

# **ADCS 2004**

*Proceedings of the Ninth  
Australasian Document Computing Symposium,*

*December 13, 2004*

*Edited by  
Peter Bruza, Alistair Moffat, and Andrew Turpin*

---

---

Proceedings of the Ninth Australasian Document Computing Symposium,  
held at the University of Melbourne,  
December 13, 2004.

Published by the  
Department of Computer Science and Software Engineering,  
The University of Melbourne,  
Victoria 3010, Australia.

Editors:  
Peter Bruza,  
Alistair Moffat,  
Andrew Turpin.

ISBN 0-9757172-0-0

<http://www.cs.mu.oz.au/~alistair/adcs2004/>

# ADCS 2004



## **The Ninth Australasian Document Computing Symposium**

**Melbourne, Australia  
13 December 2004**

### **Chairs' Preface**

These proceedings contain the papers of the Ninth Australasian Document Computing Symposium hosted by, and held within, the Department of Computer Science and Software Engineering at The University of Melbourne.

The keynote address, seven long papers, and five short papers cover a wide range of topics and activities, and reflect the breadth of interest of the Australasian research community in the area of document computing. One quarter of the papers submitted came from outside Australia.

The twelve research papers included here were selected from 21 submissions. Every paper had at least two peer reviews, and many of them three. 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 ADCS 2004. Reviewers not listed among the program committee include John Ducrou, Dawei Song, Xindong Wu, Falk Scholer, and Steven Garcia. We would also like to thank the University of Melbourne for their involvement in 2004 as sponsor and host of this event.

This 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. We look forward to meeting you in Melbourne.

Peter Bruza  
Alistair Moffat  
Andrew Turpin

15 November 2004

## **Symposium Chair**

Alistair Moffat

University of Melbourne

## **Program Co-Chairs**

Peter Bruza

DSTC

Andrew Turpin

University of Melbourne

## **Program Committee**

Vo Ngoc Anh

University of Melbourne

Theresa Dirndorfer Anderson

UTS

Bob Colomb

University of Queensland

Nick Craswell

Microsoft Research

Peter Eklund

University of Wollongong

Tamas D. Gedeon

Australian National University

David Hawking

CSIRO

Judy Kay

University of Sydney

Raymond Lau

QUT and City University of Hong Kong

Rob McArthur

DSTC

Gitesh K. Raikundalia

Victoria University

James Thom

RMIT University

Ross Wilkinson

CSIRO

Hugh Williams

RMIT University

## **ADCS Advisory Committee**

Peter Bruza

DSTC

David Hawking

CSIRO

Judy Kay

University of Sydney

Alistair Moffat

University of Melbourne

Ross Wilkinson

CSIRO

Justin Zobel

RMIT University

## Contents

### Invited Presentation

- Personal Information Management: Stuff I've Seen and Beyond* 1  
Susan Dumais (Microsoft Research)

### Long Papers (Fully Refereed)

- Focused Crawling in Depression Portal Search: A Feasibility Study* 2  
Thanh Tin Tang (Australian National University), David Hawking (CSIRO ICT Centre),  
Nick Craswell (Microsoft Research), and Ramesh S. Sankaranarayana (Australian  
National University)
- On the Effectiveness of Relevance Profiling* 10  
David J. Harper and David Lee (The Robert Gordon University)
- Optimal Structure Weighted Retrieval* 17  
Andrew Trotman (University of Otago)
- Collection-Independent Document-Centric Impacts* 25  
Vo Ngoc Anh and Alistair Moffat (The University of Melbourne)
- Novel Group Awareness Mechanisms for Real-Time Collaborative Document Authoring* 33  
Gitesh K. Raikundalia and Hao Lan Zhang (Victoria University)
- Is CORI Effective for Collection Selection? An Exploration of Parameters, Queries, and  
Data* 41  
Daryl D'Souza, Justin Zobel, and James A. Thom (RMIT University)
- Co-Training on Textual Documents with a Single Natural Feature Set* 47  
Jason Chan, Irena Koprinska, and Josiah Poon (University of Sydney)

### Short Papers (Fully Refereed)

- A Testbed for Indonesian Text Retrieval* 55  
Jelita Asian, Hugh E. Williams, and S. M. M. Tahaghoghi (RMIT University)
- Phrases and Feature Selection in E-Mail Classification* 59  
Elisabeth Crawford (Carnegie Mellon University), Irena Koprinska, and Jon Patrick  
(University of Sydney)
- Towards a New Approach to Tightly Coupled Document Collaboration* 63  
Stijn Dekeyser (University of Southern Queensland)
- GOOD Publishing System: Generic Online/Offline Delivery* 67  
Jacek Radajewski, Sally MacFarlane, and Stijn Dekeyser (University of Southern  
Queensland)
- NLPX – An XML-IR System with a Natural Language Interface* 71  
Alan Woodley and Shlomo Geva (Queensland University of Technology)