Helsinki Institute for Information Technology HIIT Annual Report 2004

Päivi Saarinen, Martti Mäntylä, and Esko Ukkonen (eds.)

Helsinki Institute for Information Technology HIIT
Tammasaarenkatu 3, Helsinki, Finland
PO BOX 9800
FIN-02015 TKK, Finland
Web http://www.hiit.fi
e-mail info@hiit.fi

Table of Contents

1 HIIT in Brief

Helsinki Institute for Information Technology HIIT is a joint research institute of the two leading research universities in Helsinki, the University of Helsinki (UH) and the Helsinki University of Technology (TKK). HIIT was founded in 1999 and has been operational since early 2000.

HIIT conducts internationally high-level strategic research in information technology and related multi-disciplinary topics, especially in areas where the Finnish information and communication technology (ICT) industry has or may reach a significant global role. HIIT works in close co-operation with Finnish universities, research institutes, and industry, aiming to improve the contents, visibility, and impact of Finnish IT research to benefit the competitiveness and progress of the Finnish information society. HIIT also aims at creating a strong network of international partnerships with leading foreign research universities and institutions.

HIIT is directed by a Board consisting of members from the universities, industry, and HIIT personnel. The Industrial Advisory Board manages the HIIT-industry liaison. The Scientific Advisory Board advises the Board on strategic planning of HIIT's research directions.

HIIT consists of two autonomous units. The Advanced Research Unit (ARU) mainly operates through two- to three-year long industry co-funded research projects. The present projects cover mobile and ubiquitous computing, user experience research, complex system computation, community-based digital content, semantic computing, and digital economy. Professor Martti Mäntylä is the Research Director of the ARU.

The Basic Research Unit (BRU) focuses on long-term basic research issues in information technology, including challenges arising in major novel application areas in sciences and engineering. The present research groups work on data analysis, adaptive computing and neuroinformatics. Professor Heikki Mannila was the Research Director of BRU until the end of August 2004. As Professor Mannila received the nomination for an Academy Professor, Professor Esko Ukkonen was appointed Research Director as of 1 September, 2004.

ARU is located in the High Tech Center Helsinki in Ruoholahti, Helsinki. BRU is located at the premises of the Department of Computer Science of UH in Kumpula, Helsinki, and the Department of Computer Science and Engineering of TKK in Otaniemi, Espoo.

2 Review of Year 2004

2.1 Advanced Research Unit (Professor Martti Mäntylä)

The year 2004 was the fifth year of operation for HIIT and its Advanced Research Unit. It was characterized more by qualitative than quantitative growth – even though the number of personnel and the volume of research also continued to grow at a steady rate.

A new phase of internationalization was reached in many respects during 2004. ARU launched two EU funded projects – *ALVIS*, where it acts as the coordinator, and *MobiLife*, a large integrated project that also covers research in the Basic Research Unit. In addition, the *Wireless Festival* Eureka project was launched under the CELTIC initiative. Through these activities, the unit has started to build the practical skills required to play a proactive role in EU- funded research in the future. A sign of this is that the unit also participated in the preparation of several new proposals during the year.

Outside Europe, the long-term relationship that the ARU has built with the University of California at Berkeley also entered a new phase as HIIT received several visitors from UCB during the year, most notably Professor Anthony Joseph who visited ARU for three months from August to November. The more recently launched co-operation with the Tsinghua University also progressed rapidly. HIIT received two visitors from Tsinghua during late 2004, notably Professor Zhou Qiang for the month of November.

In Finland, the unit continued to enjoy a strong position even though the competition for research funding appears to become fiercer every year. Six nationally funded projects – four co-funded by TEKES and industry and two funded by the Academy of Finland – were launched during 2004. The largest of the new activities is the *Infrastructure for Host Identity Protocol* project that aims to make HIIT a global leader in the development of HIP- based technology.

More importantly, the scientific output of the unit also developed favourably both in quantity and quality. HIIT researchers are publishing more articles and papers in better forums than ever. The same may be said about the more practical impact of HIIT-released software. A sign of this are the awards received by HIIT research projects during the year, most notably the Semantic Web Challenge Award 2004 given at ISWC 2004 (Hiroshima, Japan) to the semantic portal MuseumFinland created by the *Intelligent Web Services* project.

HIIT was originally founded by agreement of the UH and the TKK for five years (with one-year extensions) from 1 August 1999 to 31 July 2004. As the original period ran to an end during 2004, the HIIT board launched an extensive strategy process in February 2004, aiming at defining HIIT's overall strategy for the next 5-year period from 1 August 2005. In June 2004, the Board accepted a baseline strategy document focusing in the institutional form and external interface of HIIT. HIIT's mission and the scope of its research were discussed at the meetings of the Industrial Advisory Board in September and the Board in October. HIIT's future was also a major topic of the Scientific Advisory Board meeting in November 2004.

The final decisions were reached in April 2005 when it was decided that the two units of the research centre would be joined from 1 January 2009. During the intervening three years 2006-2008, the research programmes, practical operations, formal position, and funding structure of the two units will be homogenized.

2.2 Basic Research Unit (Professor Esko Ukkonen)

Year 2004 was the third year of operation of the Basic Research Unit (BRU). Esko Ukkonen was appointed as the Research Director of BRU as of 1 September, to follow Heikki Mannila. The research concentrated on three main themes: data analysis, adaptive computing, and neuroinformatics.

In the area of data analysis, the work in theory and applications continued under the leadership of Academy Professor Heikki Mannila. The basic challenge is to develop computationally efficient methods that can be used to obtain useful information from large masses of data. The work is characterized by the interplay of theoretical work on the methods and applied research in collaboration with researchers from other sciences and from industry. In the area of the theory of data mining there were several interesting developments. The idea of ranking based on random walks was extended to databases, creating a good framework for handling the problem of ranking answers to arbitrary queries. The work on different types of segmentations and orderings continued very actively: advances were made both on the theoretical aspects and the applications in paleontology and genetics. The Altti project, lead by Professor Hannu Toivonen, developed haplotyping methods for high-throughput genotyping. The gene mapping algorithm HPM developed in the project was successfully applied to locate a novel asthma gene, and the discovery was published in Science.

In the area of adaptive computing, BRU participates in a major international project MobiLife, an EU Integrated Project with 22 partners and a total budget of about 17,3 million Euro. The HIIT activity within this project is lead by Dr Patrik Floréen. The Proactive Computing (PROACT) research programme of the Academy of Finland was continued under the coordination of BRU (Prof. Mannila, Dr Linden). One of the projects funded by this programme within BRU, the Context project (Professor Hannu Toivonen), released ContextPhone software, the first open source context-aware research and prototyping platform for off-the-shelf mobile phones. ContextPhone has rapidly attracted a great deal of interest and is now used by research groups at MIT, Berkeley, and VTT, in addition to HIIT.

Neuroinformatics is broadly defined as the intersection of information technology and neuroscience. The leader of the neuroinformatics group, Dr Aapo Hyvärinen, joined BRU in May 2003, and 2004 was the first year of full operation. Two post-doctoral fellows, Patrik Hoyer and Jarmo Hurri, joined the group, as well as two new PhD students. A new interdisciplinary Academy of Finland project 'Advanced data analysis in vision research: feature interaction and nonlinear classification of images' was started in collaboration with the Department of Psychology. New multivariate data analysis methods were developed in collaboration with a long-term visitor, Shohei Shimizu, from Osaka University.

In the funding arena, the growth continued. The external funding of BRU in 2004 was about 1.1 million Euros, the largest share coming from the Academy of Finland. Heikki Mannila was instated as Academy Professor in August 2004.

The 2004 meeting of the HIIT Scientific Advisory Board (SAB) again produced a great deal of useful feedback to both units of HIIT. Overall, the SAB was quite satisfied with the focus and results of BRU.

International recruiting has been one of the most important goals of BRU. Dr Bart Goethals and Dr Aristides Gionis continued their work in BRU. New recruitments in 2004 were Dr Alexander Hinneburg (Halle) and Dr Panayiotis Tsaparas (Toronto). The postdoctoral researchers from abroad contribute considerably to the research environment at BRU, and the hiring from elsewhere will continue in the future.

BRU operates at the Department of Computer Science at UH and at the Laboratory of Computer and Information Science at TKK, the major part of the activities taking place at UH. In August 2004, the UH subsection of BRU moved with the Department of Computer Science to the new Exactum building at the Kumpula campus. The operation at two sites works well, and the different backgrounds of the departments provide fertile ground for innovative research.

One of the basic principles of BRU is active participation in teaching. In 2004, the links to teaching were kept active at both universities.

3   Important Dates

21–23 January: The kick-off meeting of the project ALVIS: *Superpeer Semantic Search Engine* in Paris. HIIT coordinates the joint research effort involving nine partners from Europe and also Tsinghua University in Beijing, China.

29 January: Guest lecture by Professor Dan Geiger (Superlink: A new Genetic Linkage Analysis Software).

11 February: The Board of HIIT launched a strategy process, aiming at a revised research agenda of HIIT with related institutional changes. The objective was to draft a new agreement to enter into force as of 1.8.2005.

17 March: Guest lecture by Dr Wakako Nakamura (Analysis of natural images by independent quadratic forms and temporally coherent quadratic forms).

26-29 April: HeCSE course, lecturer Professor Paul Vitanyi (An Introduction to Kolmogorov complexity and Its Applications).

6 May: Guest lecture by Post-Doctorate Fellow Themis Palpanas (Online Amnesic Approximations of Streaming Time Series).

24 May: Creative Commons and HIIT's DE Core project published the official Finnish translations of Creative Commons licenses, a set of freely available content licenses aiming at expanding the range of creative work available for others to build upon and share. The initiator of Creative Commons, Stanford Law School Professor Lawrence Lessig gave a presentation in the publication ceremony held at HIIT.

31 May – 2 June: HIIT organised a Retreat meeting under the theme "HIIT Retreat: Mobile Computing @ HIIT" at Sjökulla Education Center, Kirkkonummi, for representatives from industry and for HIIT senior researchers.

4 June: Guest lecture by Professor Dr Hans Burkhardt (Invariants in Pattern Recognition - New Results for a Classical Problem).

8 June: President of the Republic of Finland Tarja Halonen visited ARU.

21 – 24 June: The fourth Helsinki-Berkeley PhD student Workshop on Telecommunication Software Architecture was held at University of California at Berkeley.

19-21 July: Relocating to Exactum. In summer 2004 the Department of Computer Science (including Helsinki Institute for Information Technology HIIT, Basic Research Unit) moved to the new Kumpula Campus area.

23 August: HIIT's Mobile Content Communities (MC2) project organised a Mobile Media Workshop at HTC, Helsinki with Professor Marc Davis from University of California, Berkeley. The workshop explored the growing possibilities of mobile technology as a new infrastructure for creating, sharing and finding content describing metadata for mobile multimedia.

1 September: Professor Esko Ukkonen was appointed the Director of HIIT Basic Research Unit (BRU). The former director of BRU, Professor Heikki Mannila, received the nomination for an Academy Professor.

2 September: Guest lecture by Heikki Hyyrö ,PhD, (Bit-parallel techniques in computing longest common subsequences).

2-3 September: HIIT was one of the co-organisers of the Web Intelligence Symposium held at the Finnish Artificial Intelligence Conference at Heureka Science Centre, Vantaa, Finland. The symposium was organised by the Finnish AI Society, World Wide Web Consortium W3C, XML Finland Association, Tekes FENIX-programme, University of Helsinki, Finnish Science Centre Heureka and HIIT.

17 September: Inauguration of the Exactum Building in Kumpula. The new BRU offices are located in the building.

29 September: The Industrial Advisory Board meeting.

7 October: Guest lecture by PhD student Debora Donato (On the Bowtie Structure of the Web).

12 October: The HIIT Advanced Research Unit Portfolio Seminar.

11 November: HIIT organized the OPENMIND 2004 seminar in Tampere with the Centre for Open source software and Fenix Technology programme. The seminar concentrated on challenges and solutions in open source software and open content production.

11 November: The DiMaS system won the title "The Best e-Business Application in Finland" and the nomination to represent Finland in World Summit Award 2005.

23 November: The Prime Minister of Finland rewarded MuseumFinland as the most innovative application in the national Quality on the web competition (Laatua verkkoon).

15-17 November: The HIIT Scientific Advisory Board met for the second time in Helsinki.

9 December: Guest lecture by Professor Dr Gustavo Deco (A Computational Neuroscience Approach to Visual Cognition, Attention, and Working Memory).

4 Research

4.1 Advanced Research Unit Activities

The mission of ARU is to conduct strategic research in close co-operation with leading ICT companies, aiming at a significant impact on the future progress of ICT technologies and applications and the progress of the information society. To achieve this mission, the research in ARU focuses on four thematic areas, each including one or several research groups that also may participate in several areas. The research areas are: Future Mobile and Ubiquitous Computing (led by Professor Kimmo Raatikainen), Intelligent Systems (Professor Henry Tirri), Media Convergence (Professor Petri Vuorimaa, Professor Eero Hyvönen, and Dr Marko Turpeinen), and Digital Economy (Professor Jukka Kemppinen).

The research groups and their senior researchers involved in ARU's work during 2004 were as follows: Mobile Computing (Professor Kimmo Raatikainen, Dr Ken Rimey, Dr Pekka Nikander, Dr Jan Lindström, Dr Andrei Gurtov); User Experience Research (Dr Timo Saari, Dr Giulio Jacucci, Professor Martti Mäntylä); Complex System Computation (Professor Henry Tirri, Professor Petri Myllymäki, Dr Wray Buntine, Dr Jorma Rissanen); Digital Content Communities (Dr Marko Turpeinen, Dr Giulio Jacucci, Dr Timo Saari); Semantic Computing (Professor Eero Hyvönen); and Digital Economy (Professor Jukka Kemppinen, Dr Pekka Nikander, Dr Pekka Himanen, Dr Markku Stenborg, Professor Martti Mäntylä).

The following sections provide a more detailed account of the research activities in each research area.

4.1.1 Future Mobile and Ubiquitous Computing (Fuego)

The 1990s were marked by two simultaneous significant trends that changed the role of information and communications technology in the everyday lives of most people living in the developed part of the world: the Internet and mobile telephone. In the coming years these two explosive technologies are expected to merge in the Mobile Internet, fulfilling the vision of ubiquitous computing and communications providing access to digital services any time and anywhere. With this, computing seems to be destined to invade and inhabit, for better or worse, every nook and cranny of our environment and everyday life: home, office, car, school, library, sports facility, shopping mall, etc.

This vision poses immense challenges to computing research. What will the end users use their Mobile Internet devices for? Which new services can be created on the basis of technologies such as positioning, context sensitivity, and adaptive multimodal interaction? What kinds of technical infrastructures and platforms are needed?

The Future Mobile and Ubiquitous Computing (Fuego) research area at HIIT covers a selection of research topics within this general framework. Work in the research area aims to combine two major viewpoints to the Mobile Internet: the technology view and the user view (while the equally important economical and societal views are delegated to the Digital Economy research area discussed in its own section below).

The technology-oriented research line, conducted in the Mobile Computing Group led by Professor Kimmo Raatikainen, builds on expertise in areas such as Internet protocols, middleware, peer-to-peer computing, and operating systems to study enabling technologies, infrastructures, and platforms for the Mobile Internet. The user-oriented research line, conducted in the User Experience Research Group co-led by Professor Martti Mäntylä and

Dr Timo Saari, builds on a combined expertise in cognitive science, psychology, media science, and computing to study end users and their needs, novel application concepts, and the resulting user experience. The two lines of work interact in creating joint prototypes, demonstrations, and field tests.

4.1.1.1 Fuego Core: Middleware for Mobile Wireless Internet

Project leader: Raatikainen, Kimmo
Research group(s): Mobile Computing
Researchers: Tarkoma, Sasu; Kangasharju, Jaakko; Lindholm, Tancred; Saaresto, Marko; Kousa, Mika; Komu, Miika; Mäkelä, Mikko; Slavov, Kristian; Thalainayar B, Ramya
Schedule: 2002-02-01 … 2004-12-31
Cooperation units: Department of Computer Science (UH); UC Berkeley, USA
Funding: Tekes; Elisa Communications; TeliaSonera Finland; Nokia Research Center; Ericsson; Movial
Keywords: Mobile wireless Internet, middleware services, mobile computing, adaptive applications
www-page and publications: see http://www.hiit.fi/fuego/fc/index.html

The objective of this research project is to specify the set of fundamental enabling middleware services for mobile applications on future mobile environments and to implement two research prototypes. The project has three work areas, which describe the content of the research. The work areas are Adaptive Applications, Dynamic Reconfigurable Services, and Mobile Distributed Information Base.

The specification of the middleware service set for applications in future Mobile Wireless Internet is experimented and evaluated using two prototype implementations. The implemented prototypes are used to show how services can be deployed on the proposed middleware services. The project will be executed in close co-operation with the research group of Professor Randy Katz at the University of California at Berkeley (UCB). In addition, the project monitors and contributes to relevant middleware standardization bodies; for example IETF, OMG, W3C, and OntoWeb.

During 2004 the project implemented the second set of middleware services and experimented with the prototype implementation. The second middleware service set featured and enhanced the following key enablers for applications: distributed events, wireless SOAP messaging, a synchronizing XML-aware file system, a presence service, and the Host Identity Protocol (HIP). The HIP development was started in the Fuego Core project and completed in the HIPinfra project.

4.1.1.2 Personal Distributed Information Store (PDIS)

Project leader: Rimey, Kenneth
Research group(s): Mobile Computing
Researchers: Oksanen, Kenneth; Rüger, Torsten; Kanerva, Pekka; Hasu, Tero; Päivärinta, Juha
Schedule: 2003-01-01 ... 2004-12-31
Funding: Tekes; Nokia; Movial; Innofactor; Fathammer (2004); Hewlett-Packard (2003)
Keywords: Mobile data management, data synchronization, mobile computing, peer-to-peer computing
www-page and publications: see http://pdis.hiit.fi/

The PDIS project at HIIT ran through 2003 and 2004 and was sponsored by Tekes and a consortium of industrial sponsors. It focused on next-generation data synchronization, or in plain terms, on how to enable people to store replicas of their data on several electronic devices, and several computers, and to keep these replicas in sync.

