



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL
TRAINING AT THE UNIVERSITY OF HELSINKI 2005–2010

RC-Specific Evaluation of ALKO – Algorithms and Data Analysis

Seppo Saari & Antti Moilanen (Eds.)



Evaluation Panel: Natural Sciences

INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL
TRAINING AT THE UNIVERSITY OF HELSINKI 2005–2010

RC-Specific Evaluation of ALKO – Algorithms and Data Analysis

Seppo Saari & Antti Moilanen (Eds.)

University of Helsinki
Administrative Publications 80/70
Evaluations
2012

Publisher: University of Helsinki Editors: Seppo Saari & Antti Moilanen
--

Title: International Evaluation of Research and Doctoral Training at the University of Helsinki 2005–2010 : RC-Specific Evaluation of ALKO – Algorithms and Data Analysis	Type of publication: Evaluations
Summary: Researcher Community (RC) was a new concept of the participating unit in the evaluation. Participation in the evaluation was voluntary and the RCs had to choose one of the five characteristic categories to participate. Evaluation of the Researcher Community was based on the answers to the evaluation questions. In addition a list of publications and other activities were provided by the TUHAT system. The CWTS/Leiden University conducted analyses for 80 RCs and the Helsinki University Library for 66 RCs. Panellists, 49 and two special experts in five panels evaluated all the evaluation material as a whole and discussed the feedback for RC-specific reports in the panel meetings in Helsinki. The main part of this report is consisted of the feedback which is published as such in the report. Chapters in the report: 1. Background for the evaluation 2. Evaluation feedback for the Researcher Community 3. List of publications 4. List of activities 5. Bibliometric analyses The level of the RCs' success can be concluded from the written feedback together with the numeric evaluation of four evaluation questions and the category fitness. More conclusions of the success can be drawn based on the University-level report.	
RC-specific information:	
Main scientific field of research: Natural Sciences	RC-specific keywords: computer science, algorithms, machine learning, data mining, data analysis, complex systems, string methods, neuroinformatics, bioinformatics
Participation category: 1. Research of the participating community represents the international cutting edge in its field	
RC's responsible person: Ukkonen, Esko	
Keywords: Research Evaluation, Meta-evaluation, Doctoral Training, Bibliometric Analyses, Researcher Community	

Series title and number: University of Helsinki, Administrative Publications 80/70, Evaluations	
ISSN: 1795-5513 (Online)	ISBN: 978-952-10-7490-5 (PDF)
Total number of pages: 124	Language: English
Additional information: Cover graphics: Päivi Talonpoika-Ukkonen Enquiries: seppo.o.saari@helsinki.fi	Internet address: http://www.helsinki.fi/julkaisut/aineisto/rc_evaluation_2012/hallinnon_julkaisu_80_70_2012.pdf

Contents

Panel members	1
1 Introduction to the Evaluation	5
1.1 RC-specific evaluation reports	5
1.2 Aims and objectives in the evaluation	5
1.3 Evaluation method	5
1.4 Implementation of the external evaluation	6
1.5 Evaluation material	7
1.6 Evaluation questions and material	8
1.7 Evaluation criteria	10
1.8 Timetable of the evaluation	13
1.9 Evaluation feedback – consensus of the entire panel	13
2 Evaluation feedback.....	15
2.1 Focus and quality of the RC's research	15
2.2 Practises and quality of doctoral training	16
2.3 The societal impact of research and doctoral training.....	16
2.4 International and national (incl. intersectoral) research collaboration and researcher mobility	17
2.5 Operational conditions	18
2.6 Leadership and management in the researcher community	18
2.7 External competitive funding of the RC	19
2.8 The RC's strategic action plan for 2011–2013	19
2.9 Evaluation of the category of the RC in the context of entity of the evaluation material (1-8)	20
2.10 Short description of how the RC members contributed the compilation of the stage 2 material ...	20
2.11 How the UH's focus areas are presented in the RC's research	20
2.12 RC-specific main recommendations	21
2.13 RC-specific conclusions.....	21
2.14 Preliminary findings in the Panel-specific feedback	21
3 Appendices	23

Foreword

The evaluation of research and doctoral training is being carried out in the years 2010–2012 and will end in 2012. The steering group appointed by the Rector in January 2010 set the conditions for participating in the evaluation and prepared the Terms of Reference to present the evaluation procedure and criteria. The publications and other scientific activities included in the evaluation covered the years 2005–2010.

The participating unit in the evaluation was defined as a Researcher Community (RC). To obtain a critical mass with university-level impact, the number of members was set to range from 20 to 120. The RCs were required to contain researchers in all stages of their research career, from doctoral students to principal investigators (PIs). All in all, 136 Researcher Communities participated in this voluntary evaluation, 5857 persons in total, of whom 1131 were principal investigators. PIs were allowed to participate in two communities in certain cases, and 72 of them used this opportunity and participated in two RCs.

This evaluation enabled researchers to define RCs from the “bottom up” and across disciplines. The aim of the evaluation was not to assess individual performance but a community with shared aims and researcher-training activities. The RCs were able to choose among five different categories that characterised the status and main aims of their research. The steering group considered the process of applying to participate in the evaluation to be important, which led to the establishment of these categories. In addition, providing a service for the RCs to enable them to benchmark their research at the global level was a main goal of the evaluation.

The data for the evaluation consisted of the RCs’ answers to evaluation questions on supplied e-forms and a compilation extracted from the TUHAT – Research Information System (RIS) on 12 April 2011. The compilation covered scientific and other publications as well as certain areas of scientific activities. During the process, the RCs were asked to check the list of publications and other scientific activities and make corrections if needed. These TUHAT compilations are public and available on the evaluation project sites of each RC in the TUHAT-RIS.

In addition to the e-form and TUHAT compilation, University of Leiden (CWTS) carried out bibliometric analyses from the articles included in the Web of Science (WoS). This was done on University and RC levels. In cases where the publication forums of the RC were clearly not represented by the WoS data, the Library of the University of Helsinki conducted a separate analysis of the publications. This was done for 66 RCs representing the humanities and social sciences.

The evaluation office also carried out an enquiry targeted to the supervisors and PhD candidates about the organisation of doctoral studies at the University of Helsinki. This and other documents describing the University and the Finnish higher education system were provided to the panellists.

The panel feedback for each RC is unique and presented as an entity. The first collective evaluation reports available for the whole panel were prepared in July–August 2011. The reports were accessible to all panel members via the electronic evaluation platform in August. Scoring from 1 to 5 was used to complement written feedback in association with evaluation questions 1–4 (scientific focus and quality, doctoral training, societal impact, cooperation) and in addition to the category evaluating the fitness for participation in the evaluation. Panellists used the international level as a point of comparison in the evaluation. Scoring was not expected to go along with a preset deviation.

Each of the draft reports were discussed and dealt with by the panel in meetings in Helsinki (from 11 September to 13 September or from 18 September to 20 September 2011). In these meetings the panels also examined the deviations among the scores and finalised the draft reports together.

The current RC-specific report deals shortly with the background of the evaluation and the terms of participation. The main evaluation feedback is provided in the evaluation report, organised according to the evaluation questions. The original material provided by the RCs for the panellists has been attached to these documents.

On behalf of the evaluation steering group and office, I sincerely wish to thank you warmly for your participation in this evaluation. The effort you made in submitting the data to TUHAT-RIS is gratefully acknowledged by the University. We wish that you find this panel feedback useful in many ways. The bibliometric profiles may open a new view on your publication forums and provide a perspective for discussion on your choice of forums. We especially hope that this evaluation report will help you in setting the future goals of your research.

Johanna Björkroth
Vice-Rector
Chair of the Steering Group of the Evaluation

Steering Group of the evaluation

Steering group, nominated by the Rector of the University, was responsible for the planning of the evaluation and its implementation having altogether 22 meetings between February 2010 and March 2012.

Chair

Vice-Rector, professor **Johanna Björkroth**

Vice-Chair

Professor **Marja Airaksinen**

Chief Information Specialist, Dr **Maria Forsman**

Professor **Arto Mustajoki**

University Lecturer, Dr **Kirsi Pyhälä**

Director of Strategic Planning and Development, Dr **Ossi Tuomi**

Doctoral candidate, MSocSc **Jussi Vauhkonen**

Panel members

CHAIR

Professor Jan-Otto Carlsson

Materials science in chemistry and physics, nanotechnology, inorganic chemistry
Uppsala University, Sweden

VICE-CHAIR

Professor Jan van Leeuwen

Computer science, information technology
University of Utrecht, the Netherlands

Professor Caitlin Buck

Probability and statistics, archeology, palaeoenvironmental science
University of Sheffield, Great Britain

Professor David Colton

Mathematics, inverse problems of acoustic and electromagnetic scattering
University of Delaware, USA

Professor Jean-Pierre Eckmann

Mathematics, dynamical systems, mathematical physics
University of Geneva, Switzerland

Professor Ritske Huismans

Geosciences, geodynamics
University of Bergen, Norway

Professor Jukka Jurvelin

Medical physics and engineering
University of Eastern Finland

Professor Lea Kauppi

Environmental sciences, water research
The Finnish Environment Institute, Finland

Professor Riitta Keiski

Chemical engineering, heterogeneous catalysis, environmental technology, mass and heat transfer processes
University of Oulu, Finland

Professor Mats Larsson

Experimental molecular physics, chemical dynamics, molecular spectroscopy, astrobiology
Stockholm University, Sweden

Professor Holger Stark

Medicinal, organic and pharmaceutical chemistry, pharmacology
Johann Wolfgang Goethe Universität, Germany

The panel, independently, evaluated all the submitted material and was responsible for the feedback of the RC-specific reports. The panel members were asked to confirm whether they had any conflict of interests with the RCs. If this was the case, the panel members disqualified themselves in discussion and report writing.

Added expertise to the evaluation was contributed by the members from the other panels.

Experts from the Other Panels

Professor Barbara Koch, from the Panel of Biological, Agricultural and Veterinary Sciences

Professor Peter York, from the Panel of Medicine, Biomedicine and Health Sciences

EVALUATION OFFICE

Dr Seppo Saari, Doc., Senior Adviser in Evaluation, was responsible for the entire evaluation, its planning and implementation and acted as an Editor-in-chief of the reports.

Dr Eeva Sievi, Doc., Adviser, was responsible for the registration and evaluation material compilations for the panellists. She worked in the evaluation office from August 2010 to July 2011.

MSocSc Paula Ranne, Planning Officer, was responsible for organising the panel meetings and all the other practical issues like agreements and fees and editing a part the RC-specific reports. She worked in the evaluation office from March 2011 to January 2012.

Mr Antti Mollanen, Project Secretary, was responsible for editing the reports. He worked in the evaluation office from January 2012 to April 2012.

TUHAT OFFICE

Provision of the publication and other scientific activity data

Mrs Alja Kaltera, Project Manager of TUHAT-RIS served the project ex officio providing the evaluation project with the updated information from TUHAT-RIS. The TUHAT office assisted in mapping the publications with CWTS/University of Leiden.

MA Liisa Ekebon, Assisting Officer, served in TUHAT-RIS updating the publications for the evaluation. She also assisted the UH/Library analyses.

BA Liisa Jäppinen, Assisting Officer, served in TUHAT-RIS updating the publications for the evaluation.

HELSINKI UNIVERSITY LIBRARY

Provision of the publication analyses

Dr Maria Forsman, Chief Information Specialist in the Helsinki University Library, managed with her 10 colleagues the bibliometric analyses in humanities, social sciences and in other fields of sciences where CWTS analyses were not applicable.

Acronyms and abbreviations applied in the report

External competitive funding

AF – Academy of Finland
TEKES - Finnish Funding Agency for Technology and Innovation
EU - European Union
ERC - European Research Council
International and national foundations
FP7/6 etc. /Framework Programmes/Funding of European Commission

Evaluation marks

Outstanding (5)
Excellent (4)
Very Good (3)
Good (2)
Sufficient (1)

Abbreviations of Bibliometric Indicators

P - Number of publications
TCS – Total number of citations
MCS - Number of citations per publication, excluding self-citations
PNC - Percentage of uncited publications
MNCS - Field-normalized number of citations per publication
MNJS - Field-normalized average journal impact
THCP10 - Field-normalized proportion highly cited publications (top 10%)
INT_COV - Internal coverage, the average amount of references covered by the WoS
WoS – Thomson Reuters Web of Science Databases

Participation category

Category 1. The research of the participating community represents the international cutting edge in its field.
Category 2. The research of the participating community is of high quality, but the community in its present composition has yet to achieve strong international recognition or a clear break-through.
Category 3. The research of the participating community is distinct from mainstream research, and the special features of the research tradition in the field must be considered in the evaluation.
Category 4. The research of the participating community represents an innovative opening.
Category 5. The research of the participating community has a highly significant societal impact.

Research focus areas of the University of Helsinki

Focus area 1: The basic structure, materials and natural resources of the physical world
Focus area 2: The basic structure of life
Focus area 3: The changing environment – clean water
Focus area 4: The thinking and learning human being
Focus area 5: Welfare and safety
Focus area 6: Clinical research
Focus area 7: Precise reasoning
Focus area 8: Language and culture
Focus area 9: Social justice
Focus area 10: Globalisation and social change

1 Introduction to the Evaluation

1.1 RC-specific evaluation reports

The participants in the evaluation of research and doctoral training were Researcher Communities (hereafter referred to as the RC). The RC refers to the group of researchers who registered together in the evaluation of their research and doctoral training. Preconditions in forming RCs were stated in the Guidelines for the Participating Researcher Communities. The RCs defined themselves whether their compositions should be considered well-established or new.

It is essential to emphasise that the evaluation combines both meta-evaluation¹ and traditional research assessment exercise and its focus is both on the research outcomes and procedures associated with research and doctoral training. The approach to the evaluation is enhancement-led where self-evaluation constituted the main information. The answers to the evaluation questions formed together with the information of publications and other scientific activities an entity that was to be reviewed as a whole.

The present evaluation recognizes and justifies the diversity of research practices and publication traditions. Traditional Research Assessment Exercises do not necessarily value high quality research with low volumes or research distinct from mainstream research. It is challenging to expose the diversity of research to fair comparison. To understand the essence of different research practices and to do justice to their diversity was one of the main challenges of the present evaluation method. Understanding the divergent starting points of the RCs demanded sensitivity from the evaluators.

1.2 Aims and objectives in the evaluation

The aims of the evaluation are as follows:

- to improve the level of research and doctoral training at the University of Helsinki and to raise their international profile in accordance with the University's strategic policies. The improvement of doctoral training should be compared to the University's policy.²
- to enhance the research conducted at the University by taking into account the diversity, originality, multidisciplinary nature, success and field-specificity,
- to recognize the conditions and prerequisites under which excellent, original and high-impact research is carried out,
- to offer the academic community the opportunity to receive topical and versatile international peer feedback,
- to better recognize the University's research potential.
- to exploit the University's TUHAT research information system to enable transparency of publishing activities and in the production of reliable, comparable data.

1.3 Evaluation method

The evaluation can be considered as an enhancement-led evaluation. Instead of ranking, the main aim is to provide useful information for the enhancement of research and doctoral training of the participating RCs. The comparison should take into account each field of science and acknowledge their special character.

¹ The panellists did not read research reports or abstracts but instead, they evaluated answers to the evaluation questions, tables and compilations of publications, other scientific activities, bibliometrics or comparable analyses.

² [Policies on doctoral degrees and other postgraduate degrees at the University of Helsinki.](#)

The comparison produced information about the present status and factors that have lead to success. Also challenges in the operations and outcomes were recognized.

The evaluation approach has been designed to recognize better the significance and specific nature of researcher communities and research areas in the multidisciplinary top-level university. Furthermore, one of the aims of the evaluation is to bring to light those evaluation aspects that differ from the prevalent ones. Thus the views of various fields of research can be described and research arising from various starting points understood better. The doctoral training is integrated into the evaluation as a natural component related to research. Operational processes of doctoral training are being examined in the evaluation.

Five stages of the evaluation method were:

1. Registration – Stage 1
2. Self-evaluation – Stage 2
3. TUHAT³ compilations on publications and other scientific activities⁴
4. External evaluation
5. Public reporting

1.4 Implementation of the external evaluation

Five Evaluation Panels

Five evaluation panels consisted of independent, renowned and highly respected experts. The main domains of the panels are:

1. biological, agricultural and veterinary sciences
2. medicine, biomedicine and health sciences
3. natural sciences
4. humanities
5. social sciences

The University invited 10 renowned scientists to act as chairs or vice-chairs of the five panels based on the suggestions of faculties and independent institutes. Besides leading the work of the panel, an additional role of the chairs was to discuss with other panel chairs in order to adopt a broadly similar approach. The panel chairs and vice-chairs had a pre-meeting on 27 May 2011 in Amsterdam.

The panel compositions were nominated by the Rector of the University 27 April 2011. The participating RCs suggested the panel members. The total number of panel members was 50. The reason for a smaller number of panellists as compared to the previous evaluations was the character of the evaluation as a meta-evaluation. The panellists did not read research reports or abstracts but instead, they evaluated answers to the evaluation questions, tables and compilations of publications, other scientific activities, bibliometrics and comparable analyses.

The panel meetings were held in Helsinki:

- On 11–13 September 2011: (1) biological, agricultural and veterinary sciences, (2) medicine, biomedicine and health sciences and (3) natural sciences.
- On 18–20 September 2011: (4) humanities and (5) social sciences.

³ TUHAT (acronym) of Research Information System (RIS) of the University of Helsinki

⁴ Supervision of thesis, prizes and awards, editorial work and peer reviews, participation in committees, boards and networks and public appearances.

1.5 Evaluation material

The main material in the evaluation was the RCs' self-evaluations that were qualitative in character and allowed the RCs to choose what was important to mention or emphasise and what was left unmentioned.

The present evaluation is exceptional at least in the Finnish context because it is based on both the evaluation documentation (self-evaluation questions, publications and other scientific activities) and the bibliometric reports. All documents were delivered to the panellists for examination.

Traditional bibliometrics can be reasonably done mainly in medicine, biosciences and natural sciences when using the Web of Science database, for example. Bibliometrics, provided by CWTS/The Centre for Science and Technology Studies, University of Leiden, cover only the publications that include WoS identification in the TUHAT-RIS.

Traditional bibliometrics are seldom relevant in humanities and social sciences because the international comparable databases do not store every type of high quality research publications, such as books and monographs and scientific journals in other languages than English. The Helsinki University Library has done analysis to the RCs, if their publications were not well represented in the Web of Science databases (RCs should have at least 50 publications and internal coverage of publications more than 40%) – it meant 58 RCs. The bibliometric material for the evaluation panels was available in June 2011. The RC-specific bibliometric reports are attached at the end of each report.

The panels were provided with the evaluation material and all other necessary background information, such as the basic information about the University of Helsinki and the Finnish higher education system.

Evaluation material

1. Registration documents of the RCs for the background information
2. Self evaluation material – answers to the evaluation questions
3. Publications and other scientific activities based on the TUHAT RIS:
 - 3.1. statistics of publications
 - 3.2. list of publications
 - 3.3. statistics of other scientific activities
 - 3.4. list of other scientific activities
4. Bibliometrics and comparable analyses:
 - 4.1. Analyses of publications based on the verification of TUHAT-RIS publications with the Web of Science publications (CWTS/University of Leiden)
 - 4.2. Publication statistics analysed by the Helsinki University Library - mainly for humanities and social sciences
5. University level survey on doctoral training (August 2011)
6. University level analysis on publications 2005–2010 (August 2011) provided by CWTS/University of Leiden

Background material

University of Helsinki

- [Basic information about the University of the Helsinki](#)
- [The structure of doctoral training at the University of Helsinki](#)
- Previous evaluations of research at the University of Helsinki – links to the reports: [1998](#) and [2005](#)

The Finnish Universities/Research Institutes

- [Finnish University system](#)
- [Evaluation of the Finnish National Innovation System](#)
- [The State and Quality of Scientific Research in Finland. Publication of the Academy of Finland 9/09.](#)

The evaluation panels were provided also with other relevant material on request before the meetings in Helsinki.

1.6 Evaluation questions and material

The participating RCs answered the following evaluation questions which are presented according to the evaluation form. In addition, TUHAT RIS was used to provide the **additional material** as explained. For giving the feedback to the RCs, the panellists received the evaluation feedback form constructed in line with the evaluation questions:

1. Focus and quality of the RC's research

- Description of
 - the RC's research focus.
 - the quality of the RC's research (incl. key research questions and results)
 - the scientific significance of the RC's research in the research field(s)
- Identification of the ways to strengthen the focus and improve the quality of the RC's research

The additional material: TUHAT compilation of the RC's publications, analysis of the RC's publications data (provided by University of Leiden and the Helsinki University Library)

A written feedback from the aspects of: scientific quality, scientific significance, societal impact, innovativeness

- Strengths
- Areas of development
- Other remarks
- Recommendations

Numeric evaluation: OUTSTANDING (5), EXCELLENT (4), VERY GOOD (3), GOOD (2), SUFFICIENT (1)

2. Practises and quality of doctoral training

- Organising of the doctoral training in the RC. Description of the RC's principles for:
 - recruitment and selection of doctoral candidates
 - supervision of doctoral candidates
 - collaboration with faculties, departments/institutes, and potential graduate schools/doctoral programmes
 - good practises and quality assurance in doctoral training
 - assuring of good career perspectives for the doctoral candidates/fresh doctorates
- Identification of the RC's strengths and challenges related to the practises and quality of doctoral training, and the actions planned for their development.

The additional material: TUHAT compilation of the RC's other scientific activities/supervision of doctoral dissertations

A written feedback from the aspects of: processes and good practices related to leadership and management

- Strengths
- Areas of development
- Other remarks
- Recommendations

Numeric evaluation: OUTSTANDING (5), EXCELLENT (4), VERY GOOD (3), GOOD (2), SUFFICIENT (1)

3. The societal impact of research and doctoral training

- Description on how the RC interacts with and contributes to the society (collaboration with public, private and/or 3rd sector).
- Identification of the ways to strengthen the societal impact of the RC's research and doctoral training.

The additional material: TUHAT compilation of the RC's other scientific activities.

A written feedback from the aspects of: societal impact, national and international collaboration, innovativeness

- Strengths
- Areas of development
- Other remarks
- Recommendations

Numeric evaluation: OUTSTANDING (5), EXCELLENT (4), VERY GOOD (3), GOOD (2), SUFFICIENT (1)

4. International and national (incl. intersectoral) research collaboration and researcher mobility

- Description of
 - the RC's research collaborations and joint doctoral training activities
 - how the RC has promoted researcher mobility
- Identification of the RC's strengths and challenges related to research collaboration and researcher mobility, and the actions planned for their development.

A written feedback from the aspects of: scientific quality, national and international collaboration

- Strengths
- Areas of development
- Other remarks
- Recommendations

Numeric evaluation: OUTSTANDING (5), EXCELLENT (4), VERY GOOD (3), GOOD (2), SUFFICIENT (1)

5. Operational conditions

- Description of the operational conditions in the RC's research environment (e.g. research infrastructure, balance between research and teaching duties).
- Identification of the RC's strengths and challenges related to operational conditions, and the actions planned for their development.

A written feedback from the aspects of: processes and good practices related to leadership and management

- Strengths
- Areas of development
- Other remarks
- Recommendations

6. Leadership and management in the researcher community

- Description of
 - the execution and processes of leadership in the RC
 - how the management-related responsibilities and roles are distributed in the RC
 - how the leadership- and management-related processes support
 - high quality research
 - collaboration between principal investigators and other researchers in the RC
 - the RC's research focus
 - strengthening of the RC's know-how
- Identification of the RC's strengths and challenges related to leadership and management, and the actions planned for developing the processes

7. External competitive funding of the RC

- The RCs were asked to provide information of such external competitive funding, where:
 - the funding decisions have been made during 1.1.2005-31.12.2010, and
 - the administrator of the funding is/has been the University of Helsinki
- On the e-form the RCs were asked to provide:
 - 1) The relevant funding source(s) from a given list (Academy of Finland/Research Council, TEKES/The Finnish Funding Agency for Technology and Innovation, EU, ERC, foundations, other national funding organisations, other international funding organisations), and
 - 2) The total sum of funding which the organisation in question had decided to allocate to the RCs members during 1.1.2005-31.12.2010.

Competitive funding reported in the text is also to be considered when evaluating this point.

A written feedback from the aspects of: scientific quality, scientific significance, societal impact, innovativeness, future significance

- Strengths
- Areas of development
- Other remarks
- Recommendations

8. The RC's strategic action plan for 2011-2013

- RC's description of their future perspectives in relation to research and doctoral training.

A written feedback from the aspects of: scientific quality, scientific significance, societal impact, processes and good practices related to leadership and management, national and international collaboration, innovativeness, future significance

- Strengths
- Areas of development

- Other remarks
- Recommendations

9. Evaluation of the category of the RC in the context of entity of the evaluation material (1-8)

The RC's fitness to the chosen participation category

A written feedback evaluating the RC's fitness to the chosen participation category

- Strengths
- Areas of development
- Other remarks
- Recommendations

Numeric evaluation: OUTSTANDING (5), EXCELLENT (4), VERY GOOD (3), GOOD (2), SUFFICIENT (1)

10. Short description of how the RC members contributed the compilation of the stage 2 material

Comments on the compilation of evaluation material

11. How the UH's focus areas are presented in the RC's research?

Comments if applicable

12. RC-specific main recommendations based on the previous questions 1-11

13. RC-specific conclusions

1.7 Evaluation criteria

The panellists were expected to give evaluative and analytical feedback to each evaluation question according to their aspects in order to describe and justify the quality of the submitted material. In addition, the evaluation feedback was asked to be pointed out the level of the performance according to the following classifications:

- outstanding (5)
- excellent (4)
- very good (3)
- good (2)
- sufficient (1)

Evaluation according to the criteria was to be made with thorough consideration of the entire evaluation material of the RC in question. Finally, in questions 1-4 and 9, the panellists were expected to classify their written feedback into one of the provided levels (the levels included respective descriptions, 'criteria'). Some panels used decimals in marks. The descriptive level was interpreted according to the integers and not rounding up the decimals by the editors.

Description of criteria levels

Question 1 – FOCUS AND QUALITY OF THE RC'S RESEARCH

Classification: Criteria (level of procedures and results)

Outstanding quality of procedures and results (5)

Outstandingly strong research, also from international perspective. Attracts great international interest with a wide impact, including publications in leading journals and/or monographs published by leading international publishing houses. The research has world leading qualities. The research focus, key research questions scientific significance, societal impact and innovativeness are of outstanding quality.

In cases where the research is of a national character and, in the judgement of the evaluators, should remain so, the concepts of "international attention" or "international impact" etc. in the grading criteria above may be replaced by "international comparability".

Operations and procedures are of outstanding quality, transparent and shared in the community. The improvement of research and other efforts are documented and operations and practices are in alignment with the documentation. The ambition to develop the community together is of outstanding quality.

Excellent quality of procedures and results (4)

Research of excellent quality. Typically published with great impact, also internationally. Without doubt, the research has a leading position in its field in Finland.

Operations and procedures are of excellent quality, transparent and shared in the community. The improvement of research and other efforts are documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of excellent quality.

Very good quality of procedures and results (3)

The research is of such very good quality that it attracts wide national and international attention.

Operations and procedures are of very good quality, transparent and shared in the community. The improvement of research and other efforts are documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of very good quality.

Good quality of procedures and results (2)

Good research attracting mainly national attention but possessing international potential, extraordinarily high relevance may motivate good research.

Operations and procedures are of good quality, shared occasionally in the community. The improvement of research and other efforts are occasionally documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of good quality.

Sufficient quality of procedures and results (1)

In some cases the research is insufficient and reports do not gain wide circulation or do not have national or international attention. Research activities should be revised.

Operations and procedures are of sufficient quality, shared occasionally in the community. The improvement of research and other efforts are occasionally documented and operations and practices are to some extent in alignment with the documentation. The ambition to develop the community together is of sufficient quality.

Question 2 – DOCTORAL TRAINING

Question 3 – SOCIETAL IMPACT

Question 4 – COLLABORATION

Classification: Criteria (level of procedures and results)

Outstanding quality of procedures and results (5)

Procedures are of outstanding quality, transparent and shared in the community. The practices and quality of doctoral training/societal impact/international and national collaboration/leadership and management are documented and operations and practices are in alignment with the documentation. The ambition to develop the community together is of outstanding quality. The procedures and results are regularly evaluated and the feedback has an effect on the planning.

Excellent quality of procedures and results (4)

Procedures are of excellent quality, transparent and shared in the community. The practices and quality of doctoral training/societal impact/international and national collaboration/leadership and management are documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of excellent quality. The procedures and outcomes are evaluated and the feedback has an effect on the planning.

Very good quality of procedures and results (3)

Procedures are of very good quality, transparent and shared in the community. The practices and quality of doctoral training/societal impact/international and national collaboration/leadership and

management are documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of very good quality.

Good quality of procedures and results (2)

Procedures are of good quality, shared occasionally in the community. The practices and quality of doctoral training/societal impact/international and national collaboration/leadership and management are documented and operations and practices are to large extent in alignment with the documentation. The ambition to develop the community together is of good quality.

Sufficient quality of procedures and results (1)

Procedures are of sufficient quality, transparent and shared in the community. The practices and quality of doctoral training/societal impact/international and national collaboration/leadership and management are occasionally documented and operations and practices are to some extent in alignment with the documentation. The ambition to develop the community together is of sufficient quality.

Question 9 – CATEGORY

Participation category – fitness for the category chosen

The choice and justification for the chosen category below should be reflected in the RC's responses to the evaluation questions 1–8.

1. *The research of the participating community represents the international cutting edge in its field.*
2. *The research of the participating community is of high quality, but the community in its present composition has yet to achieve strong international recognition or a clear break-through.*
3. *The research of the participating community is distinct from mainstream research, and the special features of the research tradition in the field must be considered in the evaluation.* The research is of high quality and has great significance and impact in its field. However, the generally used research evaluation methods do not necessarily shed sufficient light on the merits of the research.
4. *The research of the participating community represents an innovative opening.* A new opening can be an innovative combination of research fields, or it can be proven to have a special social, national or international demand or other significance. Even if the researcher community in its present composition has yet to obtain proof of international success, its members can produce convincing evidence of the high level of their previous research.
5. *The research of the participating community has a highly significant societal impact.* The participating researcher community is able to justify the high social significance of its research. The research may relate to national legislation, media visibility or participation in social debate, or other activities promoting social development and human welfare. In addition to having societal impact, the research must be of a high standard.

An example of outstanding fitness for category choice (5) ⁵

The RC's representation and argumentation for the chosen category were convincing. The RC recognized its real capacity and apparent outcomes in a wider context to the research communities. The specific character of the RC was well-recognized and well stated in the responses. The RC fitted optimally for the category.

- Outstanding (5)
- Excellent (4)
- Very good (3)
- Good (2)
- Sufficient (1)

The above-mentioned definition of outstanding was only an example in order to assist the panellists in the positioning of the classification. There was no exact definition for the category fitness.

⁵ The panels discussed the category fitness and made the final conclusions of the interpretation of it.

1.8 Timetable of the evaluation

The main timetable of the evaluation:

- | | |
|--|-----------------------|
| 1. Registration | November 2010 |
| 2. Submission of self-evaluation materials | January–February 2011 |
| 3. External peer review | May–September 2011 |
| 4. Published reports | March–April 2012 |
| - University level public report | |
| - RC specific reports | |

The entire evaluation was implemented during the university's strategy period 2010–2012. The preliminary results were available for the planning of the following strategy period in late autumn 2011. The evaluation reports will be published in March/April 2012. More detailed time schedule is published in the University report.

1.9 Evaluation feedback – consensus of the entire panel

The panellists evaluated all the RC-specific material before the meetings in Helsinki and mailed the draft reports to the evaluation office. The latest interim versions were on-line available to all the panellists on the Wiki-sites. In September 2011, in Helsinki the panels discussed the material, revised the first draft reports and decided the final numeric evaluation. After the meetings in Helsinki, the panels continued working and finalised the reports before the end of November 2011. The final RC-specific reports are the consensus of the entire panel.

The evaluation reports were written by the panels independently. During the editing process, the evaluation office requested some clarifications from the panels when necessary. The tone and style in the reports were not harmonized in the editing process. All the reports follow the original texts written by the panels as far as it was possible.

The original evaluation material of the RCs, provided for the panellists is attached at the end of the report. It is essential to notice that the exported lists of publications and other scientific activities depend how the data was stored in the TUHAT-RIS by the RCs.

2 Evaluation feedback

2.1 Focus and quality of the RC's research

- *Description of*
 - *the RC's research focus*
 - *the quality of the RC's research (incl. key research questions and results)*
 - *the scientific significance of the RC's research in the research field(s)*
- *Identification of the ways to strengthen the focus and improve the quality of the RC's research*

ASPECTS: Scientific quality, scientific significance, societal impact, innovativeness

Strengths

The sizeable RC consists of about 100 researchers and constitutes one of the world-leading centres in the area of algorithmic and computational data analysis. The top-level research of the RC deals specifically with the computational methods for machine learning and data mining, combinatorial pattern matching and string algorithms, together with the multi-disciplinary applications in science (e.g. in biology, medicine) and industry. The RC has an outstanding international track record and represents one of the leading, core areas of HIIT. The RC is widely acclaimed as a leading centre. The RC benefits from outstanding scientific leadership and a very international network.

The RC is widely recognized also through ALGODAN, the Finnish Centre of Excellence in algorithms and data analysis that is already recognized for a second term by the AF.

The bibliometric record shows that the RC has excellent performance in peer-reviewed conference publications in well-recognized, high-level peer-reviewed conferences. For the field in question this is an excellent standard. However, quite remarkably the RC is also performing very well in publishing articles in excellent journals.

The quality of the scientific staff in the RC is very high, as seen from the citations, awards, patents and so on of its members. The RC is highly visible in organizing committees, in program committees of numerous conferences and in other leading roles in the research world.

The RC is an asset for the Department.

Areas of development

The RC sets an outstanding example of how a field within Computer Science can be practiced. The collaboration between the constituent groups and the joint research projects across discipline boundaries will be maintained and enhanced.

Other remarks

The scientific importance of data analysis research and of the computational challenges in this field is rising continuously, as applications in science and industry are becoming increasingly data-intensive.

Recommendations

The RC is excellently positioned, with an outstanding scientific record in which foundational and application-oriented research are both excellently represented. The research agendas of the constituent groups are clearly at the core of the success of the RC. Selecting among the many challenging problems in the field and maintaining the high quality in research and (doctoral) training will be important.

Numeric evaluation: 5 (Outstanding)

2.2 Practises and quality of doctoral training

- *Organising of the doctoral training in the RC. Description of the RC's principles for:*
 - *recruitment and selection of doctoral candidates*
 - *supervision of doctoral candidates*
 - *collaboration with faculties, departments/institutes, and potential graduate schools/doctoral programmes*
 - *good practises and quality assurance in doctoral training*
 - *assuring of good career perspectives for the doctoral candidates/fresh doctorates*
- *Identification of the RC's strengths and challenges related to the practises and quality of doctoral training, and the actions planned for their development.*
- *Additional material: TUHAT compilation of the RC's other scientific activities/supervision of doctoral dissertations*

ASPECTS: Processes and good practices related to leadership and management

Strengths

The RC is responsible for the specialized education in 'Algorithms and Machine Learning' in the Department and coordinates several other graduate programs, such as the joint MSc programme in Bio-informatics (with Aalto University). Its doctoral program is of high international repute. The organization of the PhD program, and the recruitment and supervision of candidates for it follow the established, excellent practices of the Department of Computer Science.

The strong cross-disciplinary collaboration within the RC is also reflected in the doctoral training. The doctoral program is part of the Helsinki Graduate School in Computer Science and Engineering and of the Finnish Graduate School in Computational Sciences.

The career perspectives of the graduates are very good, not in the least because of the many opportunities in Finland's IT companies but also through e.g. postdoc positions to pursue a possible career in science. PhD students are encouraged to take additional courses to prepare them for careers in industry.

The quality of the scientific staff is excellent and of high international standard. There also are an excellent number of postdocs in the RC.

Areas of development

The areas of development are clearly embedded in the research groups. The strong link between research and teaching will be maintained. For an outstanding program as this, it is important to be able to attract very talented PhD students.

Other remarks

Some additional information on specific courses for PhD students would be helpful.

Recommendations

Some further information on the cursory component of the PhD program would be helpful, viz. on how this component is kept up-to-date with the developments in the RC and in its research area in general. What role will ICT Labs play for the graduate training?

Numeric evaluation: 5 (Outstanding)

2.3 The societal impact of research and doctoral training

- *Description on how the RC interacts with and contributes to the society (collaboration with public, private and/or 3rd sector).*
- *Identification of the ways to strengthen the societal impact of the RC's research and doctoral training.*
- *Additional material: TUHAT compilation of the RC's other scientific activities.*

ASPECTS: Societal impact, national and international collaboration, innovativeness

Strengths

The RC has a rich network in science and industry. Its methods, tools, software, patents, courses and graduates have a clear societal impact. The RC also has an excellent and extensive record of academic and industrial cooperations.

Areas of development

The current lines of action are excellent for a sustained societal impact. Increased media visibility is among the actions foreseen for further improvement of the impact.

Other remarks

The impact of data analysis and 'discovery science' is seen everywhere and offers clear pathways to societal impact. There are many opportunities for PR in this area, especially with for this multi-disciplinary RC.

Recommendations

The ever on-going development of data-intensive applications in science and industry will continue to lead to demands for knowledge and expertise from an expert and renowned party like the RC. An active strategy for the PR may be helpful.

Numeric evaluation: 5 (Outstanding)

2.4 International and national (incl. intersectoral) research collaboration and researcher mobility

- *Description of*
 - *the RC's research collaborations and joint doctoral training activities*
 - *how the RC has promoted researcher mobility*
- *Identification of the RC's strengths and challenges related to research collaboration and researcher mobility, and the actions planned for their development.*

ASPECTS: Scientific quality, national and international collaboration

Strengths

The RC has a very strong international and cross-disciplinary tradition. The RC has an excellent record in collaboration in EU-projects and in many other international and intersectoral projects, in Europe and worldwide. Nationally and locally, the RC is also very well networked in collaborative structures. It is a key participant in HIIT. Researcher mobility, also to industry, is at a very good level.

Areas of development

The RC has wide international and national collaborative networks, also across discipline boundaries and with industry. These collaborations must be nurtured and maintained.

Other remarks

It would be helpful to know more about the role and benefits foreseen for the RC in the ICT Labs project of the EIT.

Recommendations

The RC should maintain its excellent level of external collaboration in research and in PhD training. Mobility of heavily occupied scientific staff should be facilitated in any reasonable way possible.

Numeric evaluation: 5 (Outstanding)

2.5 Operational conditions

- *Description of the operational conditions in the RC's research environment (e.g. research infrastructure, balance between research and teaching duties).*
- *Identification of the RC's strengths and challenges related to operational conditions, and the actions planned for their development.*

ASPECTS: Processes and good practices related to leadership and management

Strengths

The RC operates in the excellent research environment as provided in the Department of Computer Science. The benefits are clearly stated and convincing.

In addition, the RC has an excellent level of international and industry collaborations and projects. It gives optimal conditions for the high-level and internationally oriented PhD program.

Areas of development

The RC's activities and responsibilities are extensive.

Other remarks

Some additional remarks concerning the balance between teaching, research supervision, and research acquisition for the RC members at different levels of seniority would be helpful.

Recommendations

The strengths of the RC are clearly the excellent research environment, the excellent scientific staff, the extensive international and local collaborative networks, all the excellent graduate program. The Department should cherish this excellent RC.

2.6 Leadership and management in the researcher community

- *Description of*
 - *the execution and processes of leadership in the RC*
 - *how the management-related responsibilities and roles are distributed in the RC*
 - *how the leadership- and management-related processes support*
 - *high quality research*
 - *collaboration between principal investigators and other researchers in the RC*
 - *the RC's research focus*
 - *strengthening of the RC's know-how*
- *Identification of the RC's strengths and challenges related to leadership and management, and the actions planned for developing the processes*

ASPECTS: Processes and good practices related to leadership and management

Strengths

The scientific leadership of the RC is very strong. The RC has several PI's which are very well established. The RC has a clear leading role in its area (e.g. in HIIT).

The RC also participates in two Graduate Schools (Hecse and FICS), both of which are coordinated by directors from the RC. The RC is also responsible for the master programme in bio-informatics.

The RC has an exemplary range of actions to support and enhance its research focus and internal collaboration.

Areas of development

The interest of foreign students for studies in this field can be expected to grow. The interest of postdocs can be expected to develop likewise.

Other remarks

It is important that the acquisition- and management-related duties of the RC remain sufficiently balanced with the duties in research and education within the RC.

Recommendations

The challenges of maintaining the high quality of the teams and of the, internal and external, collaborations are non-trivial and should be supported.

2.7 External competitive funding of the RC

- *The RCs were asked to provide information of such external competitive funding, where:*
 - *the funding decisions have been made during 1.1.2005–31.12.2010, and*
 - *the administrator of the funding is/has been the University of Helsinki*
- *On the e-form the RCs were asked to provide:*
 - 1) *The relevant funding source(s) from a given list (Academy of Finland/Research Council, TEKES/The Finnish Funding Agency for Technology and Innovation, EU, ERC, foundations, other national funding organisations, other international funding organizations), and*
 - 2) *The total sum of funding which the organisation in question had decided to allocate to the RCs members during 1.1.2005–31.12.2010.*

Competitive funding reported in the text is also to be considered when evaluating this point.

ASPECTS: Scientific quality, scientific significance, societal impact, innovativeness and future significance

Strengths

The RC is funded very well from a variety of sources, reflecting the prominent standing of the RC. Funding comes from both national (e.g. Academy of Finland and Tekes) and international sources (e.g. EU, NIH) and many other organizations. The sources reflect the nature of the research in the RC.

Areas of development

The research portfolio is likely to grow in this vibrant area.

Other remarks

The RC would benefit from an ERC starting or advanced investigator grant.

Recommendations

Research funding e.g. at the national level (e.g. for ALGODAN) or at the EU-level is an ever important issue. The RC may need to develop a long-term e.g. 5-year perspective on the funding of its research effort, taking expected opportunities and uncertainties into account.

(Funding is of critical importance to any RC. For an RC of this excellent standing, the international funding will likely make the difference in the future, of course while maintaining the level of funding from Finnish sources.)

2.8 The RC's strategic action plan for 2011–2013

- *RC's description of their future perspectives in relation to research and doctoral training.*

ASPECTS: Scientific quality, scientific significance, societal Impact, processes and good practices related to leadership and management, national and international collaboration, innovativeness, future significance

Strengths

The strategic action plan is clearly focused on the commitment to do world-class research. The research agenda's of the teams are impressive and leading, and central to the actions of the RC.

Areas of development

The RC aims to keep its operations at the high levels that have become its standard. The RC aims to strengthen its activities in several directions, incl its industry-oriented research and several other initiatives. These include e.g. the plan to maintain a good distribution in the seniority levels in its leadership.

Recommendations

The RC operates at an excellent level. The strategic actions will improve its position even further. The RC is a top group in Computer science by any international standard.

2.9 Evaluation of the category of the RC in the context of entity of the evaluation material (1-8)

The RC's fitness to the chosen participation category.

Category 1. The research of the participating community represents the international cutting edge in its field.

Strengths

The excellent international position and status of the RC is clear. Considering its excellent and consolidated level of operation, it is entirely appropriate that the RC chose for Participation Category 1: 'Research of the participating community represents the international cutting edge in its field'.

Recommendations

The RC should be facilitated by the Department in order to be able to maintain its exemplary level of operation.

Numeric evaluation: 5 (Outstanding)

2.10 Short description of how the RC members contributed the compilation of the stage 2 material

The material was prepared following an excellent procedure, aptly using wiki-technology to adequately accommodate the scale of the RC and allow all members to follow the process and comment on the contents of the document. The members of the RC were all consulted on the material in final form and could submit their feedback before the final submission of the material.

2.11 How the UH's focus areas are presented in the RC's research

Focus area 7: Precise reasoning

Algorithmic data analysis is a field which fits the science-/mathematical traditions of Computer Science. It fits in the UH focus area 'Exact Thinking' (also called: Precise Reasoning), with a definite slant towards applications in a multi-disciplinary setting (e.g. touching on aspects of the UH focus area 'The Basic Structure of Life').

2.12 RC-specific main recommendations

The RC is one of the world-leading centres in the area of algorithmic and computational data analysis. It is operating at an outstanding level. The department should cherish this excellent RC.

The RC is excellently positioned in its research field, with an outstanding scientific record in which foundational and application-oriented research are both excellently represented. The research agendas of the constituent groups are clearly at the core of the success of the RC. Selecting among the many challenging problems in the field and maintaining the high quality in research and (doctoral) training will be important.

The ever on-going development of data-intensive applications in science and industry will continue to lead to demands for knowledge and expertise from an expert and renowned party like the RC. An active strategy for the Public Relations may be helpful.

The RC may need to develop a long-term e.g. 5-year perspective on the funding of its research effort, taking expected opportunities and uncertainties into account. Especially the leading role and status at the European level could be developed further.

2.13 RC-specific conclusions

The RC operates at an outstanding level and the overall focus is very strong. The RC should maintain its excellent level of external and cross-disciplinary collaboration in research and in PhD training. The RC is an asset for the department.

2.14 Preliminary findings in the Panel-specific feedback

Panel-specific feedback

The (meta-)evaluation is based solely on the documentation.

