Helsinki Institute for Information Technology HIIT Annual Report 2003

Mervi Rantanen, Martti Mäntylä, and Heikki Mannila (eds.)

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 (HUT). 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 is the Research Director of the BRU.

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

2 Review of Year 2003

2.1 Advanced Research Unit (Professor Martti Mäntylä)

The year 2003 was the fourth calendar year of operation of the Advanced Research Unit (ARU). Whereas the previous years were characterised by rapid growth and the launching of various types of activity, the year 2003 was marked by organic development within the already established structure. Yet also a major new form of operation was launched during the year with the first meeting of the Scientific Advisory Board of HIIT, a joint entity of the ARU and the BRU. The meeting provided important guidance and feedback on the work of the entire institute. A more detailed report on the meeting is given in Section 6.3 of this document.

Several initiatives launched originally in 2002 and earlier ripened and bore fruit during 2003. A major sample is the Next Generation Information Retrieval (NGIR) line of research launched within the Intelligent Systems research area led by Professor Henry Tirri. The research line is executed in three closely related projects: PROSE funded by the Academy of Finland, Search-in-a-Box (SIB) funded by the National Technology Agency of Finland (Tekes) and companies and ALVIS funded by the Commission of European Communities within its 6th framework programme (prepared during 2003, launched 1.1.2004). This focusing of effort means that the NGIR line of work will shape the entire ARU for years to come.

Another new line of research was also established during 2003 with the launching of the Mobile Content Communities (MC2) project under the leadership of Dr. Marko Turpeinen, who had joined HIIT in late 2002. The project is ARU's first entry in community-oriented computing, and as such an important extension of the themes addressed in the Media Convergence research area. With MC2, also the new Digital Content Communities research group was established in the ARU.

NGIR and MC2 demonstrate an important qualitative development in how the ARU conducts its work in that both are executed by consortia covering a wide range of Finnish and international research partners such as University of Tampere (UTA), Helsinki School of Economics, University of California at Berkeley (USA), Media Lab Europe (Ireland), and Tsinghua University (China). In particular, the ALVIS EU project will be executed by a consortium of ten European and one Chinese partner, coordinated by the ARU. Of course, extensive networking with experts elsewhere gives an opportunity for the ARU to focus on its own competences and research groups.

EU activities were also otherwise an active thread of ARU's work throughout the year 2003, in particular participation in the preparation of the MobiLife integrated project coordinated by the Nokia Research Center. The fact that thanks to the experience gained with ALVIS, ARU now possesses the capability to coordinate EU research projects when needed, naturally offers useful flexibility also in the future.

Close co-operation with the University of California at Berkeley (UC Berkeley) has been an important part of HIIT's operation since 2000 when the first direct research co-operation was launched. The co-operation also flourished in 2003, and expanded to cover new areas in addition to existing Mobile Internet and Digital Economy research. In particular, both SIB and MC2 projects involve active UC Berkeley co-operation, thus expanding the interface between HIIT and UC Berkeley. These new activities are also likely to include two-way exchange of research personnel, to commence during 2004.

Based on the successful experience with UC Berkeley, during 2003 the ARU started to build a similar connection to a leading Chinese university, the Tsinghua University in Beijing. As a result, the Department of Computer Science and Engineering of Tsinghua (Professor Zhou Lizhu) decided to join the ALVIS and SIB consortia, with expectation of personnel exchange to commence in 2004. Preparation of other joint activities also commenced during 2003, and continues in 2004.

Yet another international activity in 2003 was the organisation and execution of the First International Mobile IPR Workshop in Helsinki in August, in association with the MobileIPR project in the Digital Economy research area. The workshop had some 60 participants, including more than 20 foreign attendees.

Two new doctoral-level researchers joined the ARU during 2003. Dr. Jan Lindström joined the Future Mobile and Ubiquitous Computing (Fuego) research area as a postdoctoral researcher, focusing in particular on mobile data. Dr. Timo Saari agreed to join HIIT and become the co-leader (with Professor Martti Mäntylä) of the User Experience Research Group, while he also continues as a principal scientist at the Center for Knowledge and Innovation Research (CKIR) of the Helsinki School of Economics. The resulting "personal union" between activities in ARU and CKIR is likely to bring further fruit in the future.

The general atmosphere of the ICT industry continued to be challenging throughout the year 2003. The business news were dominated by announcements of losses and cuts - not least of research budgets. Even so, the ARU was able to maintain and extend its position in a highly competitive environment, thanks to the continued devotion and commitment of ARU personnel. Tekes also deserves our gratitude for its increased flexibility and commitment, reflected in increased funding shares aimed to maintain Finnish research competences also through difficult times. As a result, the ARU is well positioned to continue fulfilling its mission in the future – a future that we look forward to.

2.2   Basic Research Unit (Professor Heikki Mannila)

Year 2003 was the second year of operation of the Basic Research Unit (BRU). The basic research themes established in 2002, data analysis and adaptive/proactive computing, were developed throughout the year, while a third theme, neuroinformatics, was added in August 2003.

In the area of data analysis, the work in theory and applications continued. 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. On the theoretical side, advances were made especially in the study of sequence segmentation and in the discovery of frequent patterns. On the applications side, the work on onomastics, computational genetics, gene expression data analysis, and genome structure continued, producing interesting results for the applications sciences and new methodological themes for the theoretical research. A new collaboration with researchers in paleontology yielded promising first results. New themes were also initiated in the area of spatial and spatio-temporal data mining, where collaboration started in the area of analyzing spatial data on bird nesting.

In the area of adaptive (or proactive) computing managed by Dr. Patrik Floréen, the bulk of the work was connected to the Proactive Computing (PROACT) research programme of the Academy of Finland. BRU is in charge of the coordination of the research programme, and two projects (Context, NAPS) receive their funding from the programme. Professor Hannu Toivonen heads the Context project, and Patrik Floréen heads NAPS. The adaptive

computing research theme also started an industrial project SPACE4U, funded by Nokia Research Center.

Dr. Aapo Hyvärinen joined the BRU in May 2003, and in August 2003 he officially initiated the third research theme of the BRU, neuroinformatics. The goal of the work is to build models of brain function and to develop new multivariate statistical techniques. Dr. Hyvärinen is one of the leading experts in ICA (independent component analysis), and in 2003 he quickly built a team of about six persons.

Year 2003 contained some nice successes for the BRU in the funding arena. The EU-project APrIL II was formally approved, and another (MobiLife) advanced to the final negotiations. In the Academy of Finland research programme for systems biology and bioinformatics BRU got funding for two projects, one consortium headed by Heikki Mannila and a project headed by Jaakko Hollmén. The Academy also granted an academy professorship to Heikki Mannila.

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

International recruiting has been one of the most important goals of the BRU. We were happy to be able to recruit Drs. Bart Goethals and Floris Geerts from Limburg University (Belgium) and Dr. Aristides Gionis from Stanford University (USA). The postdoctoral researchers from abroad contribute considerably to the research environment at BRU, and the hiring from elsewhere will continue in the future.

The BRU operates at the Department of Computer Science at the UH and at the Laboratory of Computer and Information Science at HUT. The operation at two sites works well, and the different backgrounds of the departments provide fertile ground for the development of new types of methods.

One of the basic principles of BRU is active participation in teaching. In this, year 2003 was a success, as the links to teaching were strengthened at both universities.

3 Important Dates

8-9 April: The Digital Economy (DE) Research Group of ARU organised an invitation-based workshop on Digital Divide and Digital Commons.

26-28 May: HIIT organised a Retreat meeting in the Sjökulla Education Center, Kirkko-nummi, for representatives from industry and for HIIT senior researchers.

18-20 August: The third HIIT-Berkeley Summer School on Telecommunications Software Architectures organised by HIIT and the UC Berkeley was held at Berkeley, USA. Main organisers of the Summer School were Professor Kimmo Raatikainen and Professor Randy Katz.

27-28 August: HIIT organised The First International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet. The workshop was organised in Helsinki with the support of Tekes through the Mobile IPR project. The keynote speakers were Professor Hal Varian (UC Berkeley) and Professor Ross Anderson (Cambridge).

2 October: ARU held a Research Portfolio Seminar presenting project ideas to be developed further for submission to various funding organisations in 2003.

14 October: The new rectors Matti Pursula of HUT and Ilkka Niiniluoto of UH visited ARU.

16-17 October: The first HIIT Scientific Advisory Board meeting was held in Helsinki.

23 October: The SIMS-BCIS-HIIT Seminar on Information Society Research was held in Berkeley, California. The seminar was co-sponsored by the School of Information Management and Systems (SIMS), the Berkeley Center for the Information Society (BCIS), and HIIT. Speakers included Professor Hal Varian, Professor Annalee Saxenian, and Professor Marc Davis from UC Berkeley; Dr. Pekka Himanen from BCIS/HIIT; and Professor Henry Tirri and Dr. Marko Turpeinen from HIIT.

21 November-3 December: The Director of ARU, Professor Martti Mäntylä visited China as a member of the delegation accompanying Mr Mauri Pekkarinen, Minister of Trade and Industry. Apart from the official program in Beijing, Shanghai, and Hong Kong, Professor Mäntylä also visited the Tsinghua University where he gave a guest lecture and discussed potential future co-operation areas.

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 were as follows: Mobile Computing (Prof. Kimmo Raatikainen, Dr. Ken Rimey, Dr. Pekka Nikander, Dr. Jan Lindström); User Experience Research (Dr. Timo Saari, Dr. Anu Kankainen, Dr. Jan Blom, Prof. Martti Mäntylä); Complex System Computation (Prof. Henry Tirri, Prof. Petri Myllymäki, Dr. Wray Buntine, Dr. Jorma Rissanen, Dr. Patrik Floréen); Digital Content Communities (Dr. Marko Turpeinen, Dr. Jan Blom, Dr. Timo Saari); Interactive Digital Media (Prof. Petri Vuorimaa); Semantic Computing (Prof. Eero Hyvönen); and Digital Economy (Prof. Jukka Kemppinen, Dr. Pekka Nikander, Dr. Pekka Himanen, Prof. 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 "killer applications" 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 anytime 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 challenge 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; Rantanen, Matti; 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; More Magic Software
Publications: 43, 63, 72, 96, 97, 116
Keywords: Mobile wireless Internet, middleware services, mobile computing, adaptive applications

Abstract and Progress in 2003

The objective of this research project is to specify the set of fundamental enabling middleware services for mobile applications on future environments and to implement two research prototypes. The project has four focus areas, which describe the content of the research: Adaptive Applications, Dynamic Reconfigurable Services, Mobile Distributed Information Base, and Mobility, Multi-Homing, and Cryptographic Host Identification. 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 monitors and contributes to relevant middleware standardization bodies; for example W3C, IETF, and OMG.

During 2003 the project implemented the first middleware service set and experimented with the prototype implementation. The first middleware service set featured 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). A second service set will be specified and implemented in 2004.

4.1.1.2 Personal Distributed Information Store (PDIS)

Project leader: Rimey, Kenneth
Research group(s): Mobile Computing
Researchers: Rüger, Torsten; Oksanen, Kenneth; Kanerva, Pekka; Hasu, Tero
Schedule: 2003-01-01 ... 2004-12-31
Funding: Tekes; Nokia; Creanor Movial; Innofactor; HP
Keywords: Data synchronization, mobile computing, peer-to-peer computing

Abstract and Progress in 2003

The PDIS project is about next-generation data synchronization. Its central goal is to enable people to store replicas of their data on several devices, and several computers, and to keep these replicas in sync. More specifically, the project is creating a new kind of system service

for storing structured data in an application-independent way. The PDIS repository looks to applications much like a local XML database, but it additionally implements a versioning and labeling scheme that enables synchronizing the contents of any two repositories at any time. Changes made at one repository propagate in an epidemic fashion to the rest.

PDIS models data as collections of objects, where each object is represented as a small XML document. This is ideal for everyday Personal Information Management applications such as calendaring, as well as for storing file metadata, as in digital media applications. For instance, calendars are simply collections of calendar entries, represented as little XML documents.

4.1.1.3 Context Recognition by User Situation Data Analysis (CONTEXT)

Project leader: Mäntylä, Martti
Research group(s): User Experience
Researchers: Kankainen, Anu; Tiitta, Sauli; Rantanen, Matti; Oulasvirta, Antti
Schedule: 2002-11-01 ... 2005-12-31
Cooperation units: BRU/HIIT
Funding: Academy of Finland
Publications: 12, 32, 33, 34, 35, 88, 95, 99, 129
Keywords: Proactive computing, ubiquitous computing, presence service, instant messaging, data analysis, data mining
Research programme: Academy of Finland / Proactive Computing (PROACT)

Abstract and Progress in 2003

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 to 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 messaging, a representative ubiquitous application whose usability greatly depends on how context-sensitively communication is 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 aimed at 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 will also produce 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. In 2003, the project built a working prototype of proactive context-communication, running on Nokia Series 60 mobile phone. The prototype is going to be subjected to a series of longitudinal field studies in 2004 to examine its effects on group communication behavior. At the same time, ethnomethodological studies by the research group revealed how social and temporal contexts could be modelled in proactive computers. Moreover, a humanistic research strategy was developed for pinpointing the traditional interaction design process for mobile use contexts.

