here, the riddles can be structured and filled with more detail. If the user is happy with the result, the mission can be published and becomes available for all players of Tidy City.

Together with the newspaper publisher in Waiblingen, Germany, (ZVW), Tidy City was staged for local families in March 2011. The mission for this event was created with the help of four local kids, aged 10 to 12, who, under guidance, used the authoring tools to create engaging riddles about their home town. In this case the riddles included,

for instance, questions about the typical chimera that are found on some of the houses in the old town, an art piece on the main square, the medieval clock tower of the city hall and a historic fountain.

Creating new missions for Tidy City is an interesting way to promote cultural heritage. As a mobile game that is played outdoors, it gets the players directly to the relevant areas of the city and motivates the exploration of the surroundings. This lets the players deal with cultural heritage in an interactive and playful way. This, in addition to the fact that the creation of new missions is very simple, allows Tidy City to be used in a large range of different contexts and makes it appealing to diverse user groups, eg tourists, students, teachers and families.

Tidy City and the Tidy City authoring tools are being developed as part of the German-French research project "TOTEM – Theories and Tools for Distributed Authoring of Mobile Mixed Reality Games". The three-year project started in September 2009 and the research is undertaken by Fraunhofer Institute for Applied Information Technology in Sankt Augustin, Germany, and Telecom & Management SudParis in Evry, France. The game design for Tidy City was conceived by Michael Straeubig of i3Games. The goal of the project is to provide a flexible and powerful framework for the creation of mobile mixed reality games, ranging from experienced users to novices. The TOTEM software consists of several mobile and web-based applications that provide the necessary infrastructure for the creation of such games. Furthermore, a variety of such games is being created during the course of the project to be evaluated and analyzed. The work in the project is funded by the Programme Inter Carnot Fraunhofer from BMBF and ANR (grant 01SF0804).

Tidy city game on a smart phone.

Links:
http://www.totem-games.org/
http://totem.fit.fraunhofer.de/tidycity

Please contact:
Richard Wetzel
Fraunhofer FIT, Germany
Tel: +49 2241 14 2148
E-mail:
richard.wetzel@fit.fraunhofer.de

Design, Implementation and Evaluation of a User Generated Content Service for Europeana

by Nicola Aloia, Cesare Concordia, Anne van Gerwen, Carlo Meghini and N. Zeni

In its 2011-2015 Strategic Plan, Europeana announced User Engagement, ie new ways for end users to participate in cultural heritage, to be one of the primary tracks by which the organization will deliver value. Europeana intends to enhance the user experience and offer services that allow users to interact and participate.

Tools for the creation, management and access to user-generated content are currently very popular, as witnessed by the success of sites such as YouTube and Facebook, amongst others. The Europeana network comprises communities of archivists, curators and librarians who show a growing interest in exploring new methods of access and dialogue. User-Generated-Content (UGC) is one aspect of this renewed way of participating. Information about cultural heritage exists outside the heritage institutions; artifacts, written sources and memories of individuals complement collections held within the institutions. UGC services are designed to provide users with a means to support and interpret content. They will be involved in storytelling, curating of virtual exhibitions, reviews and even the creation of new collections. Greater participation will increase users' interest and loyalty. Europeana is therefore devoting increasing resources to initiatives that bring out the value of the contribution those users can make.

In response to these needs, the ASSETS Consortium has included the support of user-generated content amongst the services it will develop for Europeana. ASSETS is a two-year Best Practice Network that aims at improving the accessibility and usability of Europeana. The ASSETS Consortium comprises 24 partners, including institutions from ten different European countries and Japan, active in the field of cultural heritage and digital libraries.

Rather than focusing on a specific set of UGC applications, ASSETS is developing a general purpose, back end component that aims at supporting any UGC service that Europeana offers its users. To this end, the ASSETS back end component implements an Application Programming Interface (API) for creating, storing and manipulating UGC Units of Work, and for submitting these Units of Work to Europeana, in the form of Europeana Submission Information Packages (SIPs). End users will interact with their Units of Work through client interfaces, which hide from them the unnecessary technical details and complexities of the back end, providing the level of representation that is most suitable for the specific UGC task at hand. It is expected that every UGC task will be

supported by a different end user interface. This will have no impact on Europeana, since each front end will talk to Europeana through the same API.

The API will relieve future UGC applications from implementing any server side functionality; they will have to code only the client side, connecting the user world to the API. At the same time, the API will take away from Europeana the technical interoperability problems that would arise from integrating into its database objects coming from future UGC applications. The service will rely on the Europeana Data Model (being developed by the Europeana version 1.0 project) in order to tackle the more serious semantic interoperability problems.

The definition of the conceptual model underlying the UGC API is the most difficult challenge that the ASSETS UGC team is facing. The model has to strike an optimal balance between simplicity, ie easy-to-learn and easy-to-code by future UGC service developers, and generality, ie to satisfy the needs of any future UGC service. This conceptual model was defined during the first year of the ASSETS project on the basis of an analysis of different types of requirement. The model was subsequently used to define the UGC API, the first version of which should be ready by June 2011. A few initial UGC tasks will be implemented on top of the API through specific front ends. These tasks will allow Europeana users to contribute to the contents of the digital library in several ways, eg uploading new objects along with simple descriptions, annotating existing objects, or enriching existing descriptions.

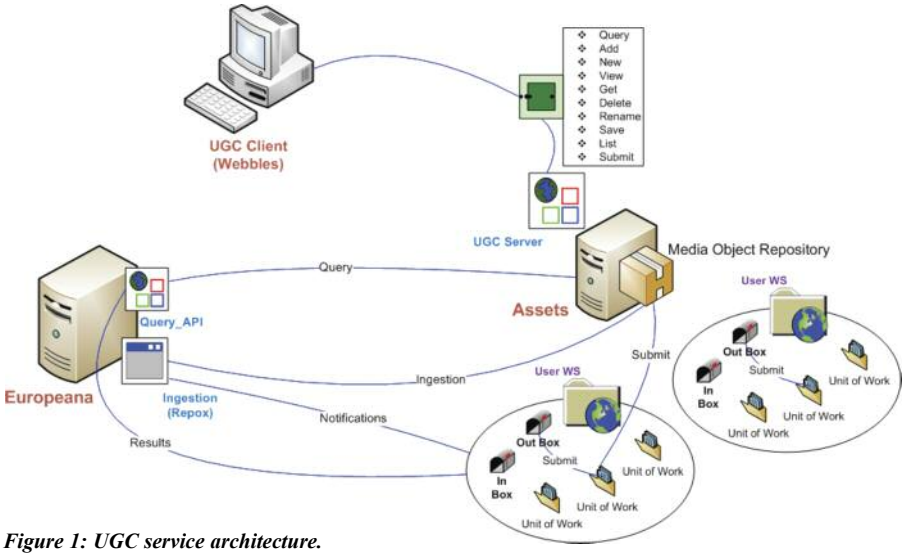


Figure 1: UGC service architecture.

The UGC service will be evaluated by the task developers and by the end users within the project's lifetime. The former will evaluate the adequacy of the API, while the latter will evaluate the adequacy of the front ends. The outcome of the evaluation will lead to a refinement of both the back end component API and implementation, and the front end UGC applications.

Each user WS is endowed with an Inbox. Messages in the Inbox are of two kinds: query results and notifications that communicate the result of submissions. Positive notifications will report the successful ingestion of a SIP, negative notifications report the reasons why a given SIP could not be ingested. Rejected SIPs can be retrieved and re-transformed into a UoW so that users can perform necessary reparations.

It is important to note that these concepts define a general-purpose schema, whose machinery need not be used by every UGC application. For instance, a simple UGC task that takes place in a single session, such as an image upload, may be implemented by directly building the corresponding SIP, by-passing the UoW stage. On the other hand, another UGC application may decide to publish a finished UoW to a community of users in order to perform a socially oriented form of mediation before submitting the UoW to Europeana. These decisions will be taken by the client side of the applications, relying on appropriate shortcuts offered by the UGC API.

Links:
http://www.europeana.eu
http://www.assets4europeana.eu
http://version1.europeana.eu/web/europeana-project/
http://europeanalabs.eu/wiki/

Please contact:
Carlo Meghini
ISTI-CNR, Italy
E-mail: carlo.meghini@isti.cnr.it

Visual Semantic Browsing - A New Way to Access Digital Collections

by Alan Payne and Peter Fry

We are in the process of developing a system for taking museum and gallery visitors on a journey of discovery through a digital collection. It is entirely image-led and allows the visitor to wander through a collection following semantic relationships between objects in the collection.

In recent years, many collection owners from museums, art galleries and libraries have put a lot of effort into digitizing their collections. This involves either document scanning or careful photography of the artefacts then combining the resulting digital images in a database with textual descriptions of the artefacts. Following this exhaustive process, the standard method of accessing the newly-formed digital collection is by keyword searching, often along with some kind of structured categorization and browsing of image thumbnails. This suits the expert and

other visitors who know exactly what they are looking for, but for many visitors who just want a general exploration of a collection, current systems don't work well.

We have developed a system for visual semantic browsing - ViziQuest. This allows the visitor to explore a collection in a very different way. It is entirely image-led and takes the visitor on a journey of discovery through a collection based on the visitor's individual selection of images that they find interesting. The visitor is presented with a montage of images; Figure 1, for instance, shows a central image surrounded by other images from the collection which are related by varying degrees to the central image. By selecting one of the surrounding images that interests them, this image moves to the centre of the display and is surrounded by new images that are related to this selected one (Figure 2). The textual descriptions are readily accessed by "flipping" the image (Figure 3).

The collection owner simply supplies their database of images and descriptive texts. There is no further "pre-work" to be done on the data by the collection owner. We carefully examine the collection and build an ontology to describe the scope and depth of the collection. Ideally, this ontology is reviewed and refined with the collection owner or other domain expert. The ontology is then used to direct the building of relationships between the artefacts using Natural Language Processing (NLP) techniques.

A system based on this work has been installed at the Scott Polar Museum in Cambridge, UK. It is proving very popular with visitors, both "experts" and casual visitors, and with all age groups. The image-centric nature of the system (it requires no textual input whatever) acts as a clear draw. The museum curators are delighted with it, particularly in the fact that it is enabling visitors to unearth images from all parts of the collection, not just a narrow slice as would be found with a traditional search.

The basic system has now been considerably enhanced by the addition of audio and movie files that were part of an archive from a further polar collection. This new system has recently been installed in the museum, where again visitor feedback is very positive.

Figure 1: A central image surrounded by other images from the collection which are related by varying degrees to the central image.

Figure 2: A selected image moves to the centre of the display and is surrounded by new images that are related to this selected one.

Figure 3: The textual descriptions are readily accessed by "flipping" the image.

A current development project is to build some targeted educational elements into the browsing experience. We are testing a modified version of ViziQuest with a group of children (aged 9-11 years old) from a local primary school. A crucial element of this stage in their learning is the ability to create and format narratives, both fictional and factual. The system provides nuance to the narrative by calling up related images and exploits the textual descriptions in the museum digital collections. This modified version requires the children to select images for hard copy that act as props for them to build a story around. In this way they are producing their personalized stories and enriching them in ways not possible with ordinary search methods. Serendipitous browsing, favoured by the semantic browser, promotes lateral thinking which in turn has a strong role in creativity. By providing the tools described, the students will better understand the formal aspects of narrative and appreciate what constitutes good practice in storytelling and report writing. As well as assisting the pupils with their story-telling skills, exploring the collection in this way inevitably teaches them about the collection as a whole which in many cases has stimulated them to pay a live visit to the museum where the collection resides.

Visual semantic browsing provides a new way to access digital collections, opening up collections to a wider audience, unearthing hidden depths of collections and extending the use of the hard work involved in digitization projects.

Please contact:
Alan W Payne
Deep Visuals Limited, Cambridge, UK
Tel: +44 (0)1223 437163
E-mail: alan.payne@deepvisuals.com
www.deepvisuals.com

Data Quality of Public Heritage Data in Professional Applications

by Jacco van Ossenbruggen

Professional users in the heritage domain have started to deploy data published on the Web: data that often comes in different forms, from different sources and is controlled by different organizations. In addition to the data integration problems in the back end, CWI is also studying the user interface challenges in the front end.

In the context of the Dutch MultimediaN e-culture project, we have been looking at ICT-related tasks of professional users from the cultural heritage domain. Many of these tasks are now performed by using a single information source at a time, typically a web site of a trusted organization or, more often, on "in house" data: data collected and curated either by the users themselves or their direct colleagues working in the same institute. For our research, we selected tasks that could obviously benefit from including multiple data sources, especially if those sources were publically available from the Web. We then looked at the implications of this shift from single source, trusted data to multiple source, untrusted data had on the user interface.

One thing we learned is that heritage data is never 'sound and complete'. Even well-curated data sets of any substantial size contain many omissions, imprecise values and plain errors. However, in practice, in the single source setting, this hardly seems an issue: users typically know what is in 'their' database and what is not, and which parts of the data are in line with the institute's guidelines and which parts are not. Users take this into account, for example, when performing search tasks in a museum's collection management system. Our team at CWI conducted several user experiments, asking heritage professionals to think aloud during their search tasks. Remarks along the lines of: "You would think I could just search on this keyword, but I know this part of the collection has not been properly annotated yet, so instead I will ..." clearly indicate that users explicitly work around the problems in their data.

However, this situation completely changed when we confronted users with similar quality issues in data sets they were either unfamiliar or less familiar with. For example, the same type of errors users seamlessly worked around in a single source setting, often led to confusion in experiments with multiple data sources. Even worse, when our prototypes revealed conflicting claims about the same art object or artist without explaining the source or provenance of these claims, it led to users totally distrusting the entire system.

Figure: Art annotation application developed by CWI in cooperation with the Rijksmuseum. While the user is typing, the interface suggests terms from different public web data sources, in this case RKD's IconClass and Princeton's WordNet.

Another striking example was a prototype annotation interface we developed to enter metadata of artworks as part of a museum's collection registration process. We studied this process extensively in the print room of the Rijksmuseum in Amsterdam, where literally hundreds of thousands of historical prints are currently being digitized and described. In the single source setting, one specific interface worked fine because it was optimized for finding, as quickly as possible, an artist's name that the user knew to be in the database: it ranked search results by frequency of use. The same interface, however, proved useless in the multiple source setting, because now the frequency-based ranking made it impossible to judge whether a specific artist was not present at all or was just ranked too low to appear in the top-N search results presented. In the multiple source setting, a simple alphabetically ordered list of search results proved to work much better than the frequency-based ranking.

These and other insights of the MultimediaN E-culture project are now being used in the Europeana project. The issues around dealing with varying data quality in the user interface will only become more relevant now more and more user generated content finds its way into the cultu ral heritage field. The successor of MultimediaN, the Dutch national research project COMMIT, will start this summer. This project will focus on provenance of, and trust in, curated and non-curated heritage data as one of its key research topics. These topics also apply to domains other than cultural heritage.

In the European Fish4knowledge project, CWI is studying the interface implications of data quality in large scale data observations of coral reef fish. Since the database is filled with observations made by automatic feature detectors analyzing large quantities of underwater video camera footage, the quality of the data varies with the quality of the detection algorithms used, the turbidity of the water, the movement of the fish, etc. How to convey the strengths and limitations of this data without overwhelming biologists with the technical details of automatic video analysis algorithms is one of the challenges in this project.