The project has developed the PDIS repository, an update-anywhere replicated XML database that runs on Series 60 mobile phones as well as personal computers. It allows any device to be synchronized with any other at any time and is ideal for storing personal databases (contacts, calendars, etc.) and metadata for digital media files. The software is written in Python and is available at http://pdis.hiit.fi/.

We began work in 2003 on a PC-based implementation of the design, for which we developed two different calendaring test clients in order to validate it with a realistic application. A first internal release was demonstrated for the project steering group in October 2003 and made available to the partner companies in November 2003.

In 2004, the project rewrote much of this, added conflict resolution and sync management, and demonstrated the repository, a to-do-list application, and a new top-level *PDIS Manager* application running on Series 60 phones with GPRS and Bluetooth. We also demonstrated the use of PDIS to store photo metadata in a web-based photo archive browser, as well as the use of PDIS in controlling file synchronization.

A separate *PDIS on Symbian* project was funded by Nokia during 2004 to port the PDIS repository to the Symbian operating system. We chose to port the PDIS repository prototype to the phone as directly as possible using Nokia's *Python for Series 60*. This project was one of the most ambitious early-access test users of Python for Series 60, which Nokia released in December 2004.

4.1.1.3 Context Recognition by User Situation Data Analysis (CONTEXT)

Project leader: Mäntylä, Martti
Research group(s): User Experience
Researchers: Tiitta, Sauli; Oulasvirta, Antti
Schedule: 2002-11-01 ... 2005-12-31
Cooperation units: BRU/HIIT
Funding: Academy of Finland
Keywords: Proactive computing, ubiquitous computing, data analysis, data mining, user-centered design, Smartphone, human-human interaction
Research programme: Academy of Finland / Proactive Computing (PROACT)
www-page and publications: see http://www.hiit.fi/fuego/context/

The Context project studies characterization and analysis of information about user context and its use in proactive adaptivity. In mobile and ubiquitous applications and systems, reacting to user context is a key component of proactivity: changes in the user's situation are rapid and they are strongly reflected in the user's needs and preferences.

The project focuses on the utilization of user context: how does the context reflect the user's motivations, how to make automatic inferences about the contexts, and how to characterize contexts to users and design user interaction about contexts? These questions are considered in the framework of an example application: mobile communication, a representative ubiquitous application whose usability greatly depends on how context-sensitively communication decisions are managed. The project has adopted a multidisciplinary approach where the research problems are approached by qualitative user studies, data analysis algorithm development, and empirical testing in a prototype

environment. The key results are (1) methods for utilizing qualitative user situation descriptions in the development of context-sensitive applications, (2) algorithms for context analysis and characterization, and (3) models for user interaction about context. The project has produced a prototype of a context-sensitive mobile communication application.

The project started in November 2002 with qualitative user studies aimed at revealing how users interpret context descriptions and generally context structures. A humanistic research strategy has been deployed to develop a working prototype of proactive context-communication, called ContextPhone, running on a Nokia Series 60 mobile phone. At the same time, ethnomethodological studies by the research group revealed how social and temporal contexts affect interaction and how contexts could be modelled in proactive computers. The prototype has been subjected to a series of longitudinal field studies in 2004 to examine its effects on group communication behavior.

4.1.1.4 Dynamic Composition and Sharing of Context-aware Mobile Services – DYNAMOS

Project leader: Raatikainen Kimmo
Research group(s): Mobile Computing
Researchers: Riva, Oriana
Schedule: 2004-06-01 ... 2005-05-31
Cooperation units: VTT Technical Research Centre of Finland
Funding: Tekes; ICT Turku; Suunto; TeliaSonera Finland; VTT
Keywords: mobile
Research programme: Tekes/ Fenix
www-page and publications: see http://www.vtt.fi/tte/proj/dynamos/

The main objective of the Dynamos project is to design and evaluate a framework for providing context-aware services for mobile users, with the focus on proactively notifying the users about potentially interesting services. In addition, the project investigates issues related to storing and managing service descriptions and how users may share these descriptions with each other.

The first system prototype has been developed. The selected use case scenario targets the needs of a marine community of recreational boaters. The requirements study was carried out by means of boater interviews conducted in October-December 2004. The conceptual architecture of the DYNAMOS system has been designed, and implementation for a proof-of-concept demonstration (due by March 2005) has been started.

The system prototype is planned to support:

- provisioning of relevant services based on matching between service descriptions and user activity and location. Specifically, in the creation of the personal profile, the user can associate a set of interests to each activity. These interests are matched against service categories. The activity is manually entered by the user while the location is acquired by means of GPS;
- the possibility of creating, storing, and sharing service annotations. Service annotations consist of textual descriptions, location and time of creation, author signature, numerical rating, and a link to the original service description; and
- the possibility of creating, storing and sharing user notes. User notes consist of text, location and time of creation, and author information.

4.1.1.5 DRAMA – Scenario Methods for User Centered Product Concept Design

Project leader: Mäntylä, Martti
Research group(s): User Experience
Researchers: Tiitta, Sauli; Kankainen, Tomi; Kantola, Vesa; Mehto, Kati; Sädekallio, Outi; Pennanen, Merja
Schedule: 2004-01-01 ... 2007-12-31
Cooperation units: Helsinki Polytechnic Stadia; University of Art and Design Helsinki
Funding: Academy of Finland
Keywords: User-centric product concept design, Devised theatre, Forum theatre, Narrative theatre, Scenario-based design
Research programme: SA/ Industrial Design Programme
www-page and publications: see http://www.hiit.fi/fuego/drama/

DRAMA - Scenario Methods for User-Centered Product Concept Design (UCPCD) is a multidisciplinary research project, in which UCPCD's scenario-based working methodology and process are reinforced with methods of devised theatre. The ultimate goal is to make the human voice loud and clear in the new product development to create humane products. This is achieved by improving current working methods, tools, and processes.

In DRAMA we study possibilities to use devised theatre methods with UCPCD to create a new, rich and more profound approach to user centricity and to create dramatic scenarios, a new form of scenarios. In DRAMA this is done by comparing several devised theatre and UCPCD methods and defining the relevant application areas for each one.

The following methodological areas of UCPCD will be studied. 1. User research: user research methods, collecting user narratives and other dramatic compositions, and notation of user narratives. 2. Analysis of user data: methods for analyzing user narratives and interpreting user needs. 3. Product or service concept design: techniques for utilizing user narratives as a basis for design, writing the manuscript of a scenario on the basis of user narratives, the form and output of scenarios. 4. Concept evaluation: evaluation of concepts with users using devised theatre and dramatic scenarios.

The project was launched on January 1, 2004. As the first step, we defined the research contexts and user groups together with the National Consumer Research Center. Taking a wide view of relevant societal issues, we decided to focus holistically in professional people of the age 55±5 years, a period of life where people often begin to be concerned with the nearing retirement age.

During the rest of 2004 the project performed two rounds of experimental drama production (narrative theatre and forum theatre) aimed at specific user groups (nurses' teachers; church workers; policemen) and also studied the same groups with traditional methods of user research. Through this, we have collected an extensive set of qualitative user data that allows us to compare qualitatively the data obtainable by these two parallel methods. During 2005, these data will be carefully analysed and various methods of dramatic scenario generation explored.

4.1.1.6 MobiLife – Mobile Life

Project leader: Mäntylä, Martti
Research group(s): User Experience, Digital Economy
Researchers: Kurvinen, Esko; Salovaara, Antti; Mäntylä, Teemu; Pitkänen, Olli; Ylitalo, Katri
Schedule: 2004-09-01 ... 2006-12-31
Cooperation units: Nokia Corporation; Alcatel-CIT; Ericsson AB; Oy LM Ericsson Ab; Hewlett-Packard Italiana srl; Motorola Ltd.; Motorola SAS; NEC Europe Ltd.; Siemes Mobile Communications SpA; DoCoMo Communications Laboratories Europe GmbH; Elisa Corporation; Fraunhofer Gesellschaft zur Foerderung der angewandten Forschung e.V.; University of Helsinki; Universität Kassel, Fachgebiet Kommunikationstechnik; The University of Surrey; NEOS Engineering SRL; Stichting Telematica Instituut; UNIS, spol s.r.o.; Suunto Oy; BellStream SP.z.o.o.; Telecom Italia
Funding: European Union
Keywords: application, service, mobile, wireless, user-centricity
Research programme: Information Society Technologies (EU FP6), FP6-2003-IST-2 (IST-511607 MobiLife)
www-page and publications: see https://www.ist-mobilife.org/

People are used to being able to contact anyone, anywhere, at any time. However, the challenge of enabling mass-market-scale ubiquitous services and applications remains. The strategic goal of MobiLife is to bring advances in mobile applications and services within the reach of users in their everyday life by innovating and deploying new applications and services based on the evolving capabilities of 3G systems and beyond.

Future environments affected by the ICT convergence give new possibilities, but also new challenges due to increasing heterogeneity, user needs and expectations. The research challenge of MobiLife is to address problematics related to different end-user devices, available networks, interaction modes, applications and services.

MobiLife will have both technological and societal global impact by addressing the technical viability and user-acceptance of solutions that belong to three main Focus Areas: self-awareness provides support for automatic configuration arrangement of devices, services, and local connectivity in the user's local environment. It also enables automatic and multi-modal interfaces that enhance the user experience and minimize the active user effort needed in man aging the local environment.

Group-awareness comprises context and presence support enabling individuals to relate to, share, and interact with each other and common artefacts. Also novel privacy and trust models are addressed, which are mandatory for users to be able to rely on and use these new services and applications.

World-awareness offers automatic support for seamless access to and delivery of services across different domains that enable individuals to retain the accustomed use of their services, as they move between different environments and infrastructures.

The MobiLife consortium consists of application owners, manufacturers, operators, solution providers and academia. MobiLife is part of the Wireless World Initiative, which comprises several projects for IST.

The work of the ARU focuses on Work Package 1 of the project, generally responsible for providing a sound basis for user centricity of the entire project. Specifically, ARU maintains the iterative user-centric approach and coordinates user research activities within MobiLife. It has also contributed extensively in the generation, analysis, and validation of the driving

scenarios of the project. Currently, MobiLife is conducting a second round of user research. A set of eleven application mock-ups are being evaluated with Finnish and Italian families. In addition, ARU also contributes to analyzing the legal issues related to the scenarios and the privacy/trust issues.

The project was launched on 1.9.2004.

4.1.1.7 Infrastructure for HIP –InfraHIP

Project leader: Mäntylä, Martti
Research group(s): Mobile Computing
Researchers: Gurtov, Andrei; Kousa, Mika; Komu, Miika; Koponen, Teemu; Vehmersalo, Essi; Lindqvist, Janne; Karlsson, Niklas
Schedule: 2004-08-01 ... 2005-07-31
Cooperation units: Laboratory of Telecommunications Software and Multimedia (TKK)
Funding: Tekes, Ericsson, Nokia, Finnish Defence Forces, Elisa
Keywords: Internet, architecture, security, mobility, communication protocols
www-page and publications: see http://infrahip.hiit.fi/

The Host Identity Protocol (HIP) and the related architecture form a proposal to change the TCP/IP stack to better support mobility and multi-homing. Additionally, they provide for enhanced security and privacy and advanced network concepts, such as moving networks and mobile ad hoc networks. The InfraHIP project studies application-related aspects of HIP, including APIs, rendezvous service, operating system security, multiple end-points within a single host, process migration, and issues related to enterprise-level solutions. Through this, the project maintains HIIT (and thereby Finland) as one of the leading research centres doing HIP related work. "Infra" in the project name stands for Infrastructure. As the basic HIP protocol is almost ready, the project focuses on developing the missing infrastructure pieces such as DNS, NAT, and firewall support to enable a widespread deployment of HIP.

The project work started on 1 August 2004 with a small focused effort on implementation and measurements of the Host Identity Indirection Infrastructure (Hi3). In addition, the HIP for Linux implementation work and a Master's thesis on Native HIP API have proceeded in cooperation with the Fuego Core project. The project funding decision arrived in November and the rest of the year was spent building the full project team.

4.1.2 Intelligent Systems

Computer science is the science of studying how things can be automated. Consequently computational studies of intelligent systems also focus on automating intelligent behaviour, a quest where modelling plays a central role in formalizing the properties of such systems: learning, inference and intervention (actions). Due to the uncertainty and incompleteness of available information in application domains of computer science (AI in particular) such models are commonly based on probabilities. The aim of our research is fundamental understanding and development of computationally efficient probabilistic and information-theoretic modelling techniques, and their multi-disciplinary applications from engineering to sciences.

The work has a strong basic research component, being at the intersection of computer science, information theory and mathematical statistics. The results of this methodological work are applied both in science and industrial applications resulting in advanced prototypes and fully fielded applications. The recent main applied research areas include PetaByte level scalable next generation information search, future networked learning environments,

location-aware personalized services and multi-disciplinary applications of modelling in social sciences, medicine and biology.

4.1.2.1 Proactive Information Retrieval by Adaptive Models of Users' Attention and Interest (PRIMA)

Project leader: Myllymäki, Petri
Research group(s): Complex System Computation (CoSCo)
Researchers: Tuulos, Ville; Miettinen, Miikka; Wettig, Hannes
Schedule: 2003-01-01 … 2005-12-31
Cooperation units: Neural Networks Research Centre (TKK), CKIR (Helsinki School of Economics)
Funding: Academy of Finland
Keywords: Proactivity, probabilistic
Research programme: Academy of Finland / PROACT
www-page and publications: see http://www.cis.hut.fi/projects/mi/prima

Successful proactivity, i.e. anticipation, in varying contexts requires generalization from past experience. Generalization, on its part, requires suitable powerful (stochastic) models and a collection of data about relevant past history to learn the models. Our goal is to build probabilistic models that learn from the actions of people to model their intentions and expectations, and use the models for disambiguating the users' vague commands and to anticipate their actions. The actions and interests are monitored by measuring eye fixations and movements that exhibit both voluntary and involuntary signs of both the cognitive state of the user and his intentions. In addition we have investigated the possibility of using real-world Web usage patterns to signal relevance.

During the second year of the project we have concentrated on taking the previous experiences to a practical level. We have focused on web usage and reading patterns and their utilization to aid proactive information retrieval. This line of research has benefitted from synergy with our ongoing efforts to build content-based search engines.

4.1.2.2 Scalable Probabilistic Methods for Next Generation Internet Search Engines (PROSE)

Project leader: Tirri, Henry
Research group(s): Complex System Computation (CoSCo)
Researchers: Buntine, Wray; Perttu, Sami; Löfström, Jaakko; Silander, Tomi; Tuominen Antti
Schedule: 2003-01-01… 2006-12-31
Funding: Academy of Finland
Keywords: Internet search, concept maps, probabilistic modeling, genre, multinomial Principal Components Analysis
www-page and publications: see http://cosco.hiit.fi/search/prose.html

It is evident that with hundreds of millions of pages of information on the Internet, search has become a fundamental service. The abundance of available information sets new challenges for even the best current search engines, and what is needed is qualitatively better ways to answer user queries. The context for our research is the development of a kernel for supporting a subject-specific node in a distributed, hierarchical system for supporting navigation and search on Internet pages. The node may have tens of millions of pages, and needs to automatically build its own hierarchies for topic, genre, and terminology - aspects of

the document set that we call a concept map. The objective of the project is to provide the statistical computing techniques and their implementations needed to build a search engine kernel for the next generation Internet search services. The topics studied focus on developing statistical modelling techniques such as the multinomial Principal Component Analysis (mPCA), and addresses both the theoretical development and the applied aspects for very large (giga and terabyte) document data sets.

Development of the Multinomial PCA code base and matching theory progressed significantly during the year. Testing of the system as a query engine on newswire data lead to a better understanding of how to use the system as a basis for the language modelling approach to information retrieval. Initial experiments looked promising. A basic strategy was developed for scaling the system, and implementation of that strategy was started.

4.1.2.3 Minimum Description Length Modelling in Computer Science and Statistics (MINOS)

Project leader: Tirri, Henry
Research group(s): Complex System Computation (CoSCo)
Researchers: Rissanen, Jorma; Myllymäki, Petri; Roos, Teemu; Kontkanen, Petri
Schedule: 2002-01-01 ... 2005-12-31
Funding: Academy of Finland
Keywords: Minimum description length principle (MDL), stochastic complexity, universal modeling, model selection, predictive inference
www-page and publications: see http://cosco.hiit.fi/Projects/MINOS/

The objective of this research is to develop and study the Minimum Description Length (MDL) approach to modelling and its relationship to other probabilistic approaches used in computer science and statistics. Formally MDL modelling is approached from various directions including Fisher's Maximum Likelihood Principle and minimax formulations leading to universal coding which extends Shannon's Source Coding Theorem. In particular we will focus on the recent Normalized Maximum Likelihood formulation of MDL, as well as in the predictive form (PMDL), properties of which are not yet well understood. The emphasis on applying the theoretical work will be in computationally efficient model selection and prediction problems in computer science.

In 2004 the research continued on studying computational issues related to the Normalized Maximum Likelihood (NML) formulation of the MDL principle, and the practical applications of the results. One of the key applications studied was denoising of images with wavelet models.

4.1.2.4 Search-In-a-Box (SIB)

Project leader: Myllymäki, Petri
Research group(s): Complex System Computation (CoSCo)
Researchers: Buntine, Wray; Tuominen, Antti; Tuulos, Ville; Löfström, Jaakko; Perkiö, Jukka; Porochine, Vladimir
Schedule: 2003-03-01 ... 2006-06-30
Cooperation units: Department of Computer Sciences (UTA)
Funding: Tekes; Novosat; Alma Media Interactive; M-Brain
Keywords: Open source, search engines, modeling
Research programme: Tekes / FENIX
www-page and publications: see http://cosco.hiit.fi/search/sib.html

The project began with basic research and experimentation with our content models and search infrastructure. Already during the early phases of the project we built several demonstrations to give us and the partners first-hand experience of the behavior of the models. The demonstrations were built using both realistic web data and data sets provided by the partners, sizes varying from 20,000 to 800,000 documents.

Based on experiences gained from the first demonstrations, we continued improving and packaging our main deliverables: the MPCA content model, which is already distributed publicly, and Ydin search back-end. These continue to be in the hard core of Search-In-a-Box, providing the needed infrastructure for demonstrations. Together these packages contain about 95,000 lines of code and currently form the basis of a demonstration using Wikipedia.