4.1.1.4 European Virtual Testbed for Location–Sensitive Ubiquitous Services (EVIUS)

Project leader: Mäntylä, Martti
Research group(s): Mobile Computing, Digital Economy
Researchers: Kontiainen, Mikko
Schedule: 2002-06-01 ... 2003-05-31
Cooperation units: Industrial Information Technology Laboratory (HUT); Otaverkko Oy; VTT Technical Research Centre of Finland; Finnish Geodetic Institute
Funding: Tekes; Navinova Oy
Keywords: Mobile services, location-sensitive services

Abstract and Progress in 2003

The ultimate objective of the project was to prepare the establishment of a Europe-wide network of test sites for future mobile services. The project was built on the basis of the NAVI-programme (2000-2002) that focused on location-based services. Building on the basis of the NAVI network established in the programme, the project aimed at (i) establishing a testbed in the Helsinki area for new mobile services by networking manufacturers, operators, content providers, and service builders; (ii) linking this activity with other Finnish testbeds such as the Octopus testbed in Oulu to establish a coordinated entity termed "Testbed Finland"; and (iii) seeking contact with other European similar test sites to provide facilities for multi-site testing. The envisaged testing environment was to support service testing from multiple viewpoints, including technical interoperability, business interoperability, and end-user acceptance. The project was completed in May 2003; however, the networking activities of the companies involved in the project has continued also after the project in the "Helsinki Area Testbed" consortium.

4.1.1.5 Ubicomp Bubbles Enhancing Human–Human and Human–Computer Interaction (Between)

Project leader: Mäntylä, Martti
Research group(s): User Experience
Researchers: Kankainen, Anu; Kankainen, Tomi; Oulasvirta, Antti; Rantanen, Matti; Tiitta, Sauli; Blom, Jan; Samarine, Mikhail
Schedule: 2001-06-01 ... 2003-03-31
Cooperation units: Industrial Information Technology Laboratory (HUT); Department of Industrial and Strategic Design (University of Art and Design, Helsinki)
Funding: Tekes; Nokia; TeliaSonera Finland; Elisa Communications; Alma Media; Sanoma-WSOY
Publications: 64, 98, 124, 147
Keywords: Ubiquitous computing, mobile computing, scenario methods, concept design, user-centric product design
Research programme: Tekes / NETS

Abstract and Progress in 2003

The aim of the Between project was to study and create ubiquitous computing product and service scenarios and prototypes that would meet people's needs in home, work and public contexts. The research methodology of the project followed the principles of user-centred product concept design (UCPCD), including work tasks related to understanding user needs, creating product concept scenarios, and building and evaluating experience prototypes.

The basis of the work was a user need study that focused on studying 25 adult city-dwellers living in Helsinki. Several different data-collection methods were used, including focus groups, photo diaries, contextual inquiry, and shadowing. All observations were written down as travel episodes, depicting a meaningful journey between two places. From the over 1300 travel episodes, situations related to social interaction were inspected. User needs were identified by looking at situations problematic for the participant and where their routine course of action was not possible because of an obstacle in reaching a goal. The main result of the project was hence the total set of user and concept data gathered and generated during the process.

The project aimed at concept creation and validation, not applications, products, or services. It therefore culminated with the creation and field testing of the Information Radar location-based communication prototype that merged several concepts created in the earlier stages of the project. The Information Radar is based on a PDA (Compaq IPAQ) connected with an external GPS device with an internal electronic compass and a GPRS transceiver attached to the card-slot of the PDA. For indoor use, WLAN positioning with accuracy of up to one metre on the basis of access node signal strengths measured by the devices. In the field tests, the prototype proved to motivate group and public location-based communication by enriching existing communications practices of groups by facilitating easy reference to physical context while providing for location- as well as time-independent communication. The project was finished at the end of March 2003.

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 formalization of 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 Multivariate Biological Data Analysis (BIDMA)

Project leader: Tirri, Henry
Research group(s): Complex System Computation (CoSCo)
Researchers: Myllymäki, Petri; Valtonen, Kimmo; Lahtinen, Jussi; Wettig, Hannes; Kontkanen, Petri; Tuominen, Antti
Schedule: 2002-08-01 ... 2003-07-31
Funding: Kibron Inc. Oy
Keywords: Multivariate data analysis

Abstract and Progress in 2003

The pharmocokinetic ADME (absorption, distribution, metabolism, excretion) parameters are of prime interest in searching for the desired beneficial effects of drugs with minimal adverse effects. The goal of the BIDMA project is to develop methods for early ADME prediction and study how the methods developed can be used in drug design.

In 2003 a wide range of methods were developed and validated in the prediction task, and the results suggest that the data has some predictive potential so that some of the ADME parameters can be estimated from the available data. However, the amount of available data was too small for reliable estimation of prediction accuracy, and it is clear that the results could be further improved by increasing the amount and quality of the data used for model development.

4.1.2.2 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 (HUT), CKIR (Helsinki School of Economics)
Funding: Academy of Finland
Keywords: Proactivity, probabilistic
Research programme: Academy of Finland / PROACT

Abstract and Progress in 2003

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 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 the first year of the project the focus was very much on the problem of how to pre-process the eye movement data so that it can be used for modelling purposes. As this turned out to be more difficult than anticipated, the group developed software tools that can be used for collecting "simulated" eye movement data: in one tool the user reads text through an "artificial eye", a computer-generated "lens" that blurs text outside a small area indicated by the mouse cursor. In another tool the eye movements are not measured directly, but all the actions related to reading (scrolling, window focus control etc.) are recorded. Another data source used in the first modelling attempts performed during the first year was provided by the movie databases found on the Internet.

As a concrete demonstration of the potentiality of modern data-analysis methods in modelling complex multidimensional environments, the group launched a public tool for analyzing the voting behaviour of the Finnish parliament members in 2003. This tool is available (in Finnish only) at http://cosco.hiit.fi/eduskunta/.

4.1.2.3 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
Publications: 18, 39, 40, 102
Keywords: Internet search, concept maps, probabilistic modeling, genre, multinomial Principal Components Analysis

Abstract and Progress in 2003

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 proposed 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 modeling 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.

4.1.2.4 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
Publications: 15, 66, 90, 107
Keywords: Minimum description length principle (MDL), stochastic complexity, universal modeling, model selection, predictive inference

Abstract and Progress in 2003

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 extend 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 2003 the research focused on studying computational issues related to the Normalized Maximum Likelihood (NML) formulation of the MDL principle, and the practical applications of the results. The relationship with the Bayesian reference prior approach was also preliminarily examined. The group developed a software tool for clustering textual or graphical documents (pictures) together so that the similarity metric induced by the clustering can be regarded as a simple approximation of Kolmogorov complexity.

4.1.2.5 Personalized Location-Dependent Services in Wireless Networks (WIRNE)

Project leader: Myllymäki, Petri
Research group(s): Complex System Computation (CoSCo)
Researchers: Kontkanen, Petri; Lahtinen, Jussi; Valtonen, Kimmo; Tuominen, Antti; Wettig, Hannes; Roos, Teemu
Schedule: 2003-03-01 ... 2004-02-29
Cooperation units: Neural Networks Research Centre (HUT), CKIR (Helsinki School of Economics)
Funding: Ekahau Oy
Keywords: Location-dependent services, wireless networks

Abstract and Progress in 2003

The motivation for the project is the need to develop intelligent methods that can be used as a basis for novel location-aware applications. The project focuses on studying positioning methods based on measurements of radio signal strengths in wireless networks, and on developing methods for analyzing this type of data. During 2003 the research group developed several algorithms that can be used for reaching the goals of the project. Initial empirical results look very promising.

4.1.2.6 Search-In-a-Box (SIB)

Project leader: Tirri, Henry
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
Publications: 18, 39, 40, 102
Keywords: Open source, search engines, modeling
Research programme: Tekes / FENIX

Abstract and Progress in 2003

The first generation of corporate intranets, labelled by some as enterprise information management systems, attempt to implement some of the vision of the Semantic Web coupled with personalization and groupware suites. These systems are quite primitive in their text processing and thus perform patchily. Moreover they have huge installation costs and restrictions on interoperability with other systems. Cost-effective information-management software for organizations, libraries, and educational and research institutions are thus non-

existent. With the advent of distributed search just over the horizon (for instance, allowing the database-backed hidden web to be exposed to search), the generally recognized perception that keyword search is a wasteful paradigm, and the recognition that search is a productivity tool that needs personalization, we propose to research and develop software with the following features: (1) it can operate as an appliance (i.e., plug into a network and play with minimal maintenance); (2) it is open source (encourages longevity and non-corporate usability); (3) it employs language models hidden from the user so that named entities, synonyms, topics and genre are implicitly dealt with to ease the user frustration associated with keyword search; (4) it employs probabilistic query evaluation models to overcome the lack of URL (link based) information in some intranets and to carefully integrate relevant synonym information, etc.; and (5) it integrates personalization tightly with search to further improve the user experience, but moreover so that all subsequent groupware tasks such as message routing, information alerts, filtering, etc. can be easily integrated at a later date.

The SIB project officially started in spring 2003 but the work did not reach full speed until the autumn so that the first year of the project will continue until June 2004. In 2003, the basic demonstration targets were set up with the affiliates MBrain and Alma Media, and initial work for the demonstrations was done. An architecture for the new search engine was developed, initial coding done for the basic document processing, and the results were integrated with the Multinomial PCA engine developed in the PROSE project.

4.1.2.7 Computationally Efficient Methods for Deep Computing (DeepC)

Project leader: Tirri, Henry
Research group(s): Complex System Computation (CoSCo)
Researchers: Myllymäki, Petri; Buntine, Wray; Miettinen, Miikka; Nokelainen, Petri; Rissanen, Jorma; Silander, Tomi; Lahtinen, Jussi
Schedule: 2000-07-01 … 2003-12-31
Funding: Academy of Finland
Publications: 15, 16, 40, 66, 68, 69, 82, 83, 84, 90, 107, 108, 112
Keywords: Deep computing, stochastic modelling, stochastic optimisation, data visualization
Research programme: Academy of Finland / Madame

Abstract and Progress in 2003

Deep Computing is a term for methods solving complex and large-scale modelling and analysis problems with emerging computer systems that combine ultra-fast processing with sophisticated analytical software. The methodological research objective of the DeepC project is to develop the theory and methods required for obtaining very large-scale computational, data and communications capabilities that can be used to solve "grand challenge"-level Deep Computing problems in business and science. The research focuses on stochastic approaches and is methodological and theoretical in nature, and aims at topics that can have great impact in this area in the future.

In 2003 the research focused on studying the use of modern modelling techniques in building intelligent environments for E-learning. The tools developed were tested in real-world use with very promising results. More information can be found at http://cosco.hiit.fi/edutech/. The research group also continued the maintenance and development of the unique B-Course data analysis tools available at http://b-course.hiit.fi. The latest version of B-Course (v2.0) was launched in Spring 2003, and it facilitates not only unsupervised dependency and causal modelling, but also a tool for supervised classification and regression tasks. In addition to the publicly available general-purpose tool, special versions of B-Course have been developed for health management research ("P-Course") and bioinformatics (developed by researchers at MIT). The project was finished by the end of December 2003.

4.1.2.8 Computationally Efficient Probabilistic Learning and Reasoning (Cepler)

Project leader: Myllymäki, Petri
Research group(s): Complex System Computation (CoSCo)
Researchers: Rissanen, Jorma; Silander, Tomi; Kontkanen, Petri; Roos, Teemu
Schedule: 2000-02-01 … 2004-12-31
Funding: Academy of Finland
Publications: 15, 16, 39, 66, 90, 107, 108
Keywords: Uncertain reasoning, machine learning, probability theory, Bayesian networks

---

Abstract and Progress in 2003

Probability theory offers a theoretically solid and versatile framework for building models from empirical data and/or expert knowledge. Although the advantages of probabilistic models have been largely recognized, the approach has often been neglected as the theoretically correct, but computationally infeasible methodology. The goal of this research is to explore theoretically valid, but still computationally efficient methods for probabilistic modelling and to develop techniques that can be used for solving both predictive (regression, classification) and explorative (data mining, visualization, causal reasoning) modelling tasks.

In 2003 the research focused on the problem of how to determine the parameters of Bayesian network models in supervised settings. It was discovered that for many non-trivial network structures, the parameters maximizing the conditional (supervised) maximum likelihood can be found efficiently. The research also revealed some interesting connections between supervised Bayesian network models and logistic regression models.

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).

A 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. The Interactive Digital Media group carries out another major line of research in XML technologies. Here the emphasis is on XML-enabled user terminals.

A significant extension in the activities of the research area took place when the Digital Content Communities (DCC) research group was formed in 2003. The DCC group researches 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.

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
Schedule: 2003-06-01 ... 2005-11-30
Cooperation units: Software Business and Engineering Institute (SoberIT) (HUT); UC Berkeley, USA; Media Lab Europe, Ireland; CKIR (Helsinki School of Economics); Hypermedialab (UTA)
Funding: Tekes; Alma Media; TeliaSonera Finland; Veikkaus; Nokia; Starcut; Sulake Labs; Accenture; Codetoys; Sumea
Publications: 91, 101
Keywords: Mobility, community, digital media content, social user experience
Research programme: Tekes / FENIX

Abstract & Progress in 2003

Launched in June 2003, the project MC2 has the goal of understanding the roles of mobility, community and user experience in playing games. The project studies the social meaning and impact of new communications technology for communities that are interested in mobile gaming. It builds an environment for monitoring and modelling the ecosystem of communities and commercial actors around open and collaboratively created game-related content. The results expected of the projects include evaluated and tested scenarios of mobile community gaming, new template-based design tools that allow people to easily make their own games and game-related content, new open source tools to empower the community activity, company-specific case studies to help the industry partners to benefit from community-created content, an academia/industry network that is self-sustaining, and guidelines for best practices for further use.