Quality in research and doctoral training

- **Research focus.** Algorithmic data analysis is one of Helsinki's widely recognized focus areas in Computer Science. The RC operates at an outstanding level, with an outstanding scientific record in which both fundamental and applications-oriented research are represented in an excellent manner. (The core of the RC is the CoE ALGODAN which is now even in its second term of AoF funding.)
- **Practices and quality of doctoral training.** The organization of the PhD program follows the established, excellent practices of the Department of Computer Science. The quality of the scientific staff is excellent. The number of postdocs is good but may increase in the future. Some more information on the cursory component of the PhD program would have been helpful.
- **Societal impact.** The RC has an excellent societal impact through its methods, tools, software, patents, courses and graduates. The ever-ongoing development towards greater data-intensiveness in all IT-applications in science and society will accelerate the importance of 'discovery science' in the future, and this RC is excellently positioned for it.
- **International and national collaboration.** The RC has a wide international and national network of collaborators, also across discipline boundaries and with industry. Mobility of heavily occupied scientific staff should be facilitated in any reasonable way possible.
- **Leadership and management.** The scientific leadership of this RC is very strong. It is important that the duties of research, education and management remain sufficiently balanced within the RC, also for the PI's.
(It is not clear whether personnel management is an issue for the department or for an RC.)
- **External funding.** The RC is funded very well from a variety of sources (including AoF, TEKES, industry etc), reflecting the prominent standing of the RC. Funding is of critical importance for all

RC's. The RC may need to develop a long-term perspective on the funding of its research, taking expected opportunities and uncertainties into account. Special interest may be given to more EU-funding, while maintaining the level of funding from Finnish sources.

- **Strategic action plan.** The strategic action plan is clearly focused on the commitment to do world-class research.
 - **Findings.** The RC operates at an outstanding level. It is an example of how RC's should work in an innovative field.
 - **Strenght.** The RC demonstrates an outstanding ability to pursue its high-level research in both fundamental and interdisciplinary contexts. The academic status is excellent. The RC is an asset for the department.
 - **Potential development areas.** Maintaining the high level in research and (doctoral) training will be important.

3 Appendices

- A. Original evaluation material
 - a. Registration material – Stage 1
 - b. Answers to evaluation questions – Stage 2
 - c. List of publications
 - d. List of other scientific activities
- B. Bibliometric analyses
 - a. Analysis provided by CWTS/University of Leiden
 - b. Analysis provided by Helsinki University Library (66 RCs)



International evaluation of research and doctoral training
at the University of Helsinki 2005-2010

RC-SPECIFIC MATERIAL FOR THE PEER REVIEW

NAME OF THE RESEARCHER COMMUNITY:
Algorithms and Data Analysis (ALKO)

LEADER OF THE RESEARCHER COMMUNITY:
Professor Esko Ukkonen, Department of Computer Science, HIIT, and Algodan Centre of
Excellence, University of Helsinki

RC-SPECIFIC MATERIAL FOR THE PEER REVIEW:

- Material submitted by the RC at stages 1 and 2 of the evaluation
 - STAGE 1 material: RC's registration form (incl. list of RC participants in an excel table)
 - STAGE 2 material: RC's answers to evaluation questions
- TUHAT compilations of the RC members' publications 1.1.2005-31.12.2010
- TUHAT compilations of the RC members' other scientific activities 1.1.2005-31.12.2010
- UH Library analysis of publications data 1.1.2005-31.12.2010 – results of UH Library analysis will be available by the end of June 2011

NB! Since Web of Science(WoS)-based bibliometrics does not provide representative results for most RCs representing humanities, social sciences and computer sciences, the publications of these RCs will be analyzed by the UH Library (results available by the end of June, 2011)



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 1 MATERIAL (registration form)

1 RESPONSIBLE PERSON

Name: Ukkonen, Esko

E-mail:

Phone: 09-19151280

Affiliation: Department of Computer Science, HIIT, and Algodan Centre of Excellence, Universi

Street address: Gustaf Hållströmin katu 2b

2 DESCRIPTION OF THE PARTICIPATING RESEARCHER COMMUNITY (RC)

Name of the participating RC (max. 30 characters): Algorithms and Data Analysis

Acronym for the participating RC (max. 10 characters): ALKO

Description of the operational basis in 2005-2010 (eg. research collaboration, joint doctoral training activities) on which the RC was formed (MAX. 2200 characters with spaces): The Algorithms and Data Analysis community (ALKO) is one of the world-leading groups in data analysis research. The community investigates computational methods for machine learning and data mining, pattern matching, and string algorithms. It is one of three natural units of the Department of Computer Science. It consists of researchers working on computer science-oriented Data Analysis, one of the three focal areas of research at the department. The community is also responsible for education in Algorithms and Machine Learning, one of the three specialization areas of the department.

At the same time, the researcher community consists of two out of four programs of Helsinki Institute for Information Technology HIIT. HIIT is a joint institute of the University of Helsinki and Aalto University, and almost all of the community members have a double affiliation with HIIT. The ALKO researcher community at the University of Helsinki forms a majority of HIIT's Algorithmic Data Analysis and Algorithmic Systems programs. Other members of these HIIT programs are an essential part of the researcher community, but being outside the University of Helsinki they are not fully covered by this assessment.

The core of this researcher community is Algodan, a Center of Excellence in Algorithmic Data Analysis. Most researchers of the community also belong to this center.

Joint doctoral training of the community is carried out not only at the department level, but also in two graduate schools coordinated within the community. They include collaboration especially with other parts of HIIT (in Helsinki Graduate School in Computer Science and Engineering) and with other computational sciences (in Finnish Graduate School in Computational Sciences).



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 1 MATERIAL (registration form)

3 SCIENTIFIC FIELDS OF THE RC

Main scientific field of the RC's research: natural sciences

RC's scientific subfield 1: Computer Science, Artificial Intelligence

RC's scientific subfield 2: Computer Science, Information Systems

RC's scientific subfield 3: Computer Science, Interdisciplinary Applications

RC's scientific subfield 4: Computer Science, Theory and Methods

Other, if not in the list: data analysis

4 RC'S PARTICIPATION CATEGORY

Participation category: 1. Research of the participating community represents the international cutting edge in its field

Justification for the selected participation category (MAX. 2200 characters with spaces): The Algorithms and Data Analysis researcher community has an established status as an international leader in its field. Indications of this from external evaluations include the following.

(1) Algodan has been appointed as a Center of Excellence by the Academy of Finland for a second six year term already (the first was under the name From Data to Knowledge).

(2) In the International Evaluation of Computer Science Research in Finland (Academy of Finland, 2007) the department was described as "an obvious leader in Finnish computer science", largely due to research in algorithms and data analysis.

(3) In the Recent Research Assessment Exercise of Aalto University (2009), HIIT's research was found to be on "Outstanding International Level". The report concludes that "The research on the Algorithmic Data Analysis theme is outstanding" and that "The work in the Probabilistic and Adaptive Systems (now: Algorithmic Systems) theme is internationally acclaimed".

5 DESCRIPTION OF THE RC'S RESEARCH AND DOCTORAL TRAINING

Public description of the RC's research and doctoral training (MAX. 2200 characters with spaces): The research of the Algorithms and Data Analysis community lies in the algorithmic and modeling problems of combinatorial pattern matching, data mining, and machine learning. The work is strongly interdisciplinary: we cooperate constantly with application experts in various application areas, formulating novel computational concepts and ways of attacking the scientific and industrial problems of the application areas. Developing new concepts, algorithms and models is an iterative process consisting of interacting extensively with the application experts, formulating computational concepts and models, analyzing their properties, designing algorithms and analyzing their performance, implementing and experimenting with the algorithms and models, and applying the results in practice. The main application areas are in biology, medicine, telecommunications, environmental studies, linguistics, and neuroscience.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 1 MATERIAL (registration form)

Doctoral training by the Algorithms and Data Analysis researcher community emphasizes a solid knowledge of the central concepts and methods of computer science. Doctors are trained for research positions both in the academia and the industry. Expertise in the core areas of computer science is essential for the innovativeness and competitiveness of the IT industry, as well as for scientific advances and industrial applications of information technology. While the framework for doctoral training is set by degree requirements of the university and faculty, the department provides additional support in the form of mentoring and a PhD seminar. A key instrument for organizing internationalization, domestic contacts and peer support, are graduate schools. The researcher community coordinates two such national schools, encompassing practically all of the PhD students of the community.

Significance of the RC's research and doctoral training for the University of Helsinki (MAX. 2200 characters with spaces): The community is a substantial resource of core computer science research and education at the national level. Besides the expected high quality research and doctoral training and the associated results (publications, projects, software, degrees), the community actively contributes to other departments and sciences at the university in the following ways.

The Algorithms and Data Analysis researcher community contributes exceptionally strongly to cross-disciplinary research jointly with other departments of the university. The community has had joint projects and publications with the Departments of Biological and Environmental Sciences, Comparative Religion, Ecology and Systematics, Education, English, Finnish Language and Literature, Finno-Ugric Studies, Geology, History, Mathematics and Statistics, Medical Genetics, Physical Sciences, and Psychology, as well as the Haartman Institute, Helsinki Institute of Physics, Institute of Biomedicine, and Institute of Biotechnology.

Doctors trained by the community are often employed by other departments in the biomedical sciences as method specialists.

Keywords: computer science, algorithms, machine learning, data mining, data analysis, complex systems, string methods, neuroinformatics, bioinformatics

6 QUALITY OF RC'S RESEARCH AND DOCTORAL TRAINING

Justified estimate of the quality of the RC's research and doctoral training at national and international level during 2005-2010 (MAX. 2200 characters with spaces): The quality of the research of the Algorithms and Data Analysis researcher community is at the highest international level. Its publications are highly cited and many have become text-book material. The results are utilized extensively in other sciences and in industry. Quotes from recent evaluations:

"The panel was very impressed with the focus and depth of the Department, with its ability to integrate theory and applications, and with its international leadership in several important fields. " (International Evaluation of Finnish Computer Science, Academy of Finland, 2007).



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 1 MATERIAL (registration form)

"In summary, we were very impressed by the CoE and its activities. We believe that Algodan is among the best CoEs worldwide, that the research carried out in the CoE is of high quality." (Scientific Advisory Board of Algodan, Center of Excellence in Algorithmic Data Analysis, 2008)

"In this program (Algorithmic Systems, then called 'Probabilistic Adaptive Systems'), we assess the Institute's activities as excellent. The groups have a strong publication record, and its researchers have assumed important leadership and organizational positions within their technical fields". – "This world-class program (Algorithmic Data Analysis) is under the leadership of a widely recognized scientific leader. The collaboration with domain experts is impressive, and provides further evidence of the impact and usefulness of the algorithms being developed." (Scientific Advisory Board of HIIT, 2008)

"The work in the PAS theme (now Algorithmic Systems) is internationally acclaimed, both fundamental research as well as scientific/engineering applications. The research on the ADA theme is outstanding, and is closely related to the Academy of Finland Centre of Excellence with the same name." (Aalto University Research Assessment Exercise, 2009)

The doctoral training has a high international quality. This is evidenced, e.g., by the excellent positions of the PhD graduates. The community has invested a lot into developing researcher training, in order to shorten PhD study times: e.g., support for planning of studies and research, mentoring, internationalization, PhD seminars, summer schools, and peer support.

Comments on how the RC's scientific productivity and doctoral training should be evaluated (MAX. 2200 characters with spaces): For the assessment of scientific productivity, the normal scientific indicators apply: international publications, citations, competitive external funding, number of degrees, etc. However, computer science research also has some strong characteristics that must be noted in any assessment of research in the field (see also the report "Research Evaluation in Computer Science" by Informatics Europe: www.informatics-europe.org/docs/research_evaluation.pdf):

* All peer-reviewed publications should be considered, also peer-reviewed conference articles. Conference publications are usually rigorously peer-reviewed, they have low acceptance rates, and they are considered original publications just like journal papers. Citation rates for papers in good conferences are typically higher than for papers in good journals.

* Citation analysis is useful, but should be exercised with equal care since conference publications and citations to and from them are poorly covered by many indexes. In computer science, Google Scholar seems to have the best coverage and should be used to complement other sources of citation information, such as Web of Science.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 1 MATERIAL (registration form)

In doctoral training, it is important to assess the PhD study and supervision processes, as well as placement of graduated in PhD students, in addition to simple statistics. This includes systematic and supportive methods to select, supervise, mentor, internationalize, and graduate PhD students.

The publishing strategy is to publish methodological results in best venues of computer science (both journals and conferences), applied results within other scientific fields in their best journals, and additionally to publish research prototype software where appropriate. Further dissemination to computer science community takes place through active participation in and organization of international projects, networks, and workshops; to other sciences and to industry through research co-operation, technology transfer (including software, IPRs), and people.

LIST OF RC MEMBERS

NAME OF THE RESEARCHER COMMUNITY:			Algorithms and Data Analysis (ALKO)		
RC-LEADER			E. Ukkonen		
CATEGORY			1		
	Last name	First name	PI-status (TUHAT, 29.11.2010)	Title of research and teaching personnel	Affiliation
1	Ukkonen	Esko	x	Professor	Faculty of Science, Department of Computer Science & HIIT
2	Hoyer	Patrik	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
3	Hyvärinen	Aapo	x	Professor	Faculty of Science, Department of Computer Science & HIIT
4	Kärkkäinen	Juha	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
5	Kaski	Petteri	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
6	Kaski	Samuel	x	Professor, Research Director	Faculty of Science, Department of Computer Science & HIIT
7	Kivinen	Jyrki	x	Professor	Faculty of Science, Department of Computer Science & HIIT
8	Koivisto	Mikko	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
9	Lemström	Kjell	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
10	Mäkinen	Veli	x	Professor	Faculty of Science, Department of Computer Science & HIIT
11	Mannila	Heikki	x	Professor, Vice Rector	Faculty of Science, HIIT
12	Myllymäki	Petri	x	Professor	Faculty of Science, Department of Computer Science & HIIT
13	Polishchuk	Valentin	x	Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
14	Rousu	Juho	x	Senior Researcher	Faculty of Science, Department of Computer Science
15	Sillanpää	Mikko	x	Senior Researcher	Faculty of Science, Department of Mathematics and Statistics and Faculty of Biosciences, Department of Agricultural Sciences
16	Toivonen	Hannu	x	Professor	Faculty of Science, Department of Computer Science & HIIT
17	Yangarber	Roman	x	Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
18	Ahonen-Myka	Helena		Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
19	Buntine	Wray		Senior Researcher	Faculty of Science, Department of Computer Science & HIIT
20	Rissanen	Jorma		Senior Researcher	Faculty of Science, HIIT
21	Tirri	Henry		Professor	Faculty of Science, Department of Computer Science & HIIT
22	Bingham	Ella		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
23	Garcia Moreno-Esteva	Enrique		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
24	Garriga	Gemma		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
25	Gionis	Aristides		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
26	Gutmann	Michael		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
27	Hurri	Jarmo		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
28	Hyvönen	Saara		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
29	Järvisalo	Matti		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
30	Kääriäinen	Matti		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
31	Kivioja	Teemu		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
32	Mahler	Sebastian		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
33	Na	Joong Chae		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
34	Nicolas	François		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
35	Oikarinen	Emilia		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
36	Pizzi	Cinzia		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
37	Salmela	Leena		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
38	Schönauer	Stefan		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT

39	Sevon	Petteri		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
40	Tamm	Hellis		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
41	Tsaparas	Panayiotis		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
42	Yu	Huizhen		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
43	Zhang	Kun		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
44	Haiminen	Niina		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
45	Hämäläinen	Wilhelmiina		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
46	Kontkanen	Petri		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
47	Lehtonen	Miro		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
48	Leino	Antti		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
49	Mielikäinen	Taneli		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
50	Miettinen	Pauli		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
51	Palin	Kimmo		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
52	Rantanen	Ari		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
53	Rastas	Pasi		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
54	Roos	Teemu		Postdoctoral Researcher	Faculty of Science, Department of Computer Science & HIIT
55	Aunimo	Lili		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
56	Autio	Ilkka		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
57	Doucet	Antoine		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
58	Klami	Arto		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
59	Kollin	Jussi		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
60	Köster	Urs		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
61	Laasonen	Kari		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
62	Lindgren	Jussi		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
63	Makkonen	Juha		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
64	Pitkänen	Esa		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
65	Raento	Mika		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
66	Silander	Tomi		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
67	Terzi	Evimaria		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
68	Ahonen	Teppo		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
69	Astikainen	Katja		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
70	Entner	Doris		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
71	Eronen	Lauri		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
72	Galbrun	Esther		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
73	Heinonen	Markus		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
74	Heinonen	Oskari		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
75	Hintsanen	Petteri		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
76	Hyttinen	Antti		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
77	Junttila	Esa		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
78	Kallio	Aleksi		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
79	Korhonen	Janne		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT

80	Kuuskoski	Reeta		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
81	Langohr	Laura		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
82	Luosto	Panu		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
83	Narayanan	Krishnan		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
84	Paalasmaa	Joonas		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
85	Parviainen	Pekka		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
86	Perkiö	Jukka		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
87	Pihlaja	Miika		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
88	Puuronen	Jouni		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
89	Siren	Jouni		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
90	Toivonen	Jarkko		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
91	Tripathi	Abhishek		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
92	Välimäki	Niko		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
93	Vuorinen	Matti		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
94	Wessman	Jaana		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT
95	Zhou	Fang		Doctoral candidate	Faculty of Science, Department of Computer Science & HIIT



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

BACKGROUND INFORMATION

Name of the RC's responsible person: Ukkonen, Esko

E-mail of the RC's responsible person:

Name and acronym of the participating RC: Algorithms and Data Analysis, ALKO

The RC's research represents the following key focus area of UH: 7. Eksakti ajattelu – Exact thinking

Comments for selecting/not selecting the key focus area: It should be clear that the ALKO community primarily belongs to the Exact thinking focus area; even the building we are located in on the Kumpula Campus is called Exactum (sic!). However, via our multidisciplinary collaborations we also contribute to many other focus areas, most notably to Basic structure of life.

1 FOCUS AND QUALITY OF RC'S RESEARCH (MAX. 8800 CHARACTERS WITH SPACES)

- **Description of the RC's research focus, the quality of the RC's research (incl. key research questions and results) and the scientific significance of the RC's research for the research field(s).**

The Algorithms and Data Analysis researcher community (ALKO) is one of the world-leading groups in data analysis research. The community investigates computational methods for machine learning and data mining, pattern matching, and string algorithms as well as their multidisciplinary applications in other sciences as well as in the industry.

ALKO is a community of about 100 members including 17 PI's of which eight are professors. Most members are affiliated with the Department of Computer Science or with the Helsinki Institute for Information Technology HIIT, a joint institute of UH and Aalto University. The core of ALKO is Algodan, a national Centre of Excellence in Algorithmic Data Analysis research, granted by the Academy of Finland. Our organizational position is explained in more detail on the Registration Form (Sect. 2).

Focus

The importance of data analysis in science and in industry is increasing continuously, as our ability to measure and store data grows dramatically. While data analysis is as old as science itself, the new methods of collecting raw data pose unprecedented challenges and opportunities to data analysis and to the algorithms of data analysis.

The ALKO community develops new concepts, algorithms, principles, and frameworks for data analysis. The overall objective includes large-scale data-intensive computational modeling and inference: how to model the currently ubiquitous data banks and streams to infer what is relevant in the vast data masses. The work combines strong basic research in computer science with interdisciplinary work in a wide variety of scientific disciplines and industrial problems. The research of the community lies in the areas of combinatorial pattern matching, data mining, and machine learning.

Developing new concepts and algorithms is an iterative process consisting of interacting extensively with the application experts, formulating computational concepts, analyzing the properties of the concepts, designing algorithms and analyzing their performance, implementing and experimenting with the algorithms, and applying the resulting software. The main application areas of ALKO are in biology (bioinformatics), medicine, information retrieval, environmental studies, linguistics, and neuroscience (neuroinformatics).



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Here is a selection of our recent results:

- Problem formulations, algorithms, and freely available software for learning linear latent variable models, by A. Hyvärinen, P. Hoyer and others, have significantly advanced the field, especially independent component analysis (ICA), non-negative matrix factorization, and causal discovery (e.g., Journal of Machine Learning Research 2006: 82 citations in Google Scholar).. Recently, we have developed nonlinear extensions of ICA to model learning of basic perceptual representations in the visual areas of the brain. This work culminated in the monograph A. Hyvärinen, J. Hurri and P.O.Hoyer: Natural Image Statistics (Springer 2009), the first book on the topic.
- ContextPhone, by M. Raento, H. Toivonen et al., is a unique prototyping and research software platform for context-aware applications running on Smartphones. ContextPhone has been used in numerous projects world-wide (e.g. Reality Mining at MIT), and was being commercialized by Jaiku Ltd. Later, Google acquired Jaiku. This was successful cross-disciplinary research, with wide visibility in press. Main publication IEEE Pervasive Computing 2005 (304 citations).
- Novel algorithms for finding orders from data, by A. Gionis and H. Mannila, significantly improve solutions to interesting paleontological problems while related complexity results are theoretically interesting. ACM Transactions on Knowledge Discovery from Data 2007 (145 citations).
- A novel model for so-called gene enhancer elements in mammalian genomes, by E. Ukkonen in collaboration with J. Taipale, was used to predict several new enhancers, some of which have then been successfully verified in vivo. Publications in Nature Genetics 2009 (84 citations), Cell 2006 (211 citations).
- Several algorithmic advances in full-text -indexing, by J. Kärkkäinen, E. Ukkonen and V. Mäkinen, have been celebrated results in the combinatorial pattern matching community. Publications in Journal of the ACM 2006 (329 citations), ACM Transactions on Algorithms 2007 (138 citations)
- The Tutte polynomial of a graph is a most fundamental invariant in graph theory, with connections to many areas. M. Koivisto et al. have developed a substantially improved evaluation algorithm of Tutte polynomials [FOCS 2008]. Our fast subset convolution framework and related results [STOC 2007; 71 citations] occupy a substantial portion of a recent book "Exact Exponential Algorithms" by Fomin and Kratch (Springer 2010).
- The probabilistic positioning algorithms by P. Myllymäki et al. have been successfully commercialized by the spin-off Ekahau Ltd, and the technology has won numerous international awards.

Significance and quality

The high quality of the research of ALKO is indicated by high number of citations, high level of external competitive funding and by high marks in previous evaluations.

The eight ALKO professors have had strong scientific impact as evidenced by their high citation rates. Their total number of citations is over 62000 (Publish or Perish with Google Scholar), with average h-index is 32 (Hyvärinen: 16850 citations, h=36; Kaski: 5415/34; Kivinen 1959/19; Mäkinen 1251/19; Mannila 17762/57; Myllymäki 1668/19; Toivonen 11200/34; Ukkonen 5917/37). Aapo Hyvärinen is recognized by Thomson ISI as a Highly Cited researcher, as one of only 20 scholars in Finland, all disciplines combined. Jorma Rissanen, a Shannon award recipient (not listed among ALKO professors but a member in ALKO) has about 17000 citations and h=43.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Corresponding figures of group leader professors in computer science in comparable units (e.g. SICS, CWI, LICT) are clearly smaller. ETH Zurich scores highly on the Shanghai ranking and has citation figures comparable to ALKO: total citations app. 44814 and average $h=32$.

The external competitive funding of ALKO in 2005-2010 is 14.8 million euros for research and 7.7 million euros for PhD programmes coordinated by ALKO members.

Our impact is visible for instance in some recent textbooks, or in the popularity of research software such as B-Course, a web-based data analysis tool for multivariate Bayesian and causal modeling, with hundreds of users world-wide monthly.

Through its wide collaborations, ALKO has a significant impact on other sciences. Our new computational methods have produced new results for various applications (published, e.g., in Nature Genetics (impact factor 34.2), Nature Biotechnology (29.5), Cell (29.4)). The novel methods are changing the way many researchers in other sciences operate: study designs are changing when new possibilities for data analysis are opened. In such fields as genetics, paleontology, and linguistics our work is influencing research plans.

In the 2005 research assessment of UH, the Department of Computer Science and HIIT both got the highest mark. In the 2010 assessment of Aalto University, HIIT was among the six best units. The Algodan CoE within ALKO got its CoE status from the Academy of Finland in a highly competitive process covering all fields of research. The recent report of the Algodan SAB (J. Gehrke (Cornell), V. Tresp (TU Munich)) states: "We believe that Algodan is among the leading CoEs worldwide, that the research carried out in the CoE is of high quality, and that it already had and very likely will continue to have high impact in the future." See Registration Form, Sect. 4

- **Ways to strengthen the focus and improve the quality of the RC's research.**

The success of ALKO style of research essentially depends on the contribution of talented individuals and not so much on coordinated effort of a large group of people. To improve the quality it is therefore crucial to hire good people and to find the best collaborators who can bring in novel computational questions. We already publish on the most prestigious forums (STOC, FOCS, SODA, Journal of the ACM, ACM Transactions on Knowledge Discovery from Data, Journal of Machine Learning Research, PNAS, Cell, Nature Genetics,...), but we need to concentrate still more on high-quality publishing. Our methodological focus stems from a shared computational toolbox and from concentration on data. Our focus in data analysis applications can naturally be quite wide. Explicitly selecting some grand challenge problems to work on would help improving the focus.

2 PRACTISES AND QUALITY OF DOCTORAL TRAINING (MAX. 8800 CHARACTERS WITH SPACES)

- How is doctoral training organised in the RC? Description of the RC's principles for recruitment and selection of doctoral candidates, supervision of doctoral candidates, collaboration with faculties, departments/institutes, and potential graduate schools/doctoral programmes, good practises and quality assurance in doctoral training, and assuring good career perspectives for the doctoral candidates/fresh doctorates.

Doctoral training is formally organized by the Department of Computer Science in close collaboration with doctoral programmes. The department (together with the faculty) coordinates admission, degree requirements, teaching, supervision, as well as follow-up of students. Most ALKO PhD students are enrolled in of the two doctoral programmes (Hecse and FICS) coordinated by ALKO PIs. These programmes provide additional training in their focus areas (summer schools, short courses),



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

supervision (support teams, peer support), joint recruitment efforts, and financial support (travel, some salaries). The processes of the department and these doctoral programmes have been carefully aligned.

Practical researcher training in ALKO takes place in research groups. Doctoral students work on national or international research projects, learning by working with more experienced researchers. Projects and team work improve important transferable skills. International activities (conferences, research visits) are strongly encouraged and financially supported, to give our students wider perspectives. We next review some of the more detailed practices.

Recruitment to PhD studies is mainly through supervisors and their networks. Additionally, our doctoral programmes organize and advertise calls for application, to attract students from a wider pool. The increasing international MSc education at the department will be used to minimize risks in recruiting international doctoral students.

Selection and admission to PhD studies is coordinated at departmental level, by a joint PhD Studies Committee with a representative from each area, including ALKO. The committee evaluates the research plan, study plan and the funding plan as well as supervision, and routinely also asks for improvements to them before acceptance. Admission to doctoral programmes is programme-specific, but always includes a review phase by three impartial experts and an open discussion among a wide board of a programme.

Supervision of doctoral students is understood in ALKO as the processes of guidance and support at large. We naturally require each new student to have a supervisor who is fully committed to the student, his or her research topic, as well as finding funding for four years. Often there is a second, younger supervisor working with the PhD student on day-to-day research. Additional supervision is provided by the research group. – An important form of guidance and support is offered by two mentors. They are typically professors or docents, and they always come from other research groups or universities. Their role is to provide general advice and encouragement, as well as quality control external to the research group. Mentoring is kept lightweight to avoid overwhelming the mentors. – Peer support is also part of supervision as we understand it. Mechanisms to support it include a PhD seminar and various activities by our doctoral programmes.

Collaboration in doctoral training is intensive. Formal forms include doctoral programmes and networks. Hecse is a joint programme on computer science with Aalto University, FICS is a wider national network on computational sciences. Activities span across universities almost on a daily basis. For instance, mentors often come from other universities and students can easily take courses from another university. Other collaboration includes organization of joint special courses and planning and implementation of the student application processes. On the international level, we are a core member of the EU Network of Excellence Pascal which supports international collaboration, such as student visits, joint research, and joint summer schools and workshops.

Cross-disciplinary collaboration in doctoral training of ALKO is also substantial. Some take place in FICS, but much is informal and purely based on mutual interests of the supervisors and students. PhDs have been co-supervised in the intersections of computer science and biology, linguistics, and medicine. – Cross-sector collaboration with industry is also common. Many of our students already have professional experience and many carry out their research in a project with industrial partners. – Co-operation with the department and faculty is easy and natural, as there is a clear division of work and strong culture of co-operation and coordination.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Good practices have already been covered above so we just list them here: supervision at large, learning by working in projects and teams, clear division of work and coordination between organizations, collaboration across units and disciplines, recruitment and admission procedures, active and concrete support for internationalization.

The quality of the doctoral training environment is the highest possible in Finland: our research and teaching staff have been found to be at the top level in several evaluations, and also the infrastructures for research and learning are excellent. Quality of doctoral training is assured from several viewpoints. The student selection processes guarantee fair selection of best candidates. Their progress is followed on a regular basis at the PhD student seminar, by annually updated research plans, and by an annual poster presentation. The SAB's of Algodan and HIIT, as well as the current and other evaluations provide quality control for the research at large, while constant publication at peer reviewed venues during PhD studies gives feedback on individual work. Mentoring is not used for control, but for supportive actions in case of difficulties in research. Financial difficulties are handled by the Algodan CoE or the department. Other potential issues are handled by the PhD Studies Committee. National benchmarking of doctoral training takes place in the doctoral programmes, international by active mutual participations in PhD committees. As a rule, we have three external examiners for each PhD, of which at least one usually comes from abroad.

Career perspectives of our doctoral students and graduates PhDs are very good. From the close collaboration across sectors and disciplines we are well aware of the needs for computer science graduates and have been able shift emphasis where needed. In addition to computer science substance, project and team work skills are actively built, and we encourage our students to take studies in topics such as project management and industrial economics. A proof of good career perspectives is the placement of our recent graduates in excellent positions, whether in R&D units of large companies (e.g., Nokia, Ericsson, NVIDIA, Yahoo, Google), in exciting start-ups, or as post-docs in computer science or in other sciences. The doctoral programme Hecse can already list 18 computer science professors among its former students.

- RC's strengths and challenges related to the practises and quality of doctoral training, and the actions planned for their development.

The main strength of our doctoral training lies in the high quality of the research and teaching staff, good coordination and collaboration in doctoral training, as well as the good practices. The PIs are active in the research community, creating opportunities for collaboration and academic career development, also with other sciences and industry.

A challenge is how to continuously succeed in attracting talented students, as the number of MSc graduates is decreasing. We will further emphasize international recruitment.

Another challenge are the long study times and relative high drop-out rates in computer science. We have improved the student selection and supervision processes and will monitor the success.

A new opportunity for international recruitment and co-operation is provided by the recently established European Institute for Innovation and Technology (EIT) ICT Labs. Hecse and FICS are both affiliated with it.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

3 SOCIETAL IMPACT OF RESEARCH AND DOCTORAL TRAINING (MAX. 4400 CHARACTERS WITH SPACES)

- Description of how the RC interacts with and contributes to the society (collaboration with public, private and/or 3rd sector).

The ALKO community has an exceptionally rich network of research collaborators in the public and private sector. We will describe collaborations and partners in more detail in Section 4 below. Here, we focus on the interaction and impact outside computer science research.

Impact of scientific co-operation

Through collaborations, ALKO has a large indirect impact on society as methods and tools developed in ALKO are used to solve important problems. Examples include the following.

- New algorithms and software (patented) for gene mapping and haplotyping have been used successfully to find novel genes, potentially leading to novel diagnostics and medications.
- Computational reconstruction of the dissemination history of the Legend of St. Henry, i.e., identification of the paths along which different versions have been created by copying and altering earlier versions, has provided a basis for the first scientific edition of this historical text. The study has also provided crucial information about medieval literature and the history of Baltic regions.

During the reporting period, we have collaborated with in total over 300 partners in academia and industry.

Industrial impact

Many research projects of ALKO are funded by companies (in 2005-10, total 270 000 euros), often jointly with Tekes, or collaboration takes place in European projects with strong involvement of companies (2 projects in 2010). Industrial partners include enterprises in medicine, broadcasting and media, IT solutions, language services, etc. We have had joint projects with 68 companies (10 from abroad).

Our researchers also provide consulting and courses to companies. Several spin-offs have been founded with basis in our research (e.g., Bayes Information Technology, Ekahau, Jaiku, Whitevector).

- Ekahau has successfully commercialized probabilistic positioning algorithms developed in ALKO. The company has over 10 000 customers worldwide, and the patented technology has won numerous international awards.

- Jaiku commercialized research that was carried out in the ContextPhone project, on context-awareness and microblogging in mobile phones. Jaiku was eventually acquired by Google.

Impact through software and patents

Software produced in ALKO is used both in industry and in academia. For instance, both Bayes-IT and ICA have been downloaded over 10 000 times. HaploRec has 130 registered users.

Our researchers have filed about twenty patents, e.g., in gene mapping, location positioning, data visualization, and pattern discovery, many of which have been utilized commercially by the start-ups or in the software.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Impact on Finnish academia

ALKO also has tight collaborations with many Finnish universities, especially Aalto, and also with polytechnics (Stadia and Haaga-Helia), and it this way spreads its expertise and results into wider use in the Finnish society.

Societal impact of doctoral training

The Finnish and European society has a pressing need for highly trained computer scientists. Many our PhD graduates are employed in exciting positions in industry or in other sciences.

Of our 28 PhD graduates, 6 are employed in industry, 3 are entrepreneurs, and 16 are researchers in units outside ALKO. Out of these 28, 13 are currently abroad.

- Ways to strengthen the societal impact of the RC's research and doctoral training.

The industrial impact could be further increased by more research collaboration. Given the recent increase in PhD output and the good placement of them in industry, it is becoming easier to find natural collaborations.

Our doctoral programmes have started adding components for entrepreneurship in their curricula. This direction will be strengthened to increase the societal as well as economical impact of research and PhD training.

An area for improvement is our media visibility. This would help make our impact and skills better visible, and attract more collaboration.

We are in the process of building up our alumni operations. An alumni association for computer science was recently established, and we will use them to reach to the society (and to let the society reach to us). Our annual open house day will be made better known to alumni and others interested, in order to improve the contacts.

4 INTERNATIONAL AND NATIONAL (INCL. INTERSECTORAL) RESEARCH COLLABORATION AND RESEARCHER MOBILITY (MAX. 4400 CHARACTERS WITH SPACES)

- Description of the RC's research collaborations and joint doctoral training activities and how the RC has promoted researcher mobility.

ALKO has a strong international and cross-disciplinary tradition. Here we give a brief summary of the collaborations during the evaluation period.

We have conducted research in international collaboration in numerous EU-funded projects (SMART, Bison, IQ, VisMaster, CLASS, Alvis; total MEUR 1.2 granted during 2005-10), in two Networks of Excellence (Pascal 2 with 65 partners, and Biosapiens with 24 partners), and in projects funded by NIH and Frontex (total MEUR 0.6). The total number of research partners (excluding the networks) is 74.

A concrete evidence of successful international collaboration is jointly authored articles. During the evaluation period, we have authored peer-reviewed papers with co-authors from 93 foreign research units (including top institutes such as University of California, Berkeley, Carnegie Mellon University and University of Toronto). Of these units, 72 are in the area of computer science, while 21 are in other fields, including bioinformatics, statistics, biology, chemistry, neuroscience, aviation, medicine, philosophy and economics.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Intersectoral research is even stronger nationally, where we have collaborated with departments or research institutes in the areas of computer science, bioinformatics, biology, neuroscience, history, medicine, forensics, software innovation and business intelligence technology, education, language technology, geography, geology, and physics with co-authors from 15 domestic units in about 30 projects.

Examples of very successful international or interdisciplinary collaboration include:

- New methods for estimating linear and nonlinear causal models from data, jointly with Shohei Shimizu at the University of Osaka, Bernhard Schölkopf at MPI Tübingen, Peter Spirtes at Carnegie Mellon University. We showed that if the data is non-Gaussian one can obtain much more information on the underlying data generating process.
- Storage and retrieval of individual genomes, jointly with Gonzalo Navarro, University of Chile. For data analysis of next-generation DNA sequencing we developed space-efficient self-indexes that provide fast extraction of any substring and fast searching of exact/approximate occurrences of a pattern.
- Finding orders from data, jointly with Mikael Fortelius at Department of Geosciences and Geography. In the analysis of paleontological presence/absence data, an important task is to find an ordering for the sites so that each species occurs in consecutive observations. We have developed novel algorithms for this seriation task.

Researcher mobility is strongly encouraged, using funding from the CoE, department, and doctoral programmes. Approximately 30% of our PhD students and post-docs are international, and 65% of our PhD graduates have had a substantial stay abroad.

There is also significant researcher mobility across sectors and disciplines. Many of our PhD students have worked in the industry and some take their PhD while in industry. Three of our professors have had positions at Nokia Research Center, including the current Head of Nokia Research. Several PhD's have been taken in the intersection of computer science and some other field (biology, medicine, linguistics), and several of our PhD graduates are now employed by other departments.

- RC's strengths and challenges related to research collaboration and researcher mobility, and the actions planned for their development.

Our strengths clearly are in the wide collaboration networks, also across scientific disciplines and to industry, the good level of international staff, and the English-speaking environment and culture at the department.

Longer research visits to abroad still need financial and other incentives to keep the international activity on satisfactory level in the future.

5 OPERATIONAL CONDITIONS (MAX. 4400 CHARACTERS WITH SPACES)

- Description of the operational conditions in the RC's research environment (e.g. research infrastructure, balance between research and teaching duties).

The Department of Computer Science and HIIT provide an excellent research environment for the ALKO community. The administration, working space and library are very good, and the computing environment is excellent with a dedicated IT team and modern infrastructure.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Much of the research of ALKO is collaborative. The research institute HIIT has been and will continue to be a major platform for collaboration between these universities. ALKO PIs Kaski, Myllymäki and Mannila also have an Aalto affiliation.

In the Helsinki area, ALKO has extensive network of collaboration partners in the University of Helsinki, the Aalto University and the VTT Technical Research Centre of Finland. Here, the collaborators include eight national Centres of Excellence in research. Collaboration conditions with such relatively well-funded units are easy and smooth, without high bureaucratic barriers. The funding of ALKO mostly comes from competitive sources (Academy of Finland, EU, Tekes). The longer-term funding of the Algodan CoE balances shorter term funding from other sources. Doctoral programmes are also a part of the overall funding structures.

The work load of PIs in applying for external funding retrieval is relative high. The teaching load, on the other hand, is negotiable and allows research active staff to teach less. The ALKO community is jointly responsible for the teaching of the Algorithms and Machine Learning subprogramme of the department and of the MBI programme.

ALKO participates and is responsible for organizing two national PhD programmes funded by the Academy of Finland: Helsinki Graduate School in Computer Science and Engineering (Hecse), and Finnish Doctoral Programme in Computational Sciences (FICS) is coordinated in ALKO.

- RC's strengths and challenges related to operational conditions, and the actions planned for their development.

Our strengths are an excellent research infrastructure, extensive international and local collaboration, and excellent researchers and students. The teaching load of research active PIs can be reasonably low. We strive to provide a strong link between our research and teaching.

The main challenges include the administrative load of struggling with short term external funding, common in Finland, to maintain our current good funding level. With increasing number of graduated PhDs from our doctoral programmes, creating an attractive academic career beyond the PhD becomes a more and more serious challenge. Hopefully the new tenure track systems will provide some assistance in this respect, by providing better career paths for young researchers.

6 LEADERSHIP AND MANAGEMENT IN THE RESEARCHER COMMUNITY (MAX. 4400 CHARACTERS WITH SPACES)

- Description of the execution and processes of leadership in the RC, how the management-related responsibilities and roles are distributed in the RC and how the leadership- and management-related processes support high quality research, collaboration between principal investigators and other researchers in the RC, the RC's research focus and strengthening of the RC's know-how.

ALKO is a community of about 15 research groups of varying size and in varying stage of their evolution, sharing the mission of combinatorial-probabilistic data analysis and having joint funding for research and PhD education. Joint leaderships and management of research takes place in the Algodan CoE and in HIIT. Joint doctoral training is primarily lead in doctoral programmes Hecse and FICS, while teaching of PhD and MSc level courses is managed in the Algorithms and Machine Learning subprogramme of the department and the MBI master's programme. The leaderships are described below. For the daily management the community utilizes administrative service units of the Department of Computer Science, HIIT and Kumpula Campus.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

Professor Esko Ukkonen is the director of the Algodan CoE, Professor Heikki Mannila being the vice-director. Most ALKO groups are members of the Algodan centre. The centre has a board that coordinates joint activities and develops the research agenda and fund raising. The director is responsible for the allocation of the joint funding to the teams, new hirings, and developing the infrastructure of the CoE. When taking these decisions, he consults with the board.

Most groups of ALKO also participate in the research programmes of HIIT. Here, Professor Samuel Kaski is the director of the Algorithmic Data Analysis program, and Professor Petri Myllymäki is directing the Algorithmic Systems program. The programmes have boards and some joint funds.

The ALKO community is responsible for the teaching on the Algorithms and Machine Learning subprogramme at the Department of Computer Science. This is organized by Professor Jyrki Kivinen who is the director of this line.

The PhD programme Helsinki Graduate School in Computer Science and Engineering (Hecse) is coordinated in ALKO, Professor Hannu Toivonen being the director of the school.

The Finnish Doctoral Programme in Computational Sciences (FICS) is coordinated in ALKO, Samuel Kaski being the director. Our former PhD programme Graduate School on Computational Biology, Bioinformatics and Biometry (ComBi) was merged with FICS.

The joint Masters Programme in Bioinformatics (MBI) of the University of Helsinki and Aalto University is coordinated in ALKO, Professor Veli Mäkinen being the director.

To support high quality research, internal collaboration, research focus and know-how of ALKO, our actions include:

- Competitive allocation of the internal funds by using international open calls to find best people to open positions.
- Regular internal full-day review and brainstorming seminars to give feed-back to the groups and to encourage new collaborations as well as to develop up-to-date and focused research agenda.
- Regular weekly research seminar to report own research as well as to invite visitors.
- Internal mentoring of young PIs.
- Joint hiring and exchange of researchers with our application partners.
- Organizing multidisciplinary seminars comprising the subjects of our application partners.
- The researchers of ALKO regularly participate in teaching by giving lectures, arranging seminars, and supervising of project work and theses.
- Participation in international consortia in research (e.g. the EU programmes) and in PhD education (e.g. the BREW collaboration).
- RC's strengths and challenges related to leadership and management, and the actions planned for developing the processes.

The ALKO community is a result of a natural evolution that started in the 1980's. Among our 17 PIs, four have PhD from Helsinki University of Technology (Aalto University), two from the US (Stony Brook, NYU), and the rest have been educated in the UH. All the senior members have a long history of collaboration and good confidential personal relations. This is our clear strength on which we can build in leadership



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

and management. A further strength is the very good infrastructure provided by the Department of Computer Science and HIIT. Almost all members of the community work physically in the same building.

The main challenges include how to maintain sufficient coherence and focus in the large community comprising of several strong and scientifically independent PIs, and how the entire community can most efficiently utilize the special competencies of individual groups. Here transparency of planning could help, for example by organizing open seminars in which the PIs communicate regularly to the community each new project that is about to start.

7 EXTERNAL COMPETITIVE FUNDING OF THE RC

- Listing of the RCs external competitive funding, where:
 - the funding decisions have been made during 1.1.2005-31.12.2010, and
 - the administrator of the funding is/has been the University of Helsinki
- Academy of Finland (AF) - total amount of funding (in euros) AF has decided to allocate to the RC members during 1.1.2005-31.12.2010: 10450000
- Finnish Funding Agency for Technology and Innovation (TEKES) - total amount of funding (in euros) TEKES has decided to allocate to the RC members during 1.1.2005-31.12.2010: 1470000
- European Union (EU) - total amount of funding (in euros) EU has decided to allocate to the RC members during 1.1.2005-31.12.2010: 1330000
- European Research Council (ERC) - total amount of funding (in euros) ERC has decided to allocate to the RC members during 1.1.2005-31.12.2010:
- International and national foundations – names of international and national foundations which have decided to allocate funding to the RC members during 1.1.2005-31.12.2010, and the amount of their funding (in euros).
 - names of the foundations: Finnish Cultural Foundation, Alexander von Humboldt Foundation, Technology Industries of Finland Centennial Foundation
 - total amount of funding (in euros) from the above-mentioned foundations: 350000
- Other international funding - names of other international funding organizations which have decided to allocate funding to the RC members during 1.1.2005-31.12.2010, and the amount of their funding (in euros).
 - names of the funding organizations: National Institutes of Health (NIH), Frontex
 - total amount of funding (in euros) from the above-mentioned funding organizations: 570000
- Other national funding (incl. EVO funding and Ministry of Education and Culture funded doctoral programme positions) - names of other national funding organizations which have decided to allocate funding to the RC members during 1.1.2005-31.12.2010, and the amount of their funding (in euros).
 - names of the funding organizations: Ministry of Education (doctoral programmes 7680000), University of Helsinki, Orion, VTT Technical Research Centre of Finland, Nokia, Finnish Association on Intellectual and Developmental Disabilities
 - total amount of funding (in euros) from the above-mentioned funding organizations: 8320000



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

8 RC'S STRATEGIC ACTION PLAN FOR 2011–2013 (MAX. 4400 CHARACTERS WITH SPACES)

- Description of the RC's future perspectives in respect to research and doctoral training.