In the latter part of the year we decided to crystallize the main lessons learned in a large scale, full-fledged Internet search engine, providing content-based search for the whole .FI domain, consisting of about 5 million documents. An effort of this scale requires strong emphasis on distribution of computational load, which was achieved by streamlining the needed components to small, independent modules. We took the opportunity to experiment with a novel content-based ranking scheme of ours, which should be especially robust in noisy enviroments like the Web. Together the new ranking scheme and the modular back-end form the basis for our Aino Search Engine. In addition this effort gave birth to an efficient web crawler, HooWWWer. Together Aino and HooWWWer power a publicly available, regularly updated, search engine for the Finnish Web at http://aino.hiit.fi.

4.1.2.5 ALVIS – Superpeer Semantic Search Engine

Project leader: Myllymäki, Petri
Research group(s): Complex System Computation (CoSCo)
Researchers: Buntine, Wray; Valtonen, Kimmo; Silander, Tomi
Administrative Manager: Kontiainen, Mikko
Schedule: 2004-01-01 ... 2006-12-31
Cooperation units: Unite Mathematique, Informatique et Genome, Institut National de la; Recherche Agronomique (INRA-CRJJ); Ecole Polytechnique Federale de Lausanne, Distributed Information Systems Lab (EPFL); Lund University, Department of Information Technology (ULUND); Technical University of Denmark, Center of Knowledge Technology (DTU); Index Data Aps (Index Data); Exalead SA (Exalead); Universite Paris-Nord, Laboratoire d'Informatique (Paris 13); ALMA Bioinformatica, S.L. (AB); Jozef Stefan Institute (JSI), Department of Intelligent Systems and Department of Knowledge Technologies; Tsinghua University (TU), Department of Computer Science and Technology
Funding: European Union
Keywords: semantic based search, machine learning, peer-to-peer, probabilistic models
Research programme: EU FP6-IST-1
www-page and publications: see http://cosco.hiit.fi/search/alvis.html

The ALVIS objectives are: to provide a powerful, free, stand-alone semantic-based search system so that application-domain experts can readily build topic-specific search sites without needing to become information retrieval experts or computer systems gurus; and further to develop complementary distributed components, together with bridges to existing topic-specific search sites, so that the individual sites can be linked up to form a search network. The semantic-based search engine is intended to automatically build and maintain its own semantic structure with named entities, topics and so forth, and to input primitive ontologies. It is not a Semantic Web engine, and does not rely on the existence of Semantic Web ontologies or build its own ontologies. The semantic structure is created semi-automatically using statistical and machine learning methods for the purpose of returning

better search results. The distributed system is intended to be able to operate with heterogeneous search servers, using query topics as a routing mechanism, and using distributed methods for ranking and semantic-based processing.

HIIT coordinates the project, develops the central relevance component that supports ranking of documents retrieved by a query, for instance, and plays a major role in the integration and testing stages. In 2004 the project began, and major design decisions were made. HIIT co-developed a document-processing pipeline, integrated some core technology into the pipeline, and developed a demonstration of the relevance technology using the Wikipedia as content.

4.1.3 Media Convergence

The Media Convergence research area is intended to cover the entire production, distribution, and use process of future digital media services from content production to service delivery through various channels such as mobile terminals, digital TV, or the Internet. Its core research topics include context creation, content management (archival, indexing, structuring, semantics), service management, and content delivery (content adaptation, XML technologies).

The activities of the Digital Content Communities (DCC) research group, formed in 2003, constitute the major research thrust in the Media Convergence area. The group studies the social activities in digital media and computer-mediated communication. The emphasis is on understanding the "media consumer as a producer", i.e. studying the evolving relationship between professionally created and user-created media content.

The current research focuses on societal computing, i.e. information systems that enable and support social creativity, participatory media and distributed problem solving. However, to develop successful new technologies, and bear the responsibility for design decisions, we as developers should understand and anticipate the dynamics of technology-society interaction. This requires multi-disciplinary end-to-end research from technological platforms to various viewpoints to their impact on the use environment.

Another major line of research within this area deals with machine-processable semantics, i.e., representing data and knowledge in such a way that machines can "understand" its meaning, and developing algorithmic methods for creating intelligent applications based on such representations. The Semantic Computing research group carries out this work.

During 2004, discussions between the Media Convergence and Digital Economy researchers took place, with the intention of merging the increasingly parallel lines of research in a coherent whole. The decision was reached that the two research areas will be merged from 2005 onwards as the Network Society research programme.

4.1.3.1 Mobile Content Communities (MC2)

Project leader: Turpeinen, Marko
Research group(s): Digital Content Communities, User Experience
Researchers: Herrera, Fernando; Kuikkaniemi, Kai; Rantanen, Matti; Saari, Timo; Salovaara, Antti; Sarvas, Risto; Vuorenmaa, Janne; Hietanen, Herkko; Lehdonvirta, Vili
Schedule: 2003-06-01 ... 2005-05-31
Cooperation units: Software Business and Engineering Institute (SoberIT) (TKK); UC Berkeley, USA; MIT Media Lab, USA
Funding: Tekes; Alma Media; TeliaSonera Finland; Veikkaus; Nokia; Starcut; Sulake Labs; Accenture; Futurice; Digital Chocolate
Keywords: Mobility, community, digital media content, social user experience
Research programme: Tekes / FENIX
www-page and publications: see http://pong.hiit.fi/

The Mobile Content Communities (MC2) project, which started in June 2003, studies the social meaning and impact of new communication technology for communities that are interested in mobile gaming. The results expected of the MC2 project include evaluated and tested scenarios of mobile community gaming, template-based design tools that allow people to create their own games and game-related content, new open source tools to empower the community activity, and company-specific case studies to help the industry partners to benefit from community-created content.

We also study how the integration of a camera into a mobile phone affects people and their photographing. The first phase of camera-phone research included a mobile picture annotation system named *Mobile Media Metadata (MMM)* that was built and tested. The MMM system demonstrates a novel approach in combining contextual information, content-based image retrieval, and human-computer interaction to create media content metadata. More recent work related to camera phones is focused on what kind of pictures people take and with whom they share them. The approach has been to study literature and people's picture-taking habits in traditional photography as well as digital photography. Based on these studies, we constructed a mobile picture-sharing application *MobShare*. The system is currently going through user tests where we try to find out how the system was adopted by the users as well as how it fits into the users' picture-taking and sharing habits.

4.1.3.2 Rich Semantic media for personal and professional users –RISE

Project leader: Turpeinen Marko
Research group(s): Digital Content Communities
Researchers: Reti, Tommo, Sarvas, Risto
Schedule: 2004-08-01 ... 2005-07-31
Cooperation units: Technical Research Centre of Finland (VTT), Information Technology
Funding: Tekes; Alma Media; Sanoma WSOY; VTT; YLE
Keywords: semantic metadata, media content template, licensing
Research programme: Tekes / FENIX
www-page and publications: see http://www.vtt.fi/tte/tte41/rise/

The project studies the possibilities in using rich semantic descriptions to combine professional-quality media content with user-created media content. In particular, the project studies media captured with mobile devices, semi-automatic template-based methods to compose media content and in creating new metadata for media, as well as different technical options for implementation. The project also studies the legal issues related to combining personal and professional media content.

4.1.3.3 Intelligent Web Services

Project leader: Hyvönen, Eero
Research group(s): Semantic Computing
Researchers: Apiola, Mikko; Lindgren, Petri; Mäkelä, Eetu; Saarela, Samppa; Salminen, Mirva; Sidoroff, Teemu; Viljanen, Kim;
Schedule: 2003-08-01 ... 2005-07-31
Cooperation units: Department of Computer Science (UH)
Funding: Tekes; Fonecta; TeliaSonera Finland; TietoEnator; The Ministry of Finance; Stakes
Keywords: Semantic web, web services
Research programme: Tekes / FENIX
www-page and publications: see http://www.cs.helsinki.fi/group/iwebs/

The Intelligent Web Services (IWebS) research project studies the possibilities of the Semantic Web and Web Services technologies in both annotating services and delivering relevant services to the end-users. The project contains three major case studies. Firstly, a demonstrational intelligent yellow pages service is created, where the services can be annotated and found easily using ontology techniques. Second, we study application of semantic portal techniques using content from the eGovernment portal Suomi.fi. Thirdly, ontological representation and discovery of eHealth services are studied. The results of the IWebS project will include an intelligent annotation editor for the service providers, ontologies for describing and storing the services, and an intelligent search engine for finding services for the end-users.

4.1.3.4 Wireless Woodstock Services in Finland

Project leader: Mäntylä, Martti
Research group(s): Digital Economy, User Experience, Digital Content Communities
Researchers: Pitkänen, Olli; Salovaara, Antti; Jacucci, Giulio
Schedule: 2004-05-01 ... 2007-02-28
Cooperation units: TeliaSonera Finland and international Celtic consortium including Ericsson AB (Sweden), Blekinge Inst. of Technology (Sweden), Musiclink AB (Sweden), Migoli (Sweden), Stockholm School of Economics (Sweden), TeliaSonera Sweden (Sweden), WIT-Software (Portugal), University of Coimbra (Portugal) and Light Minds (UK).
Funding: Tekes
Keywords: mobile solutions, large-scale events, user experience, pricing models, legal issues
Research programme: Eureka/ CELTIC
www-page and publications: see http://www.hiit.fi/wf/index.html

Wireless Woodstock Services in Finland is the Finnish part of the multinational CELTIC project Wireless Festival. It is managed, carried out and funded separately in Finland, but is implemented in close co-operation with participants in Sweden, Portugal and in the UK. The two and a half year project will study, prototype and evaluate mobile solutions for large-scale events, such as music festivals and sports events. In Finland, the main emphasis is on user experience studies, business and pricing models, value chains, and legal issues.

The project studies, prototypes and evaluates communication and applications solutions for large-scale events to assess their business, user, service and communication aspects. It will give clear recommendations for exploitation of the business opportunities before, during, and after large-scale events, including how the results apply to other markets, e.g., conferences, public gatherings, and disaster areas. The project will apply an iterative approach involving

users and stakeholders throughout the process to prove the feasibility of new technology, service, and business concepts. This leads to two large field trials validating and demonstrating the benefits and advantages of the developed communication platform and new mobile multi-media services.

In Finland, the research focus is on user experiences, business and pricing models, value chains, and legal issues. The Finnish part complements the main study by examining more deeply and piloting event logics, interesting service areas, and the Finnish consumer behaviour. The first Finnish study took place in Jyväskylä, in August 2004 at Neste Rally. Its focus was on understanding visitor expectations regarding video clip mms-services.

4.1.4 Digital Economy

Digital Economy (DE) refers to an integrated approach to legal, technical, societal, and economic research issues that are specific to the network society. The rapid development of information and communication technologies challenges the traditional ways to understand, structure, organize, analyze, develop, and regulate the activities in a society. Especially, regulation in the form of legislation and self-adopted bylaws seems to contain surprising difficulties. The strong link between the information society and globalization poses difficult problems to policy-makers both in public administration and in companies.

The DE group consists at present of about fifteen persons. Professor Jukka Kemppinen is the leader in charge of the group, and Olli Pitkänen is the Programme Coordinator. The group's strengths especially include issues in intellectual property rights, digital rights management, open source licensing, and security. Information society research, led by Dr Pekka Himanen, is another strong line of work within the group.

The DE group has published a number of papers and reports including several journal articles. The projects have also developed working prototypes to study certain aspects of the future network services. In August 2003, MobileIPR organized the First International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet. Co-operation with other research institutes, especially with the UC Berkeley, USA, and the Lappeenranta University of Technology has continued actively.

4.1.4.1 Structures of Mobile Digital Economy (DE Core)

Project leader: Kemppinen, Jukka
Research group(s): Digital Economy
Researchers: Virtanen, Perttu (project manager); Mäntylä, Martti; Reti, Tommo Soininen, Aura; Hietanen, Herkko
Schedule: 2002-01-01 ... 2004-12-31
Cooperation units: UC Berkeley, USA, Lappeenranta University of Technology
Funding: Tekes; Nokia Corporation; TeliaSonera Finland; Elisa Communications; Asianajajien tekonologiaoikeudellinen tutkimusyhdistys
Keywords: Digital economy, computer law, electronic commerce, mobile computing, virtual communities, metadata, content management, digital content distribution
Research programme: Tekes / UTT
www-page and publications: see http://www.hiit.fi/de/core/

The project studied the structures of digital economy based on mobile computing particularly from legal and technical viewpoints. The DE Core research work progressed in 2004 according to the project plan. The researchers carried out the study in Finland and at UC Berkeley. The research has yielded a number of research papers, articles, conference

papers and presentations both in Finland and abroad together with the planned seminar days and research topic get-togethers with funding partners and other research institutions and universities including subjects such as digital content distribution, corporate intellectual property strategies, software and business method patenting, protection of databases and alternative licensing methods in the electronic environment. All this has also paved the way for the project pilot work by means of a working demonstration and related research on Digital Content Distribution Management System, DiMaS, that was successfully completed, exhibited and awarded a prize in 2004. An extensive final report on the main findings and research was also prepared.

4.1.4.2 The Welfare of Nations: The Interaction of Technology, Economy and Society through the Cases of Silicon Valley, Finland/EU and Singapore/China

Project leader: Mäntylä, Martti
Research group(s): Digital Economy
Researchers: Himanen, Pekka; Kalliokoski, Matti; Rantakokko, Mika
Schedule: 2002-01-01 ... 2004-04-30
Cooperation units: Berkeley Center for the Information Society, UC Berkeley, USA
Funding: Tekes
Keywords: Information society, digital economy
www-page and publications: see http://www.hiit.fi/de/welfare/

The Welfare of Nations project studied the interaction of technology, economy, and society through a comparison of three successful information societies: Silicon Valley/USA, Singapore/Greater China, and Finland/EU. The project was carried out in close co-operation with the Berkeley Center for Information Society at University of California, Berkeley. It was intellectually a continuation of the research project carried out by Professor Manuel Castells and Dr Pekka Himanen which resulted in a book *"The Information Society and the Welfare State - The Finnish Model"* (Oxford University Press, 2002).

During 2004, the project prepared its final deliverable *Globaali tietoyhteiskunta - Kehityssuuntia Piilaaksosta Singaporeen* and published it as TEKES report 155/2004. Also some other public reports and presentations on the outcome of the project were prepared and given.

The project was finished at the end of April 2004.

4.1.4.3 Managing Privacy and Trust in P2P Communication (Muppet)

Project leader: Mäntylä, Martti
Research group(s): Digital economy
Researchers: Kortesniemi, Yki; Heikkilä, Juho; Ylitalo, Katri; Stenborg, Markku; Päivärinta, Juha
Schedule: 2004-01-01 ... 2006-12-31
Cooperation units: Laboratory for Theoretical Computer Science (TKK); University of California at Berkeley, USA
Funding: Tekes; Ericsson; Nokia; Yleisradio
Keywords: trust, privacy, P2P, mobile
www-page and publications: see https://isoveli.hiit.fi/Muppet/

Peer-to-Peer technologies can offer radically new possibilities for communication, be it broadcast, group or person-to-person communication. At the same time, the distributed

nature of this technology presents us with many trust and privacy challenges especially in the mobile environment.

In 2004 we looked at game-theory-based incentives, collected an extensive state-of-the-art report on existing reputation mechanisms, synthesized an initial reputation management engine that can work in a peer-to-peer fashion, and built a simple demonstration for Nokia Series 60 phones based on the findings. The demonstration focused on a simple e-market application.

In 2005 the work is planned to continue with reputation in multiple contexts. We will now be taking a closer look at digital identities and privacy aspects of reputation systems, and implementing a prototype over a real P2P network.

4.2 Basic Research Unit Activities

The mission of BRU is to do basic research in computer science in areas in which there are applications visible, either in other sciences or in industry. The research themes are adaptive computing (led by Dr Patrik Floréen and Professor Hannu Toivonen), data analysis (Professor Heikki Mannila, Professor Hannu Toivonen), and neuroinformatics (Dr Aapo Hyvärinen).

4.2.1 Adaptive Computing Systems

Adaptive computing focuses on the methodology and implementation of systems that adjust to different situations. An adaptive system may change its own behaviour to the goals, tasks, interests, and other features of individual users. It may be context-aware adapting to changes in location, time and user activity. Adaptive computing plays an important role in ubiquitous and pervasive computing as well as in intelligent and user-friendly applications.

4.2.1.1 Coordination of Research Programme on Proactive Computing

Programme leader: Mannila, Heikki
Programme coordinator: Lindén, Greger
Schedule: 2002-01-01 … 2005-12-31
Funding: Academy of Finland
Research programme: Academy of Finland / Proactive Computing (PROACT)
www-page and publications: see http://www.aka.fi

BRU coordinates the Research Programme on Proactive Computing (PROACT 2002-2005) jointly funded by the Academy of Finland, Tekes and the French Ministry of Research. The objective of the coordination is to guarantee that the projects form a coherent programme by co-operating and benefiting from each other's work. In addition to the coordination, HIIT participates in three PROACT projects, namely Context Recognition by User Situation Data Analysis (CONTEXT), Networking and Architecture for Proactive Systems (NAPS) and Proactive Information Retrieval by Adaptive Models of User's Attention and Interests (PRIMA).

During 2004, the coordinator and the programme leader made a second round of site visits to the projects. The coordination also organised a two-day seminar targeted at the PhD students working in the PROACT projects. The students gave 27 presentations at the workshop. They also helped organise the workshop, reviewing other students' papers and chairing sessions at the workshop. Several other related workshops and events were organised during 2004 by researchers in the programme. Summaries of the site visits,

workshop proceedings and other information of the projects have been published at the web site of the research programme www.aka.fi/proact/.

4.2.1.2 Context Recognition by User Situation Data Analysis (CONTEXT)

Project leader: Toivonen, Hannu
Research group(s): Adaptive Computing Systems
Researchers: Laasonen, Kari; Petit, Renaud; Raento, Mika; Toivonen, Hannu
Schedule: 2002-11-01 … 2005-12-31
Co-operation units: ARU/HIIT
Funding: Academy of Finland, graduate schools
Keywords: Context recognition, mobile devices, data mining
Research programme: Academy of Finland / Proactive Computing (PROACT)
www-page and publications: see http:// www.cs.helsinki.fi/group/context/

Led by Professor Hannu Toivonen, this project looks at ways of analysing context information to enable mobile devices to recognize the situation in which they are used and to present this information to the users. This joint BRU/ARU project is part of the Research Programme on Proactive Computing of the Academy of Finland (see above), and it started in November 2002.