4.1.3.2 XML Devices

Project leader: Vuorimaa, Petri
Research group(s): Interactive Digital Media
Researchers: Honkala, Mikko; Vierinen, Juha; Sikha, Para Brahmendra Rao
Schedule: 2001-09-01 ... 2003-11-30
Cooperation units: Laboratory of Telecommunications Software and Multimedia (HUT)
Funding: Tekes
Publications: 89
Keywords: XML

Abstract and Progress in 2003

The XML Devices project's main goal was to implement a client-side environment for XML-based services. The project was based on the open source XML browser X-Smiles, which is written in Java. The project developed the browser further and ported it to different kinds of user devices, such as PDAs and MHP-enabled digital television set-top-boxes. The project also studied the utilization of different XML-based presentation languages, such as XForms and SMIL. The project attended XML standardization work by being part of the W3C XForms Working Group and producing a reference implementation of the XForms standard.

In the project's final year the project developed further the X-Smiles browser, which is designed for presenting dynamic multimedia content in XML format. In addition, the project

has researched embedded Java environments and adapted them to run the X-Smiles browser. The project also developed further the support for different presentation XML languages in the X-Smiles browser and studied their utilization in different environments. The project was also a member of the W3C XForms working group and has significantly contributed to the specification work. XForms became a final Recommendation at the end of 2003. The XForms implementation that was done in this project was one of the two reference implementations. One focus of the research has been the integration of different XML languages, notably, SVG, XForms, XHTML, CSS and SMIL. This has been done on the abstract level as well as implementing the proposed integrated languages.

Additionally the project studied the utilization of the XML browser in Digital Television. The browser was ported to the MHP (Multimedia Home Platform) environment. This environment has a specific user interface library that is specifically designed to be used with a remote control.

The results of the project are released in the form of journal and conference articles. The program and its source code has also been released as part of the Open Source X-Smiles distribution at http://www.x-smiles.org/. Tens of thousands of users have downloaded the browser and it is mainly used in research and development work. The browser can be used to develop future XML-based networked services. In addition, the browser can be used for research about future multimedia terminals. The project was finished in November, 2003.

4.1.3.3   Intelligent Web Services

Project leader: Hyvönen, Eero **Research group(s):** Semantic Computing **Researchers**: Apiola, Mikko; Lindgren, Petri; Viljanen, Kim **Schedule:** 2003-08-01 ... 2004-07-30 **Cooperation units:** Department of Computer Science (UH) **Funding:** Tekes; Fonecta; TeliaSonera Finland; TietoEnator **Keywords:** Semantic web, web services **Research programme:** Tekes / FENIX

Abstract and Progress in 2003

The Intelligent Web Services (IWebS) research project studies the possibilities of the Semantic Web and Web Services technologies in both annotating the services and delivering the relevant services to the end-users. The goal of IWebS is to create an intelligent yellow pages service, where the semantically annotated services cover both static advertisement as in traditional yellow page services and dynamic services whose availability to the end-user depends on the context in which the user and the services are. The main outputs of the IWebS project are an intelligent annotation editor for the service providers, ontologies for describing and storing the services, and an intelligent reasoning engine for matching the static and dynamic services for the end-user.

The IWebs-project started at the beginning of August 2003. During autumn 2003 the problem domain and relevant previous research were studied. Work on creating the needed ontologies underlying the demonstration system started.

4.1.3.4 Semantic Web

Project leader: Hyvönen, Eero
Research group(s): Semantic Computing
Researchers: Viljanen, Kim; Kettula, Suvi-Päivikki; Saarela, Samppa; Valo, Arttu; Junnila, Miikka; Salminen, Mirva; Syreeni, Ahti; Mäkelä, Eetu; Holi, Markus
Schedule: 2002-01-01 ... 2003-12-31
Cooperation units: Department of Computer Science (UH)
Funding: Tekes; Nokia; TietoEnator; Espoo City Museum; National Board of Antiquities, Foundation of Helsinki University Museum; Antikvaria-group
Publications: 58, 59, 60, 111, 120
Keywords: Semantic Web

Abstract and Progress in 2003

The goal of the project was to investigate from the information storage and retrieval viewpoints, how semantically rich data should be annotated with metadata in order to facilitate content-based information retrieval. Cultural artifact databases and photograph repositories were used as case studies.

The first implementation of the system MuseumFinland - Finnish Museums on the Semantic Web was finished and demonstrated in Finland and abroad. The second case study software, the semantic image retrieval system "Promoottori" was finished and deployed at the University museum for public usage. Implementation of the mobile telephone interface started. Technical and scientific articles concerning the case studies were written.

4.1.3.5 Adanets

Project leader: Hyvönen, Eero
Research group(s): Semantic Computing
Researchers: Leinonen, Jani
Schedule: 2003-01-01 … 2004-12-31
Cooperation units: Department of Computer Science (UH)
Funding: Nokia Research Centre
Keywords: Semantic web, web services
Research programme: ITEA

Abstract and Progress in 2003

The goal of the Adanets project is to investigate the scope of problems related to the mobility in applications and in network services. Particular attention is to be paid to the adaptiveness, i.e. the capability of networks and services to provide mobility features. Specifically, Adanets elaborates a generic service/network model with adaptive properties enabling the provisioning of mobile services and end-user applications, with the guarantees of the required Quality of Service (QoS) levels and according to Service Level Agreements (SLA).

HIIT participated in the project as a subcontractor of Nokia Research Centre. The research was related to the problem of finding mobile services based on their content.

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 responsible leader of the group, and Olli Pitkänen is the Program Coordinator. The group's strengths include especially 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 Digital Rights Management in Mobile Services (MobileIPR)

Project leader: Kemppinen, Jukka
Research group(s): Digital Economy
Researchers: Pitkänen, Olli; Välimäki, Mikko; Oksanen, Ville; Reti, Tommo; Soininen, Aura
Schedule: 2000-09-01 ... 2003-12-31
Cooperation units: UC Berkeley, USA
Funding: Tekes; Nokia; TeliaSonera Finland; Elisa Communications, Finnish Broadcasting Corporation (YLE)
Publications: 21, 85, 86, 87, 105, 106, 110, 114, 115, 121, 122, 123, 133
Keywords: Intellectual property rights (IPR), digital rights management (DRM), wireless networks, mobile Internet, information products

---

Abstract and Progress in 2003

The MobileIPR project studied intellectual property rights (IPR) management in the changing business and technology environment. Digital Rights Management (DRM) is becoming a key issue in the distribution of information products. Wireless networks and mobile terminals along with the digital media bring forward new legal problems. The research focus was on copyright and patent, but issues like the liability of intermediaries, agreements on rights and liabilities as well as their management were also studied. The project had two locations: HIIT in Helsinki, Finland, and the School of Information Management and Systems (SIMS) at UC Berkeley, California.

In 2003, the project was carried out in accordance with the project plan, the researchers produced a number of publications, and the project organized the First International Mobile IPR Workshop. The project ended in December 2003.

4.1.4.2 Security Technologies and Attitudes in Mobile IPR (STAMI)

Project leader: Mäntylä, Martti
Research group(s): Digital Economy
Researchers: Kortesniemi, Yki; Heikkilä, Juho; Kanerva, Pekka; Päivärinta, Juha; Latva-Koivisto, Antti
Schedule: 2001-03-01 ... 2003-12-31
Funding: Tekes; Nokia; Elisa Communications
Publications: 70, 148, 150
Keywords: Security, privacy, certificates, intellectual property rights

Abstract and Progress in 2003

The STAMI project aimed at managing information validity as well as end user privacy in the distribution of digital content to mobile terminals. Mobile networks enable new business models, where the user fetches the required content from the network on a need basis. On the other hand, this raises questions about the validity of information as well as how to maintain sufficient end user privacy against usage monitoring. The deliverables were concepts, scientific papers, licentiate's thesis, and research prototypes. In 2003 the project looked at giving end-users the chance to act as content creators with their mobile terminals and the project developed a prototype demonstrating this in the context of a news service. The project was completed in December 2003.

4.1.4.3 Structures of Mobile Digital Economy (DE Core)

Project leader: Kemppinen, Jukka
Research group(s): Digital Economy
Researchers: Virtanen, Perttu; Soininen, Aura; Sarvas, Risto; Hietanen, Herkko; Reti, Tommo
Schedule: 2002-01-01 ... 2004-12-31
Cooperation units: UC Berkeley, USA
Funding: Tekes; Nokia; TeliaSonera Finland; Elisa Communications; Ericsson; Sanoma WSOY; Finland Post (Suomen Posti); Nixu Oy
Keywords: Digital economy, computer law, electronic commerce, mobile computing, virtual communities, metadata, content management
Research programme: Tekes / UTT

Abstract and Progress in 2003

The project studies the structures of digital economy based on mobile computing particularly from the legal viewpoint. Part of the work is performed at UC Berkeley. The DE Core research work progressed in 2003 according to the project plan. The preparatory work for the pilot project in 2004 has already commenced. The researchers carried out the study in Finland and at UC Berkeley. The research has yielded a number of research papers, articles and presentations both in Finland and abroad together with the planned seminar days and research topic get-togethers with partners and other research institutions and universities. All this has also paved the way for the project pilot work and research on Digital Content Distribution Management System to be completed in 2004.

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

---

Project leader: Kemppinen, Jukka (1.1.-31.8.); Mäntylä, Martti (1.9.-)
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

Abstract and Progress in 2003

The Welfare of Nations project studies 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 is carried out in close co-operation with the Berkeley Center for Information Society at University of California, Berkeley. It is 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). In 2003 the project produced manuscripts for research reports and related publications, and also developed plans for a continuation of the project.

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 systems (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- or situation-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)

Abstract and Progress in 2003

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 program 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 2003, the coordinator and programme leader made site visits to all projects. During the site visits, the project researchers presented themselves and their research. Short summaries of the site visits are available on the programme's web pages. The programme has been presented at several events to the researcher community and other funding organisations. An email list has been established for informing researchers about news and events related to the programme.

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; Raento, Mika; Toivonen, Hannu
Schedule: 2002-11-01 … 2005-12-31
Co-operation units: ARU/HIIT
Funding: Academy of Finland
Keywords: Context recognition, mobile devices, data mining
Research programme: Academy of Finland / Proactive Computing (PROACT)

Abstract and Progress in 2003

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 adapt their behaviour appropriately. This joint BRU/ARU project is part of the Research Program on Proactive Computing of the Academy of Finland (see above), and it started in November 2002.

In 2003 we have investigated discovery and recognition of personally important places, based on traces of GSM cells of the user. Besides recognizing a meaningful location, the aim is also to make predictions of the next important location in case the user is moving. 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. First results include concepts and algorithms for these tasks.

We have also developed software for practical testing of the ideas in Series60 mobile phones. The implementation allows collection of context data (location, communication, calendar, and profile), automatic recognition and prediction of important places, as well as a presence-aware phone book. This sofware will be used to verify the methods as well as in user studies in collaboration with our partner project at HIIT/ARU. We plan to publish the software and make it available for interested users and other researchers.

4.2.1.3 Networking and Architecture for Proactive Systems (NAPS) – Algorithms and Protocols

Project leader: Floréen, Patrik
Research group(s): Adaptive Computing Systems
Researchers: Floréen, Patrik; Kohonen, Jukka; Ukkonen, Johannes
Schedule: 2003-01-01 … 2005-12-31
Co-operation units: Laboratory for Theoretical Computer Science (HUT); Networking Laboratory (HUT)
Funding: Academy of Finland
Publications: 45, 149
Keywords: Ad hoc networking, algorithms
Research programme: Academy of Finland / Proactive Computing (PROACT)

Abstract and Progress in 2003

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 optimization problems in topology control and hierarchical routing of ad hoc networks by the well-established methods of systematic design and analysis of algorithms, specifically approximative and randomized techniques applied in the context of distributed online computational problems.

During 2003 the project has studied topology control problems in ad hoc networks with the emphasis on the effect of energy constraints. Jukka Kohonen presented the results from this research at the DIALM-POMC Joint Workshop on Foundations of Mobile Computing, which was held in connection with MobiCom 2003. Related to this research, Petteri Kaski (HUT) submitted a journal paper on Steiner tree packing. The group started research on wireless sensor networks. The focus is again on the effect of energy constraints. Also, work on clustering for hierarchical ad hoc networks was conducted. Johannes Ukkonen wrote his Master's Thesis on this topic.