Links:
http://e-culture.multimedian.nl/
http://www.europeana.eu/
http://www.commit-nl.nl/
http://www.fish4knowledge.org/

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

Discovering Europeana Users' Search Behaviour

by Diego Ceccarelli, Sergiu Gordea, Claudio Lucchese, Franco Maria Nardini, Raffaele Perego and Gabriele Tolomei

Europeana is a strategic project funded by the European Commission with the goal of making Europe's cultural and scientific heritage accessible to the public. ASSETS is a two-year Best Practice Network co-funded by the CIP PSP Programme to improve performance, accessibility and usability of the Europeana search engine. Here we present a characterization of the Europeana logs by showing statistics on common behavioural patterns of the Europeana users.

The strong inclination for culture and beauty in Europe has created a rich abundance of artifacts, starting with antiquity and continuing to today. This cultural strength is recognized globally, making Europe the preferred destination for 50% of world tourism.

Europeana [1] is a long-term project funded by the European Commission with the goal of making Europe's cultural and scientific heritage accessible to the public. Since 2008, about 1 500 institutions have contributed to Europeana, enabling people to explore Europe's museums, libraries and archives. This huge amount of multilingual and multimedia data is made available through the Europeana Portal, a search engine that allows users to explore the content by means of textual queries.

Due to the increasing amount of information published, access to the description of a specific masterpiece is becoming increasingly difficult, in particular when the user is not able to formulate a sufficiently discriminating query. For example, if we search today for general terms like "renaissance" or "art nouveau" we obtain more than 10 000 results. If we search for the term "Gioconda" we find a couple of hundred of items, while the query "Mona Lisa, Da Vinci" provides us twenty images of the well known painting. These examples show how important it is to have a good formulation of users' information needs in textual queries when looking for very specific information on the web by using a portal like Europeana. This is even more challenging since Europeana documents are cross-domain, multi-lingual, and multi-cultural.

Figure 1: Most popular queries submitted by Italian users. Size is proportional to frequency.

In order to improve the Europeana users' search experience, the ASSETS project [2] has the overall goal of enhancing the performance of the Europeana search engine. One of the most important resources for enhancing the users' search experience in large information spaces is the use of the information stored in query logs. The knowledge extracted from query logs can be used to enhance both efficiency (ie, response time, throughput) and efficacy (ie, quality of results) of information retrieval platforms.

The Europeana logs show the behaviour of users interacting with the portal. We present here a preliminary characterization of this behaviour, and a comparison with that of general Web users. Figure 2 shows the frequency distribution of submitted queries. As expected, the popularity of the queries follows a power-law distribution ($p(x) \propto k \times x^{-\alpha}$), where x is the popularity rank. The best fitting α parameter is $\alpha = 0.86$, which indicates the skew of the frequency distribution. The larger the α, the larger the portion of the log covered by the top most frequent queries. The same analysis conducted on query logs coming from commercial Web search engines shows larger values of α (2.4 and 1.84 respectively for an Excite and a Yahoo! query log).

Such a small value of α means that the most popular queries submitted to Europeana do not account for a signifi-

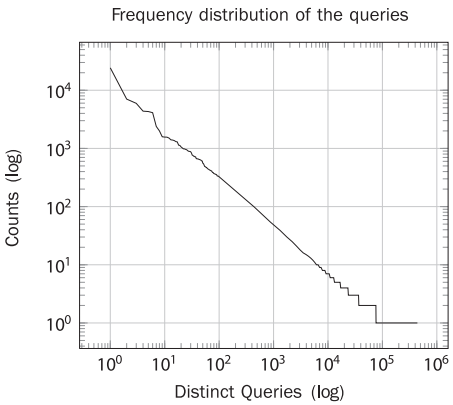


Figure 2: Frequency distribution of queries.

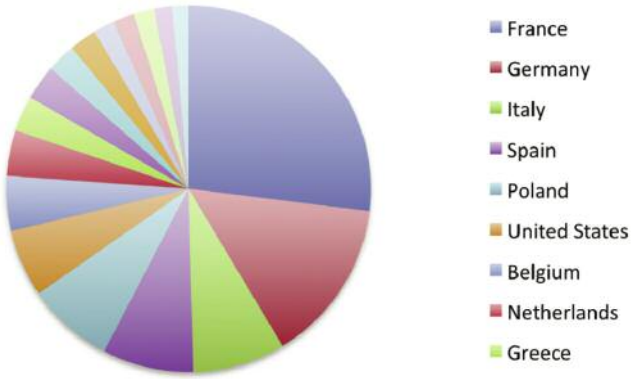


Figure 3: Distribution of the queries over the countries.

cantly large portion of the query log. This might be explained by considering the characterizing features of Europeana. Indeed, since Europeana is strongly focused on the specific context of cultural heritage, its users are likely to have a greater vocabulary, and therefore use a more diverse vocabulary. In addition, we found that the average length of queries is 1.86 terms, which is again lower than the typical value observed in Web search engine logs. We can argue that the Europeana users use a richer vocabulary, with discriminative queries made of specific domain terms.

Figure 3 shows the distribution of the queries grouped by country. France, Germany, and Italy are the three major countries accounting for about the 50% of the total traffic of the Europeana portal.

Furthermore, Figure 4 reports the number of queries submitted per day. We observe a periodic behaviour on a weekly basis, with a number of peaks probably related to some Europeana dissemination or advertisement activities. For example, we observe several peaks between 18 and 22 November,

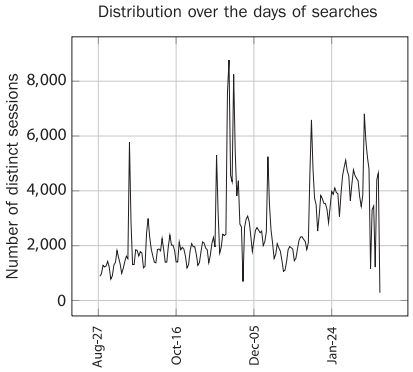


Figure 4: Distribution of the searches over the days.

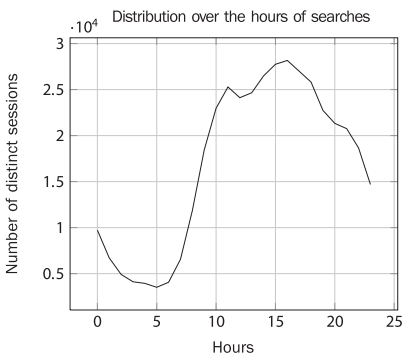


Figure 5: Distribution of the searches over the hours.

probably due to the fact that, during this period, Europeana announced the indexing of new collections and contents of 14 million documents.

Figure 5 shows the load on the Europeana portal on an hourly basis. We observe a particular trend. The peak of load on the Europeana portal is in the afternoon, between 15.00 and 17.00. This is slightly different from commercial Web search engines where the peak is reached in the evening, between 19.00 and 23.00. A possible explanation for this phenomenon could be that the Europeana portal is intensively used by people working in the cultural heritage field and thus mainly accessed during working hours, whereas a commercial Web search engine is used by a wider range of users with the most disparate information needs throughout the entire day.

Links:
[1] http://www.europeana.eu/portal/
[2] http://www.assets4europeana.eu/

Please contact:
Raffaele Perego, ISTI–CNR, Italy
Tel: +39 050 3152993

Preparing the Ground for the German Digital Library

by Kai Stalmann, Marion Borowski and Sven Becker

Providing users with information and knowledge about cultural heritage objects has been a core business of libraries, museums, archives, and other institutions for centuries. Accessing the information, however, has always been limited to those who literally step through the portals of these elevated places of acquired knowledge. Digitizing and publishing of digitized objects to the web brings knowledge to a much broader audience. It is even conceivable that cultural heritage might once again play a significant role in society, providing it is easy enough to explore, openly accessible and applicable to range of purposes.

With the use of the right key terms, a search engine can offer the user a veritable treasure trove of knowledge. The "treasure", however, can vary depending on what the individual is searching for, ranging from an object located somewhere on the globe to an image of or deeper knowledge about a concept or object.

Which technologies might be useful in enabling the public to access their cultural heritage in a variety of ways? The approach taken for implementing the Deutsche Digitale Bibliothek (DDB) at Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS in conjunction with the Federal Government Representative for Culture and Media (BKM) relies on a modular and scalable, light weight architecture, a clear cut API based on open standards, a flexible model, and the use of given datasets and smart algorithms.

The DDB may hold binaries like images, scanned book pages, digitized 3D objects and video material, thus allowing users to work directly with these assets. This offers a range of fantastic uses, including opportunities for those who are artistically inclined who wish to work with the bare material and use it for their own creative productions.

Finding items turns out to be amazingly easy on the "Cortex platform" which underlies the DDB. The state of the art search engine supports keyword search and facetted search which are usually found on commercial search interfaces. Facets provide a means to drill down iteratively, and thus allow for easy management of the largest result sets. What makes a difference here is the semantic

The DDB as a linked data project: see the woods and see the trees...

richness of the facets which depends on both the quality of the model and of the mapping of the input data to that model. The DDB uses a harmonized model based mainly on CIDOC CRM and the new Europeana Data Model (EDM). These models were designed explicitly with a vision of a network of knowledge in mind.

Creating a network of knowledge is at the heart of the DDB project. Over the years, millions of things, events, actors, places and time spans will become entangled by different kinds of relationships between these entities. Thus the DDB contributes to the semantic web. It also depends on the semantic web: vocabularies and ontologies are needed for recognizing relationshipss and identifiers. A two-way relationship exists between users and the DDB; users profit from a DDB but are also the most powerful resource for enabling the network of knowledge to come into existence. The unique selling point of the DDB is not only the original content, but also what given persistent identifiers, 'artificial intelligence' and 'the crowd' may reveal as hidden links between the objects of cultural heritage.

Analyzing single objects (like museum pieces or dossiers in an archive) may still require travelling to the place where these objects reside whenever high quality digital representations can not be made accessible online due to technical or legal restrictions. But exploring relationships between entities that are part of the knowledge model encourages a different kind of exploration that is not restricted by physical borders. The DDB will therefore support different levels of queries. Besides keyword search combined with semantically rich facets, more advanced queries based on SPARQL will allow for data mining the cultural heritage. Data mining in general unveils correlations that are extremely difficult to detect without those technologies.

The DDB consists of a middleware (Cortex) for managing access, search, and the ingestion of objects. The ingest service binds objects to identifiers, resolves resources already in place, and links objects to internal and external resources. The ingest process is the tricky part since cultural heritage metadata files differ tremendously in semantic richness, format (DC, EAD, Lido, Marc, Mods, to name a few), and size (from less than 1KB to over 100MB for one item). Despite approaches towards standardizing, the metadata formats of one kind may well still be used in many different ways. In our approach, the incoming data is prepared for ingest by a separate tool called ASC or Augmented SIP Creator (SIP standing for Submission Information Package, the term is borrowed from the OAIS Reference Model). The Cortex middleware, a set of metadata mappings, and the ASC are currently being built at Fraunhofer IAIS (a first release was accepted in April 2011). The middleware makes use of various widely accepted technologies like Spring, REST, JSON, XML). As backends Cortex uses a triple store, Solr as a search index, and a storage cloud. The architecture was explicitly designed for reuse in similar projects. Since the underlying model and the mappings can be adapted to other needs, this system could be applied to any in which digital objects form a knowledge network based on either given or automatically extracted metadata.

The DDB is a national project that will be steered by a "Kompetenznetzwerk" (competence network) DDB consisting of members of many institutions that contribute content to the DDB. But the DDB may also be embraced by an ecosystem that still has to be evocated. A first exciting glimpse of an ecosystem around the DDB was presented in Berlin in April 2011. Therefore a set of distributed services where mashed up to extend the core DDB functionality. We anticipate that the DDB could stimulate activities around cultural heritage objects. Beyond 'looking for something' this might include 'exploring the betweens', and ultimately also providing functionality that does something new to both objects and betweens.

Links:

http://www.bundesregierung.de/Content/DE/Pressemitteilungen/BPA/2009/12/2009-12-02-bkm-deutsche-digitale-bibliothek.html

DDB project website:
http://www.deutsche-digitale-bibliothek.de/

Fraunhofer IAIS Netmedia:
http://www.iais.fraunhofer.de/index.php?id=191&L=1

Documents on the Cortex platform (currently only in German):
http://www.iais.fraunhofer.de/cortex.html?&L=1

Please contact:
Kai Stalmann, Marion Borowski, Sven Becker
Fraunhofer IAIS, Germany
E-mail: {kai.stalmann, marion.borowski, sven.becker}@iais.fraunhofer.de

Managing and Archiving Digital Audio: The GAMELAN Project

by Jérôme Barthélemy

The GAMELAN project is devoted to the development of an environment for management and archival of digital audio and music (in French, Gamelan stands for "Gestion et Archivage de la Musique Et de L'Audio Numérique"). It will provide an intelligent archival system for music, based on the history and the semantics of the production.

The production of a music piece – either art music or pop music - using digital technology is a very complex process involving many people, an array of technologies and many different processes. It generally results in a very complex set of digital objects, and the organization of this set of files for the purpose of archiving and eventually republication or repurposing is not a straightforward task.

At this stage, it should be evident that it is not sufficient to archive the result of the production, the "master". In reference to Acousmatic music, composer Daniel Teruggi states that "It may seem that the essential issue for an Acousmatic work is to be preserved as such; in the same way a traditional instrumental score seems to convey the necessary information for music performance (which it does not, it conveys partial information, the missing information is reconstructed through use tradition). However, the situation is much more complex due to the characteristics of the production environment, which interfere with the nature of the result and the number of possible elements to be preserved."

The new uses facilitated by digital technologies make it possible to envisage new productions - what is known as "repurposing", even in the domain of popular music, such as the creation of a version for karaoke. For art music, other forms of repurposing, such as musicological studies, have to be anticipated. This is true too, for popular music, in which (as pointed out by Marc Danger from EMI), "a lot of fans would hear the ninety-one mixes for 'Billie Jean' and understand why Jackson and audio engineer Bruce Swedien chose to go with the second mix". The general absence of semantic qualifiers on the objects produced makes these studies and new productions almost impossible to realize.

This is where GAMELAN intends to develop a new approach, by developing an environment that will track the activities of users during the production process - including the activity done in end-users' tools like sequencers - audio imports and exports, mixing and applying effects.

Gamelan takes care of the future of digitally encoded music.

By recording the main events, the environment will be able to track the activities. But this is evidently insufficient for applying semantic qualifiers to digital objects and for detecting the relationships between the different objects of the audio production (original recordings, modifications, results of mixes, etc.). For this purpose, we need to develop a language for the description and representation of musical production processes, as well as for the musical object itself. The definition of this language is at the very core of the GAMELAN project, and is the subject of a PhD thesis (advisors Bruno Bachimont, UTC, and Alain Bonardi, IRCAM).

The main difficulty in developing this language is to define the right level of abstraction. As explained by Bruno Bachimont: "We are looking for a level of representation that enables to hold [sic] the content invariants without being burdened by the incidental aspects of its technical implementation. This issue is very close to the so-called 'Knowledge Level' in knowledge engineering, where this level would get the treatment imposed to representations, being more general than the code implementing this treatment, but precise and prescriptive enough to reproduce that treatment. The purpose is finally to mark out the variety of processes involved and to find out the invariants, enabling on the one hand to drive a musical production, and on the other hand to manage musical contents as results of a production."