In 2004 we have continued developing the ContextPhone software. It is a prototyping platform for context-aware systems: it allows collection, analysis, transmission and presentation of user context on Series60 mobile phones. ContextPhone was released as free software, and it already has a number of research uses. ContextContacts is a service for automatic communication and representation of the context of your friends, and it is used in field research with HIIT/ARU to study how such automatic publication of one's context can be integrated into the way people manage privacy: what data can safely be communicated, or how should the service be controlled and what level of accountability should be provided.

The ContextMedia application automatically annotates and shares photos and other media captured with the phone, used in co-operation with Media Lab, University of Art and Design Helsinki, to experiment with mobile, locative and situated media practices. University of California Berkeley's Garage Cinema Research Group at the School of Information Management Science is using ContextMedia to study computer-aided annotation of mobile media. MIT Media Lab uses ContextLogger to gather—for the first time—actual measurements of the strength, dynamics and evolution of social networks. VTT Technical Research Center of Finland is also using ContextLogger, in order to study workgroup social patterns. The Oslo National Academy of the Arts is staging a surveillance exhibit with the aid of a custom ContextPhone application.

We have also continued developing methods for analyzing context data, in particularly discovery and recognition of personally important places and routes based on traces of GSM cells of a user. An immediate application is in presence services: "John left the office ten minutes ago and is heading towards home" would be an informative description of John's status. Results include new concepts and algorithms for these tasks. We have also investigated user interaction and privacy issues for context aware mobile computing.

4.2.1.3 Networking and Architecture for Proactive Systems – Algorithmics (NAPS)

Project leader: Floréen, Patrik
Research group(s): Adaptive Computing Systems
Researchers: Floréen, Patrik ; Kohonen, Jukka; Nurmi, Petteri; Suomela, Jukka
Schedule: 2003-01-01 … 2005-12-31
Co-operation units: Laboratory for Theoretical Computer Science (TKK); Networking Laboratory (TKK)
Funding: Academy of Finland
Keywords: Ad hoc networking, algorithms
Research programme: Academy of Finland / Proactive Computing (PROACT)
www-page and publications: see http://www.cs.helsinki.fi/hiit_bru/projects/naps/

This project at HIIT/BRU is part of the Networking and Architecture for Proactive Systems (NAPS) consortium, to which belong the research groups of Professor Pekka Orponen (Laboratory for Theoretical Computer Science, TKK) and Professor Jorma Virtamo (Networking Laboratory, TKK). It is part of the Proactive Computing (PROACT) research programme of the Academy of Finland.

The network computing and communication models underlying proactive applications give rise to new opportunities and challenges in the fields of algorithm design and analysis. The project approaches, among others, multiobjective optimisation problems in topology control and routing of ad hoc and sensor networks.

In 2004 the project has continued research on energy-efficient designs, now concentrating on data-gathering problems in sensor networks. We have studied balanced data gathering, where the task is to maximise the amount of data received at the sink, while simultaneously ensuring that a minimum amount of data is forwarded from all sensors. The results were published at the workshop Algorithmic Aspects of Wireless Sensor Networks (Algosensors 2004) and also presented by Jukka Kohonen at the workshop of the PROACT programme. A faster algorithm for the balanced data-gathering problem has since been obtained and the results have been accepted for publication in the journal Theoretical Computer Science.

In 2004, work on applying game theory to routing in ad hoc networks was initiated by Petteri Nurmi. Based on earlier work on lifetime maximisation in ad-hoc networks, an article was published in the IEEE Journal on Selected Areas in Communications at the beginning of 2005.

4.2.1.4 SPACE4U Terminal Software Management System Design/Development

Project leader: Floréen, Patrik
Research group(s): Adaptive Computing Systems
Researchers: Floréen, Patrik; Przybilski, Michael; Kurppa, Teemu
Schedule: 2003-07-01 … 2005-06-30
Co-operation units: Industrial Information Technology Laboratory (TKK); Nokia Research Center
Funding: Nokia Research Center
Keywords: Context-awareness; component frameworks; middleware; embedded systems
www - page and publications: http://www.cs.helsinki.fi/group/space4u/

The EUREKA ITEA project Space4U (Software Platform and Component Environment for yoU) builds on the foundations established by the previous ITEA project ROBOCOP. Whereas ROBOCOP provided a component-based software architecture for the middleware

of embedded appliances, Space4U aims to extend the framework in the areas of power management, fault management and remote terminal management. HIIT/BRU works as a subcontractor to Nokia Research Center. The work focuses on the terminal management and context-aware configuration part of Space4U. Specifically, this part of the project develops context-aware selection, download and execution of software components on resource-constrained terminals.

The project has been undertaken in two parts: the first "Terminal Software Management System Design/Development" during 07/03-06/04, and the second "Terminal Management Demonstrator Development" during 07/04-06/05. In 2004, the group at HIIT/BRU developed a context-aware extension of the ROBOCOP middleware platform for Symbian terminals, based on a blackboard architecture approach, together with prototype applications to demonstrate the provided functionality.

Results have been published in a workshop paper at the workshop Component-Oriented Approaches to Context-Aware Computing at the European Conference on Object-Oriented Programming (ECOOP 2004).

4.2.1.5 MobiLife – Mobile Life

Project leader: Floréen, Patrik
Research group(s): Adaptive Computing Systems
Researchers: Floréen, Patrik **;** Nurmi, Petteri; Przybilski, Michael; Raento, Mika
Schedule: 2004-09-01 ... 2006-12-31
Cooperation units: Nokia Corporation; Alcatel-CIT; Ericsson AB; Oy LM Ericsson Ab; Hewlett-Packard Italiana srl; Motorola Ltd.; Motorola SAS; NEC Europe Ltd.; Siemes Mobile Communications SpA; DoCoMo Communications Laboratories Europe GmbH; Elisa Corporation; Fraunhofer Gesellschaft zur Foerderung der angewandten Forschung e.V.; University of Helsinki; Universität Kassel, Fachgebiet Kommunikationstechnik; The University of Surrey; NEOS Engineering SRL; Stichting Telematica Instituut; UNIS, spol s.r.o.; Suunto Oy; BellStream SP.z.o.o.; Telecom Italia
Funding: European Union
Research programme: Information Society Technologies (EU FP6), FP6-2003-IST-2 (IST-511607 MobiLife)
www - page and publications: http:// www.cs.helsinki.fi/group/mlife/

The objective of MobiLife is to bring advances in mobile applications and services within the reach of users in their everyday life by innovating and deploying new applications and services, applying a user-centred approach. The MobiLife project is coordinated by Nokia and has 22 partners from nine countries. The group at HIIT / BRU focuses on the context-awareness aspects of the project and researches context-reasoning methods and context-aware software architectures. HIIT/BRU is task leader of the task "Context management".

Besides collecting and reporting the state of the art, the work in 2004 has concentrated on the initial design of the Context Management Framework, including work on requirements and architecture. The areas we are working on in addition to context management are personalisation, group awareness and proactive service provisioning.

MobiLife is part of the Wireless World Initiative, which comprises several projects for IST.

4.2.2 Data Analysis

The developments in measurement and data collection technologies have made it possible to gather and store large amounts of information in many areas of science and industry. The ability to analyze these masses of raw data has increased at a much slower pace, however. The BRU research program on data analysis develops data mining and computational statistics methods for various application tasks.

4.2.2.1 Altti: Computational Methods for Gene Mapping

Project leader: Toivonen, Hannu
Research group(s): Data analysis
Researchers: Eronen, Lauri; Geerts, Floris; Goethals, Bart; Hintsanen, Petteri; Muhonen, Juho; Onkamo, Päivi; Sevon, Petteri; Toivonen, Hannu
Schedule: 2003-08-01 … 2005-02-28
Co-operation units: GeneOS Ltd; Jurilab Ltd; Cyberell Ltd; Karolinska Institute, Sweden; Department of Medical Genetics (UH)
Funding: Tekes; companies; BRU; UH
Keywords: Gene mapping, genetics, data mining
www-page and publications: see http://www.cs.helsinki.fi/group/genetics/

We develop and apply methods for gene mapping, and for analysing genotypes, haplotypes, and phenotypes and their relationships in more general. Locating genes that predispose to diseases is highly important in understanding the etiology of complex common diseases, such as heart disease, or asthma. For association analysis, the sample of patients, controls, and their relatives is genotyped and haplotyped, i.e. the two alleles at each marker locus in each individual are ordered according to their parental origin. Then, alleles and short strings of alleles of nearby (consecutive) markers correlating with the patient-control status are searched by means of association methods. The aim is to pinpoint the location of the disease susceptibility (DS) mutation as accurately as possible.

In 2004 we continued developing computationally efficient methods for different stages of analysis of large genetic data sets. We developed new methods for haplotyping large marker maps. These methods allow efficient and accurate haplotyping of modern high-throughput genetic data, and they have been applied with good success to some very large data sets of our partners. We improved our gene mapping method, HPM, and witnessed a breakthrough in the mapping of an asthma gene using HPM. Significant efforts were invested into developing new population simulation technology and using it to study the effects of different gene mapping study designs, sampling methods, mapping methods, and haplotyping methods on gene mapping accuracy. The results indicate that our pipeline of computationally efficient tools is also accurate and statistically powerful in gene mapping, and it allows easier and more economical study designs.

4.2.2.2 Data Analysis for Functional Genomics

Project leader: Hollmén, Jaakko
Research group(s): Data analysis
Researchers: Ruosaari, Salla; Seppänen, Jouni; Patrikainen, Anne
Schedule: 2002-01-01... 2005-31-12
Co-operation units: Finnish Institute of Occupational Health, Department of Medical genetics (UH), National Public Health Institute (KTL)
Funding: Academy of Finland, European Commission; graduate schools
www-page and publications: see http://www.cis.hut.fi/jhollmen/hiit/genomics.html

The project has concentrated on developing computational methods for analyzing gene expression data in combination with other sources of data, such as CGH microarrays measuring the gene copy number changes. Such copy number number alterations play an important role in cancer research. Also, the context, such as the spatial location of the genes, has been used as an additional source of information. One of the goals of the research has been to identify a few important genes or factors behind the disease, that is, the identification of diagnostic markers. Methodologically, research towards learning of parsimonious or sparse models has been initiated. The work is mainly carried out in collaboration with domain specialists in biology and medicine.

4.2.2.3 CompGenome: New Computational Methods for Analyzing the Structural and Functional Landscapes of Mammalian Genomes

Project leader: Mannila, Heikki **Research group(s):** Data analysis **Researchers:** Gionis, Aristides; Haiminen, Niina; Koivisto, Mikko; Kollin, Jussi; Heino, Jaana **Schedule:** 2003-11-01 ... 2007-12-31 **Co-operation units:** National Public Health Institute (KTL); Finnish Genome Center; Karolinska Institute, Sweden; Department of Medical Genetics Research (UH) **Funding:** Academy of Finland **Research programme:** Academy of Finland / Systems Biology and Bioinformatics **www-page and publications:** see http://www.cs.helsinki.fi/group/genetics/

The availability of a large mass of genomic data will make it possible to study in detail the genomic landscape in humans and other mammalian organisms and to investigate the variation both within and between species. The CompGenome project will develop and apply computational tools for describing the genomic and functional variation between individuals and between species, and study the significance of these variations for the functions of genes. This information will eventually be used to understand the potential relationships of such variations in the genome landscape and complex diseases. In particular, the project will develop (a) tools for the hierarchical description of the haplotype structure of human (and other mammalian) genomes, (b) techniques for the identification of rearrangements, duplications, and other large-scale variations in genomes; (c) methods for trait-cluster-based multiple locus testing in complex traits, and (d) tools for analyzing the relationship between transcript profiles and sharing of genome segments.

The project started in late 2003 and will last four years. Earlier results in the area include methods for finding haplotype blocks and efficient computational tools for the computation of genetic risks. The work in 2004 has concentrated on segmentation algorithms and their applications in genomes. There are strong ties to work in the areas of gene mapping and algorithmic data mining.

4.2.2.4 Algorithmic and Probabilistic Methods in Data Mining

Project leader: Mannila, Heikki
Research group(s): Data analysis
Researchers: Toivonen, Hannu; Hollmén, Jaakko; Gionis, Aristides; Geerts, Floris; Goethals, Bart; Terzi, Evimaria; Leino, Antti; Mielikäinen, Taneli; Seppänen, Jouni; Tatti, Nikolaj; Bingham, Ella;
Schedule: 2002-01-01…
Funding: Academy of Finland; graduate schools; European Commission funding from TKK; BRU basic funding
www-page and publications: see http://www.cs.helsinki.fi/research/fdk/datamining/

The project develops methods for the exploratory data analysis of large and high-dimensional data sets. One of the themes has been finding frequent patterns in large collections of data. The pattern classes include ordered and unordered patterns. Currently areas of interest include condensed representations and the combination of combinatorial and probabilistic techniques for approximating distributions. For sequential data, interests are in algorithms for sequence segmentation under various restrictions and in discovery of order from unordered data sets. Also issues in subspace clustering and spectral methods have been studied.

In 2004 there were several interesting developments. The methods for relational ranking and clustering aggregation seem to be quite strong. Work on condensed representations of large (0-1) datasets continued on several fronts, with many interesting results. The activities on combining probabilistic and algorithmic aspects of data mining led to new models and algorithms for 0-1 data.

4.2.2.5 Spatial and Temporal Data Mining

Project leaders: Salmenkivi, Marko; Gionis, Aristides
Research group(s): Data analysis
Researchers: Leino, Antti; Hyvönen, Saara; Gionis, Aristides; Mannila, Heikki
Co-operation units: Research Institute for the Languages of Finland; Division of Atmospheric Sciences (UH); Finnish Museum of Natural History (UH); Department of Geology (UH); Institute of Biotechnology (UH)
Funding: Graduate schools, BRU basic funding
www-page and publications: see http://www.cs.helsinki.fi/research/fdk/datamining/

Study of place names, dialects, biodiversity, and climate, for example, results in data sets that have strong spatial and (possibly) temporal components. The research project looks at data mining methods that can be used to find spatial and temporal relationships in high-dimensional data. The project works in very close collaboration with the "Algorithmic and probabilistic methods in data mining" project.

The main application areas in 2004 were closely related to linguistics: firstly, the study of the distribution of place names in Finland, and secondly, the investigation of spatial distributions of Finnish dialect words. In addition to research on compound names, which was initiated in the group in 2002-03, we adopted methods from language technology to extract individual name elements from compound names. Distributions and concentrations of name elements, as well as relationships between them are issues of great interest in, e.g., linguistics, study of the history of settlement, cultural history, and comparative religions.

We have developed efficient methods for detecting attraction and repulsion in point patterns in large data sets. The methods have been applied to the place name data in, e.g., a case study on the Iron Age Finnish culture. Spatial autocorrelation and concentration of the occurrences of a single name element have been studied by utilizing spatial clustering. It has turned out that there are remarkable differences between name elements as to the local accumulation, terms referring to animals, for instance, having a strong tendency to form local clusters. The reasons for the observed phenomena are not fully understood yet, and they are under investigation.

Clustering and dimension reduction techniques have been applied to a large set of Finnish dialect words, each word being associated with the set of municipalities where the word is known to be used. A goal of the research has been the exploration and evaluation of dialectically coherent regions. The results mostly confirm the prevailing understanding of the major Finnish dialectical regions with some interesting exceptions.

We have employed semi-supervised clustering techniques and probabilistic modelling with spatial dependence structure in the study of the distribution of breeding bird species in Finland.

In close collaboration with the Division of Atmospheric Sciences we have analyzed meteorological and micrometeorological data sets to detect factors influencing the formation of atmospheric aerosol particles. Clustering and classification methods have been used. Also the applicability of kernel methods to this task has been under study.

4.2.3 Neuroinformatics

Leader: Hyvärinen, Aapo
Researchers: Hoyer, Patrik; Hurri, Jarmo; Hyvärinen, Aapo; Kurki, Ilmari; Köster, Urs; Perkiö, Jukka; Shimizu, Shohei
Schedule: 2003-08-01...
Co-operation units: Department of Psychology (UH); Neural Networks Research Centre (TKK); Osaka University, Naples University, Riken Brain Science Institute (Japan)
Funding: Academy of Finland; BRU basic funding, Japanese and German foundations
www-page and publications: see http://www.cs.helsinki.fi/hiit_bru/neuroinf/

Neuroinformatics is broadly defined as the intersection of Information technology and neuroscience. Our research goals are 1) to build mathematical models of brain function. In computational visual neuroscience, our approach is to consider how the brain performs a sophisticated statistical and probabilistic analysis of the environment. To this end we also need 2) to develop new multivariate statistical models. A fundamental mathematical method that we use is independent component analysis (ICA) and some of its extensions. As a collaborative effort, we also 3) apply advanced statistical methods on neuroscientific data.

The leader of the group, Dr Aapo Hyvärinen, joined BRU in 2003, and 2004 was the first year of full operation. Two post-doctoral fellows, Patrik Hoyer and Jarmo Hurri joined the group in January 2004, as well as new PhD students.

Regarding computational neuroscience, a PhD student, Ilmari Kurki, started working in a project funded by the Academy of Finland "Advanced data analysis in vision research: feature interaction and nonlinear classification images", which is conducted in collaboration with the department of psychology. Urs Köster from the University of St Andrews started a PhD in computational neuroscience in August, on the topic of multi-layer models of natural images and cortical visual processing.

We developed new multi-layer models of natural image statistics with a visiting student from ETH Zürich, Michael Gutmann. Models on the probabilistic underpinnings of segmentation were also investigated. Statistical models of natural language were developed in collaboration with the Neural Networks Research Centre of TKK. New methods for brain imaging data analysis were developed in collaboration with brain imaging experts in Naples University.