In general, we are happy with the current mode and focus of operations of ALKO. The Algodan CoE's SAB stated in their 2010 report: "The general technical direction of this confluence of people and research areas – the combination of combinatorial and statistical techniques in data analysis – is highly innovative." "The research groups are integrated in a way that comes naturally out of overlap in the projects that they work on. We do not think that further overlap needs to be necessary or forced."

Specific actions in 2011-13:

- Strengthen industry-oriented work by taking a central role in the forthcoming major research programme 'From data to intelligence' of the ICT cluster of the Finnish Strategic Centres for Science, Technology and Innovation (ICT-SHOK). Petri Myllymäki is the academic coordinator of this initiative.
- Maintain active international hiring, and support researcher careers by increasing postdoctoral and higher levels.
- Keep the computing infrastructure powerful.
- Implement new incentives to boost international activity, in particular longer research visit to and from abroad.
- Our general policy is to make research software publicly available. Systematize this by creating coherent practices and a portal for software distribution, to improve our visibility and impact.
- The current term of the Algodan CoE expires in 2013. This CoE, with HIIT, has been the main source of long-term funding. We need to decide how to continue the CoE and how to participate in other CoEs.
- Strengthen the leadership and administrative role of younger PIs.
- In doctoral training, keep running the two doctoral programmes. One of them (Hecse) is currently under review for prolongation by another 4 year term.
- The PI's of ALKO have formulated mission statements and research goals of their groups for 2011-13. The groups share key elements of their mission but fruitful variability is fostered. Here are some condensed examples of plans:

Hyvärinen: The most practical goal is development of machine learning methods tailored for analysis of brain imaging data such that the specific properties of the data are taken into account. For example, EEG and MEG data are characterized by oscillations, and thus methods based on time-frequency decompositions are necessary. Another example relevant to brain imaging is that we usually record data from many subjects. New methods are needed to properly combine information from all the subjects. On more theoretical level, the goal is to formalize new information-processing principles used by the brain. Another theoretical goal is development of computationally efficient estimation theory.

Kaski: In computational biology and medicine we address the grand challenge of how to keep biology maximally cumulative when first principles are seldom accessible as in physics, and the number of published experiments is beyond comprehension of a single researcher. As recently shown, a new experiment can be put in the context of earlier ones, by viewing it as an information retrieval problem, of retrieving earlier relevant experiments. Making biology cumulative is thus transformed into the



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC STAGE 2 MATERIAL

computational problem of combining prior knowledge and data-driven learning in a model which the analyst can apply for retrieval on a massive scale.

Myllymäki: We focus on flexible modeling approaches, using Bayesian and the Minimum Description Length approaches, where the model complexity is automatically regularized to match the complexity of the learning task. We also work on variable-order Markov models, with applications in areas involving heterogeneous data. One such area is intelligent information access, where one has to gather information at the user end about the context and information need, use this data for accessing relevant information through an intelligent retrieval engine, and finally to present the gathered information in a useful, non-intrusive form.

Mäkinen: The group's focus is shifting towards bioinformatics applications, especially to the new challenges set by high-throughput sequencing technologies. We are working on new solutions to the classical de novo fragment assembly problem using our own approach. A challenge is to design index structures for Markov representations of collections of individual genomes allowing efficient similarity searches.

9 SHORT DESCRIPTION OF HOW THE RC MEMBERS HAVE CONTRIBUTED TO THE COMPILATION OF THE STAGE 2 MATERIALS (MAX. 1100 CHARACTERS WITH SPACES).

In the first phase, we established a group of 12 volunteers from researchers of all levels and from all three RC's at the department (ALKO, NODES, SOFTSYS). This group met three times during December and January to discuss the evaluation and to fill in initial content using bullet points. Some of these activities were organized in subgroups specific to ALKO or NODES, whereas one subgroup worked on doctoral training, a shared activity between the RC's.

The writing of all stage 2 material has taken place in Wiki, allowing all members of the working group as well as all PI's of the communities to follow the process and to contribute to it.

The working group presented its results to the whole department of computer science in its strategy seminar.

In a second phase, the PI's of the RC's took over the responsibility to write the full versions of the texts. Collaboration continued in Wiki where the draft texts were written and edited.

A third and final phase of the preparation was to circulate the final drafts to all members of the RC's for feedback.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

1 Analysis of publications

- Associated person is one of Esko Ukkonen , Patrik Hoyer , Aapo Hyvärinen , Juha Kärkkäinen , Petteri Kaski , Samuel Kaski , Jyrki Kivinen , Mikko Koivisto , Kjell Lemström , Veli Mäkinen , Heikki Mannila , Mikko Sillanpää , Petri Myllymäki , Hannu Toivonen , Roman Yangarber , Helena Ahonen-Myka , Wray Lindsay Buntine , Jorma Johannes Rissanen , Henry Tirri , Ella Bingham , Enrique Garcia Moreno-Esteva , Teemu Kalevi Hurri , Saara Annele Hyvönen , Gemma Garriga, Aristides Gionis , Michael Urs Gutmann , Jarmo Kalevi Hurri , Saara Annele Hyvönen , Sebastian Jean Mahler , Stefan Schönauer , Emilia Oikarinen , Cinzia Pizzi , Leena Salmela , Huizhen Yu , Niina Susanna Haiminen , Suvi Wilhelmina Hämäläinen , Petri Kontkanen , Miro Petteri Lehtonen , Antti Leino , Taneli Johannes Mielikäinen , Pauli Aleksii Miettinen , Kimmo Juhani Palin , Ari Juhani Rantanen , Pasi Rastas , Teemu Teppo Roos , Lili Aunimo, Ilkka Kalevi Autio, Antoine Doucet , Arto Olavi Klami , Jussi Jalmarin Kollin , Kari Tapani Laasonen, Jussi Tapio Lindgren, Juha Antero Makkonen , Esa Pitkänen , Mika Petteri Raento, Tomi Silander , Evimaria Terzi, Teppo Ahonen , Katja Astikainen , Doris Entner , Lauri Eronen , Esther Galbrun , Markus Heinonen , Janne Henrik Korhonen , Reeta Kuuskoski, Laura A Langohr , Antti Hyttinen , Esa Juhani Juntila , Krishnan Narayanan , Joonas Paalasmaa, Pekka Parviainen , Jouni Sirén , Panu Luosto , Miika Pihlaja , Jouni Stephan Puuronen , Niko Välimäki , Jukka Petteri Perkiö , Jarkko Toivonen , Abhishek Tripathi , Fang Zhou , Matti Vuorinen , Jaana Wessman ,

Publication type	Publication year						Total Count 2005 - 2010
	2005	2006	2007	2008	2009	2010	
A1 Refereed journal article	34	33	37	40	38	41	223
A2 Review in scientific journal			1		1	1	3
A3 Contribution to book/other compilations (refereed)	18	11	2	12	9	8	60
A4 Article in conference publication (refereed)	56	46	40	58	44	63	307
B1 Unrefereed journal article		1				1	2
B2 Contribution to book/other compilations (non-refereed)				1			1
B3 Unrefereed article in conference proceedings	3			2	2	2	9
C1 Published scientific monograph	3	5	4	4	3	3	22
C2 Edited book, compilation, conference proceeding or special issue of journal	3	5	1	6	3	4	22
D1 Article in professional journal						2	2
D2 Article in professional hand or guide book or in a professional data system, or text book material						5	5
D4 Published development or research report	2	1		1		1	5
E1 Popular article, newspaper article	1	1	1	1			4
H1 Patents	3		1	1			5



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Publication type	2005	2006	2007	2008	2009	2010	Total Count 2005 - 2010
I1 Audiovisual materials						1	1
I2 ICT programs or applications					2	3	5



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

2 Listing of publications

A1 Refereed journal article

2005

- Borodin, A, Roberts, GO, Rosenthal, JS, Tsaparas, P **2005**, 'Link analysis ranking: algorithms, theory, and experiments', **ACM Transactions on Internet Technology**, vol 5, no. 1, pp. 231-297.
- Buntine, W **2005**, 'Open source search: a data mining platform', **Forum**, vol 39, no. 1, pp. 4-10.
- Esposito, F, Scarabino, T, Hyvärinen, A, Himberg, J, Formisano, E, Comani, S, Tedeschi, G, Goebel, R, Seifritz, E, Di Salle, F **2005**, 'Independent component analysis of fMRI group studies by self-organizing clustering', **NeuroImage**, vol 25, pp. 193-205.
- Fontaine, M, Burkhardt, S, Kärkkäinen, J **2005**, 'BDD-based analysis of gapped q-gram filters', **International Journal of Foundations of Computer Science**, vol 16, no. 6, pp. 1121-1134.
- Fredriksson, K, Navarro, G, Ukkonen, E **2005**, 'Sequential and indexed two-dimensional combinatorial template matching allowing rotations', **Theoretical Computer Science**, vol 347, no. 1-2, pp. 239-275.
- Gasbarra, D, Sillanpää, MJ, Arjas, E **2005**, 'Backward simulation of ancestors of sampled individuals', **Theoretical Population Biology**, vol 67, no. 2, pp. 75-83.
- Hyvärinen, A **2005**, 'Estimation of non-normalized statistical models by score matching', **Journal of machine learning research.**, vol 6, pp. 695-709.
- Hyvärinen, A, Gutmann, MU, Hoyer, PO **2005**, 'Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2', **BMC Neuroscience**, vol 6, pp. -.
- Hyvärinen, A **2005**, 'A unifying model for blind separation of independent sources', **Signal Processing**, vol 85, pp. 1419-1427.
- Hyvönen, S, Junninen, H, Laakso, L, Dal Maso, M, Grönholm, T, Bonn, B, Keronen, P, Aalto, P, Hiltunen, V, Pohja, T, Launiainen, S, Hari, P, Mannila, H, Kulmala, M **2005**, 'A look at aerosol formation using data mining techniques', **Atmospheric Chemistry and Physics**, vol 5, pp. 3345-3356.
- Kaski, S, Nikkilä, J, Sinkkonen, J, Lahti, L, Knuuttila, JEA, Roos, C **2005**, 'Associative clustering for exploring dependencies between functional genomics data sets', **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, vol 2, no. 3, pp. 203-216.
- Kaski, S, Sinkkonen, J, Klami, A **2005**, 'Discriminative clustering', **Neurocomputing**, vol 69, pp. 18-41.
- Kivioja, T, Arvas, M, Saloheimo, M, Penttilä, M, Ukkonen, E **2005**, 'Optimization of cDNA-AFLP experiments using genomic sequence data', **Bioinformatics**, vol 21, no. 11, pp. 2573-2579.
- Kurten, T, Biczysko, M, Rajamäki, T, Laasonen, K, Halonen, L **2005**, 'Computational study of the adsorption energetics and vibrational wavenumbers of NH₃ adsorbed on the Ni(111) surface', **Journal of Physical Chemistry B**, vol 109, no. 18, pp. 8954-8960.
- Laur, S, Lipmaa, H, Mielikäinen, T **2005**, 'Private itemset support counting', **Information and communications security**, pp. 97-111.
- Lemström, K, Mäkinen, V **2005**, 'On minimizing pattern splitting in multi-track string matching', **Journal of Discrete Algorithms**, vol 3, no. 2-4, pp. 248-266.
- Lemström, K, Navarro, G, Pinzon, Y **2005**, 'Practical algorithms for transposition-invariant string-matching', **Journal of Discrete Algorithms**, vol 3, no. 2-4, pp. 267-292.
- Martinez, V, Thorgaard, G, Robison, B, Sillanpää, MJ **2005**, 'An application of Bayesian QTL mapping to early development in double haploid lines of rainbow trout including environmental effects', **Genetical Research**, vol 86, no. 3, pp. 209-221.
- Miettinen, M, Nokelainen, P, Kurhila, J, Silander, T, Tirri, H **2005**, 'EDUFORM: a tool for creating adaptive questionnaires', **International Journal on E-learning**, vol 4, no. 3, pp. 365-373.
- Mäkinen, V, Navarro, G **2005**, 'Succinct suffix arrays based on run-length encoding', **Nordic Journal of Computing**, vol 12, no. 1, pp. 44-66.
- Mäkinen, V, Navarro, G, Ukkonen, E **2005**, 'Transposition invariant string matching', **Journal of Algorithms**, vol 56, pp. 124-153.
- Nikkilä, J, Roos, C, Savia, E, Kaski, S **2005**, 'Exploratory modeling of yeast stress response and its regulation with gCCA and associative clustering', **International Journal of Neural Systems**, vol 15, no. 4, pp. 237-246.
- Nokelainen, P, Miettinen, M, Kurhila, J, Floreen, P, Tirri, H **2005**, 'A shared document-based annotation tool to support learner-centred collaborative learning', **British Journal of Educational Technology**, vol 36, no. 5, pp. 757-770.
- Oja, M, Sperber, GO, Blomberg, J, Kaski, S **2005**, 'Self-organizing map-based discovery and visualization of human endogenous retroviral sequence groups', **International Journal of Neural Systems**, vol 15, no. 3, pp. 163-179.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Pajala, A, Jakulin, A, Buntine, W **2005**, 'Eduskuntaryhmien äänestyskäyttäytyminen ja -koheesio vuoden 2003 valtiopäivillä', **Politiikka**, vol 47, no. 3, pp. 205-217.

Pekkinen, M, Varvio, S, Kulju, K, Kärkkäinen, H, Smolander, S, Viherä-Aarnio, A, Koski, V, Sillanpää, M **2005**, 'Linkage map of birch, *Betula pendula* Roth, based on microsatellites and amplified fragment length polymorphisms', **Genome**, vol 48, no. 4, pp. 619-625.

Peltonen, J, Kaski, S **2005**, 'Discriminative components of data', **IEEE Transactions on Neural Networks**, vol 16, no. 1, pp. 68-83.

Raento, M, Oulasvirta, A, Petit, R, Toivonen, H **2005**, 'ContextPhone: a prototyping platform for context-aware mobile applications', **IEEE Pervasive Computing**, vol 4, no. 2, pp. 51-59.

Roos, T, Wettig, H, Grunwald, P, Myllymäki, P, Tirri, H **2005**, 'On discriminative Bayesian network classifiers and logistic regression', **Machine Learning**, vol 59, no. 3, pp. 267-296.

Rousu, JH, Rantanen, A, Ketola, RA, Kokkonen, JT **2005**, 'Isotopomer distribution computation from tandem mass spectrometric data with overlapping fragment spectra', **Spectroscopy**, vol 19, pp. 53-67.

Rousu, J, Shawe-Taylor, J **2005**, 'Efficient computation of gapped substring kernels on large alphabets', **Journal of machine learning research**, vol 6, pp. 1323-1344.

Salmenkivi, M, Mannila, H **2005**, 'Using Markov chain Monte Carlo and dynamic programming for event sequence data', **Knowledge and Information Systems**, vol 7, no. 3, pp. 267-288.

Sillanpää, MJ, Bhattacharjee, M **2005**, 'Bayesian association-based fine mapping in small chromosomal segments', **Genetics**, vol 169, no. 1, pp. 427-439.

Waldmann, P, Garcia-Gil, M, Sillanpää, M **2005**, 'Comparing Bayesian estimates of genetic differentiation of molecular markers and quantitative traits: an application to *Pinus sylvestris*', **Heredity**, vol 94, no. 6, pp. 623-629.

2006

Bas, P, Hurri, J **2006**, 'Vulnerability of DM watermarking of non-iid host signals to attacks utilising the statistics of independent components', **IEE proceedings information security** : **IEE**, vol 153, no. 3, pp. 127-139.

Doucet, A, Ahonen-Myka, H **2006**, 'Probability and expected document frequency of discontinued word sequences', **TAL : traitement automatique des langues**, vol 46, no. 2, pp. 13-37.

Eronen, L, Geerts, F, Toivonen, H **2006**, 'HaploRec: efficient and accurate large-scale reconstruction of haplotypes', **BMC Bioinformatics**, vol 7, no. 542, 38 s.

Evans, J, Hemment, D, Humphies, T, Raento, M **2006**, 'LOCA: location oriented critical arts', **Leonardo electronic almanac**.

Gasbarra, D, Sillanpää, MJ **2006**, 'Constructing the parental linkage phase and the genetic map over distances <1 cM using pooled haploid DNA', **Genetics**, vol 172, pp. 1325-1335.

Greig, M, Haanpää, H, Kaski, P **2006**, 'On the coexistence of conference matrices and near resolvable $2-(2k+1, k, k-1)$ designs', **Journal of Combinatorial Theory, Series A**, vol 113, no. 4, pp. 703-711.

Haanpää, H, Jarvisalo, M, Kaski, P **2006**, 'Hard satisfiable clause sets for benchmarking equivalence reasoning techniques', **Journal of Satisfiability, Boolean Modeling and Computation**, vol 2, pp. 27-46.

Hallikas, O, Palin, K, Sinjushina, N, Rautiainen, R, Partanen, JM, Ukkonen, EJ, Taipale, J **2006**, 'Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity', **Cell**, vol 124, no. 1, pp. 47-59.

Hintsanen, P, Sevon, P, Onkamo, P, Eronen, L, Toivonen, H **2006**, 'An empirical comparison of case-control and trio based study designs in high throughput association mapping', **Journal of Medical Genetics**, vol 43 (2006), 8 s.

Hoti, F, Sillanpää, M **2006**, 'Bayesian mapping of genotype x expression interactions in quantitative and qualitative traits', **Heredity**, vol 97, pp. 4-18.

Hyvärinen, A **2006**, 'Consistency of pseudolikelihood estimation of fully visible Boltzmann machines', **Neural Computation**, no. 18, pp. 2283-2292.

Jaeger, M, Nielsen, JD, Silander, T **2006**, 'Learning probabilistic decision graphs', **International Journal of Approximate Reasoning**, vol 42, pp. 84-100.

Kaski, P, Östergård, PRJ, Pottonen, O **2006**, 'The Steiner quadruple systems of order 16', **Journal of Combinatorial Theory, Series A**, vol 113, no. 8, pp. 1764-1770.

Kivinen, J, Warmuth, MK, Hassibi, B **2006**, 'The p -Norm Generalization of the LMS Algorithm for Adaptive Filtering', **IEEE Transactions on Signal Processing**, vol 54, no. 5, pp. 1782-1793.

Koivisto, M **2006**, 'Optimal 2-constraint satisfaction via sum-product algorithms', **Information Processing Letters**, vol 98, no. 1, pp. 24-28.

Kurki, I, Hyvärinen, A, Laurinen, P **2006**, 'Collinear context (and learning) change the profile of the perceptual filter', **Vision Research**, vol 46, pp. 2009-2014.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Kärkkäinen, J, Sanders, P, Burkhardt, S **2006**, 'Linear work suffix array construction', **Journal of the ACM**, vol 53, no. 6, pp. 918-936.
- Lehtonen, M **2006**, 'Preparing heterogeneous XML for full-text search', **ACM Transactions on Information Systems**, vol 24, no. 4, pp. 455-474.
- Leino, A, Hyvönen, S, Salmenkivi, M **2006**, 'Mitä murteita suomessa onkaan: murreseanaston levikin kvantitatiivista analyysiä', **Virtittäjä**, vol 110, no. 1, pp. 26-45.
- Mielikäinen, T **2006**, 'Frequency-based views to pattern collections', **Discrete Applied Mathematics**, vol 154, pp. 1113-1139.
- Mielikäinen, T, Ukkonen, E **2006**, 'The complexity of maximum matroid-greoid intersection and weighted greoid maximization', **Discrete Applied Mathematics**, vol 154, pp. 684-691.
- Onkamo, P, Toivonen, H **2006**, 'A survey of data mining methods for linkage disequilibrium mapping', **Human Genomics**, vol 2, no. 5, pp. 336-340.
- Palin, K, Taipale, J, Ukkonen, E **2006**, 'Locating potential enhancer elements by comparative genomics using the EEL software', **Nature Protocols**, vol 1, no. 1, pp. 368-374.
- Raento, M **2006**, 'The data subject's right of access and to be informed in Finland: an experimental study', **International Journal of Law and Information Technology**, vol 14, no. 3, pp. 390-409.
- Rantanen, A, Mielikäinen, T, Rousu, J, Maaheimo, H, Ukkonen, E **2006**, 'Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes', **Bioinformatics**, vol 22, no. 10, pp. 1198-1206.
- Rousu, J, Saunders, C, Szedmak, S, Shawe-Taylor, J **2006**, 'Kernel-based learning of hierarchical multilabel classification models', **Journal of machine learning research.**, vol 7, pp. 1601-1626.
- Ryynänen, O, Puhakka, M, Myllymäki, P, Palomäki, P, Anttonen, V, Jukola, R, Takala, J **2006**, 'Sairaalaan lähettämisen arviointi Bayesin verkkomallilla', **Suomen lääkärilehti**, vol 61, no. 51, pp. 5353-5358.
- Sevon, P, Toivonen, H, Ollikainen, V **2006**, 'TreeDT: tree pattern mining for gene mapping', **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, vol 3, no. 2, pp. 174-185.
- Shimizu, S, Hoyer, PO, Hyvärinen, A, Kerminen, A **2006**, 'A linear non-gaussian acyclic model for causal discovery', **Journal of machine learning research.**, vol 7, pp. 2003-2030.
- Shimizu, S, Hyvärinen, A, Hoyer, PO, Kano, Y **2006**, 'Finding a causal ordering via independent component analysis', **Computational Statistics & Data Analysis**, vol 50, pp. 3278-3293.
- Sillanpää, MJ, Bhattacharjee, M **2006**, 'Association mapping of complex trait loci with context-dependent effects and unknown context variable', **Genetics**, vol 174, no. 3, pp. 1597-1611.
- Tamm, H, Nykänen, M, Ukkonen, E **2006**, 'On size reduction techniques for multitape automata', **Theoretical Computer Science**, vol 363, pp. 234-246.
- Yee, WG, Beigbeder, M, Buntine, W **2006**, 'SIGIR06 workshop report: open source information retrieval systems (OSIR06)', **Forum**, vol 40, no. 2, pp. 61-65.
- 2007**
- Arkin, EM, Mitchell, JSB, Polishchuk, V **2007**, 'Two new classes of Hamiltonian graphs (extended abstract)', **Electronic Notes in Discrete Mathematics.**, vol 29, pp. 565-569.
- Arvas, M, Kivioja, T, Mitchell, A, Saloheimo, M, Ussery, D, Penttilä, M, Oliver, S **2007**, 'Comparison of protein coding gene contents of the fungal phyla Pezizomycotina and Saccharomycotina', **BMC Genomics**, vol 8, pp. 325.
- Asuncion Vicente, M, Hoyer, PO, Hyvärinen, A **2007**, 'Equivalence of some common linear feature extraction techniques for appearance-based object recognition tasks', **IEEE Transactions on Pattern Analysis and Machine Intelligence**, vol 29, no. 5, pp. 896-900.
- Bingham, E, Kaban, A, Fortelius, M **2007**, 'The aspect Bernoulli model: multiple causes of presences and absences', **Pattern analysis & applications**, vol 12, no. 1, pp. 55-78.
- Ferragina, P, Manzini, G, Mäkinen, V, Navarro, G **2007**, 'Compressed representations of sequences and full-text indexes', **ACM transactions on algorithms.**, vol 3, no. 2, pp. Art. 20.
- Fredriksson, K, Mäkinen, V, Navarro, G **2007**, 'Rotation and lighting invariant template matching', **Information and Computation**, vol 205, no. 7, pp. 1096-1113.
- Gasbarra, D, Pirinen, M, Sillanpää, MJ, Arjas, E **2007**, 'Estimating genealogies from linked marker data: a Bayesian approach', **BMC Bioinformatics**, vol 8, no. 411, 56 s.
- Gasbarra, D, Pirinen, M, Sillanpää, MJ, Salmela, E, Arjas, E **2007**, 'Estimating genealogies from unlinked marker data: a Bayesian approach', **Theoretical Population Biology**, vol 72, no. 3, pp. 305-322.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Gionis, A, Mannila, H, Mielikäinen, TJ, Tsaparas, P 2007, 'Assessing data mining results via swap randomization', **ACM Transactions on Knowledge Discovery from Data**, vol 1, no. 3.
- Gionis, A, Mannila, H, Tsaparas, P 2007, 'Clustering aggregation', **ACM Transactions on Knowledge Discovery from Data**, vol 1, no. 1.
- Haiminen, N, Mannila, H, Terzi, E 2007, 'Comparing segmentations by applying randomization techniques', **BMC Bioinformatics**, vol 171, no. 8, 8 s.
- Haiminen, N, Mannila, H 2007, 'Discovering isochores by least-squares optimal segmentation', **Gene**, vol 394, pp. 53-60.
- Heikinheimo, H, Fortelius, M, Eronen, JT, Mannila, H 2007, 'Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters', **Journal of Biogeography**, vol 34, no. 6, pp. 1053-1064.
- Hyvärinen, A, Köster, U 2007, 'Complex cell pooling and the statistics of natural images', **Network: Computation in Neural Systems**, vol 18, no. 2, pp. 81 - 100.
- Hyvärinen, A 2007, 'Connections between score matching, contrastive divergence, and pseudolikelihood for continuous-valued variables', **IEEE Transactions on Neural Networks**, vol 18, no. 5, pp. 1529-1531.
- Hyvärinen, A 2007, 'Some extensions of score matching', **Computational Statistics & Data Analysis**, vol 51, no. 5, pp. 2499-2512.
- Hyvönen, S, Leino, A, Salmenkivi, M 2007, 'Multivariate analysis of Finnish dialect data: an overview of lexical variation', **Literary and Linguistic Computing**, vol 22, no. 3, pp. 271-290.
- Kaski, P, Östergård, PRJ 2007, 'There exists no symmetric configuration with 33 points and line size 6', **Australasian Journal of Combinatorics**, vol 38, pp. 273-277.
- Kontkanen, P, Wetteg, H, Myllymäki, P 2007, 'NML computation algorithms for tree-structured multinomial Bayesian networks', **EURASIP Journal on Bioinformatics and Systems Biology**, vol 2007, pp. Article ID 90947.
- Kontkanen, P, Myllymäki, P 2007, 'A linear-time algorithm for computing the multinomial stochastic complexity', **Information Processing Letters**, vol 103, no. 6, pp. 227-233.
- Krozel, J, Mitchell, JSB, Polishchuk, V, Prete, J 2007, 'Maximum flow rates for capacity estimation in level flight with convective weather constraints', **Air Traffic Control Quarterly**, vol 15, no. 3.
- Kärkkäinen, J 2007, 'Fast BWT in small space by blockwise suffix sorting', **Theoretical Computer Science**, vol 387, no. 3, pp. 249-257.
- Landwehr, N, Mielikäinen, T, Eronen, L, Toivonen, H, Mannila, H 2007, 'Constrained hidden Markov models for population-based haplotyping', **BMC Bioinformatics**, vol 8, no. 2, 9 s.
- Leino, A 2007, 'Peruskartalta kieliooppiin: lektiot', **Virittäjä**, vol 2007/2.
- Lemström, K, Pienimäki, A 2007, 'On comparing edit distance and geometric frameworks in content-based retrieval of symbolically encoded polyphonic music', **Musicae Scientiae**, vol 4A, pp. 135-152.
- Mäkinen, V, Navarro, G 2007, 'Rank and select revisited and extended', **Theoretical Computer Science**, vol 387, no. 3, pp. 332-342.
- Mäkinen, V 2007, 'Peak alignment using restricted edit distances', **Biomolecular engineering**, vol 24, no. 3, pp. 337-342.
- Navarro, G, Mäkinen, V 2007, 'Compressed full-text indexes', **ACM Computing Surveys**, vol 39, no. 1, pp. Article No. 2.
- Nokelainen, P, Silander, T, Ruohotie, P, Tirri, H 2007, 'Investigating the number of non-linear and multi-modal relationships between observed variables measuring growth-oriented atmosphere', **Quality and Quantity**, vol 41, no. 6, pp. 869-890.
- Oja, M, Peltonen, J, Blomberg, J, Kaski, S 2007, 'Methods for estimating human endogenous retrovirus activities from EST databases', **BMC Bioinformatics**, vol 8, no. Suppl 2, pp. S11.
- Oulasvirta, A, Petit, R, Raento, MP, Tiitta, S 2007, 'Interpreting and Acting on Mobile Awareness Cues', **Human - Computer Interaction (Mahwah)**, vol 22, no. 1-2, pp. 97-135.
- Rautio, JJ, Bailey, M, Kivioja, T, Söderlund, H, Penttilä, M, Saloheimo, M 2007, 'Physiological evaluation of the filamentous fungus *Trichoderma reesei* in production processes by marker gene expression analysis', **BMC Biotechnology**, vol 7, no. 28, 16 s.
- Rissanen, J, Grunwald, P, Heikkonen, J, Myllymäki, P, Roos, T, Rousu, J 2007, 'Information theoretic methods for bioinformatics', **EURASIP Journal on Bioinformatics and Systems Biology**, vol 2007, no. Article ID 79128.
- Sillanpää, MJ, Hoti, F 2007, 'Mapping quantitative trait loci from a single-tail sample of the phenotype distribution including survival data', **Genetics**, vol 177, no. 4, pp. 2361-2377.
- Soikkeli, J, Lukk, M, Nummela, P, Virolainen, S, Jahkola, T, Katainen, R, Harju, L, Ukkonen, E, Saksela, O, Hölttä, E 2007, 'Systematic search for the best gene expression markers for melanoma micrometastasis detection', **Journal of Pathology**, vol 213, no. 2, pp. 180-189.
- Välimäki, N, Gerlach, W, Dixit, K, Mäkinen, V 2007, 'Compressed suffix tree: a basis for genome-scale sequence analysis', **Bioinformatics**, vol 23, no. 5, pp. 629-630.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Wettig, J, Kontkanen, P, Myllymäki, P **2007**, 'Calculating the normalized maximum likelihood distribution for Bayesian forests', **IADIS International Journal on Computer Science and Information Systems**, vol 2, no. 2, pp. 1-12.

2008

Astikainen, K, Holm, L, Pitkänen, E, Szedmak, S, Rousu, J **2008**, 'Towards structured output prediction of enzyme function', **BMC proceedings**, vol 2(Suppl 4), no. S2, 10 s.

Bhattacharjee, M, Botting, C, Sillanpää, M **2008**, 'Bayesian biomarker identification based on marker-expression-proteomics data', **Genomics**, vol 92, no. 6, pp. 384-392.

Böcker, S, Mäkinen, V **2008**, 'Combinatorial approaches for mass spectra recalibration', **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, vol 5, no. 1, pp. 91-100.

Chee, YM, Kaski, P **2008**, 'An enumeration of graphical designs', **Journal of combinatorial designs**, vol 16, pp. 70-85.

De Raedt, L, Kersting, K, Kimmig, A, Revoredo, K, Toivonen, H **2008**, 'Compressing probabilistic Prolog programs', **Machine Learning**, vol 70, no. 2-3, pp. 151-168.

Dementiev, R, Kärkkäinen, J, Mehnert, J, Sanders, P **2008**, 'Better external memory suffix array construction', **Journal of Experimental Algorithmics**, vol 12, pp. 3.4:1-3.4:24.

Giegerich, R, Brazma, A, Jonassen, I, Ukkonen, E, Vingron, M **2008**, 'The BREW workshop series: a stimulating experience in PhD education', **Briefings in Bioinformatics**, vol 9, no. 3, pp. 250-253.

Haiminen, N, Gionis, A, Laasonen, K **2008**, 'Algorithms for unimodal segmentation with applications to unimodality detection', **Knowledge and Information Systems**, vol 14, no. 1, pp. 39-57.

Haiminen, N, Mannila, H, Terzi, E **2008**, 'Determining significance of pairwise co-occurrences of events in bursty sequences', **BMC Bioinformatics**, vol 9, no. 336, 10 s.

Heikinheimo, H, Fortelius, M, Eronen, J, Mannila, H **2008**, 'Clusterings should not be compared by visual inspection: response to Gagne & Proulx', **Journal of Biogeography**, vol 36, no. 3, pp. 563-565.

Heinonen, M, Rantanen, A, Mielikäinen, T, Kokkonen, J, Kiuru, J, Ketola, RA, Rousu, J **2008**, 'FiD: a software for ab initio structural identification of product ions from tandem mass spectrometric data', **Rapid Communications in Mass Spectrometry**, vol 22, no. 19, pp. 3043-3052.

Henriksson, L, Nurminen, L, Hyvärinen, A, Vanni, S **2008**, 'Spatial frequency tuning in human retinotopic visual areas', **Journal of vision**, vol 8, no. 10, pp. 1-13.

Hintsanen, P, Toivonen, H **2008**, 'Finding reliable subgraphs from large probabilistic graphs', **Data Mining and Knowledge Discovery**, vol 17, no. 1, pp. 3-23.

Hoyer, PO, Shimizu, S, Kerminen, A, Palviainen, M **2008**, 'Estimation of causal effects using linear non-Gaussian causal models with hidden variables', **International Journal of Approximate Reasoning**, vol 49, no. 2, pp. 362-378.

Hyvärinen, A **2008**, 'Optimal approximation of signal priors', **Neural Computation**, vol 20, no. 12, pp. 3087-3110.

Kaban, A, Bingham, E **2008**, 'Factorisation and denoising of 0-1 data: a variational approach', **Neurocomputing**, vol 71, no. 10-12, pp. 2291-2308.

Kaski, P, Östergård, PRJ **2008**, 'There are exactly five biplanes with $k = 11$ ', **Journal of combinatorial designs**, vol 16, pp. 117-127.

Kaski, P, Östergård, PRJ, Topalova, S, Zlatarski, R **2008**, 'Steiner triple systems of order 19 and 21 with subsystems of order 7', **Discrete Mathematics**, vol 308, no. 13, pp. 2732-2741.

Kaski, P, Penttinen, A, Suomela, J **2008**, 'Coordinating concurrent transmissions: A constant-factor approximation of maximum-weight independent set in local conflict graphs', **Ad-Hoc & Sensor Wireless Networks**, vol 6, no. 3-4, pp. 239-263.

Kulathinal, S, Gasbarra, D, Kinra, S, Ebrahim, S, Sillanpää, MJ **2008**, 'Estimation of additive genetic and environmental sources of quantitative trait variation using data on married couples and their siblings', **Genetics Research**, vol 90, no. 3, pp. 269-279.

Leino, A, Hyvönen, S **2008**, 'Comparison of Component Models in Analysing the Distribution of Dialectal Features', **International Journal of Humanities and Arts Computing**, vol 2, no. 1-2, pp. 173-187.

Lindgren, JT, Hurri, J, Hyvärinen, A **2008**, 'Spatial dependencies between local luminance and contrast in natural images', **Journal of vision**, vol 8, no. 12, pp. 1-13.

Liow, LH, Fortelius, M, Bingham, E, Lintulaakso, K, Mannila, H, Flynn, L, Stenseth, NC **2008**, 'Reply to Vilar et al: Sleep or hide, better for survival anytime', **Proceedings of the National Academy of Sciences of the United States of America**, vol 105, no. 35, pp. E57.

Liow, LH, Fortelius, M, Bingham, E, Lintulaakso, K, Mannila, H, Flynn, L, Stenseth, NC **2008**, 'Higher origination and extinction rates in larger mammals', **Proceedings of the National Academy of Sciences of the United States of America**, vol 105, no. 16, pp. 6097-6102.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Miettinen, P **2008**, 'The Boolean column and column-row matrix decompositions', **Data Mining and Knowledge Discovery**, vol 17, no. 1-2, pp. 39-56.
- Miettinen, P **2008**, 'On the Positive-Negative Partial Set Cover problem', **Information Processing Letters**, vol 108, no. 4, pp. 219-221.
- Miettinen, P, Mielikäinen, T, Gionis, A, Das, G, Mannila, H **2008**, 'The discrete basis problem', **IEEE Transactions on Knowledge and Data Engineering**, vol 20, no. 10, pp. 1348-1362.
- Mitchell, JSB, Polishchuk, V **2008**, 'Minimum-perimeter enclosures', **Information Processing Letters**, vol 107, pp. 120-124.
- Mäkinen, V, Navarro, G **2008**, 'Dynamic entropy-compressed sequences and full-text indexes', **ACM transactions on algorithms**, vol 4, no. 3, pp. art. 32.
- Pirinen, M, Kulathinal, S, Gasbarra, D, Sillanpää, MJ **2008**, 'Estimating population haplotype frequencies from pooled DNA samples using PHASE algorithm', **Genetics Research**, vol 90, pp. 509-524.
- Pitkänen, E, Åkerlund, A, Rantanen, A, Jouhten, P, Ukkonen, E **2008**, 'ReMatch: a web-based tool to construct, store and share stoichiometric metabolic models with carbon maps for metabolic flux analysis', **Journal of Integrative Bioinformatics : JIB**, vol 5, no. 102, 13 s.
- Pizzi, C, Ukkonen, E **2008**, 'Fast profile matching algorithms: a survey', **Theoretical Computer Science**, vol 395, pp. 137-157.
- Raento, MP, Oulasvirta, A **2008**, 'Designing for privacy and self-presentation in social awareness', **Personal and Ubiquitous Computing**, vol 12, no. 7, pp. 527-542.
- Rantanen, A, Rousu, J, Jouhten, P, Zamboni, N, Maaheimo, H, Ukkonen, E **2008**, 'An analytic and systematic framework for estimating metabolic flux ratios from ¹³C tracer experiments', **BMC Bioinformatics**, vol 9, no. 266, 19 s.
- Sevon, P, Eronen, L **2008**, 'Subgraph queries by context-free grammars', **Journal of Integrative Bioinformatics : JIB**, vol 5, no. 100, 16 s.
- Sillanpää, MJ, Noykova, N **2008**, 'Hierarchical modeling of clinical and expression quantitative trait loci', **Heredity**, vol 101, pp. 271-284.
- Tripathi, A, Klamí, A, Kaski, S **2008**, 'Simple integrative preprocessing preserves what is shared in data sources: methodology article', **BMC Bioinformatics**, vol 9, no. 111, 13 s.
- Varjosalo, M, Björklund, M, Cheng, F, Syvänen, H, Kivioja, T, Kilpinen, S, Sun, Z, Kallioniemi, O, Stunnenberg, HG, He, W, Ojala, P, Taipale, J **2008**, 'Application of active and kinase-deficient kinome collection for identification of kinases regulating hedgehog signaling', **Cell**, vol 133, no. 3, pp. 537-548.
- Waldmann, P, Hallander, J, Hoti, F, Sillanpää, MJ **2008**, 'Efficient Markov chain Monte Carlo implementation of Bayesian analysis of additive and dominance genetic variances in noninbred pedigrees', **Genetics**, vol 179, no. 2, pp. 1101-1112.
- Yu, H, Bertsekas, DP **2008**, 'On near optimality of the set of finite-state controllers for average cost POMDP', **Mathematics of Operations Research**, vol 33, no. 1, 1-11 s.
- 2009**
- Ala-Härkönen, V, Brunberg, J, Lemström, K, Mikkilä, N **2009**, 'A content-based music retrieval engine: JMIR-Mozart', **Musicae Scientiae**, vol 4B (2009), pp. 235-256.
- Arkin, EM, Fekete, SP, Islam, K, Meijer, H, Mitchell, JSB, Nunez-Rodriguez, Y, Polishchuk, V, Rappaport, D, Xiao, H **2009**, 'Not being (super)thin or solid is hard: a study of grid Hamiltonicity', **Computational Geometry**, vol 42, no. 6-7, pp. 582-605.
- Arkin, EM, Bae, SW, Efrat, A, Okamoto, K, Mitchell, JSB, Polishchuk, V **2009**, 'Geometric stable roommates', **Information Processing Letters**, vol 109, no. 4, pp. 219-224.
- Bauer, AM, Hoti, F, von Korff, M, Pillen, K, Leon, J, Sillanpää, MJ **2009**, 'Advanced backcross-QTL analysis in spring barley (*H. vulgare* ssp. *spontaneum*) comparing a REML versus a Bayesian model in multi-environmental field trials', **Theoretical and Applied Genetics**, vol 119, no. 1, pp. 105-123.
- Bauer, AM, Hoti, F, Reetz, TC, Schuh, W, Leon, J, Sillanpää, MJ **2009**, 'Bayesian prediction of breeding values by accounting for genotype-by-environment interaction in self-pollinating crops', **Genetics Research**, vol 91, pp. 193-207.
- Benachou, F, Jern, P, Oja, M, Sperber, G, Blikstad, V, Somervuo, P, Kaski, S, Blomberg, J **2009**, 'Evolutionary Conservation of Orthoretroviral Long Terminal Repeats (LTRs) and ab initio Detection of Single LTRs in Genomic Data', **PLoS One**, vol 4, no. 4, pp. e5179.
- Bertsekas, DP, Yu, H **2009**, 'Projected equation methods for approximate solution of large linear systems', **Journal of Computational and Applied Mathematics**, vol 227 (2009), pp. 27-50.
- Bingham, E, Mannila, H **2009**, 'Complexity control in a mixture model by the Hardy-Weinberg equilibrium', **Computational Statistics & Data Analysis**, vol 53, no. 5, pp. 1711-1719.
- Björklund, A, Husfeldt, T, Koivisto, M **2009**, 'SET PARTITIONING VIA INCLUSION-EXCLUSION', **SIAM Journal on Computing**, vol 39, no. 2, pp. 546-563.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Fischer, J, Mäkinen, V, Navarro, G **2009**, 'Faster entropy-bounded compressed suffix trees', **Theoretical Computer Science**, vol 410 (2009), pp. 5354-5364.
- Gasbarra, D, Pirinen, M, Arjas, E, Sillanpää, MJ **2009**, 'Bayesian quantitative trait locus mapping based on reconstruction of recent genetic histories', **Genetics**, vol 183, no. 2, pp. 709-721.
- Hao, LQ, Yli-Pirilä, P, Tiitta, P, Romakkaniemi, S, Vaattovaara, P, Kajos, MK, Rinne, J, Heijari, J, Kortelainen, A, Miettinen, P, Kroll, JH, Holopainen, JK, Smith, JN, Joutsensaari, J, Kulmala, M, Worsnop, DR, Laaksonen, A **2009**, 'New particle formation from the oxidation of direct emissions of pine seedlings', **Atmospheric Chemistry and Physics**, vol 9, no. 20, pp. 8121-8137.
- Henriksson, L, Hyvärinen, A, Vanni, S **2009**, 'Representation of cross-frequency spatial phase relationships in human visual cortex', **Journal of Neuroscience**, vol 29, no. 45, pp. 14342-14351.
- Jouhten, P, Pitkänen, E, Pakula, T, Saloheimo, M, Penttilä, M, Maaheimo, H **2009**, '¹³C-metabolic flux ratio and carbon path analyses confirmed that *Trichoderma reesei* uses primarily the respirative pathway also on the preferred carbon source glucose', **BMC Systems Biology**, vol 3, no. 104, 16 s.
- Kaski, P, Östergård, PRJ **2009**, 'There are 1,132,835,421,602,062,347 nonisomorphic one-factorizations of K_n ', **Journal of combinatorial designs**, vol 17, pp. 147-159.
- Kaski, P, Östergård, PRJ, Pottönen, O, Kiviluoto, L **2009**, 'A catalogue of the Steiner triple systems of order 19', **Bulletin of the Institute of Combinatorics and its Applications**, vol 57, pp. 35-41.
- Kaski, P, Östergård, PRJ **2009**, 'Classification of resolvable balanced incomplete block designs: the unitals on 28 points', **Mathematica Slovaca**, vol 59, pp. 121-136.
- Korhonen, JH, Martinmäki, P, Pizzi, C, Rastas, P, Ukkonen, E **2009**, 'MOODS: fast search for position weight matrix matches in DNA sequences', **Bioinformatics**, vol 25, no. 23, pp. 3181-3182.
- Kurki, I, Peromaa, T, Hyvärinen, A, Saarinen, JJ **2009**, 'Visual features underlying perceived brightness as revealed by classification images', **PLoS One**, vol 4, no. 10, pp. e7432.
- Kurki, I, Peromaa, T, Hyvärinen, A, Saarinen, J, Kurki, I **2009**, 'Visual features underlying perceived brightness as revealed by classification images', **PLoS One**, vol 4, no. 10, pp. e7432.
- Linge, J, Steinberger, R, Weber, T, Yangarber, R, van der Goot, E, Al Khudhairy, D, Stilianakis, N **2009**, 'Internet surveillance systems for early alerting of health threats', **Eurosurveillance**, vol 14, no. 13.
- Liow, LH, Fortelius, M, Lintulaakso, K, Mannila, H, Stenseth, NC **2009**, 'Lower Extinction Risk in Sleep-or-Hide Mammals', **American Naturalist**, vol 173, no. 2, pp. 264-272.
- Nicolas, F, Mäkinen, V, Ukkonen, E **2009**, 'Efficient construction of maximal and minimal representations of motifs of a string', **Theoretical Computer Science**, vol 410 (2009), pp. 2999-3005.
- Ojala, M, Vuokko, N, Kallio, A, Haiminen, N, Mannila, H **2009**, 'Randomization methods for assessing data analysis results on real-valued matrices', **Statistical analysis and data mining**, vol 2, no. 4, pp. 209-230.
- Pikkuhookana, P, Sillanpää, M **2009**, 'Correcting for relatedness in Bayesian models for genomic data association analysis', **Heredity**, vol 103, no. 3, pp. 223-237.
- Pitkänen, E, Jouhten, P, Rousu, J **2009**, 'Inferring branching pathways in genome-scale metabolic networks', **BMC Systems Biology**, vol 3, no. 103, 22 s.
- Polishchuk, V, Suomela, J **2009**, 'A simple local 3-approximation algorithm for vertex cover', **Information Processing Letters**, vol 109, no. 12, pp. 642-645.
- Roos, TT, Heikkilä, T **2009**, 'Evaluating methods for computer-assisted stemmatology using artificial benchmark data sets', **Literary and Linguistic Computing**, vol 24, no. 4, pp. 417-433.
- Roos, T, Myllymäki, P, Rissanen, J **2009**, 'MDL denoising revisited', **IEEE Transactions on Signal Processing**, vol 57, no. 9, pp. 3347-3360.
- Savola, S, Klam, A, Tripathi, A, Niini, T, Serra, M, Picci, P, Kaski, S, Zambelli, D, Scotlandi, K, Knuutila, S **2009**, 'Combined use of expression and CGH arrays pinpoints novel candidate genes in Ewing sarcoma family of tumors', **BMC Cancer**, vol 9, pp. 17.
- Shimizu, S, Hoyer, PO, Hyvärinen, A **2009**, 'Estimation of linear non-Gaussian acyclic models for latent factors', **Neurocomputing**, vol 72, no. 7-9, pp. 2024-2027.
- Sillanpää, MJ **2009**, 'Detecting interactions in association studies by using simple allele recoding', **Human Heredity**, vol 67, no. 1, pp. 69-75.
- Tuupanen, S, Turunen, M, Lehtonen, R, Hallikas, O, Vanharanta, S, Kivioja, T, Björklund, M, Wei, G, Yan, J, Niittymäki, I, Mecklin, J, Järvinen, H, Ristimäki, A, Di-Bernardo, M, East, P, Carvajal-Carmona, L, Houlston, RS, Tomlinson, I, Palin, K, Ukkonen, E, Karhu, A, Taipale, J, Aaltonen, LA **2009**, 'The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling', **Nature Genetics**, vol 41, no. 8, pp. 885-890.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Ukkonen, E 2009, 'Maximal and minimal representations of gapped and non-gapped motifs of a string', **Theoretical Computer Science**, vol , 410, no. 43, pp. 4341-4349.