4.2.1.4 SPACE4U Terminal Software Management System Design/Development

Project leader: Floréen, Patrik
Research group(s): Adaptive Computing Systems
Researchers: Floréen, Patrik; Virtanen, Otso
Schedule: 2003-07-01 … 2004-06-30
Co-operation units: Industrial Information Technology Laboratory (HUT); Nokia Research Center
Funding: Nokia Research Center
Keywords: Context-awareness; component frameworks; middleware; embedded systems

Abstract and Progress in 2003

The ITEA project Space4U (Software Platform and Component Environment for you) builds on the foundations established by the ITEA project ROBOCOP. Whereas ROBOCOP defined a component-based software architecture for the middleware layer of high-volume embedded appliances, Space4U aims to extend the framework in the areas of power management, fault management and remote terminal management based on clear demands from the application domain. HIIT works in this project as a subcontractor to Nokia Research Center. During the autumn of 2003 HIIT participated in the Space4U task 1.4 and contributed to the design of a context-aware re-configuration framework for ROBOCOP terminals.

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 speed, 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; Onkamo, Päivi; Sevon, Petteri
Schedule: 2003-08-01 … 2005-02-28
Co-operation units: GeneOS Ltd; Jurilab Ltd; Cyberell Ltd; Karolinska Institute, Sweden; National Public Health Institute (KTL); Department of Medical Genetics (UH)
Funding: Tekes; companies; BRU; UH
Publications: 19, 25, 36, 44, 51, 93, 100, 119, 135, 136
Keywords: Gene mapping, genetics, data mining

Abstract and Progress in 2003

We develop and apply methods for gene mapping, for analysing genotypes, haplotypes, and phenotypes and their relationships in more general, as well as for data mining in bioinformatics. 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 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 2003 we have investigated three problems related to linkage disequilibrium-based gene-mapping: novel methods for haplotyping large marker maps, models and methods for haplotype similarity and clustering, and generalization of our earlier gene-mapping methods to extend the scope of their applications. An emphasis has been in utilisation of genotype data. Many gene-mapping methods assume haplotype data, whereas genotypes are often all that is available. We introduced a novel approach for haplotyping a set of given non-related genotypes, based on Markov chains of variable order. Further, we introduced HPM-G, a gene-mapping method that searches for plausible haplotype patterns directly in the genotype data. These approaches potentially allow significant savings in wet labs and an improved mapping accuracy.

In the autumn of 2003 we started a new project, Altti, funded by Tekes and three genetics companies. In the context of this project we have started to study "phenotype mining", analysis of large and complex medical datasets that contain rich information about the study subjects: genotypes of thousands of markers, biophysical and biomedical measurements, questionnaires, etc.

4.2.2.2 Computational Tools in Gene Expression Data Analysis

Project leader: Hollmén, Jaakko
Research group(s): Data analysis
Researchers: Ruosaari, Salla; Seppänen, Jouni; Patrikainen, Anne
Schedule: 2002-01-01... 2005-31-12
Funding: European Commission; graduate schools
Publications: 24, 25, 130, 131

Abstract and Progress in 2003

The project has concentrated on developing methods for assessing the significance of differences in gene expression profiles between groups. Typically such data sets have only a small number of observations, and thus advanced techniques are needed for significance testing. In 2003, the research has resulted in several publications. The main methodological themes have been techniques for comparing the spatial location of genes and their expression, and the study of heritability of expression.

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
Publications: 145
Research programme: Academy of Finland / Systems Biology and Bioinformatics

Abstract and Progress in 2003

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. 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; Patrikainen, Anne
Schedule: 2002-01-01…
Funding: Academy of Finland; graduate schools; European Commission funding from HUT; BRU basic funding
Publications: 1, 2, 3, 4, 13, 28, 29, 38, 41, 42, 46, 47, 48, 49, 50, 51, 53, 75, 76, 77, 78, 79, 80, 81, 92, 103, 104, 117, 118

Abstract and Progress in 2003

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 2003 there were several interesting developments. The methods for finding recurrent sources from sequences and discovering fragments of order 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: Department of Geology (UH); Institute of Biotechnology (UH); Research Institute for the Languages of Finland
Funding: Graduate schools, BRU basic funding
Publications: 28, 48, 49, 71, 74, 109

Abstract and Progress in 2003

Biodiversity, climate studies, and onomastics, for example, result in data sets that have a strong spatial component. 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 area in 2003 was the study of the distribution of place names in Finland. The results indicate that some of the traditional properties of the occurrence of place names need rethinking: e.g., the repulsion phenomena do not seem as strong as assumed. In late 2003, a study was initiated on the distribution of bird species in Finland.

In the study of temporal phenomena in data, work was done in paleontology, in finding good orderings for fossil sites. Spectral methods proved to be quite useful in this task, giving good accuracies in the estimation of the age of sites. Current themes of interest included the approximation qualities of spectral and other methods.

4.2.3 Neuroinformatics

Neuroinformatics is broadly defined as the intersection of Information technology and neuroscience. Our research goals are to 1) 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 to 2) 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 May 2003.The year 2003 was mainly spent in finishing previous projects not affiliated with HIIT, planning and organization of new projects to be started in BRU, and recruitment of new researchers. Funding was applied for and obtained from the Academy of Finland for a new collaborative project "Advanced data analysis in vision research: feature interaction and nonlinear classification images" starting in January 2004. Another Academy of Finland project, "Independent Component Analysis and its extensions", was transferred to BRU starting in January 2004. Two post-doctoral researchers were recruited to join the group in January 2004. A long-term visitor from Osaka University (Japan) arrived in September. A PhD student was recruited in collaboration with the ARU.

Collaborative partners include the Department of Psychology at UH, Neural Networks Research Centre at HUT, Department of Behaviormetrics at Osaka University, and Department of Neurological Sciences at the University of Naples (Italy).

Publications: 5, 6, 7, 8, 9, 10, 52, 54, 55, 56, 57

4.3 EU-activities

HIIT has aimed to be active in EU-funded research and although only part of the proposals HIIT participated in during the year turned out to be successful, HIIT became involved in a number of Information Society Technologies (IST) projects in 2003. The year 2003 was the first year for the EU 6th framework programme for research. 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. Of these, three IPs were accepted for funding in the first call of the IST priority. 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. This application was submitted to the second call of the IST priority in October 2003. The evaluation of the proposal was favourable and MobiLife received an invitation to contract negotiations. The consortium consists of many major companies and research organisations in Europe. Both units of HIIT will be active in the project.

The Superpeer Semantic Search Engine (ALVIS) Specific Targeted Research Project (STREP), led by Professor Henry Tirri in the ARU, aims at building the foundation for a distributed semantic search engine. The proposal was submitted to the 1st call of the IST programme in April 2003, the contract negotiations were undertaken in Luxembourg in July, and the Commission contract and consortium agreement were signed in December. This three-year-project starting on 1.1.2004 is of particular importance, as 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 works on probabilistic logic learning systems and their application to real-life problems and it starts on 1.1.2004. The BRU is a partner in this project.

In addition to these, HIIT participated in EUREKA-activities and is preparing further EUREKA participation. Also, HIIT-affiliated researchers were active in projects still running under the 5th framework programme, as well as in some new 6th framework programme projects, but with an affiliation to the corresponding university department. It should also be noted that of HIIT's researchers, Dr. Marko Turpeinen participated in the evaluation of research proposals in Brussels and Dr. Patrik Floréen is a member of the Tekes IST-support/strategy group for the Finnish delegates to the IST Programme Committee.

5 Research Training and Research Visits

5.1 Doctoral Degrees Earned by HIIT Personnel

12.12.2003: Ella Bingham, *Advances in Independent Component Analysis with Applications to Data Mining*, Helsinki University of Technology, supervised by Professor Heikki Mannila, Dr. Aapo Hyvärinen and Professor Erkki Oja.

5.2 Post-graduate Courses Arranged by HIIT

Spring 2003	Spatial Data Mining (Antti Leino)
Spring 2003	Research Seminar on Algorithmics (Patrik Floréen)
Spring 2003	Research Seminar on Causal Models (Petri Myllymäki)
Spring 2003	Network Application Frameworks & XML (Pekka Nikander)
Spring 2003	Management of IPR Strategy and Implementation in SW Organization (Olli Pitkänen, Aura Soininen)
Spring 2003	Law and Information Technology (Olli Pitkänen, Aura Soininen)
August 2003	Genetics for Computer Scientists (Päivi Onkamo)
Autumn 2003	Context-Aware Computing (Patrik Floréen, Greger Lindén, Hannu Toivonen, Henry Tirri)
Autumn 2003	Mining Biomolecular Data (Hannu Toivonen)
Autumn 2003	Algorithms for Ad Hoc Networking (Patrik Floréen)
Autumn 2003	Graphical Models (Wray Buntine, Petri Myllymäki)
Autumn 2003	Special Course on Data Mining (Bart Goethals)
Autumn 2003	Game Theoretic Concepts in Computer Science (Floris Geerts)
Autumn 2003	The Metalanguage XML (Mika Raento)
Autumn 2003	Web Application (Torsten Rüger)

5.3 Visits to HIIT

Foto Afrati, PhD, Professor, National Technical University of Athens, Greece (29.10.-4.11.2003).

Luc De Raedt, PhD, Professor, Albert-Ludwigs-Universitaet, Freiburg, Germany (9.-15.6.2003).

Alexander Hinneburg, PhD, The Institute of Computer Science of the Martin-Luther University, Halle/Wittenberg, Germany (27.1.-15.2.2003).

Manfred Jaeger, PhD, Max-Planck-Institut für Informatik, Saarbrücken, Germany (1.-8.6.2003).

Tuomas Sandholm, Professor, Carnegie Mellon University, Pittsburgh, USA (18.-22.1.2003).

Mohammed Zaki, PhD, Assoc. Professor, Rensselaer Polytechnic Institute, Troy, New York, USA (10.5.-8.6.2003).

5.4 Visits from HIIT

Pekka Himanen, PhD, Berkeley Center for the Information Society, International Computer Science Institute, USA (several visits during 2003)

Juha Muilu, PhD, European Bioinformatics Institute, Hinxton, UK (June-July 2003)

Juha Muilu, PhD, IBM e-business solutions center, LaGaude, France (August 2003)

Martti Mäntylä, Professor, University of California at Berkeley, USA, (several short visits during 2003)

Martti Mäntylä, Professor, Tsinghua University, China (15.-26.3.2003, 26.-28.11.2003).

Antti Oulasvirta, PhD student, Center for Cognitive Science, University of California at Berkeley, USA (1.1.-25.5.2003).

Olli Pitkänen, Lic.Sc., LL.M., School of Information Management and Systems, University of California at Berkeley, USA (25.5.-22.6., 8.9.-25.9. and 16.10.-25.11.2003)

Risto Sarvas, M.Sc. (Tech.), School of Information Management and Systems, University of California at Berkeley, USA, (1.1.-30.11.2003)

Aura Soininen, LL.M., School of Information Management and Systems, University of California at Berkeley, USA (6.8.2003 –)

Evimaria Terzi, M.Sc., University of Milan, Italy (July 2003)

Henry Tirri, Professor, Department of Computer Science, Stanford University, USA (1.-30.7.2003).

Henry Tirri, Professor, University of California at Berkeley, USA (several visits during 2003).

6 Administration

6.1 Overview

According to the management bylaws established in 2002, HIIT is a joint research institute of HUT 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 HUT. Similarly and symmetrically, 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 resulting structure is shown in the organisation chart below.
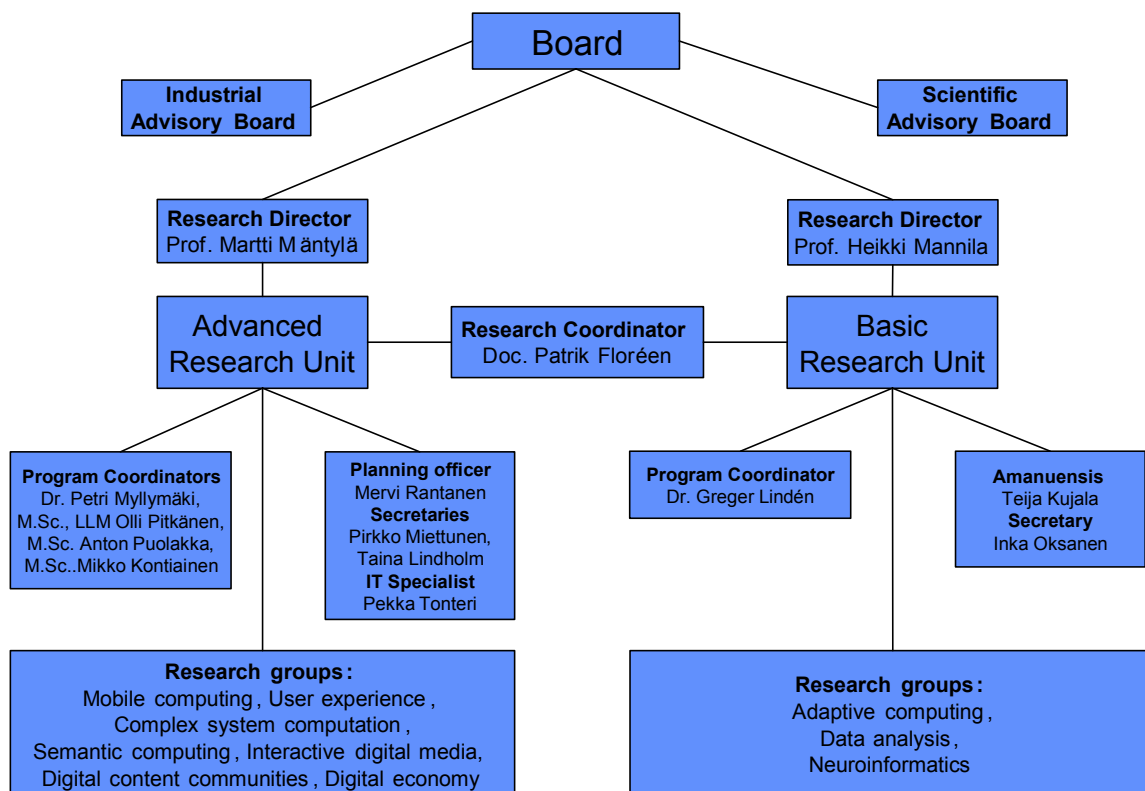
Internally, ARU consists of seven research groups, facilitated by a small administration team. In addition, each of the four thematic research areas has a Program Coordinator who facilitates the planning and management of research activities. The Management Board of ARU, headed by the Research Director 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 HUT and UH. Heikki Mannila is in overall charge of the unit's general management, budgeting, and planning. The administrative personnel consists of Planning Secretary Inka Oksanen, 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 2003 the Board convened four 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 completed the drafting and negotiating of HIIT research contract templates, a process started in 2002. In its first meeting of 2003, the Board ratified the so-called Open Source Research Agreement. That was the third and the last of the set of HIIT research contract templates. Consequently, HIIT now has two research agreements (the above-mentioned Open Source Research Agreement and a Business Critical Agreement) as well as a Contract for the Assignment of the Intellectual Property Rights of Research Results.