On the theoretical front, in collaboration with Shohei Shimizu, a long-term visitor from Osaka University, new methods for causal discovery for continuous-valued data were developed. These are based on post-processing and re-interpreting results obtained by ICA. Application of ICA to classification was investigated in cooperation with a visitor from the University of Valencia, Asun Vicente. A new method for estimating non-normalized statistical models as well as new methods combining ICA with non-negative constraints were developed.

4.3 EU-activities

HIIT has continued being active in EU-funded research. In 2004 HIIT was involved in a number of EU projects and submitted many proposals. The part of the framework programme most suitable for HIIT is the IST-priority, but also some other parts are relevant: the "Life science, genomics and biotechnology for health" priority is relevant for the bioinformatics research and the Marie Curie programme is relevant in all fields for researcher mobility.

The biggest EU-funded research effort HIIT participates in is the Wireless World Initiative (WWI). WWI consists of five Integrated Projects (IP) and a coordination action. HIIT is participating in the Nokia-coordinated application MobiLife – Mobile Life, which concentrates on ubiquitous mobile applications and services from a user-centric point of view. The consortium consists of many major companies and research organisations in Europe. Both units of HIIT are active in the project.

The Superpeer Semantic Search Engine (ALVIS) Specific Targeted Research Project (STREP), led by Professor Henry Tirri in the ARU, has completed its first year of research. The project aims at building the foundation for a distributed semantic search engine. The project continues for another two years. HIIT is the coordinator of the project.

Application of Probabilistic Inductive Logic Programming II (APrIL II) is also a three-year STREP, funded as part of the Future and Emerging Technologies (FET) scheme of the IST priority. The project has completed its first year of research. The project works on probabilistic logic-learning systems and their application to real-life problems. BRU is a partner in this project.

Inductive Queries for Mining Patterns and Models (IQ) is a three-year STREP, funded as a part of the Future and Emerging Technologies (FET) scheme of the IST priority. The goal of the project is to develop a sound theoretical understanding of inductive querying that enables one to develop effective inductive database systems and to apply them to significant real-life applications. The successful proposal was submitted in September 2004 and the project starts in September 2005. BRU is a partner in this project.

Academy Professor Heikki Mannila of BRU is a member of the EU's European Strategy Forum on Research Infrastructures (ESFRI) Steering Committee on Physical Sciences and Engineering. ESFRI is preparing a European roadmap for new research infrastructures of pan-European interest.

In addition to these, HIIT participated in EUREKA-activities and is preparing further EUREKA participation.

5 Research Training and Research Visits

5.1 Doctoral Degrees Earned by HIIT Personnel

30.1.2004: Mikko Koivisto, *Sum-Product Algorithms for the Analysis of Genetic Risks*, University of Helsinki, Department of Computer Science, supervised by Professor Heikki Mannila.

16.7.2004: Petteri Sevon, *Algorithms for association-based gene mapping*, University Of Helsinki, Department of Computer Science, supervised by Professor Hannu Toivonen.

22.10.2004: Matti Kääriäinen, *Learning small trees and graphs that generalize*, University of Helsinki, Department of Computer Science, supervised by Professor Tapio Elomaa.

12.12.2004: Giulio Jacucci. *Cases of Configuring Physical Interfaces in Mixed Media in the public dissertation event*, University of Oulu, Department of Information Processing Science, supervised by Professor Kari Kuutti.

5.2 Post-graduate Courses Arranged by HIIT

Spring 2004 Courses

Modeling of Perception (Aapo Hyvärinen)

Laskennallisen data-analyysin seminaari (Marko Salmenkivi)

Research Themes in Context-Aware Computing (Greger Lindén)

Data Mining Seminar: Analysis of sequential, temporal and spatial data (Aristides Gionis)

Special course on attentional resources in mobile interaction (Antti Oulasvirta)

Seminar on interruptions and intelligent environments Seminar on interruptions and intelligent environments (Antti Oulasvirta)

Laboratory course in software technology (Kenneth Oksanen)

Computer law (Olli Pitkänen)

Network application frameworks and XML (Pekka Nikander and Sasu Tarkoma)

Seminar in new business in digital economy: Computer mediated communities (Marko Turpeinen)

Course on probability theory (Petri Myllymäki)

Research Seminar on Hot Topics in Internet Protocols (Kimmo Raatikainen)

Linux Seminar (Kimmo Raatikainen)

Mobile systems programming (Sasu Tarkoma)

Summer 2004 Courses

Introduction to Independent Component Analysis (Patrik Hoyer, Jarmo Hurri, Aapo Hyvärinen)

4th Berkeley-Helsinki Workshop on Telecommunication Software Architectures (Kimmo Raatikainen)

Information theory and statistics (Jorma Rissanen)

Autumn 2004 Courses

Molecular Genetics and Gene Mapping for Methodological Sciences (Päivi Onkamo)

Approximation Algorithms (Aristides Gionis)

Classification (Patrik Floréen)

Information Visualization (Alexander Hinneburg)

Puukieliopit (Miro Lehtonen, Mika Raento)

Special Course on Data Mining (Marko Salmenkivi)

Computational Neuroscience (Patrik Hoyer, Jarmo Hurri)

Pattern Analysis in Sequences (Veli Mäkinen, Esko Ukkonen)

Research Seminar on Computational Data-Analysis (Aristides Gionis)

Special course on mobile field experiments (Antti Oulasvirta)

Research seminar: ID/Loc split and Distributed Hash Tables (Pekka Nikander and Andrei Gurtov)

Research seminar on intelligent information retrieval (Petri Myllymäki)

Nodes Research Seminar (Kimmo Raatikainen)

5.3 Visits to HIIT

Rudi Cilibrasi, MSc, The Centre for Mathematics and Computer Science, Holland (16-29.8.2004)

Marc Davis, Professor, University of California at Berkeley, USA (18-24.8.2004)

Debora Donato, PhD student, University of Rome, Italy (11.9.-11.10.2005 and 26.11.2004-27.1.2005)

Christos Faloutsos, Professor, Carnegie Mellon University, USA (7.-14.8.2004)

Jens Grossklags, MSc, University of California at Berkeley, USA (27.6-1.7.2004)

Michael Gutmann, PhD student, Tokyo University, Japan (14.8.-5.9.2004)

Mikio Handa, Senior Consultant at Nomura Research Institute, Japan (5.2.2004)

Guo Hang, MSc, Tsinghua University in Beijing (1.11.2004-30.4.2005)

Simon Jones, Professor, Managing Director of MediaLab Europe (20.10.2004)

Anthony Joseph, Professor, University of California at Berkeley, USA (25-30.5. 2004 and 19.8-30.11.2004)

Simon Lacoste-Julien, M.Sc., University of California at Berkeley, USA (20.5-20.6.2004)

Lawrence Lessig, Professor, Stanford Law School, USA (24.5.2004)

Wakako Nakamura, Dr Hokkaido University, Japan (11.-27.3.2004)

Guillame Obozinski, MSc, University of California at Berkeley, USA (1.8-5.9.2004)

Tero Ojanperä, Head of Nokia Research Center, Finland (30.8.2004)

Spiros Papadimitriou, PhD student, Carnegie-Mellon University, Pittsburgh, USA (7.8.-16.10.2004)

Shohei Shimizu, Research Fellow, Osaka University, Japan (1.1.-31.12.2004)

Sam Steinhardt, Director of Stanford Center for Innovations in Learning (SCIL), USA (28.10.2004)

Asuncion Vicente, PhD student, Miguel Hernandez University, Alicante, Spain (9.8.-30.10.2004)

Anita Wilhelm, Researcher, University of California at Berkeley, USA (13.-16.9.2004)

Ydo Wexler, PhD student, Israel Institute of Technology, Haifa, Israel (6.8.-3.10.2004)

Qiang Zhou, Professor, Tsinghua University in Beijing, China (1-30.11.2004)

Shoshana Zukoff, Professor, Harvard Business School, USA (27.10.2004)

5.4 Visits from HIIT

Wray Buntine, PhD, Tsinghua University in Beijing, China (10.-14.5.2004)

Mikko Kontiainen, MSc, Tsinghua University in Beijing, China (10.-14.5.2004)

Martti Mäntylä, Professor, Tsinghua University in Beijing, China (17.3-24.3.2004)

Martti Mäntylä, Professor, University of California at Berkeley, USA (1.-5.11.2004)

Pekka Nikander, Tsinghua University in Beijing, China (27.2.2004)

Jukka Perkiö, MSc, Tsinghua University in Beijing, China (18.9.-25.9.2004)

Sami Perttu, MSc, Tsinghua University in Beijing, China (18.9.-25.9.2004)

Kimmo Raatikainen, Professor, Tsinghua University in Beijing, China (25.2.2004)

Kimmo Raatikainen, Professor, Waseda University, Tokyo, Japan (22.3.2004, 4.10.2004)

Matti Rantanen, MSc, University of California at Berkeley, USA (17.3.-31.12.2004)

Risto Sarvas, MSc, University of California at Berkeley, USA (23.3.-2.4.2004)

Risto Sarvas, MSc, Media Lab Europe, Dublin, Ireland (11.5.-14.5.2004)

Risto Sarvas, MSc, MIT Media Lab, Cambridge, USA (5.6.-14.6.2004)

Tomi Silander, MSc, University of California at Berkeley, USA (1.-3.7.2004)

Aura Soininen, LLM, University of California at Berkeley, USA (1.1-30.11.2004)

Henry Tirri, Professor, University of California at Berkeley, USA (several visits during 2004)

Henry Tirri, Professor, Tsinghua University in Beijing, China (several visits during 2004)

Ville Tuulos, MSc, University of California at Berkeley, USA (27.6.-4.7.2004)

Ville Tuulos, MSc, Tsinghua University in Beijing, China (18.9.-25.9.2004)

Katri Ylitalo, MSc, University of California at Berkeley, USA (1.-5.11.2004)

6 Administration

6.1 Overview

According to the management bylaws established in 2002, HIIT is the joint research institute of TKK and UH that consists of two units, the Advanced Research Unit (ARU) and the Basic Research Unit (BRU). HIIT is directed by a Board. The Scientific Advisory Board (SAB) advises the HIIT Board on strategic planning of HIIT's research activities, and the institute's industrial liaison is managed by the Industrial Advisory Board (IAB).

ARU receives its administrative services from TKK. Similarly BRU is administratively associated with UH and follows its administrative rules and procedures. In practice, BRU is hosted by the Department of Computer Science of the UH. The structure is shown in the organisation chart below that depicts the status at the end of the year 2004.
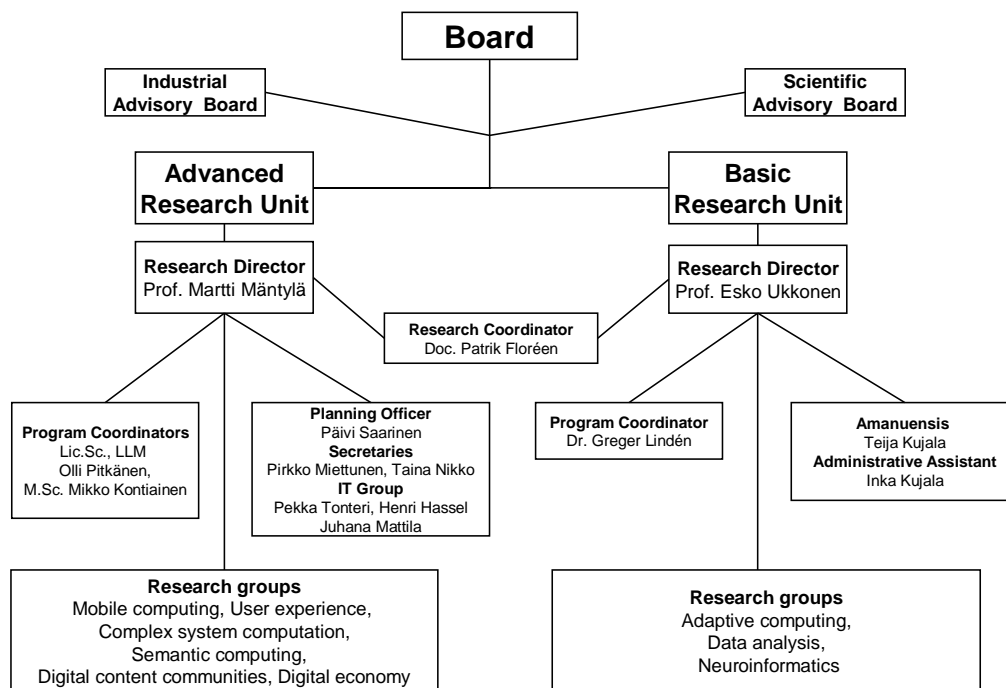
Internally, ARU consists of six research groups, facilitated by a small administration team. In addition, some thematic research areas have a Programme Coordinator who facilitates the planning and management of research activities. The Management Board of ARU, headed by the Research Director and consisting of senior researchers and administrative personnel, handles most daily management and coordination activities.

The BRU administrative structure is very light, as the unit can use the services of the departments of computer science of TKK and UH. Heikki Mannila and, from September 2004, Esko Ukkonen is in overall charge of the unit's general management, budgeting, and planning. The administrative personnel consists of Administrative Assistant Inka Kujala, responsible for personnel, budgeting, events, and miscellaneous secretarial tasks, and

Amanuensis Teija Kujala, who participates in administration with about 50 % of her total work time.

6.2 Board

The highest decision-making body of HIIT is the Board. The Board consists of nine full members of which eight are appointed by the parent universities and represent the academic community and the main industrial partners of HIIT. One member of the board represents and is elected by HIIT personnel. In addition, the Board invites members from industrial companies with whom HIIT co-operates to participate in the work of the Board. The decision-making power is invested in the full members, whereas the invited members have the right to attend and to speak at the meetings.

In 2004 the Board convened six times. Apart from dealing with the statutory tasks (i.e. approving the annual budgets and activity plans of the both units, approving the HIIT annual report, following up the work of the units through the regular activity updates given by the two Research Directors of HIIT, etc.), the Board launched and managed the planning of HIIT's future after the initial 5-year period. This process remained an item for the Board's agenda throughout the year. The Board also nominated Professor Esko Ukkonen as the Research Director of the Basic Research Unit from 1.9.2004, after Professor Heikki Mannila had been appointed Academy Professor.

In 2004 the Board members (and their personal deputies) were as follows:

Members:
Vice Rector, Professor Mauri Airila, TKK, Chairman of the Board
(Vice Rector, Professor Olavi Nevanlinna)
Professor Olli Simula, TKK (Professor Markku Syrjänen)
Vice Rector, Professor Thomas Wilhelmsson, UH, Vice-Chairman of the Board
(Vice Rector, Professor Marja Makarow)
Academy Professor Esko Ukkonen, UH. From 5.10.2004 Professor Jukka Paakki
(Professor Jukka Paakki . From 5.10.2004 Professor Jyrki Kivinen)
Raimo Vuopionperä, LM Ericsson (Björn Melén)
Aimo Maanavilja, Elisa (Pertti Hölttä)
Veikko Hara, TeliaSonera Finland. From 24.3.2004 Juha Aaltonen
(Juha Aaltonen. From 24.3.2004 Martin Mäklin)
Heikki Saikkonen, Nokia. From 14.12.2004 Jan Bosch
(Harry Herlin)
Petri Myllymäki, HIIT (Greger Lindén)
*Invited Member*s:
Ari Hirvonen, TietoEnator (Olli Lötjönen)
Seppo Vanhatalo, Alma Media (Marko Turpeinen)
Eskoensio Pipatti, Sanoma-WSOY
Juha Vesaoja, Yleisradio
Raimo Näätsaari, Nordea (Juha Toivari)

The two Research Directors of HIIT are responsible for preparing and submitting propositions to the Board. In addition, the Research Coordinator and the Program Coordinators have a right to attend the meetings. In 2004 these expert members were as follows:

Expert member:
Research Director Martti Mäntylä, HIIT ARU
Research Director Heikki Mannila, HIIT BRU. From 5.10.2004 Esko Ukkonen
Research Coordinator Patrik Floréen, HIIT ARU & BRU
Program Coordinator Olli Pitkänen, HIIT ARU
Program Coordinator Mikko Kontiainen, HIIT ARU
Program Coordinator Anton Puolakka, HIIT ARU until 30.4.2004

6.3 Scientific Advisory Board

The Scientific Advisory Board (SAB) of HIIT consists of internationally prominent scholars who are invited by the HIIT Board. The objective of the SAB is to provide critical guidance about HIIT's research activities and to advise the HIIT Board on strategic planning of the future research directions of HIIT. The following scholars are members of the SAB:

Dr Ross Anderson, University of Cambridge
Professor Alberto Apostolico, Purdue University
Professor Richard Buxbaum, University of California at Berkeley
Professor Christos Faloutsos, Carnegie Mellon University
Professor Randy Katz, University of California at Berkeley
Professor Bengt Jonsson, Uppsala University
Professor Martin Kersten, University of Amsterdam and CWI[1]
Professor Kari-Jouko Räihä, University of Tampere
Professor Mart Saarma, University of Helsinki
Professor John Shawe-Taylor, University of Southampton
Professor Hal Varian, University of California at Berkeley
Dr Martin Vingron, Director, Max Planck Institute for Molecular Genetics

[1]) National Research Institute for Mathematics and Computer Science in the Netherlands

HIIT's Scientific Advisory Board met for its second meeting on 15-17 November 2004. The members present were Randy Katz (UCB), Alberto Apostolico (Purdue Univ. and Univ. of Padua), Christos Faloutsos (Carnegie Mellon), Bengt Jonsson (Uppsala Univ.), Martin Kersten (CWI) and John-Shawe-Taylor (Univ. of Southampoton). During the three days in Helsinki, the board reviewed all the research lines of HIIT. Research group leaders gave presentations and the board met with students in poster and demonstration sessions in both Kumpula and Ruoholahti. The SAB also met with the research directors of HIIT and the computer science department heads of the university and TKK.

The SAB presented a written report of its findings. Each research line's research achievements and plans received comments and, in addition, the SAB had a number of general comments. The SAB called for a unified vision of HIIT to support the merging of ARU and BRU into one unit. Also better administrative transparency was called for in allocation of funds and selection of research lines. The report was used in the strategy discussion for HIIT.