Vingron, M, Brazma, A, Coulson, R, van Helden, J, Manke, T, Palin, K, Sand, O, Ukkonen, E 2009, 'Integrating sequence, evolution and functional genomics in regulatory genomics', **Genome Biology**, vol 10, no. 1, 202 , 8 s.

Välimäki, N, Mäkinen, V, Gerlach, W, Dixit, K 2009, 'Engineering a Compressed Suffix Tree Implementation', **Journal of Experimental Algorithmics**, vol 14, no. 4.2, pp. 4.2:1-4.2:23.

Wessman, J, Paunio, T, Tuulio-Henriksson, A, Koivisto, M, Partonen, T, Suvisaari, J, Turunen, JA, Wedenoja, J, Hennah, W, Pietiläinen, O, Lönnqvist, J, Mannila, H, Peltonen, L 2009, 'Mixture model clustering of phenotype features reveals evidence for association of DTNBP1 to a specific subtype of schizophrenia', **Biological Psychiatry**, vol 66, no. 11, pp. 990-996.

Yu, H, Bertsekas, DP 2009, 'Convergence results for some temporal difference methods based on least squares', **IEEE Transactions on Automatic Control**, vol 54, no. 7, pp. 1515-1531.

2010

Ahopuro, P, Björklund, M, Sammalkorpi, HP, Turunen, M, Tuupanen, SJ, Bistrom, M, Niittymäki, I, Lehtonen, H, Kivioja, T, Launonen, V, Saharinen, JMN, Nousiainen, KJ, Hautaniemi, S, Nuorva, K, Mecklin, J, Järvinen, H, Orntoft, T, Arango, D, Lehtonen, RJ, Karhu, AI, Taipale, J, Aaltonen, LA 2010, 'Mutations in the Circadian Gene CLOCK in Colorectal Cancer', **Molecular Cancer Research**, vol 8, no. 7, pp. 952-960.

Arkin, E, Mitchell, J, Polishchuk, V 2010, 'Maximum Thick Paths in Static and Dynamic Environments', **Computational Geometry**, vol 43, no. 3, pp. 279-294.

Arvas, M, Haiminen, NS, Smit, B, Rautio, J, Vitikainen, M, Wiebe, M, Martinez, D, Chee, C, Kunkel, J, Sanchez, C, Nelson, MA, Pakula, N, Saloheimo, M, Penttilä, M, Kivioja, T 2010, 'Detecting novel genes with sparse arrays', **Gene**, vol 467, no. 1-2, pp. 41-51.

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M 2010, 'Trimmed Moebius Inversion and Graphs of Bounded Degree', **Theory of Computing Systems**, vol 47, no. 3, pp. 637-654.

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M 2010, 'Evaluation of permanents in rings and semirings', **Information Processing Letters**, vol 110, no. 20, pp. 867-870.

Brummayer, R, Järvisalo, M 2010, 'Testing and Debugging Techniques for Answer Set Solver Development', **Theory and Practice of Logic Programming**, vol 10, no. 4-6, pp. 741-758.

Colbourn, CJ, Forbes, AD, Grannell, MJ, Griggs, TS, Kaski, P, Östergård, PRJ, Pike, DA, Pottönen, O 2010, 'Properties of the Steiner triple systems of order 19', **The Electronic Journal of Combinatorics**, vol 17, no. 1, pp. R98.

Dias, G, Moraliyski, R, Cordeiro, J, Doucet, A, Ahonen-Myka, H 2010, 'Automatic Discovery of Word Semantic Relations using Paraphrase Alignment and Distributional Lexical Semantics Analysis', **Natural Language Engineering**, vol 16, no. 4, pp. 439-467.

Doucet, A, Ahonen-Myka, H 2010, 'An efficient any language approach for the integration of phrases in document retrieval', **Language Resources and Evaluation**, vol 44, no. 1-2, pp. 159-180.

Floréen, P, Kaski, P, Polishchuk, V, Suomela, J 2010, 'Almost Stable Matchings by Truncating the Gale-Shapley Algorithm', **Algorithmica**, vol 58, no. 1, pp. 102-118.

Garriga, GC, Junttila, E, Mannila, H 2010, 'Banded structure in binary matrices', **Knowledge and Information Systems**, vol 2010.

Haiminen, NS, Mannila, H 2010, 'Evaluation of BIC and cross validation for model selection on sequence segmentations', **International Journal of Data Mining and Bioinformatics**, vol 4, no. 6, pp. 675-700.

Hallander, J, Waldmann, P, Wang, C, Sillanpää, MJ 2010, 'Bayesian inference of genetic parameters based on conditional decompositions of multivariate normal distributions', **Genetics**, vol 185, no. 2, pp. 645-654.

Hartley, D, Nelson, N, Walters, R, Arthur, R, Yangarber, R, Madoff, L, Linge, J, Mawudeku, A, Collier, N, Brownstein, J, Thinus, G, Lightfoot, N 2010, 'The landscape of international event-based biosurveillance', **Emerging Health Threats Journal**, vol 2010/3, pp. 2-18.

Honkela, T, Hyvärinen, A, Väyrynen, JJ 2010, 'WordICA-emergence of linguistic representations for words by independent component analysis', **Natural Language Engineering**, vol 16, no. 3, pp. 277-308.

Huopaniemi, I, Suvitaval, T, Nikkilä, J, Oresic, M, Kaski, S 2010, 'Multivariate multi-way analysis of multi-source data', **Bioinformatics**, vol 26, no. 12, pp. i391-i398.

Hyvärinen, A, Zhang, K, Shimizu, S, Hoyer, PO 2010, 'Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity', **Journal of machine learning research**, vol 11 , pp. 1709-1731 .

Hyvärinen, A, Ramkumar, P, Parkkonen, L, Hari, R 2010, 'Independent component analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis', **NeuroImage**, vol 49, no. 1, pp. 257-271.

Hämäläinen, SW 2010, 'StatApriori: an efficient algorithm for searching statistically significant association rules', **Knowledge and Information Systems**, vol 23, no. 3, pp. 373-399.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Jolma, A, Kivioja, T, Toivonen, J, Cheng, L, Wei, G, Enge, M, Taipale, M, Vaquerizas, JM, Yan, J, Sillanpää, MJ, Bonke, AWM, Palin, K, Talukder, S, Hughes, TR, Luscombe, NM, Ukkonen, E, Taipale, J **2010**, 'Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities', **Genome Research**, vol 20, no. 6, pp. 861-873.
- Köster, U, Hyvärinen, A **2010**, 'A Two-Layer Model of Natural Stimuli Estimated with Score Matching', **Neural Computation**, vol 22, no. 9, pp. 2308-2333.
- Lahti, L, Knuuttila, JEA, Kaski, S **2010**, 'Global modeling of transcriptional responses in interaction networks', **Bioinformatics**, vol 26, pp. 2713-2720.
- Lemstrom, K, Mikkilä, N, Mäkinen, V **2010**, 'Filtering methods for content-based retrieval on indexed symbolic music databases', **Information retrieval**, vol 13, pp. 1-21.
- Lukk, M, Kapushesky, M, Nikkilä, JT, Parkinson, H, Goncalves, A, Huber, W, Ukkonen, E, Brazma, A **2010**, 'A global map of human gene expression', **Nature Biotechnology**, vol 28, no. 4, pp. 322-324.
- Mavroudis, D, Bingham, E **2010**, 'Enhancing the stability and efficiency of spectral ordering with partial supervision and feature selection', **Knowledge and Information Systems**, vol 23, no. 2, pp. 243-265.
- Michael, M, Nicolas, F, Ukkonen, E **2010**, 'On the complexity of finding gapped motifs', **Journal of Discrete Algorithms**, vol 8, no. 2, pp. 131-142.
- Mutshinda, CM, Sillanpää, MJ **2010**, 'Extended Bayesian LASSO for multiple quantitative trait loci mapping and unobserved phenotype prediction', **Genetics**, vol 186, pp. 1067-1075.
- Mäkinen, V, Navarro, G, Sirén, J, Välimäki, N **2010**, 'Storage and Retrieval of Highly Repetitive Sequence Collections', **Journal of Computational Biology**, vol 17, no. 3, pp. 281-308.
- Parkkinen, J, Kaski, S **2010**, 'Searching for functional gene modules with interaction component models', **BMC Systems Biology**, vol 4, no. 4.
- Peltonen, J, Yaslan, Y, Kaski, S **2010**, 'Relevant subtask learning by constrained mixture models', **Intelligent Data Analysis**, vol 14, no. 6, pp. 641-662.
- Perkiö, J, Tuominen, A, Vähäkangas, T, Myllymäki, P **2010**, 'Image Similarity: From Syntax to Weak Semantics', **Multimedia Tools and Applications**.
- Pitkänen, E, Rousu, J, Ukkonen, E **2010**, 'Computational methods for metabolic reconstruction', **Current Opinion in Biotechnology**, vol 21, no. 1, pp. 70-77.
- Rissanen, J, Roos, T, Myllymäki, P **2010**, 'Model selection by sequentially normalized least squares', **Journal of Multivariate Analysis**, vol 101, pp. 839-849.
- Saarinen, J, Oikarinen, E, Fortelius, M, Mannila, H **2010**, 'The living and the fossilized: how well do unevenly distributed points capture the faunal information in a grid?', **Evolutionary Ecology Research**, vol 12, pp. 363-376.
- Salmela, L **2010**, 'Correction of sequencing errors in a mixed set of reads', **Bioinformatics**, vol 26, no. 10, pp. 1284-1290.
- Silander, T, Roos, TT, Myllymäki, P **2010**, 'Learning locally minimax optimal Bayesian networks', **International Journal of Approximate Reasoning**, vol 51/2010, no. 5, pp. 544-557.
- Tanner, T, Toivonen, H **2010**, 'Predicting and preventing student failure - using the k-nearest neighbour method to predict student performance in an online course environment', **International Journal of Learning Technology**, vol 5, no. 4, pp. 356-377.
- Ukkonen, E **2010**, 'Geometric Point Pattern Matching in the Knuth-Morris-Pratt Way', **Journal of Universal Computer Science**, vol 16, no. 14, pp. 1902-1911.
- Vesala, T, Launiainen, S, Kolari, P, Pumpanen, J, Sevanto, S, Hari, P, Nikinmaa, E, Kaski, P, Mannila, H, Ukkonen, E, Piao, SL, Ciais, P **2010**, 'Autumn temperature and carbon balance of a boreal Scots pine forest in Southern Finland', **Biogeosciences**, vol 7, no. 1, pp. 163-176.
- Wei, G, Badis, G, Berger, MF, Kivioja, T, Palin, K, Enge, M, Bonke, M, Jolma, A, Varjosalo, M, Gehrke, AR, Yan, J, Talukder, S, Turunen, M, Taipale, M, Stunnenberg, HG, Ukkonen, E, Hughes, TR, Bulik, ML, Taipale, J **2010**, 'Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo', **EMBO Journal**, vol 29, no. 13, pp. 2147-2160.
- Yu, H, Bertsekas, DP **2010**, 'Error Bounds for Approximations from Projected Linear Equations', **Mathematics of Operations Research**, vol 35, no. 2, pp. 306-329.

A2 Review in scientific journal

2007

- Kaski, S, Rousu, J, Ukkonen, E **2007**, 'Probabilistic modeling and machine learning in structural and systems biology', **BMC Bioinformatics**, vol 8, no. Suppl 2, pp. S1.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

2009

O'Hara, R, Sillanpää, MJ **2009**, 'A review of Bayesian variable selection methods: what, how and which', **Bayesian analysis**, vol 4, no. 1, pp. 85-118.

2010

Hyvärinen, A **2010**, 'Statistical models of natural images and cortical visual representation', **Topics in Cognitive Science**, vol 2, no. 2, pp. 251-264.

A3 Contribution to book/other compilations (refereed)

2005

Ahonen-Myka, H, Doucet, A **2005**, 'Data mining meets collocations discovery', **Inquiries into words, constraints and contexts. Festschrift in the honour of Kimmo Koskenniemi on his 60th birthday.**, [s.n.], [S.l.], pp. 194-203.

Ahonen-Myka, H **2005**, 'Text analysis by discovering frequent phrases', **Modernin informaatioteknologian menetelmätutkimusta, Petroskoin yliopisto, Petroskoi**, pp. 96-112.

Aunimo, L, Kuuskoski, R, Makkonen, J **2005**, 'Finnish as source language in bilingual question answering', **Multilingual information access for text, speech and images, Springer, Berlin**.

Eronen, L, Geerts, F, Toivonen, H **2005**, 'Efficient Markovian algorithms for haplotype reconstruction', **Modernin informaatioteknologian menetelmätutkimusta, Petroskoin yliopisto, Petroskoi**, pp. 16-39.

Goethals, B, Laur, S, Lipmaa, H, Mielikäinen, T **2005**, 'On private scalar product computation for privacy-preserving data mining', **Information security and cryptology-ICISC 2004, Springer, cop., Berlin**, pp. 104-120.

Kaski, S, Nikkilä, J, Savia, E, Roos, C **2005**, 'Discriminative clustering of yeast stress response', **Bioinformatics using computational intelligence paradigms, Springer, Berlin**, pp. 75-92.

Kontkanen, P, Myllymäki, P, Buntine, W, Rissanen, J, Tirri, H **2005**, 'An MDL framework for data clustering', **Advances in minimum description length, MIT Press cop., Cambridge (Mass.)**, pp. 323-353.

Laasonen, K **2005**, 'Clustering and prediction of mobile user routes from cellular data', **Knowledge discovery in databases, Springer, cop., Berlin**, pp. 569-576.

Lehtonen, M **2005**, 'EXTIRP 2004: towards heterogeneity', **Advances in XML information retrieval, Springer, cop., Berlin**, pp. 372-381.

Leino, A **2005**, 'In search of naming patterns: a survey of Finnish lake names', **Rivista italiana di onomastica; Supplemento al n. XI, 1 di RION (QuadRION), Quaderni Internazionali di RION, vol. 1, Societa Editrice Romana, Roma**, pp. 355-367.

Mielikäinen, T **2005**, 'An automata approach to pattern collections', **Knowledge discovery in inductive databases, Springer, cop., Berlin**, pp. 130-149.

Mielikäinen, T **2005**, 'Implicit enumeration of patterns', **Knowledge discovery in inductive databases, Springer, cop., Berlin**, pp. 150-172.

Salmenkivi, M, Mannila, H **2005**, 'Piecewise Constant Modeling of Sequential Data Using Reversible Jump Markov Chain Monte Carlo: chapter 5', in JT Wang, MJ Zaki, HT Toivonen, D Shasha (eds), **Data Mining in Bioinformatics**, 1. edn, **Advanced Information and Knowledge Processing, Springer, New York**, pp. 85-103.

Salojärvi, J, Puolamäki, K, Kaski, S **2005**, 'Implicit relevance feedback from eye movements', **Artificial neural networks, Springer, New York, NY**, pp. 513-518.

Salojärvi, J, Puolamäki, K, Kaski, S **2005**, 'On discriminative joint density modeling', **Machine learning, Springer, New York, NY**, pp. 341-352.

Sevon, P, Toivonen, HTT, Onkamo, P **2005**, 'Gene Mapping by Pattern Discovery: chapter 6', **Data mining in bioinformatics, Springer, New York**, pp. 105-126.

Toivonen, H, Onkamo, P, Hintsanen, P, Terzi, E, Sevon, P **2005**, 'Data mining for gene mapping', **Next generation of data-mining applications, Wiley-Interscience, Hoboken, N.J.** pp. 263-293.

Wang, JTL, Zaki, MJ, Toivonen, HTT, Shasha, D **2005**, 'Introduction to data mining in bioinformatics: chapter 1', **Data mining in bioinformatics, Springer, New York**, pp. 3-8.

2006

Buntine, W, Jakulin, A **2006**, 'Discrete component analysis', **Subspace, latent structure and feature selection, Springer, cop., Berlin**, pp. 1-33.

Doucet, A **2006**, 'Advanced document description, a sequential approach', **Forum, New York (NY)**, pp. 71-72.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Kulovesi, K, Muhonen, J, Lappalainen, I, Riikonen, PT, Vihinen, M, Toivonen, H, Pasanen, T **2006**, 'Visualisation of associations between nucleotides in SNP neighbourhoods', **Intelligent data analysis in biomedicine and pharmacology, University of Verona, Verona**, pp. 61-62.

Kääriäinen, M **2006**, 'Active learning in the non-realizable case', **Algorithmic learning theory, Springer, Berlin**, pp. 63-77.

Landwehr, N, Mielikäinen, T, Eronen, L, Toivonen, H, Mannila, H **2006**, 'Constrained Hidden Markov Models for population-based haplotyping', **Probabilistic modeling and machine learning in structural and systems biology, Helsinki University Printing House, Helsinki**, pp. 38-43.

Lehtonen, M **2006**, 'Designing user studies for XML retrieval', **SIGIR 2006, ACM, New York (NY)**.

Lehtonen, M **2006**, 'When a few highly relevant answers are enough', **Advances in XML information retrieval and evaluation, Springer, Berlin**, pp. 296-305.

Mielikäinen, T **2006**, 'Transaction databases, frequent itemsets, and their condensed representations', **Knowledge discovery in inductive databases, Springer, cop., Berlin**, pp. 139-164.

Mielikäinen, T **2006**, 'Tietämyksen yhteenvedot', **Tietojenkäsittelytieteen päivät 2006, Helsingin Yliopisto, Helsinki**, pp. 12-17.

Rantanen, A, Maaheimo, H, Pitkänen, E, Rousu, J, Ukkonen, E **2006**, 'Equivalence of metabolite fragments and flow analysis of isotopomer distributions for flux estimation', **Transactions on computational systems biology VI, Springer, cop., Berlin**, pp. 198-220.

Sevon, P, Eronen, L, Hintsanen, P, Kulovesi, K, Toivonen, H **2006**, 'Link discovery in graphs derived from biological databases', **Data integration in the life sciences, Springer, Berlin**, pp. 35-49.

2007

Leino, A **2007**, 'Regional variation in Finnish lake and hill names', **Nordiske navnes centralitet og regionalitet, NORNA-förlaget, Uppsala**, pp. 123-144.

Rousu, J, Saunders, C, Szedmak, S, Shawe-Taylor, J **2007**, 'Efficient algorithms for Max-Margin structured classification', **Predicting structured data, MIT Press cop., Cambridge, Mass**, pp. 105-128.

2008

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M **2008**, 'The travelling salesman problem in bounded degree graphs', **Automata, languages and programming, Springer, Berlin**, pp. 198-209.

Hintsanen, P, Toivonen, H **2008**, 'Finding reliable subgraphs from large probabilistic graphs', **Machine learning and knowledge discovery in databases, Springer, cop., Berlin**.

Kärkkäinen, J, Ukkonen, E **2008**, 'Multidimensional string matching', **Encyclopedia of Algorithms, Springer-Verlag, Boston, MA**, pp. 559-562.

Kärkkäinen, J **2008**, 'Suffix array construction', **Encyclopedia of Algorithms, Springer-Verlag, Boston, MA**, pp. 919-921.

Myllymäki, P, Roos, T, Silander, T, Kontkanen, P, Tirri, H **2008**, 'Factorized NML models', **Festschrift in honor of Jorma Rissanen on the occasion of his 75th birthday, Tampere International Center for Signal Processing, Tampere**, pp. 189-204.

Mäkinen, V, Ukkonen, E **2008**, 'Point pattern matching', **Encyclopedia of Algorithms, Springer-Verlag, Boston, MA**, pp. 657-660.

Mäkinen, V, Navarro, G **2008**, 'Compressed text indexing', **Encyclopedia of Algorithms, Springer-Verlag, Boston, MA**, pp. 176-178.

Mäkinen, V **2008**, 'Compressed suffix array', **Encyclopedia of Algorithms, Springer-Verlag, Boston, MA**, pp. 174-176.

Nock, R, Luosto, P, Kivinen, J **2008**, 'Mixed Bregman clustering with approximation guarantees', **Machine learning and knowledge discovery in databases, Springer, Berlin**, pp. 154-169.

Rastas, P, Koivisto, M, Mannila, H, Ukkonen, E **2008**, 'Phasing genotypes using a hidden Markov model', **Bioinformatics algorithms, Wiley-Interscience cop., Hoboken, N.J**, pp. 355-372.

Rissanen, JJ **2008**, 'Minimum description length', in M Hutter, S Chennu (eds), **Scholarpedia. The peer-reviewed open-access encyclopedia.**, 3(8):6727 edn.

Yu, H, Bertsekas, DP **2008**, 'New error bounds for approximations from projected linear equations', **Recent advances in reinforcement learning. Sertan Girgin ... [et al.] (eds.)**, **Lecture notes in computer science, Springer, Berlin**, pp. 253-267.

2009

Bhattacharjee, M, Sillanpää, M **2009**, 'Bayesian joint disease-marker-expression analysis applied to clinical characteristics of chronic fatigue syndrome', in P McConnell, S Lim, A Cuticchia (eds), **Methods of Microarray Data Analysis VI, CreateSpace Publishing, Scotts Valley, California**, pp. 15-34.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M **2009**, 'Counting paths and packings in Halves', **Algorithms - ESA 2009 edited by Amos Fiat, Peter Sanders, Lecture notes in computer science, Springer Berlin Heidelberg, Berlin, Heidelberg**, pp. 578-586.

Coulson, R, Manke, T, Palin, K, Roeder, H, Sand, O, van Helden, J, Ukkonen, E, Vingron, M, Brazma, A **2009**, 'From gene expression profiling to gene regulation', **Modern genome annotation. Dmitrij Frishman, Alfonso Valencia, editors., Springer, cop., Wien**, pp. 105-119.

Koivisto, M **2009**, 'Partitioning into sets of bounded cardinality', **Parameterized and exact computation. Jianer Chen, Fedor V. Fomin (eds.), Lecture notes in computer science, Springer-Verlag, Berlin**, pp. 258-263.

Lemström, K, Tzanetakis, T **2009**, 'Music Information Retrieval in Polyphonic Music', in MJ Bates, MN Maack (eds), **Encyclopedia of Library and Information Sciences, Third Edition, CRC Press**.

Lemström, K, Wiggins, GA **2009**, 'Formalizing invariances for content-based music retrieval', **ISMIR 2009. edited by Keiji Hirata, George Tzanetakis and Kazuyoshi Yoshii. - The International Society for Music Information Retrieval, 2009., pp. 591-596.**

Rastas, P **2009**, 'A general framework for local pairwise alignment statistics with gaps', **Algorithms in bioinformatics. Steven L. Salzberg, Tandy Warnow (eds.), Lecture Notes in Bioinformatics, Springer, cop., Berlin**, pp. 233-245.

Rissanen, JJ **2009**, 'Model Selection and Testing by the MDL Principle: Chapter 2', in F Emmert-Streib, M Dehmer (eds), **Information Theory and Statistical Learning, Springer-Verlag**.

Rissanen, J **2009**, 'Model selection and testing by the MDL principle', in [FEMD (ed.), **Information theory and statistical learning, Springer, cop., New York**, pp. 25-43.

2010

De Raedt, L, Kimmig, A, Gutmann, B, Kersting, K, Santos Costa, V, Toivonen, H **2010**, 'Probabilistic Inductive Querying Using ProbLog', in S Dzeroski, B Goethals, P Panov (eds), **Inductive Databases and Constraint-Based Data Mining, Springer**, pp. 229-262.

Hämäläinen, W, Vinni, M **2010**, 'Classifiers for educational technology', in C Romero, S Ventura, M Pechenizkiy, RS Baker (eds), **Handbook on Educational Data Mining, Chapman & Hall/CRC Data Mining and Knowledge Discovery Series , CRC Press**, pp. 54-74.

Kaski, S **2010**, 'Three Paths to Relevance', in A Hanazawa, T Miki, H Keichi (eds), **Brain-inspired information technology, Studies in Computational Intelligence, no. 11-13, vol. 266, Springer**, pp. 11-13.

Lemström, K **2010**, 'Transposition and time-scale invariant geometric music retrieval', in T Elomaa, H Mannila, P Orponen (eds), **Algorithms and applications. Essays dedicated to Esko Ukkonen on the occasion of his 60th birthday., Lecture Notes in Computer Science, no. 6060, vol. 2010, Heidelberg, Berlin, Springer Verlag.,**

Mäkinen, V, Välimäki, N, Laaksonen, A, Katainen, R **2010**, 'Unified view of backward backtracking in short read mapping', in T Elomaa, H Mannila, P Orponen (eds), **Algorithms and applications. Essays dedicated to Esko Ukkonen on the occasion of his 60th birthday., Lecture Notes in Computer Science, no. 6060, vol. 2010, Heidelberg, Berlin, Springer Verlag.,**

Rissanen, J **2010**, 'Minimum Description Length Principle', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning. Part 14., Springer-Verlag**, pp. 666-668.

Salmela, L, Tarhio, J **2010**, 'Approximate string matching with reduced alphabet', in T Elomaa, H Mannila, P Orponen (eds), **Algorithms and applications. Essays dedicated to Esko Ukkonen on the occasion of his 60th birthday., Lecture Notes in Computer Science, no. 6060, vol. 2010, Heidelberg, Berlin, Springer Verlag.,**

Zarzoso, V, Hyvärinen, AJ **2010**, 'Iterative algorithms', **Handbook on Independent Component Analysis and Blind Source Separation , Academic Press**.

A4 Article in conference publication (refereed)

2005

Afrati, F, Das, G, Gionis, A, Mannila, H, Mielikäinen, T, Tsaparas, P **2005**, 'Mining chains of relations', in **Proceedings**.

Ahonen-Myka, H **2005**, 'Mining all maximal frequent word sequences in a set of sentences', in **Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM 2005, October 31- November 5, 2005, Bremen, Germany**, pp. 255-256.

Aunimo, L **2005**, 'A Question typology and feature set for QA', in **Proceedings of the IJCAI-05: IJCAI 2005**, pp. 53-56.

Aunimo, L, Kuuskoski, R **2005**, 'Reformulations of Finnish questions for question answering', in **Nodalida 2005**, pp. ???.

Autio, I, Borrás, J, Immonen, I, Jalli, P, Ukkonen, E **2005**, 'A voting margin approach for the detection of retinal microaneurysms', in **Proceedings of the Fifth IASTED International Conference on Visualization, imagine, and Image Processing, September 7-9, 2005, Benidorm, Spain**, pp. 511-517.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Bas, P, Hurri, JK 2005, 'Security of DM Quantization Watermarking Schemes: A Practical Study for Digital Images', in **Proceedings of the 4th International Workshop on Digital Watermarking (IWDW 2005)**, pp. 186-200 **Lecture Notes in Computer Science**, vol. 3710.
- Bertino, E, Kamra, A, Terzi, E, Vakali, A 2005, 'Intrusion detection in RBAC-administered databases', in **Proceedings of the 21st Annual Computer Security Applications Conference (ACSAC 2005)**.
- Buntine, W, Löfström, J, Perttu, S, Valtonen, K 2005, 'Topic-specific scoring of documents for relevant retrieval', in **Proceedings of the Workshop on Learning in Web Search (LWS 2005)**, 7-11 August 2005 in Bonn, Germany: 2005, pp. 34-41.
- Buntine, W 2005, 'Static ranking of web pages, and related ideas', in **First workshop on Open Source Web Information Retrieval**, pp. 23-26.
- Buntine, W, Aberer, K, Podnar, I, Rajman, M 2005, *Opportunities from open source search*, Paper presented at **International Conference on Web Intelligence, Piscataway, N.J.**
- Dementiev, R, Kärkkäinen, J, Mehnert, J, Sanders, P 2005, 'Better external memory suffix array construction', in **Proceedings of the Seventh Workshop on Algorithm Engineering and Experiments and the Second Workshop on Analytic Algorithmics and Combinatorics**, pp. 86-97.
- Donato, D, Leonardi, S, Tsaparas, P 2005, 'Stability and Similarity of Link Analysis Ranking Algorithms', in **Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Proceedings**, pp. 717-729 **Lecture Notes in Computer Science**, vol. 3580.
- Doucet, A, Ahonen-Myka, H 2005, 'A method to calculate probability and expected document frequency of discontinued word sequences', in **ELECTRA Workshop on Methodologies and Evaluation of Lexical Cohesion Techniques in Real-World Applications (Beyond Bag of Words): ACM 2005**, pp. 33-40.
- Elomaa, T, Kujala, J, Rousu, J 2005, 'Approximation algorithms for minimizing Empirical Error by Axis-Parallel Hyperplane', in **16th European Conference on Machine Learning (ECML-2005): Lecture Notes in Artificial Intelligence 3720**, pp. 547-555.
- Gionis, A, Hinneburg, A, Papadimitriou, S, Tsaparas, P 2005, 'Dimension-induced clustering', in **Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining (KDD '05)**.
- Gionis, A, Mannila, H, Tsaparas, P 2005, 'Clustering aggregation', in **The 21st International Conference on Data Engineering (ICDE 2005) April 5-8, 2005 National Center of Sciences, Tokyo, Japan**.
- Goethals, B, Muhonen, J, Toivonen, H, Muhonen, J 2005, 'Mining non-derivable association rules', in **Proceedings of the Fifth SIAM International Conference on Data Mining**, pp. 239-249.
- Honkela, T, Hyvärinen, AJ, Vährynen, J 2005, 'Emergence of linguistic features: Independent component analysis of contexts', in **Proc. of the Neural Computation and Psychology Workshop 9**.
- Hyvärinen, AJ, Hoyer, PO, Hurri, J, Gutmann, MU 2005, 'Statistical models of images and early vision', in **Proc. Int. Conf. Adaptive Knowledge Representation and Reasoning**, pp. 1-14.
- Kaski, S 2005, 'From learning metrics towards dependency exploration', in **Proceedings**, pp. 307-314.
- Kaski, S, Myllymäki, P, Kojo, I 2005, 'User models from implicit feedback for proactive information retrieval', in **Workshop 4**, pp. 25-26.
- Klami, A, Kaski, S 2005, 'Non-parametric dependent components', in **ICASSP 2005**, pp. v209-v212.
- Koivisto, M, Sood, K 2005, 'Computational aspects of Bayesian partition models', in **Proceedings: ACM cop. 2005**, pp. 433-440.
- Kontkanen, P, Myllymäki, P 2005, 'A fast normalized maximum likelihood algorithm for multinomial data', in **IJCAI-05**, pp. 1613-1615.
- Kärkkäinen, J 2005, 'Alphabets in generic programming', in **Proceedings of the Prague Stringology Conference, Prague, Czech Republic, August 29-31, 2005: Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University 2005**, pp. 163-173.
- Laasonen, K 2005, 'Route prediction from cellular data', in **Proceedings of the workshop on context awareness for proactive systems CAPS 2005**, pp. 147-157.
- Mazhelis, O, Raento, M, Puuronen, S 2005, 'Evaluating classifiers for mobile-masquerader detection', in **Security and privacy in the age of ubiquitous computing**, pp. 271-283.
- Mielikäinen, TJ, Ravantti, J 2005, 'Sinogram denoising of cryo-electron microscopy images', in **Computational Science and Its Applications – ICCSA 2005: International Conference, Singapore, May 9-12, 2005, Proceedings, Part IV**, pp. 1251-1261 **Lecture Notes in Computer Science**, vol. 3483.
- Miettinen, M, Tuulos, V, Myllymäki, P 2005, 'A testbed for proactive information retrieval', in **Proceedings of the workshop on context awareness for proactive systems CAPS 2005**, pp. 137-146.
- Nikkilä, J, Kaski, S, Roos, C 2005, 'Integration of transcription factor binding and gene expression by associative clustering', in **Proceedings of KRBIO'05, International Symposium on Knowledge Representation in Bioinformatics, Helsinki University of Technology, Espoo, Finland, June 15-17, 2005**, pp. 22-29.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Oulasvirta, A, Raento, M, Tiitta, S 2005, 'Contextcontacts: re-designing smartphone's contact book to support mobile awareness and collaboration', in **Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services, MOBILEHCI'05**, pp. 167-174.
- Perkiö, JP, Tuulos, VH, Buntine, WL, Tirri, H 2005, 'Multi-Faceted Information Retrieval System for Large Scale Email Archives', in **WI '05 Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence**, pp. 557-564.
- Perkiö, J, Buntine, W, Tirri, H 2005, 'A temporally adaptive content-based relevance ranking algorithm', in **SIGIR 2005: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval**, pp. 647-648.
- Pitkänen, E, Rantanen, A, Rousu, JH, Ukkonen, E 2005, 'Finding feasible pathways in metabolic networks', in **Proceedings of the Panhellenic Conference on Informatics**, pp. 123-133 **Lecture Notes in Computer Science, vol. 3746**.
- Puolamäki, K, Salojärvi, J, Savia, E, Simola, J, Kaski, S 2005, 'Combining eye movements and collaborative filtering for proactive information retrieval', in **SIGIR 2005**, pp. 146-153.
- Raento, MP, Petit, R 2005, 'ContextMedia', in **Proceedings of the Workshop on Context Awareness for Proactive Systems (CAPS 2005)**, pp. 181-182 **HIIT Publications, no. 2005-1**.
- Raento, MP, Oulasvirta, A 2005, 'Privacy Management for Social Awareness Applications', in **Proceedings of the Workshop on Context Awareness for Proactive Systems (CAPS 2005)**, pp. 105-114 **HIIT Publications, no. 2005-1**.
- Rantanen, A, Rousu, J, Pitkänen, E, Maaheimo, H, Ukkonen, E 2005, 'Flow analysis of metabolite fragments for flux estimation', in **Third international workshop on computational methods in systems biology: 2005**, pp. 242-255.
- Rantanen, A, Mielikäinen, T, Rousu, JH, Ukkonen, E 2005, 'Planning isotopome measurements for estimation of metabolic fluxes', in **German Conference on Bioinformatics**.
- Rastas, P, Koivisto, M, Mannila, H, Ukkonen, E 2005, 'A hidden Markov technique for haplotype reconstruction', in **Algorithms in Bioinformatics (WABI-2005)**, pp. 140-151 **Lecture Notes in Computer Science, vol. 3692**.
- Roos, T, Myllymäki, P, Tirri, H 2005, 'On the behavior of MDL denoising', in **AISTATS 2005**, pp. 309-316.
- Rousu, J, Saunders, C, Szedmak, S, Shawe-Taylor, J 2005, 'Learning Hierarchical Multi-Category Text Classification Models', in **22nd International Conference on Machine Learning (ICML-2005)**, pp. 745-752.
- Salojärvi, J, Puolamäki, K, Kaski, S 2005, 'Expectation maximization algorithms for conditional likelihoods', in **Proceedings: ACM cop. 2005**, pp. 753-760.
- Savia, E, Puolamäki, K, Sinkkonen, J, Kaski, S 2005, 'Two-way latent grouping model for user preference prediction', in **Uncertainty in artificial intelligence**, pp. 518-525.
- Shimizu, S, Hyvärinen, AJ, Kano, Y 2005, 'A generalized least squares approach to blind separation of sources which have variance dependencies', in **Proc. IEEE Workshop on Statistical Signal Processing (SSP2005)**.
- Shimizu, S, Hyvärinen, AJ, Kano, Y, Hoyer, PO 2005, 'Discovery of non-gaussian linear causal models using ICA', in **Proc. 21st Conference on Uncertainty in Artificial Intelligence (UAI2005)**, pp. 526-533.
- Tamm, H, Nykänen, M, Ukkonen, E 2005, 'Size reduction of multitape automata', in **Implementation and application of automata**, pp. 329-330.
- Toivonen, H, Hyvönen, S, Sevon, P 2005, 'Combining phenotypic and genotypic data to discover multiple disease genes', in **Proceedings of KRBO'05, International Symposium on Knowledge Representation in Bioinformatics, Helsinki University of Technology, Espoo, Finland, June 15-17, 2005**, pp. 7-14.
- Tuulos, V, Silander, T 2005, 'Language pragmatics, contexts and a search engine', in **Proceedings of AKRR'05, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning, Helsinki University of Technology, Espoo, Finland, June 15-17, 2005**, pp. 114-120.
- Ukkonen, A, Fortelius, M, Mannila, H 2005, 'Finding partial orders from unordered 0-1 data', in **KDD-2005**, pp. 285-293.
- Venna, J, Kaski, S 2005, 'Local multidimensional scaling with controlled tradeoff between trustworthiness and continuity', in **Proceedings**, pp. 695-702.
- Venna, J, Kaski, S 2005, 'Visualized atlas of a gene expression databank', in **Proceedings of KRBO'05, International Symposium on Knowledge Representation in Bioinformatics, Helsinki University of Technology, Espoo, Finland, June 15-17, 2005**, pp. 30-36.
- Yangarber, R, Jokipii, L, Rauramo, A, Huttunen, SM 2005, 'Information Extraction from Epidemiological Reports', in **In Proceedings Human Language Technology Conference/ Conference on Empirical Methods in Natural Language Processing: HLT/EMNLP-2005**.
- Yangarber, R, Jokipii, L 2005, 'Redundancy-based Correction of Automatically Extracted Facts', in **Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing**.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Yangarber, R, Bogatyrev, K **2005**, 'Use of Deep Syntax Parsing in Cross-Language Information Extraction', in **Proceedings of the Workshop on Intelligent Linguistic Technologies: International Conference on Machine Learning: Models, Technologies and Applications MLMTA-2005**, pp. 18-24.

Yangarber, R **2005**, 'Mining the Semantics of Text via Counter-Training', in **Progress in Artificial Intelligence: 12th Portuguese Conference on Artificial Intelligence, EPIA 2005, Covilha, Portugal, December 5-8, 2005. Proceedings**, pp. 647-657 **Lecture Notes in Computer Science**, vol. 2005/3808.

2006

Agrawal, R, Terzi, E **2006**, 'On honesty in sovereign information sharing', in **Advances in database technology, EDBT 2006**.

Agrawal, R, Rantau, R, Terzi, E **2006**, 'Context-sensitive ranking', in **SIGMOD 2006**.

Aunimo, L, Kuuskoski, R **2006**, 'Question Answering Experiments for Finnish and French', in **Accessing Multilingual Information Repositories**, pp. 477-487 **Lecture Notes in Computer Science**, vol. 4022.

Autio, I, Lindgren, JT **2006**, 'Online learning of discriminative patterns from unlimited sequences of candidates', in **Proceedings of the 18th International Conference on Pattern Recognition**, pp. 437-440.

Bingham, E, Gionis, A, Haiminen, N, Hiisilä, H, Mannila, H, Terzi, E **2006**, 'Segmentation and dimensionality reduction', in **Proceedings of the Sixth SIAM International Conference on Data Mining**, pp. 372-383.

Buntine, W, Taylor, MP, Lagunas, F **2006**, 'Standards for Open Source Information Retrieval', in **OSIR 2006**, pp. 68-72.

Doucet, A, Ahonen-Myka, H **2006**, 'Fast extraction of discontinuous sequences in text: a new approach based on maximal frequent sequences', in **Proceedings of 5th Slovenian and 1st International Language Technologies Conference 2006**, pp. 186-191.

Elomaa, T, Kujala, J, Rousu, J **2006**, 'Practical approximation of optimal multivariate discretization', in **Foundations of intelligent systems**, pp. 612-621.

Fokker, J, Buntine, W, Pouwelse, J **2006**, 'Tagging in Peer-to-Peer Wikipedia: a method to induce cooperation', in **OSIR 2006**.

Gionis, A, Mannila, H, Mielikäinen, T, Tsaparas, P **2006**, 'Assessing data mining results via swap randomization', in **KDD-2006**, pp. 167-176.

Gionis, A, Mannila, H, Puolamäki, K, Ukkonen, A **2006**, 'Algorithms for discovering bucket orders from data', in **KDD-2006**, pp. 561-566.

Heinonen, M, Rantanen, A, Mielikäinen, T, Pitkänen, E, Kokkonen, J, Rousu, J **2006**, 'Ab initio prediction of molecular fragments from tandem mass spectrometry data', in **German Conference on Bioinformatics, GCB 2006, Tübingen, Germany, September 19-22, 2006**, pp. 40-53.

Hermersdorf, M, Nyholm, H, Salminen, J, Tirri, H, Perkiö, J, Tuulos, V **2006**, 'Sensing in rich Bluetooth environments', in **WSW'2006: 2006**, pp. 27-31.

Hoyer, PO, Shimizu, S, Kerminen, A **2006**, 'Estimation of linear, non-gaussian causal models in the presence of confounding latent variables', in **Proceedings of the third European workshop on Probabilistic Graphical Models**, pp. 155-162.

Hoyer, PO, Shimizu, S, Hyvärinen, A, Kerminen, A, Kano, Y **2006**, 'New permutation algorithms for causal discovery using ICA', in **ICA 2006: ICA 2006**, pp. 115-122.

Hurri, J **2006**, 'Learning cue-invariant visual responses', in **Advances in neural information processing systems 18**, pp. 539-546.

Hyvärinen, AJ, Perkiö, JP **2006**, 'Learning to segment any random vector', in **IEEE International Joint Conference on Neural Networks 2006: Proceedings**, pp. 4167-4172.

Hyvärinen, AJ, Köster, U **2006**, 'FastISA: A fast fixed-point algorithm for independent subspace analysis', in **In Proc. European Symposium on Artificial Neural Networks**.

Hyvärinen, A, Shimizu, S **2006**, 'A quasi-stochastic gradient algorithm for variance-dependent component analysis', in **Proc. Int. Conf. on Artificial Neural Networks**, pp. 211-220.

Kaban, A, Bingham, E **2006**, 'ICA-based binary feature construction', in **ICA 2006: ICA 2006**, pp. 140-148.

Koivisto, M **2006**, 'Parent assignment is hard for the MDL, AIC, and NML costs', in **Learning theory**, pp. 289-303.

Koivisto, M **2006**, 'An $O^*(2^n)$ algorithm for graph coloring and other partitioning problems via inclusion-exclusion', in **47th Annual IEEE Symposium on Foundations of Computer Science**, pp. 583-590.

Koivisto, M **2006**, 'Advances in exact Bayesian structure discovery in Bayesian networks', in **Uncertainty in artificial intelligence**, pp. 241-248.

Kollin, J, Koivisto, M **2006**, 'Bayesian learning with mixtures of trees', in **Machine learning**, pp. 294-305.