In 2003 the Board members (and their personal deputies) were as follows:

Members:
Vice Rector, Professor Mauri Airila, HUT, Chairman of the Board (Vice Rector, Professor Matti
         Pursula. From Oct. 2003: Vice Rector, Professor Olavi Nevanlinna)
Professor Olli Simula, HUT (Professor Markku Syrjänen)
Vice Rector, Professor Thomas Wilhelmsson, UH, Vice-Chairman of the Board (Vice Rector,
         Professor Raija Sollamo. From June 2003: Vice Rector, Professor Marja Makarow)
Academy Professor Esko Ukkonen, UH (Professor Jukka Paakki)
Rolf Svanbäck, LM Ericsson (Raimo Vuopionperä)
Aimo Maanavilja, Elisa (Minttu Lintera)
Veikko Hara, TeliaSonera Finland (Juha Aaltonen)
Heikki Saikkonen, Nokia (Harry Herlin)
Petri Myllymäki, HIIT (Yki Kortesniemi)

*Invited Member*s:
Anneli Rantanen, TietoEnator. From Dec. 2003: Ari Hirvonen (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 2003 these expert members were as follows:

Expert member:
Research Director Martti Mäntylä, HIIT ARU
Research Director Heikki Mannila, HIIT BRU
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
Board Secretary:
Institute Secretary Pirkko Miettunen, HIIT ARU. From Oct. 2003: Planning Officer Mervi Rantanen, HIIT ARU.

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, Royal Holloway, University of London
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

Year 2003 was an important milestone in HIIT's history as the SAB convened for the first time. The meeting was held 16-17 October in Helsinki. Of the members of the board, the following were present for the full meeting: Professor Christos Faloutsos, Professor Randy Katz, Professor Martin Kersten, Professor Kari-Jouko Räihä and Professor John Shawe-Taylor. In addition, Professor Mart Saarma was present for part of the meeting.

In its concluding comments presented to the Chairman of the HIIT Board, Vice Rector Mauri Airila, and to the senior researchers of HIIT, the SAB focused on the scientific level of the research, as well as on the quality of graduate and postdoctoral education. Neither industrial nor societal impact was commented as it was considered hard for SAB to determine. The overall evaluation was that HIIT's research is generally of world class with impressive collaborations. The graduate and postdoctoral education was considered impressive, having excellent facilities and great projects to work on, and excellent students.

The main recommendations of SAB were that collaboration between the research groups in HIIT should be further strengthened, that the strategic goals of HIIT should be clarified and that more emphasis should be put on PR-activity and publicity. Of the different research areas, the SAB especially appreciated the research done in data mining, mobile computing, intelligent systems and computational neuroscience.

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 2003 in connection with the HIIT Retreat on May 27, 2003. The meeting focused on the present status and competences of HIIT, the future outlook of HIIT's research areas and competences, some strategic issues, the financing of its operations and infrastructure, and HIIT's interface with the industry.

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 HUT 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 2003 the ARU staff completed 54 person-years, out of which approximately 90 per cent was with external funding and approximately 10 per cent with budget funding. There is a clear increase in person-years compared to 2002 (40 person-years) and 2001 (24 person-years), respectively. The BRU staff completed 33 person-years in 2003.

7 Funding and Costs

7.1 Advanced Research Unit Finances

The finances of the Advanced Research Unit for 2003 are shown in the table below. For reference, the comparable numbers of 2001 and 2002 are also given.

	2001	2002	2003
Total funding	2 095 664 €	3 025 570 €	3 498 410 €
HUT funding	258 874 €	258 188 €	168 188 €
UH funding	240 636 €	188 188 €	168 188 €
Academy of Finland	118 825 €	189 724 €	379 462 €
National Technology Agency Tekes	947 062 €	1 644 796 €	2 103 451 €
Industry	508 403 €	643 724 €	600 699 €
Ministries and other public funding	21 864 €	100 950 €	71 694 €
Other national funding			6 728 €
Total expenses	2 051 592 €	3 121 607 €	3 545 099 €
Salaries	1 035 096 €	1 819 884 €	2 376 864 €
Other operational expenses	652 829 €	586 327 €	660 739 €
Rents	217 616 €	476 895 €	343 216 €
Service charge to HUT	146 051 €	238 501 €	164 280 €

	2001	2002	2003
Balance	44 072 €	-96 037 €	-46 689 €
Cumulative balance	150 616 €	54 578 €	7 889 €

	2001	2002	2003
University funding % of total funding	24 %	15 %	10 %
External funding % of total funding	76 %	85 %	90 %
Academy funding % of total funding	6 %	6 %	11 %
Tekes funding % of total funding	45 %	54 %	60 %
Industry funding % of total funding	24 %	21 %	17 %
Other public funding % of total funding	1 %	3 %	2 %

	2001	2002	2003
Salaries % of total expenses	50 %	58 %	67 %
Other expenses % of total expenses	32 %	19 %	19 %
Rents % of total expenses	11 %	15 %	10 %

The numbers indicate gradual, yet significant growth also in 2003, especially in the funding from the Academy of Finland.

7.2 Basic Research Unit Finances

The finances of the Basic Research Unit for 2003 are shown in the table below.

Own basic funding	631 800 €
Academy of Finland	
ProAct Program Coordination	90 000 €
ProAct –projects	140 000 €
Life 2000	55 000 €
Tekes	115 000 €
Industry	45 000 €
Graduate schools	260 000 €
Center of excellence funding from HUT	75 000 €
Total	1 411 800 €

Appendices

A Publications

Articles in international scientific journals with referee practice

1. M. Datar, T. Feder, A. Gionis, R. Motwani, R. Panigrahy: A Combinatorial Algorithm for MAX SCP, *Journal of Information Processing Letters*, 2003, vol. 85.

2. M. N. Garofalakis, A. Gionis, R. Rastogi, S. Seshadri, K. Shim: XTRACT: Learning Document Type Descriptors from XML Document Collections, *Data Mining and Knowledge Discovery*, 2003, vol. 7(1), pp. 23-56.

3. F. Geerts: Expressing the Box Cone Radius in the Relational Calculus with Real Polynomial Constraints, *Discrete and Computational Geometry*, 2003, vol. 30 (4), pp. 607-622.

4. D. Gunopulos, R. Khardon, H. Mannila, S. Saluja, H. Toivonen, R.S. Sharma: Discovering All Most Specific Sentences, *ACM Transactions on Database Systems,* 2003, vol. 28 (2), pp. 140 - 174.

5. J. Hurri, A. Hyvärinen: Simple-Cell-Like Receptive Fields Maximize Temporal Coherence in Natural Video, *Neural computation*, 2003, vol. 15 (3), pp. 663-691.

6. J. Hurri, A. Hyvärinen: A Two-Layer Temporal Generative Model of Natural Video Exhibits Complex-Cell-Like Pooling of Simple Cell Outputs, *Neurocomputing*, 2003, vol. 52-54, pp. 553-559.

7. J. Hurri, A. Hyvärinen: Temporal and Spatiotemporal Coherence in Simple-Cell Responses: A Generative Model of Natural Image Sequences, *Network: Computation in Neural Systems*, 2003, vol. 14 (3), pp. 527-551.

8. A. Hyvärinen, E. Bingham: Connection Between Multilayer Perceptrons and Regression Using Independent Component Analysis, *Neurocomputing*, 2003, vol. 50 (C), pp. 211-222.

9. A. Hyvärinen, J. Hurri, J. Väyrynen: Bubbles: A Unifying Framework for Low-Level Statistical Properties of Natural Image Sequences, *Journal of the Optical Society of America A*, 2003, vol. 20 (7), pp. 1237-1252. Reprinted in *Virtual Journal of Biological Physics Research*, 2003. vol. 6 (2).

10. V. Kiviniemi, J.-H. Kantola, J. Jauhiainen, A. Hyvärinen, O. Tervonen: Independent Component Analysis of Nondeterministic fMRI Signal Sources, *NeuroImage*, 2003, vol. 19(2), pp. 253-260.

11. J.E. Litton, J. Muilu, A. Bjorklund, A. Leinonen, N.L. Pedersen: Data Modeling and Data Communication in GenomEUtwin, *Twin Research,* 2003, vol. 6 (5), pp. 383-390.

12. A. Oulasvirta, E. Kurvinen, T. Kankainen: Understanding Contexts by Being There: Case Studies in Bodystorming, *Personal and Ubiquitous Computing*, 2003, no. 7, pp. 125-134.

13. D. Pavlov, H. Mannila, P. Smyth: Beyond Independence: Probabilistic Methods for Query Approximation on Binary Transaction Data, *IEEE Transactions on Data and Knowledge Engineering*, 2003, vol. 15 (6), pp. 1409-1421.

14. O. Pitkänen, M. Mäntylä, M. Välimäki, J. Kemppinen: Assessing Legal Challenges on the Mobile Internet, *International Journal of Electronic Commerce*, 2003, vol. 8 (1), pp. 101-120.

15. J. Rissanen: Complexity of Simple Nonlogarithmic Loss Functions, *IEEE Transactions on Information Theory*, 2003, vol. 49 (2), pp. 476-484.

16. J. Rousu, L. Flander, M. Suutarinen, K. Autio, P. Kontkanen, A. Rantanen: Novel Computational Tools in Bakery Process Data Analysis: A Comparative Study, *Journal of Food Engineering*, 2003, pp. 45-56.

17. T.A. Thanaraj, F. Clark, J. Muilu: Conservation of Human Alternative Splice Events in Mouse, *Nucleic Acids Research,* 2003, vol. 31 (10), pp. 2544-2552.

18. H. Tirri: Search in Vain: Challenges for Internet Search, *Computer*, 2003, vol. 36 (1), pp. 115-116.

19. H.T.T. Toivonen, A. Srinivasan, R.D. King, S. Kramer, C. Helma: Statistical Evaluation of the Predictive Toxicology Challenge 2000-2001, *Bioinformatics*, 2003, vol. 19 (10), pp. 1183-1193.

20. A. Vakali, E. Terzi, E. Bertino, A.K. Elmagarmid: Hierarchical Data Placement for Navigational Multimedia Applications, *Data Knowledge Eng.*, 2003, vol. 44 (1), pp. 49-80.

21. M. Välimäki: Dual Licensing in Open Source Software Industry, *Systemes d´Information et Management*, 2003, vol. 8 (1), pp. 63-75.

Accepted for publication

22. F. Geerts, B. Kuijpers: Topological Formulation of Termination Properties of Iterates of Functions, *Information Processing Letters*, Elsevier. To appear.

23. Hyvärinen, J. Hurri: Blind Separation of Sources that Have Spatiotemporal Variance Dependencies, *Signal Processing*. To appear.

24. 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, H. Wikman: Differentially Expressed Genes in Non-Small Cell Lung Cancer (NSCLC) – The Expression Profiling of Cancer-Related Genes in Squamous Cell Lung Cancer, *Cancer Genetics and Cytogenetics*. To appear.

25. S. Koskenmies, E. Widén, P. Onkamo, M. Zucchelli, P. Sevón, H. Julkunen, J. Kere: Haplotype Associations Define Target Regions for Susceptibility Loci in Systemic Lupus Erythromatus, *Eur J Hum Genetics*. To appear.

26. E. Kurvinen, A. Oulasvirta: Towards Socially Aware Pervasive Computing: A Turntaking Approach, *2nd International IEEE Conference on Pervasive Computing and Communications*, Orlando, Florida, March 2004. Accepted.

27. J. Makkonen, H. Ahonen-Myka, M. Salmenkivi: Simple Semantics in Topic Detection and Tracking, *Information Retrieval Journal.* To appear.

28. H. Mannila, M. Salmenkivi: Using Markov Chain Monte Carlo Methods and Dynamic Programming for Event Sequence Data, *Knowledge and Information Systems*. To appear.