6.4 Industrial Advisory Board

The Industrial Advisory Board (IAB) consists of representatives of companies that work closely with HIIT. The IAB had one meeting in 2004 on September 29, 2004. The meeting focused on the research strategy of HIIT, especially its vision and mission and their realization in the main research lines of HIIT. Presentations on the four main lines (Future Internet, Network Society, Intelligent Systems, Data Mining) were presented. Also issues related to the HIIT-Industry interface, especially the forms of industry participation in research projects, were discussed.

6.5 Personnel

The personnel directly employed by HIIT is formally employed by the two parent universities, i.e. the ARU personnel is employed by TKK and the BRU personnel by UH. In addition, there is a number of persons working in HIIT with some other form of funding, such as postgraduate students with funding from Helsinki Graduate School of Computer Science and Engineering (HeCSE) and researchers with academic positions. Many of HIIT's personnel have double or even triple affiliations. Most common is the affiliation to one or both of the parent universities, but there are also some who share their time between HIIT and some other organisation. The diversity of affiliations is characteristic to HIIT personnel.

In 2004 the ARU staff completed 56 person-years, out of which approximately 90 per cent was with external funding and approximately 10 per cent with budget funding. There is a slight increase in person-years compared to 2003 (54 person-years). In 2004 BRU staff completed 45 person-years, out of which approximately 35% was external funding and approximately 65% was budget funding.

7 Funding and Costs

7.1 Advanced Research Unit Finances

The finances of the Advanced Research Unit for 2004 are shown in the table below. For reference, the comparable numbers of 2002 and 2003 are also given.

	2002	2003	2004
Total funding	**3 025 570 €**	**3 498 410 €**	**3 727 072 €**
TKK funding	258 188 €	168 188 €	165 688 €
UH funding	188 188 €	168 188 €	168 200 €
Academy of Finland	189 724 €	379 462 €	392 087 €
National Technology Agency TEKES	1 644 796 €	2 103 451 €	2 135 821 €
European Union			116 951 €
Industry	643 724 €	600 699 €	693 910 €
Ministries and other public funding	100 950 €	71 694 €	44 779 €
Other domestic funding		6 728 €	9 636 €
Total expenses	**3 121 607 €**	**3 545 099 €**	**3 731 995 €**
Salaries	1 819 884 €	2 376 864 €	2 423 516 €
Other operational expenses	586 327 €	660 739 €	761 813 €
Rents	476 895 €	343 216 €	380 404 €
Service charge to TKK	238 501 €	164 280 €	166 262 €

	2002	2003	2004
University funding % of total funding	15 %	10 %	9 %
External funding % of total funding	85 %	90 %	91 %
Academy funding % of total funding	6 %	11 %	11 %
TEKES funding % of total funding	54 %	60 %	57 %
Industry funding % of total funding	21 %	17 %	19 %
Other public funding % of total funding	3 %	2 %	1 %

	2002	2003	2004
Salaries % of total expenses	58 %	67 %	65 %
Other expenses % of total expenses	19 %	19 %	20 %
Rents % of total expenses	15 %	10 %	10 %

The first projects (ALVIS, MobiLife) funded by the European Union were launched during 2004. In the coming years, the share of EU-funded research is expected to rise considerably. The continuing high level of operational expenses is explained by investments to IT infrastructure (file server, computational grid).

7.2 Basic Research Unit Finances

The finances of the Basic Research Unit for 2004 are shown in the table below. For reference, the comparable numbers of 2003 are also given.

	2003	2004
Own basic funding	**631 800 €**	**631 800 €**
Academy of Finland	**285 000 €**	**496 000 €**
ProAct Program Coordination	90 000 €	92 960 €
ProAct –projects	140 000 €	141 000 €
Other projects	55 000 €	262 040 €
Tekes	**115 000 €**	**124 000 €**
Industry	**45 000 €**	**130 000 €**
Graduate schools	**260 000 €**	**247 500 €**
EU		**80 500 €**
Center of excellence funding from TKK	**75 000 €**	
Total	**1 411 800 €**	**1 709 800 €**

Appendices

A   Publications

Publications 2002–2004

	2002	2003	2004
Articles in international scientific journals with referee practice	15	21	19
Articles in international edited works and conference proceedings with referee practice	51	71	98
Articles in Finnish scientific journals with referee practice	2	2	1
Articles in Finnish edited works and conference proceedings with referee practice	9	6	5
Scientific monographs published abroad	4	3	4
Scientific monographs published in Finland	10	2	0
Other scientific publications	22	14	19
Computer programs (and algorithms)	0	8	2
Patents	1	0	0
Notifications on invention	2	1	0
Degrees			
PhD or DSc thesis	4	1	4
Licentiate thesis	1	0	1
Masters thesis	8	5	14

Articles in international scientific journals with referee practice

Abie, H.; Foyn, B.; Bing, J.; Blobel, B.; Pharow, P.; Delgado, J.; Karnouskos, S.; Pitkänen, O.; Tzovaras, D.: The need for a digital rights management framework for the next generation of e-goverment services. *Electronic Goverment*, 2004. Vol. 1, nro 1, 8-28.

J. Himberg, A. Hyvärinen, and F. Esposito: Validating the independent components of neuroimaging time-series via clustering and visualization. *NeuroImage*, 22(3):1214--1222, 2004.

P. O. Hoyer: Non-negative Matrix Factorization with sparseness constraints. *Journal of Machine Learning Research*, 5:1457--1469, 2004.

J. Hurri, J. Väyrynen, and A. Hyvärinen: Spatiotemporal receptive fields maximizing temporal coherence in natural image sequences. *Neurocomputing*, 58--60:815--820, 2004.

A. Hyvärinen and J. Hurri: Blind separation of sources that have spatiotemporal variance dependencies. *Signal Processing*, 84(2):247--254, 2004.

A. Hyvärinen, J. Hurri, and J. Väyrynen: A unifying framework for natural image statistics: spatiotempotal activity bubbles. *Neurocomputing*, 58--60:801--806, 2004.

E. Kettunen, S. Anttila, J.K. Seppänen, A. Karjalainen, H. Edgren, I. Lindström, R. Salovaara, A-M Nissén, J. Salo, K. Mattson, J. Hollmén, S. Knuutila, and H. Wikman: Differentially expressed genes in nonsmall cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer. *Cancer Genetics and Cytogenetics*, Vol. 149, Issue 2, pp. 98-106, March 2004. doi:10.1016/S0165-4608(03)00300-5.

M. Koivisto, and K. Sood: Exact Bayesian Structure Discovery in Bayesian Networks. *Journal of Machine Learning Research*, 5(May):549-573, 2004.

S. Koskenmies, E. Widen, P. Onkamo, P. Sevon, H. Julkunen, and J. Kere: Haplotype associations define target regions for susceptibility loci in systemic lupus erythematosus. *European Journal of Human Genetics* 12, 6 (2004), 489-494.

M. Kääriäinen, T. Malinen, and T. Elomaa: Selective Rademacher penalization and reduced error pruning of decision trees. *Journal of Machine Learning Research* 5 (2004), 1107-1126.
S. Luyssaert, M. Sulkava, H. Raitio, and J. Hollmén**:** Evaluation of forest nutrition based on large-scale foliar surveys: are nutrition profiles the way of the future? *Journal of Environmental Monitoring*, 6(2):160–167, 2004.

S. Luyssaert, M. Sulkava, H. Raitio, and J. Hollmén**:** Evaluation of forest nutrition based on large-scale foliar surveys: are nutrition profiles the way of the future? *Journal of Environmental Monitoring*, 6(2):160–167, 2004.

J. Makkonen, H. Ahonen-Myka, and M. Salmenkivi: Simple semantics in topic detection and tracking. *Information Retrieval* 7(3): 347-368, Kluwer Academic Publishers, 2004.

Oulasvirta, Antti; Saariluoma, Pertti: Long term working memory and interrupting messages in human-computer interaction. *Behavior & Information Technology*, 2004. Vol. 23, nro 1, 53-64.

Oulasvirta, Antti: Task-processing demands and memory in web interaction: A levels-of-processing approach. Interacting with Computers, 2004. Vol. 16, nro 2, 217-241.

M. Peyrard-Janvid, H. Anthoni, P. Onkamo, P. Lahermo, M. Zucchelli, N. Kaminen, K. Hannula-Jouppi, J. Nopola-Hemmi, A. Voutilainen, H. Lyytinen, and J. Kere: Fine mapping of the 2p11 dyslexia locus and exclusion of TACR1 as a candidate gene. *Human Genetics* 114, 5 (2004), 510-516.

J. Pitkäniemi, P. Onkamo, J. Tuomilehto, and E. Arjas: Increasing incidence of type I diabetes - role for genes? [electronic]. *BMC Medical Genetics* 5, 5 (2004).

Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A.: Understanding mobile contexts. Personal and Ubiquitous Computing, 2004. Vol. 8, 135-143.

Välimäki, Mikko; Hietanen, Herkko: The Challenges of Creative Commons Licensing. Computer Lav Review, 2004. Vol. 5, nro 6.

H. Wikman, J.K. Seppänen, V. K. Sarhadi, E. Kettunen, K. Salmenkivi, E. Kuosma, K. Vainio-Siukola, B. Nagy, A. Karjalainen, T. Sioris, J. Salo, J. Hollmén, S. Knuutila, and S. Anttila: Caveolins as tumor markers in lung cancer detected by combined use of cDNA and tissue microarrays. *The American Journal of Pathology*, Vol. 203, pp. 584-593, 2004.

Articles in international edited works and conference proceedings with referee practice

F. Afrati, A. Gionis, and H. Mannila: Approximating a collection of frequent sets. 10th International Conference on Knowledge Discovery and Data Mining (KDD) 2004.

P. Andritsos, R. J. Miller, and P. Tsaparas: Information-Theoretic Tools for Mining Database Structure from Large Data Sets. In SIGMOD, Paris, June 2004.

P. Andritsos, P. Tsaparas, R. J. Miller, and K. C. Sevcik: LIMBO: Scalable Clustering of Categorical Data. In 9th International Conference on Extending DataBase Technology (EDBT), Heraklion, Greece, 2004.

M. Bawa, H. Garcia-Molina, A. Gionis, and R. Motwani: The price of validity in dynamic networks. 23rd International Conference on Management of Data (SIGMOD) 2004.

E. Bertino, T. Leggieri, and E. Terzi: Securing DBMS: Characterizing and Detecting Query Floods. 7th Information Security Conference (ISC) 2004, Palo Alto, USA, September 27-29, 2004.

F. Bonchi, and B. Goethals: FP-Bonsai: The Art of Growing and Pruning Small FP-Trees. In Proceedings of the 8th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD 2004), Lecture Notes in Computer Science, Volume 3056, Springer-Verlag, Sydney, Australia, May 26-28, 2004, pp. 155-160.

Buntine, Wray; Löfström, Jaakko; Perkiö, Jukka; Perttu, Sami; Poroshin, Vladimir; Silander, Tomi; Tirri, Henry; Tuominen, Antti; Tuulos, Ville: A Scalable Topic-Based Open Source Search Engine. The IEEE/WIC/ACM Conference on Web Intelligence (WI 2004), Beijing, China, September 2004. Pp. 228-234. [http://cosco.hiit.fi/Articles/hiit-2004-14.pdf].

Buntine, Wray; Perttu, Sami; Tuulos, Ville: Using Discrete PCA on Web Pages. Statistical Approaches for Web Mining (SAWM), Pisa, Italy, 2004. Pp. 99-110. [http://cosco.hiit.fi/Articles/hiit-2004-15.pdf].

A. Bykowski, J.K. Seppänen, and J. Hollmén: Model-independent bounding of the supports of Boolean formulae in binary data. In R. Meo, P. L. Lanzi, and M. Klemettinen (eds.): Database Support for Data Mining Applications: Discovering Knowledge with Inductive Queries, LNCS 2682, Springer 2004. This is a revised version of the KDID'02 paper.

Davis, Marc; King, Simon; Good, Nathan; Sarvas, Risto: From Context to Content: Leveraging Context to Infer Media Metadata. The 12th International Conference on Multimedia (MM'04), New York, USA, ACM Press 2004. 188-195.

Davis, Marc; Good, Nathan; Sarvas, Risto: From Context to Content: Leveraging Context for Mobile Media Metadata. Workshop on Context Awareness at MobiSys 2004, 6.6.2004, Boston, MA, USA, pp. 1-6.

Davis, Marc; Sarvas, Risto: Mobile Media Metadata for Mobile Imaging. IEEE International Conference on Multimedia & Expo (ICME), 27.-30.6.2004, Taipei, Taiwan.

L. Eronen, F. Geerts, and H. Toivonen: A Markov chain approach to reconstruction of long haplotypes. In Pacific Symposium on Biocomputing (PSB 2004), pp. 104-115, Hawaii, USA, January 2004. World Scientific.

E. Falck, P. Floréen, P. Kaski, J. Kohonen, and P. Orponen: "Balanced data gathering in energy-constrained sensor networks". Algorithmic Aspects of Wireless Sensor Networks (First International Workshop, ALGOSENSORS 2004, Turku, Finland, July 16, 2004), Lecture Notes in Computer Science, Vol. 3121, Springer-Verlag, Berlin Heidelberg, 2004, pp. 59-70.

F. Geerts, B. Goethals, and T. Mielikäinen: Tiling Databases. In Proceedings of the 7th International Conference on Discovery Science (DS 2004). Lecture Notes in Artificial Intelligence, Volume 3245, Springer-Verlag, October 2-5, Padova, Italy, 2004, pp. 278-289.

F. Geerts, H. Mannila, and E. Terzi: Relational link-based ranking. 30th International Conference on Very Large Databases (VLDB) 2004, Toronto, Canada.

A. Gionis, H. Mannila, and J.K. Seppänen: Geometric and combinatorial tiles in 0-1 data. 8th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD) 2004.

A. Gionis, H. Mannila, and E. Terzi: Clustered segmentations. 3rd Workshop on Mining Temporal and Sequential Data (TDM) 2004.

B. Goethals: Memory Issues in Frequent Itemset Mining. In Proceedings of the 2004 ACM Symposium on Applied Computing (SAC'04), Nicosia, Cyprus, March 14-17, 2004, pp.530-534.

B. Goethals, and M. Zaki: Advances in Frequent Itemset Mining Implementations: report on FIMI'03. ACM SIGKDD Explorations Newsletter, Volume 6 , Issue 1 (June 2004), pp. 109-117.

N. Haiminen, and A. Gionis: Unimodal segmentation of sequences. In Proceedings of the 4th IEEE International Conference on Data Mining (ICDM), Brighton, UK, November 2004, pp. 106-113.

Hietanen, Herkko: Legal metadata and open content distribution. eBRF 2004 Conference, 20-22 September 2004, Tampere. [http://www.ebrc.info/kuvat/2108_04.pdf].

H. Hiisilä, and E. Bingham: Dependencies between transcription factor binding sites: comparison between ICA, NMF, PLSA and frequent sets. In Proceedings of the 4th IEEE International Conference on Data Mining, November 1-4, 2004, Brighton, UK, pp. 114-121.

Holi, Markus; Hyvönen, Eero: A Method for Modeling Uncertainly in Semantic Web Taxonomies. WWW2004, Alternate Trac Papers abd Posters, New York, USA. [http://www.cs.helsinki.fi/u/eahyvone/publications/www2004/ontoBayes_poster.pdf]

T. Honkela and A. Hyvärinen: Linguistic feature extraction using independent component analysis. In Proceedings of International Joint Conference on Neural Networks (IJCNN2004), Budapest, Hungary, 2004.

J. Hurri, J. Väyrynen, and A. Hyvärinen: Spatiotemporal linear simple-cell models based on temporal coherence and independent component analysis. In Proceedings of the Neural Computation and Psychology Workshop 8, Canterbury, UK, 2004.

A. Hyvärinen: Why would factors of components be non-normal and square-correlated? In Proceedings of Factor Analysis Centennial Symposium, Osaka, Japan, 2004. In press.

Hyvönen, Eero; Holi, Markus; Viljanen, Kim: Designing and Creating a Web Site Based on RDF Content. CEUR Workshop, New Yorkk, May 18th. [http://www.cs.helsinki.fi/u/ eahyvone/publications/swehg_www2004.pdf]

Hyvönen, Eero; Junnila, Miikka; Kettula, Suvi; Mäkelä, Eetu; Saarela, Samppa; Salminen, Mirva; Syreeni, Ahti; Valo, Arttu; Viljanen, Kim: Finnish Museums on the Semantic Web. User's Perspective on MuseumFinland. Museums and the Web 2004, Arlington, Virginia, USA, March 31 - April 3. [http://www.archimuse.com/mw2004/papers/hyvonen/hyvonen.html]

Hyvönen, Eero; Junnila, Miikka; Mäkelä, Eetu; Saarela, Samppa; Salminen, Mirva; Syreeni, Ahti; valo, Arttu; Viljanen, Kim: Publishing Museum Collections on the Semantic Web - the MuseumFinland Portal. WWW2004, New York, USA. [http://www.cs.helsinki.fi/u/ eahyvone/publications/www2004/museumFinland_poster.pdf]

Hyvönen, Eero; Saarela, Samppa; Viljanen, Kim; Mäkelä, Eetu; Valo, Arttu; Salminen, Mirva; Kettula, Suvi; Junnila, Miikka: A Cultural Community Portal for Publishing Museum Collections on the Semantic Web. 16th European Conference on Artical Intelligence (ECAI 2004), Workshop on Application of Semantic Web Technologies to Web Communities, Valencia, Spain, August 23.-27. 2004. [http://www.cs.helsinki.fi/u/eahyvone/publications/ MuseumFinlandHyvonen2004.pdf]

Hyvönen, Eero; Saarela, Samppa; Viljanen, Kim; Mäkelä, Eetu; Valo, Arttu; Salminen, Mirva; Kettula, Suvi; Junnila, Miikka: A Semantic Portal for Publishing Museum Collections on the Web. ECAI/PAIS 2004, Valencia, Spain.