Kääriäinen, M **2006**, 'Semi-supervised model selection based on cross-validation', in **2006 International Joint Conference on Neural Networks, Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada, July 16-21, 2006**, pp. 1894-1899.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Lahtinen, J, Myllymäki, P, Ryyänen, O **2006**, 'P-Course: medical applications of Bayesian classification with informative priors', in **ECAI'2006: Università di Trento 2006**, pp. 38-40.
- Laur, S, Lipmaa, H, Mielikäinen, T **2006**, 'Cryptographically private support vector machines', in **KDD-2006**, pp. 618-624.
- Lemström, K, Pienimäki, A **2006**, 'Approaches for content-based retrieval of symbolically encoded polyphonic music', in **9th International Conference on Music Perception and Cognition, August 22-26, Bologna, Italy**.
- Mielikäinen, T, Terzi, E, Tsaparas, P **2006**, 'Aggregating time partitions', in **KDD-2006**, pp. 347-356.
- Mielikäinen, T, Panov, P, Dzeroski, S **2006**, 'Itemset support queries using frequent itemsets and their condensed representations', in **Discovery science**, pp. 161-172.
- Miettinen, P, Mielikäinen, T, Gionis, A, Das, G, Mannila, H **2006**, 'The discrete basis problem', in **Knowledge discovery in databases**, pp. 335-346.
- Muhonen, J, Toivonen, H **2006**, 'Closed non-derivable itemsets', in **Knowledge discovery in databases**, pp. 601-608.
- Mäkinen, V, Navarro, G **2006**, 'Dynamic entropy-compressed sequences and full-text indexes', in **Combinatorial pattern matching**, pp. 306-317.
- Mäkinen, V, Navarro, G **2006**, 'Position-restricted substring searching', in **LATIN 2006**, pp. 703-714.
- Parikka, P, Pitkänen, E, Rantanen, A, Åkerlund, A, Ukkonen, E **2006**, 'Pathway assistant: a web portal for metabolic modelling', in **NETTAB 2006: Sixth International Workshop on Distributed Applications, Web Services, Tools and GRID Infrastructures for Bioinformatics, July, 10th-13th, 2006, Santa Margherita di Pula, Cagliari, Italy : 2006**, pp. 90-96.
- Perkiö, J, Tuulos, V, Hermersdorf, M, Nyholm, H, Salminen, J, Tirri, H **2006**, 'Utilizing rich Bluetooth environments for identity prediction and exploring social networks as techniques for ubiquitous computing', **Paper presented at IEEE/WIC/ACM International Conference on Web Intelligence, Los Alamitos, CA..**
- Roos, TT, Heikkilä, T, Myllymäki, P **2006**, 'A Compression-Based Method for Stemmatic Analysis', in **Proceeding of the 2006 conference on ECAI 2006**, pp. 805-806.
- Roos, T, Myllymäki, P, Grunwald, P, Tirri, H **2006**, 'Generalization to unseen cases', in **Advances in neural information processing systems 18**, pp. 1129-1136.
- Shimizu, S, Hyvärinen, A, Hoyer, PO, Kerminen, A, Kano, Y **2006**, 'Testing significance of mixing and demixing coefficients in ICA', in **ICA 2006: ICA 2006**, pp. 901-908.
- Silander, T, Myllymäki, P **2006**, 'A simple approach for finding the globally optimal Bayesian network structure', in **Uncertainty in artificial intelligence**, pp. 445-452.
- Tamm, H, Nykänen, M, Ukkonen, E **2006**, 'Size reduction of multitape automata', in **Proceedings of the Tenth International Conference on Implementation and Application of Automata (CIAA 2005)**, pp. 307-318 **Lecture Notes in Computer Science, vol. 3845**.
- Tatti, N, Mielikäinen, T, Gionis, A, Mannila, H **2006**, 'What is the dimension of your binary data?', in **Sixth IEEE International Conference on Data Mining, ICDM 2006, 18-22 December 2006, Hong Kong, China**.
- Terzi, E, Tsaparas, P **2006**, 'Efficient algorithms for sequence segmentation', in **Proceedings of the Sixth SIAM International Conference on Data Mining**.
- Vallin, A, Magnini, B, Giampiccolo, D, Aunimo, L, Ayache, C, Osenova, P, Peñas, A, de rijke, M, Sacaleanu, B, Santos, D, Sutcliffe, R **2006**, 'Overview of the CLEF 2005 Multilingual Question Answering Track', in **Accessing Multilingual Information Repositories 6th Workshop of the Cross-Language Evaluation Forum (CLEF 2005), Lecture Notes in Computer Science, vol. 4022**.
- Yangerber, R **2006**, 'Verification of Facts across Document Boundaries', in **Proceedings of International Workshop on Intelligent Information Access (IIA-2006)**.
- Zhou, L, Buntine, W **2006**, 'Preface to the "web search technology - from search to semantic search" workshop (SWET'06)', in **Proceedings of The Semantic Web ASWC 2006 Workshops: Jilin University Press 2006**, pp. 221-222.
- 2007**
- Astikainen, K, Holm, L, Pitkänen, E, Szedmak, S, Rousu, J **2007**, 'Towards structured output prediction of enzyme function (extended abstract)', in **Proceedings of the International Workshop on Machine Learning in Systems Biology 2007, MLSB'07, September 24-25, 2007, Evry, France: Université d'Evry-Val d'Essonne 2007**, pp. 27-32.
- Bender, MA, Fekete, SP, Kröller, A, Liberatore, V, Mitchell, JSB, Polishchuk, V, Suomela, J **2007**, 'The minimum-backlog problem', in **MACIS 2007: 2007**.
- Bhattacharjee, M, Botting, CH, Sillanpää, MJ **2007**, 'Biomarker identification using Bayesian variable selection based on marker-expression-proteomics data', in **Proceeding of the CAMDA 2007 workshop: December 13-14, 2007 at Valencia, Spain**.
- Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M **2007**, 'Fourier meets Möbius: fast subset convolution', in **STOC'07**, pp. 67-74.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Cheung, A, Grandison, T, Johnson, C, Schönauer, S 2007, 'Infinity: a generic platform for application development and information sharing on mobile devices', in **MobiDE '07: Proceedings of the 6th ACM international workshop on Data engineering for wireless and mobile access**, pp. 25-32.
- De Raedt, L, Kimmig, A, Toivonen, H 2007, 'ProbLog: a probabilistic prolog and its application in link discovery', in **IJCAI-07, proceedings of the Twentieth International Joint Conference on Artificial Intelligence**, pp. 2468-2473.
- De Raedt, L, Kimmig, A, Toivonen, H 2007, 'Probabilistic Explanation Based Learning', in **Machine learning**, pp. 176-187.
- Doucet, A, Lehtonen, M 2007, 'Unsupervised classification of text-centric XML document collections', in **Comparative evaluation of XML information retrieval systems**, pp. 497-509.
- Doucet, A, Lehtonen, M 2007, 'EXTIRP: baseline retrieval from Wikipedia', in **Comparative evaluation of XML information retrieval systems**, pp. 115-120.
- Floréen, P, Kaski, P, Suomela, J 2007, 'A distributed approximation scheme for sleep scheduling in sensor networks', in **Sensor, Mesh and Ad Hoc Communications and Networks, 2007, SECON '07: IEEE 2007**, pp. 152-161.
- Hintsanen, P 2007, 'The most reliable subgraph problem', in **Knowledge discovery in databases**, pp. 471-478.
- Hyvärinen, A 2007, 'Unsupervised learning of an embodied representation for action selection', in **Proceedings of EuroCogSci07**, pp. 658-665.
- Hyvärinen, A 2007, 'Behavioural priors: learning to search efficiently in action planning', in **Proceedings of EuroCogSci07**, pp. 324-328.
- Hyvönen, S, Junttila, E, Salmenkivi, M 2007, 'Pre-processing large spatial data sets with Bayesian methods', in **Knowledge discovery in databases: PKDD 2007 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, September 17-21, 2007. Proceedings**, pp. 498-505 **Lecture Notes in Computer Science**, vol. 4702/2007.
- Junttila, T, Kaski, P 2007, 'Engineering an efficient canonical labeling tool for large and sparse graphs', in **Proceedings of the ninth Workshop on Algorithm Engineering and Experiments and the fourth Workshop on Analytic Algorithmics and Combinatorics**, pp. 135-149.
- Junttila, E, Salmenkivi, M 2007, 'Modeling missing data with Markov random fields in large data sets', in **IADIS European Conference Data Mining: IADIS Press cop. 2007**, pp. 73-80.
- Kaski, P, Penttinen, A, Suomela, J 2007, 'Coordinating concurrent transmissions: A constant-factor approximation of maximum-weight independent set in local conflict graphs', in **Ad-Hoc, Mobile, and Wireless Networks: 6th International Conference, ADHOC-NOW 2007, Morelia, Mexico, September 24-26, 2007. Proceedings**, pp. 74-86 **Lecture Notes in Computer Science**, vol. 4686.
- Koch, K, Schönauer, S, Jansen, I, van den Bussche, J, Burzykowski, T 2007, 'Finding clusters of positive and negative coregulated genes in gene expression data', in **Proceedings of the IEEE 7th International Conference on Bioinformatics and Bioengineering: IEEE Press 2007**, pp. 93-99.
- Kontkanen, P, Myllymäki, P 2007, 'MDL histogram density estimation', in **AISTATS 2007: The Society for Artificial Intelligence and Statistics 2007**.
- Krozel, J, Mitchell, JSB, Polishchuk, V, Prete, J 2007, 'Capacity estimation for airspaces with convective weather constraints', in **A collection of technical papers**.
- Kärkkäinen, J, Na, JC 2007, 'Faster Filters for approximate string matching', in **Proceedings of the Workshop on Algorithm Engineering and Experiments, ALENEX 2007, New Orleans, Louisiana, USA, January 6, 2007 2007**, pp. 84-90.
- Köster, U, Hyvärinen, A 2007, 'A two-layer ICA-like model estimated by score matching', in **ICANN '07**, pp. 798-807.
- Lehtonen, M 2007, 'Vocabulary-independent methods for XML information retrieval', in **Proceedings of AMICT'2006**, pp. 53-61.
- Leino, A 2007, 'Construction Grammar in onomastics: the case of Finnish hydronyms', in **Atti del XXII Congresso Internazionale di Scienze Onomastiche**, pp. 297-309.
- Lindgren, JT, Hurri, J, Hyvärinen, A 2007, 'The statistical properties of local log-contrast in natural images', in **SCIA 2007**, pp. 354-363.
- Lindgren, JT, Hyvärinen, A 2007, 'Emergence of conjunctive visual features by quadratic independent component analysis', in **Advances in neural information processing systems 19**, pp. 897-904.
- Mitchell, JSB, Polishchuk, V 2007, 'Thick non-crossing paths and minimum-cost flows in polygonal domains', in **Proceedings of the Twenty-third Annual ACM Symposium on Computational Geometry**, pp. 56-65.
- Mononen, T, Myllymäki, P 2007, 'Fast NML computation for naive Bayes models', in **Discovery science**, pp. 151-160.
- Mäkinen, V, Navarro, G 2007, 'Implicit compression boosting with applications to self-indexing', in **String Processing and Information Retrieval**, pp. 229-241 **Lecture notes in computer science**, vol. 4726.
- Nyrkkö, S, Carlson, L, Keijola, M, Ahonen-Myka, H, Niemi, J, Piitulainen, JO, Viitanen, S, Meri, M, Seitsonen, L, Mannonen, P, Juvonen, J, TEKES 4M language technology project 2007, 'Ontology-based Knowledge in Interactive Maintenance Guide', in **Proceedings of the 40th Hawaii International Conference on System Sciences**, pp. 47.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Pizzi, C, Rastas, P, Ukkonen, E **2007**, 'Fast search algorithms for position specific scoring matrices', in **Bioinformatics Research and Development**, pp. 239-250 **Lecture Notes in Computer Science**, vol. 4414.

Rastas, P, Ukkonen, E **2007**, 'Haplotype inference via hierarchical genotype parsing', in **Algorithms in Bioinformatics (WABI 2007)**, pp. 85-97 **Lecture Notes in Computer Science**, vol. 4645.

Silander, T, Kontkanen, P, Myllymäki, P **2007**, 'On sensitivity of the MAP Bayesian network structure to the equivalent sample size parameter', in **Uncertainty in artificial intelligence: Proceedings of the Twenty-Third Conference Conference on Uncertainty in Artificial Intelligence (2007)**, pp. 360-367.

Ukkonen, E **2007**, 'Structural analysis of gapped motifs of a string', in **Proceedings of the 32nd International Symposium on Mathematical Foundations of Computer Science (MFCS 2007)**, pp. 681-690 **Lecture Notes in Computer Science**, vol. 4708.

Välimäki, N, Mäkinen, V **2007**, 'Space-efficient algorithms for document retrieval', in **Combinatorial pattern matching**, pp. 205-215.

Waldmann, P, Hallander, J, Hoti, F, Sillanpää, MJ **2007**, 'The effect of dominance geneitive parameters on rank of breeding value predictions', **Paper presented at Nordic-Baltic Biometric Conference**, Denmark. 06. - 08. June, 2007..

Wettig, H, Kontkanen, P, Myllymäki, P **2007**, 'Calculating the normalized maximum likelihood distribution for Bayesian forests', in **Proceedings IADIS International Conference on Intelligent Systems and Agents (ISA 2007)**, July 3-5, 2007, Lisbon, Portugal: IADIS Press cop. **2007**.

Yangerber, R, Best, C, von Etter, P, Fuat, F, Horby, D, Steinberger, R **2007**, *Combining Information about Epidemic Threats from Multiple Sources*, **Paper presented at In Proceedings Multi-source, Multilingual Information Extraction and Summarization at RANLP-2007, Borovets, Bulgaria.**

Yangerber, R, Steinberger, R, Best, C, von Etter, P, Fuat, F, Horby, D **2007**, *Combining Information Retrieval and Information Extraction for Medical Intelligence*, **Paper presented at Mining Massive Data Sets for Security, NATO Advanced Study Institute , Gazzada, Italy..**

Yu, H, Bertsekas, DP **2007**, 'Q-learning algorithms for optimal stopping based on least squares', in **European Control Conference 2007 cop. 2007**, pp. 2368-2375.

2008

Ahonen, TE, Lemström, K **2008**, 'Identifying cover songs using normalized compression distance', in **International Workshop on Machine Learning and Music, MML'08, July 5-12, 2008, Helsinki, Finland.**

Ala-Härkönen, V, Brunberg, J, Lemström, K, Mikkilä, N **2008**, 'JMIR serves Mozart', in **CMMR 2008**, pp. 148-157.

Arkin, EM, Mitchell, JSB, Polishchuk, V **2008**, 'Maximum thick paths in static and dynamic environments', in **Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry (SCG-08)** , June 9-11, 2008, College Park, Maryland, USA, pp. 20-27.

Bauer, A, Hoti, F, Korff, MV, Pillen, K, Léon, J, Sillanpää, MJ **2008**, *Bayesian multi-environmental advanced backcross-QTL analysis in Spring Barley*, **Paper presented at International Barley Genetics Symposium, Aleksandria, Egypt. 05. - 10. April, 2008..**

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M **2008**, 'Computing the Tutte polynomial in vertex-exponential time', in **49th Annual IEEE Symposium on Foundations of Computer Science**, pp. 677-686.

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M **2008**, 'Trimmed moebius inversion and graphs of bounded degree', in **Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science**, pp. 85-96.

Doucet, A, Lehtonen, M **2008**, 'Let's phrase it: INEX topics need keyphrases', in **Proceedings of the SIGIR 2008 Workshop on Focused Retrieval, 24 July, 2008, Singapore**, pp. 9-14.

Efrat, A, Fekete, SP, Gaddehosur, PR, Mitchell, JSB, Polishchuk, V, Suomela, J **2008**, 'Improved approximation algorithms for relay placement', in **Algorithms - ESA 2008: 16th Annual European Symposium. Karlsruhe, Germany, September 15-17, 2008. Proceedings**, pp. 356-367 **Lecture Notes in Computer Science**, vol. 5193.

Fischer, J, Mäkinen, V, Välimäki, N **2008**, 'Space-efficient string mining under frequency constraints', in **Eighth IEEE International Conference on Data Mining, ICDM 2008, 15-19 December, 2008, Pisa, Italy: IEEE Computer Society 2008**, pp. 193-202.

Fischer, J, Mäkinen, V, Navarro, G **2008**, 'An(other) entropy-bounded compressed suffix tree', in **Combinatorial pattern matching**, pp. 152-165.

Floreen, P, Hassinen, M, Kaski, P, Suomela, J **2008**, 'Tight local approximation results for max-min linear programs', in **Algorithmic Aspects of Wireless Sensor Networks: Fourth International Workshop. ALGOSENSORS 2008. Reykjavik, Iceland, July 2008. Revised Selected Papers**, pp. 2-17 **Lecture Notes in Computer Science**, vol. 5389.

Floreen, P, Kaski, P, Musto, T, Suomela, J **2008**, 'Local approximation algorithms for scheduling problems in sensor networks', in **Algorithmic Aspects of Wireless Sensor Networks: Third International Workshop, ALGOSENSORS 2007. Wrocław, Poland, July 14, 2007. Revised Selected Papers**, pp. 99-113 **Lecture Notes in Computer Science**, vol. 4837.

Floreen, P, Kaski, P, Musto, T, Suomela, J **2008**, 'Approximating max-min linear programs with local algorithms', in **proceedings of the 2008 IEEE International Parallel and Distributed Processing Symposium, April 14-18, 2008, Miami, Florida: IEEE 2008.**



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Gallo, A, Miettinen, P, Mannila, H **2008**, 'Finding subgroups having several descriptions: algorithms for redescription mining', in **Proceedings of the Eighth SIAM International Conference on Data Mining, April 24-26, 2008, Atlanta, Georgia**, pp. 334-345.
- Garriga, GC, Junttila, E, Mannila, H **2008**, 'Banded structure in binary matrices', in **KDD 08**, pp. 292-300.
- Grandoni, F, Gupta, A, Leonardi, S, Miettinen, P, Sankowski, P, Singh, M **2008**, 'Set covering with our eyes closed', in **49th Annual IEEE Symposium on Foundations of Computer Science**, pp. 347-356.
- Gutmann, M, Hyvärinen, AJ, Aihara, K **2008**, 'Learning encoding and decoding filters for data representation with a spiking neuron', in **Proc. Int. Joint Conference on Neural Networks (IJCNN2008)**.
- Hassinen, M, Polishchuk, V, Suomela, J **2008**, 'Local 3-approximation algorithms for weighted dominating set and vertex cover in quasi unit-disk graphs', in **Adjunct workshop proceedings: 2008**, pp. V.9-V.12.
- Hoyer, PO, Hyvärinen, A, Scheines, R, Spirtes, P, Ramsey, J, Lacerda, G, Shimizu, S **2008**, 'Causal discovery of linear acyclic models with arbitrary distributions', in **Uncertainty in Artificial Intelligence**, pp. 282-289.
- Hyvärinen, A, Shimizu, S, Hoyer, PO **2008**, 'Causal modelling combining instantaneous and lagged effects: an identifiable model based on non-Gaussianity', in **Proceedings of the 25th International Conference on Machine Learning**, pp. 424-431.
- Hyvärinen, A **2008**, 'Estimation theory and information geometry based on denoising', in **Proceedings of the First Workshop on Information Theoretic Methods in Science and Engineering, August 18-20, 2008, Tampere, Finland**.
- Hyvönen, S, Miettinen, P, Terzi, E **2008**, 'Interpretable nonnegative matrix decompositions', in **KDD 08**, pp. 345-353.
- Hämäläinen, W, Nykänen, M **2008**, 'Efficient discovery of statistically significant association rules', in **Eighth IEEE International Conference on Data Mining, ICDM 2008, 15-19, December 2008, Pisa, Italy: editors Fosca Giannotti ... [et al.]**, pp. 203-212.
- Karvonen, M, Lemström, K **2008**, 'Using mathematical morphology for geometric music information retrieval', in **International Workshop on Machine Learning and Music, MML'08, July 5-12, 2008, Helsinki, Finland**.
- Kim, J, Mitchell, JSB, Polishchuk, V, Vihavainen, A **2008**, 'Routing a maximum number of disks through a scene of moving obstacles', in **Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry (SCG-08), June 9-11, 2008, College Park, Maryland, USA**, pp. 230-231.
- Krasnoshchekov, D, Polishchuk, V **2008**, 'Robust curve reconstruction with k-order -shapes', in **Proceedings of the IEEE International Conference on Shape Modeling and Applications 4-6 June, 2008, Stony Brook, New York, USA: IEEE 2008**, pp. 279-280.
- Kärkkäinen, J, Rantala, T **2008**, 'Engineering radix sort for strings', in **String Processing and Information Retrieval: 15th International Symposium, SPIRE 2008, Melbourne, Australia, November 10-12, 2008. Proceedings**, pp. 3-14 **Lecture Notes in Computer Science**.
- Lacerda, G, Spirtes, P, Ramsey, J, Hoyer, PO **2008**, 'Discovering cyclic causal models by Independent Components Analysis', in **Uncertainty in Artificial Intelligence**.
- Lehtonen, M, Doucet, A **2008**, 'XML-aided phrase indexing for hypertext documents', in **SIGIR 2008**, pp. 843-844.
- Lehtonen, M, Doucet, A **2008**, 'Phrase detection in the Wikipedia', in **Focused access to XML documents**, pp. 115-121.
- Leino, A **2008**, 'On the Semantic Basis of Heraldic Propaganda, or What do Arms Mean, and How?', in **Proceedings of the XXVII International Congress of Genealogical and Heraldic Sciences**, pp. 407-418.
- Lemström, K, Mikkilä, N, Mäkinen, V **2008**, 'Fast index based filters for music retrieval', in **ISMIR 2008**, pp. 677-682.
- Lindgren, JT, Hurri, J, Hyvärinen, A **2008**, 'Unsupervised learning of dependencies between local luminance and contrast in natural images', in **Proceedings of the International Joint Conference on Neural Networks, IJCNN2008, June 1-6, 2008, Hong Kong, China**, pp. 356-362.
- Lindgren, JT, Hyvärinen, A **2008**, 'On the learning of nonlinear visual features from natural images by optimizing response energies', in **Proceedings of the International Joint Conference on Neural Networks, IJCNN2008, June 1-6, 2008, Hong Kong, China**, pp. 1026-1033.
- Mavroudis, D, Bingham, E **2008**, 'Enhancing the stability of spectral ordering with sparsification and partial supervision: application to paleontological data', in **Eighth IEEE International Conference on Data Mining, ICDM 2008, 15-19, December 2008, Pisa, Italy: IEEE Computer Society 2008**.
- Mononen, T, Myllymäki, P **2008**, 'Computing the multinomial stochastic complexity in sub-linear time', in **Proceedings of the 4th European Workshop on Probabilistic Graphical Models (PGM-08), September 17-19, 2008, Hirtshals, Denmark: The authors 2008**, pp. 209-216.
- Mononen, T, Myllymäki, P **2008**, 'On recurrence formulas for computing the stochastic complexity', in **Proceedings of the 2008 International Symposium on Information Theory and its Applications, Auckland, New Zealand, 7th-10th December 2008: IEEE 2008**, pp. 281-286.
- Mononen, T, Myllymäki, P **2008**, 'Computing the NML for Bayesian forests via matrices and generating polynomials', in **Proceedings of the 2008 IEEE Information Theory Workshop: IEEE 2008**, pp. 276-280.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Mononen, T., Myllymäki, P 2008, 'On the multinomial stochastic complexity and its connection to the birthday problem', in **ITSL 2008: CSREA Press 2008**, pp. 17-22.
- Myllymäki, P 2008, 'Recent advances in computing the NML for discrete Bayesian networks', in **Proceedings of the First Workshop on Information Theoretic Methods in Science and Engineering, August 18-20, 2008, Tampere, Finland**.
- Mäkinen, V, Navarro, G 2008, 'On self-indexing images: image compression with added value', in **DDC 2008**, pp. 422-431.
- O'Hara, R, Sillanpää, MJ 2008, 'A latent variable approach to multivariate quantitative trait loci', in **Proceeding of the International Biometric Conference: Dublin 13-18 July, 2008**.
- Ojala, M, Vuokko, N, Kallio, A, Haiminen, N, Mannila, H 2008, 'Randomization of real-valued matrices for assessing the significance of data mining results', in **Proceedings of the Eighth SIAM International Conference on Data Mining, April 24-26, 2008, Atlanta, Georgia**, pp. 494-505.
- Perkiö, J, Myllymäki, P, Tuulos, VH, Boda, P 2008, 'Magrathea: a mobile agent- and sensing platform', in **Proceedings of the 2008 International Conference on Wireless Networks: CSREA Press 2008**.
- Pernestål, A, Wettig, H, Silander, T, Nyberg, M, Myllymäki, P 2008, 'A Bayesian approach to learning in fault isolation', in **Proceedings of the 19th International Workshop on Principles of Diagnosis (DX-08), September 22-24, 2008, Blue Mountains, NSW, Australia: 2008**, pp. 143-150.
- Pitkänen, E, Rantanen, A, Rousu, J, Ukkonen, E 2008, 'A computational method for reconstructing gapless metabolic networks', in **Bioinformatics Research and Development**, pp. 288-302.
- Polishchuk, V, Suomela, J 2008, 'Optimal backlog in the plane', in **Algorithmic Aspects of Wireless Sensor Networks: Fourth International Workshop. ALGOSENSORS 2008. Reykjavik, Iceland, July 2008. Revised Selected Papers**, pp. 141-150 **Lecture Notes in Computer Science**, vol. 5389.
- Rastas, P, Kollin, J, Koivisto, M 2008, 'Fast Bayesian haplotype inference via Context Tree Weighting', in **Proceedings of the 8th International Workshop, WABI 2008, Karlsruhe, Germany, September 15-19, 2008**, pp. 259-270.
- Rizo, D, Lemström, K, Inesta, JM 2008, 'Tree structured and combined methods for comparing metered polyphonic music', in **CMMR 2008**, pp. 263-278.
- Roos, T, Silander, T, Kontkanen, P, Myllymäki, P 2008, 'Bayesian network structure learning using factorized NML universal models', in **Proceedings of the 2008 Information Theory and Applications Workshop, San Diego, California, USA, January-February 2008: IEEE 2008**, pp. 272-276.
- Roos, T, Rissanen, J 2008, 'On sequentially normalized maximum likelihood models', in **Proceedings of the First Workshop on Information Theoretic Methods in Science and Engineering, August 18-20, 2008, Tampere, Finland**.
- Roos, T 2008, 'Monte Carlo estimation of minimax regret with an application to MDL model selection', in **Proceedings of the 2008 IEEE Information Theory Workshop: IEEE 2008**, pp. 284-288.
- Shimizu, S, Hyvärinen, A 2008, 'Discovery of linear non-gaussian acyclic models in the presence of latent classes', in **Proceedings of 14th International Conference on Neural Information Processing, ICONIP 2007, Kitakyushu, Japan, November 13-16, 2007**, pp. 752-761.
- Silander, T, Roos, T, Kontkanen, P, Myllymäki, P 2008, 'Factorized normalized maximum likelihood criterion for learning Bayesian network structures', in **Proceedings of the 4th European Workshop on Probabilistic Graphical Models (PGM-08), September 17-19, 2008, Hirtshals, Denmark: The authors 2008**, pp. 257-272.
- Sirén, J, Välimäki, N, Mäkinen, V, Navarro, G 2008, 'Run-length compressed indexes are superior for highly repetitive sequence collections', in **String Processing and Information Retrieval: 15th International Symposium, SPIRE 2008**, pp. 164-175 **Lecture Notes in Computer Science**, no. 5280.
- Steinberger, R, Fuat, F, van der Groot, E, Best, C, von Etter, P, Yangarber, R 2008, 'Text Mining from the Web for Medical Intelligence', in **Mining Massive Data Sets for Security**, pp. 295-310.
- Wong, D, Bingham, E, Hyvönen, S 2008, 'Recommendation of multimedia items by link analysis and collaborative filtering', in **Proceedings of the Second International Conference on Weblogs and Social Media, March 30 - April 2, 2008, Seattle, Washington, USA: AAAI Press 2008**.
- Yangarber, R, von Etter, P, Steinberger, R 2008, 'Content Collection and Analysis in the Domain of Epidemiology', in **In Proceedings of DrMED-2008: International Workshop on Describing Medical Web Resources, at MIE-2008: the 21st International Congress of the European Federation for Medical Informatics**.
- 2009**
- Ahonen, TE 2009, 'Measuring harmonic similarity using PPM-based compression distance', in **Proceedings of the Workshop on Exploring Musical Information Spaces, ECDL 2009, Corfu, Greece, October 2009. - cop. 2009**, pp. 52-55.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Atkinson, M, Piskorski, J, Tanev, H, van der Goot, E, Yangarber, R, Zavarella, V 2009, 'Automated Event Extraction in the Domain of Border Security', in **User Centric Media: First International Conference, UCMedia 2009, Venice, Italy, December 9-11, 2009, Revised Selected Papers**, pp. 321-326 **Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering**, vol. 40.
- Elovaara, E, Myllymäki, P 2009, 'MDL-based attribute models in naive Bayes classification', in **Proceedings of the Second Workshop on Information Theoretic Methods in Science and Engineering, August 17-19, 2009, Tampere, Finland**.
- Floreen, P, Kaasinen, J, Kaski, P, Suomela, J 2009, 'An optimal local approximation algorithm for max-min linear programs', in **SPAA '09: Proceedings of the Twenty-First Annual Symposium on Parallelism in Algorithms and Architectures. August 11-13, 2009, Calgary, Alberta, Canada**, pp. 260-269.
- Gutmann, MU, Hyvärinen, A 2009, 'Unsupervised learning by discriminating data from artificial noise', in **Proceedings of NIPS 2009 Workshop**.
- Gutmann, MU, Hyvärinen, A 2009, 'Learning features by contrasting natural images with noise', in **Proceedings of the 19th International Conference on Artificial Neural Networks (ICANN-09), Limassol, Cyprus, 14-17 September, 2009**, pp. 623-632.
- Gutmann, MU, Hyvärinen, A, Aihara, K 2009, 'Learning reconstruction and prediction of natural stimuli by a population of spiking neurons', in **European 17th Symposium on Artificial Neural Networks**.
- Hoyer, PO, Hyttinen, A 2009, 'Bayesian discovery of linear acyclic causal models', in **UAI '09: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence**, pp. 240-248.
- Hoyer, PO, Janzing, D, Mooij, J, Peters, J, Schölkopf, B 2009, 'Nonlinear causal discovery with additive noise models', in **Advances in neural information processing systems 21 : 2nd Annual Conference on Neural Information Processing Systems 2008**, pp. 689-696.
- Hämäläinen, W 2009, 'Lift-based search for significant dependencies in dense data sets', in **Proceedings of the KDD-09 Workshop on Statistical and Relational Learning in Bioinformatics, June, 28th, 2009, Paris, France / Conference Chairs: Christophe Costa Florencio, Fabrizio Costa Jan Ramon, Joost Kok. - ACM, cop. 2009**.
- Jursic, M, Lavrac, N, Mozetic, I, Podpecan, V, Toivonen, H 2009, 'Constructing information networks from textual documents', in **Proceedings of the Workshop on Explorative Analytics of Information Networks 2009 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Bled, Slovenia, September 11, 2009**, pp. 23-38.
- Kumpulainen, K, Myllymäki, P, Smeds, R, Kronqvist, J, Pöyry-Lassila, P 2009, 'VISCI virtual intelligent space for collaborative innovation: project description', in **Proceedings of 1st. International Symposium on Tangible Software Engineering Education, October 9, 2009, Tokyo International Forum, Tokyo, Japan : (STANS09)**, pp. 75-82.
- Kärkkäinen, J, Manzini, G, Puglisi, SJ 2009, 'Permuted Longest-Common-Prefix Array', in **Combinatorial Pattern Matching: 20th Annual Symposium, CPM 2009 Lille, France, June 22-24, 2009 / Proceedings**, pp. 181-192 **Lecture Notes in Computer Science**, vol. 5577.
- Köster, U, Hyvärinen, A 2009, 'Natural image statistics: energy-based models estimated by score matching', in **International Workshop on Local and Non-Local Approximation in Image Processing: General Chair: Jaakko Astola. - IEEE Xplore, cop. 2009**, pp. 16-25.
- Köster, U, Lindgren, JT, Gutmann, M, Hyvärinen, A 2009, 'Learning natural image structure with a horizontal product model', in **Independent Component Analysis and Signal Separation**, pp. 507-514 **Lecture Notes in Computer Science**, vol. 5441.
- Köster, U, Lindgren, JT, Hyvärinen, A 2009, 'Estimating Markov random field potentials for natural images', in **Proc. Int. Conference on Independent Component Analysis and Blind Signal Separation (ICA2009)**, pp. 515-522.
- Langohr, L, Toivonen, H 2009, 'Finding representative nodes in probabilistic graphs', in **Proceedings of the Workshop on Explorative Analytics of Information Networks 2009 Andreas Nurnberger, Michael R. Berthold, Tobias Köter, Kilian Thiel (eds.). - 2009**, pp. 65-76.
- Leino, A 2009, 'Multilingual Names on the Finnish Basic Map', in **Proceedings of the 23rd International Congress of Onomastic Sciences**, pp. 665-676.
- Merivuori, T, Roos, T 2009, 'Some observations on the applicability of normalized compression distance to stemmatology', in **Proceedings of the Second Workshop on Information Theoretic Methods in Science and Engineering, August 17-19, 2009, Tampere, Finland**.
- Myllymäki, P 2009, 'A framework for MDL clustering', in **Proceedings of the Second Workshop on Information Theoretic Methods in Science and Engineering, August 17-19, 2009, Tampere, Finland**.
- Mäkinen, V, Navarro, G, Sirén, J, Välimäki, N 2009, 'Storage and retrieval of individual genomes', in **Research in Computational Molecular Biology: 13th Annual International Conference, RECOMB 2009**, pp. 121-137 **Lecture Notes in Computer Science**, no. 5541.
- Parviainen, P, Koivisto, M 2009, 'Exact structure discovery in Bayesian networks with less space', in **Proceedings of the Proceedings of the Twenty-Fifth Conference Conference on Uncertainty in Artificial Intelligence (2009)**, pp. 436-443.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Perkiö, J, Myllymäki, P 2009, 'Magrathea: building and analyzing ubiquitous and social systems', in **Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology**: - IEEE, [2009], pp. 66-75.
- Perkiö, J, Tuominen, A, Myllymäki, P 2009, 'Image similarity: from syntax to weak semantics using multimodal features with application to multimedia retrieval', in **The First International Conference on Multimedia Information Networking and Security, MINES 2009: Proceedings**, pp. 213-219.
- Perkiö, J, Hyvärinen, A 2009, 'Modelling image complexity by independent component analysis, with application to content-based image retrieval', in **Proceedings of the 19th International Conference on Artificial Neural Networks (ICANN-09)**, Limassol, Cyprus, 14-17 September, 2009, pp. 704-714.
- Perkiö, J, Tuytelaars, T, Buntine, WL 2009, 'Exploring scale-induced feature hierarchies in natural images', in **Machine Learning and Applications, 2009. ICMLA '09**, pp. 25-31.
- Polishchuk, V, Pääkkö, A, Mitchell, JSB, Krozel, J 2009, 'Planning routes with wiggle room in en route weather-impacted airspaces', in **Proceedings of AIAA Guidance, Navigation, and Control Conference, 10-13 August 2009, Chigaco, IL, USA. - American Institute of Aeronautics and Astronautics, cop. 2009**.
- Rizo, D, Inesta, JM, Lemström, K 2009, 'Ensemble of state-of-the-art methods for polyphonic music comparison', in **Proceedings of the Workshop on Exploring Musical Information Spaces, ECDL 2009, Corfu, Greece, October 2009. - cop. 2009**, pp. 46-51.
- Rizo, D, Lemström, K, Inesta, JM 2009, 'Tree representation in combined polyphonic music comparison', in **Computer Music Modeling and Retrieval. Genesis of Meaning in Sound and Music: 5th International Symposium, CMMR 2008 Copenhagen, Denmark, May 19-23, 2008 Revised Papers**, pp. 177-195 *Lecture notes in computer science*, vol. 5493/2009.
- Roos, T, Yu, B 2009, 'Estimating sparse models from multivariate discrete data via transformed Lasso', in **Proceedings of the 2009 Information Theory and Applications Workshop, University of California, San Diego, USA, February 8-13, 2009. - IEEE, 2009**.
- Roos, T, Yu, B 2009, 'Sparse Markov source estimation via transformed Lasso', in **Proceedings of the 2009 IEEE Information Theory Workshop on Networking and Information Theory (ITW-09)**, Volos, Greece, June 10-12, 2009. - IEEE, 2009, pp. 241-245.
- Roos, T, Yu, B 2009, 'Recovering sparse models by parameter transformations: applications in Markov models and logistic regression', in **Proceedings of the Second Workshop on Information Theoretic Methods in Science and Engineering, August 17-19, 2009, Tampere, Finland**.
- Rousu, J 2009, 'Bioinformatics: technologies and challenges', in **Proceedings of AMICT'2008: Vol. 10 / editors: Jussi Kangasharju, Yury A. Bogoyavlenskiy**, pp. 50-57.
- Shimizu, S, Hyvärinen, A, Kawahara, Y, Washio, T 2009, 'A direct method for estimating a causal ordering in a linear non-gaussian acyclic model', in **Proc. 25th Conference on Uncertainty in Artificial Intelligence (UAI2009)**.
- Silander, T, Roos, T, Myllymäki, P 2009, 'Locally minimax optimal predictive modeling with Bayesian networks', in **JMLR workshop and conference proceedings**, pp. 504-511.
- Sirén, J 2009, 'Compressed suffix arrays for massive data', in **String Processing and Information Retrieval: 16th International Symposium, SPIRE 2009**, pp. 63-74 *Lecture Notes in Computer Science*, no. 5721.
- Tripathi, A, Klami, A, Kaski, S 2009, 'Using Dependencies to Pair Samples for Multi-View Learning', in **ICASSP 2009: IEEE International Conference on Acoustics, Speech and Signal Processing, 2009**, pp. 1561-1564.
- Yangarber, R, Steinberger, R 2009, 'Automatic Epidemiological Surveillance from On-line News in MedISys and PULS', in **Proceedings of IMED-2009: International Meeting on Emerging Diseases and Surveillance**.
- Yu, H, Bertsekas, DP 2009, 'Basis function adaptation methods for cost approximation in MDP', in **IEEE International Symposium on Adaptive Dynamic Programming and Reinforcement Learning: [Program Chair: Derong Liu]. - IEEE, 2009**.
- Zhang, K, Peng, H, Chan, L, Hyvärinen, AJ 2009, 'ICA with sparse connections: Revisited', in **Proc. Int. Conference on Independent Component Analysis and Blind Signal Separation (ICA2009)**, pp. 195-202.
- Zhang, K, Hyvärinen, A 2009, 'Causality discovery with additive disturbances: An informationtheoretical perspective', in **Proc. European Conference on Machine Learning (ECML2009)**, pp. 570-585.
- Zhang, K, Hyvärinen, A 2009, 'On the identifiability of the post-nonlinear causal model', in **Proc. 25th Conference on Uncertainty in Artificial Intelligence (UAI2009)**.
- Zhou, F, Mahler, S, Toivonen, H 2009, 'Review of network abstraction techniques', in **Proceedings of the Workshop on Explorative Analytics of Information Networks 2009 Andreas Nurnberger, Michael R. Berthold, Tobias Kötter, Kilian Thiel (eds.). - 2009**, pp. 50-63.
- Åstrand, M, Floreen, P, Polishchuk, V, Rybicki, J, Suomela, J, Uitto, J 2009, 'A local 2-approximation algorithm for the vertex cover problem', in **Distributed Computing: 23rd International Symposium, DISC 2009. Elche, Spain, September 23-25, 2009. Proceedings**, pp. 191-205 *Lecture Notes in Computer Science*, vol. 5805.

2010

- Ahonen, T 2010, 'Combining Chroma Features for Cover Version Identification', in **ISMIR 2010**.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Ahonen, T 2010, 'Compressing Lists for Audio Classification', in **MML'10**.

Ahonen-Myka, H, Doucet, A 2010, 'Statistical Methods for the Evaluation of Indexing Phrases', in **Proceedings of the International Conference on Knowledge Discovery and Information Retrieval (KDIR 2010)**.

Arkin, E, Efrat, A, Mitchell, J, Polishchuk, V, Ramasubramanian, S, Sankararaman, S, Taheri, J 2010, 'Data Transmission and Base-Station Placement for Optimizing Network Lifetime', in **DIALM-POMC '10 Proceedings of the 6th International Workshop on Foundations of Mobile Computing**, pp. 23-32.

Arroyuelo, D, Claude, F, Maneth, S, Mäkinen, V, Navarro, G, Nguyen, K, Siren, JLT, Välimäki, N 2010, 'Fast in-memory XPath search using compressed indexes', in **ICDE 2010: 26th IEEE International Conference on Data Engineering**, pp. 417-428.

Astikainen, K, Pitkänen, E, Rousu, J, Holm, L, Szedmak, S 2010, 'Reaction Kernels: Structured Output Prediction Approaches for Novel Enzyme Function', in **Proceedings of the First International Conference on Bioinformatics**, pp. 48-55.

Atkinson, M, Belyaeva, J, Zavarella, V, Piskorski, J, Huttunen, S, Vihavainen, A, Yangarber, R 2010, 'News Mining for Border Security Intelligence', in **Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on**.

Atkinson, M, Piskorski, J, Belyaeva, J, Huttunen, S, Yangarber, R 2010, 'Real-Time Text Mining in Multilingual News for the Creation of a Pre-frontier Intelligence Picture', in **Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**.

Bertsekas, DP, Yu, H 2010, 'Distributed Asynchronous Policy Iteration in Dynamic Programming', in **Proceedings of 2010 Allerton Conference on Communication, Control, and Computing**.

Björklund, A, Husfeldt, T, Kaski, P, Koivisto, M 2010, 'Covering and packing in linear space', in **Proceedings of the 37th International Colloquium on Automata, Languages and Programming (Bordeaux, July 6--10, 2010) : Part I**, pp. 727-737 **Lecture Notes in Computer Science**, vol. 6198.

Caldas, J, Kaski, S 2010, 'Hierarchical generative biclustering for microRNA expression analysis', in **Research in Computational Molecular Biology: Proceedings of 14th Annual International Conference RECOMB 2010, Lisbon, Portugal, April 25-28**, pp. 65-79 **Lecture Notes in Computer Science**, vol. 6044.

Durian, B, Peltola, H, Salmela, L, Tarhio, J 2010, 'Bit-parallel search algorithms for long patterns', in **Proceedings of the 9th International Symposium on Experimental Algorithms**, pp. 129-140 **Lecture Notes in Computer Science**, no. 6049, vol. 2010.

Eberhardt, F, Hoyer, PO, Scheines, R 2010, 'Combining experiments to discover linear cyclic models with latent variables', in **Proceedings of the 13th International Conference on Artificial Intelligence and Statistics**, pp. 185-192.

Entner, D, Hoyer, PO 2010, 'On causal discovery from time series data using FCI', in **Proceedings of the 5th European Workshop on Probabilistic Graphical Models**, pp. 121-128.

Gutmann, UM, Hyvärinen, A 2010, 'Noise-contrastive estimation: A new estimation principle for unnormalized statistical models', in **Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS2010)**, pp. 297-304 **JMLR Workshop and Conference Proceedings**, vol. 9.

Heule, M, Järvisalo, M, Biere, A 2010, 'Clause Elimination Procedures for CNF Formulas', in **Logic for Programming, Artificial Intelligence, and Reasoning - LPAR-17**, pp. 357-371 **Lecture Notes in Computer Science**, vol. 6397.

Heule, M, Järvisalo, M, Biere, A 2010, 'Covered Clause Elimination', in **Short Paper Proceedings of the 17th International Conference on Logic for Programming, Artificial Intelligence and Reasoning**, Yogyakarta, Indonesia, October 10-15, 2010.

Hintsanen, P, Toivonen, H, Sevón, P 2010, 'Fast Discovery of Reliable Subnetworks', in **2010 International Conference on Advances in Social Network Analysis and Mining: ASONAM 2010**, pp. 104-111.

Hirayama, J, Hyvärinen, A, Ishii, S 2010, 'Sparse and low-rank estimation of time-varying markov networks with alternating direction method of multipliers', in **Neural Information Processing. Theory and Algorithms: 17th International Conference, ICONIP 2010, Sydney, Australia, November 22-25, 2010, Proceedings, Part I**, pp. 371-379 **Lecture Notes in Computer Science**.

Huopaniemi, I, Suvitaival, T, Oresic, M, Kaski, S 2010, 'Graphical Multi-Way Models', in **Machine Learning and Knowledge Discovery in Databases: Proceedings of European Conference, ECML PKDD 2010, Barcelona, Spain, September 20-24, 2010**, pp. 538-553 **Lecture Notes in Computer Science**, vol. 6321.

Hyttinen, A, Eberhardt, F, Hoyer, PO 2010, 'Causal discovery for linear cyclic models with latent variables', in **Proceedings of the Fifth European Workshop on Probabilistic Graphical Models: PGM 2010**, pp. 153-160 [163-170] **HIIT Publications**, no. 2, vol. 2010.

Hyvärinen, A 2010, 'Pairwise measures of causal direction in linear non-gaussian acyclic models', in **JMLR: Workshop and Conference Proceedings Volume 13: 2nd Asian Conference on Machine Learning**, pp. 1-16.

Häggman, J, Juga, J, Sillanpää, MJ, Thompson, R 2010, 'Genetic parameters for claw health and foot and leg conformation traits in Finnish Ayrshire cows', in **Proceedings of the 9th World Congress on Genetics Applied to Livestock Production**.