29. T. Mielikäinen, E. Ukkonen: The Complexity of Maximum Matroid-Greedoid Intersection and Weighted Greedoid Maximization, *Discrete Applied Mathematics*. To appear.

30. M. Miettinen, P. Nokelainen, J. Kurhila, T. Silander, H. Tirri: EDUFORM – A Tool for Creating Adaptive Questionnaires, *Journal of Technology and Teacher Education*. To appear.

31. P. Nokelainen, M. Miettinen, J. Kurhila, P. Floréen, H. Tirri: A Shared Document-Based Annotation Tool to Support Lerner-Centered Collaborative Learning, *British Journal of Educational Technology.* To appear.

32. A. Oulasvirta: Finding Meaningful Uses for Context-Aware Technologies: The humanistic Research Strategy, *Proceedings of ACM CHI 2004*, Vienna, Austria, April 2004.

33. A. Oulasvirta: Task-Processing Demands and Memory in Web Interaction: A Levels-of-Processing Approach, *Interacting with Computers*, 2004, vol. 16 (2), pp. 217-241. To appear.

34. A. Oulasvirta, L. Kärkkäinen, J. Laarni: Expectations and Memory in Link Search, *Computers in Human Behavior*, Elsevier. To appear.

35. A. Oulasvirta, P. Saariluoma: Long-Term Working Memory and Interrupting Messages in Human-Computer Interaction, *Behavior & Information Technology*, 2004, vol. 23 (1), pp. 53-64. To appear.

36. P. Sevon, V. Ollikainen, H.T.T. Toivonen: Tree Pattern Mining for Gene Mapping, *Information Sciences*. To appear

37. T.A. Thanaraj, S. Stamm, F. Clark, J.J. Riethoven, V. Le Texier, J. Muilu: The Alternative Splicing Database, *Nucleic Acids Research*, 2004 Jan 1;32(1):D64-9 ASD.

Articles in international edited works and conference proceedings with referee practice

38. S. Agrawal, S. Chaudhuri, G. Das, A. Gionis: Automated Ranking of Database Query Results, *First Biennial Conference on Innovative Data Systems Research (CIDR)*, Asilomar, California, USA, January 2003.

39. W. Buntine, P. Myllymäki, S. Perttu: Language Models for Intelligent Search Using Multinomial PCA, in B. Berendt, A. Hotho, D. Mladenic, M. van Someren, M. Spiliopoulou, G. Stumme (eds.), *Proceedings of the First European Web Mining Forum at the 14th European Conference on Machine Learning and the 7th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Cavtat-Dubrovnik, Croatia, 2003, Ruder Boskovic Institute, pp. 37-50.

40. W. Buntine, S. Perttu: Is Multinomial PCA Multi-faceted Clustering or Dimensionality Reduction?, in C.M. Bishop, B.J. Frey (eds.), *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, 2003, Society for Artificial Intelligence and Statistics, pp. 300-307.

41. A. Bykowski, J. Seppänen, J. Hollmén: Model-Independent Bounding of the Supports of Boolean Formulae in Binary Data, in P. Lanzi, R. Meo (eds.), *Database technologies for data mining*, 2003, Springer-Verlag.

42. T. Calders, B. Goethals: Minimal k-free Representations of Frequent Sets, in Proceedings of the 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'03), September 2003, Cavtat, Croatia, *Lecture Notes in Artificial Intelligence*, vol. 2838, Springer-Verlag, pp. 71-82.

43. C. Candolin, M. Komu, M. Kousa, J. Lundberg: An Implementation of HIP for Linux, *Linux Symposium 2003 Proceedings (The Ottawa Linux Symposium)*, Ottawa, Canada, July 23-26, 2003.

44. L. Eronen, F. Geerts, H.T.T. Toivonen: A Markov Chain Approach to Reconstruction of Long Haplotypes, *Pacific Symposium on Biocomputing (PSB2004)*, Hawaii, USA, January 2004, World Scientific, pp. 104-115.

45. P. Floréen, P. Kaski, J. Kohonen, P. Orponen: Multicast Time Maximization in Energy Constrained Wireless Networks, in *Proceedings of the 2003 Joint Workshop on Foundations of Mobile Computing (DIALM-POMC'03)*, San Diego, USA, September 2003.

46. F. Geerts, B. Goethals, T. Mielikäinen: What You Store is What You Get (extended abstract), in J-F. Boulicaut, S. Dzeroski (eds.), *Proceedings of the 2nd International Workshop on Knowledge Discovery in Inductive Databases*, 2003, pp. 60-69.

47. F. Geerts, B. Kuijpers: Deciding Termination of Query-Evaluation in Transitive Closure Logics for Constraint Databases, in *Proceedings of the 9th International Conference on Database Theory (ICDT 2003)*, Sienna, Italy, January 2003, pp. 190-206.

48. A. Gionis, T. Kujala, H. Mannila: Fragments of Orders, in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2003)*, Washington, DC, USA, August 2003, pp. 129-136.

49. A. Gionis, H. Mannila: Finding Recurrent Sources in Sequences, *RECOMB 2003*, in W. Miller, M. Vingron, S. Istrail, P. Pevzner, M. Waterman (eds.), April 2003, Berlin, Germany, pp. 123-130.

50. B. Goethals, M.J. Zaki: Advances in Frequent Itemset Mining Implementations, FIMI03, in *Proceedings of the FIMI'03 Workshop on Frequent Itemset Mining Implementations*, Melbourne, Florida, USA, November 2003

51. J. Heino, H.T.T. Toivonen: Automated Detection of Epidemics from the Usage Logs of a Physicians' Reference Database, *7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD-2003)*, Cavtat-Dubrovnik, Croatia, September 2003, Springer, pp. 180-191,

52. J. Himberg, A. Hyvärinen: Icasso: Software for Investigating the Reliability of Ica Estimates by Clustering and Visualization, in *Proceedings of International Workshop on Neural Networks for Signal Processing (NNSP 2003)*, Toulouse, France, 2003.

53. J. Hollmén, J. Seppänen, H. Mannila: Mixture Models and Frequent Sets: Combining Global and Local Methods for 0-1 Data, in *Proceedings of the Third SIAM International Conference on Data Mining (SDM'03)*, pp. 289-293.

54. P.O. Hoyer, A. Hyvärinen: Interpreting Neural Response Variability as Monte Carlo Sampling of the Posterior, in *Advances in Neural Information Processing Systems (NIPS)*, volume 15. MIT Press, 2003.

55. J. Hurri, A. Hyvärinen: Temporal Coherence, Natural Image Sequences, and the Visual Cortex, in *Advances in Neural Information Processing Systems (NIPS)*, 2003, vol. 15, MIT Press.

56. A. Hyvärinen, P.O. Hoyer, J. Hurri: Extensions of ICA as Models of Natural Images and Visual Processing, in *Proceedings of International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, Nara, Japan, pp. 963-974.

57. A. Hyvärinen, Y. Kano: Independent Component Analysis for Non-Normal Factor Analysis, in *New Developments in Psychometrics (proc. IMPS2001)*, 2003, pp. 649-656

58. E. Hyvönen, S. Kettula, V. Raatikka, S. Saarela, K. Viljanen: Finnish Museums on the Semantic Web, *The Twelfth International World Wide Web Conference*, Budapest, Hungary, May 2003, World Wide Web Consortium.

59. E. Hyvönen, S. Saarela, A. Styrman, K. Viljanen: Ontology-Based Image Retrieval, *The Twelfth International World Wide Web Conference*, Budapest, Hungary, May 2003, World Wide Web Consortium.

60. E. Hyvönen, K. Viljanen, A. Hätinen: Yellow Pages on the Semantic Web, *The Twelfth International World Wide Web Conference*, Budapest, Hungary, May 2003, World Wide Web Consortium.

61. P. Isto, M. Mäntylä, J. Tuominen: On Addressing the Run-Cost Variance in Randomized Motion Planners, in *Proceedings of 2003 IEEE International Conference on Robotics and Automation*, IEEE, pp. 2934-2939.

62. P. Isto, J. Tuominen, M. Mäntylä: Adaptive Strategies for Probabilistic Roadmap Construction, in *Proceedings of 2003 International Conference on Advanced Robotics*, 2003, pp. 682-687.

63. J. Kangasharju, S. Tarkoma, K. Raatikainen: Comparing SOAP Performance for Various Encodings, Protocols, and Connections, in M. Conti, S. Giordano, E. Gregori, S. Olariu, *Personal Wireless Communications, Series: Lecture Notes in Computer Science*, IFIP-TC6 8th International Conference, Venice, Italy, September 2003, vol. 2775, pp. 397-406.

64. A. Kankainen: UCPCD: User-Centered Product Concept Design, DUX, San Francisco, USA, June 5-7, 2003, pp. 1-13.

65. M. Koivisto, M. Perola, T. Varilo, W. Hennah, J. Ekelund, M. Lukk, L. Peltonen, E. Ukkonen, H. Mannila: An MDL Method for Finding Haplotype Blocks and for Estimating the Strength of Haplotype Block Boundaries, *Pacific Symposium on Biocomputing 2003 (PSB'03)*, Hawaii, USA, January 2003, R.B. Altman, A.K. Dukner, L. Hunter, T.A. Jung, T.E. Klein (eds.), World Scientific 2002, pp. 502-513.

66. P. Kontkanen, W. Buntine, P. Myllymäki, J. Rissanen, H. Tirri: Efficient Computation of Stochastic Complexity, in C.M. Bishop, B.J. Frey (eds.), *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, Society for Artificial Intelligence and Statistics, 2003, pp. 181-188.

67. P. Kontkanen, P. Myllymäki, W. Buntine, J. Rissanen, H. Tirri: An MDL Framework for Data Clustering, to appear in P. Grünwald, I. J. Myung, M. Pitt (eds.), *Advances in Minimum Description Length: Theory and Applications*, 2004, MIT Press.

68. J. Kurhila, M. Miettinen, P. Nokelainen, P. Floréen, H. Tirri: Joint Annotation and Knowledge Building in Collaborative E-Learning, in *Proceedings of the E-Learn 2003 Conference*, Phoenix, USA, November 2003, pp. 2249-2252.

69. J. Kurhila, M. Miettinen, P. Nokelainen, P. Floréen, H. Tirri: Peer-To-Peer Learning with Open-Ended Writable Web, in *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, The 8th Annual Conference on Innovation and Technology in Computer Science Education, Thessaloniki, Greece, June 2003, pp. 173-177.

70. A. Latva-Koivisto: A Design that Meets User's Goals Creates Usable Security. *Proceedings of the HCI International 2003*, Crete, Greece, June 2003.

71. A. Leino, H. Mannila, R-L Pitkänen: Rule Discovery and Probabilistic Modeling for Onomastic Data, in Knowledge Discovery in Databases: The 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD) 2003, Springer, pp. 291-302.

72. T. Lindholm: XML Three-Way Merge as a Reconciliation Engine for Mobile Data, in S. Banerjee, M. Cherniack, A. Labrinidis (eds.), *Proceedings of MobiDE 2003*, Third ACM International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE), September 2003, ACM Press, pp. 93-97.

73. J. Makkonen, H. Ahonen-Myka, M. Salmenkivi: Topic Detection and Tracking with Spatio-Temporal Evidence, in *Proceedings of the 25th European Conference on International Retrieval Research (ECIR)*, Pisa, Italy, 2003.

74. H. Mannila, M. Salmenkivi: Intensity Modeling of Genome Data. To appear in: J. Wang, D. Shasha, H.T.T. Toivonen, M. Zaki (eds.), *Data Mining in Bioinformatics*, London, England 2003, Springer-Verlag.

75. T. Mielikäinen: Frequency-based Views to Pattern Collections, in P. L. Hammer (ed.): *Proceedings of the IFIP/SIAM Workshop on Discrete Mathematics and Data Mining*, SIAM International Conference on Data Mining (2003), San Francisco, CA, USA, May 2003.

76. T. Mielikäinen: Finding All Occurring Sets of Interest, in J-F. Boulicaut, S, Dzeroski (eds.): *Proceedings of the 2nd International Workshop on Knowledge Discovery in Inductive Databases*, 2003, pp. 97-106.

77. T. Mielikäinen: Chaining Patterns, in G. Grieser, Y. Tanaka, A. Yamamoto (eds.), *Discovery Science - 6th International Conference, DS 2003*, Sapporo, Japan, October 2003, Proceedings, vol. 2843 of Lecture Notes in Artificial Intelligence, Springer 2003, pp. 233-244.

78. T. Mielikäinen: Change profiles, in X. Wu, A. Tuzhilin (eds.): *Proceedings of the 2003 IEEE International Conference on Data Mining (ICDM 2003)*, Melbourne, Florida, USA, November 2003, IEEE Computer Society, pp. 219-226.

79. T. Mielikäinen: On Inverse Frequent Set Mining, in W. Du, C. Clifton (eds.), in *Proceedings of the 2nd Workshop on Privacy Preserving Data Mining (PPDM)*, 2003, IEEE Computer Society, pp. 18-23.

80. T. Mielikäinen: Intersecting Data to Closed Sets With Constraints, in B. Goethals, M. J. Zaki (eds.), *Proceedings of the FIMI'03 Workshop on Frequent Itemset Mining Implementations*, Melbourne, Florida, USA, November 2003, vol. 90 of CEUR Workshop Proceedings, 2003, online CEUR-WS.org/Vol-90/.