Hyvönen, E.; Salminen, M.; Junnila, M.: Annotation of Heterogeneous Databasee Content for the Semantic Web. SemAnnot 2004, Hiroshima, Japan, Nov. 8, 2004. [http://www.cs.helsinki.fi/u/eahyvone/publications/iswc2004/contentCreation.pdf]

Hyvönen, Eero; Salminen, Mirva; Kettula, Suvi; Junnila, Miikka: A Content Creation Process for the Semantic Web. OntoLex 2004, Lisbon, Portugal. [http://www.cs.helsinki.fi/u/ eahyvone/publications/contentCreation.pdf]

Hyvönen, Eero; Valo, Arttu; Viljanen, Kim; Holi, Markus: A Logic-Based Semantic Web HTML Generator - A Poor Man's Publishing Approach. WWW2004, New York, Alternate Track Papers and Posters, USA, 2004. [http://www.cs.helsinki.fi/u/eahyvone/publications/ www2004/swehg_poster.pdf]

W. Hämäläinen, H. Toivonen, and V. Poroshin: Mining relaxed graph properties in internet. IADIS International Conference WWW/Internet 2004, Madrid, Spain, October 2004, pp. 152-159.

A. Kaban, E. Bingham, and T. Hirsimäki: Learning to read between the lines: The aspect Bernoulli model. In Proceedings of the 4th SIAM International Conference on Data Mining, April 22-24, 2004, Lake Buena Vista, Florida, pp. 462-466.

M. Koivisto, T. Kivioja, P. Rastas, H. Mannila, and E. Ukkonen: Hidden Markov modelling techniques for haplotype analysis. Algorithmic Learning Theory: 15th International Conference, ALT 2004, S. Ben-David, J. Case, and A. Maruoka (eds.), Lecture Notes in Computer Science, 3244, pp. 37-52, Springer, 2004.

Kontkanen, Petri; Myllymäki, Petri; Buntine, Wray; Rissanen, Jorma; Tirri, Henry: An MDL Framework for Data Clustering. In: Grünwald, P; Myung, I. J.; Pitt, M, Advances in Minimum Description Length: Theory and Applications. England, UK 2004, The MIT Press,

Kontkanen, Petri; Myllymäki, Petri; Roos, Teemu; Tirri, Henry; Valtonen, Kimmo; Wettig, Hannes: Probabilistic Methods for Location Estimation in Wireless Networks. In: Ganesh, Rajamani; Kota, Sastri; Pahlavan, Kaveh; Agustí, Ramón, Emerging Location Aware Broadband Wireless Adhoc Networks. Amsterdam 2004, Kluwer Academic Publishers, [http://cosco.hiit.fi/Articles/hiit-2004-3.pdf].

Kontkanen, Petri; Myllymäki, Petri; Roos, Teemu; Tirri, Henry, Valtonen, Kimmo; Wettig, Hannes: Topics in Probabilistic Location Estimation in Wireless Networks. The 15th IEEE Symposium on Personal, Indoor and Mobile Radio Communications, Barcelona, Spain, 2004. 2004, IEEE Press, [http://cosco.hiit.fi/Articles/hiit-2004-4.pdf].

Korpilahti, Tuomas; Hyvönen, Eero: An Architecture for Collaborative OntologyLibrary Development. 16th European Conference on Artificial Intelligence (ECAI 2004), Workshop on Application of Semantic Web Technologies to Web Communities, Valencia, Spain, August 23-27. [http://www.cs.helsinki.fi/u/eahyvone/publications/korpilahti.pdf]

Kurhila, Jaakko; Miettinen, Miikka; Nokelainen, Petri; Tirri, Henry: The Role of the Learning Platform in Student-Centered E-Learning. The 4th IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, August 2004. Pp. 540-544. [http://cosco.hiit.fi/edutech/publications/icalt2004_1.pdf]

Kurvinen, Esko; Oulasvirta, Antti: Towards socially aware pervasive computing: a turntaking approach. Proceedings of the 2nd International IEEE Conference on Pervasive Computing and Communications, Orlando, Florida, March 2004. pp. 346-351.

K. Laasonen, M. Raento, and H. Toivonen: Adaptive On-device Location Recognition. Pervasive Computing: Second International Conference (Pervasive 2004), LNCS 3001, 287-304, Vienna, Austria, April 2004. Springer Verlag.

Laukkanen, Mikko; Viljanen, Kim; Apiola, Mikko; Lindgren, Petri, Hyvönen, Eero: Towards Ontology-Based Yellow Page Services. CEUR Workshop, New York, May 18th. [http://www.cs.helsinki.fi/group/iwebs/publications/iwebs-www2004.pdf]

A. Leino: Computational Overview of Finnish lake names. Dzintra Hirsa (ed.), Onomastica Lettica. II. LU Latviesu valodas instituts.

A. Leino: Pikes and perches go together: A data-analytical view on Finnish lake names. In Marja Nenonen (ed.): Papers from the 30th Finnish Conference of Linguistics, pp. 79-84. Studies in Languages 39. University of Joensuu.

J. Lindgren and A. Hyvärinen: Learning high-level independent components of images through a spectral representation. In Proceedings of International Conference on Pattern Recognition ICPR2004), Cambridge, UK, 2004.

Lindholm, Tancred: Three-way Merge for XML Documents. DocEng 2004, ACM Symposium on Document Engineering, Wisconsin, October 28-30, 2004. 2004, ACM.

Lindström, Jan: Distributed Optimistic Concurrency Control for Real-Time Database Systems Proceedings of the 10th International Conference on Real-Time and Embedded Computing Systems and Applications. In: Lecture Notes on Computer Science. Göteborg, Sweden 2004, 664-682. [http://www.cs.helsinki.fi/u&jplindst/rtcsa04.pdf].

Lindström, Jan: Performance of Distributed Optimistic Concurrency Control in Real-Time Databases. The 7th International Conference on Information Technology, December 20-23, Hyderabad, India. Hyderabad, India 2004.

Lugano, Giuseppe; Nokelainen, Petri; Miettinen, Miikka; Kurhila, Jaakko; Tirri, Henry: On the Relationship Between Learners' Orientations and Activity in CSCL. The 4th IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, August 2004. Pp. 759-761. [http://cosco.hiit.fi/edutech/publications/icalt2004_2.pdf]

T. Mielikäinen, and Janne Ravantti: Sinogram Denoising of Cryo-Electron Microscopy Images. In A. Laganà, M.L. Gavrilova, V. Kumar, Y. Mun, C.J.K. Tan, and O. Gervasi (eds.): Computational Science and Its Applications - ICCSA 2004, International Conference, Assisi, Italy, May 14-17, 2004. Springer, 2004. (The article is not printed in the proceedings due to an editorial error.)

T. Mielikäinen: Separating structure from interestingness. In H. Dai, R. Srikant, and C. Zhang (eds.): Advances in Knowledge Discovery and Data Mining - 8th Pacific-Asia Conference, PAKDD 2004, Sydney, Australia, May 26-28, 2004, Proceedings. Volume 3056 of Lecture Notes in Artificial Intelligence, pages 476-485. Springer, 2004.

T. Mielikäinen: Discovery of Serial Episodes from Streams of Events. In M. Hatzopoulos, and Y. Manolopoulos (eds.): Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM 2004), 21-23 June, 2004, Santorini Island, Greece, pages 447-448. IEEE Computer Society, 2004.

T. Mielikäinen, J. Ravantti, and Esko Ukkonen: The computational complexity of orientation search in cryo-electron microscopy. In M. Bubak, G.D. van Albada, P.M.A. Sloot, and J.J. Dongarra (eds.): Computational Science - ICCS 2004: 4th International Conference, Kraków, Poland, June 6-9, 2004, Proceedings, Part I. Volume 3036 of Lecture Notes in Computer Science, pages 231-238. Springer, 2004.

T. Mielikäinen: Inductive databases as ranking. In Y. Kambayashi, M.K. Mohania, and W. Wöß (eds.): Data Warehousing and Knowledge Discovery, 6th International Conference, DaWaK 2004, Zaragoza, Spain, September 1-3, 2004, Proceedings. Volume 3181 of Lecture Notes in Computer Science, pages 149-158. Springer, 2004.

T. Mielikäinen: Privacy problems with anonymized transaction databases. In S. Arikawa, and E. Suzuki (eds.): Discovery Science - 7th International Conference, DS 2004, Padova, Italy, October 2-5, 2004, Proceedings, Volume 3245 of Lecture Notes in Artificial Intelligence.Springer, 2004, pp. 219-229.

T. Mielikäinen: Implicit enumeration of patterns. In B. Goethals, and A. Siebes (eds.): Proceedings of the Third International Workshop on Knowledge Discovery in Inductive Databases, pages 73-84. 2004.

T. Mielikäinen: An automata approach to pattern collections. In B. Goethals, and A. Siebes (eds.): Proceedings of the Third International Workshop on Knowledge Discovery in Inductive Databases, pages 85-96. 2004.

J. Muskens, O. Virtanen, M. Chaudron, and R. MacLaverty: Maintaining terminal integrity and context-aware reconfiguration. In the workshop of Component-Oriented Approaches to Context-Aware Computing at the European Conference on Object-Oriented Programming (ECOOP 2004), Oslo, Norway, June 2004.

Mäkelä, Eetu; Hyvönen, Eero; Saarela, Samppa; Viljanen, Kim: OntoViews – A Tool for Creating Semantic Web Portals. 3rd International Semantic Web Conference (ISWC 2004), Hiroshima, Japan. 2004, Springer-Verlag, [http://www.cs.helsinki.fi/u/eahyvone/publications/iswc/fmscocoon.pdf].

P. Nurmi: Modelling routing in wireless ad hoc networks with dynamic Bayesian games. In proceedings of the first IEEE International Conference on Sensor and Ad Hoc Communications and Networks (IEEE Secon), Santa Clara, California, US, October 2004.

Oulasvirta, Antti: Finding Meaningful Uses for Context-Aware Technologies: The Humanistic Research Strategy. ACM Conference on Human Factors in Computing Systems CHI 2004, Vienn, Austria, April 24-29, 2004. 2004, ACM Press, full paper 8 pages.

Oulasvirta, Antti; Salovaara, Antti: A cognitive meta-analysis of design approaches to interruptions in intelligent environments. Proceedings of the 2004 Conference on Human Factors in Computing Systems, Austria, Vienna. 1155-1158.

Oulasvirta, Antti; Salovaara, Antti: Six modes of proactive resource management: a user-centric typology for proactive behaviors. NordiCHI 04, November 23-27 2004, Tampere. New York, NY, USA 2004, ACM Press, 57-60.

Perkiö, Jukka; Buntine, Wray; Perttu, Sami: Exploring Independent Trends in a Topic-Based Search Engine. The IEEE/WIC/ACM Conference on Web Intelligence (WI 2004). Pp. 664-668. [http://cosco.hiit.fi/Articles/hiit-2004-12.pdf]

Pitkänen, O.; Saarinen, V.; Anttila, J.; Lauronen, P.; Välimäki, M.: Implementing a DRM Extension to an XML Browser. ODRL International Workshop 2004, Wien, Itävalta, 22.-23.4.2004. [http://www.hiit.fi/de/mobileipr/NoniusODRL2004.pdf]

M. Przybilski, S. Campadello, and T. Saridakis: Mobile, on Demand Access of Service-Annotated 3D Maps. The IASTED International Conference on Software Engineering,February 17-19, 2004 Innsbruck, Austria.

M. Raento: Mobile Communication and Context Dataset, Benchmarks and a Database for Context Recognition. Pervasive 2004 Workshop Proceedings (2004).

M. Raento: Kill your personal data dead. MobileHCI 04 Workshop on Location Systems Privacy and Control, Glasgow, Scotland, September 2004.

M. Raento: Ethic of choice for Location based system. In Proceedings of the Workshop on Usability, Utility and Ethics of Location Based Services, in NordiCHI 2004. University of Tampere, 2004.

Rantanen, Matti; Oulasvirta, Antti; Blom, Jan; Tiitta, Sauli; Mäntylä, Martti : InfoRadar: Group and public messaging in mobile context. NordiCHI'04, November 23-27 2004, Tampere.
New York 2004, ACM Press, 131-140. [http://www.cs.helsinki.fi/u/oulasvir/scipubs/omfpradar _nchi_camready_v03.pdf]

Reti, Tommo; Sarvas, Risto: The DiMaS System for Authoring Communities: Distributing Semantic Multimedia Content on Peer-to-Peer File Sharing Networks. Web Intelligence Symposium - STeP 2004; Vantaa, Finland, 2.-3.9.2004. [http://www.hiit.fi/u/reti/ publications/Reti_Web_Intelligence_Symposium_2004].

Reti, Tommo; Sarvas, Risto: DiMaS: System for Multimedia Producers to Distribute Semantic Content on Peer-to-Peer File Sharing Networks. Intellectual Property Beyond Rights Conference, Espoo, Finland, October 2004. [http://www.hiit.fi/u/reti/dimas/downloads/ Reti_Intellectual_Property_Beyond_Rights_2004.pdf]

Reti, Tommo; Sarvas, Risto: DiMaS: Distributing Multimedia on Peer-to-Peer File Sharing Networks. ACM Multimedia 2004; New York, N.Y., U.S.A., 10.-17.10.2004;. [http://www.hiit.fi/u/reti/publications/Reti_ACM_Multimedia_2004.pdf].

Saari, Timo; Turpeinen, Marko: Psychological Customisation of Information. Applications for Personalizing the Form of News. 54th Annual Conference of the International Communication Association, May 27-31 2004, New Orleans. USA, New Orleans 2004, International Communication Association.

Saari, Timo; Turpeinen, Marko: Towards Psychological Customization of Information for Individuals and Social Groups. In: Karat, J. and Karat. M.-C. (eds.), Personalization of User Experiences for eCommerce. Saksa 2004, Kluwer.

Saari, T., Turpeinen, M., Ravaja, N.; Laarni, J. and Kallinen, K.: Psychologically targeted persuasive advertising and product information presentation in eCommerce. ICEC 2004, the international conference for electronic commerce, Delft, The Netherlands, 25-27 October 2004. 22-36.

Saari, T., Turpeinen, M., Ravaja, N.; Laarni, J. and Kallinen, K.: Psychologically intelligent mobile multimedia messaging systems. SID 2004, Social Intelligence in Design, July 2004, Twente, Netherlands. 177-184.

M. Salmenkivi: Evaluating attraction in spatial point patterns with an application in the field of cultural history. In Proceedings of the Fourth IEEE International Conference on Data Mining (ICDM'04), Brighton, UK, November 2004.

M. Salmenkivi, and H. Mannila: Piecewise constant modeling of sequential data using reversible jump Markov chain Monte Carlo. In J. Wang, M. Zaki, H. Toivonen, and D. Shasha (eds.): Data Mining in Bioinformatics, Springer London Ltd., 2004, pp. 85-103.

Sarvas, Risto; Herrarte, Eric; Wilhelm, Anita; Davis, Marc: Metadata Creation for Mobile Images. The Second International Conference on Mobile Systems, Applications, and Services (MobiSys), Boston, MA, USA, 6.-9.6.2004 *Julkaisuvuosi: 2004. 36-48.

Sarvas, Risto; Viikari, Mikko; Pesonen, Juha; Nevanlinna, Hanno: MobShare: Controlled and Immediate Sharing of Mobile Images. The 12th International conference on Multimedia (MM'04), New York, NY, USA. ACM Press 2004. pp. 724-731.

Savia, E.; Kaski, S.; Tuulos, V.; Myllymäki, P.: On Text-Based Estimation of Document Relevance. International Joint Conference on Neural Networks 2004, Budapest, Hungary, July 25-29, 2004. pp. 3275-3280.

J. K. Seppänen, and H. Mannila: Dense Itemsets. In R. Kohavi, J. Gehrke, W. DuMouchel, and J. Ghosh (eds.): Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), pp. 683-688, Seattle, WA, USA 2004.

P. Sevon, H.T.T. Toivonen, and P. Onkamo: Gene Mapping by Pattern Discovery. In J. Wang et al (eds.), Data Mining in Bioinformatics, 105-126. Springer, 2004.

S. Shimizu, A. Hyvärinen, and Y. Kano: Independent component analysis and its application to causal analysis. In Proceedings of Factor Analysis Centennial Symposium, Osaka, Japan, 2004. In press.

M. Sulkava, J. Tikka, and J. Hollmén: Sparse regression for analyzing the development of foliar nutrient concentrations in coniferous trees. In S. Džeroski, B. Ženko, and M. Debeljak (eds.): Proceedings of the Fourth International Workshop on Environmental Applications of Machine Learning (EAML 2004), pp. 57–58, Bled, Slovenia, 2004.

Tarkoma, Sasu: Event Dissemination Service for Pervasive Computing. Pervasive 2004 Doctoral Colloquim, 18-19.4. Linz/Vienna, Austria. 2004, Austrian Computer Society, 155-160.

J. Tikka, and J. Hollmén: Learning linear dependency trees from multivariate time-series data. In ICDM 2004 Workshop on Temporal Data Mining: Algorithms, Theory and Applications (in conjunction with The Fourth IEEE International Conference on Data Mining), Brighton, UK, 2004. In Press.

P. Tsaparas: Using Non-Linear Dynamical Systems for Web Searching and Ranking. In Principles of Database Systems (PODS), Paris, June 2004.

Turpeinen, Marko; Saari, Timo: System Architecture for Psychological Customization of Communication Technology. 37th Hawaii International Conference on System Sciences (HICSS-37, 2004), Hawaii, USA 5.-8.1.2004. 2004, University of Hawaii.

Tuulos, Ville; Tirri, Henry: Combining Topic Models and Social Networks for Chat Data Mining. The IEEE/WIC/ACM Conference on Web Intelligence (WI 2004), Beijing, China, September 2004. Pp. 206-213. [http://cosco.hiit.fi/Articles/hiit-2004-13.pdf].

K. Vasko, H. Toivonen, and A, Korhola: Segmentation of paleoecological spatio-temporal count data. In Proceedings of the Fourth International Workshop on Environmental Applications of Machine Learning (EAML 2004), 61-62, Bled, Slovenia, Sep-Oct 2004.

J. Väyrynen, T. Honkela, and A. Hyvärinen: Independent component analysis of word contexts and comparison with traditional categories. In Proceedinsg of NORSIG 2004, Espoo, Finland, 2004.

Wilhelm, Anita; Takhteyev, Yuri; Sarvas, Risto; Van House, Nancy; Davis, Marc : Photo Annotation on a Camera Phone. Conference on Human Factors in Computing Systems (CHI), Wien, Itävalta 24.29.2004. 1403-1406.

Articles in Finnish scientific journals with referee practice

R. Pulkkinen, M. Salmenkivi, A. Leino, and H. Mannila: What was the Finnish _hiisi_? Applying Computational Methods to the study of Folk Religion. Temenos: Studies in Comparative Religion. Suomen Uskontotieteellinen Seura.

Articles in Finnish edited works and conference proceedings with referee practice

Hyvönen, E.; Kauppinen, T.; Salminen, M.; Ala-Siuru, P.: Proceedings of the 11th Finnish AI Conference, vol 2.. Älyä verkossa - Web Intelligence, Symposio Suomen Tekoälytutkimuksen päivillä 2004 (STeP 2004) 2.-3.9.2004, Tiedekeskus Heureka, Vantaa, Suomi. [http://www.cs.helsinki.fi/group/seco/conference/step2004/].

Kauppinen, Tomi; Hyvönen, Eero: Bridging the Semantic Gap Between Ontology Versions. Älyä verkossa - Web Intelligence, Tiedekeskus Heureka, Vantaa, 2.-3.9.2004. 2004, Finnish AI Society, Finland, 2004, [http://www.cs.helsinki.fi/u/eahyvone/publications/wis2004/bridge2004.pdf].

Nokelainen, Petri; Tirri, Henry: Bayesian Methods that Optimize Cross-cultural Data Analysis. In: Campbell, J. R.; Tirri, Kirsi; Ruohotie, Pekka; Walberg, H, Cross-cultural Research: Basic Issues, Dilemmas, and Strategies. Tampere 2004, Research Centre for Vocational Education, University of Tampere, pp. 141-158.

Oulasvirta, Antti: Developing future technologies that support social practices and community building. Psykocenter Symposium, Jyväskylä, Nov 3.-4., 2003, Jyväskylä, Finland. Jyväskylä 2004, Jyväskylä Psykocenter, 177-184.

Sarvas, Risto: Media Content Metadata and Mobile Picture Sharing. Älyä Verkossa - Web Intelligence Symposio, 2.-3.9.2004, Vantaa, Suomi.

Scientific monographs published abroad

R. Bayardo, B. Goethals, and M.J. Zaki (eds.): Proceedings of the FIMI'04 Workshop on Frequent Itemset Mining Implementations. November 1, 2004, Brighton, UK.

B. Goethals, and A. Siebes (eds.): Proceedings of the KDID'04 Workshop on Knowledge Discovery from Inductive Databases. September 20, Pisa, Italy.

J. Wang, M. Zaki, H. Toivonen, and D. Shasha (eds.): Data Mining in Bioinformatics. Springer 2004). ISBN 1-85233-671-4.

Zhong, N.; Tirri, Henry; Yao, Y.; Zhou, L.: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence. Beijing, China, September 2004.

Other scientific publications

A. Gurtov, A. D. Joseph, Friends or Rivals: Insights from Integrating HIP and i3, Workshop on HIP and Related Architectures, Washington D.C., USA, November 2004.

Hyvönen, Eero: Semanttinen Web - mitä se on käytännössä?. Helsinki: Helsingin yliopisto, tietojenkäsittelytieteen laitos, 2004. 38-42 (ATK-Tietotekniikkaa yliopistoille 2).

G. Iachello, M. Raento, and I.E. Smith (eds.): MobileHCI 2004: workshop on location systems privacy and control: held at MobileHCI 04, 13 September 2004, University of Strathclyde in Glasgow, Scotland. University of Strathclyde, 2004.

J. Kohonen: Data Gathering in Sensor Netowkrs. Proceedings of the proactive computing workshop PROW: 2004 25-26 November 2004, Helsinki, Finland, pp. 32-36.

M. Kääriäinen: Relating the Rademacher and VC Bounds. Report C-2004-57, Department of Computer Science, Department of Computer Science, University of Helsinki, 2004.

M. Koivisto, P. Rastas, and E. Ukkonen: Recombination systems. In J. Karhumaki, H. Maurer, G. Paun, G. Rozenberg (eds.): Theory is Forever (Salomaa Festschrift). Lecture Notes in Computer Science, 3113, pp. 159-169, Springer-Verlag, Berlin, Heidelberg, 2004.

A. Korhola, H. Toivonen, and K. Vasko: Paleoekologia – kadonneen aarteen metsästys. Tietoyhteys 3/2004 (lokakuu), 26-28.

K. Laasonen: Where Are You Going? Predicting User Movement from Cellular Data. Proceedings of the proactive computing workshop PROW: 2004 25-26 November 2004, Helsinki, Finland, 121-124.

G. Lindén (ed.): Proceedings of the Proactive Computing Workshop PROW 2004, Helsinki 25-26 November 2004, Helsinki, University Press, 2004.

T. Mielikäinen, and E. Ukkonen: The Complexity of Maximum Matroid-Greedoid Intersection and Weighted Greedoid Maximization. Report C-2004-2. Department of Computer Science, University of Helsinki, 2004.

T. Mielikäinen, J. Ravantti, and E. Ukkonen: The Computational Complexity of Orientation Search Problems in Cryo-Electron Microscopy. Report C-2004-3. Department of Computer Science, University of Helsinki, 2004.

Nokelainen, Petri; Miettinen, Miikka; Kurhila, Jaakko; Floréen, Patrik; Tirri, Henry: A Shared Document-Based Annotation Tool to Support Learner-Centered Collaborative Learning. British Journal of Educational Technology, 2004.

R. Pulkkinen, M. Salmenkivi, A. Leino, and H. Mannila: What was the Finnish hiisi? Applying computational methods to the study of folk religion. Temenos, 2004. In press.

M. Raento: Context Recognition by User Situation Data Analysis --- the Context prototype software. 12th International Symposium on Electronic Arts ISEA 2004, Helsinki, August 2004.

M. Raento: Context software: a prototype platform for contextual mobile applications. Proceedings of the proactive computing workshop PROW: 2004 25-26 November 2004, Helsinki, Finland, 103-111.

Rimey, Ken: Version Headers for Flexible Synchronization and ConflictResulution. Helsinki: TKK/HIIT, 2004. 12 (HIIT Technical Reports 2004-3). [http://www.hiit.fi/publications/pub_files/hiit-2004-3.pdf].

Roto, Virpi; Oulasvirta, Antti; Haikarainen, Tuulia; Lehmuskallio, Harri; Nyyssönen, Tuomo: Examining mobile phone use in the wild with quasi-experimentation. Helsinki: 2004. (HIIT Technical report 2004-1, August 13).

Sarvas, Risto; Turpeinen, Marko; Jokela, Sami (eds.): New Business in Computer-Mediated Communities. Espoo: Helsinki University of Technology, Software Business and Engineering Institute, 2004. 137 p. (Teknillisen korkeakoulun ohjelmistoliiketoiminnan ja -tuotannon laboratorion tekninen raportti C7).

Soininen, Aura: Yhdysvaltain patenttijärjestelmä murroksessa. IPRInfo, 2004. Nro 3, 23-25. [http://www.iprinfo.com/showPage.php].

Computer programs (and algorithms)

M. Raento and R. Petit: ContextPhone: a prototyping platform for context-aware applications on Symbian phones (http://www.cs.helsinki.fi/group/context/)

L. Eronen: HaploRec: haplotyping software for large genetic data sets (http://www.cs.helsinki.fi/group/genetics/haplotyping.html)

Degrees

PhD thesis

M. Koivisto: Sum-Product Algorithms for the Analysis of Genetic Risks. PhD Thesis. University of Helsinki, Department of Computer Science. 2004.

M. Kääriäinen: Learning small trees and graphs that generalize. PhD Thesis. University of Helsinki, Department of Computer Science. 2004.

P. Sevon: Algorithms for association-based gene mapping. PhD Thesis. University of Helsinki, Department of Computer Science. 2004.

G. Jacucci: Cases of Configuring Physical Interfaces in Mixed Media in the public dissertation event. PhD Thesis. University of Oulu, Department of Computer Science. 2004.

Licentiate thesis

A. Oulasvirta: Human-computer interaction in mobile context: a cognitive resources perspective. Licentiate Thesis. University of Helsinki, Department of Psychology, Cognitive Science. 2004.

Master's thesis

L. Eronen: Models and algorithms for haplotype reconstruction on genetically long marker maps. Master's Thesis. University of Helsinki, Department of Computer Science. 2004.

N. Haiminen: Discovering Recurrent Sources in Genomic Sequences. Master's thesis. University of Helsinki, Department of Computer Science. 2004.

F. Herrera: Value configuration in the European mobile multiplayer gaming industry. Master's thesis, Helsinki University of Technology, Department of Industrial Engineering and Management, 2004.

M. Holi: A Method for Modelling Uncertainty in Semantic Web Ontologies. Master's thesis, University of Helsinki, Department of Computer Science. 2004.

T. Kauppinen: An Ontology Versioning Framework. Master's thesis, University of Helsinki, Department of Computer Science. 2004.

M. Komu: Application Programming Interfaces for the Host Identity Protocol. Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering, 2004.

T. Korpilahti: Architecture for Distributed Development of an Ontology Library. Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering. 2004.

J. Muhonen: Non-Redundant Association Rules – Deriving the Confidence of an Association Rule. Master's Thesis. University of Helsinki, Department of Computer Science. 2004.

S. Saarela: Näkymäpohjainen RDF-haku. Master's thesis, University of Helsinki, Department of Computer Science. 2004.

J. Saijos: Kontekstitietoinen tiedonhaku. Master's thesis. University of Helsinki, Department of Computer Science. 2004.

T. Sidoroff: Semanttiset portaalit. Master's thesis, University of Helsinki, Department of Computer Science. 2004.

K. Slavov: Implementing HIP Algorithms in Linux Kernel. Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering, 2004.

A. Varis: WordNet sanatietokannan hyödyntäminen luonnollisen kielen sovelluksissa. Master's thesis, University of Helsinki, Department of Computer Science. 2004.

M. Viikari: Kamerapuhelimella otettujen kuvien jakaminen lähipiirin kesken, Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering, 2004.

B   List of Personnel

Apiola, Mikko[1]	Research Assistant	ARU
Bingham, Ella	Researcher	BRU
Buntine, Wray	Senior Research Scientist	ARU
Eronen, Lauri	Researcher	BRU
Floréen, Patrik	Research Coordinator	ARU/BRU
Geerts, Floris[2]	Researcher	BRU
Gionis, Aristides	Researcher	BRU
Goethals, Bart [2]	Researcher	BRU
Guo, Hang[2]	Researcher	ARU
Gurtov, Andrei[2]	Research Scientist	ARU
Haiminen, Niina	Researcher	BRU
Hakli, Raul	Researcher	ARU
Hassel, Henri[1]	IT Specialist	ARU
Hasu, Tero	Researcher	ARU
Heikkilä, Juho	Researcher	ARU
Heino, Jaana	Researcher	BRU
Herrera Avila, Fernando	Researcher	ARU
Hietanen, Herkko[1]	Researcher	ARU
Hiisilä, Heli	Researcher	BRU
Himanen, Pekka[2]	Senior Research Scientist	ARU
Hinneburg, Alexander[2]	Researcher	BRU
Hintsanen, Petteri	Researcher	BRU
Hollmén, Jaakko	Professor	BRU
Hoyer, Patrik	Researcher	BRU
Hurri, Jarmo	Researcher	BRU
Hyvärinen, Aapo	Senior Research Fellow (Academy of Finland)	BRU
Hyvönen, Eero[1]	Principal Scientist	ARU
Hyvönen, Saara	Researcher	BRU
Jacucci, Giulio[2]	Researcher	ARU
Kanerva, Pekka	Researcher	ARU
Kangasharju, Jaakko	Researcher	ARU
Kankainen, Tomi[1]	Researcher	ARU
Kantola, Vesa[3]	Researcher	ARU
Karp, Elina[3]	Assistant Researcher	BRU
Kaski, Sami	Researcher	BRU
Kemppinen, Jukka[1]	Principal Scientist	ARU
Kettula, Suvi-Päivikki[2]	Researcher	ARU
Kohonen, Jukka	Researcher	BRU
Koivisto, Mikko	Researcher	BRU
Kollin, Jussi	Assistant Researcher	BRU
Komu, Miika	Researcher	ARU
Kontiainen, Mikko	Researcher/Coordinator	ARU
Kontkanen, Petri[2]	Researcher	ARU
Koponen, Teemu[3]	Researcher	ARU
Kortesniemi, Yki[2]	Researcher	ARU
Kousa, Mika	Research Assistant	ARU
Kuikkaniemi, Kai[1]	Research Assistant	ARU
Kujala, Inka	Project Assistant	BRU
Kujala, Teija	Researcher	BRU
Kulikova, Yevgeniya	Researcher	BRU

Kurki, Ilmari	Researcher	BRU
Kurppa, Teemu[2]	Researcher	BRU
Kurvinen, Esko[3]	Researcher	ARU
Kääriäinen, Matti	Researcher	BRU
Köster, Urs[2]	Researcher	BRU
Laasonen, Kari	Researcher	BRU
Lacoste-Julien, Simon[2]	Researcher	ARU
Lahtinen, Jussi[2]	Researcher	ARU
Lehdonvirta, Vili[3]	Research Assistant	ARU
Leino, Antti	Researcher	BRU
Lepola, Tuomas	Researcher/ HeCSE	ARU
Lindén, Greger	Co-ordinator	BRU
Lindgren, Petri[1]	Research Assistant	ARU
Lindholm Tancred	Researcher	ARU
Lindström, Jan[2]	Senior Research Scientist	ARU
Löfström, Jaakko	Research Assistant	ARU
Mannila, Heikki	Academy Professor	BRU
Mattila, Juhana[2]	IT Specialist	ARU
Mehto, Katri[3]	Researcher	ARU
Mielikäinen, Taneli	Researcher	BRU
Miettinen, Miikka	Researcher	ARU
Miettunen, Pirkko	Institute Secretary	ARU
Mononen, Tommi	Researcher	ARU
Muhonen, Juho	Researcher	BRU
Muilu, Juha[1]	Database Manager	BRU
Myllymäki, Petri[3]	Principal Scientist	ARU
Mäkelä, Eetu	Research Assistant	ARU
Mäkelä, Mikko[2]	Research Assistant	ARU
Mäntylä, Martti	Research Director	ARU
Mäntylä, Teemu[1]	Research Assistant	ARU
Nikander, Pekka	Ericsson Visiting Senior Research Scientist	ARU
Nikko, Taina	Project Secretary	ARU
Nurmi, Petteri	Researcher	BRU
Obozinski, Guillaume[2]	Researcher	ARU
Oksanen, Kenneth[2]	Researcher	ARU
Oksanen, Ville[3]	Researcher	ARU
Onkamo, Päivi	Researcher	BRU
Oulasvirta, Antti	Researcher	ARU
Patrikainen, Anne	Researcher	BRU
Pelkonen, Ari-Pekka[2]	IT- Administrator	ARU
Pennanen, Merja[3]	Research Assistant	ARU
Perkiö, Jukka	Researcher	ARU/BRU
Perttu, Sami[1]	Researcher	ARU
Petit, Renaud	Researcher	BRU
Pitkänen, Olli	Program Coordinator	ARU
Porochine, Vladimir	Research Assistant	ARU
Przybilski, Michael[2]	Researcher	BRU
Puolakka, Anton[3]	Program Coordinator	ARU
Puolamäki, Kai[2]	Researcher	BRU
Päivärinta, Juha	Research Assistant	ARU
Raatikainen, Kimmo[1]	Principal Scientist	ARU
Raento, Mika	Researcher	BRU
Rantanen, Matti	Researcher	ARU
Rantanen, Mervi[2]	Planning Officer	ARU

Reti, Tommo	Researcher	ARU
Rimey, Kenneth	Senior Research Scientist	ARU
Rissanen, Jorma[1]	HIIT Fellow	ARU
Riva, Oriana[2]	Researcher	ARU
Roos, Teemu[1]	Researcher/HeCSE	ARU
Rüger, Torsten	Researcher	ARU
Ruosaari, Salla	Researcher	BRU
Ruotsalainen, Laura[2]	Researcher	BRU
Saarela, Samppa[1]	Research Assistant	ARU
Saaresto, Marko	Research Assistant	ARU
Saari, Timo[1]	Senior Research Scientist	ARU
Saarinen, Päivi[2]	Planning Officer	ARU
Salmenkivi, Marko	Researcher	BRU
Salminen, Mirva[3]	Research Assistant	ARU
Salovaara, Antti	Researcher	ARU
Seppänen, Jouni	Researcher	BRU
Sevon, Petteri	Researcher	BRU
Shimizu, Shohei	Researcher	BRU
Sidoroff, Teemu[2]	Research Assistant	ARU
Silander, Tomi	Researcher	ARU
Slavov, Kristian[3]	Researcher	ARU
Soininen, Aura[3]	Researcher	ARU
Sood, Kismat[3]	Researcher	BRU
Stenborg, Markku[2]	Senior Research Scientist	ARU
Sädekallio, Outi[3]	Research Assistant	ARU
Tarkoma, Sasu	Researcher	ARU
Tatti, Nikolaj	Researcher	BRU
Terzi, Evimaria	Researcher	BRU
Thalainayar Balasubramanian, Ramya	Research Assistant	ARU
Tiitta, Sauli	Researcher	ARU
Tirri, Henry[2]	Nokia Visiting Principal Scientist	ARU
Toivonen, Hannu	Professor	BRU
Tonteri, Pekka	IT Manager	ARU
Tsaparas, Panayiotis[2]	Researcher	BRU
Tuominen, Antti	Researcher	ARU
Tuominen, Heikki[3]	Researcher	BRU
Turpeinen, Marko[1]	Senior Research Scientist	ARU
Tuulos, Ville	Researcher	ARU
Ukkonen, Antti	Researcher	BRU
Ukkonen, Esko	Research Director	BRU
Valtonen, Kimmo	Researcher	ARU
Viljanen, Kim[1]	Research Assistant	ARU
Virtanen, Otso[2]	Researcher	BRU
Virtanen, Perttu	Project Manager	ARU
Vuorenmaa, Janne[1]	Research Assistant	ARU
Wettig, Hannes[2]	Researcher	ARU
Ylitalo, Katri[2]	Researcher	ARU
Zhou, Qiang[2]	Researcher	ARU

---

[1] working part time
[2] working part of the year
[3] working part time and part of the year