Janzing, D, Hoyer, PO, Schölkopf, B 2010, 'Telling cause from effect based on high-dimensional observations', in **Proceedings of the 27th International Conference on Machine Learning**.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

- Järvisalo, M, Biere, A, Heule, M **2010**, 'Blocked Clause Elimination', in **Tools and Algorithms for the Construction and Analysis of Systems - TACAS 2010**, pp. 129-144 **Lecture Notes in Computer Science**, vol. 6015.
- Järvisalo, M, Biere, A **2010**, 'Reconstructing Solutions after Blocked Clause Elimination', in **Theory and Applications of Satisfiability Testing – SAT 2010**, pp. 340-345 **Lecture Notes in Computer Science**, vol. 6175.
- Karvonen, M, Laitinen, M, Lemström, K, Vikman, J **2010**, *Applying mathematical morphology for content-based music retrieval*, **Paper presented at International Symposium on Computer Music Modeling and Retrieval, Malaga, Spain. 21. - 24. June, 2010.**
- Kasari, M, Toivonen, H, Hintsanen, P **2010**, 'Fast Discovery of Reliable k-terminal Subgraphs', in **Advances in Knowledge Discovery and Data Mining: The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining, Part II**, pp. 168-177 **Lecture Notes in Artificial Intelligence**, no. 6119.
- Klami, A, Virtanen, S, Kaski, S **2010**, 'Bayesian exponential family projections for coupled data sources', in **Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence (2010)**, pp. 286-293.
- Koivisto, M, Parviainen, P **2010**, 'A space-time tradeoff for permutation problems', in **Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms**, pp. 484-492.
- Koolen, WM, Warmuth, MK, Kivinen, J **2010**, 'Hedging Structured Concepts', in **COLT 2010 : Proceedings of the 23rd Annual Conference on Learning Theory**, pp. 93-105.
- Kostitsyna, I, Polishchuk, V **2010**, 'Simple Wiggling is Hard unless you are a Fat Hippo', in **Fun with Algorithms**, pp. 272-283 **Lecture Notes in Computer Science**, vol. 6099.
- Krasnoshchekov, DN, Polishchuk, V, Vihavainen, A **2010**, 'Shape approximation using k-order alpha-hulls.', in **SoCG '10 Proceedings of the 10th annual symposium on Computational geometry**, pp. 109-110.
- Krozel, J, Mitchell, J, Polishchuk, V, Pääkkö, A **2010**, 'Throughput/Complexity Tradeoffs for Routing Traffic in the Presence of Dynamic Weather', in **ICRAT 2010 Fourth International Conference on Research in Air Transportation**.
- Kärkkäinen, J, Puglisi, SJ **2010**, 'Medium-Space Algorithms for Inverse BWT', in **Algorithms - ESA 2010: 18th Annual European Symposium, Liverpool, UK, September 6-8, 2010. Proceedings, Part I**, pp. 451-462 **Lecture Notes in Computer Science**, no. 6346, vol. 2010.
- Lai, P, Roos, TT, O'Sullivan, J **2010**, 'MDL hierarchical clustering for stemmatology', in **Information Theory Proceedings (ISIT)**, pp. 1403-1407.
- Lejeune, G, Doucet, A, Yangarber, R, Lucas, N **2010**, *Filtering news for epidemic surveillance: towards processing more languages with fewer resources*, **Paper presented at CLIA: Fourth International Workshop On Cross Lingual Information Access, at COLING 2010 (2010) , Beijing, China.**
- Lemström, K **2010**, 'Towards More Robust Geometric Content-Based Music Retrieval', in **ISMIR 2010**, pp. 577-582.
- Linge, J, Steinberger, R, Fuat, F, Bucci, S, Belyaeva, J, Gemo, M, Al-Khudhairy, D, Yangarber, R, van der Goot, E **2010**, 'MediSys: Medical Information System', in **Advanced ICTs for Disaster Management and Threat Detection: Collaborative and Distributed Frameworks**, pp. 131-142.
- Luosto, P, Kivinen, J, Mannila, H **2010**, 'Gaussian Clusters and Noise: An Approach Based on the Minimum Description Length Principle', in **Discovery Science: 13th International Conference, DS 2010, Canberra, Australia, October 6-8, 2010. Proceedings**, pp. 251-265 **Lecture Notes in Computer Science**, vol. 6332.
- Mozetic, I, Lavarac, N, Podpecan, V, Novak, PK, Motaln, H, Petek, M, Gruden, K, Toivonen, H, Kulovesi, K **2010**, 'Bisociative knowledge discovery for microarray data analysis', in **Proceedings of the International Conference on Computational Creativity: ICC-C-X, 7-9 January 2010, Lisbon, Portugal**, pp. 190-199.
- Norta, A, Yangarber, R, Carlson, L **2010**, 'Utility Evaluation of Tools for Collaborative Development and Maintenance of Ontologies', in **14th IEEE International Enterprise Distributed Object Computing Conference Workshops: VORTE 2010/MOST 2010**, pp. 207-214.
- Nöllenburg, M, Polishchuk, V, Sysikaski, M **2010**, 'Dynamic One-Sided Boundary Labeling', in **GIS '10 Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems**, pp. 310-319.
- Oikarinen, E, Woltran, S **2010**, 'Characterizing Strong Equivalence for Argumentation Frameworks', in **Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010**, pp. 123-133.
- Parkkinen, J, Nybo, K, Peltonen, J, Kaski, S **2010**, 'Graph Visualization With Latent Variable Models', in **Proceedings of the Eighth Workshop on Mining and Learning with Graphs: MLG-2010**, pp. 94-101.
- Parviainen, P, Koivisto, M **2010**, 'Bayesian structure discovery in Bayesian networks with less space', in **Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS) 2010**, pp. 589-596.
- Pihlaja, M, Gutmann, MU, Hyvärinen, AJ **2010**, 'A Family of Computationally Efficient and Simple Estimators for Unnormalized Statistical Models', in **Conference on Uncertainty in Artificial Intelligence**.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Polishchuk, V., Vihavainen, A 2010, 'Periodic Multi-Labeling of Public Transit Lines', in **GIScience: 6th International Conference, GIScience 2010, Zurich, Switzerland, September 14-17, 2010. Proceedings**, pp. 175-188 **Lecture Notes in Computer Science**, vol. 6292/2010.

Ramkumar, P., Hyvärinen, A., Parkkonen, L., Hari, R 2010, 'Characterization of spontaneous neuromagnetic brain rhythms using independent component analysis of short-time Fourier transforms', in **IFMBE Proceedings: 17th International Conference on Biomagnetism Advances in Biomagnetism – Biomag2010 March 28 – April 1, 2010 Dubrovnik, Croatia**.

Sirén, J 2010, 'Sampled longest common prefix array', in **Combinatorial Pattern Matching: 21st Annual Symposium, CPM 2010**, pp. 227-237 **Lecture Notes in Computer Science**, no. 6129.

Sogawa, Y., Shimizu, S., Hyvärinen, A., Washio, T., Shimamura, T., Imoto, S 2010, 'Discovery of exogenous variables in data with more variables than observations', in **Artificial Neural Networks – ICANN 2010: 20th International Conference Thessaloniki, Greece, September 15-18, 2010, Proceedings, Part I**, pp. 67-76 **Lecture Notes in Computer Science**, vol. 6352.

Su, H., Heinonen, M., Rousu, J 2010, 'Multilabel Classification of Drug-like Molecules via Max-Margin Conditional Random Fields', in **Proceedings of the Fifth European Workshop on Probabilistic Graphical Models (PGM-2010)**, pp. 265-272 **HIIT Publications**, no. 2, vol. 2010.

Su, H., Heinonen, M., Rousu, J 2010, 'Structured Output Prediction of Anti-Cancer Drug Activity', in **Pattern Recognition in Bioinformatics**, pp. 38-49 **Lecture Notes in Computer Science**, vol. 6282.

Toivonen, H., Mahler, S., Zhou, F 2010, 'A Framework for Path-Oriented Network Simplification', in **Advances in Intelligent Data Analysis IX: 9th International Symposium, IDA 2010**, pp. 220-231 **Lecture Notes in Computer Science**, no. 6065.

Tripathi, A., Klami, A., Virpioja, S 2010, 'Bilingual sentence matching using Kernel CCA', in **IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2010**, pp. 130-135.

Viinikanoja, J., Klami, A., Kaski, S 2010, 'Variational Bayesian mixture of robust CCA models', in **Machine Learning and Knowledge Discovery in Databases: Proceedings of European Conference, ECML PKDD 2010, Barcelona, Spain, September 20-24, 2010**, pp. 370-385 **Lecture Notes in Computer Science**, vol. 6323.

Välimäki, N., Ladra, S., Mäkinen, V 2010, 'Approximate All-Pairs Suffix/Prefix Overlaps', in **CPM 2010: Proceedings of the 21st Annual Symposium on Combinatorial Pattern Matching**, pp. 76-87.

Wettig, J., Hiltunen, S., Yangarber, R 2010, *Hidden Markov Models for Induction of Morphological Structure of Natural Language*, **Paper presented at WITMSE-2010: Workshop on Information Theoretic Methods in Science and Engineering, Tampere, Finland. 16. - 18. August, 2010.**

Yu, H 2010, 'Convergence of Least Squares Temporal Difference Methods Under General Conditions', in **The 27th International Conference on Machine Learning, Haifa, Israel.**

Zhang, K., Hyvärinen, A 2010, 'Source separation and higher-order causal analysis of MEG and EEG', in **Proceedings of the Twenty-Sixth Conference (UAI 2010): Uncertainty in Artificial Intelligence**, pp. 709-716.

Zhang, K., Hyvärinen, A 2010, 'Distinguishing causes from effects using nonlinear acyclic causal models', in **JMLR Workshop and Conference Proceedings Volume 6: Causality: Objectives and Assessment (NIPS 2008)**, pp. 157-164.

Zhou, F., Mahler, S.J., Toivonen, H 2010, 'Network Simplification with Minimal Loss of Connectivity', in **Proceedings of the 10th IEEE International Conference on Data Mining (ICDM)**, pp. 659-668.

von Etter, P., Huttunen, S., Vihavainen, A., Vuorinen, M., Yangarber, R 2010, 'Assessment of Utility in Web Mining for the Domain of Public Health', in **Proceedings of the NAACL HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents**.

B1 Unrefereed journal article

2006

Mäkinen, V 2006, 'Itseindeksit: kun tiivistetty teksti ja sen indeksi ovatkin sama asia', **Tietojenkäsittelytiede**, vol 2006, no. 25, pp. 28-37.

2010

Salmela, L 2010, 'Merkkijonoalgoritmeja monen hahmon hakuun', **Tietojenkäsittelytiede**, vol 2010, no. 31, pp. 70-83.

B2 Contribution to book/other compilations (non-refereed)

2008

Kivinen, J 2008, 'Attribute-efficient learning', in M Kao (ed.), **Encyclopedia of Algorithms**, Springer.

B3 Unrefereed article in conference proceedings



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

2005

Buntine, WL, Valtonen, K, Taylor, MP **2005**, 'The ALVIS Document Model for a Semantic Search Engine', in **Proceedings of the 2nd European Semantic Web Conference (ESWC 2005)**.

Roos, TT, Heikkilä, T, Cilibrasi, R, Myllymäki, P **2005**, 'Compression-based stemmatology: A study of the legend of St. Henry of Finland', in **17th European Conference on Artificial Intelligence (ECAI'06), Pascal Eprints**.

Tuulos, VH, Perkiö, JP, Tirri, H **2005**, 'Multi-faceted information retrieval system for large scale email archives', in **SIGIR 2005: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval**.

2008

Bonifati, A, Leighton, G, Mäkinen, V, Maneth, S, Navarro, G, Pugliese, A **2008**, 'An in-memory XQuery/XPath engine over a compressed structured text representation', in **Dagstuhl seminar proceedings, Dagstuhl Seminar Proceedings, no. 08261**.

Mäkinen, V, Navarro, G, Siren, J, Välimäki, N **2008**, 'Storage and retrieval of individual genomes (extended abstract)', in **Dagstuhl seminar proceedings, Dagstuhl seminar proceedings, no. 08261**.

2009

Ahonen, TE **2009**, 'Cover Song Identification Based on Data Compression', in **Proceedings of Music Information Retrieval Evaluation eXchange (MIREX 2009)**.

Gutmann, MU, Hyvärinen, A **2009**, 'Learning spike-timings based representations of sensory stimuli with leaky integrate-and-fire neurons', in **Eighteenth Annual Computational Neuroscience Meeting**, pp. P144.

2010

Leino, A **2010**, 'Experiences in teaching computer usability in Second Life', in **Blended Learning in Finland**, pp. 154–166.

Roos, T, Schmidt, DF **2010**, 'On the consistency of sequentially normalized least squares', in **The Third Workshop on Information Theoretic Methods in Science and Engineering : WITMSE 2010, TICSP Series, no. 55**.

C1 Published scientific monograph

2005

Kollin, J, Koivisto, M **2005**, *Bayesian Learning with Mixtures of Trees*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki**.

Kääriäinen, M **2005**, *On generalization error bounds using unlabeled data*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Hki**.

Wang, JTL, Toivonen, H **2005**, *Data mining in bioinformatics*, **Advanced information and knowledge processing, Springer, New York**.

2006

Kaski, P, Östergård, PRJ **2006**, *Classification algorithms for codes and designs*, **Algorithms and computation in mathematics, no. 15, Springer, Berlin**.

Navarro, G, Mäkinen, V **2006**, *Compressed full-text indexes*, **University of Chile, Department of Computer Science, Santiago**.

Yu, H, Bertsekas, DP **2006**, *Convergence results for some temporal difference methods based on least squares*, **[Massachusetts Institute of Technology], [USA]**.

Yu, H, Bertsekas, DP **2006**, *A least squares Q-learning algorithm for optimal stopping problems*, **[Massachusetts Institute of Technology], [USA]**.

Yu, H, Bertsekas, DP **2006**, *On near-optimality of the set of finite-state controllers for average cost POMDP*, **[Massachusetts Institute of Technology], [USA]**.

2007

Lu, C, Rousu, J **2007**, *Experiments for Completion of Biological Networks with Output Kernel Trees*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki**.

Rissanen, J **2007**, *Information and complexity in statistical modeling*, **Information Science and Statistics, vol. VIII, Springer, New York**.

Yu, H, Bertsekas, DP **2007**, *Solution of large systems of equations using approximate dynamic programming methods*, **[Massachusetts Institute of Technology], [USA]**.

Yu, H, Rousu, J **2007**, *An efficient method for large margin parameter optimization in structured prediction problems*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Hki**.



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

2008

Kaski, P, Pottonen, O **2008**, *libexact user's guide: version 1.0*, **HIIT technical reports, no. 2008-1, HIIT, Helsinki.**

Mäkinen, V, Siren, J, Välimäki, N **2008**, *Storage and retrieval of individual genomes and other repetitive sequence collections*, **Department of Computer Science Series of Publications C, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

Mäkinen, V, Navarro, G, Siren, J, Välimäki, N **2008**, *Run-length compressed indexes for repetitive sequence collections*, **Department of Computer Science Series of Publications C, vol. C-2008-42, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

Yu, H, Bertsekas, DP **2008**, *New error bounds for approximations from projected linear equations*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

2009

Astikainen, K, Holm, LUT, Pitkänen, E, Rousu, J **2009**, *Reaction kernels: predicting enzyme functions you have never seen before*, **Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

Bertsekas, DP, Yu, H **2009**, *A unifying polyhedral approximation framework for convex optimization*, **[Massachusetts Institute of Technology], [USA].**

Hyvärinen, A, Hurri, J, Hoyer, PO **2009**, *Natural image statistics: a probabilistic approach to early computational vision*, **Springer, New York.**

2010

Bertsekas, D, Yu, H **2010**, *Q-Learning and Enhanced Policy Iteration in Discounted Dynamic Programming (Revised)*, **Technical Report, no. C-2010-10, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

Yu, H **2010**, *Convergence of Least Squares Temporal Difference Methods Under General Conditions*, **Department of Computer Science Series of Publications C Report C-2010-1, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

Yu, H **2010**, *Least Squares Temporal Difference Methods: An Analysis Under General Conditions*, **Technical Report, no. C-2010-39, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

C2 Edited book, compilation, conference proceeding or special issue of journal

2005

Bloedorn, S, Buntine, WL, Hotho, A **2005**, *Proceedings of Learning in Web Search (LWS 2005)*.

Bloedorn, S, Buntine, W, Hotho, A (eds) **2005**, *Proceedings of the Workshop on Learning in Web Search (LWS 2005), 7-11 August 2005 in Bonn, Germany*, **260.**

Bounsaythip, C, Hollmen, J, Kaski, S, Oresic, M (eds) **2005**, *Proceedings of KRBO'05, International Symposium on Knowledge Representation in Bioinformatics, Helsinki University of Technology, Espoo, Finland, June 15-17, 2005*, **TKK, Espoo.**

2006

Califf, ME, Greenwood, MA, Stevenson, M, Yangarber, R (eds) **2006**, *Proceedings of the Workshop on Information Extraction Beyond The Document*, **Association for Computational Linguistics.**

Beigbader, M, Buntine, W, Yee, WG (eds) **2006**, *OSIR 2006: Second Workshop on Open Source Information Retrieval: August 10, 2006, Seattle, WA: in conjunction with the 2006 ACM SIGIR Conference.*

Lemström, K (ed.), Tindale, A, Dannenberg, R **2006**, *[Proceedings of the] 7th International Conference on Music Information Retrieval, Fairmont Empress Hotel, Victoria, BC, Canada, October 8-12, 2006*, **University of Victoria.**

Rousu, J, Kaski, S, Ukkonen, E (eds) **2006**, *Probabilistic modeling and machine learning in structural and systems biology: workshop proceedings, Tuusula, Finland, June 17-18, 2006*, **Series of Publications B / Helsingin yliopisto, no. 4, Helsinki University Printing House, Helsinki.**

Rousu, J, Kaski, S, Ukkonen, E (eds) **2006**, *Probabilistic Modeling and Machine Learning in Structural and Systems Biology: Workshop Proceedings; Tuusula, Finland, June 17-18*, **BMC Bioinformatics, no. Supplement 2, vol. 8, University of Helsinki.**

2007

Mäkinen, V, Linden, G, Toivonen, H (eds) **2007**, *Summer School on Algorithmic Data Analysis (SADA 2007) and Annual Hecse Poster Session*, **Series of publications / Department of Computer ScienceB, no. 2007-4, Helsingin yliopisto, tietojenkäsittelytieteen laitos, Helsinki.**

2008



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

Blohm, S, Brefeld, U, Jungermann, F, Yangarber, R (eds) **2008**, *Proceedings of the ECML Workshop on High-level Information Extraction*,.

Grunwald, P, Myllymäki, P, Tabus, I, Weinberger, M, Yu, B (eds) **2008**, *Festschrift in honor of Jorma Rissanen on the occasion of his 75th birthday*, **TICSP series, no. 38, Tampere International Center for Signal Processing, Tampere**.

Heikkonen, J, Kontoyiannis, I, Liski, EP, Myllymäki, P, Rissanen, J, Tabus, I (eds) **2008**, *Proceedings of the First Workshop on Information Theoretic Methods in Science and Engineering, August 18-20, 2008, Tampere, Finland*, **Tampere International Center for Signal Processing, TICSP series, no. 43, Tampere International Center for Signal Processing, Tampere**.

McAllester, D, Myllymäki, P (eds) **2008**, *Uncertainty in Artificial Intelligence: proceedings of the Twenty-Fourth Conference (2008), July 9-12, 2008, Helsinki, Finland*, **AUAI Press, Corvallis, Or**.

Mäkinen, V, Lindén, G, Kujala, I (eds) **2008**, *From data to knowledge: final report 2002-2007*, **Julkaisusarja / Helsingin yliopisto, tietojenkäsittelytieteen laitosB, no. 2008-1, University of Helsinki, Helsinki**.

Poibeau, T, Saggion, H, Yangarber, R (eds) **2008**, *Multi-source, Multilingual Information Extraction and Summarization: MMIES-2*, **Association for Computational Linguistics**.

2009

Dzeroskig, S, Geurts, P, Rousu, J (eds) **2009**, *Machine learning in systems biology: proceedings of The Third International Workshop, September 5-6, 2009, Ljubljana, Slovenia*, **Series of Publications B / Report, University of Helsinki, Department of Computer ScienceB, no. B-2009-1, Helsinki University Printing House, Helsinki**.

Heikkonen, J, Kontoyiannis, I, Liski, EP, Myllymäki, P, Rissanen, J, Tabus, I (eds) **2009**, *Proceedings of the Second Workshop on Information Theoretic Methods in Science and Engineering, August 17-19, 2009, Tampere, Finland*, **Tampere International Center for Signal Processing, TICSP series, no. 49, Tampere International Center for Signal Processing, Tampere**.

Kucherov, G, Ukkonen, E (eds) **2009**, *Combinatorial pattern matching: 20th annual symposium, CPM 2009, Lille, France, June 22-24, 2009 : proceedings*, **LNCS sublibrarySL 1, no. 5577, Springer, Berlin**.

2010

Arita, M, Heinonen, M, Rousu, J (eds) **2010**, *Mass Spectrometry Informatics in Systems Biology: Abstracts of the Workshop*, **Series of Publications C, University of Helsinki, Department of Computer Science**.

Kaski, S, Miller, DJ, Oja, E, Honkela, A (eds) **2010**, *Proceedings of the 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010)*, **IEEE**.

Myllymäki, P, Roos, T, Jaakkola, T **2010**, *Proceedings of the Fifth European Workshop on Probabilistic Graphical Models, HIIT Publications, no. 2010-2, Helsinki Institute for Information Technology, Helsinki*.

Yamanishi, K, Kontoyiannis, I, Liski, EP, Myllymäki, P, Rissanen, JJ, Tabus, I (eds) **2010**, *Proceedings of the Third Workshop on Information Theoretic Methods in Science and Engineering, August 16-18, 2010, Tampere, Finland*, **Tampere International Center for Signal Processing, TICSP series, no. 55, Tampere International Center for Signal Processing**.

D1 Article in professional journal

2010

Roos, T **2010**, 'Terveisiä huippuyliopistoista', **Tietojenkäsittelytiede**, no. 30, pp. 7-12.

Ukkonen, E **2010**, 'Tila ja tulevaisuus : pysähtyneisyydestä uuteen vauhtiin', **Tietojenkäsittelytiede**, no. 30, pp. 4-6.

D2 Article in professional hand or guide book or in a professional data system, or text book material

2010

Toivonen, H **2010**, 'Frequent Pattern', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning**, 1. edn, **Springer**.

Toivonen, H **2010**, 'Frequent Itemset', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning**, 1. edn, **Springer**.

Toivonen, H **2010**, 'Association Rule', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning**, 1. edn, **Springer**.

Toivonen, H **2010**, 'Basket Analysis', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning**, 1. edn, **Springer**.

Toivonen, H **2010**, 'Apriori Algorithm', in C Sammut, GI Webb (eds), **Encyclopedia of Machine Learning**, 1. edn, **Springer**.

D4 Published development or research report



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

2005

Kontkanen, P, Myllymäki, P **2005**, *Analyzing the stochastic complexity via tree polynomials*, **HIIT Technical Reports 2005-4**.

Kontkanen, P, Myllymäki, P **2005**, *Computing the regret table for multinomial data*, **HIIT Technical Reports, no. 2005-1, vol. 2005**, vol. 2005.

2006

Kontkanen, P, Myllymäki, P **2006**, *Information-theoretically optimal histogram density estimation*, **HIIT Technical Reports, no. 2006-2, vol. 2006, Helsinki Institute for Information Technology, Espoo**.

2008

Yangerber, R, Salmenkivi, M, Välsälo, M **2008**, *A Database of the Uralic Language Family for Etymological Research*, **Technical Report C, no. C-2008-38, University of Helsinki, Department of Computer Science**.

2010

Belov, A, Järvisalo, M **2010**, *Depth-Based Local Search Heuristics for Boolean Circuit Satisfiability*, **HIIT Technical Reports, vol. 2010-2, vol. 2010-2, Helsinki Institute for Information Technology HIIT, Helsinki, Finland**.

E1 Popular article, newspaper article

2005

Raento, M, Oulasvirta, A, Toivonen, H, Mäntylä, M **2005**, 'Sosiaalista tilatietoa kontekstipuhelimella', **Proessori**, vol 2005, no. 1, pp. 54-56.

2006

Doucet, A **2006**, 'Prendre les mots dans le bon sens, une question d'ordre', **Universitas Helsingiensis : a quarterly**, vol 25, no. 4, pp. 36-38.

2007

Doucet, A **2007**, 'Opponentti, kustos, karonkka jne', **Yliopistolainen : Helsingin yliopiston henkilöstölehti**, no. 2, pp. 10.

2008

Mäkinen, V **2008**, 'Kauniita algoritmeja väärin ongelmiin', **Luova : tiedettä & teknologiaa**.

H1 Patents

2005

Kontkanen, P, Lahtinen, J, Myllymäki, P, Silander, T, Tirri, H, Valtonen, K **2005**, *Visualization method and visualization system*, US6873325.

Myllymäki, P, Kontkanen, P, Roos, T, Valtonen, K, Lahtinen, J, Wettig, J, Tuominen, AT, Tirri, H **2005**, *Error estimate concerning a target device's location operable to move in a wireless environment*, US2005/0131635 (A1).

Toivonen, H, Onkamo, P, Vasko, K, Ollikainen, VV, Sevon, P, Mannila, H, Kere, J May. 20 2010, *Method for gene mapping from chromosome and phenotype data*, 6909971.

2007

Misikangas, P, Myllymäki, P **2007**, *Probabilistic model for a positioning technique*, US2007117568 (A1).

2008

Misikangas, P, Myllymäki, P **2008**, *Sequence-based positioning technique*, US7349683.

I1 Audiovisual materials

2010

MapLab



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF PUBLICATIONS DATA 2005-2010

ALKO/Ukkonen

I2 ICT programs or applications

2009

MOODS

Run-Length Compressed Suffix Array

2010

InvCoal

Hybrid SHREC

Generalized Compressed Suffix Array



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

1 Analysis of activities 2005-2010

- Associated person is one of Esko Ukkonen , Patrik Hoyer , Aapo Hyvärinen , Juha Kärkkäinen , Petteri Kaski , Samuel Kaski , Jyrki Kivinen , Mikko Koivisto , Kjell Lemström , Veli Mäkinen , Heikki Mannila , Mikko Sillanpää , Petri Myllymäki , Hannu Toivonen , Roman Yangarber , Helena Ahonen-Myka , Wray Lindsay Buntine , Jorma Johannes Rissanen , Henry Tirri , Ella Bingham , Enrique Garcia Moreno-Esteve , Teemu Kivioja , Sebastian Jean Mahler , Gemma Garriga, Aristides Gionis , Michael Urs Gutmann , Jarmo Kalevi Hurri, Saara Annele Hyvönen , Matti Järvisalo , Matti Kääriäinen , Teemu Kivioja , Stefan Schönauer , Emilia Oikarinen , Cinzia Pizzi , Leena Salmela , Petteri Sevón , Hellis Tamm, Panayiotis Tsaparas , Huizhen Yu , Niina Susanna Haiminen , Suvi Wilhelmina Hämäläinen , Petri Kontkanen , Miro Petteri Lehtonen , Antti Leino , Taneli Johannes Mielikäinen , Pauli Aleksii Miettinen , Kimmo Juhani Palin , Ari Juhani Rantanen, Pasi Rastas , Teemu Teppo Roos , Lili Aunimo, Ilkka Kalevi Autio, Antoine Doucet , Arto Olavi Klami , Jussi Jalmar Kolli , Kari Tapani Laasonen, Jussi Tapio Lindgren, Juha Antero Makkonen , Esa Pitkänen , Mika Petteri Raento, Tomi Silander , Evimaria Terzi, Teppo Ahonen , Katja Astikainen , Doris Entner , Lauri Eronen , Esther Galbrun , Markus Heinonen , Petteri Hintsanen , Antti Hyttinen , Esa Juhani Junttila , Janne Henrik Korhonen , Reeta Kuuskoski, Laura A Langohr , Panu Luosto , Krishnan Narayanan , Joonas Paalasmaa, Pekka Parviainen , Jukka Petteri Perkiö , Miika Pihlaja , Jouni Stephan Puuronen , Jouni Sirén , Jarkko Toivonen , Abhishek Tripathi , Niko Välimäki , Matti Vuorinen , Jaana Wessman , Fang Zhou ,

Activity type	Count
Supervisor or co-supervisor of doctoral thesis	62
Prizes and awards	22
Editor of research journal	72
Editor of research anthology/collection/conference proceedings	47
Peer review of manuscripts	457
Editor of series	4
Editor of special theme number	3
Assessment of candidates for academic posts	25
Membership or other role in review committee	43
Membership or other role in research network	12
Membership or other role in national/international committee, council, board	42
Membership or other role in public Finnish or international organization	16
Membership or other role of body in private company/organisation	3
Other tasks of an expert in private sector	3
Participation in interview for written media	48
Participation in radio programme	5



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Activity type	Count
Participation in TV programme	2



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

2 Listing of activities 2005-2010

Supervisor or co-supervisor of doctoral thesis

Esko Ukkonen ,

Supervision of Doctoral thesis of Hellis Tamm, Esko Ukkonen, 2005
PhD Supervisor of Jarkko Toivonen, Esko Ukkonen, 2006 → ..., Finland
Supervision of Doctoral thesis of Ari Rantanen, Esko Ukkonen, 2006
Supervision of Doctoral thesis of Kimmo Palin, Esko Ukkonen, 2007
PhD Supervisor of Otto Solin, Esko Ukkonen, 2008 → ..., Finland
Supervision of Doctoral thesis of Ilkka Autio, Esko Ukkonen, 2008
Supervision of Doctoral thesis of Juha Makkonen, Esko Ukkonen, 2009
Supervision of Doctoral thesis of Pasi Rastas, Esko Ukkonen, 2009
Supervision of Doctoral thesis of Margus Lukk, Esko Ukkonen, 2010
Supervision of Doctoral thesis of Esa Pitkänen, Esko Ukkonen, 2010

Patrik Hoyer ,

PhD Supervisor of Antti Hyttinen, Patrik Hoyer, 2008 → ..., Finland
PhD Supervisor of Doris Entner, Patrik Hoyer, 2008 → ..., Finland

Aapo Hyvärinen ,

PhD supervision: Jussi Lindgren, Aapo Hyvärinen, 2004 → 2008
PhD supervision: Urs Köster, Aapo Hyvärinen, 2004 → 2009

Samuel Kaski ,

PhD Thesis supervisor of 7 PhDs in 2005-2010, Samuel Kaski, 2005 → 2010
PhD supervisor of Abhishek Tripathi, Samuel Kaski, 2006 → 2011

Jyrki Kivinen ,

Doctoral thesis supervision, Jyrki Kivinen, 01.07.2008 → 30.06.2012, Finland

Mikko Koivisto ,

Co-supervisor of Jussi Kollin's PhD, Mikko Koivisto, 2005 → 10.2010
Supervisor of Pekka Parviainen's PhD, Mikko Koivisto, 2007 → 2011
Supervisor of Janne Korhonen's PhD, Mikko Koivisto, 2009 → 2013
Co-supervisor of Esther Galbrun's PhD, Mikko Koivisto, 2010 → 2014

Kjell Lemström ,

Supervision of Teppo Ahonen's Doctoral Studies, Kjell Lemström, 2008 → ..., Finland

Veli Mäkinen ,

PhD Supervisor of Jouni Siren, Veli Mäkinen, 2007 → ..., Finland
PhD Supervisor of Niko Välimäki, Veli Mäkinen, 2008 → ..., Finland

Heikki Mannila ,

PhD Supervisor of Evimaria Terzi, Heikki Mannila, 2002 → 2007, Finland
PhD Supervisor of Taneli Miellikäinen, Heikki Mannila, 2002 → 2005, Finland
PhD Supervisor of Antti Leino, Heikki Mannila, 2003 → 2007, Finland
PhD Supervisor of Niina Haiminen, Heikki Mannila, 2004 → 2008, Finland
PhD Supervisor of Pauli Miettinen, Heikki Mannila, 2006 → 2009, Finland



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

PhD Supervisor of Aleksis Kallio, Heikki Mannila, 2007 → 2011, Finland

Petri Myllymäki ,

Supervision of Teemu Roos's doctoral thesis, Petri Myllymäki, 2003 → 2007, Finland

Supervision of Jukka Perkiö's doctoral thesis, Petri Myllymäki, 2004 → ..., Finland

Supervision of Petri Kontkanen's doctoral thesis, Petri Myllymäki, 2004 → 2009, Finland

Supervision of Tomi Silander's doctoral thesis, Petri Myllymäki, 2004 → 2009, Finland

Supervision of Tommi Mononen's doctoral thesis, Petri Myllymäki, 2004 → 2009, Finland

Juho Rousu ,

Doctoral thesis supervisor, Juho Rousu, 01.2001 → 11.2006, Finland

Doctoral thesis supervisor, Juho Rousu, 07.2004 → 11.2010, Finland

Doctoral thesis supervisor, Juho Rousu, 01.2008 → ..., Finland

Doctoral thesis supervisor, Juho Rousu, 01.2009 → ..., Finland

Mikko Sillanpää ,

PhD Supervisor of Matti Pirinen, Mikko Sillanpää, 2005 → 2009, Finland

PhD Supervisor of Pinja Pikkuhookana, Mikko Sillanpää, 2006 → ..., Finland

PhD Supervisor of Timo Knurr, Mikko Sillanpää, 2008 → ..., Finland

PhD Supervisor of Hanni Kärkkäinen, Mikko Sillanpää, 2009 → ..., Finland

PhD Supervisor of Mahlako Makgahlela, Mikko Sillanpää, 2009 → ..., Finland

PhD Supervisor of Zitong Li, Mikko Sillanpää, 2009 → ..., Finland

Supervisor of the PhD thesis, Department of Mathematics and Statistics, UH, Mikko Sillanpää, 06.2009, Finland

Hannu Toivonen ,

PhD supervisor of Mika Raento, Hannu Toivonen, 2002 → 2007, Finland

PhD supervisor of Petteri Hintsanen, Hannu Toivonen, 2003 → 2011, Finland

PhD supervisor of Kari Laasonen, Hannu Toivonen, 2004 → 2009, Finland

PhD supervisor of Lauri Eronen, Hannu Toivonen, 2004 → 2011, Finland

PhD supervisor of Kimmo Hätönen, Hannu Toivonen, 2007 → 2009, Finland

PhD supervisor of Fang Zhou, Hannu Toivonen, 2008 → ..., Finland

PhD supervisor of Laura Langohr, Hannu Toivonen, 2008 → ..., Finland

PhD supervisor of Wilhelmiina Hämäläinen, Hannu Toivonen, 2008 → 2010, Finland

PhD supervisor of Esther Galbrun, Hannu Toivonen, 2010 → ..., Finland

PhD supervisor of Joonas Paalasmaa, Hannu Toivonen, 2010 → ..., Finland

Helena Ahonen-Myka ,

PhD supervisor of Oskari Heinonen, Helena Ahonen-Myka, 1997 → ..., Finland

PhD supervisor of Antoine Doucet, Helena Ahonen-Myka, 2005, Finland

PhD supervisor of Lili Aunimo, Helena Ahonen-Myka, 2007, Finland

PhD supervisor of Miro Lehtonen, Helena Ahonen-Myka, 2007, Finland

PhD supervisor of Juha Makkonen, Helena Ahonen-Myka, 2009, Finland

Pasi Rastas ,

Supervisor, Pasi Rastas, 2005 → 2010

Prizes and awards

Esko Ukkonen ,



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Medix-prize, Esko Ukkonen, 2007

Science Prize of the City of Helsinki 2007, Esko Ukkonen, 2007

The first honorary member of bioinformatics society Bioinformatiikan seura ry, Esko Ukkonen, 13.05.2009

Silver Medal of the University of Helsinki, Esko Ukkonen, 2010

Aapo Hyvärinen ,

Highly Cited researcher, Aapo Hyvärinen, 2010

Petteri Kaski ,

2007 Kirkman medal, Petteri Kaski, 2007, Canada

Veli Mäkinen ,

Hyvä tutkija -palkinto 2007, Veli Mäkinen, 19.12.2007, Finland

Hannu Toivonen ,

Best Paper Award, European Conference on Machine Learning, Hannu Toivonen, 2008

Honorary Member of TKO-äly, Hannu Toivonen, 2010, Finland

Jorma Johannes Rissanen ,

Kolmogorov medal, Jorma Johannes Rissanen, 2006

2009 IEEE Claude E. Shannon Award, Jorma Johannes Rissanen, 2009

Leena Salmela ,

Award for best paper submitted to HITSeq 2010, Leena Salmela, 10.07.2010

Petri Kontkanen ,

Palkinto artikkelista, Petri Kontkanen, 2007

Väitöskirjapalkinto, Petri Kontkanen, 2010, United States

Pauli Aleksi Miettinen ,

Certificate of Recognition, ACM SIGKDD Doctoral Dissertation Award 2010, Pauli Aleksi Miettinen, 2010

Teemu Teppo Roos ,

2009 Cor Baayen Award, Teemu Teppo Roos, 06.11.2009, France

Katja Astikainen ,

Best student paper award, Katja Astikainen, 23.01.2010, Spain

Janne Henrik Korhonen ,

Good Teacher Award for junior teachers, Janne Henrik Korhonen, 19.12.2007, Finland

Pekka Parviainen ,

The Best Student Paper runner-up award (UAI'09), Pekka Parviainen, 19.06.2009, Canada

Jukka Petteri Perkiö ,

International Neural Network Society Best Student Paper Award, Jukka Petteri Perkiö, 17.09.2009

Abhishek Tripathi ,

Honorary mention at IEEE ICASSP 2009., Abhishek Tripathi, 20.04.2009, Taiwan

Niko Välimäki ,

Nuorempien tutkijoiden Hyvä tutkija -palkinto, Niko Välimäki, 2009, Finland

Editor of research journal

Esko Ukkonen ,

Member of the Board of Editors of the Journal of Universal Computer Science, Esko Ukkonen, 01.01.1999 → ..., Germany



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

IEEE/ACM Transactions on Computational Biology and Bioinformatics, Esko Ukkonen, 2004 → ..., United States

Petteri Kaski ,

Discrete Mathematics, Petteri Kaski, 01.01.2006 → 31.12.2006

Graphs and Combinatorics, Petteri Kaski, 01.01.2006 → 31.12.2006

IEEE Transactions on Knowledge and Data Engineering, Petteri Kaski, 01.01.2006 → 31.12.2006

Journal of Combinatorial Designs, Petteri Kaski, 01.01.2006 → 31.12.2006

Journal of Combinatorial Mathematics and Combinatorial Computing, Petteri Kaski, 01.01.2006 → 31.12.2006

Rendiconti del Circolo Matematico di Palermo, Petteri Kaski, 01.01.2006 → 31.12.2006

15th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB'07), Petteri Kaski, 01.01.2007 → 31.12.2007

Current Trends in Theory and Practice of Computer Science (SOFSEM'08), Petteri Kaski, 01.01.2007 → 31.12.2007

Discrete Applied Mathematics, Petteri Kaski, 01.01.2007 → 31.12.2007

Discrete Mathematics, Petteri Kaski, 01.01.2007 → 31.12.2007

IEEE Transactions on Information Technology in Biomedicine, Petteri Kaski, 01.01.2007 → 31.12.2007

IEEE Transactions on Knowledge and Data Engineering, Petteri Kaski, 01.01.2007 → 31.12.2007

Journal of Combinatorial Designs, Petteri Kaski, 01.01.2007 → 31.12.2007

ACM Transactions on Algorithms, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

ACM Transactions on Knowledge Discovery from Data, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

Algorithm Theory SWAT 2008 11th Scandinavian Workshop on Algorithm Theory, Gothenburg, Sweden, July 2-4, 2008. Proceedings Book Series Lecture Notes in Computer Science Publisher Springer Berlin / Heidelberg ISSN 0302-9743 (Print) 1611-3349 , Petteri Kaski, 01.01.2008 → 31.12.2008, Germany

Discrete Mathematics, Petteri Kaski, 01.01.2008 → 31.12.2008, Netherlands

Graphs and Combinatorics, Petteri Kaski, 01.01.2008 → 31.12.2008, Japan

International Journal of Computer Mathematics, Petteri Kaski, 01.01.2008 → 31.12.2008, United Kingdom

Journal of Combinatorial Designs, Petteri Kaski, 2008 → 2012

Journal of Combinatorial Designs, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

Journal of Combinatorial Designs, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

Journal of Machine Learning Research, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

Journal of Statistics and Applications, Petteri Kaski, 01.01.2008 → 31.12.2008, United States

Machine Learning and Knowledge Discovery in Databases European Conference, ECML PKDD 2008, Antwerp, Belgium, September 15-19, 2008, Proceedings Book Series Lecture Notes in Computer Science Publisher Springer Berlin / Heidelberg ISSN 0302-9743, Petteri Kaski, 01.01.2008 → 31.12.2008, Germany

Samuel Kaski ,

Intelligent Data Analysis, Samuel Kaski, 2003 → 2010

International journal of neural systems, Samuel Kaski, 01.01.2004 → ...

Cognitive Neurodynamics, Samuel Kaski, 2007 → ...

International Journal of Knowledge Discovery in Bioinformatics, Samuel Kaski, 2009 → ...

Mikko Koivisto ,

Reviewer for IEEE/ACM Transactions on Computational Biology and Bioinformatics, Mikko Koivisto, 2006 → ...

Reviewer for Journal of Machine Learning Research, Mikko Koivisto, 2006 → ...

Kjell Lemström ,

Member of the Editorial Board of the Musiikki Magazine, Kjell Lemström, 03.2004 → 02.2009, Finland

Petri Myllymäki ,

Entropy, Petri Myllymäki, 1998 → 2006



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Journal of Artificial Intelligence Research, Petri Myllymäki, 2010 → ...

Statistics and Computing, Petri Myllymäki, 2010 → ...

Valentin Polishchuk ,

15th International Conference on Telecommunications ICT June 2008, Valentin Polishchuk, 01.01.2008 → 31.12.2008

Computational Geometry: Theory and Applications, Valentin Polishchuk, 01.01.2008 → 31.12.2008

Journal of Systems and Software, Valentin Polishchuk, 01.01.2008 → 31.12.2008

The 19th International Symposium on Algorithms and Computation ISAAC 2008, Valentin Polishchuk, 01.01.2008 → 31.12.2008

The 28th Conference on Computer Communications, INFOCOM 2009, Valentin Polishchuk, 01.01.2008 → 31.12.2008

Mikko Sillanpää ,

Theoretical and Applied Genetics, Mikko Sillanpää, 2006 → ...

BMC Genetics, Mikko Sillanpää, 2009 → ...

Frontiers in Livestock Genomics, Mikko Sillanpää, 2010 → ...