At the current stage, the project has developed a first prototype of the environment that is able to track some of the activities coming from end-user applications, record these events, and apply some very simple semantic rules to the set of events in order to define new relationships. The prototype is, at this stage, able to determine the audio sources of a mix and to present that relationship in a graphical interface to the end user. In parallel, studies have been done in order to define the bases for the future language for musical production. The next step will be to enter into an iterative process to refine the language and the environment.

The expected outcomes of the project are the development of a language for musical production and of an environment that is able to track events and build an intelligent, semantically qualified archive.

GAMELAN started in November 2009 and is scheduled to end in February 2013. The project is partially supported by the French ANR (National Agency for Research). It involves four partners: IRCAM (Institut de Recherche et de Coordination Acoustique/Musique), INA (Institut National de l'Audiovisuel), UTC (Université de Technologie de Compiègne), and EMI Music France.

Link:
http://www.gamelan-projet.fr

Please contact:
Jérôme Barthélemy, IRCAM, France
Tel: +33 1 44 78 15 53
E-mail: jerome.barthelemy@ircam.fr

The Open Kunstkammer Data Project

by Vladimir Ivanov

This project has led to the publication of metadata(as Linked Data) relating to more than 40 000 photos from Russia's oldest museum.

In 2009 general activity on digitization carried out by Peter the Great Museum of Anthropology and Ethnography Russian Academy of Sciences (the Kunstkammer or MAE RAS) had led to the creation of dataset consisting of more than 40 000 digital images and their descriptions in English and Russian. The dataset supplied with a search tool is accessible on the MAE RAS's website in HTML format. In 2010 the Computational Linguistics Laboratory of the Kazan Federal University initiated the Open Kunstkammer Data project (OKD) aimed at standardization of metadata representation and publication the dataset as Linked Data.

Figure 1: The website of MAE RAS.

The OKD is a research and development project inspired by Tatyana Bogomazova, head of IT department of MAE RAS. Bogomazova articulated the project's aim, which is to enrich the existing dataset using the International Committee for Documentation of the International Council of Museums (ICOM/CIDOC) recommendations and the best practices of the Semantic Web. To this end, two steps were taken. First, a legacy database schema was mapped to the Web Ontology Language (OWL) representation of the CIDOC Conceptual Reference Model (CIDOC CRM). The mapping process raised a few semantic issues which were resolved only after consultations with museum staff and IT-experts. The mapping was defined in a machine readable format which allows automatic generation of target (CIDOC CRM) representation for each source metadata record. The result of this step is an Resource Description Framework (RDF) dataset that contains the CIDOC CRM classes, properties and their interrelated instances. The dataset describes the meaning of the source data in terms of the CIDOC CRM.

An additional task carried out in the OKD project was to "SKOSify" MAE RAS's controlled bilingual vocabularies ("SKOS" stands for Simple Knowledge Organization System). The list of vocabularies includes "Places", "Actors (Authors and Collectors)", "Expeditions", "Ethnicity", "Subjects" and "Genres". SKOS-based versions of these vocabularies are also represented in RDF format. The total amount of RDF-triples (ie statements in the form of "subject-predicate-object") in the RDF-dataset is more than 5 million (for both languages). A special software component supporting regular updating of the RDF-dataset was also developed. Its implementation was quite straightforward because updates made by users never affect the source database schema. There is therefore, no need to update the mapping. Finally, all the RDF-triples were loaded into the Virtuoso Universal Server RDF-store. Public SPARQL Protocol and RDF Query Language endpoint (http://data.kunstkamera.ru/sparql) is supported by Open Source edition of the Virtuoso Universal Server. The server's software also supports dereferencing of URIs, which means that any URI in the RDF-store not only identifies a resource (eg image or vocabulary concept), but also provides access to corresponding RDF-triples which describe the resource.

The second step is the definition of mappings between concepts from MAE

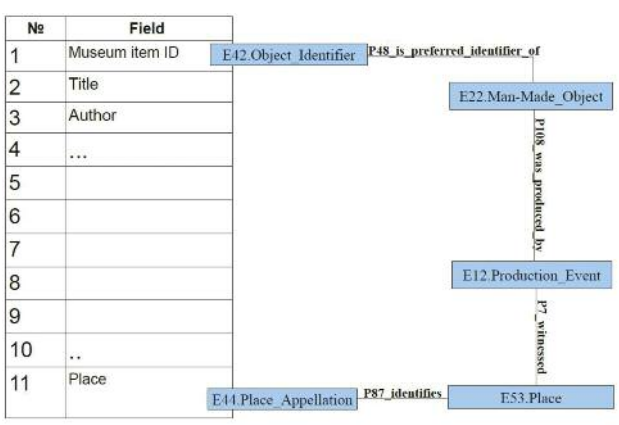


Figure 2: A graphical representation of the mapping fragment.

Figure 3: A map with creation places and museum items.

RAS controlled vocabularies and concepts from external datasets. DBPedia (http://dbpedia.org) and Geonames (http://geonames.org) were chosen as target datasets for their multilingualism and global coverage. The main issue we faced here was ambiguity. For example, concepts from the "Places" vocabulary can have up to hundreds of "equivalents" when matched to the Geonames concepts using labels. A set of special heuristic disambiguation procedures were developed. For both target datasets these procedures significantly decreased the ambiguity. More than 60% of matched concepts (in the "Place" vocabulary) have a single equivalent and only 6% of matched concepts have more than ten equivalents after applying the procedure. However, manual disambiguation is still required.

External datasets enrich the source database with useful information. Geographic coordinates from Geonames could be useful when showing the MAE RAS dataset on the global map (see Figure 3). Multiple languages of DBPedia resources allow multilingual searching.

The MAE RAS RDF-dataset could be useful in external services and semantic web applications. For instance, triples from the dataset are already used by the Sig.ma (http://sig.ma) semantic mashup application (see Figure 4).

Other applications can also consume the MAE RAS RDF-triples online. A screenshot of such a software component based on Information Workbench framework (http://iwb.fluidops.com/) is shown in Figure 5.

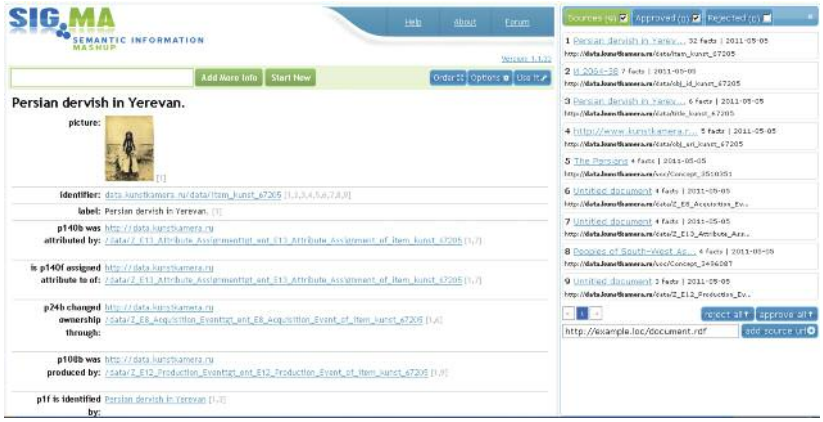


Figure 4: MAE RAS description used by the Sig.ma.

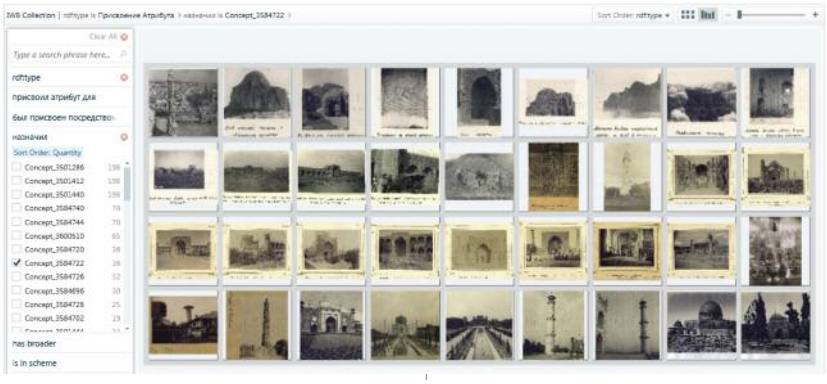


Figure 5: A set of photos to which two attributes ("Islam" and "Cultural landscape") have been assigned.

Future research and development activities are related to improvement of automatic matching and disambiguation quality, integration with the original search tool and web application deployed on the www.kunstkamera.ru, and providing end-users with new services based on the published dataset.

Links:
http:// www.kunstkamera.ru
http://data.kunstkamera.ru/sparql

Please contact:
Vladimir Ivanov
Kazan Federal University,
Computational Linguistics Laboratory,
Kazan, Russia
E-mail: nomemm@gmail.com

Pat.mapa: Linking Catalan Cultural Heritage

by Mercè López, Oriol Almirall and Sergi Fernández

Pat.mapa is an ambitious project that presents the Catalan cultural heritage on the Internet in an integrated and innovative format. The project addresses the semantic web challenge of aggregating cross-domain cultural heritage content into a semantically rich intelligent system. Pat.mapa is an interdisciplinary project that for the first time in the Catalan area, implements the principles of linked data, ontologies and complex data visualization on cultural heritage content.

The mission of Pat.mapa is to present the Catalan cultural heritage content in a different way. Heritage is not an isolated area and only related to the past. Cultural heritage is alive, connected with the real world, the science, the gastronomy, the companies, the people. We can help to bring cultural heritage alive with the implementation of technologies like the Semantic Web. Through the ontological approach, we can see what links exist between the Monastery of Sant Benet de Bages with Ferran Adrià of El Bulli restaurant, or between the Roman Church of Sant Climent de Taüll and the modernist architect Lluís Domènech i Montaner.

Key concepts on which the project is built:

- Representation and vision: Pat.mapa will represent a set of cultural values and assets and offer their differing views and interpretations.
- Evolution and creativity: The project aims to show the evolution of the Catalan cultural heritage in order to promote knowledge generation and innovation as an economic engine and source of cultural wealth.
- Innovation and participation: This project represents innovation in data capture and content presentation. It makes extensive use of digital tools and collaboration in order to develop a Catalan digital culture.
- Quality and proximity. Pat.mapa will become a reference point for the dissemination of cultural heritage and local civil society.

At a technical level, the project represents a major challenge both in the structuring and standardizing of data assets such as technology development (enrichment of ontologies tool and visualization of complex data).

The process of structuring data is based on an intensive mapping between different databases and metadata management (thesauri, controlled vocabularies, etc.). The main innovation and challenge in this work process is the application of the principles of the semantic web. We are working on the implementation of an ontology based on the standard CIDOC Conceptual Reference Model (CRM), that provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation (official standard ISO 21127:2006).

At the same time, given that much of the relevant information about the history of cultural heritage artefacts (monuments, archaeological sites etc) is in free text fields on the databases, a new tool for natural language processing has been developed. Thus, we can treat complex and abundant sources of text, identify characters and events and incorporate them semi-automatically into the existing ontological structure.

Furthermore, we face the challenge of making visible the entire network of relationships that gives us the ontology. For this reason, the i2cat Foundation is developing a visualization tool that allows complex data visualization of any ontology.. Technically, this means working with converter triple-labelled graphs, a graph viewer and a system of widgets customized according to the ontology. The navigation tool for cultural heritage data is a viewer based on standard HTML5 using processing.js javascript libraries, jQuery and a physics engine developed by Jeff Traer. The viewer reads data from knowledge bases classified on ontologies and shows their relationship graphs. Thanks to the ontology visualization tool, we can provide an interface that allows the user to browse the map of ontologies on cultural heritage. With the presentation of content through this tool, we contribute to the dissemination of knowledge about Catalan cultural heritage.

Pat.mapa is a long term project. The starting point is the cultural heritage information but it will continue growing. We are aggregating content from databases on architecture, archaeology and museums, and this will be extended to other areas of knowledge, culture and society (performing arts, and tourism for instance). Pat.mapa is the first step in the Catalan government's strategy on Linked Data for cultural content. In the near future, we will also face the web 2.0 challenge and explore the possibilities of memory organizations and citizen's contributions.

Pat.mapa is a project led by the Directorate General for the Cultural Heritage (Department of Culture, Government of Catalonia), developed jointly with the i2cat Foundation and ISOCO Company. The project began in 2008 and is currently in beta. The public presentation is scheduled for November 2011.

Link:
http://patmapa.gencat.cat (beta version):

Please contact:
Mercè López
i2cat Foundation, Spain
Tel: +34 93.551.06.81
E-mail: merce.lopez@i2cat.net

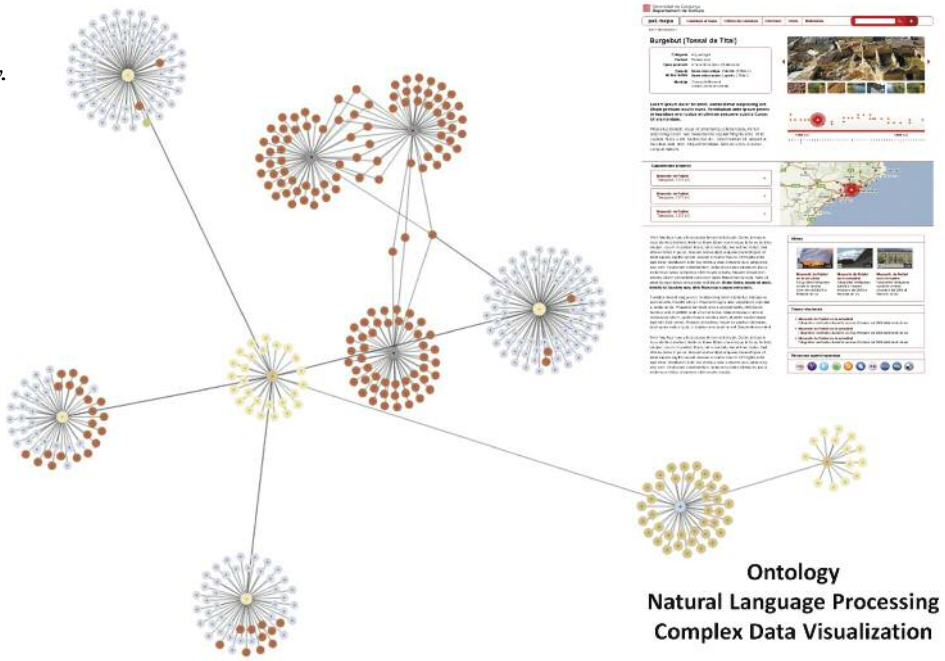


Figure 1: Pat.mapa overview.

GeoMemories: A Spatial-Temporal Atlas of the Italian Landscape

by Andrea Marchetti, Elizabeth Jane Shepherd and Maurizio Tesconi

We present an interactive Web 2.0 application that is being developed with the aim of showing the evolution of the Italian landscape through an important historical archive of aerial photographs.

The AeroFototeca Nazionale (AFN) of the Italian Ministry of Cultural Heritage in Rome maintains an extensive set of some million aerial photographs constituting an important memory archive of the Italian territory throughout the 20th century. This huge archive provides a picture of Italy as it was 70 years ago and during its transformation by the post-war reconstruction and the economic boom, and also by natural disasters such as severe earthquakes and floods.