81. T. Mielikäinen, Heikki Mannila: The Pattern Ordering Problem, in N. Lavrac, D. Gamberger, L. Todorovski, H. Blockeel (eds.): *Knowledge Discovery in Databases: PKDD 2003 - 7th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Cavtat-Dubrovnik, Croatia, September 2003, Proceedings, vol. 2838 of Lecture Notes in Artificial Intelligence, Springer 2003, pp. 327-338.

82. M. Miettinen, J. Kurhila, P. Nokelainen, P. Floréen, H. Tirri: EDUCOSM - Personalized Writable Web for Learning Communities, in *Proceedings of the ITCC 2003 Conference*, Las Vegas, USA, April 2003, pp. 37-42.

83. P. Nokelainen, J. Kurhila, M. Miettinen, P. Floréen, H. Tirri: Evaluating the Role of a Shared Document-Based Annotation Tool in Learner-Centered Collaborative Learning, *Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies*, Athens, Greece, July 2003, pp. 200-203.

84. P. Nokelainen, P. Ruohotie, M. Miettinen, J. Kurhila, H. Tirri: The Role of Inservice Teachers' Motivation, Learning Strategy and Social Ability Profiles in CSCL, in *The Proceedings of the SITE 2003 Conference*, Albaquerque, USA, March 2003, pp. 1518-1522.

85. V. Oksanen, M. Välimäki: Country Report – Finland, in I. Brown (ed.), *Implementing the European Union Copyright Directive*, London, England 2003, FIPR.

86. V. Oksanen, M. Välimäki: Music Collecting Societies and Webcasting, *Annual Conference of the Society for Economic Research on Copyright Issues (SERCI)*, June 2003.

87. V. Oksanen, M. Välimäki: Some Economic Aspects of the European Harmonization of Intellectual Property Rights in Software and its Impact to Eastern EU, *Conference on the Economic Integration of the Eastern EU*, Warsaw, Poland, July 2003.

88. A. Oulasvirta, H. Kuoppala: Reducing Task-Switch Errors by User Interface Design, *Proceedings of the 10th International Conference on Human – Computer Interaction*, 10th International Conference on Human-Computer Interaction, Crete, Greece, June 2003.

89. K. Pihkala, M. Honkala, P. Vuorimaa: Multimedia Web Forms, *SMIL Europe Conference*, Paris, France, February 2003.

90. J. Rissanen: Complexity and Information in Data, in A. Greven, G. Keller, G. Warnecke (eds.), *Entropy*, Princeton and Oxford 2003, Princeton University Press, pp. 299-312.

91. T. Saari, M. Turpeinen: Psychological Customization of Information - Basic Concepts, System Architecture and Applications, in *Proceedings of Interdisciplinary World Congress on Mass Customization and Personalization*, 2nd Interdisciplinary World Congress on Mass Customization and Personalization, Munich, Germany, October 2003.

92. J. Seppänen, E. Bingham, H. Mannila: A Simple Algorithm for Topic Identification in 0-1 Data, in 7th *European Conference on Principles and Practice of Knowledge discovery in Databases (PKDD'03)*, Dubrovnik, Croatia, September 2003.

93. P. Sevon, H. Toivonen, P. Onkamo: Gene Mapping by Pattern Discovery. To appear in J. Wang *et al* (eds.), *Data Mining in Bioinformatics*, Springer.

94. M. Sulkava, J. Hollmén: Finding Profiles of Forest Nutrition by Clustering of the Self-Organizing Map, in *Proceedings of the Workshop on Self-organizing Maps (WSOM'03)*, Hibikino, Japan, September 2003.

95. S. Tamminen, A. Oulasvirta, K. Toiskallio, A. Kankainen: Understanding Mobile Context, in L. Chittaro (ed.), *Proceedings of Mobile HCI 2003 (Mobile HCI 2003)*, Udine, Italy, Springer LNCS, pp. 18-35.

96. S. Tarkoma: Distributed Event Dissemination for Ubiquitous Agents, in R. Jardim-Gonçalves, J. Cha, A. Steiger-Garção, *Concurrent Engineering, Enhanced Interoperable Systems*, 10th ISPE International Conference on Concurrent Engineering (CE-2003), Madeira, Portugal, ISPE, UNINOVA, pp. 105-110.

97. S. Tarkoma, J. Kangasharju, K. Raatikainen: Client Mobility in Rendezvous-Notify, *Proceedings of the 2nd International Workshop on Distributed Event-Based Systems (DEBS'03)*, in conjunction with the ACM SIGMOD/PODS Conference, San Diego, California, USA, June 2003.

98. S. Tiitta: Identifying Elderly People's Needs for Communication and Mobility, *Include 2003 - Inclusive Design for Society and Business Conference*, London, UK, March 2003, The Helen Hamlyn Research Centre, pp. 7: 266-271.

99. S. Tiitta, A. Kankainen: Exploring Everyday Needs of Teenagers Related to Context-Aware Mobile Services, in *Proceedings of the 19th International Symposium on Human Factors in Telecommunication*, Berlin, Germany, December 2003, pp. 19-25.

100. H.T.T. Toivonen, P. Onkamo, P. Hintsanen, E. Terzi, P. Sevon: Data Mining for Gene Mapping. To appear in J. Zurada, M. Kantardzic (eds.), *New Generation of Data Mining Applications*, IEEE Press.

101. M. Turpeinen: Co-Evolution of Broadcast, Customized and Community-Created Media, in G.F. Lowe, T. Hujanen (eds.), *Broadcasting & Convergence: New Articulations of the Public Service Remit*, Göteborg, Sweden 2003, Nordicom, pp. 301-312.

102. J. Veijalainen, V. Terziyan, H. Tirri: Transaction Management for M-Commerce at a Mobile Terminal, *Proceedings of 36th Hawaii International Conference on Systems Sciences*, Big Island, Hawaii, USA, January 2003.

103. J. Vesanto, J. Hollmén: An Automated Report Generation Tool for the Data Understanding Phase, in A. Abraham, L. Jain (eds), *Innovations in Intelligent Systems: Design, Management and Applications, Studies in Fuzziness and Soft Computing*, 2003, Springer (Physica) Verlag, chapter 5.

104. J. Vesanto, M. Sulkava, J. Hollmén: On the Decomposition of the Self-Organizing Map Distortion Measure, in *Proceedings of the Workshop on Self-Organizing Maps (WSOM'03)*, Hibikino, Japan, September 2003, pp. 11-16.

105. M. Välimäki: Defending the Freedom of Speech in an Advanced Information Society: The Finnish Story, in C. Hardy, C. Möller (eds.): *Spreading the Word on the Internet, Reflections on Freedom of the Media and the Internet*, Vienna, Austria, 2003, OSCE, Office of the Representative on the Freedom of the Media, pp. 63-68.

106. M. Välimäki, V. Oksanen: The Impact of Free and Open Source Components on the Competition in Microcomputer Operating System Software Markets, in *Proceedings of 8th Annual Center for Tele-Information Conference*, Copenhagen, Denmark, 2003.

107. H. Wettig, P. Grünwald, T. Roos, P. Myllymäki, H. Tirri: When Discriminative Learning of Bayesian Network Parameters Is Easy, in G. Gottlob, T. Walsh (eds.), *Proceedings of the 18th International Conference on Artificial Intelligence*, 2003, Morgan Kaufmann, pp. 491-496.

108. H. Wettig, J. Lahtinen, T. Lepola, P. Myllymäki, H. Tirri: Bayesian Analysis of Online Newspaper Log Data, in *Proceedings of the 2003 Symposium on Applications and the Internet Workshops (SAINT 2003 Workshops)*, 2003, Los Alamitos, USA, IEEE Computer Society, pp. 282-287.

Articles in Finnish scientific journals with referee practice

109. R. Pulkkinen, M. Salmenkivi, A. Leino, H. Mannila: Louhi ja naistenmaa - kalevalaisten runojen Pohjolan sijainti peruskartan nimistön valossa, *Sananjalka: Suomen kielen seuran vuosikirja*, 45 (2003). (In Finnish.)

110. M. Välimäki: Vapaammista kirjallisten ja taiteellisten teosten lisenssiehdoista, *Defensor Legis*, 2003, No. 6, pp. 1064-1073. (In. Finnish.)

Articles in Finnish edited works and conference proceedings with referee practice

111. E. Hyvönen: Johdatus ohjelmistoalaan, in E. Hyvönen (ed.), *Ohjelmistoliiketoiminta*, Helsinki, Finland 2003, WSOY, pp. 1-7. (In Finnish.)

112. J. Kurhila, M. Miettinen, P. Nokelainen, H. Tirri: Enhancing Groupwork with Social Navigation in Collaborative Learning Environment, in *Proceedings of the 3rd Annual Finnish/Baltic Sea Conference on Computer Science Education*, Joensuu, Finland, October 2003, pp. 15-23.

113. O. Pitkänen: Julkisuus ja yksityisyys teknologioiden kehittyessä, in Letto-Vanamo P. (ed.), *Mitä saa sanoa. Viestintäoikeuden vuosikirja* (the Annals of Communications Law), Helsinki, Finland 2003, University of Helsinki, pp. 17 - 25. (In Finnish.)

114. O. Pitkänen: Oikeudet tietokoneohjelmistoihin, in E. Hyvönen (ed.), *Ohjelmistoliike-toiminta*, Helsinki, Finland 2003, WSOY, pp. 77-86. (In Finnish.)

115. T. Reti, M. Välimäki: Business Models for Content Distribution on Mobile Peer-to-Peer Networks, in *Proceedings of Frontiers in eBusiness Conference*, Tampere, Finland, 2003.

116. S. Tarkoma, M. Laukkanen: Adaptive Agent-based Service Composition for Wireless Terminals, in M. Klusch, A. Omicini, S. Ossowski, H. Laamanen, *Cooperative Information Agents VII*, LNAI 2728, Seventh International Workshop on Cooperative Information Agents (CIA-2003), Helsinki, Finland, August 2003, Springer, pp. 16-29.

Scientific monographs published abroad

117. B. Goethals, M.J. Zaki (eds.): *Proceedings of the FIMI'03 Workshop on Frequent Itemset Mining Implementations,* Melbourne, Florida, USA, November 2003.

118. D. Hand, H. Mannila, P. Smyth: Principles of Data Mining, MIT press 2001. ISBN: 0-262-98290-X. Second printing, 2002. Chinese Edition, China machine Press, 2003. ISBN: 7-111-11577-5. Polish Edition, to appear, 2004.

119. M. Zaki, J.T.L. Wang, H. Toivonen (eds.): *Proceedings of BIOKDD'03*, 3rd ACM SIGKDD Workshop on Data Mining in Bioinformatics. Washington DC, USA, August 2003, Report No. 03-11, Rensselaer Polytechnic Institute, Troy, NY. 2003.

Scientific monographs published in Finland

120. E. Hyvönen (ed.): *Ohjelmistoliiketoiminta*, Helsinki, Finland 2003, WSOY. (In Finnish.)

121. O. Pitkänen (ed.): Proceedings of the First International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet, Helsinki, Finland, Helsinki Institute for Information Technology HIIT Publications 2003-2.

Other scientific publications

122. H. Abie, J. Bing, B. Blobel, J. Delgado, B. Foyn, S. Karnouskos, P. Pharow, O. Pitkänen, D. Tzovaras: DigiRight: Network of Excellence for a Framework for Policy, Privacy, Security, Trust and Risk Management for Digital Rights Management, in O. Pitkänen (ed.), *Proceedings of the First International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet,* Helsinki, Finland, 27-28 August, 2003, Helsinki Institute for Information Technology HIIT, pp. 117-126 (HIIT Publications 2003-2).

123. H. Abie, B. Blobel, J. Delgado, S. Karnouskos, R. Marti, P. Pharow, O. Pitkänen, D. Tzovaras: DigiRight: Relevance to and Potential Impact on Europe's Need to Strengthen the Science and Technology Excellence on DRM, in O.Pitkänen (ed.), *Proceedings of the First International Mobile IPR Workshop: Rights Management of Information Products on the Mobile Internet*, Helsinki, Finland, 27-28 August, 2003, Helsinki Institute for Information Technology HIIT, pp. 127-134 (HIIT Publications 2003-2).

124. J. Blom, A. Kankainen, T. Kankainen, S. Tiitta: *Location-Aware Multiuser Messaging: Exploring the Evolution of Mobile Text-Based Communication Services*, Helsinki,

Finland 2003, Helsinki Institute for Information Technology HIIT (HIIT Technical Reports 2003-2).

125. P. Floréen, G. Lindén (eds.): *Context-Aware Methods,* Course on Context-Aware Computing 2003, Report C-2003-71, Department of Computer Science, University of Helsinki, December 2003.

126. P. Floréen, G.Lindén (eds*.): Context-Aware Scenarios*. Course on Context-Aware Computing 2003. Report C-2003-70, Department of Computer Science, University of Helsinki, December 2003.

127. A. Kankainen, A. Oulasvirta, M. Rantanen, S. Tiitta, T. Kankainen, M. Mäntylä: Arkipäivän läsnä-älyä, *Prosessori (Elektroniikan suunnittelu –erikoisjulkaisu)*, 2003, November, pp. 31-33. (In Finnish.)

128. M. Katajamaa, J. Hollmén: Simulation Model for Gene Expression Data, in R. Spang, P. Beziat, M. Vingron (eds.): *Currents in Computational Molecular Biology 2003*, pp. 249-250, April 2003. Poster presentation at the 7th Annual International Conference on Research in Computational Molecular Biology, RECOMB 2003.

129. E. Kurvinen, A. Oulasvirta: *Towards Social Awareness in Ubiquitous Computing: A Turntaking Approach*, Helsinki Institute for Information Technology HIIT, Helsinki, Finland 2003 (HIIT Technical Report 2003-1).

130. S. Ruosaari, J. Hollmén: Identifying Differentially Expressed Genes with Bootstrap-Based Testing, in *Bioinformatics 2003*, May 2003, Helsinki, Finland. A poster.

131. S. Ruosaari, J. Hollmén: Identifying Differentially Expressed Genes. In TICSP Workshop on Computational Systems Biology, June 2003, Tampere, Finland. A poster.

132. J. Saarela, J. Hollmén, D. Chen, P. Tainola, A. Jokiaho, A. Palotie, H. Mannila, L. Peltonen: Inheritance of Expression profiles: A Family Based Analysis Reveals Higher Similarity in the Gene Expression Profiles of Related Individuals. Poster presentation at the Annual meeting of the American Society of Human Genetics (ASHG), November 2003.

133. A. Soininen (ed.), O. Pitkänen, M. Välimäki, V. Oksanen, T.Reti: *Mobile IPR Final Report*, Helsinki, Finland 2003, Helsinki Institute for Information Technology HIIT (HIIT Publications 2003-3).

134. H. Wikman, K. Salmenkivi, J. Seppänen, E. Kettunen, K. Vainio-Siukola, J. Hollmén, A. Karjalainen, S. Knuutila, S. Anttila: *Down-Regulation of Cavelion 1 and Cavelion 2 in Lung Cancer Revealed by the Combined Use of cDNA and Tissue Microarreys*, poster presentation at the ESF Program in Functional Genomics: 1st European Conference, Prague, Czech Republic, May 2003.

135. M.J. Zaki, J.T.L. Wang, H.T.T.Toivonen: BIOKDD 2002: Recent Advances in Data Mining for Bioinformatics, January 2003, *SIGKDD Explorations* 4 (2): 112 - 114.

Computer programs (and algorithms)

136. L. Eronen, F. Geerts, H.T.T. Toivonen: HaploRec. A Markov Chain Approach to Reconstruction of Long Haplotypes. Available at: http://www.cs.helsinki.fi/group/genetics/haplotyping,html

137. A. Gionis: Implementation of Sequent Segmentation with Recurrent Sources. Available at: http://www.cs.helsinki.fi/u/gionis/khcode.html

138. J. Himberg, A. Hyvärinen: Icasso: Software for Investigating the Reliability of ICA Estimates by Clustering and Visualization. Available at http://www.cis.hut.fi/jhimberg/icasso/

139. M. Koivisto, H. Mannila, M. Lukk, K. Sood, E. Ukkonen: MDLBlockFinder -- A Minimum Description Length -Based Method for the Analysis of Haplotype Blocks (c) 2003. University of Helsinki and HIIT/BRU. Available at http://www.cs.helsinki.fi/u/mkhkoivi/software/blocks/MDLBlockFinder.html

140. M.Miettinen, P. Nokelainen, H.Tirri, J.Kurhila: Software Tools for E-learning. Available at: http://cosco.hiit.fi/edutech/

141. T.Silander, P.Uronen, P.Myllymäki, H.Tirri: B-Course, a Web-Based Data Analysis Tool for Bayesian Modeling (v2.0). Available at: http://b-course.hiit.fi/

142. V. Tuulos: Finnish Parliament Member Clustering Engine. Available at: http://cosco.hiit.fi/eduskunta/index.jsp

143. V. Tuulos: FastPCA: A Fast, Distributed, Pure C, Iterative PCA. Available at: http://cosco.hiit.fi/search/software.html

Degrees

PhD thesis

144. E. Bingham: *Advances in Independent Component Analysis with Applications to Data Mining*. PhD Thesis. Helsinki University of Technology, Laboratory of Computer and Information Science. 2003.

Master's theses

145. J. Heino: *Automated Detection of Epidemics from the Usage Logs of a Physicians' Reference Database*. Master's thesis. University of Helsinki, Department of Computer Science. 2003.

146. H. Hiisilä: *Finding Components In Discrete Biosequences*. Master's thesis. Helsinki University of Technology, Department of Computer Science and Engineering. 2003.

147. T. Kankainen: *User Need Driven Product Concept Development - Ubiquitous Computing Scenario Creation and Evaluation*. Master of Arts thesis. University of Art and Design Helsinki UIAH, Department of Product and Strategic Design, 2003.

148. J. Päivärinta: *Customer Care Tool for WWW Services*. Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering, 2003.

149. J. Ukkonen: *Klusterointi ad –hod verkkojen reitityksessä*. Master's thesis. University of Helsinki, Department of Computer Science. 2003. (In Finnish.)

Notifications on invention

150. K. Rimey, J. Heikkilä, Y. Kortesniemi: *Anonymous Stochastic Census, a Method for Estimating the Total Population of Entities in a Distributed System*.

B List of Personnel

Apiola, Mikko[3]	Research Assistant	ARU
Bingham, Ella	Researcher	BRU
Blom, Jan[2]	Senior Research Scientist	ARU
Buntine, Wray	Senior Research Scientist	ARU
Eronen, Lauri	Researcher	BRU
Floréen, Patrik	Research Coordinator	ARU/BRU
Geerts, Floris	Researcher	BRU
Gionis, Aristides[2]	Researcher	BRU
Goethals, Bart	Researcher	BRU
Haiminen, Niina[2]	Researcher	BRU
Hakli, Raul[1]	Researcher	ARU
Hassel, Henri[3]	IT Specialist	ARU
Hasu, Tero[3]	Researcher	ARU
Heikkilä, Juho	Research Assistant	ARU
Heino, Jaana[2]	Researcher	BRU
Herrera Avila, Fernando[2]	Research Assistant	ARU
Hietanen, Herkko[1]	Research Assistant	ARU
Hiisilä, Heli	Researcher	BRU
Himanen, Pekka[2]	Senior Research Scientist	ARU
Hintsanen, Petteri	Assistant Researcher	BRU
Holi, Markus[2]	Researcher	ARU
Hollmén, Jaakko	Professor	BRU
Honkala, Mikko[2]	Project Manager	ARU
Huvio, Eero[2]	Researcher	BRU
Hyvärinen, Aapo[2]	Senior Research Fellow (Academy of Finland)	BRU
Hyvönen, Eero[1]	Senior Research Scientist	ARU
Hyvönen, Saara (on maternity leave)	Researcher	BRU
Juhala, Jari-Pekka[1]	Assistant Researcher	BRU
Junnila, Miikka[3]	Research Assistant	ARU
Kalliokoski, Matti[1]	Project Manager	ARU
Kanerva, Pekka	Researcher	ARU
Kangasharju, Jaakko	Researcher	ARU
Kankainen, Anu[2]	Researcher	ARU
Kankainen, Tomi[2]	Project Manager	ARU
Karp, Elina[3]	Assistant Researcher	BRU
Kemppinen, Jukka	Principal Scientist	ARU
Kettula, Suvi-Päivikki[1]	Researcher	ARU
Kiiskinen, Miika[3]	Research Assistant	ARU
Kohonen, Jukka	Researcher	BRU
Koivisto, Mikko	Researcher	BRU
Kollin, Jussi[2]	Assistant Researcher	BRU
Komu, Miika[1]	Research Assistant	ARU
Kontiainen, Mikko	Researcher/Coordinator	ARU
Kontkanen, Petri	Researcher	ARU
Korpiaho, Kalle[1]	Assistant Researcher	BRU
Kortesniemi, Yki	Project Manager	ARU
Kousa, Mika	Research Assistant	ARU
Kuikkaniemi, Kai[3]	Research Assistant	ARU
Kujala, Teija	Researcher	BRU
Laasonen, Kari	Researcher	BRU

Lahtinen, Jussi	Researcher	ARU
Latva-Koivisto, Antti[2]	Researcher	ARU
Leino, Antti	Researcher	BRU
Leinonen, Jani[2]	Researcher	ARU
Lepola, Tuomas	Researcher/HeCSE	ARU
Lindén, Greger	Co-ordinator	BRU
Lindgren, Petri[3]	Research Assistant	ARU
Lindholm Tancred	Researcher	ARU
Lindholm, Taina	Project Secretary	ARU
Lindström, Jan[2]	Senior Research Scientist	ARU
Löfström, Jaakko	Research Assistant	ARU
Mannila, Heikki	Research Director	BRU
Mielikäinen, Taneli	Researcher	BRU
Miettinen, Miikka	Researcher	ARU
Miettunen, Pirkko	Institute Secretary	ARU
Mononen, Tommi	Researcher	ARU
Muilu, Juha[1]	Database Manager	BRU
Myllymäki, Petri[1]	Senior Research Scientist	ARU
Mäkelä, Eetu[2]	Research Assistant	ARU
Mäkelä, Mikko[2]	Research Assistant	ARU
Mäntylä, Martti	Research Director	ARU
Mäntylä, Teemu[3]	Research Assistant	ARU
Nikander, Pekka	Ericsson Visiting Senior Research Scientist	ARU
Nokelainen, Petri[3]	Researcher	ARU
Nyman, Jan[2]	Research Assistant	ARU
Oksanen, Inka	Project Assistant	BRU
Oksanen, Kenneth[2]	Researcher	ARU
Oksanen, Ville	Researcher	ARU
Onkamo, Päivi	Researcher	BRU
Oulasvirta, Antti	Researcher	ARU
Patrikainen, Anne	Researcher	BRU
Perkiö, Jukka[2]	Research Assistant	ARU/BRU
Perttu, Sami	Researcher/HeCSE	ARU
Pitkänen, Olli	Program Coordinator	ARU
Porochine, Vladimir[2]	Research Assistant	ARU
Puolakka, Anton	Program Coordinator	ARU
Päivärinta, Juha	Research Assistant	ARU
Raatikainen, Kimmo[1]	Principal Scientist	ARU
Raento, Mika	Researcher	BRU
Rantakokko, Mika[3]	Researcher	ARU
Rantanen, Matti	Researcher	ARU
Rantanen, Mervi[2]	Planning Officer	ARU
Reti, Tommo	Researcher	ARU
Rimey, Kenneth	Senior Research Scientist	ARU
Rissanen, Jorma[2]	HIIT Fellow	ARU
Roos, Teemu	Researcher/HeCSE	ARU
Ruosaari, Salla	Researcher	BRU
Rüger, Torsten[2]	Researcher	ARU
Saarela, Samppa	Research Assistant	ARU
Saaresto, Marko	Research Assistant	ARU
Saari, Timo[3]	Senior Research Scientist	ARU
Saijos, Jani[2]	Assistant Researcher	BRU
Salmenkivi, Marko	Researcher	BRU
Salminen, Mirva[2]	Research Assistant	ARU

Salovaara, Antti[2]	Research Assistant	ARU
Samarine, Mikhail[3]	Researcher	ARU
Sarvas, Risto	Researcher	ARU
Seppänen, Jouni	Researcher	BRU
Sevon, Petteri	Researcher	BRU
Shimizu, Shohei[2]	Researcher	BRU
Sikha, Para Brahmendra Rao[2]	Researcher	ARU
Silander, Tomi	Researcher	ARU
Slavov, Kristian[2]	Research Assistant	ARU
Soininen, Aura	Researcher	ARU
Sood, Kismat[1]	Researcher	BRU
Styrman, Avril[3]	Research Assistant	ARU
Syreeni, Ahti[3]	Research Assistant	ARU
Tarkoma, Sasu	Project Manager	ARU
Tatti, Nikolaj	Researcher	BRU
Terzi, Evimaria	Researcher	BRU
Thalainayar Balasubramanian, Ramya[2]	Research Assistant	ARU
Tiitta, Sauli	Researcher	ARU
Tirri, Henry[1]	Principal Scientist	ARU
Toivonen, Hannu	Professor	BRU
Tonteri, Pekka	IT Manager	ARU
Tuominen, Antti	Researcher	ARU
Tuominen, Heikki	Researcher	BRU
Turpeinen, Marko[3]	Senior Research Scientist	ARU
Tuulos, Ville	Researcher	ARU
Ukkonen, Johannes[2]	Researcher	BRU
Uronen, Pekka	Researcher/HeCSE	ARU
Valo, Arttu[3]	Research Assistant	ARU
Valtonen, Kimmo	Researcher	ARU
Wettig, Hannes	Researcher	ARU
Vierinen, Juha-Pekka	Research Assistant	ARU
Viljanen, Kim	Research Assistant	ARU
Virtanen, Otso[2]	Researcher	BRU
Virtanen, Perttu	Project Manager	ARU
Vuorenmaa, Janne[3]	Research Assistant	ARU
Vuorimaa, Petri[3]	Senior Research Scientist	ARU
Välimäki, Mikko	Researcher	ARU

----

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