Hannu Toivonen ,

Editor: Data Mining and Knowledge Discovery, Hannu Toivonen, 2005 → 2009, Netherlands

Editorial Board Member: International Journal of Data Mining and Bioinformatics, Hannu Toivonen, 2006 → 2009, Switzerland

Editorial Board Member: BioData Mining, Hannu Toivonen, 2008 → 2010, United Kingdom

Associate Editor: Computational Intelligence, Hannu Toivonen, 2009 → ..., United States

Editorial Board Member: Data Mining and Knowledge Discovery, Hannu Toivonen, 2009 → ..., Netherlands

Editorial Board Member: Machine Learning, Hannu Toivonen, 2010 → ..., Netherlands

Huizhen Yu ,

Discrete Event Dynamic Systems: Theory and Applications, Huizhen Yu, 2010 → 2012

Taneli Johannes Mielikäinen ,

ACM Transactions on Database Systems (TODS), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005

Data Mining and Knowledge Discovery (DMKD), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005

Data and Knowledge Engineering (DKE), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005

Fifth International Conference on Computational Science (ICCS 2005), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005, United States

Fourth International Workshop on Knowledge Discovery in Inductive Databases (KDID 2005), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005, Portugal

IEEE Transactions on Knowledge and Data Engineering (TKDE), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005

International Workshop on Dependability Aspects on Data Warehousing and Mining Applications (DAWAM 2006), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005, Austria

Knowledge and Information Systems (KAIS), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005

Workshop on Privacy and Security Aspects of Data Mining (PSDM 2005), Taneli Johannes Mielikäinen, 01.01.2005 → 31.12.2005, United States

ACM Transactions on Database Systems (TODS), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

ACM/IEEE Transactions on Computational Biology and Bioinformatics (TCBB), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

Data Mining and Knowledge Discovery (DMKD), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

Data and Knowledge Engineering (DKE), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

Discrete Applied Mathematics (DAM), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

Fifth International Workshop on Knowledge Discovery in Inductive Databases (KDID 2006), Taneli Johannes Mielikäinen, 18.09.2006 → 31.12.2006, Germany

IEEE Transactions on Knowledge and Data Engineering (TKDE), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Information Sciences (INS), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

International Workshop on Dependability Aspects on Data Warehousing and Mining Applications (DAWAM 2006), Taneli Johannes Mielikäinen, 20.04.2006 → 22.04.2006, Austria

The 17th European Conference on Machine Learning and the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2006), Berlin, Taneli Johannes Mielikäinen, 18.09.2006 → 22.09.2006, Germany

2007 ACM Symposium on Applied Computing (SAC 2007), Special Track on Data Mining (DM), Taneli Johannes Mielikäinen, 11.03.2007 → 15.03.2007, South Korea

Editor of research anthology/collection/conference proceedings

Samuel Kaski ,

KRBIO05, Symposium on Knowledge Representation in Bioinformatics, Samuel Kaski, 01.01.2005 → 31.12.2005, Finland

Petri Myllymäki ,

Co-Chair of the First International Workshop on Information Theoretic Methods in Science and Engineering (WITMSE-2008), Petri Myllymäki, 01.01.2008 → 31.12.2008, Finland

Editor of the book "Festschrift in Honour of Jorma Rissanen", Petri Myllymäki, 01.01.2008 → 31.12.2008

PC Chair of The Fifth European Workshop on Probabilistic Graphical Models (PGM-2010), Petri Myllymäki, 01.01.2008 → 31.12.2008, Finland

PC Chair of the International Conference on Uncertainty in Artificial Intelligence (UAI2008), Petri Myllymäki, 01.01.2008 → 31.12.2008, United States

Co-Chair of the Second International Workshop on Information Theoretic Methods in Science and Engineering (WITMSE-2009), Petri Myllymäki, 01.01.2009 → 31.12.2009, Finland

Co-Chair of the Third International Workshop on Information Theoretic Methods in Science and Engineering (WITMSE-2010), Petri Myllymäki, 01.01.2010 → 31.12.2010, Finland

Hannu Toivonen ,

PC Member: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Hannu Toivonen, 2005

PC Member: International Conference on Machine Learning (ICML), Hannu Toivonen, 2005

PC Member: SIAM International Conference on Data Mining (SDM), Hannu Toivonen, 2005

PC Area Chair: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Hannu Toivonen, 2006

PC Member: IEEE International Conference on Data Mining (ICDM), Hannu Toivonen, 2006

PC Member: International Symposium on Computational Life Science (CompLife), Hannu Toivonen, 2006

PC Member: SIAM International Conference on Data Mining (SDM), Hannu Toivonen, 2006

Senior PC Member: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Hannu Toivonen, 2006

Senior PC Member: International Conference on Machine Learning (ICML), Hannu Toivonen, 2006

PC Area Chair: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Hannu Toivonen, 2007

PC Member: ACM Conference on Information and Knowledge Management (CIKM), Hannu Toivonen, 2007

PC Member: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Hannu Toivonen, 2007

PC Member: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hannu Toivonen, 2007

PC Member: IEEE International Conference on Data Mining (ICDM), Hannu Toivonen, 2007

PC Member: IEEE Symposium on Computational Intelligence and Data Mining (CIDM), Hannu Toivonen, 2007

PC Member: International Conference on Data Engineering (ICDE), Hannu Toivonen, 2007

PC Member: International Symposium on Intelligent Data Analysis (IDA), Hannu Toivonen, 2007

PC Member: Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), Hannu Toivonen, 2007

PC Member: SIAM International Conference on Data Mining (SDM), Hannu Toivonen, 2007

PC Area Chair: Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), Hannu Toivonen, 2008



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

PC Member: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Hannu Toivonen, 2008
PC Member: ACM Symposium on Applied Computing (ACM SAC), Hannu Toivonen, 2008
PC Member: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Hannu Toivonen, 2008
PC Member: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hannu Toivonen, 2008
PC Member: International Symposium on Bioinformatics Research and Applications (ISBRA), Hannu Toivonen, 2008
PC Member: SIAM International Conference on Data Mining (SDM), Hannu Toivonen, 2008
PC Member: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hannu Toivonen, 2009
PC Member: International Conference on Machine Learning (ICML), Hannu Toivonen, 2009
PC Member: International Symposium on Intelligent Data Analysis (IDA), Hannu Toivonen, 2009
PC Vice Chair: IEEE International Conference on Data Mining (ICDM), Hannu Toivonen, 2009
Senior PC Member: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Hannu Toivonen, 2009
PC Member: AAAI Conference on Artificial Intelligence (AAAI), Hannu Toivonen, 2010
PC Member: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Hannu Toivonen, 2010
PC Member: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hannu Toivonen, 2010
PC Member: International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Hannu Toivonen, 2010
PC Member: International Conference on Machine Learning (ICML), Hannu Toivonen, 2010
PC Member: International Symposium on Intelligent Data Analysis (IDA), Hannu Toivonen, 2010
PC Member: SIAM International Conference on Data Mining (SDM), Hannu Toivonen, 2010
PC Vice Chair: IEEE International Conference on Data Mining (ICDM), Hannu Toivonen, 2010

Teemu Teppo Roos ,

Proceedings of the Fifth European Workshop on Probabilistic Graphical Models, Teemu Teppo Roos, 09.2010, Finland

Peer review of manuscripts

Esko Ukkonen ,

Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications, Esko Ukkonen, 2010, United States

Patrik Hoyer ,

IEEE Transactions on Knowledge and Data Engineering, Patrik Hoyer, 2004 → ...
IEEE Transactions on Pattern Analysis and Machine Intelligence, Patrik Hoyer, 2004 → ...
Neural Computation, Patrik Hoyer, 2004 → ...
IEEE Signal Processing Letters, Patrik Hoyer, 01.01.2005 → 31.12.2005
IEEE Trans. on Knowledge and Data Engineering, Patrik Hoyer, 01.01.2005 → 31.12.2005
IEEE Transactions on Image Processing, Patrik Hoyer, 01.01.2005 → 31.12.2005
IEEE Transactions on Neural Networks, Patrik Hoyer, 01.01.2005 → 31.12.2005
IEEE Transactions on Speech and Audio, Patrik Hoyer, 01.01.2005 → 31.12.2005
Image and Vision Computing, Patrik Hoyer, 01.01.2005 → 31.12.2005
International Journal of Neural Systems, Patrik Hoyer, 01.01.2005 → 31.12.2005
International Journal of Pattern Recognition and Artificial Intelligence, Patrik Hoyer, 01.01.2005 → 31.12.2005
Journal of Machine Learning Research, Patrik Hoyer, 01.01.2005 → 31.12.2005
Neurocomputing, Patrik Hoyer, 01.01.2005 → 31.12.2005
Bioinformatics, Patrik Hoyer, 01.01.2006 → 31.12.2006
IEEE Transactions on Knowledge and Data Engineering, Patrik Hoyer, 01.01.2006 → 31.12.2006



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

IEEE Transactions on Neural Networks, Patrik Hoyer, 01.01.2006 → 31.12.2006
IEEE Transactions on Pattern Analysis and Machine Intelligence, Patrik Hoyer, 01.01.2006 → 31.12.2006
International Journal of Neural Systems, Patrik Hoyer, 01.01.2006 → 31.12.2006
Journal of Machine Learning Research, Patrik Hoyer, 2006 → ...
Journal of Machine Learning Research, Patrik Hoyer, 01.01.2006 → 31.12.2006
Machine Learning, Patrik Hoyer, 01.01.2006 → 31.12.2006
Neural Computation, Patrik Hoyer, 01.01.2006 → 31.12.2006
Neurocomputing, Patrik Hoyer, 01.01.2006 → 31.12.2006
Pattern Recognition Letters, Patrik Hoyer, 01.01.2006 → 31.12.2006

Aapo Hyvärinen ,

Area Chair in Program Committee of NIPS2009 conference, Aapo Hyvärinen, 2009
Program Committee Member: COSYNE2010 conference, Aapo Hyvärinen, 2010

Juha Kärkkäinen ,

Reviewer for Algorithmica, Juha Kärkkäinen, 2005
Reviewer for IEEE Transactions on Knowledge and Data Engineering, Juha Kärkkäinen, 2005
Reviewer for Information and Computation, Juha Kärkkäinen, 2005
PC Member: 13th Symposium on String Processing and Information Retrieval, SPIRE 2006, Juha Kärkkäinen, 2006
Reviewer for ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, Juha Kärkkäinen, 2006
Reviewer for Algorithmica, Juha Kärkkäinen, 2006
Reviewer for EURASIP Journal on Bioinformatics and Systems Biology, Juha Kärkkäinen, 2006
Reviewer for International Colloquium on Automata, Languages and Programming, ICALP 2006, Juha Kärkkäinen, 2006
Reviewer for Journal of Discrete Algorithms, Juha Kärkkäinen, 2006
Reviewer for Nordic Journal of Computing, Juha Kärkkäinen, 2006
Reviewer for Software: Practice and Experience, Juha Kärkkäinen, 2006
Reviewer for Theoretical Computer Science, Juha Kärkkäinen, 2006
PC Member: 14th Symposium on String Processing and Information Retrieval, SPIRE 2007, Juha Kärkkäinen, 2007
Reviewer for ACM Journal of Experimental Algorithmics, Juha Kärkkäinen, 2007
Reviewer for Algorithmica, Juha Kärkkäinen, 2007
Reviewer for Combinatorial Pattern Matching, CPM 2007, Juha Kärkkäinen, 2007
Reviewer for ECML/PKDD 2007, Juha Kärkkäinen, 2007
Reviewer for European Symposium on Algorithms, ESA 2007, Juha Kärkkäinen, 2007
Reviewer for Journal of Discrete Algorithms, Juha Kärkkäinen, 2007
Reviewer for Mathematics in Computer Science, Juha Kärkkäinen, 2007
Reviewer for Theoretical Computer Science, Juha Kärkkäinen, 2007
Reviewer for Workshop on Algorithm Engineering & Experiments, ALENEX 2008, Juha Kärkkäinen, 2007
Reviewer for Workshop on Algorithms and Data Structures, WADS 2007, Juha Kärkkäinen, 2007
Reviewer for SOFSEM 2009, Juha Kärkkäinen, 2008
Reviewer for Scandinavian Workshop on Algorithm Theory, SWAT 2008, Juha Kärkkäinen, 2008
Reviewer for String Processing and Information Retrieval, SPIRE 2008, Juha Kärkkäinen, 2008
Reviewer for Symposium on Theoretical Aspects of Computer Science, STACS 2009, Juha Kärkkäinen, 2008
Reviewer for Workshop on Algorithm Engineering & Experiments, ALENEX 2009, Juha Kärkkäinen, 2008



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

PC Member: 17th Annual European Symposium on Algorithms, ESA 2009, Juha Kärkkäinen, 2009

Reviewer for ACM Journal of Experimental Algorithmics, Juha Kärkkäinen, 2009

Reviewer for ACM Symposium on Theory of Computing, STOC 2009, Juha Kärkkäinen, 2009

Reviewer for Combinatorial Pattern Matching, CPM 2009, Juha Kärkkäinen, 2009

Reviewer for Discrete Mathematics & Theoretical Computer Science, Juha Kärkkäinen, 2009

Reviewer for ISMB/ECCB 2009, Juha Kärkkäinen, 2009

Reviewer for Journal of Discrete Algorithms, Juha Kärkkäinen, 2009

Reviewer for String Processing and Information Retrieval, SPIRE 2009, Juha Kärkkäinen, 2009

Reviewer for Symposium on Theoretical Aspects of Computer Science, STACS 2010, Juha Kärkkäinen, 2009

Reviewer for Theoretical Computer Science, Juha Kärkkäinen, 2009

Reviewer for Workshop on Algorithms and Applications, WAA 2010, Juha Kärkkäinen, 2009

PC Member: 12th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2010, Juha Kärkkäinen, 2010

Reviewer for European Symposium on Algorithms, ESA 2010, Juha Kärkkäinen, 2010

Reviewer for IEEE Symposium on Foundations of Computer Science, FOCS 2010, Juha Kärkkäinen, 2010

Reviewer for Information processing Letters, Juha Kärkkäinen, 2010

Reviewer for International Colloquium on Automata, Languages and Programming, ICALP 2010, Juha Kärkkäinen, 2010

Reviewer for String Processing and Information Retrieval, SPIRE 2010, Juha Kärkkäinen, 2010

Reviewer for Workshop on Algorithm Engineering & Experiments, ALENEX 2011, Juha Kärkkäinen, 2010

Samuel Kaski ,

PC Member, Machine Learning for Signal Processing, Samuel Kaski, 01.01.2005 → 31.12.2005, France

PC Member: 4th International Workshop on Web Semantics (WebS 2005), Samuel Kaski, 01.01.2005 → 31.12.2005, Denmark

PC Member: AWIC'05, 2005 Atlantic Web Intelligence Conference, Samuel Kaski, 01.01.2005 → 31.12.2005, Poland

PC Member: IDEAL'05, International Symposium on Intelligent Data Engineerin, Samuel Kaski, 01.01.2005 → 31.12.2005, Australia

PC Member: IIPWM'05, New Trends in Intelligent Information Processing and Web Mining, Samuel Kaski, 01.01.2005 → 31.12.2005, Poland

PC Member: International Conference on Natural Computation (ICNC'05), Samuel Kaski, 01.01.2005 → 31.12.2005, China

PC Member: MLSP'05, Machine Learning for Signal Processing, Samuel Kaski, 01.01.2005 → 31.12.2005, United States

PC Member: PKDD'05, European Conference on Principles and Practice of Knowledge Discovery in Databases, Samuel Kaski, 01.01.2005 → 31.12.2005, Portugal

PC Member: WSOM 05, 5th Workshop On Self-Organizing Maps, Samuel Kaski, 01.01.2005 → 31.12.2005, France

Proceedings of Workshop at UM'05, Machine Learning for User Modeling: Challenges, Samuel Kaski, 01.01.2005 → 31.12.2005, United Kingdom

IEEE PAMI, Samuel Kaski, 2010

Journal for Artificial Intelligence Research, Samuel Kaski, 2010

Neural Processing Letters, Samuel Kaski, 2010

PC Chair: ICML2010 Workshop on Reinforcement learning and search in very large spaces, Samuel Kaski, 2010

PC Chair: Machine Learning for Signal Processing (MLSP 2010), Samuel Kaski, 2010

PC Member: Asia-Pacific Bioinformatics Conference (APBC 2010), Samuel Kaski, 2010

PC Member: European Conference on Machine Learning and European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2010), Samuel Kaski, 2010

PC Member: European Symposium on Artificial Neural Networks (ESANN 2010), Samuel Kaski, 2010

PC Member: International Conference on Machine Learning (ICML 2010), Samuel Kaski, 2010

PC Member: International Conference on Pattern Recognition (ICPR 2010), Samuel Kaski, 2010



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

PC Member: Mining and Learning with Graphs (MLG 2010), Samuel Kaski, 2010

PC Member: the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2010), Samuel Kaski, 2010

PC Vice Chair: 2010 IEEE/WIC/ACM International Conference on Web Intelligence (WI'10), Samuel Kaski, 2010

Peer reviewer for several journals and conferences, Samuel Kaski, 2010

Scholarpedia, Samuel Kaski, 2010

Jyrki Kivinen ,

IEEE Transactions on Information Theory, Jyrki Kivinen, 01.01.2005 → 31.12.2005, United States

IEEE Transactions on Signal Processing, Jyrki Kivinen, 01.01.2005 → 31.12.2005, United States

Journal of Machine Learning Research, Jyrki Kivinen, 01.01.2005 → 31.12.2005, United States

Proceedings of the 22nd International Machine Learning Conference, Jyrki Kivinen, 01.01.2005 → 31.12.2005, United States

Advances in Neural Information Processing Systems 19, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Data Mining and Knowledge Discovery, Jyrki Kivinen, 01.01.2006 → 31.12.2006

IEEE Signal Processing Letters, Jyrki Kivinen, 01.01.2006 → 31.12.2006

IEEE Transactions on Information Theory, Jyrki Kivinen, 01.01.2006 → 31.12.2006

IEEE Transactions on Signal Processing, Jyrki Kivinen, 01.01.2006 → 31.12.2006

IEEE Transactions on Systems, Man and Cybernetics - Part, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Journal of Machine Learning Research, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Machine Learning, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Proceedings of the 19th International Conference on Learning Theory, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Theoretical Computer Science, Jyrki Kivinen, 01.01.2006 → 31.12.2006

Advances in Neural Information Processing Systems 20, Jyrki Kivinen, 01.01.2007 → 31.12.2007

Data Mining and Knowledge Discovery, Jyrki Kivinen, 01.01.2007 → 31.12.2007

Discrete Applied Mathematics, Jyrki Kivinen, 01.01.2007 → 31.12.2007

IEEE Signal Processing Letters, Jyrki Kivinen, 01.01.2007 → 31.12.2007

IEEE Transactions on Pattern Analysis and Machine Intelligence, Jyrki Kivinen, 01.01.2007 → 31.12.2007

IEEE Transactions on Signal Processing, Jyrki Kivinen, 01.01.2007 → 31.12.2007

Journal of Machine Learning Research, Jyrki Kivinen, 01.01.2007 → 31.12.2007

Proceedings of the 18th International Conference on Algorithmic Learning Theory, Jyrki Kivinen, 01.01.2007 → 31.12.2007

Advances in Neural Information Processing Systems 21, Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

IEEE Transactions on Pattern Analysis and Machine Intelligence, Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

IEEE Transactions on Signal Processing, Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

IEEE Transactions on Systems, Man and Cybernetics - Part B, Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

Journal of Machine Learning Research, Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

Proceedings of the 21st Annual Conference on Learning Theory (COLT 2008), Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

Proceedings of the 25th International Conference on Machine Learning (ICML 2008), Jyrki Kivinen, 01.01.2008 → 31.12.2008, United States

Theoretical Computer Science, Jyrki Kivinen, 01.01.2008 → 31.12.2008, Netherlands

Algorithms, Jyrki Kivinen, 14.12.2009

ALT 2010 - the 21st International Conference on Algorithmic Learning Theory, Jyrki Kivinen, 31.05.2010 → 07.06.2010

IEEE Signal Processing Letters, Jyrki Kivinen, 14.05.2010, United States

IEEE Transactions on Neural Networks, Jyrki Kivinen, 15.05.2010, United States

IEEE Transactions on Signal Processing, Jyrki Kivinen, 16.11.2010, United States



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

IEEE Transactions on Signal Processing, Jyrki Kivinen, 16.10.2010, United States

Journal of Machine Learning Research, Jyrki Kivinen, 31.03.2010

Knowledge and Information Systems, Jyrki Kivinen, 25.06.2010

Neurocomputing, Jyrki Kivinen, 16.07.2010

Theoretical Computer Science, Jyrki Kivinen, 03.02.2010

Mikko Koivisto ,

Reviewer for ECAI 2004, Mikko Koivisto, 2004 → ...

Reviewer for WABI 2004, Mikko Koivisto, 2004 → ...

Reviewer for ALT 2005, Mikko Koivisto, 2005 → ...

Reviewer for CPM 2005, Mikko Koivisto, 2005 → ...

Reviewer for PSB 2005, Mikko Koivisto, 2005 → ...

Reviewer for Bioinformatics, Mikko Koivisto, 2006 → ...

Reviewer for Computational Statistics & Data Analysis, Mikko Koivisto, 2006 → ...

Reviewer for PSM 2006, Mikko Koivisto, 2006 → ...

Reviewer for RECOMB 2006, Mikko Koivisto, 2006 → ...

Reviewer for UAI 2006, Mikko Koivisto, 2006 → ...

Reviewer for AISTATS 2007, Mikko Koivisto, 2007 → ...

Reviewer for European Journal of Human Genetics, Mikko Koivisto, 2007 → ...

Reviewer for Journal of Discrete Algorithms, Mikko Koivisto, 2007 → ...

Reviewer for RECOMB 2007, Mikko Koivisto, 2007 → ...

PC member for SWAT 2008, Mikko Koivisto, 2008 → ...

Reviewer for Artificial Intelligence in Medicine, Mikko Koivisto, 2008 → ...

Reviewer for Journal of Machine Learning Research, Mikko Koivisto, 2008 → ...

Reviewer for UAI 2008, Mikko Koivisto, 2008 → ...

reviewer ECAI 2008, Mikko Koivisto, 2008 → ...

Reviewer CPM 2009, Mikko Koivisto, 2009 → ...

Reviewer for BMC Genetics, Mikko Koivisto, 2009 → ...

Reviewer for ICALP 2009, Mikko Koivisto, 2009 → ...

Reviewer for Information Processing Letters, Mikko Koivisto, 2009 → ...

Reviewer for International Journal on Artificial Intelligence Tools, Mikko Koivisto, 2009 → ...

Reviewer for Journal of Machine Learning Research, Mikko Koivisto, 2009 → ...

Reviewer for Theoretical Computer Science, Mikko Koivisto, 2009 → ...

PC member for ESA 2010, Mikko Koivisto, 2010 → ...

PC member for PGM 2010, Mikko Koivisto, 2010 → ...

Reviewer for Algorithmica, Mikko Koivisto, 2010 → ...

Reviewer for Discrete Applied Mathematics, Mikko Koivisto, 2010 → ...

Reviewer for UAI 2010, Mikko Koivisto, 2010 → ...

Reviewer for WG 2010, Mikko Koivisto, 2010 → ...

Kjell Lemström ,

Reviewer for Software-Practice & Experience, Kjell Lemström, 2005

Reviewer for the 6th International Conference on Music Information Retrieval, Kjell Lemström, 2005



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Reviewer for the Computer Music Journal, Kjell Lemström, 2005
Reviewer for the Netherlands Organisation for Scientific Research (NWO), Kjell Lemström, 2005
Program Co-Chair of the 7th International Society for Music Information Retrieval Conference, Kjell Lemström, 2006, Canada
Reviewer for Musicae Scientiae, Kjell Lemström, 2006
Reviewer for the Foundations and Trends in Information Retrieval, Kjell Lemström, 2006
Reviewer for the IEEE Transactions on Multimedia, Kjell Lemström, 2006
Book review for Springer, Kjell Lemström, 2007
Reviewer for the IEEE Transactions on Audio, Speech and Language Processing, Kjell Lemström, 2007
Reviewer for Signal Processing, Kjell Lemström, 2008
Reviewer for Algorithms and Applications, LNCS 6060, Kjell Lemström, 2009
Late Breaking and Demo Chair of the 11th International Society for Music Information Retrieval Conference, Kjell Lemström, 2010, Netherlands
Reviewer for the International Journal on Digital Libraries, Kjell Lemström, 2010
Reviewer for the Journal of Intelligent Information Systems, Kjell Lemström, 2010

Veli Mäkinen ,

ACM Transactions on Algorithms, Veli Mäkinen, 2005
Algorithmica, Veli Mäkinen, 2005
Annual Symposium on Combinatorial Pattern Matching (CPM 2005), Veli Mäkinen, 2005
IEEE Multimedia, Veli Mäkinen, 2005
IEEE/ACM Transactions on Computational Biology and Bioinformatics, Veli Mäkinen, 2005
Information Processing Letters, Veli Mäkinen, 2005
Intelligent Systems for Molecular Biology (ISMB 2005), Veli Mäkinen, 2005
International Conference on Music Information Retrieval (ISMIR 2005), Veli Mäkinen, 2005
Italian Conference on Theoretical Computer Science (ICTCS 2005), Veli Mäkinen, 2005
Journal of Discrete Algorithms, Veli Mäkinen, 2005
Pattern Recognition, Veli Mäkinen, 2005
Symposium on String Processing and Information Retrieval (SPIRE 2005), Veli Mäkinen, 2005
Annual Symposium on Combinatorial Pattern Matching (CPM 2006), Veli Mäkinen, 2006
BMC Bioinformatics, Veli Mäkinen, 2006
IEEE Transactions on Knowledge and Data Engineering, Veli Mäkinen, 2006
International Workshop on Experimental Algorithms (WEA 2006), Veli Mäkinen, 2006
Symposium on String Processing and Information Retrieval (SPIRE 2006), Veli Mäkinen, 2006
Symposium on Theoretical Aspects of Computer Science (STACS 2007), Veli Mäkinen, 2006
Tenth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2006), Veli Mäkinen, 2006
Theoretical Computer Science, Veli Mäkinen, 2006
Workshop on Algorithm Engineering and Experimentation (ALENEX 2007), Veli Mäkinen, 2006
Data and Knowledge Engineering, Veli Mäkinen, 2007
Fundamenta Informaticae, Veli Mäkinen, 2007
IEEE Transaction on Knowledge Discovery and Data Engineering, Veli Mäkinen, 2007
Information Processing Letters, Veli Mäkinen, 2007
Information Retrieval, Veli Mäkinen, 2007
Information and Computation, Veli Mäkinen, 2007



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

International Conference on Music Information Retrieval 2007 (ISMIR 2007), Veli Mäkinen, 2007
International Journal on Bioinformatics Research and Applications, Veli Mäkinen, 2007
Latin American Theoretical Informatics (LATIN 2008), Veli Mäkinen, 2007
Mathematical Foundations of Computer Science (MFCS 2007), Veli Mäkinen, 2007
Symposium on Discrete Algorithms (SODA 2008), Veli Mäkinen, 2007
Symposium on String Processing and Information Retrieval (SPIRE 2007), Veli Mäkinen, 2007
Symposium on Theoretical Aspects of Computer Science (STACS 2008), Veli Mäkinen, 2007
Workshop on Experimental Algorithms (WEA 2008), Veli Mäkinen, 2007
Algorithmica, Veli Mäkinen, 2008
Electronic Journal of Combinatorics, Veli Mäkinen, 2008
European Symposium on Algorithms (ESA 2008), Veli Mäkinen, 2008
IEEE/ACM Transactions on Computational Biology and Bioinformatics, Veli Mäkinen, 2008
Information & Computation, Veli Mäkinen, 2008
International Colloquium on Automata, Languages and Programming (ICALP 2008), Veli Mäkinen, 2008
International Conference on Music Information Retrieval (ISMIR 2008), Veli Mäkinen, 2008
International Symposium on Information Theory and its Applications (ISITA 2008), Veli Mäkinen, 2008
Journal of Discrete Algorithms, Veli Mäkinen, 2008
Pattern Recognition in Bioinformatics (PRIB 2008), Veli Mäkinen, 2008
Symposium on Combinatorial Pattern Matching (CPM 2008), Veli Mäkinen, 2008
Theoretical Computer Science, Veli Mäkinen, 2008
16th String Processing and Information Retrieval Symposium (SPIRE 2009), Veli Mäkinen, 2009
ACM Transactions on Database Systems, Veli Mäkinen, 2009
Algorithmica, Veli Mäkinen, 2009
Bioinformatics, Veli Mäkinen, 2009
Computational Intelligence, Veli Mäkinen, 2009
Data Compression Conference (DCC 2010), Veli Mäkinen, 2009
European Symposium on Algorithms (ESA 2009), Veli Mäkinen, 2009
IEEE Transactions on Knowledge and Data Engineering (TKDE), Veli Mäkinen, 2009
IEEE/ACM Transactions on Computational Biology and Bioinformatics, Veli Mäkinen, 2009
International Conference on Music Information Retrieval (ISMIR 2009), Veli Mäkinen, 2009
Journal of Computer and System Sciences, Veli Mäkinen, 2009
Journal of Experimental Algorithmics, Veli Mäkinen, 2009
Software: Practice and Experience, Veli Mäkinen, 2009
Symposium on Theoretical Aspects of Computer Science (STACS 2009), Veli Mäkinen, 2009
Theoretical Computer Science, Veli Mäkinen, 2009
12th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2010), Veli Mäkinen, 2010
17th String Processing and Information Retrieval Symposium (SPIRE 2010), Veli Mäkinen, 2010
18th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2010), Veli Mäkinen, 2010
European Symposium on Algorithms (ESA 2009), Veli Mäkinen, 2010
IEEE/ACM Transactions on Computational Biology and Bioinformatics, Veli Mäkinen, 2010
Information & Computation, Veli Mäkinen, 2010



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

International Conference on Bioinformatics Models, Methods and Algorithms (BIOINFORMATICS 2011), Veli Mäkinen, 2010

International Conference on Music Information Retrieval (ISMIR 2010), Veli Mäkinen, 2010

Journal of Discrete Algorithms, Veli Mäkinen, 2010

Theoretical Computer Science, Veli Mäkinen, 2010

Petri Myllymäki ,

PC member, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning, Petri Myllymäki, 15.06.2005 → 17.06.2005, Finland

PC member, NIPS 2005 Workshop on Machine Learning for Implicit Feedback and User Modeling, Petri Myllymäki, 05.12.2005 → 11.12.2005, Canada

PC Member of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI 2006), Petri Myllymäki, 01.01.2006 → 31.12.2006, United States

PC member, IJCAI Workshop On Cross Lingual Information Access: Addressing the Information Need of Multilingual Societies (CLIA-2007), Petri Myllymäki, 01.01.2006 → 31.12.2006

PC member, International Workshop on Intelligent Information Access (IIIA-2006), Petri Myllymäki, 01.01.2006 → 31.12.2006

Reviewer for journal IEEE Transactions on Vehicular Technology, Petri Myllymäki, 01.01.2006 → 31.12.2006

Reviewer for the 2006 IEEE Information Theory Workshop, Petri Myllymäki, 13.03.2006 → 17.03.2006, Uruguay

Reviewer for the 8th European Conference on Case-Based Reasoning, Petri Myllymäki, 01.01.2006 → 31.12.2006, Turkey

Artificial Intelligence and Statistics 2007, Petri Myllymäki, 01.01.2007 → 31.12.2007

PC member, International Conference on Uncertainty in Artificial Intelligence (UAI-2007), Petri Myllymäki, 01.01.2007 → 31.12.2007

PC member, Workshop on Cross-Lingual Information Access, Petri Myllymäki, 01.01.2007 → 31.12.2007

PC member, Workshop on Mobile User Improved Interaction, Petri Myllymäki, 01.01.2007 → 31.12.2007

Reviewer for journal IEEE Transactions on Vehicular Technology, Petri Myllymäki, 01.01.2007 → 31.12.2007

Reviewer for journal Wireless Communications and Mobile Computing, Petri Myllymäki, 01.01.2007 → 31.12.2007

Reviewer for the journal IEEE Transactions on Information Theory, Petri Myllymäki, 01.01.2007 → 31.12.2007

Area chair, European Conference on Artificial Intelligence (ECAI-08), Petri Myllymäki, 01.01.2008 → 31.12.2008

PC member, Workshop on Cross-Lingual Information Access (CLIA-2008), Petri Myllymäki, 01.01.2008 → 31.12.2008

Reviewer for journal "Statistics and Computing", Petri Myllymäki, 01.01.2008 → 31.12.2008

Reviewer for journal IEEE Transactions on Vehicular Technology, Petri Myllymäki, 01.01.2008 → 31.12.2008

Reviewer for the 2008 IEEE International Symposium on Information Theory (ISIT), Petri Myllymäki, 01.01.2008 → 31.12.2008

PC member, International Joint Conference on Artificial Intelligence (IJCAI-2009), Petri Myllymäki, 2009

Reviewer for the Journal of Machine Learning Research, Petri Myllymäki, 2010

Valentin Polishchuk ,

PC Member: Symposium on Computational Geometry (SoCG 2010), Valentin Polishchuk, 02.12.2009 → 14.02.2010

Juho Rousu ,

Bioinformatics, Juho Rousu, 22.05.2005 → 05.06.2005

IEEE Transactions on Knowledge and Data engineering, Juho Rousu, 27.08.2005 → 14.10.2005

Journal of Machine Learning Research, Juho Rousu, 14.10.2005 → 17.11.2005

Data mining and Knowledge Discovery, Juho Rousu, 19.07.2006 → 11.08.2006

IEEE Transactions of Neural Networks, Juho Rousu, 12.11.2006 → 27.03.2007

Journal of Machine Learning Research, Juho Rousu, 14.07.2006 → 08.08.2008

Neurocomputing, Juho Rousu, 28.12.2006 → 19.02.2007

Pattern Recognition, Juho Rousu, 30.08.2006 → 27.09.2006

Computational Statistics and Data Analysis, Juho Rousu, 03.01.2007 → 05.01.2007



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

IEEE Transactions on Knowledge and Data engineering, Juho Rousu, 13.02.2007 → 30.03.2007
Pattern Recognition, Juho Rousu, 25.01.2007 → 27.03.2007
Pattern Recognition, Juho Rousu, 24.05.2007 → 05.07.2007
Journal of Machine Learning Research, Juho Rousu, 16.09.2008 → 22.09.2008
Journal of Machine Learning Research, Juho Rousu, 06.04.2008 → 04.05.2008
IEEE Transactions on Knowledge and Data engineering, Juho Rousu, 19.06.2009 → 22.07.2009
Journal of Bioinformatics and Computational Biology, Juho Rousu, 21.08.2009 → 26.08.2009
Journal of Food Engineering, Juho Rousu, 08.08.2009 → 10.10.2009
Pattern Recognition, Juho Rousu, 25.03.2009 → 27.03.2009
Transactions of Computational Biology and Bioinformatics, Juho Rousu, 07.01.2009 → 10.02.2009
Artificial Intelligence, Juho Rousu, 09.06.2010
Journal of Machine Learning Research, Juho Rousu, 08.11.2010
Journal of Machine Learning Research, Juho Rousu, 28.02.2010
Machine Learning, Juho Rousu, 28.12.2010

Mikko Sillanpää ,

Genetics, Mikko Sillanpää, 1998 → ...
American Journal of Human Genetics, Mikko Sillanpää, 2000 → ...
Annals of Human Genetics, Mikko Sillanpää, 2000 → ...
BMC Bioinformatics, Mikko Sillanpää, 2000 → ...
Bioinformatics, Mikko Sillanpää, 2000 → ...
Biometrical Journal, Mikko Sillanpää, 2000 → ...
Biometrics, Mikko Sillanpää, 2000 → ...
Gene, Mikko Sillanpää, 2000 → ...
Genetic Epidemiology, Mikko Sillanpää, 2000 → ...
Genetica, Mikko Sillanpää, 2000 → ...
Genetical Research, Mikko Sillanpää, 2000 → ...
Genetics, Selection, Evolution, Mikko Sillanpää, 2000 → ...
Heredity, Mikko Sillanpää, 2000 → ...
Human Genomics, Mikko Sillanpää, 2000 → ...
Human Heredity, Mikko Sillanpää, 2000 → ...
International Statistical Review, Mikko Sillanpää, 2000 → ...
JABES, Mikko Sillanpää, 2000 → ...
Journal of Animal Breeding and Genetics, Mikko Sillanpää, 2000 → ...
Molecular Breeding, Mikko Sillanpää, 2000 → ...
Plant Breeding, Mikko Sillanpää, 2000 → ...
Scandinavian Journal of Work, Environment & Health, Mikko Sillanpää, 2000 → ...
Statistics in Medicine, Mikko Sillanpää, 2000 → ...
Theoretical and Applied Genetics, Mikko Sillanpää, 2000 → ...
Tree Genetics and Genomics, Mikko Sillanpää, 2000 → ...
Trends in Plant Science, Mikko Sillanpää, 2000 → ...
Signal Processing, Mikko Sillanpää, 2010



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Statistical Modelling, Mikko Sillanpää, 2010

Hannu Toivonen ,

Reviewer: Acta Informatica, Hannu Toivonen, 2005

Reviewer: Bioinformatics, Hannu Toivonen, 2005

Reviewer: Data Mining and Knowledge Discovery, Hannu Toivonen, 2005

Reviewer: Fuzzy Sets and Systems, Hannu Toivonen, 2005

Reviewer: IEEE Transactions on Knowledge and Data Engineering, Hannu Toivonen, 2005

Reviewer: IEEE Transactions on Knowledge and Data Engineering, Hannu Toivonen, 2005

Reviewer: BMC Bioinformatics, Hannu Toivonen, 2006

Reviewer: Data Mining and Knowledge Discovery, Hannu Toivonen, 2006

Reviewer: Data and Knowledge Engineering, Hannu Toivonen, 2006

Reviewer: International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, Hannu Toivonen, 2006

Reviewer: Software: Practice and Experience, Hannu Toivonen, 2006

Reviewer: Bioinformatics, Hannu Toivonen, 2007

Reviewer: IEEE Transactions on Information Technology in Biomedicine, Hannu Toivonen, 2007

Reviewer: BioData Mining, Hannu Toivonen, 2008

Reviewer: Human Heredity, Hannu Toivonen, 2008

Reviewer: Data Mining and Knowledge Discovery, Hannu Toivonen, 2010

Helena Ahonen-Myka ,

Program committee member, International ACM Conference on Research and Development in Information Retrieval (SIGIR), Helena Ahonen-Myka, 2005

Program committee member: ACM SIGIR Workshop ELEKTRA (Methodologies and Evaluation of Lexical Cohesion Techniques in Real-world Applications), 2005, Helena Ahonen-Myka, 2005

Reviewer for International Journal of Neural Systems, Helena Ahonen-Myka, 2005

Reviewer for Web Intelligence and Agent Systems Journal, Helena Ahonen-Myka, 2005

Program committee member, International ACM Conference on Research and Development in Information Retrieval (SIGIR), Helena Ahonen-Myka, 2006

Reviewer for Data Warehousing and Knowledge Discovery conference (DaWaK 2006), Helena Ahonen-Myka, 2006

Reviewer for International Journal of Artificial Intelligence Tools, Helena Ahonen-Myka, 2006

Reviewer for International Journal of Neural Systems, Helena Ahonen-Myka, 2006

Program committee member, International ACM Conference on Research and Development in Information Retrieval (SIGIR), Helena Ahonen-Myka, 2007

Program committee member: The 16th Nordic Conference of Computational Linguistics (NODALIDA 2007), Helena Ahonen-Myka, 2007, Estonia

Reviewer for IEEE Transactions on Systems, Man, and Cybernetics, Helena Ahonen-Myka, 2010, United States

Ella Bingham ,

PC Member: Data Warehousing and Knowledge Discovery (DaWaK) 2005, Ella Bingham, 2005 → ..., Denmark

KDD 2006, Ella Bingham, 01.01.2006 → 31.12.2006

Neurocomputing, Ella Bingham, 01.01.2006 → 31.12.2006

PC member: International Conference on Knowledge Discovery and Data Mining (KDD) 2006, Ella Bingham, 2006 → ...

IEEE Signal Processing Letters, Ella Bingham, 01.01.2007 → 31.12.2007

IEEE Transactions on Knowledge and Data Engineering, Ella Bingham, 01.01.2007 → 31.12.2007

IEEE Transactions on Signal Processing, Ella Bingham, 01.01.2007 → 31.12.2007



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

International Journal of Neural Systems, Ella Bingham, 01.01.2007 → 31.12.2007

Neurocomputing, Ella Bingham, 01.01.2007 → 31.12.2007

PLoS Computational Biology, Ella Bingham, 01.01.2007 → 31.12.2007

SDM 2008, Proceedings, Ella Bingham, 01.01.2007 → 31.12.2007

ACM Transactions on Knowledge Discovery from Data, Ella Bingham, 01.01.2008 → 31.12.2008

Complex-Valued Neural Networks, Ella Bingham, 01.01.2008 → 31.12.2008

ECML/PKDD 2008, Ella Bingham, 01.01.2008 → 31.12.2008

IEEE Transactions on Audio, Speech and Language Processing, Ella Bingham, 01.01.2008 → 31.12.2008

IEEE Transactions on Image Processing, Ella Bingham, 01.01.2008 → 31.12.2008

PC member: European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML PKDD) 2008, Ella Bingham, 2008 → ...

PC member: SIAM Data Mining Conference (SDM) 2010, Ella Bingham, 2010

Matti Järvisalo ,

Reviewer for AAAI Conference on Artificial Intelligence (AAAI 2010), Matti Järvisalo, 2010, United States

Reviewer for AI Communications, Matti Järvisalo, 2010

Reviewer for Artificial Intelligence Journal, Matti Järvisalo, 2010

Reviewer for European Conference on Logics for Artificial Intelligence (JELIA 2010), Matti Järvisalo, 2010

Reviewer for International Conference on Formal Methods in Computer-Aided Design (FMCAD 2010), Matti Järvisalo, 2010

Reviewer for International Conference on Theory and Applications of Satisfiability Testing (SAT 2010), Matti Järvisalo, 2010

Reviewer for International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2010), Matti Järvisalo, 2010

Teemu Kivioja ,

review of a manuscript for IEEE/ACM Transactions on Computational Biology and Bioinformatics, Teemu Kivioja, 2010 → ...

Leena Salmela ,

International Journal of Computer Mathematics reviewer, Leena Salmela, 21.09.2009 → ..., United Kingdom

Frontiers in Computer Science in China reviewer, Leena Salmela, 25.02.2010 → ..., Germany

IPL reviewer, Leena Salmela, 03.01.2010 → ..., Netherlands

Journal of Information Science reviewer, Leena Salmela, 22.03.2010 → ..., United Kingdom

SWAT 2010 reviewer, Leena Salmela, 21.06.2010 → 23.06.2010, Germany

TCBB reviewer, Leena Salmela, 03.01.2010 → ..., United States

Stefan Schöner ,

Pattern Recognition Letters, Stefan Schöner, 10.03.2009

Huizhen Yu ,

IEEE Transactions on Automatic Control, Huizhen Yu, 2010

Journal of Machine Learning Research, Huizhen Yu, 2010

Mathematics of Operations Research, Huizhen Yu, 2010

Petri Kontkanen ,

Ohjelmatoimikunnan jäsen, Petri Kontkanen, 2005 → 2010

Arviotsija, Petri Kontkanen, 2006

Arviotsija, Petri Kontkanen, 2007

Arviotsija, Petri Kontkanen, 2008

Arviotsija, Petri Kontkanen, 2009



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Arvioitsija, Petri Kontkanen, 2009

Arvioitsija, Petri Kontkanen, 2010

Arvioitsija, Petri Kontkanen, 2010

Ohjelmatoimikunnan jäsen, Petri Kontkanen, 2010

Pauli Aleksi Miettinen ,

IEEE Transactions on Knowledge and Data Engineering, Pauli Aleksi Miettinen, 2008

Data Mining and Knowledge Discovery, Pauli Aleksi Miettinen, 2009

IEEE Transactions on Pattern Analysis and Machine Intelligence, Pauli Aleksi Miettinen, 2009

Teemu Teppo Roos ,

IEEE Transactions on Pattern Analysis and Machine Intelligence, Teemu Teppo Roos, 01.01.2007 → 31.12.2007, United States

Journal of Machine Learning Research, Teemu Teppo Roos, 01.01.2007 → 31.12.2010, United States

Machine Learning, Teemu Teppo Roos, 01.01.2007 → 31.12.2007

Proceedings of the IEEE Information Theory Workshop 2007, Teemu Teppo Roos, 2007, United States

Proceedings of the IEEE International Symposium on Information Theory 2007, Teemu Teppo Roos, 01.01.2007 → 31.12.2007, United States

SIAM Journal on Imaging Sciences, Teemu Teppo Roos, 01.01.2007 → 31.12.2007, United States

24th Conference on Uncertainty in Artificial Intelligence (UAI2008), Teemu Teppo Roos, 09.07.2008 → 12.07.2008, Finland

Data and Knowledge Engineering, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, Netherlands

IEEE Transactions on Signal Processing, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, United States

IEEE Transactions on Wireless Communications, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, United States

Journal of Information Science, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, United Kingdom

Pattern Recognition Letters, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, Netherlands

Statistica Sinica, Teemu Teppo Roos, 01.01.2008 → 31.12.2008, China

The 18th European Conference on Artificial Intelligence (ECAI2008), Teemu Teppo Roos, 21.07.2008 → 25.07.2008, Greece

The First Workshop on Information Theoretic Methods in Science and Engineering (WITMSE2008), Teemu Teppo Roos, 18.08.2008 → 20.08.2008, Finland

The Fourth European Workshop on Probabilistic Graphical Models (PGM2008), Teemu Teppo Roos, 17.09.2008 → 19.09.2008, Denmark

25th Conference on Uncertainty in Artificial Intelligence (UAI2009), Teemu Teppo Roos, 18.06.2009 → 21.06.2009, Canada

GENETICS, Teemu Teppo Roos, 2009

IEEE Transactions on Information Theory, Teemu Teppo Roos, 2009 → ..., United States

Journal of Applied Probability, Teemu Teppo Roos, 2009 → ...

The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD2009), Teemu Teppo Roos, 07.09.2009 → 11.09.2009, Slovenia

The Second Workshop on Information Theoretic Methods in Science and Engineering (WITMSE2009), Teemu Teppo Roos, 17.08.2009 → 19.08.2009, Finland

26th Conference on Uncertainty in Artificial Intelligence (UAI2010), Teemu Teppo Roos, 08.07.2010 → 11.07.2010, United States

Fifth European Workshop on Probabilistic Graphical Models (PGM2010), Teemu Teppo Roos, 13.09.2010 → 15.09.2010, Finland

IEEE Wireless Communications Magazine, Teemu Teppo Roos, 2010 → ..., United States

Journal of Machine Learning Research, Teemu Teppo Roos, 2010 → ..., United States

Statistics and Computing, Teemu Teppo Roos, 2010 → ..., Netherlands

Teppo Ahonen ,

ISMIR 2010, Teppo Ahonen, 2010



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE
UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Panu Luosto ,

12th Scandinavian Symposium and Workshops on Algorithm Theory, Panu Luosto, 2010

Pekka Parviainen ,

Journal of Machine Learning Research, Pekka Parviainen, 2009

12th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT), Pekka Parviainen, 2010

International Journal of AI Tools (IJAIT), Pekka Parviainen, 2010

The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Pekka Parviainen, 2010

Jouni Sirén ,

Reviewer: 19th Annual Symposium on Combinatorial Pattern Matching, Jouni Sirén, 2008

Reviewer: Journal of Combinatorial Mathematics and Combinatorial Computing, Jouni Sirén, 2008

Reviewer: 16th International Symposium on String Processing and Information Retrieval, Jouni Sirén, 2009

Reviewer: Bioinformatics, Jouni Sirén, 2009

Reviewer: Journal on Experimental Algorithms, Jouni Sirén, 2009

Reviewer: 12th Scandinavian Symposium and Workshops on Algorithm Theory, Jouni Sirén, 2010

Reviewer: 17th International Symposium on String Processing and Information Retrieval, Jouni Sirén, 2010

Reviewer: 24th AAAI Conference on Artificial Intelligence, Jouni Sirén, 2010

Reviewer: EURASIP Journal on Bioinformatics and Systems Biology, Jouni Sirén, 2010

Abhishek Tripathi ,

Reviewer , Abhishek Tripathi, 04.01.2010 → ..., United States

Niko Välimäki ,

CPM 2008, Niko Välimäki, 2008

PCM 2008, Niko Välimäki, 2008

STACS 2008, Niko Välimäki, 2008

FOCS 2009, Niko Välimäki, 2009

SPIRE 2009, Niko Välimäki, 2009

Algorithms and Applications (Ukkonen Festschrift), Niko Välimäki, 2010, Finland

Information Processing Letters, Niko Välimäki, 2010

Journal of Discrete Algorithms, Niko Välimäki, 2010

SPIRE 2010, Niko Välimäki, 2010

Editor of series

Aapo Hyvärinen ,

Action Editor in J of Machine Learning Research, Aapo Hyvärinen, 2004 → ...

Action Editor in Neural Computation, Aapo Hyvärinen, 2005 → ...

Contributing Faculty Member of the Faculty of 1000, Aapo Hyvärinen, 2006 → ...

Member of Editorial Board of Foundations and Trends in Machine Learning, Aapo Hyvärinen, 2007 → ...

