

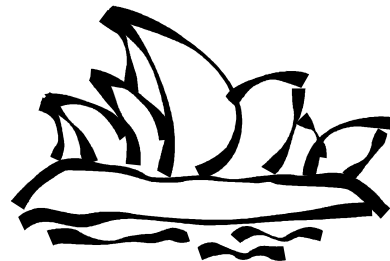
ADCS 2005

Proceedings of the Tenth Australasian
Document Computing Symposium,

12 December 2005

Edited by

Judy Kay, Andrew Turpin and Ross Wilkinson



Proceedings of the Tenth Australasian Document Computing Symposium,
held at The University of Sydney
12 December 2005.

Published by the
School of Information Technologies,
The University of Sydney,
NSW 2006, Australia.

Editors:

Judy Kay
Andrew Turpin
Ross Wilkinson

ISBN 1-86487-787-1

<http://www.cs.rmit.edu.au/~aht/adcs2005>



ADCS 2005 The Tenth Australasian Document Computing Symposium

**Sydney, Australia
12 December 2005**

Chairs' Preface

These proceedings contain the papers of the Tenth Australian Document Computing Symposium hosted by, and held within, the School of Information Technologies at The University of Sydney.

The two keynote addresses, eleven papers and six abstracts reflect the breadth and interest of the Australian research community in the area of document computing.

The eleven papers here were selected from thirteen submissions. Every paper had three peer reviews. Submitted papers were anonymously reviewed at their full length by experts in the area. Dual submissions were explicitly prohibited.

The members of the program committee and the extra reviewers deserve special thanks for their contribution to ADCS2005. Reviewers not listed among the program committee include Halil Ali, Michael Cameron, Richard Cole, Jon Ducrou and Mingfang Wu. We would also like to thank the University of Sydney for their involvement in 2005 as sponsor and host of the event. The event is also supported by the ARC HCSNet.

The symposium includes many formal presentations, but perhaps its greatest benefit lies in the opportunity it provides for document computing practitioners to get together informally and to share ideas and enthusiasm.

Judy Kay
Andrew Turpin
Ross Wilkinson

Symposium Chair

Judy Kay

The University of Sydney

Program Co-chairs

Andrew Turpin

RMIT University

Ross Wilkinson

CSIRO

Program Committee

Vo Ahn

The University of Melbourne

Theresa Dirndorfer Anderson

University of Technology, Sydney

Peter Bruza

Queensland University of Technology

Bob Colomb

University of Queensland

Stijn Dekeyser

University of Southern Queensland

Peter Eklund

University of Wollongong

Tom Gedeon

Australian National University

David Hawking

CSIRO

Judy Kay

University of Sydney

Rob McArthur

CSIRO

Alistair Moffat

University of Melbourne

Gitesh K. Raikundalia

Victoria University

Falk Scholer

RMIT University

Saied Tahagohghi

RMIT University

Jamie Thom

RMIT University

Anne-Marie Vercoustre

INRIA, France

William Webber

The University of Melbourne

Justin Zobel

RMIT University

ADCS Advisory Committee

Peter Bruza

Queensland University of Technology

David Hawking

CSIRO

Judy Kay

University of Sydney

Alistair Moffat

University of Melbourne

Ross Wilkinson

CSIRO

Justin Zobel

RMIT University

Contents

Keynote Address

- From Non-Segmenting Language Processing to Web Language Engineering* 1
Virach Sornlertlamvanich (Thai Computational Linguistics Laboratory, NICT, Thailand)

Invited Presentation

- Document Processing using Formal Concept Analysis* 3
Peter Eklund (University of Wollongong)

Papers (Fully Refereed)

- Web Searcher Interactions with Multiple Federate Content Collections* 4
Amanda Spink (Queensland University of Technology), Bernard J. Jansen (The Pennsylvania State University), Chris Blakely (Infospace, Inc), Sherry Koshman (University of Pittsburgh)
- Document modelling for customised information delivery* 11
Shijian Lu, Cecile Paris and Mingfang WU (CSIRO ICT Centre)
- Readability of French as a Foreign Language and its Uses* 19
Alexandra Uitdenbogerd (RMIT University)
- In Search of Reliable Retrieval Experiments* 26
William Webber and Alistair Moffat (The University of Melbourne)
- Document Expansion versus Query Expansion for Ad-hoc Retrieval* 34
Bodo Billerbeck and Justin Zobel (RMIT University)
- ePOC: Mobile Clinical Information Access and Diffusion in Ambulatory Care Service Settings* 42
Peter Eklund and Jason Sargent (University of Wollongong)
- An Experimental Study of Workflow and Collaborative Document Authoring in Medical Research* 48
Venkata Nallaparaju, Gitesh K. Raikundalia (Victoria University), Caroline Brand, Christopher Bain and Ana Hutchinson (Melbourne Health)
- Applying Formal Concept Analysis to Semantic File Systems Leveraging Wordnet* 56
Ben Martin (The University of Queensland) and Peter Eklund (The University of Wollongong)
- Biomedical Named Entity Recognition System* 64
Jon Patrick and Yefeng Wang (University of Sydney)
- Cross Training and Under Sampling in Categorization of Company Announcements* 72
Cheng Weng and Josiah Poon (The University of Sydney)
- Recommending Geocaches* 76
Andrew Trotman, Timothy Jones and Chris Handley (University of Otago)
- ## Posters (Refereed)
- Evaluating an ontology with OntoClean* 84
Jonathan Yu, James A. Thom and Audrey Tam (RMIT University)
- Document Ranking for Effectiveness-Efficiency Tradeoffs* 85
Vo Ngoc Anh and Alistair Moffat (The University of Melbourne)
- Document Priors for Query Prediction* 86
Steven Garcia, Nicholas Lester and Justin Zobel (RMIT University)
- A Metadata Collection Technique for Documents in WinFS* 87
Stijn Dekeyser University of Southern Queensland)
- Hosting search services for the Australian Government* 88
George Ferizis and David Hawking (CSIRO ICT Centre)
- Automatic Identification of English and Indonesian Parallel Documents* 89
Jelita Asian, Falk Scholer, S.M.M. Tahaghoghi, and Justin Zobel (RMIT University)
- SciFly - Customised Flyers on Demand* 90
Andrew Lampert (CSIRO ICT Centre)