GeoMemories

In 2010, the AFN and the Institute of Informatics and Telematics of CNR in Pisa signed an agreement to make this archive accessible via Internet through a Web 2.0 mashup application that displays the evolution of the Italian Landscape. Thanks to funding from the Italian Internet Domains Registry, the GeoMemories project was launched with the aim of creating a web platform covering spatial-temporal dimensions and also integrating multimedia data from other archives. The project is coordinated by Andrea Marchetti, Head of the IIT Lab "Web Applications for the Future Internet" and Elizabeth Jane Shepherd, Director of AFN.

Version 5.0 of Google Earth (GE) includes a timeline to display historical imagery. However, this new feature has important limitations, in particular, most of the Google images are from satellites and are relatively recent. Italy, for example, has significant coverage only from 2003. The only samples of actual historic images (1943) concern some large cities (Rome, Florence, Naples, Turin, Trieste and Venice) and the image resolution is very low. In addition, GE manages geographic data layers without considering the time variable.Our aim is to rebuild a virtual globe, similar to GE but oriented towards the management of the time.

From Photos to Historical Maps

The vertical photos made available to the project are digitized and stored to form a parallel virtual archive. In this way, the originals, on paper or film, can be protected and preserved. The images are then processed as shown in Figure 1 to create historical maps. Each digitized photo is cropped and eventually normalized to eliminate any differences in exposure; they are then orthorectified and georeferenced, using Google Imagery as the reference map. Finally, the georeferenced photos are joined together using mosaicing techniques.

Initially, we have started to process 200 photos covering the northern part of Tuscany, for the period 1943-1945; from mid-2011 the set will be increased to 1000/2000 photos. The first small set

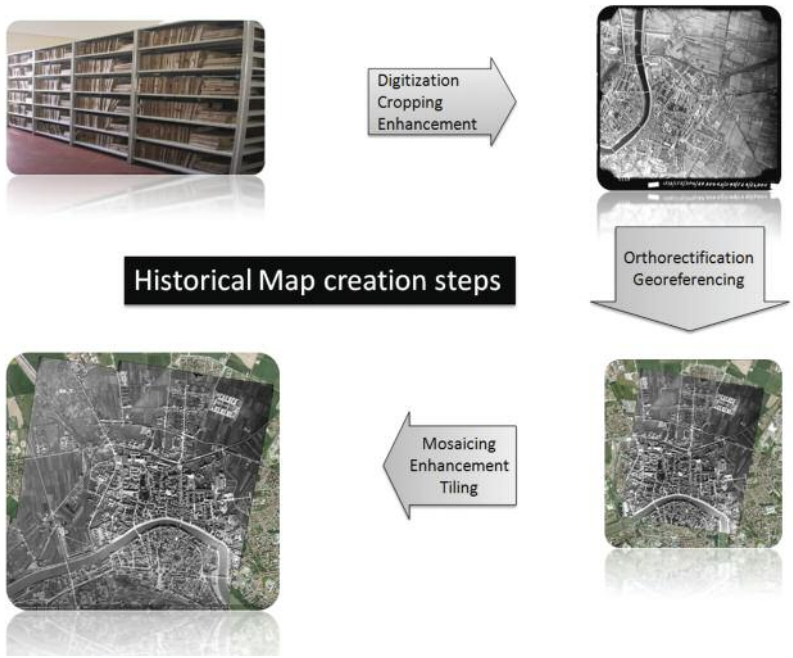


Figure: 1 The process for creating historical maps from aerial strip photography.

Figure: 2 A screenshot of the first prototype for browsing the historical maps.

has allowed us to test several procedures for processing aerial strip photos.

The workflow for creating historical maps is very heavy. Although we will develop automatic procedures in order to reduce the human effort required, the size of the archive means that it will be necessary to develop solutions based on social contributions. An important future activity of the project will concern the development of collaborative web applications for georeferencing activities.

Data Display

The historical maps - each referring to a specific historical period - will be browsable in the four dimensions via a web application based on a Google Earth plugin and some javascript libraries. Figure 2 shows the first prototype of the application.

Data integration

With the use of geographic reference (geotagging), Google Earth integrates into its maps different layers, such as video, pictures and webcam. We intend to use the same mechanism adding the time value (timetagging). The geo-historical data layers will be obtained through web mining techniques, or filtered from open archives such as Wikipedia, Youtube, Flicker, or via social contributions related to initiatives for memory preservation. The result will be a sort of Historical Geographical Atlas, where it will be possible to build spatial-temporal tours.

Future activities

This article describes the first steps of the GeoMemories project and the challenges currently being addressed in designing tools for the realization and visualization of historical maps. A major issue will be to find the resources needed to digitize and process the entire archives. We must also find ways to solve critical questions regarding the copyright of the original photos.

Our hope is that the outcome of this project will be to raise interest and to attract funding at European and international levels in activities that will ensure the protection and sharing of heritage aerial photography.

Link:
http://www.geomemories.org

Please contact:
Andrea Marchetti, IIT-CNR, Italy
Tel: +39 0503152649
E-mail: Andrea.Marchetti@iit.cnr.it

Nephele – Smart Digital Memory of Artwork Conservation

by Barbara Zitová, Miroslav Beneš and Janka Hradilová

Nephele is a comprehensive information system for processing and archiving data produced in the process of art conservation.

The benefits of using digital image processing in the area of cultural heritage were identified many years ago. New sensors and modern techniques are employed in the study and in conservation of old, and often damaged, artworks. The Nephele project introduced here, focuses on material analysis research and addresses the key issue of the art conservation - identification of the painting materials used.

The Nephele system was designed for archiving and use of material research reports. To art conservators, it offers better access to the archived material reports they use. Furthermore, its internally implemented methods for image data processing improve the outcome of data analyses and the functioning of the database. Its image retrieval system offers report retrieval based on visual similarity of studied samples. This feature eases the complicated task of retrieval of relevant experience from previous conservation cases. Such a database with a broad spectrum of reports can serve as a knowledge base for future actions.

The Nephele project is realized in close cooperation between the experts from the Institute of Information Theory and Automation, the Academy of Sciences of CR, and from the Academic Materials Research Laboratory for Painted Artworks, joint site of the Academy of Fine Arts and Institute of Inorganic Chemistry, also belonging to the Academy of Sciences of CR. All organizations are located in Prague, Czech Republic. In addition to this tight partnership, we are also incorporating suggestions and ideas from other sources, including laboratories, conservation studios (UK, USA), and conference and workshop discussions. The direct participation of the experts from the cultural heritage area proved to be very important in the development of the tools for the conservators. The project started officially in 2006 and is ongoing.

In general, the functional objectives of the Nephele system can be summarized as follows:

- Archiving – standard database services
- Content based image retrieval
- Sample preprocessing
- Sample analysis – layer segmentation, layer description and morphology
- Material description
- Material classification.

The Nephele system should serve as an archiving environment with the ability to process material reports. The data is used to interpret painting techniques and to support the choice of proper conservation method and materials. The content of the Nephele extended database reflects the structure of individual reports, which describe the process of material research of given artwork. They contain all acquired information about the object - general information about the artwork and its author, information about samples taken from the artwork, results of chemical analyses, and scanned parts analyzed in different modalities. The implicitly included imaging modalities are microscopic data in visible and ultraviolet light and scanning electron microscope (SEM) data.

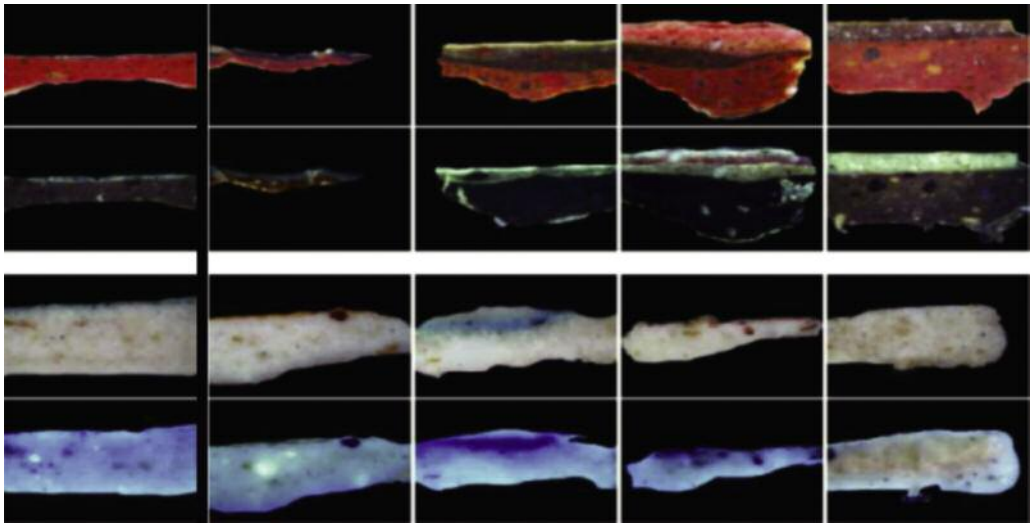


Figure 1: Results of the content based image retrieval. The left column contains query specimens; the subsequent columns in the corresponding rows are results of the retrieval in order of their similarity.

The standard database functionality, such as report creation, editing, or saving is upgraded by an image based search function, which fetches the image-wise most similar reports and its usage is often more natural than standard databases for the operator. The content-based image retrieval approach is implemented here (based on co-occurrence matrices and wavelet based descriptors), which compares the query image sample with archived image data (see Figure 1). This functionality is enabled due to the included image processing modules, which go beyond the functionality of ordinary databases and, furthermore, are intensively used in the process of artwork analysis and report creation.

The ultimate aim of the image processing modules is the identification and description of individual material layers. The Nephele system can process multimodal data, which can be consecutively enhanced (denoising, artifact removal, deconvolution), geometrically aligned (mutual information), and, finally, segmented into individual material layers (a feature-based approach). These steps produce set of base structures, which are homogenous and can be further described and analyzed. The indentified structures serve as an input for more sophisticated tasks such as the aforementioned image based retrieval, image fusion, or used material classification. An example of the Nephele system GUI is presented in Figure 2.

The Nephele project aims to facilitate the tedious work of art conservators with the use of image processing methods and database technology. The system is still under development, with the described functionality being the final aim of the project. In the long run we want to create a knowledge base of reports for use in future conservation cases. Moreover, as a byproduct of the proposed research, we plan to create a taxonomy of material descriptors.

Link:
http://zoi.utia.cas.cz/nephele.html

Please contact:
Barbara Zitová
Dept. of Image Processing , UTIA - Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic / CRCIM, Czech Republic
Tel: +420 266 052 390
E-mail: zitova@utia.cas.cz

Janka Hradilová, Academy of Fine Arts in Prague, Czech Republic
Tel: +420 233 015 324
E-mail: janka@avu.cz

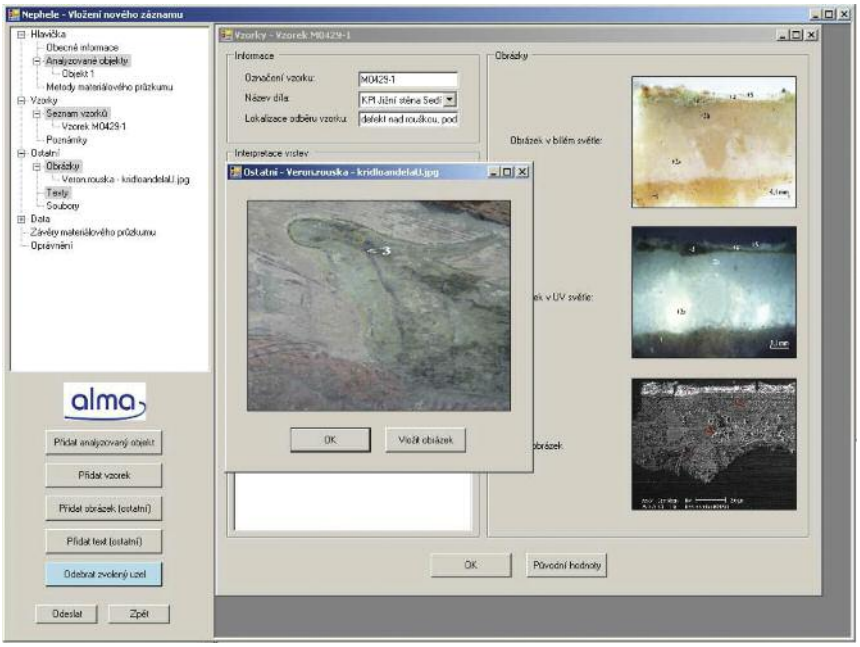


Figure 2: An example of the Nephele system GUI.

Soft Sensors: The New Frontier in Measurement for Non-Invasive Monitoring of Cultural Heritage Sites

by Umberto Maniscalco and Giovanni Pilato

Soft sensors are innovative tools for the acquisition of measurements in complex experimental conditions. We present a set of soft sensors designed for the spatial measurement forecasting of environmental atmospheric parameters in order to solve the problem of non-invasive monitoring of archaeological sites.

For the conservation of cultural heritage it is crucial to have non-invasive tools capable of monitoring both the physical and chemical conditions of the various materials composing the artefacts. The classical approach to artefact monitoring is to apply physical sensors to the artefact for a long time. These sampling campaigns must be repeated many times. This is a long process and such actions are typically hard to realize because they are expensive and/or invasive, reducing as a consequence the attractiveness and the full-fruition of the site. For this reason, in the context of the SIINDA project (an Italian national project regarding the study and the development of an integrated approach for computer aided cultural heritage conservation), we were asked by the "Agency for Cultural Heritage of the Aosta Region, Italy" to find a solution for the non-invasive monitoring of cultural heritage sites.

We propose to employ the soft sensor paradigm. Soft Sensors, known also as virtual sensors, are software tools capable of calculating quantities that cannot be measured or that are difficult to measure. They are based on technologies that provide an estimation of measurements by creating a mathematical model from real data. This approach allows the computation of a given variable value from other measurements that are related to it.

Our methodology realizes a spatial forecast of environmental parameters in the presence of one or more atmospheric monitoring stations (AMS) in the neighbourhood of the artefact.
The spatial forecast is based on a number of neural networks organized as an interconnected system. In this system, each neural network (or group of neural networks) constitutes a single soft sensor capable of capturing the associative model among the data and

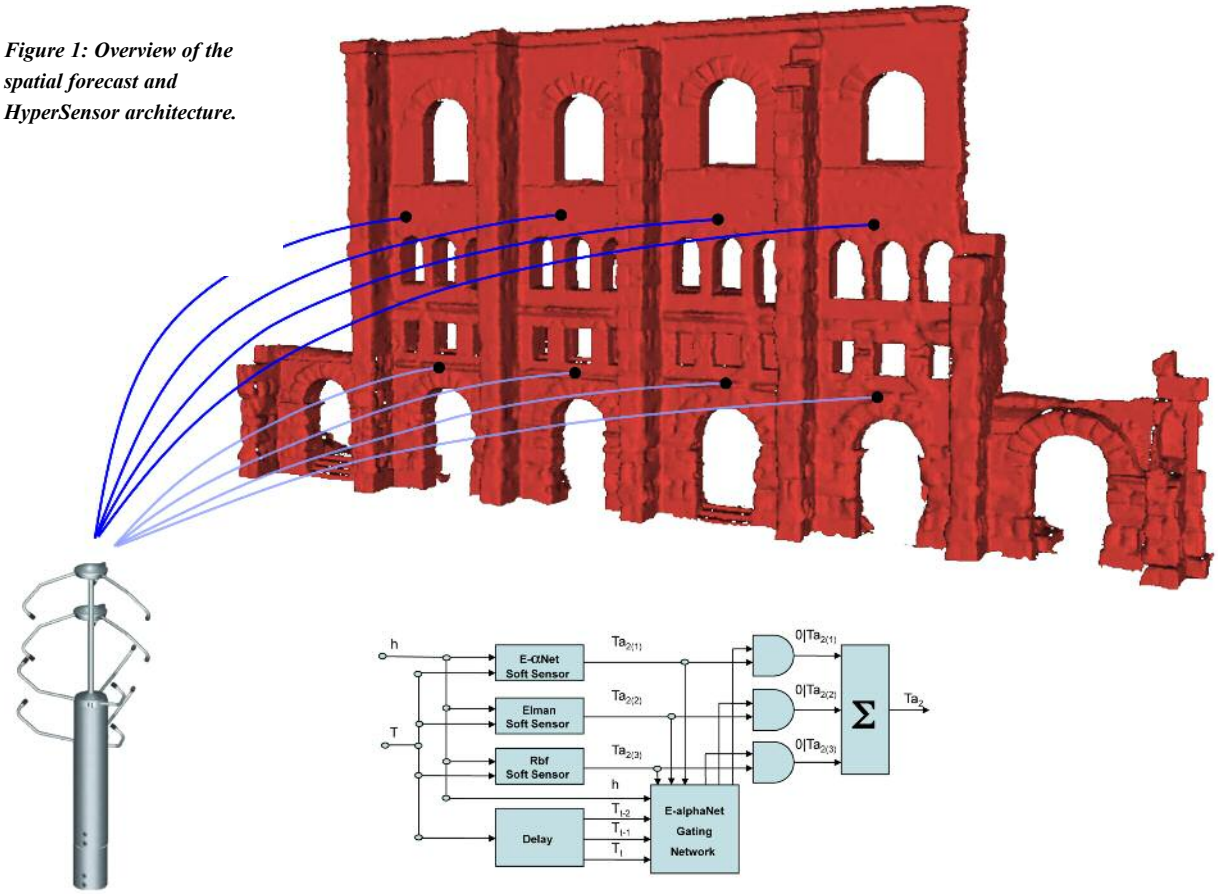


Figure 1: Overview of the spatial forecast and HyperSensor architecture.

to estimate accurately an environmental parameter value in a specific spot of the artefact. As the associative model is induced through a data driven approach, a set of sensors must be installed on the artefact for a reasonable (preferably short) training time in order to measure these parameters. After the training phase the set of sensors is removed from the artefact.

Once the soft sensors have learned the associative model, we can derive the parameter values that would have been measured by physical sensors. This is achieved using the AMS as the data source and the set of soft sensors as the forecaster. This approach is conceived as a general methodology that can be employed for the non-invasive measurement of critical parameters in cultural heritage sites.

Damage to artefacts are always the result of interaction between the materials composing the objects and the environmental factors, including pollutants, surrounding them. In particular, the physical and chemical properties of the materials composing the artefact typically change in an attempt to reach equilibrium with the environment. The state of the artefact is thus dynamic, and will change according to atmospheric parameters. This means that the recognition of a critical state in the environment, and the knowledge of the time period during which this state will remain critical, is extremely important, since it is the pre-condition of a degradation process. Therefore, we have adopted and adapted the soft sensor technology to develop a new methodology for the non-invasive measurement of atmospheric parameters.

We have designed and developed several kinds of soft sensor based on different models of neural networks and have analyzed how well they are able to estimate the measurements from a metrological point of view. The idea is to make them work as a substitute for physical sensors. We have introduced and adopted an "ad hoc" validation procedure, which consists of two phases: a statistical phase based on several statistical estimators, and a validation by comparison phase. This procedure is based on a "substitution error" that we have defined as the difference between the values estimated by a soft sensor and the measurement provided by a real sensor.

Recently, we introduced the HyperSensor architecture (see Figure 1): a mixture of several soft sensors, based on different neural network models. The performance of this new architecture has been statistically analyzed. In comparison with traditional soft sensor architectures, this architecture gives better results for each evaluator and in any measurement condition.

Our methodology has been applied to the Roman theatre (see Figure 1) in the city of Aosta for which we have a suitable set of atmospheric parameters measured both by two AMSs located close to the theatre and by real sensors placed on the theatre.

The effectiveness of the methodology in monitoring cultural heritage sites has been validated and approved by the Cultural heritage authorities of Aosta.

Please contact:
Umberto Maniscalco or Giovanni Pilato
ICAR-CNR, Palermo, Italy
E-mail:
{maniscalco,pilato}@pa.icar.cnr.it

Mathematical Models for Copper Sulphation

by Fabrizio Clarelli, Barbara De Filippo, and Roberto Natalini

Air pollution is one of the most serious causes of degradation in bronze monuments and artifacts; in particular, sulphur dioxide (SO_2) is one of the main causes of bronze disease. Mathematicians have now begun to model this kind of problem in order to assist the protection and restoration of bronze artifacts. In our research, we have introduced a free boundary model which describes the growth of corrosion on the surface of bronze monuments under SO_2 aggression, quantifying the influence of different environmental factors such as concentration of pollutants, but also humidity and temperature.

Deterioration of bronze (monuments and artifacts) is a complex problem and one of the main concerns for people working in the field of conservation and restoration of Cultural Heritage. It is extremely difficult to isolate a single factor in these processes, which are the result of the interaction between various mechanisms, many of which also occur in natural weathering. Atmospheric pollution, however, can certainly be considered as one of the key contributing factors. Despite the reduction in air pollution in European urban areas that has occurred in recent years, there is still a harmful concentration of pollutants, such as sulphur dioxide (SO_2), which negatively affects the conservation of bronze.

SO_2 can react with copper, producing several products of corrosion, in particular brochantite ($Cu_4(OH)_6SO_4$), a common copper hydroxide sulphate. This process greatly depends on the nature of the copper sample and on the presence of moisture. The condensation of moisture may depend on the structure of the material, and is critical to its reactivity to pollutants.

Within this complex framework, mathematical modelling may be useful, simultaneously providing both quantitative information and simulations of the various processes involved. Using the basic laws of electro-chemistry, models can take into account environmental conditions and their variations (air pressure, temperature, moisture, rainfall, and concentrations of pollutants) and can be used as additional tools in determining optimal strategies of intervention.

At the Istituto per le Applicazioni del Calcolo "M. Picone" (IAC – CNR) in

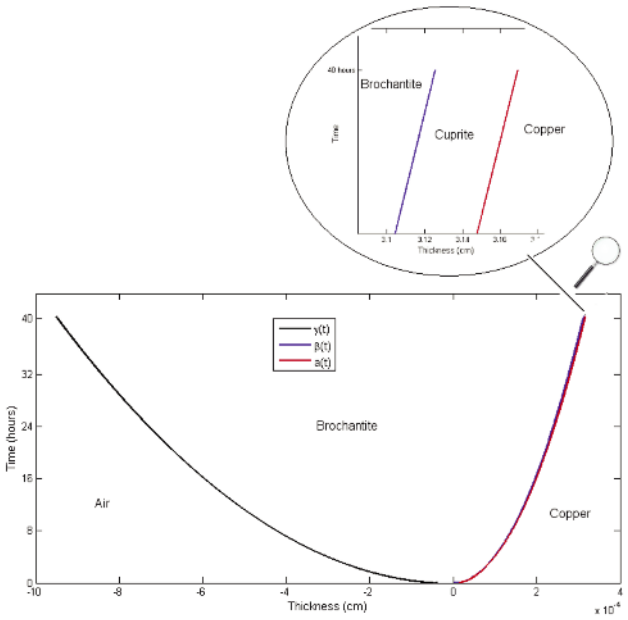


Figure 1: Formation of cuprite and brochantite layers over the copper boundary.

Rome, we are currently investigating this modelling approach, and we have proposed a differential model with two free boundaries describing the evolution of corrosion products of bronze, a copper tin alloy (88-12%).

This model includes important features such as swelling of the products of corrosion, which depends on the structure of the matter, and the total wasting of bronze. For simplicity, as a first step, we have considered only copper corrosion product development. To describe copper corrosion evolution, we consider, on the most internal layer, the transformation of copper in cuprite by oxidation, which involves the transformation of two moles of copper in one mole of cuprite (Cu2O), accompanied by a volume change caused by the changed molar volume. The cuprite external layer reacts with SO_2 in the presence of water and oxygen, thus for every mole of brochantite that is formed, two moles of cuprite and one mole of SO_2 are required. Swelling is expected also in this case, see Figure 1.

To calibrate this model, in collaboration with the Chemistry Department of the University Roma "La Sapienza", specific laboratory experiments have been performed, in a climatic chamber and atmosphere control (Erichsen Mod. 519/AUTO), with 40º Celsius of temperature, 100% of Relative Humidity, and 200 ppm of SO_2. Under these conditions, the growth of corrosion products (cuprite and brochantite) after 40 hours was about 13.2µm, with a standard deviation of 2.4µm. The model was thus calibrated using these results. In addition, using environmental data (SO_2, temperature, pressure and humidity) collected in Piazzale Fermi in Rome (Italy) during 2005, a simulation was performed to assess the formation of corrosion products under these conditions. The simulation with Piazzale Fermi data is shown in Figure 2. It must be observed that a greater cuprite thickness is due to the model calibration which is obtained under high Relative Humidity and a very high SO_2 concentration; environmental conditions have a lower concentration of SO2 and a variable Relative Humidity.

In the future, our aim is to develop standard predictive software based on our model, as an aid for monitoring outdoor and indoor monuments.

Link:
http://www.iac.rm.cnr.it/~natalini

Please contact:
Roberto Natalini
Istituto per le Applicazioni del Calcolo "M. Picone", Italy
Tel: +39 06 88470257
E-mail: r.natalini@iac.cnr.it

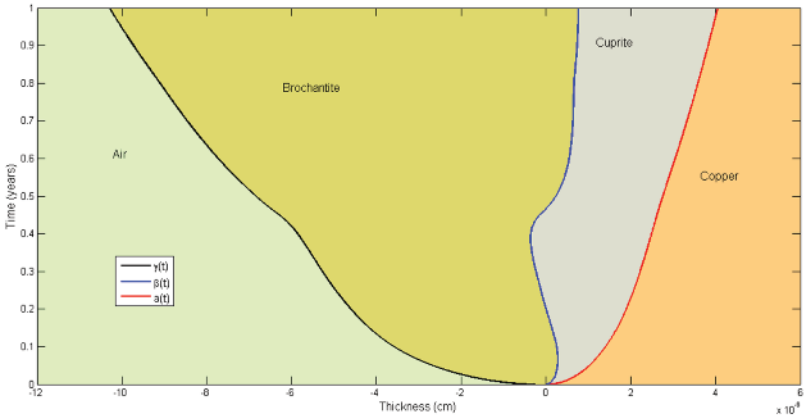


Figure 2: Simulation of Cuprite and Brochantite formation with P. Fermi (Rome) data.

Three-Dimensional Reconstruction of a Nanoparticle at Atomic Resolution

by Joost Batenburg and Sandra van Aert

Researchers from CWI, the University of Antwerp, ETH Zürich, and the Swiss Federal Laboratories for Materials Science and Technology, have developed the first practical method to reconstruct the 3D position of all atoms inside a nanoparticle. Their breakthrough required a combination of statistical image modelling, discrete tomography, and state-of-the-art electron microscopy. The research, published in Nature on 2 February 2011, will allow researchers to better understand the three-dimensional structure of materials. This is important for the development of new catalysts, solar cells and better LED lighting.

This project employed the use of tomography, an imaging technique which deals with the reconstruction of 3D images from a series of 2D projection images, acquired from a range of angles. The image reconstruction task is then modelled as a mathematical inverse problem, which can be solved using techniques from analysis and linear algebra. In the case of atomic resolution imaging, this problem is highly challenging, as only a few projection images – typically less than four - can be recorded. As a result, the reconstruction problem is highly ambiguous: there exist many different atomic configurations sharing the same projections.

The key to solving the reconstruction problem lies in the incorporation of prior knowledge within the reconstruction algorithm. In a crystal, the atoms are not randomly distributed: they closely follow a regular lattice structure. Conventional tomography methods model the image as a greylevel distribution, corresponding to the density of the material. At the atomic level, nature is actually discrete: an atom is either absent or present at a particular location, but there cannot be half an atom. Joost Batenburg, researcher at CWI, applied new combinatorial reconstruction algorithms for discrete tomography, which take into account prior knowledge of the lattice structure and the discrete nature of atoms. Using this knowledge, just a few projections are often sufficient to accurately reconstruct the atomic configuration of a nanoparticle.

Nanoparticles are structures consisting of only a few atoms up to thousands of atoms, ranging in size from one to 100 nanometres. Nanoparticles often have physical, chemical, or biological properties that are very different from the same materials at larger scale. For example, gold changes colour from yellow to red as the size of the gold particles is reduced to a few nanometres. The often unexpected properties depend on the exact three-dimensional configuration of all atoms within the nanoparticle. It has, therefore, been a long-term challenge to zoom down to the atomic level. For the first time, scientists have managed to make this dream come true .

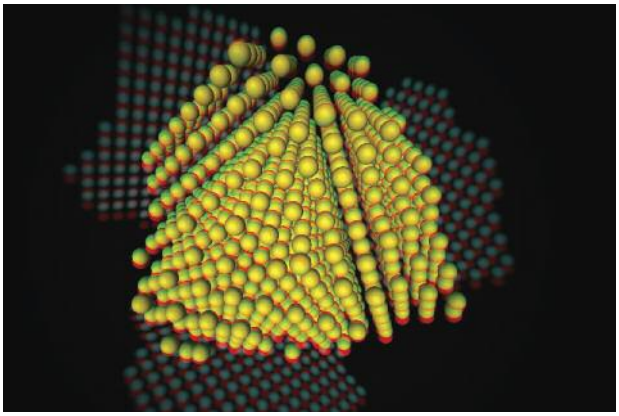


Figure 1: Visualization of the three-dimensional structure of a silver nanoparticle at an atomic level. Sophisticated measurement and reconstruction techniques are applied to images obtained with one of the most powerful electron microscopes in the world.

The researchers applied their algorithms to images of a silver nanoparticle obtained with one of the world's most powerful electron microscopes. By imaging this particle from different directions, two dimensional projection images were acquired at atomic resolution. To use these images for discrete tomography, an intermediate step was required: atom counting. Sandra van Aert from the University of Antwerp applied sophisticated statistical models to count the integral number of atoms in each projected column. The resulting discrete projections were then used as input for a discrete tomography algorithm, resulting in the atomic resolution 3D reconstruction.

International team

The research was performed at the EMAT laboratory (Electron Microscopy for Materials Science) of the University of Antwerp, under the guidance of Sandra Van Aert and Gustaaf Van Tendeloo, and at CWI in Amsterdam, under the guidance of Joost Batenburg. The research team also included Rolf Erni and Marta Rossell, scientists from Switzerland who acquired experimental images with a prototype of the new Qu-Ant-EM electron microscope which has been in use in Antwerp since June 2010.

The research conducted by this team pushes the limits of our perception and will promote further research on nanoparticles. The ability to use new algorithms to visualize the three-dimensional structure of nanoparticles at the atomic level will open up new possibilities in the development of innovative materials with fascinating and revolutionary properties. One will be able, for example, to acquire all necessary knowledge to understand the structure and activity of catalysts and to design optimal catalysts, for specific uses, for instance in the automobile industry. Furthermore, knowledge about the atomic structure holds promises for the development of more efficient solar cells, computer chips, versatile lasers and brighter LED lighting.

Links:
http://homepages.cwi.nl/~kbatenbu/
http://dx.doi.org/10.1038/nature09741

Please contact:
Joost Batenburg, CWI Amsterdam, Netherlands
Tel: +31 20 592 4037
E-mail: K.J.Batenburg@cwi.nl

Robust Methods for Image Processing in Anthropology and Biomedicine

by Jan Kalina

Common methods of 2D image analysis in anthropological and biomedical applications are too sensitive to the presence of artifacts or outliers in the data. We study robust statistical methods suitable for high-dimensional information processing. The results acquired from practical anthropological and biomedical tasks demonstrate the benefits of methods based on robust statistics.

Standard methods of image analysis applied to forensic anthropology or biomedicine (for example, analysis of medical ultrasound images) are not robust. Typical procedures start with a dimension reduction and feature extraction and proceed to a machine learning method (for example neural networks, support vector machines or Bayesian classification). Robust mathematical statistics demonstrate that these methods are non-robust in terms of their vulnerability to outlying values. Individual classifiers are often combined with ad hoc procedures to create complex black box systems with

Figure 1: Forensic anthropology aims to identify crime victims. The superprojection compares the skull with 2D photographs of missing persons.

numerous parameters, rendering impossible a simple interpretation. The noise plays a highly influential role in the information extraction from images and the methods are also sensitive to two-dimensional artifacts or to specific assumptions which cannot be verified.

In the Centre of Biomedical Informatics (project of the Ministry of Education, Youth and Sports of the Czech Republic) we propose robust approaches for 2D image analysis. Robust statistical methods bring new insights to image analysis problems as well as being very appropriate to meet their specific requirements. We apply new robust statistical methods and algorithms to the analysis of real data in anthropology and biomedicine. One of the aims is a denoising of 2D images, for which we propose a robust filter based on local smoothing by robust regression. Further, we propose a robust machine learning procedure based on a new estimator of location and scatter for multivariate data. Reflecting the most recent developments of robust statistics, the method possesses a high breakdown point, which is a statistical measure of sensitivity against outliers or noise in the data. This technique has desirable properties including robustness to noise, reliability for a high dimension of the data, computational stability, clear interpretation and model selection.

The motivation for the anthropological research is a distant aim to examine the connection between the size and shape of facial features and the genetic code. The basic task of anthropometrics is person identification, which can be interpreted as a statistical problem of analysis of variability. Namely, we work in face recognition, comparing the image of a person with images in a database and the inter-person variability must be differentiated from intra-person variability. Anthropometric measures in the face allow also the diagnosis of genetic diseases based on a craniofacial dysmorphism.

In the task of face detection, each face must be automatically localized in the image. This is again a problem of machine learning, because each part of an image is classified as a face or a non-face. We carry out the face detection by detecting symmetrical areas within an image. Symmetry is measured between two neighbouring rectangular areas in the images using a new robust correlation coefficient, which down-weights regions of the face violating the symmetry. It also yields reliable results in face localization for faces which are not entirely symmetrical. Further, we have developed a robust nonparametric discrimination technique allowing location of facial landmarks, which are biologically defined points of correspondence in each face. The use of optimal templates in face detection ensures the results to be resistant to severe noise in the image, occlusion of the face, illumination changes, asymmetric hair style or rotation of the face.

The results of 2D image analysis of faces also have applications to forensic anthropology. In this area we intend to implement a 3D image analysis method for forensic analysis of skulls and faces. Geometric morphometrics based on a certain point of landmarks is commonly used to compare a skull of a crime victim with a 2D image of the face of a missing person. Here we propose to objectify the specification of landmarks. For the sake of morphometric analysis, we will perform an optimization procedure involving the computation of weights for individual landmarks for a training set of skulls and images of the face. This will enable us to maximize the discrimination between the image corresponding to the particular skull and all other images. Robust statistical methods are underlaid by profound theoretical reasoning and yield very promising and reliable results in anthropological and biomedical applications.

Link: http://www.euromise.org

Please contact:
Jan Kalina, Centre of Biomedical Informatics, Institute of Computer Science, Academy of Sciences of the Czech Republic / CRCIM
Tel.: +420 266053099
E-mail: kalina@euromise.cz

DIRECT: the First Prototype of the PROMISE Evaluation Infrastructure for Information Retrieval Experimental Evaluation

by Nicola Ferro

PROMISE is a network of excellence focused on the experimental evaluation of multilingual and multimedia information access systems. One of its key contributions is to develop and provide an open evaluation infrastructure, which brings automation to the evaluation process, managing, curating, and providing access to the scientific data produced during the evaluation activities.

Experimental evaluation is a key activity for driving and supporting the development of multilingual and multimedia information access systems. It is an essential part of the scientific process since using shared data sets and evaluation scenarios systems can be compared, performances can be better understood, and progress can be pursued and demonstrated.

Large-scale evaluation initiatives, such as Text REtrieval Conference (TREC) in the United States, the Cross-Language Evaluation Forum (CLEF) in Europe, and the NII-NACSIS Test Collection for IR Systems (NTCIR) in Asia, contribute significantly to advancements in research and industrial innovation in the information retrieval sector, and to the building of strong research communities. A study conducted by NIST reports that "for every $1 that NIST and its partners invested in TREC, at least $3.35 to $5.07 in benefits accrued to IR researchers. The internal rate of return (IRR) was estimated to be over 250% for extrapolated benefits and over 130% for unextrapolated benefits".

Large-scale evaluation campaigns produce a huge amount of extremely valuable scientific data which provides the foundations for subsequent scientific production and system development and constitutes an essential reference for literature in the field. This data is also economically valuable, due the considerable effort devoted to its production: the NIST study estimates in about 30 million dollars the overall investment in TREC.

Nevertheless, little attention has been paid over the years to modelling, managing, curating and accessing the scientific data produced by evaluation initiatives, despite the fact that the importance of scientific data in general has been highlighted by many institutional organizations, such the European Commission, the US National Scientific Board, and the Australian Working Group on Data for Science.

Objectives

Our goal is to deliver a unified infrastructure and environment for data, knowledge, tools, methodologies, and the user community in order to advance the experimental evaluation of complex multimedia and multilingual information systems. The evaluation infrastructure will:

- manage and provide access to the scientific data produced during evaluation activities;
- support the organization of evaluation campaigns;
- increase the automation of the evaluation process;
- provide component-based evaluation;
- foster the usage and understanding of the scientific data;

A user-centered design approach will be adopted involving the different stakeholders, eg scientists, evaluation campaign organizers, system developers, students, in the development of the infrastructure.

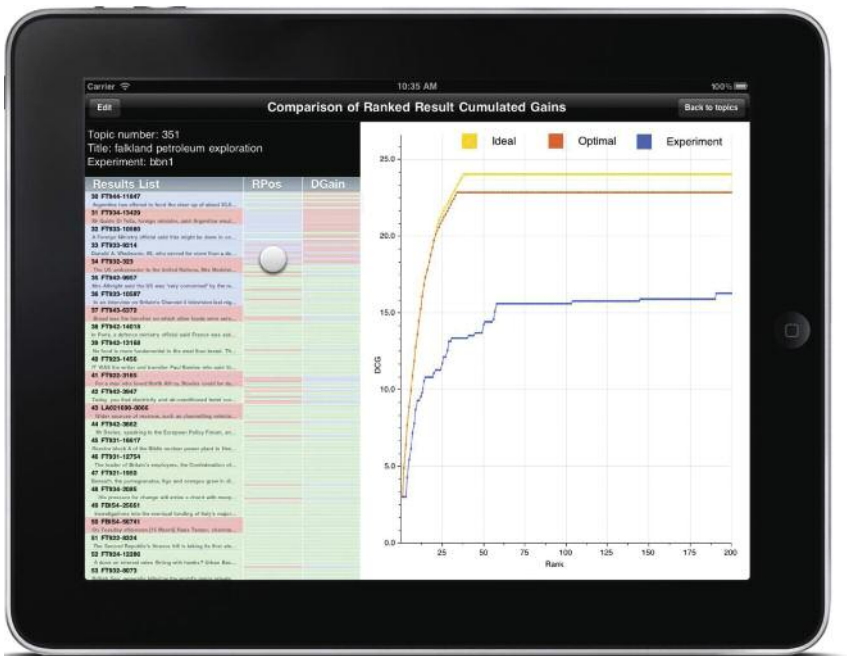


Figure 1: DIRECT provides tools for statistical analyses and reporting of results

DIRECT: the First Prototype

The outcome of this effort is the Distributed Information Retrieval Evaluation Campaign Tool (DIRECT), which

- introduces a conceptual model of the information space of an evaluation campaign;
- provides metadata describing the scientific data managed, to enable sharing and re-use;
- adopts a unique identification mechanism allowing explicit citation of and easy access to the scientific data;
- manages all the aspects of an evaluation campaign, and provides tools for statistical analyses and reporting of results.

DIRECT has been developed and tested in the course of the annual CLEF evaluation campaigns since 2005. It now manages and provides online access to much of the data produced over ten years of CLEF. It also aims at improving

interaction with the experimental results by researchers and system developers. We are now investigating the adoption of innovative devices, such as the iPad, which can allow for a natural and easy interaction with the experimental results and scientific data in real time.

Next Steps

PROMISE is a three year project beginning in September 2010. It will issue releases of the evaluation infrastructure with new functionalities, such data annotation and visual analytic techniques, annually. In order to achieve a better representation, interaction, and understanding of experimental results, we are investigating how best to exploit human-computer interaction and the principles of visual analytics. This will be the topic of the PROMISE Winter School Information Retrieval meets Information Visualization, which will be held in January 2012, Zinal, Switzerland.

The information retrieval area is now beginning to explore and exploit the scientific data produced during the evaluation studies by making use of methods typical of the database and knowledge management areas. The aim of the Data infrastructurEs for Supporting Information Retrieval Evaluation (DESIRE 2011) workshop, co-located with CIKM 2011, the 20th ACM Conference on Information and Knowledge Management, in October 2011, Glasgow, UK is to bring together experts from the three communities in order to discuss the challenges involved. The intention of the organizers is to produce a roadmap and a set of initial best practices guiding the development of evaluation infrastructures to manage experimental data.

Links:

PROMISE: http://www.promise-noe.eu/

DIRECT: http://direct.dei.unipd.it/

DEMO: http://www.youtube.com/watch?v=fDsXDCUPkiM

CLEF 2011: http://www.clef2011.org/

CLEF: http://www.clef-campaign.org/

DESIRE 2011 Workshop:
http://www.promise-noe.eu/events/desire-2011/

PROMISE Winter School 2012:
http://www.promise-noe.eu/events/winter-school-2012/

TREC Economic Impact Study:
http://trec.nist.gov/pubs/2010.economic.impact.pdf

Please contact:
Nicola Ferro, University of Padua, Italy
E-mail: ferro@dei.unipd.it

Potential Cell-phone Dangers: Latest Studies

by Harry Rudin

The results of two studies — actually collections of studies — have been published just recently. These are the Europe-wide Interphone study and the Swiss National Research Program NRP 57, Entitled "Non-Ionizing Radiation – Health and Environment". In few words, neither has identified measurable health dangers arising from cell-phone use — but both recommend further study and caution.

Swiss NRP 57 Study

The Swiss National Science Foundation financed eleven research projects in four categories: dosimetry and exposure assessment, laboratory studies and epidemiology, cell biology, and risk perception. The projects were carried out over a three-year period.

In the first category, a study was made to discover the major source of non-ionizing radiation. There are many wireless devices that radiate but, based on an extensive set of week-long measurements made in the city of Basel, the major source of radiation remains the cell phone. How much radiation depends on the transmission system used (UMTS radiates less than the older GSM) and using a headset reduces the exposure by a factor of ten. A surprising and upsetting result of the study is that the modern induction stove causes very strong electromagnetic fields and could well be a substantial danger for pregnant women and their unborn children. This is particularly true for professional cooks.

In the second category, it was clearly shown that the pulse-modulated radio-frequency fields emitted by cell phones do affect the human brain. First, these induce short-term effects observable in the brain's electro-encephalogram. Some effects continued after the exposed person went to sleep but these effects did not have a sleep-disturbing result. Changes in the cerebral blood circulation and heart rate were also observed but these were small and no negative effect on a person's well being was observed.

In the cell-biology category, DNA fragmentation in mammalian cells was shown. The frequency of the radiation was not cell-phone frequencies but normal household 50 Hz. The observed effect on cellular processes was very weak and presumably of no medical consequence.

As to results of the risk-perception study, the Swiss public accepts wide-spread cell-phone use but at the same time is somewhat apprehensive of the radiation coming from base stations. Some 5 % believe that they are hypersensitive to electromagnetic fields and suffer health problems as a result.

While the NRP 57 studies found no direct evidence for health problems caused by cell-phone radiation, the recommendation is that research be continued, particularly as cell-phone usage increases and technology changes so rapidly. The first link below is a summary of the findings.

Interphone Study

The Interphone study is the largest study yet of potential cell-phone dangers, having been carried out in thirteen European countries and having identified some 6500 cancer patients. The patients were selected on the basis of four different kinds of brain tumors. The cell-phone use of these 6500 patients was compared with the cell-phone use of a similar number of healthy individuals. Many individual (country-based) studies had been published earlier but only recently have the results of all the studies been pooled together in the hopes of providing greater statistical certainty.

The results are confusing, a likely result of imprecision inherent in the interviews used to estimate cell-phone usage.

Smartphone apps can monitor the phone's radiation level and recommend whether it's at a safe level to talk or not.

Cancer patients were identified and interviewed to learn of their cell-phone use. For comparison, non-patients from a similar environment were identified and also questioned about cell-phone usage. The study was based on 30 to 60 year olds.

If one casually looks at the results one could conclude that for very high usage, the chance of a brain glioma (one of the four tumor types investigated) resulting from cell-phone radiation is significant. Casually looking at other usage categories, one could conclude that for some types of tumors, moderate cell-phone use has a protective effect! This is hard to explain on a medical basis. These observations and the results of a calculation of the statistical confidence interval (which show relatively little confidence in the results) weaken the study's conclusions. Quoting from the report referenced below: "The possible effects of long-term, heavy use of mobile phones require further investigation."

Probably the difficulties are a result of the interviews: How accurate are a person's recollection of the intensity of his or her cell-phone use two, five, or even ten years in the past? And, what biases are introduced when a person has learned that he or she actually has a tumor?

If there is a weak conclusion, it is that prolonged high usage may cause a brain glioma. The glioma develops on the side of the head where the cell phone is held. High usage is defined as a total accumulation of more than 1640 hours of cell-phone use.

Again, quoting from the conclusion of the study, "… biases and errors limit the strength of the conclusions we can draw from these analyses and prevent a causal interpretation" (quoted from the third link below). Clearly there is a need for further study as usage --- particularly among young people --- is growing and there is a fear that adolescents may be more sensitive. These worries may be balanced by newer technologies with reduced radiation.

The second link below is a very readable summary of the results of the Interphone study --- but in German. The third link was written in English by the study's coordinator and gives details of the study.

Conclusions

What has been learned? Non-ionizing radiation from cell phones does have an effect on our bodily processes; whether this has an effect on our health remains unclear. There is a weak indication that heavy cell-phone use may cause a brain glioma but this is by no means certain. Until further results become available, we might take advantage of having learned that the use of a headset reduces radiation delivered to the head by a factor of ten.

Links:

http://www.nfp57.ch
Non-Ionizing Radiation – Health and Environment; Synthesis report (in English)

http://www.forummobil.ch/files/webcontent/documents/frequentia_10_d.pdf
G. Duerrenberger's analysis of the Interphone study (in German)

In http://www.forummobil.ch/de/suche
"Brain tumor risk in relation to mobile telephone use: results of the INTERPHONE international case–control study" (in English)

Please contact:

Harry Rudin
Swiss Editor for ERCIM News
hrudin@smile.ch

Crowdsourcing Culture with a Twist

by Pär Hansson

In Stockholm, Sweden, there is a new initiative which looks to bring a new way of creating and funding cultural projects. The idea is to let citizens influence the distribution of funding for cultural projects. The concept combines private money with public money to fund the creation of culture, using a new technical platform.

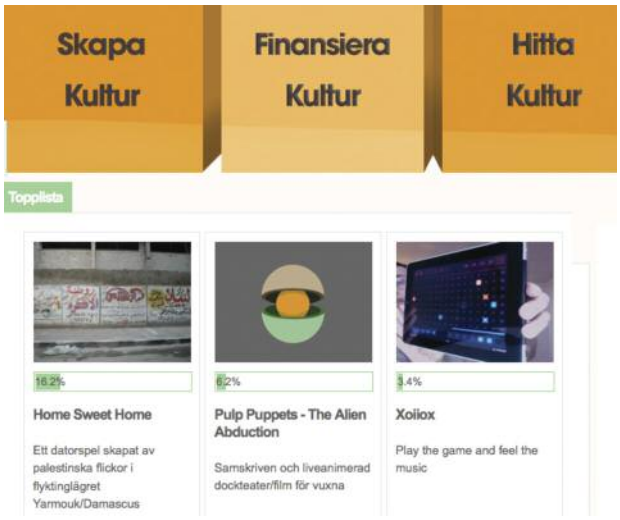
In an applied research project, called CrowdCulture, which was founded around the question of "How would we distribute cultural funding if we invented that concept today", the Swedish Institute of Computer Science (SICS) is working together with a communication company in developing new technology to combine private and public funding for the creation of a new culture. In Sweden today, most cultural investments follow a top-down political logic. One larger aim of this project is to engage more people to contribute and participate in the cultural development, and consequently delegate decision-making power to a broader public with different references and quality criteria.

We are also addressing new ideas around different crowdsourcing concepts and techniques, knowing that crowdsourcing and working with experts are equally qualitative methods in many fields - take for example the much talked about article confirming similar standards of quality of Wikipedia and The Encyclopedia Britannica. Furthermore, we know that micro funding is a great tool in creating local empowerment and initiating important processes of change; it was not without reason Muhammad Yunus got the Nobel prize for his practice with the Grameen Bank in Bangladesh. Furthermore, we believe that arts and culture respond, in their own ways, to the circumstances in which society finds itself. Considering this, we think that funding mechanisms should respond to society's circumstances as well, instead of to the logics of top politicians only.

Private symbolic funding steers larger sums

The core of the CrowdCulture project is a web based platform which allows people to invest small monthly sums in any available cultural projects they choose. Currently, the system allows an individual to contribute as little as €5 per month. An individual person's investment is spread out over time, whereby funding trickles out per time unit according to which cultural project the giver choses to support at the moment. However, as a twist on crowdfunding, the participating individuals' support for different projects will have a leveraging effect, diverting a potentially much larger sum from public funding or other contributing organizations to the supported projects.

At any time there can be any number of "passive" contributing organizations and companies, which donate larger sums to the system. This funding is not directed towards particular projects, but can be tagged to support certain categories of cultural projects. Since a lot of the parameters of the system -such as the number of individuals, number of projects, the amount of contributed money, which project the individuals support - are always changing, the leveraging effect is of course also always changing in this highly dynamic real-time system. Those projects that meet their budgetary targets within a given time frame are awarded the requested funding.



Screenshot of the CrowdCulture application.

Encouraging new cultural initiatives

In addition to funding distribution, the service allows a paying member to suggest a new project they would like to work on and submit it to the crowd to decide if it is suitable. The site also allows for, and is largely marketed by, connections to the social web, as well as in-site crowd participation mechanisms.

Used by the City of Stockholm

The project ran a successful three-month beta test during the fall of 2010, attracting 500 members using only the social web for awareness. During the test a combination of public, private, and industry funds were distributed, with the City of Stockholm as the largest contributor. The success of this trial has caused the City of Stockholm to, as of April 2011, continue distribution of cultural funding through this system. In this first stage mostly more conventional crowd participation techniques were used while examining different crowdsourcing mechanisms. The continuation of this work applied to a similar 'democracy in action' area would also look at the relationship between crowd wisdom and the rest of the democratic mix, as well as community support for producers. This very young area has caused a lot of discussion in the democratic arena in general the last couple of years, highlighted by the quite diverse opinions in the media articles this project alone has generated.

The CrowdCulture technology is being developed by the Swedish Institute of Computer Science (SICS), in cooperation with concept developer Fabel Kommunikation AB, supported by the City of Stockholm, and the Swedish innovation agency VINNOVA.

Links:
http://www.crowdculture.se
http://www.sics.se/projects/crowdculture

Please contact:
Pär Hansson, SICS, Sweden
E-mail: par@sics.se

Building a Building for Nanotechnology

by Harry Rudin

On 17 May, IBM Research and the ETH/Z (Swiss Federal Institute of Technology at Zurich) opened the Binnig and Rohrer Nanotechnology Center on the campus of the IBM Zurich Research Laboratory in Rueschlikon, Switzerland. The name and the location are well chosen as it was here that Gerd Binning and Heinrich Rohrer invented the STM (Scanning Tunneling Microscope) some 30 years ago. They were awarded the Nobel Prize for Physics for this in 1986. The timing was also ideal as the opening helps IBM celebrate its centennial.

Binnig and Rohrer's invention was effectively the birth of today's nanotechnology science. There was earlier, rather scattered work at the nano-scale (ie, at dimensions in the order of 10-9 meters) but the STM and later the AFM (Atomic Force Microscope, invented by Binnig in 1985) provided the critical turning point and effectively laid the foundation for what is now the field of nanotechnology.

Nanotechnology has to do with structures with a dimension of less than 100 nanometers --- about 1000 times smaller than a strand of human hair. But working at this dimension requires rather special conditions. In fact a new building with some 6000m^2 floorspace at a cost of roughly 70 million Euro was necessary to house the required equipment and researchers. This includes almost 1000m^2 of clean-room facilities of the sort required for the production of the most modern integrated circuits. The facilities range from class 100 (ISO 5) to class 10000 (ISO 7). The first means that there are less than 100 particles larger than 0.5 micrometer in a cubic foot of air in the clean-room, some 10,000 times less than in the atmosphere we are used to. Temperature and humidity are carefully regulated.

The planning and construction of the "noise-free" laboratories have advanced the state of the art. Unique in the world and requiring a design task that was a substantial research effort in itself, these laboratories are necessary for the extraordinary investigations that are planned. The IBM Research designer of these facilities, Emanuel Loertscher, calls the laboratories "ultra-isolated." The foundation rests on the bedrock below the laboratory. To further minimize the vibration there are active and passive techniques which were developed to isolate the 30- to 70-ton platforms on which the experiments will take place. This isolation reduces vibration from the rest of the laboratory, the nearby super-highway, and nearby railroad tunnel so that it remains below 500 nanometers/second. As far as acoustic noise is concerned, noise from outside the laboratories is reduced by roughly 60 db. Special shielding and dynamic electromagnetic cancelling fields reduce the alternating electro-magnetic fields to a value below 5 nano-Teslas and the static fields to a value below 20 nano-Teslas. Changes in temperature are kept below 0.1 C. per hour. The isolation is so complete that it is almost unnerving for a researcher to be in such an enclosure; the researchers sit outside of these laboratories in a less severe environment,

Figure 1: Clean-room facility at the Binning and Rohrer Nanotechnology Center (courtesy of IBM Research - Zurich).

controlling and observing their experiments remotely. Their presence in the room where the investigation takes place would severely disturb the experiment in any case.

The nanotechnology laboratory facilities will mainly be used by researchers from IBM and the ETH although other institutions will also have use of the laboratory. Both the ETH and IBM have already produced exciting results in the nanotechnology world. For example there has been work going on at the ETH on field-effect transistors based on carbon nanotubes and investigations of the use of graphene as a basis for quantum computers. At IBM there has been work on spintronics, using the magnetic moment of electrons for computing, as well as on transistors based on nanowires and integration of photonics for communication within integrated-circuit chips. The new facilities will facilitate further work on these topics.

Figure 2: View of stabilized platforms for nanotechnology investigations (courtesy of IBM Research - Zurich).

Looking to the future, we know that materials behave differently at the atomic level compared to their behavior at the bulk material level, creating various unique phenomena. The Center will explore these phenomena. From IBM's point of view, a major question is what comes after the transistor. Nanowires have the advantage of very small power consumption. New display and storage technologies are also exciting topics. Bringing two groups of top-notch researchers together in a uniquely suited facility will certainly produce exciting results over the coming years, not only for nanoelectronics but also for medicine and means for environmentally friendly production.

Links:
http://www.zurich.ibm.com/nanocenter/
http://www.zurich.ibm.com/pdf/nanocenter/Nano_Center_Fact_Sheet.pdf

Please contact:
Harry Rudin, Swiss editor for ERCIM News, Switzerland
E-mail: hrudin@sunrise.ch

4th Summer School on Network and Information Security (NIS'11)

"The Challenge of the Changing Risk Landscape" was the special theme of the 4th Summer School jointly organised by The European Network and Information Security Agency (ENISA) and FORTH on 27 June - 1 July 2011, Crete, Greece.

The Changing Risk Landscape refers to the dynamics, dependencies and complexity inherent to Information Technology. In this context, the changing security "ecosystem" is one of the main challenges of Information Security. The relevant stakeholders need to improve reaction, and extend collaboration and information exchange in order to achieve timely and effective responses to the posed challenges. This seems to be the only option for Information Security in the continuous arms race between defense strategies and offense attempts.

ENISA is dedicated to promoting a culture of security in Europe that will improve the ability of EU member states to respond to cyber-attacks. It does so, by pursuing a strategy of mitigating risks through awareness, studies, reports and position papers on current NIS matters. Towards this objective, ENISA and ICS-FORTH brought together in this summer school a distinguished faculty from around the world with the purpose of identifying current trends, threats and opportunities against the background of recent advances on NIS measures and policies.

Recognising the multi-dimensional facets and intricacies causing changes in the information risks landscape, an array of lectures covered a variety of key aspects on policy, economic, legal and research matters. The audience included policy makers from EU member states and EU institutions, decision makers from industry and members of the academic community.

By going through a natural evolution cycle, but also by adopting current trends in networking and exchange of knowledge, this year's Summer School helped increasing interaction among participants. Breakout sessions enhanced the dialogue and exchange of ideas, while cutting edge security issues were collected before the event through publicly accessible fora and channelled into the discussions.

Summaries of these interactive sessions were compiled and distributed during the event.

Link:
http://www.nis-summer-school.eu

Please contact:
Ioannis Askoxylakis
FORTH-ICS, Greece
E-mail: asko@ics.forth.gr

Euro-India SPIRIT Taking Research Cooperation to the next level

Facilitating an Indian Delegate's Participation at the ICT Proposers' Day 2011

by Tom Williamson

The goal of EU-funded Euro-India SPIRIT project managed by ERCIM, is to formulate between the EU and Indian Information and Communication Technology (ICT) stakeholders a mutually beneficial research and innovation agenda which can be taken up through specific bilateral initiatives. With this specific goal in mind, Euro-India SPIRIT invited the Indian ICT expert Abhishek Sharma to the ICT Proposers' Day at Budapest in May 2011. The aim was to facilitate an enabling environment for Abhishek to familiarise himself with the variety of ICT areas in which collaboration is possible, to learn about the technical and administrative aspects of being involved in FP7 projects, and to network with prospective project partners.

Abhishek Sharma is the Chairman and CEO of NetEdge Tele-Solutions, an Indian SME which develops mobile applications and provides telecom value-added and associated services. Clients of NetEdge include Vodafone and Idea Cellular, major mobile phone service providers in India. Thus, Abhishek brought with him to Budapest a network of Indian ICT operators. In addition, Abhishek is a mobile value-added-services, applications and content consultant with rich experience in operations, project management and consulting roles in the telecom/IT domains. His wide range of experience and ability to engage with a variety of research priorities as well as his extensive network and ability to disseminate/act as a focal point made him an ideal Indian partner at the event.

Abhishek was accompanied over the course of the event by Ashok Kar, Infra Technologies, one of the European project partners of Euro-India SPIRIT and Tom Williamson, ERCIM, project coordinator of Euro-India SPIRIT.

The delegation arrived at the Hungexpo centre in Budapest with a packed agenda for the two days' sessions. Meetings covered a complete cross-section of attendees both European and non-European, private industry, academics, National Contact Points (NCPs) and Commission officials.

Abhishek was also taken to sessions of particular interest, eg cloud computing, international cooperation and mobility. These proved very stimulating and useful for Abhishek, and a number of potential link-ups were identified for follow-up. The delegation also attended the session on FP7 rules and proposal-making to give a flavour of the administrative requirements involved.

While the results of such an event , from the perspective of the Euro-India SPIRIT project, will be judged over time in links maintained and stregthened, the initial aim of familiarisation with ICT proposals was more than met over the course of the event and it is hoped that Abhishek will be amongst the first of many future Indian attendees and proposers in the years to come.

Link:
http://www.euroindia-ict.org/

Please contact:
Tom Williamson
ERCIM office
E-mail: tom.williamson@ercim.eu

Call for Participation

AmI–11 - International Joint Conference on Ambient Intelligence

Amsterdam, The Netherlands, 16-18 Noveber 2006

Ambient Intelligence (AmI) represents a vision of the future in which products and services will be responsive to the user context, offering a rich variety of applications in the professional and consumer domains. AmI combines concepts of ubiquitous technology, intelligent systems and advanced user interface design putting the humans in the center of technological developments.

Starting as the European Symposium on Ambient Intelligence in 2003 the conference has grown to an annual international event that brings together researchers and serves as a forum to discuss the latest trends and solutions.

Following the publication of early scenarios on Ambient Intelligence in 2001 by the European Commission's Joint Research Centre, the theme of AmI-11 will focus on the "The Road Ahead, Shaping the Next 10 years".

Workshops

AmI-11 hosts a series of workshops:

- Aesthetic Intelligence: Designing Smart and Beautiful Architectural Spaces
- Human-Centered Design of Smart/Ambient Environments and Services: Models and Tools for Tasks Migratability
- Interactive Human Behavior Analysis in Open or Public Spaces
- Workshop on User Interaction Methods for Elderly, People with Dementia
- Empowering and integrating senior citizens with virtual coaching
- Integration of AMI and AAL platforms in the Future Internet (FI) Platform initiative
- Ambient Gaming
- 2nd Int. Workshop on Human Behavior Understanding: Inducing Behavioral Change
- Privacy, Trust and Interaction in the Internet of Things.

Workshop on Integration of AMI and AAL Platforms in the Future Internet Privacy Public Platform Initiative

This workshop will focus on the challenges of integrating Ambient Intelligence (AmI) and Ambient Assisted Living (AAL) platforms with the Future Internet Privacy Public Partnership (FI-PPP) initiative. In this respect the workshop provides information on concrete means to engage with the FI PPP future calls and experiments, either as a user or as a co-developer.The following topics will be discussed:

- Architecture of an FI platform. Challenges from an FI viewpoint
- Architecture of existing and to come AmI / AAL platforms. Challenges from an AmI/AAL viewpoint
- How can integration take place? What are the hurdles?
- What are the means of experimentation? Can AAL be a usage area for the Future Internet?
- What is the impact of evaluation? How can we move towards a European digital signle market?
- Identification of measures and instruments
- Actions.

Participants at the workshop will include both members of the AmI/AAL platform community, members of the FI community and policy makers.

More information:
http://www.ami-11.org/
http://www.trialog.com/AALworkshop.html

Call for Participation

European Gender Summit

Brussels, 8-9 November 2011

genSET - gender in science, the European Science Foundation (ESF) and the European Cooperation in Science and Technology (COST) invite you to the first European Gender Summit in Brussels .

For the first time, this Summit will bring together key stakeholders from research, industry, publications, innovation management and policy to discuss how European innovation and research systems can benefit from more effective mainstreaming of the gender dimension. The event is held under the patronage of the Polish Presidency of the European Union.

Projects, organisations and companies are invited to present their initiatives and contribute to the shaping of future models for research and innovation.

The Summit discussions will feed into the Policy Manifesto on Integrated Action on the Gender Dimension in Research & Innovation, to be presented to the European Commission and other key policy actors. Please contribute your views to the Manifesto by completing the public consultation on Gender and Innovation.

More information:
http://www.gender-summit.eu/

New Technologies and Cultural Heritage at VIEW Conference 2011

Torino, Italy, 25-28 October 2011

"New Technologies applied to Cultural Heritage" is one of the main topics of the 12th edition of the VIEW Conference, to be held on 25-28 October 2011 in Torino. The VIEW conference is the premiere international event in Italy on Computer Graphics, Interactive Techniques, Digital Cinema, 3D Animation, Gaming and VFX.

The section "New Technologies and Cultural Heritage" seeks to promote, on a high quality standard, the application of the new technologies in view of valorizing cultural heritage, intended in the vast sense, assuming initiatives in the research and promoting education.

Researchers and representatives from companies, universities, museums or schools who have an innovative idea to present and want to have a talk at VIEW about New Technologies and Cultural Heritage are invited to submit an abstract before 31st August to the address segreteria@viewconference.it.

More information:
http://viewconference.it/

Call for Participation

SERENE 2011
3rd International Workshop on Software Engineering for Resilient Systems

Geneva, Switzerland, 29-30 September , 2011

An unprecedented level of complexity of modern software makes it difficult to ensure its resilience – the ability of a system to reliably deliver its services even when undergoing change. Yet we are observing an increasingly pervasive use of software in such critical infrastructures as transportation, health care, energy production etc. Such a trend could lead to devastating accidents unless the research community develops powerful methods for assuring resilience of software-intensive systems. The SERENE 2011 workshop provides a forum for researchers and practitioners to exchange reports on advances in all areas relevant to this challenge, such as:

- Modelling of resilience properties: formal & semi-formal techniques
- Requirements, software engineering & re-engineering for resilience
- Verification and validation of resilient systems
- Resilience prediction and experimental measurement
- Error, fault and exception handling in the software life-cycle
- Frameworks, patterns and software architectures for resilience
- Resilience at run-time: metadata, mechanisms, reasoning and adaptation
- Engineering of self-healing autonomic systems
- Quantitative approaches to ensuring resilience
- CASE tools for developing resilient systems.

The workshop is organized by the ERCIM Working Group "Software Engineering for Resilient Systems (SERENE) and supported by the Laboratory for Advanced Software Systems, University of Luxembourg, and the EU-funded Deploy project .

More information:
http://serene2011.uni.lu/

Call for Participation

4th International Conference of the ERCIM Working Group on Computing & Statistics

London, 17-19 December 2011

The 4th International Conference of the ERCIM Working Group on Computing & Statistics will take place at the Senate House, University of London, 17-19 December 2011, jointly with the 5th CSDA International Conference on Computational and Financial Econometrics (CFE'11). The conference is organized by the London School of Economics, Queen Mary, University of London, and Birkbeck University of London.

Topics

Conference topics include robust methods, statistical algorithms and software, high-dimensional data analysis, statistics for imprecise data, extreme value modelling, quantile regression and semiparametric methods, model validation, functional data analysis, Bayesian methods, optimization heuristics in estimation and modelling, computational econometrics, quantitative finance, statistical signal extraction and filtering, small area estimation, latent variable and structural equation models, mixture models, matrix computations in statistics, time series modelling and computation, optimal design algorithms and computational statistics for clinical research.

Keynote speakers

Keynote speakers include:

- Peter M. Robinson, London School of Economics, UK
- Alastair Young, Imperial College, London, UK
- Stephen G Walker, University of Kent, Canterbury, UK
- Andrew Harvey, University of Cambridge, UK
- Dick van Dijk, Erasmus School of Economics, The Netherland.

More information:
http://www.cfe-csda.org/ercim11/

Call for Participation

MUSCLE Workshop on Computational Intelligence for Multimedia Understanding

Pisa, Italy, 13-15 December 2011

The purpose of this workshop, organised by the ERCIM MUSCLE Working Group, is to provide an international forum to present and discuss current trends and future directions in computational intelligence for multimedia understanding. The workshop also aims at fostering the creation of a permanent network of scientists and practitioners for an easy and immediate access to people, data and ideas. The scientific program, organized in a single-session format, consists of invited lectures, contributed talks, and poster presentations.

The workshop will feature high-quality presentations addressing multimedia processing and understanding methodologies, ranging from statistical, neural, evolutionary methods to mixed reality and multisensor interaction, as well as a variety of applications, in the areas of public services, transport/mobility, aids to disabled people, process control and diagnosis, remote sensing, and beyond. Active participation of graduate students is strongly encouraged.

Topics include, but are not limited to: multisensor system, multimodal analysis, crossmodal data analysis and clustering, mixed-reality applications, activity and object detection and recognition, text and speech recognition, multimedia labeling, semantic annotation, and metadata, multimodal indexing and searching in very large data-bases, case studies.

Invited Speakers

- Bulent Sankur, Bogazici University, Turkey
- Sanni Siltanen, VTT, Finland

The workshop is co-sponsored by CNR and INRIA.

More information:
http://muscle.isti.cnr.it/pisaworkshop2011/

Call for Participation

SOFSEM 2012 38th International Conference on Current Trends in Theory and Practice of Computer Science

Spindleruv Mlyn, Czech Republic, January 21–27, 2012

SOFSEM (SOFtware SEMinar) is the annual international winter conference devoted to the theory and practice of computer science. Its aim is to present the latest developments in research for professionals from academia and industry, working in leading areas of computer science. Proceedings containing the invited and contributed papers will be published in the series Lecture Notes in Computer Science by Springer Verlag.

The SOFSEM program consists of Invited Talks by prominent researchers, Contributed Talks selected by the Program Committee from submitted papers, Poster Session and the Student Research Forum. SOFSEM is organized in plenary and parallel tracks, giving a unique opportunity to obtain an excellent overview of the selected research areas. All tracks are devoted to original research and leading developments.

By tradition, one track of SOFSEM is devoted to Foundations of Computer Science. SOFSEM 2012 offers in addition three, outstanding tracks: Software & Web Engineering, Cryptography, Security, and Verification, and Artificial Intelligence.

Special Event: Session on Turing Machines

In memory of Alan Turing, whose 100th anniversary is celebrated in 2012, SOFSEM 2012 will host a session on Turing machines. The session will consist of invited and contributed talks on Turing machines as the basic model of computability and complexity, and reporting new developments related to models inspired by, for example, biology and physics. SOFSEM 2012 is among the official Centenary Events of The Alan Turing Year.

SOFSEM is the ideal conference for discussions, for establishing personal contacts with colleagues, and for exchanging ideas. SOFSEM is also well-known for its familiar and inspiring atmosphere and is especially suited for young computer scientists.

Invited speakers

As usual at SOFSEM, a very distinguished feature of the SOFSEM 2012 will be the higher number of invited speakers:

Foundations of Computer Science Track:

- Yuri Gurevich (University of Michigan and Microsoft Research, USA): "What's an Algorithm?"
- Giuseppe F. Italiano (University of Rome "Tor Vergata", Italy)

Special event: Session on Turing Machines:

- Felipe Cucker (City University of Hong Kong, Hong Kong)
- Peter van Emde Boas (University of Amsterdam, The Netherlands): "Turing Machines for Dummies"
- Jiri Wiedermann (Institute of Computer Science, Academy of Sciences, Czech Republic)

Software & Web Engineering Track:

- Paul De Bra (Eindhoven University of Technology, The Netherlands): "A Fully Generic Approach for Realizing the Adaptive Web"
- Pavel Zezula (Masaryk University in Brno, Czech Republic): "Multi Feature Indexing Network (MUFIN) - Similarity Search Platform for many Applications"

Cryptography, Security, and Verification Track:

- Orna Kupferman (Hebrew University in Jerusalem, Israel): "Recent Challenges and Ideas in Temporal Synthesis"
- Krzysztof Pietrzak (Cryptology Research Group, CWI): "Efficient Cryptography from Hard Learning Problems"

Artificial Intelligence

- Kevin Warwick (University of Reading, United Kingdom): "Not Another Look at the Turing Test!"
- Roberto Navigli (Sapienza University of Rome, Italy)

More information:
http://www.sofsem.cz

Marie–Paule Cani
Winner of the Eurographics Outstanding Technical Contributions Award 2011

© Eurographics association

Marie-Paule Cani received the Eurographics Outstanding Technical Contributions Award 2011. This internationally recognized distinction, given each year to an individual in computer graphics to highlight some outstanding technical achievement, rewards her work in implicit modelling, animation and interactive shape design. Marie-Paule Cani, a university professor in Grenoble, leads the EVASION research team, a joint INRIA Grenoble and Jean Kuntzmann laboratory team, which is affiliated with the French National Centre for Scientific Research (CNRS) and the universities of Grenoble. The Eurographics award is given for major contributions to the field of computer graphics that are both technologically advanced and creative. This is the case in Marie-Paule Cani's work on implicit surfaces and multiresolution deformable models. For example, she has developed several solutions for animation of nature scenes, and more recently, for animating characters' hair and clothing. Her work in creating interactive forms is also highly innovative. Based on theoretical concepts, her research has been applied in several domains: animated games and films, training simulators for surgical operations, industrial virtual prototyping, etc.

ERCIM Activity Report 2010

The recently published ERCIM annual report 2010 provides the opportunity to review ERCIM's activities and gain insight into ERCIM's structure and organisation, projects, working groups, members, etc. We hope you enjoy reading it and obtain from it ideas or stimulation. Perhaps it will inspire you to participate in ERCIM's activities!

http://www.ercim.eu/publications/annual-report

Jos Baeten
New General Director of CWI

Source: TU/e, Eindhoven

Prof. Dr. Jos C.M. Baeten will be the new general director of CWI in Amsterdam. The General Board of the Netherlands Organisation for Scientific Research (NWO) confirmed his appointment on Monday 6 June. On 1 October 2011 Baeten succeeds the current general director, Prof. Dr. Jan Karel Lenstra, who will continue working at CWI as a senior researcher until his retirement in 2012. Baeten obtained a PhD in mathematics in 1985. Since 1991, he has been affiliated with the Eindhoven University of Technology as Professor of Theoretical Computer Science, and since 2010 as Professor of Systems Engineering. The Board of CWI describes Baeten as a very prominent researcher, with an excellent understanding of both mathematics and computer science, and extensive experience in attracting external funding. He has been successful in contract research and valorization and is a very experienced manager. Jos Baeten will succeed Jan Karel Lenstra as CWI representative for ERCIM.

Twenty-five Years .nl Domain

In April 2011, the 25th birthday of the .nl domain was celebrated. On 25 April 1986,.nl was registered by CWI, the Centrum Wiskunde & Informatica in Amsterdam. It was one of the first country domains in the world. Six days later, Piet Beertema, at that time system administrator, registered the first internet domain in the Netherlands: cwi.nl, which is still in use by CWI.

Beertema was given the authority over the .nl domain and managed it for ten years. In 1996, authority was handed over to Stichting Internet Domeinregistratie Nederland (SIDN), co-founded by CWI. Currently, SIDN has administrated ca. 5.5 million unique domain names of which 4.2 million are active. Not only is this one of the highest numbers worldwide, the Netherlands also ranks second in the number of domains registered per inhabitant. SIDN celebrated 25 years of .nl with the publication of a magazine and an anniversary website, http://de25jaarvan.nl (in Dutch).

The first open internet connection between Europe and the United States originated from CWI as well. On 17 November 1988, a short e-mail correspondence confirmed the connection of CWI to NSFnet, an American academic computer network and the forerunner of today's internet. In these early days, most of the transatlantic internet traffic had to go through CWI.

http://www.godfatherof.nl/

ERCIM
European Research Consortium for Informatics and Mathematics

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

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

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

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

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

CWI
Centrum Wiskunde & Informatica
Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
http://www.cwi.nl/

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

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

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

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

Fraunhofer IUK-TECHNOLOGIE
Fraunhofer ICT Group
Friedrichstr. 60
10117 Berlin, Germany
http://www.iuk.fraunhofer.de/

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

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

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

PEG
PORTUGUESE ERCIM GROUPING
Portuguese ERCIM Grouping
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal

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

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

SpaRCIM
Spanish Research Consortium for Informatics and Mathematics
Spanish Research Consortium for Informatics and Mathematics,
D3301, Facultad de Informática, Universidad Politécnica de Madrid,
Campus de Montegancedo s/n,
28660 Boadilla del Monte, Madrid, Spain,
http://www.sparcim.es/

Swedish Institute of Computer Science
SICS
Swedish Institute of Computer Science
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

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

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

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
http://www.cs.ucy.ac.cy/

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

Order Form

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

Send, fax or email this form to:

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

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

I wish to subscribe to the

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

Name:

Organisation/Company:

Address:

Postal Code:

City:

Country

E-mail:

You can also subscribe to ERCIM News and order back copies by filling out the form at the ERCIM Web site at http://ercim-news.ercim.eu/