Editor of special theme number

Petri Myllymäki ,

EURASIP Journal on Bioinformatics and Systems Biology, special issue on Information-Theoretic Methods for Bioinformatics, Petri Myllymäki, 2006 → 2007

Juho Rousu ,



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Information Theoretic Methods for Bioinformatics, Juho Rousu, 01.01.2007 → 01.01.2008

Hannu Toivonen ,

Guest Editor: Journal of Machine Learning Research, Hannu Toivonen, 2008, United States

Assessment of candidates for academic posts

Esko Ukkonen ,

Assessment for a professorship, Esko Ukkonen, 2005, Israel

Assessment committee for a professorship, Esko Ukkonen, 2007, Norway

Assessment for a professorship, Esko Ukkonen, 2007, Germany

Assessment for a tenured Associate Professorship, Esko Ukkonen, 2007, United States

Appointment committee for an associate professorship, Esko Ukkonen, 2008, Sweden

Assessment for a position of Research Director, Esko Ukkonen, 2008, France

Assessment for a position of Senior Lecturer, Esko Ukkonen, 2008, Israel

Assessment of a Research Scientist, Esko Ukkonen, 2009, Norway

Assessment for a position of Senior Lecturer, Esko Ukkonen, 2010, Israel

Aapo Hyvärinen ,

Statement regarding promotion to full professor, POSTECH, Aapo Hyvärinen, 2007

Statement for filling a position of associate professor at NYU, Aapo Hyvärinen, 2008

Statement regarding promotion to full professor, UC Berkeley, Aapo Hyvärinen, 2009

Statement regarding filling a position of associate professor, Aapo Hyvärinen, 2010

Petri Myllymäki ,

Assessment for a senior position, Petri Myllymäki, 2006, Australia

Assessment for a professorship, Petri Myllymäki, 2006, New Zealand

Assessment for a senior position, Petri Myllymäki, 2007, United Kingdom

Assessment for a docentship, Petri Myllymäki, 2009, Finland

Assessment for a docentship, Petri Myllymäki, 2009, Finland

Assessment for a professorship, Petri Myllymäki, 2009, Finland

Assessment for a senior position, Petri Myllymäki, 2009, France

Juho Rousu ,

Assessment for Adjunct Professor position, Juho Rousu, 29.02.2008, Finland

Mikko Sillanpää ,

Reviewer of a Wellcome Trust Senior Fellowship application, Mikko Sillanpää, 08.2006 → 09.2006, United Kingdom

Hannu Toivonen ,

External reviewer for Full Professor promotion in the University of Texas at Arlington, USA, Hannu Toivonen, 2007, United States

External reviewer for Full Professor promotion in the University of Texas at Arlington, USA, Hannu Toivonen, 2008, United States

Helena Ahonen-Myka ,

Assessment of a docent candidate, Helena Ahonen-Myka, 2006, Finland

Membership or other role in review committee

Esko Ukkonen ,

Evaluation of research proposals, Esko Ukkonen, 2005, Netherlands

Evaluation of a Senior Fellowship application, Esko Ukkonen, 2006, United Kingdom



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Evaluation of a research proposal, Esko Ukkonen, 2007, Germany
Evaluation panel of Bioinformatics Initiative, Esko Ukkonen, 2007, Germany
Evaluation panel of doctoral programmes in Sweden, Esko Ukkonen, 2007 → 2008, Sweden
Expert evaluator for European Research Council, Esko Ukkonen, 2007
Evaluation of a research proposal, Esko Ukkonen, 2008, Israel
Evaluation of a postdoctoral project proposal, Esko Ukkonen, 2009, France
Review of academic educational program (TU München), Esko Ukkonen, 2009, Germany
Evaluation of a grant proposal, Esko Ukkonen, 2010, Canada
Evaluation of research proposals, Esko Ukkonen, 2010, Austria
Expert evaluator for the European Union Framework Programme 7, Esko Ukkonen, 2010 → 2011

Mikko Koivisto ,

Reviewer for the Netherlands Organisation for Scientific Research, Mikko Koivisto, 2010 → ...

Petri Myllymäki ,

European Commission, review panel for the IST programme, Petri Myllymäki, 18.04.2005 → 27.04.2005, Belgium
NWO, Petri Myllymäki, 01.11.2005 → 30.11.2005, Netherlands
European Commission, Petri Myllymäki, 01.01.2006 → 31.12.2006
Expert reviewer for the Austrian Science Fund (FWF), Petri Myllymäki, 01.01.2006 → 31.12.2006, Austria

Valentin Polishchuk ,

socg, Valentin Polishchuk, 03.12.2009 → 14.02.2010

Juho Rousu ,

Research project review, Juho Rousu, 2005 → 2010, United Kingdom
Board Member of Graduate School, Juho Rousu, 01.01.2007 → ..., Finland
Research project review, Juho Rousu, 2007, Austria
Research project review, Juho Rousu, 2009, Switzerland
Research project review, Juho Rousu, 2010, Belgium
Research project review, Juho Rousu, 2010, United Kingdom

Mikko Sillanpää ,

External Reviewer for inner evaluation purposes of the New Zealand Forest Research Institute, Mikko Sillanpää, 08.2006 → ..., New Zealand
Reviewer of a Equipment, Technology Development & Biomedical Resources Grant application, Mikko Sillanpää, 04.2007 → 06.2007, United Kingdom
Reviewer of a Wellcome Trust Research Grant application, Mikko Sillanpää, 08.2007 → 09.2007, United Kingdom
Reviewer of a Research Grant Application, Mikko Sillanpää, 05.2008 → 06.2008, Israel
Reviewer of a Research Grant Applications for Postdoctoral Researchers, Mikko Sillanpää, 08.2008 → ..., Finland
Reviewer of MRC Career Development Award, Mikko Sillanpää, 2010 → ..., United Kingdom

Hannu Toivonen ,

Reviewer for The Netherlands Organisation for Scientific Research (NWO), Hannu Toivonen, 2005, Netherlands
Reviewer for European Young Investigators Awards (EURYI)/EUROHORCS, Hannu Toivonen, 2007, Finland
Reviewer for Health Science & Technology (HST, The Netherlands), Hannu Toivonen, 2007, Netherlands
Reviewer for The National Fund for Scientific Research (FWO, Belgium), Hannu Toivonen, 2007, Belgium
Member of SIGKDD Dissertation Award Committee, Hannu Toivonen, 2008 → ...
Member of the Computer Science panel of the Swedish Science Council, Hannu Toivonen, 2009, Sweden



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Reviewer for European Research Council (ERC): ERC Starting Grants, Hannu Toivonen, 2009

Reviewer for the Netherlands Genomics Initiative, Horizon programme, Hannu Toivonen, 2009, Netherlands

Member of ERC Starting Grants evaluation panel, European Research Council (ERC), Hannu Toivonen, 2010 → 2011

Reviewer for The National Fund for Scientific Research (FWO, Belgium), Hannu Toivonen, 2010, Belgium

Helena Ahonen-Myka ,

Reviewer for the University of Antwerpen, Helena Ahonen-Myka, 2006, Belgium

Reviewer for an Advanced Study Institute program, The Croucher Foundation, Helena Ahonen-Myka, 2007, Hong Kong

Reviewer for Swiss National Science Foundation, Helena Ahonen-Myka, 2008, Switzerland

Membership or other role in research network

Esko Ukkonen ,

Director of National Center-of-Excellence "From Data to Knowledge" (Academy of Finland, 2002–2007), Esko Ukkonen, 2002 → 2007, Finland

Director of National Center-of-Excellence in Algorithmic Data Analysis research (Academy of Finland, 2008–2013), Esko Ukkonen, 2008 → 2013, Finland

Samuel Kaski ,

Steering Committee Member, Samuel Kaski, 2008 → ...

Member, Samuel Kaski, 2010 → ...

Petri Myllymäki ,

Member of the PASCAL Steering Committee, Programme Director, Node Director, Petri Myllymäki, 2003 → 2012

Valentin Polishchuk ,

complexWorld.eu, Valentin Polishchuk, 26.04.2010 → ...

halal, Valentin Polishchuk, 26.04.2010 → ...

Juho Rousu ,

Member of EU FP6 NoE PASCAL, Juho Rousu, 01.01.2004 → 28.02.2008

Coordinator of thematic programme in EU network, Juho Rousu, 01.12.2005 → 30.06.2006, United Kingdom

Member of EU FP7 NoE PASCAL2, Juho Rousu, 01.03.2008 → 28.12.2013

Mikko Sillanpää ,

Substitute of the EU COST TD0801 Management Committee in Finland, Mikko Sillanpää, 08.2009 → 12.2012

Hannu Toivonen ,

Member of PASCAL 2 (EU Network of Excellence), Hannu Toivonen, 2008 → 2013

Membership or other role in national/international committee, council, board

Esko Ukkonen ,

Member of the international Steering Committee of the bi-annual conference series Scandinavian Workshop on Algorithm Theory (SVAT), Esko Ukkonen, 1991 → ...

International steering committee of the Bioinformatics Research and Education Workshops (BREW), Esko Ukkonen, 2002 → ...

Panel chairman of the Publication Forum Project of the FEDERATION OF FINNISH LEARNED SOCIETIES, Esko Ukkonen, 2010 → 2012, Finland

Samuel Kaski ,

Member, Samuel Kaski, 2009 → ...

Member, Samuel Kaski, 2009 → ..., Finland

Member of Governing Board, Samuel Kaski, 2010 → ...

Jyrki Kivinen ,



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Association for Computational Learning, Jyrki Kivinen, 01.01.2005 → 31.12.2005, United States

Association for Computational Learning, Jyrki Kivinen, 01.01.2006 → 24.06.2006

Kjell Lemström ,

Steering Committee member of ISMIR (International Society for Music Information Retrieval), Kjell Lemström, 01.2002 → 10.2009

Member of the board of the Finnish Musicological Society, Kjell Lemström, 03.2004 → 03.2009, Finland

Member of the Faculty Council, Kjell Lemström, 01.2007 → 11.2008, Finland

Member of the University Collegium, Kjell Lemström, 03.2007 → 11.2008, Finland

Veli Mäkinen ,

Tietojenkäsittelytieteen laitoksen johtoryhmä, Veli Mäkinen, 2007 → 2009, Finland

Board of Helsinki Graduate School in Computer Science and Engineering (Hecse), Veli Mäkinen, 2008 → ..., Finland

Matemaattis-luonnontieteellisen tiedekunnan tiedekuntaneuvosto, Veli Mäkinen, 2009, Finland

Bioinformatiikan maisteriohjelman ohjausryhmä, Veli Mäkinen, 2010 → ..., Finland

Steering group for Finnish Doctoral Programme in Computational Sciences, FICS, Veli Mäkinen, 2010, Finland

Petri Myllymäki ,

Steering Committee of the Pascal EU Network of Excellence, Petri Myllymäki, 2003 → 2012, United Kingdom

Mikko Sillanpää ,

A member in Complex Trait Consortium (CTC), Mikko Sillanpää, 2004 → ...

Hannu Toivonen ,

Board Member and Steering Committee Member, Helsinki Graduate School in Computer Science and Engineering (Hecse), Hannu Toivonen, 01.2003 → ..., Finland

Director, Helsinki Graduate School in Computer Science and Engineering (Hecse), Hannu Toivonen, 01.01.2007 → 31.12.2011

Member: Committee for Development of Service Centers at Campuses, Hannu Toivonen, 04.12.2008 → 27.02.2009, Finland

Member: Committee for structural development of the university, Hannu Toivonen, 01.04.2008 → 31.12.2009, Finland

Member: Expert Committee on Academic Leadership, Hannu Toivonen, 11.2008 → 12.2009, Finland

Member: Scientific Expert Committee of the Faculty of Science, Hannu Toivonen, 01.10.2008 → 31.12.2013, Finland

Steering Group Member: Development of Research Information System at the University of Helsinki, Hannu Toivonen, 09.09.2008 → 31.12.2009, Finland

Strategic Committee Member of Finnish Doctoral Programme in Computational Sciences, FICS, Hannu Toivonen, 2009 → 2010, Finland

Board Member: Helsinki Institute for Information Technology HIIT, Hannu Toivonen, 01.10.2010 → 31.03.2014, Finland

Chairman of the board of Master's programme in Bioinformatics (MBI), Hannu Toivonen, 03.2010 → ..., Finland

Management Committee Member of COST Action IC1002, "MUMIA", Hannu Toivonen, 2010 → 2013

Member: Doctoral Education Committee at the University of Helsinki, Hannu Toivonen, 01.01.2010 → 31.12.2012, Finland

Member: Doctoral Education Committee of the Faculty of Science, Hannu Toivonen, 01.12.2010 → 31.12.2011, Finland

Helena Ahonen-Myka ,

Board member of the Helsinki Graduate School in Computer Science and Engineering (HeCSE), Helena Ahonen-Myka, 2005, Finland

Board member of the national graduate school for human language technology, Helena Ahonen-Myka, 2005, Finland

Board member of the national language technology education network (KIT), Helena Ahonen-Myka, 2005, Finland

Board member of the Helsinki Graduate School in Computer Science and Engineering (HeCSE), Helena Ahonen-Myka, 2006, Finland

Board member of the national graduate school for human language technology, Helena Ahonen-Myka, 2006, Finland

Board member of the national language technology education network (KIT), Helena Ahonen-Myka, 2006, Finland

Board member of the national graduate school for human language technology, Helena Ahonen-Myka, 2007, Finland



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Board member of the national language technology education network (KIT), Helena Ahonen-Myka, 2007, Finland

Ella Bingham ,

ACM SIGKDD (konferenssi KDD 2006), Ella Bingham, 01.01.2006 → 31.12.2006

Taneli Johannes Mielikäinen ,

PASCAL - Pattern Analysis, Statistical Modelling, and Computational Learning Network of Excellence (www.pascal-network.org), Taneli Johannes Mielikäinen, 01.01.2006 → 31.12.2006

Membership or other role in public Finnish or international organization

Esko Ukkonen ,

Member of the board of Helsinki Graduate School in Computer Science and Engineering, Esko Ukkonen, 1996 → ..., Finland

Member of the Board of the Institute of Biotechnology (University of Helsinki), Esko Ukkonen, 2002 → 2009, Finland

Member of the Board of Graduate School in Computational Biology, Bioinformatics, and Biometry, Esko Ukkonen, 2003 → 2009, Finland

Chairman of the Board of the Institute of Biotechnology (University of Helsinki), Esko Ukkonen, 2009 → ..., Finland

Samuel Kaski ,

Vice director, Samuel Kaski, 2006 → ..., Finland

Vice head of department, Samuel Kaski, 2008 → 2010, Finland

Director, Samuel Kaski, 2009 → ..., Finland

Director, Samuel Kaski, 2010 → ..., Finland

Petri Myllymäki ,

Helsinki Institute for Information Technology HIIT, member of the Board of Directors, Petri Myllymäki, 2002 → 2006, Finland

Helsinki Graduate School in Computer Science and Engineering, member of the Board of Directors, Petri Myllymäki, 2004 → ..., Finland

Steering Committee of the Department of Computer Science, University of Helsinki, Petri Myllymäki, 2007 → ..., Finland

Steering Committee of the Kumpula science library, Petri Myllymäki, 2007 → 2010, Finland

Hannu Toivonen ,

Member of the Council of Faculty of Science, Hannu Toivonen, 10.2004 → 12.2013, Finland

Vice Chair of the Council of Department of Computer Science, Hannu Toivonen, 01.2004 → 12.2006, Finland

Head of Department of Computer Science, Hannu Toivonen, 01.2007 → 12.2009, Finland

Vice Member of the Council of Department of Computer Science, Hannu Toivonen, 01.2010 → 12.2013, Finland

Membership or other role of body in private company/organisation

Kjell Lemström ,

Board member of Hecse (Helsinki Graduate School in Computer Science and Engineering, Kjell Lemström, 2002 → ..., Finland

Board member of Hiidenkivi comprehensive school, Kjell Lemström, 2009 → 2012

Petri Myllymäki ,

Cloud'N'Sci Inc., member of the Board of Directors, Petri Myllymäki, 2010 → ..., Finland

Other tasks of an expert in private sector

Hannu Toivonen ,

Referee for the Imagine Cup, Software Design, Finnish semifinal, Hannu Toivonen, 2009

Advisory Board member of Finsor Ltd, Hannu Toivonen, 01.2010 → ..., Finland

Scientific Advisory Board Member of STACC, Hannu Toivonen, 01.2010 → ..., Estonia

Participation in interview for written media



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

Esko Ukkonen ,

Interview for Perspektiivi, Esko Ukkonen, 2008

Interview for Tieteen tietotekniikka (ICT for Science), Esko Ukkonen, 2009

Samuel Kaski ,

Apropos 1/2005, magazine of the Academy of Finland, Samuel Kaski, 2005, Finland

Informaatietekniikan laboratorion 40-vuotisjuhla, Samuel Kaski, 2005, Finland

ProAcademia, News from the Academy of Finland 2/2005, Samuel Kaski, 2005, Finland

Proaktiivisen tietotekniikan kollokvio, Samuel Kaski, 2005, Finland

Kjell Lemström ,

Interview in LUOVA (Nuorten luonnontieteidenverkkoletti): "Mikä se biisi nyt oli ... ja missä?", Kjell Lemström, 05.08.2007, Finland

Interview in Forum för ekonomi och teknik: "Sökmotor som hittar låten du nynnär", Kjell Lemström, 2008, Finland

Petri Myllymäki ,

Interview on the Tietokone magazine, Petri Myllymäki, 18.10.2007, United Kingdom

Interview on the Tietokone magazine, Petri Myllymäki, 19.10.2007, United Kingdom

Hannu Toivonen ,

Kauppalehti 14.2.2005, Hannu Toivonen, 14.02.2005, Finland

Portaali 1/2005, p. 16, Hannu Toivonen, 2005, Finland

Puhelin, 2/2005, p. 32-33, Hannu Toivonen, 2005, Finland

Tietoyhteys 1/2005, s. 3, Hannu Toivonen, 2005, Finland

Illustrerad Vetenskap 6/2006, s. 44-45, Hannu Toivonen, 2006, Sweden

Illustreret Videnskab 6/2006, p. 40-41, Hannu Toivonen, 2006, Denmark

Illustrert Vitenskap 6/2006, p. 46-47, Hannu Toivonen, 2006, Norway

The New York Times, 6 July 2006, Hannu Toivonen, 06.07.2006, United States

Tieteen Kuvalehti 6/2006, Hannu Toivonen, 2006, Finland

Valitut Palat, June 2007, Hannu Toivonen, 06.2007, Finland

Helsingin Sanomat 28.8.2008, Hannu Toivonen, 28.08.2008, Finland

Helsingin Sanomat 29.8.2008, Hannu Toivonen, 29.08.2008, Finland

Hufvudstadsbladet 28.8.2008, Hannu Toivonen, 28.08.2008, Finland

Polyteekkari 15/2008, Hannu Toivonen, 2008, Finland

Tekniikan maailma 19/2008, Hannu Toivonen, 2008, Finland

Tietoviikko 19.9.2008, Hannu Toivonen, 19.09.2008, Finland

Ylioppilaslehti 11/08, Hannu Toivonen, 2008, Finland

Helsingin Sanomat 4.2.2009, Hannu Toivonen, 04.02.2009, Finland

Helsinki University Bulletin 2/10, Hannu Toivonen, 2010, Finland

Ylioppilaslehti 2010, Hannu Toivonen, 12.2010, Finland

Henry Tirri ,

Euroopan tiede- ja teknologiaviikko, Henry Tirri, 01.11.2000 → 31.12.2011, Finland

Fincon tietoisuusmessut, Henry Tirri, 01.08.2000 → 31.12.2011, Finland

Helsingin Sanomat, Henry Tirri, 01.01.2000 → 31.12.2011, Finland

Tekniikka ja talous, Henry Tirri, 01.09.2000 → 31.12.2011, Finland

Tietoviikko, Henry Tirri, 01.05.2000 → 31.12.2011, Finland



INTERNATIONAL EVALUATION OF RESEARCH AND DOCTORAL TRAINING AT THE UNIVERSITY OF HELSINKI

RC-SPECIFIC TUHAT COMPILATIONS OF OTHER SCIENTIFIC ACTIVITIES 2005-2010

ALKO/Ukkonen

YLE, Henry Tirri, 01.03.2000 → 31.12.2011, Finland
Yliopisto, Henry Tirri, 01.11.2000 → 31.12.2011, Finland
Helsingin Sanomat, Henry Tirri, 11.11.2003 → 31.12.2011, Finland
Ideapolis, Henry Tirri, 01.01.2003 → 31.12.2011, Finland
Profile, Henry Tirri, 01.01.2003 → 31.12.2011, Finland
Puhelin, Henry Tirri, 01.01.2003 → 31.12.2011, Finland
Taltio, Henry Tirri, 01.01.2003 → 31.12.2011, Finland
Tekniikka & Talous, Henry Tirri, 12.03.2003 → 31.12.2011, Finland
Tiede 03-tiedekatselmus, Henry Tirri, 07.11.2003 → 31.12.2011, Finland
Tieteen päivät 2003, Henry Tirri, 12.01.2003 → 31.12.2011, Finland
Tietopalvelu, Henry Tirri, 01.01.2003 → 31.12.2011, Finland
Tietoviikko, Henry Tirri, 21.08.2003 → 31.12.2011, Finland
YLE - IT-uutiset, Henry Tirri, 01.01.2003 → 31.12.2011, Finland

Participation in radio programme

Petri Myllymäki ,

Interview on the Finnish radio (YleX), Petri Myllymäki, 21.11.2006, Finland
Radio-ohjelma YleX-kanavalla, Petri Myllymäki, 21.11.2006, Finland

Juho Rousu ,

Radiaattori, Juho Rousu, 21.02.2007

Hannu Toivonen ,

Radio Suomi 17.3.2005, Hannu Toivonen, 17.03.2005, Finland
Radio Suomi 8.7.2008, Hannu Toivonen, 08.07.2008, Finland

Participation in TV programme

Petri Myllymäki ,

Haastattelu Verkossa-ohjelmassa, MTV3, Petri Myllymäki, 02.04.2007, United Kingdom

Teemu Teppo Roos ,

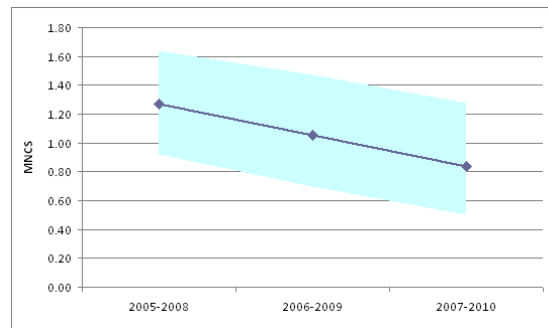
Interview on Prisma Studio, TV1, Teemu Teppo Roos, 14.11.2006, Finland

Research Group: Ukkonen E

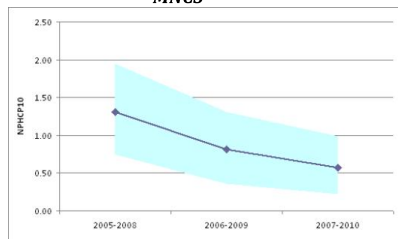
Basic statistics

Number of publications (P)	180
Number of citations (TCS)	748
Number of citations per publication (MCS)	4.21
Percentage of uncited publications	42%
Field-normalized number of citations per publication (MNCS)	1.12
Field-normalized average journal impact (MNJS)	1.43
Field-normalized proportion highly cited publications (top 10%)	1.06
Internal coverage	.64

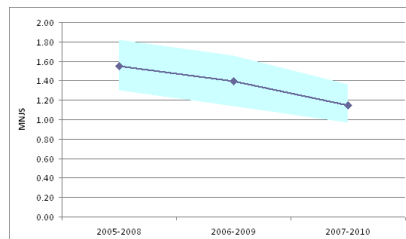
Trend analyses



MNCS

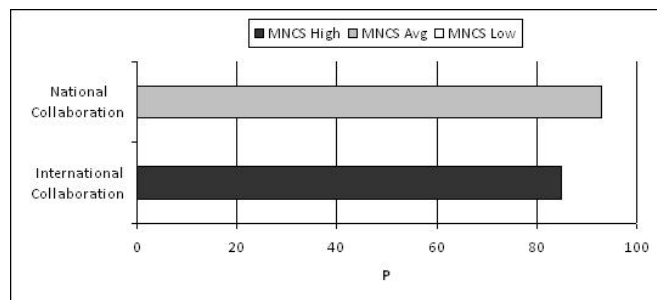


THCP10



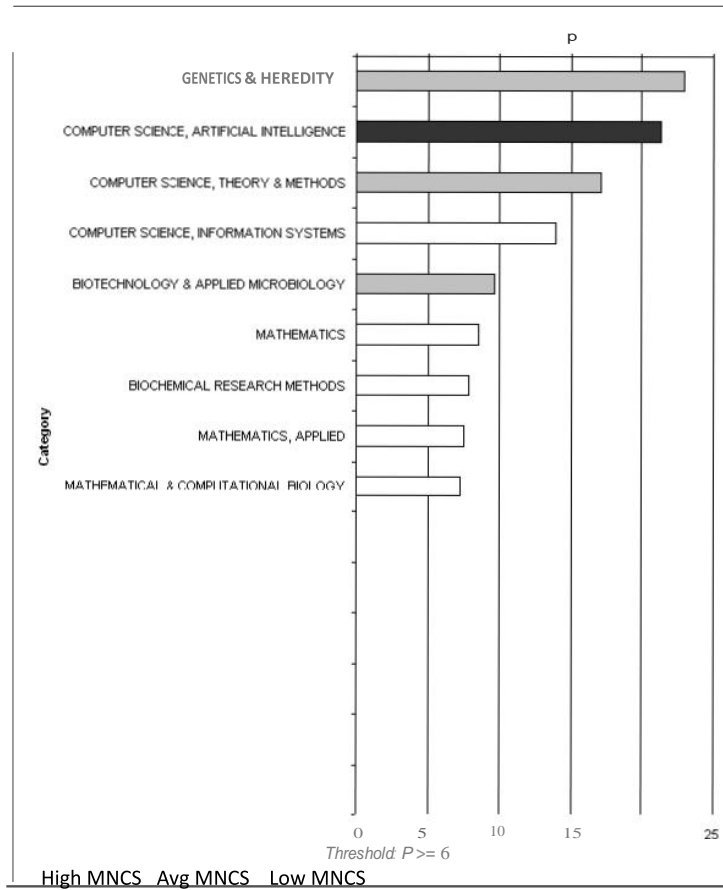
MNJS

Collaboration



Performance (MNCS) by collaboration type

Research profile



Appendix B.b.

Maria Forsman, Chief Information Specialist, DSocSc
Helsinki University Library 7.7.2011

The bibliometric analyses by Helsinki University Library (HULib)

Background: The bibliometric analyses – especially citation analyses – have raised a lot of discussion and critics among researchers in social sciences and humanities. Researchers view that bibliometric analyses are often unfair to these fields of sciences because they do not give a good enough picture of the publishing. Citation databases – Web of Science and Scopus – cover only weakly the main publications in these fields. Also, in humanities and social sciences monograph is still the main form of publishing, and it does not include in these article databases.

At the University of Helsinki, the above mentioned concerns have been taken into account in the evaluation. The Evaluation Office has ordered analyses from the Helsinki University Library (HULib) for the participating researcher communities that are weakly represented in Web of Science. The database for the HULib analyses is TUHAT (<https://tuhat.halvi.helsinki.fi/portal/en/>) including all the publications that the researchers have considered important.

Based on this data, information specialists at HULib have carried out the following analyses:

- 1) Number of authors/publication/year as a table; a pie of authors/publication in the period 2005-2010;
- 2) Language of publication/year; a pie of language of publication in the period 2005-2010;
- 3) Articles/journal/year; journals have been compared by ISSN with the Norwegian, Australian and ERIH (2007-2008) journal ranking lists; number of articles in ranked journals;
- 4) Publisher/monograph type (according to TUHAT database); monographs have been compared with the Norwegian publisher ranking list. According to this, it has been counted how many monographs are published by a leading scientific publisher (2) or a scientific publisher (1).
- 5) Conference publications (from TUHAT database) especially in computer sciences; compared with the Australian conference ranking list.

Where relevant, some additional analyses and notes concerning the publication culture of a scientific field have been added. Overall, these analyses complement the other evaluation material and lists of the publications of the participating researcher communities.

If the publications of the RCs were less than 50 or/and the internal coverage less than 40 percentage, the WoS analyses were considered not reliable. These RCs were 58 altogether.

In addition, both Leiden and Library analyses were done to the RCs if WoS analyses covered less than 40 per cent of the peer review (A+C) publications of the RC. These RCs were 8 altogether.

The appendix includes the analyses of the RC under discussion.

**Analysis of publications by Helsinki University
Library – 66 RCs altogether**

Biological, Agricultural and Veterinary Sciences

Luukkanen, Olavi – VITRI
Valsta, Lauri – SUVALUE

Natural Sciences

Abrahamsson, Pekka – SOFTSYS
Kangasharju, Jussi – NODES
Ukkonen, Esko – ALKO
Väänänen, Jouko – HLG

Humanities

Aejmelaesus, Anneli – CSTT
Anttonen, Pertti – CMVG
Dunderberg, Ismo – FC
Havu, Eva – CoCoLaC
Heikkilä, Markku – RCSP
Heinämaa, Sara – SHC
Henriksson, Markku – CITA
Janhunen, Juha – LDHFTA
Kajava Mika, – AMNE
Klippi, Anu – Interaction
Knuuttila, Simo – PPMF
Koskenniemi, Kimmo – BAULT
Lauha, Aila – CECH
Lavento, Mika – ARCH-HU
Lukkarinen, Ville – AHCI
Lyytikäinen, Pirjo – GLW
Mauranen, Anna – LFP
Meinander, Henrik – HIST
Nevalainen, Terttu – VARIENG
Pettersson, Bo – ILLC
Pulkkinen, Tuija – Gender Studies
Pyrhönen, Heta – ART
Ruokanen, Miikka – RELDIAL
Saarinen, Risto – RELSOC
Sandu, Gabriel – LMPS
Tarasti, Eero – MusSig
Vehmas-Lehto, Inkeri – TraST
Östman, Jan-Ola – LMS

Social Sciences

Airaksinen, Timo – PPH
Engeström, Yrjö – CRADLE
Granberg, Leo – TRANSRURBAN
Haila, Anne – Sociopolis
Hautamäki, Jarkko – CEA
Heinonen, Visa – KUMU
Helén, Ilpo – STS
Hukkinen, Janne – GENU
Jallinoja, Riitta – SBII
Kaartinen, Timo – SCA
Kettunen, Pauli – NordSoc
Kivinen, Markku – FCREES
Koponen, Juhani – DEVERELE
Koskenniemi, Martti – ECI
Kultti, Klaus – EAT
Lahelma, Elina – KUFÉ
Lanne, Markku – TSEM
Lavonen, Jari – RCMSE
Lehtonen, Risto – SocStats
Lindblom-Yläne, Sari – EdPsychHE
Nieminen, Hannu – MECOL
Nuotio, Kimmo – Law
Nyman, Göte – METEORI
Ollikainen, Markku – ENFIFO
Pirttilä-Backman, Anna-Maija – DYNASOBIC
Rahkonen, Keijo – CulCap
Roos, J P – HELPS
Simola, Hannu – SOCE-DGI
Sulkunen, Pekka – PosPus
Sumelius, John – AG ECON
Vaattovaara, Mari – STRUTSI
Vainio, Martti – SigMe

The next appendix includes the analyses of the
RC under discussion.

ALKO / Ukkonen

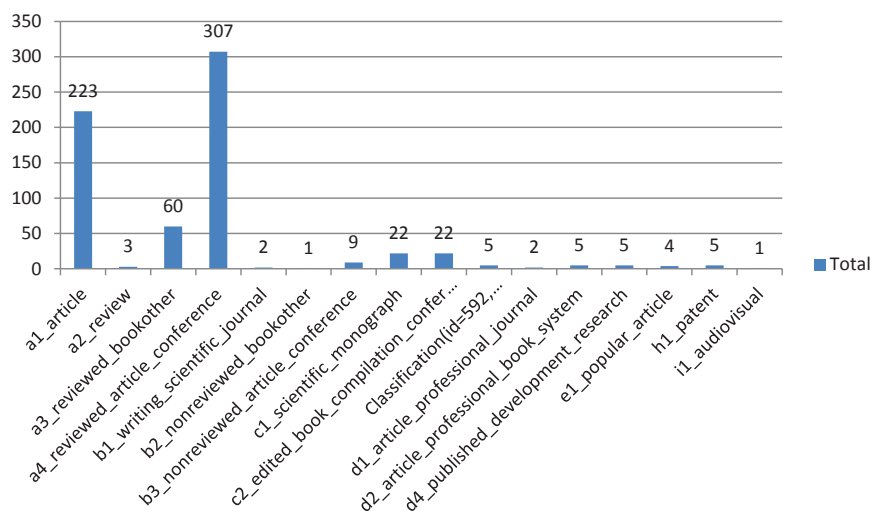
Category 1. The research of the participating community represents the international cutting edge in its field.

The principles used in evaluating Computer Science are stated in *Research Evaluation for Computer Science: An Informatics Europe report* (2008)

http://www.informatics-europe.org/docs/research_evaluation.pdf

Basic Statistics

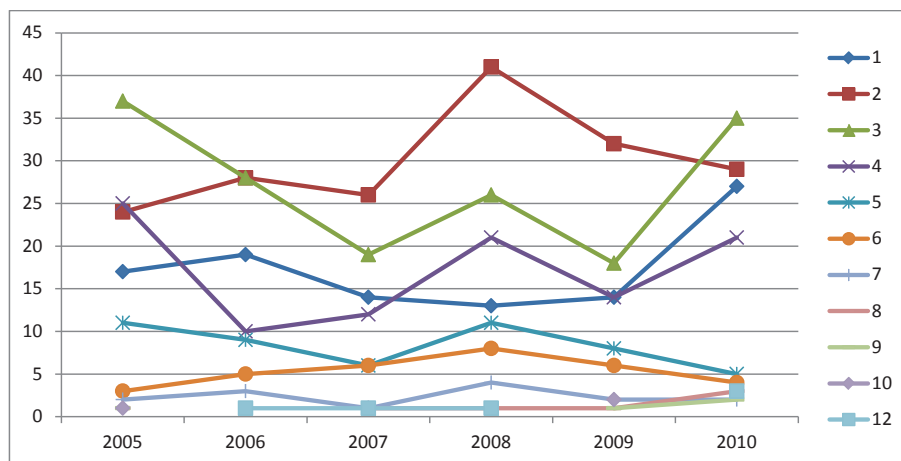
The group is fairly large, with 676 publications in TUHAT, showing a peak in A4 conference publications, typical for Computer Science, but also a largish amount of A1 refereed papers, as shown in a chart with publication counts per classification:



The papers have on average 3.3 authors per publication, as shown in the following table with breakdown of papers with 1..22 authors:

# of Authors	Count
1	104
2	180
3	163
4	103
5	50
6-9	57
10-22	19
Total	676

The following chart shows the breakdown of the number of authors for each year:



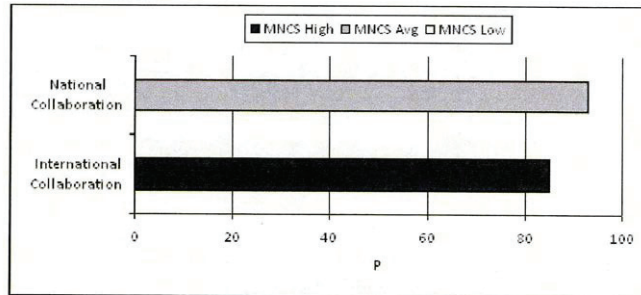
There are 184 papers in Web of Science, which makes a quantitative analysis possible. However, as conference publications have a very central role in Computer Science, WoS analysis gives only a partial view on the situation.

Collaboration

The CWTS research performance profile provides adequate information about the performance of the group by collaboration type (national and international collaboration), thus the analysis has not been duplicated here for conference publications.

The preliminary CWTS analysis graph is shown below.

Collaboration



Performance (MNCS) by collaboration type

ARC Conference Rankings

As could be expected, conference publications cover 47% of the output of this group. Conference data are not clearly indicated in TUHAT records, thus they were examined separately to find out matches with the Australian Research Council's (ARC) ranked conference list (2010):

http://www.arc.gov.au/era/era_journal_list.htm#2

The output of conference publications is large and heavily oriented towards A-ranked conferences. The non-ranked conferences were too new, local, or interdisciplinary (bioinformatics, musicology, etc.) to be found on the ARC list. The rankings found (or the lack of rankings) are listed below:

ARC Rank	Count
A	119
B	46
C	28
none found	122
Total	315

A list of conference acronyms with (and without) ARC ranks is given below.

A-ranked: ACSAC, ALNEX (2), CIKM, COLING/CLIA, COLT (2), DCC, DISC, ECAI (2), ECML (4), ECML/PKDD (7), EDBT, EMNLP, ESA (2), FOCS (3), HICSS, HLT/EMNLP, ICALP (2), ICDE (2), ICDM (6), ICML (6), ICML/LWS, ICONIP (2), IDA, IFIP SEC/TC11, IJCAI (3), IJCNN (5), IPDPS, KR, LPAR (2), MFCS, NIPS (6), PAKDD, PKDD (4), SAT, SCG (3), SDM (5), SIGIR (4), SIGIR/ELECTRA, SIGIR/OSIR (2), SIGKDD (6), SIGMOD, SODA, SPAA, SPIRE, STOC, TACAS, UAI (13)

B-ranked: ADHOC-NOW, AISTATS (6), ANALCO, CIAA (2), CPM (6), EDOC/VORTE/MOST, EPIA, ESANN (2), FUN, ICANN (5), ICPR, ISITA (2), ITW (3), LATIN, MobileHCI, RECOMB (2), SEA, SECON, SIGSPATIAL, SPIRE (3), STACS, WABI (3)

C-ranked: Algosensors (3), BIBE (2), CMSB, IASTED, ICCSA, ICMLA, IEEE ADPRL, INEX (3), ISI, ISMIS, IWDW, LOCALALGOS, MACIS, MobiDE, PSC, RANLP, WI-AT, WI-IAT (3), WKDD (3)

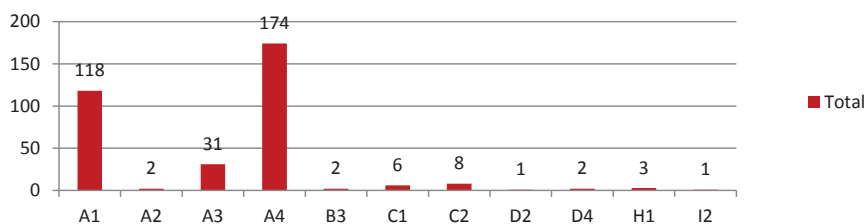
No rank listed: ACML, AIAA GN&C (2), AKRR (2), AMICT (2), ASONAM, ASWC, BIOINFORMATICS, Biomag, BIRD, CAMDA, CAPS (4), CLEF (2), CMMR (4), DIALM-POMC, DrMED/MIE, DS (3), DX, ECC, ECDL (2), EuroCogSci (2), GCB (2), GIScience, IBC, IBGS, ICA (6), ICASSP (2), ICCX-X, ICMPC, ICOS (2), ICRAT, ICWN, ICWSM, IIA none, IMED, ISA (IADIS), ISMIR (3), ISTL, ITA (2), KDIR, KRBIO (3), LNLA, Louhi, LTC, LWS, MCCSIS, MINES, MINUCS, MLG, MLMTA, MLSB, MLSP, MML (3), NBBC, NCPW, NETTAB, NODALIDA, OSWIR, PCI, PGM (6), PRIB, SCIA, SMI, SoCG, SSP, STANS, UM, WCGALP, WIMSE (2), WITMSE (5), WSOM (2), WSW

Publish or Perish (Google Scholar) data

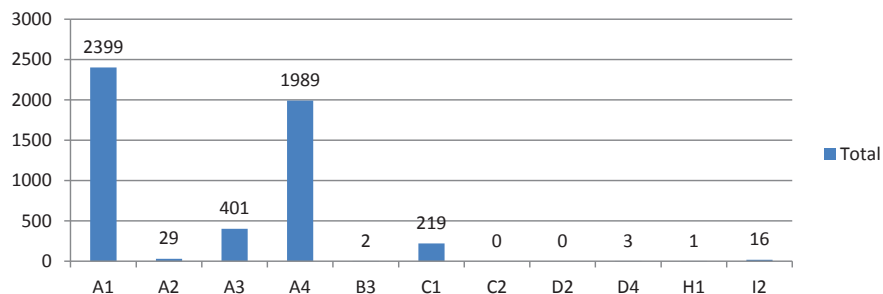
A Publish or Perish (PoP) search with names of the ALKO team members shows that the publication data from TUHAT seems incomplete. Several refereed papers are missing from the group's list of publications. On the other hand, PoP is missing some of the listed publications.

We included 348 papers/titles that are in TUHAT and can also be found with PoP. The following charts summarize the findings (citation count date: May 24,2011):

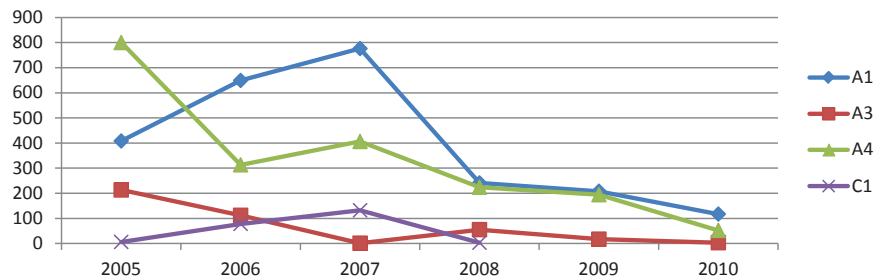
The PoP publication counts per class match the TUHAT publications distribution quite well:



Out of 483 PoP publications, 17,5% are uncited. As is to be expected, refereed A1 papers are the most cited in PoP. A4 conference papers which dominate the publication counts are not very far behind:



Four publication types have a significant amount of citations, broken down by year as follows:



ACM

The ACM database at <http://portal.acm.org> includes citations, but only a part of the papers can be found in the database. Bibliometric ACM summaries of the PIs including both citation and download counts (for ACM publications available for download) are listed below. Note that in the ACM database, one cannot choose the appropriate time range for the analysis.

Ukkonen:	http://portal.acm.org/author_page.cfm?id=81100574701
Hoyer:	http://portal.acm.org/author_page.cfm?id=81100304686
Hyvärinen:	http://portal.acm.org/author_page.cfm?id=81100100305
Kärkkäinen:	http://portal.acm.org/author_page.cfm?id=81100071822
Kaski P:	http://portal.acm.org/author_page.cfm?id=81100348743
Kaski S:	http://portal.acm.org/author_page.cfm?id=81100348810
Kivinen:	http://portal.acm.org/author_page.cfm?id=81339509885
Koivisto:	http://portal.acm.org/author_page.cfm?id=81309511472
Mannila:	http://portal.acm.org/author_page.cfm?id=81100086722
Myllymäki:	http://portal.acm.org/author_page.cfm?id=81100436253
Polishchuk:	http://portal.acm.org/author_page.cfm?id=81331501995
Rousu:	http://portal.acm.org/author_page.cfm?id=81100389431
Toivonen:	http://portal.acm.org/author_page.cfm?id=81100609333
Yangarber:	http://portal.acm.org/author_page.cfm?id=81100442002

Note that Lemström, Mäkinen and Sillanpää were not listed in the ACM database. Some key data for the PIs are summarized in the following table.

PI	Years in ACM	Publication Count	Citation Count	Docs for download	Downloads 6 weeks	Downloads 12 months	Colleagues
Ukkonen	1979-2011	94	519	8	7	105	49
Hoyer	1999-2010	14	291	5	77	463	12
Hyvärinen	1996-2010	51	851	5	16	84	26
Kärkkäinen	1994-2010	20	167	3	41	358	17
Kaski P	2002-2010	24	92	4	28	178	19
Kaski S	1994-2011	58	286	11	76	640	50
Kivinen	1989-2010	26	388	5	17	248	14
Koivisto	2004-2010	15	89	5	25	157	8
Mannila	1982-2010	135	2854	39	353	2998	104
Myllymäki	1990-2010	32	106	4	7	59	22
Polishchuk	2007-2010	19	18	7	48	226	29
Rousu	1996-2010	25	111	3	8	99	18
Toivonen	1990-2010	36	1839	8	26	304	42
Yangarber	1996-2010	15	169	13	2	187	23

According to co-author lists (“Collaborative colleagues”, linked to author’s personal summaries), this group has several very well established PIs with a large network of “collaborative colleagues”.

CiteSeer

As suggested by the Informatics Europe report, CiteSeer database at <http://citeseerx.ist.psu.edu> was checked for PIs, but the database seems not to be up to date. The following PIs were found on the CiteSeer most cited authors list:

PI	Placement	Cites
Mannila	523	4547
Toivonen	1141	3032
Ukkonen	2923	1701
Hyvärinen	4088	1355
Kaski S	4485	1265
Hoyer	8479	776
Kivinen	8648	764

University of Helsinki
Administrative Publications 80/70
Evaluations

ISBN 978-952-10-7490-5 (PDF)
ISSN 1795-5513 (Online)

Internet address:

http://www.helsinki.fi/julkaisut/aineisto/rc_evaluation2012/hallinnon_julkaisuja_80_70_2012.pdf



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI