CONFERENCE IN RESEARCH AND PRACTICE IN INFORMATION TECHNOLOGY

VOLUME 104

DATABASE TECHNOLOGIES 2010

AUSTRALIAN COMPUTER SCIENCE COMMUNICATIONS, VOLUME 32, NUMBER 3
DATABASE TECHNOLOGIES 2010

Proceedings of the
Twenty-First Australasian Database Conference
(ADC 2010), Brisbane, Australia,
January 2010

Heng Tao Shen and Athman Bouguettaya, Eds.

Volume 104 in the Conferences in Research and Practice in Information Technology Series.
Published by the Australian Computer Society Inc.

Published in association with the ACM Digital Library.
# Table of Contents

Proceedings of the Twenty-First Australasian Database Conference (ADC 2010), Brisbane, Australia, January 2010

Preface .......................................................... vii
Programme Committee ........................................ viii
Organising Committee .......................................... x
Welcome from the Organising Committee ................ xi
CORE - Computing Research & Education .................. xii
ACSW Conferences and the Australian Computer Science Communications ................ xiii
ACSW and ADC 2010 Sponsors .............................. xv

## Invited Papers

A Framework of Data Integration, Knowledge Management and User Behaviour Modelling in Health-care Applications of Diabetes ................................................................. 3  
  Yanchun Zhang, Guangdong Xu, Liping Wang and Kerry Bennett

Contrast Pattern Mining and its Applications ................ 5  
  Kotagiri Ramamohanarao

## Contributed Papers

How Consistent Are Human Judgments of Whether an Open Resource is Educational Material? .... 9  
  Michael C. Harris, James A. Thom and Falk Scholer

Dynamic Framed-Slot ALOHA Anti-Collision using Precise Tag Estimation Scheme .................. 19  
  Prapassara Pupamwiwat and Bela Stantic

Scalable Online Index Construction with Multi-core CPUs ................................................. 29  
  Hiroyuki Yamada and Motomichi Toyama

Recursive Partitioning Method for Trajectory Indexing .................................................. 37  
  Elizabeth Antoine, Kotagiri Ramamohanarao, Jie Shao and Rui Zhang

Efficient Best Path Monitoring in Road Networks For Instant Local Traffic Information .......... 47  
  Shuo Shang, Ke Deng and Kai Zheng

Towards Unifying Advances in Twig Join Algorithms .................................................. 57  
  Nils Grimsmo and Truls A. Bjorklund

A Quantum Interpretation of the View-Update Problem .................................................. 67  
  Christian Flender

Counting Distinct Objects over Sliding Windows .................................................. 75  
  Wenjie Zhang, Ying Zhang, Muhammad Aamir Cheema and Xuemin Lin
Privacy-Aware Access Control in XML Databases ............................................. 85
       Anders H. Landberg, J. Wenny Rahayu and Eric Pardede

Systematic Clustering Method for l-diversity Model ........................................ 93
       Md Enamul Kabir, Hua Wang, Elisa Bertino and Yanxiang Chi

Tuning QoD in Stream Processing Engines .................................................. 103
       Mohamed A. Sharaf, Panos K. Chrysanthis and Alexandros Labrinidis

Building a Dynamic Classifier for Large Text Data Collections ...................... 113
       Pavel Kalinov, Bela Stantic and Abdul Sattar

EP-based Robust Weighting Scheme for Fuzzy SVMs ..................................... 123
       Shaoyi Zhang, Kotagiri Ramamohanarao and James C. Bezdek

Exploit Keyword Query Semantics and Structure of Data for Effective XML Keyword Search ................. 133
       Khanh Nguyen and Jinli Cao

An Analysis of Spreadsheet-Based Services Mashup ........................................ 141
       Dat Dac Hoang, Hye-young Paik and Boualem Benatallah

Optimizing XML Data with View Fragments ............................................... 151
       Jun Liu, Mark Roantree and Zohra Bellahsene

A Real Time Hybrid Pattern Matching Scheme for Stock Time Series ................ 161
       Zhe Zhang, Jian Jiang, Xiaoyan Liu, Ricky Lau, Huaiqing Wang and Rui Zhang

Distributed Data Clustering in Multi-Dimensional Peer-To-Peer Networks .......... 171
       Stefano Lodi, Gianluca Moro and Claudio Sartori

Stock Risk Mining by News ................................................................. 179
       Qi Pan, Hong Cheng, Di Wu, Jeffrey Xu and Yu Yiping Ke

Author Index ......................................................................................... 189
Preface

The series of Australasian Database Conference is an annual forum for exploring novel technical developments and applications of database systems. The 21st Australasian Database Conference, ADC 2010, is held in Brisbane, Australia, as part of Australasian Computer Science Week.

ADC 2010 invited submissions of original contributions in all research areas of databases and its applications. The program committee received forty five submissions of full research papers; each was thoroughly reviewed by at least three PC members or external reviewers. Nineteen papers have been selected for presentation at the conference. In addition, the program committee invited two prominent researchers, Professor Kotagiri Ramamohanarao and Professor Yanchun Zhang for the traditional ADC invited talks.

The ADC PC chairs have also looked at all the accepted papers to select a paper to be awarded the conference’s Best Paper. This year’s Best Paper award goes to two papers. The first paper that won the award is “Counting Distinct Objects over Sliding Windows” by Wenjie Zhang, Ying Zhang, Muhammad Aamir Cheema and Xuemin Lin. The second paper that won the award is “Stock Risk Mining by News” by Qi Pan, Hong Cheng, Di Wu, Jeffrey Yu and Yiping Ke. Congratulations to both teams! We are grateful to the EII for donating the prize for the best paper award.

We would like to take this opportunity to thank all the authors who submitted papers and conference participants for the fruitful discussions. We are grateful to the members of the program committee and external referees for their timely expertise and effort in carefully reviewing the papers.

Heng Tao Shen  
University of Queensland

Athman Bouguettaya  
CSIRO

ADC 2010 Programme Chairs  
January 2010
Programme Committee and Review Panel

Program Chairs

Heng Tao Shen (University of Queensland, Australia)
Athman Bouguettaya (CSIRO, Australia)

Programme Committee

Annika Hinze (University of Waikato, New Zealand)
Anthony Tung (National University of Singapore, Singapore)
Aoying Zhou (East China Normal University, China)
Bin Cui (Peking University, China)
Boualem Benatallah (University of New South Wales, Australia)
Brahim Medjahed (University of Michigan, USA)
Chen Li (University of California at Irvine, USA)
Chengfei Liu (Swinburne University of Technology, Australia)
Christian Bhm (University of Munich, Germany)
Donghui Zhang (Northeastern University, USA)
Egemen Tanin (University of Melbourne, Australia)
Falk Scholer (RMIT, Australia)
Gill Dobbie (University of Auckland, New Zealand)
Guoren Wang (Northeast University, China)
Helen Huang (University of Queensland, Australia)
Helen Paik (University of New South Wales, Australia)
Hong Cheng (Chinese University of Hong Kong, Hong Kong)
Liam O’Brien (NICTA, Australia)
Jenny Zhang (RMIT, Australia)
Jialie Shen (Singapore Management University, Singapore)
Jian Yang (Macquarie University, Australia)
Jianyong Wang (Tsinghua University, China)
Jie Shao (University of Melbourne, Australia)
Jinli Cao (La Trobe University, Australia)
Kian-Lee Tan (National University of Singapore, Singapore)
Markus Stumptner (University of South Australia, Australia)
Miyuki Nakano (Tokyo University, Japan)
Mourad Ouzzani (Purdue University, USA)
Quan Sheng (University of Adelaide, Australia)
Rui Zhang (University of Melbourne, Australia)
Shan Wang (Renmin University, China)
Shazia Sadiq (University of Queensland, Australia)
Shichao Zhang (University of Technology Sydney, Australia)
Selcuk Candan (Arizona State University, USA)
Walid Aref (University of Purdue, USA)
Wei Wang (University of New South Wales, Australia)
Xiangmin Zhou (CSIRO, Australia)
Xuan Zhou (CSIRO, Australia)
Xue Li (University of Queensland, Australia)
Yoshiharu Ishikawa (Nagoya University, Japan)
Yu Qi (Rochester Institute of Technology, USA)
Yu Zheng (Microsoft Research Asia, China)
Zahir Tari (RMIT, Australia)
Zaki Malik (Wayne State University, USA)
Additional Reviewers

Ali, Mohammed Eunus
Asadzadeh, Parvin
Barukh, Moshe
Beheshti, Seyed Mehdi Reza
Behm, Alexander
Borkar, Vinayak
Nutanong, Sarana
Purohit, Vijendra
Shao, Jie
Sun, Haiyang
Talukder, Nilothpal

Tao, Aries
Thom, James
Wang, Yi
Weber, Ingo
Xiao, Xiangye
Xie, Hairuo
Yang, Zhenglu
Ye, Zhen
Zheng, Huiyuan
Zheng, Wenchen
Zhou, Xuan
Organising Committee

Co-Chairs
Dr. Wayne Kelly
Prof. Mark Looi

Budget and Facilities
Mr. Malcolm Corney

Catering and Booklet
Dr. Diane Corney

Sponsorship and Web
Dr. Tony Sahama

Senior Advisors
Prof. Colin Fidge
Prof. Kerry Raymond

Finance and Travel
Ms. Therese Currell
Ms. Carol Richter

Registration
Mr. Matt Williams

DVD and Signage
Mr. Matthew Brecknell

Satchels and T-shirts
Ms. Donna Teague
Welcome from the Organising Committee

On behalf of the Australasian Computer Science Week 2010 (ACSW2010) Organising Committee, we welcome you to this year’s event hosted by the Queensland University of Technology (QUT). Striving to be a "University for the Real World" our research and teaching has an applied emphasis. QUT is one of the largest producers of IT graduates in Australia with strong linkages with industry. Our courses and research span an extremely wide range of information technology, everything from traditional computer science, software engineering and information systems, to games and interactive entertainment.

We welcome delegates from over 21 countries, including Australia, New Zealand, USA, Finland, Italy, Japan, China, Brazil, Canada, Germany, Pakistan, Sweden, Austria, Bangladesh, Ireland, Norway, South Africa, Taiwan and Thailand. We trust you will enjoy both the experience of the ACSW 2010 event and also get to explore some of our beautiful city of Brisbane. At Brisbane’s heart, beautifully restored sandstone buildings provide a delightful backdrop to the city’s glass towers. The inner city clusters around the loops of the Brisbane River, connected to leafy, open-skied suburban communities by riverside bikeways. QUT’s Garden’s Point campus, the venue for ACSW 2010, is on the fringe of the city’s botanical gardens and connected by the Goodwill Bridge to the Southbank tourist precinct.

ACSW2009 consists of the following conferences:

- Australasian Computer Science Conference (ACSC) (Chaired by Bernard Mans and Mark Reynolds)
- Australasian Computing Education Conference (ACE) (Chaired by Tony Clear and John Hamer)
- Australasian Database Conference (ADC) (ADC) (Chaired by Heng Tao Shen and Athman Bouguettaya)
- Australasian Information Security Conference (AISC) (Chaired by Colin Boyd and Willy Susilo)
- Australasian User Interface Conference (AUIC) (Chaired by Christof Lutteroth and Paul Calder)
- Australasian Symposium on Parallel and Distributed Computing (AusPDC) (Chaired by Jinjun Chen and Rajiv Ranjan)
- Australasian Workshop on Health Informatics and Knowledge Management (HIKM) (Chaired by Anthony Maeder and David Hansen)
- Computing: The Australasian Theory Symposium (CATS) (Chaired by Taso Viglas and Alex Potanin)
- Asia-Pacific Conference on Conceptual Modelling (APCCM) (Chaired by Sebastian Link and Aditya Ghose)
- Australasian Computing Doctoral Consortium (ACDC) (Chaired by David Pearce and Rachel Cardell-Oliver).

The nature of ACSW requires the co-operation of numerous people. We would like to thank all those who have worked to ensure the success of ACSW2010 including the Organising Committee, the Conference Chairs and Programme Committees, our sponsors, the keynote speakers and the delegates. Special thanks to Justin Zobel from CORE and Alex Potanin (co-chair of ACSW2009) for his extensive advice and assistance. If ACSW2010 is run even half as well as ACSW2009 in Wellington then we will have done well.

Dr Wayne Kelly and Professor Mark Looi
Queensland University of Technology
ACSW2010 Co-Chairs
January, 2010
CORE welcomes all delegates to ACSW2010 in Brisbane. CORE, the peak body representing academic computer science in Australia and New Zealand, is responsible for the annual ACSW series of meetings, which are a unique opportunity for our community to network and to discuss research and topics of mutual interest. The original component conferences ACSC, ADC, and CATS, which formed the basis of ACSW in the mid 1990s now share the week with seven other events, which build on the diversity of the Australasian computing community.

In 2010, we have again chosen to feature a small number of plenary speakers from across the discipline: Andy Cockburn, Alon Halevy, and Stephen Kisely. I thank them for their contributions to ACSW2010. I also thank the keynote speakers invited to some of the individual conferences. The efforts of the conference chairs and their program committees have led to strong programs in all the conferences again, thanks. And thanks are particularly due to Wayne Kelly and his colleagues for organising what promises to be a strong event.

In Australia, 2009 saw, for the first time in some years, an increase in the number of students choosing to study IT, and a welcome if small number of new academic appointments. Also welcome is the news that university and research funding is set to rise from 2011-12. However, it continues to be the case that per-place funding for computer science students has fallen relative to that of other physical and mathematical sciences, and, while bodies such as the Australian Council of Deans of ICT seek ways to increase student interest in the area, more is needed to ensure the growth of our discipline.

During 2009, CORE continued to work on journal and conference rankings. A key aim is now to maintain the rankings, which are widely used overseas as well as in Australia. Management of the rankings is a challenging process that needs to balance competing special interests as well as addressing the interests of the community as a whole. ACSW2010 includes a forum on rankings to discuss this process. Also in 2009 CORE proposed a standard for the undergraduate Computer Science curriculum, with the intention that it be used for accreditation of degrees in computer science.

CORE’s existence is due to the support of the member departments in Australia and New Zealand, and I thank them for their ongoing contributions, in commitment and in financial support. Finally, I am grateful to all those who gave their time to CORE in 2009; in particular, I thank Gill Dobbie, Jenny Edwards, Alan Fekete, Tom Gedeon, Leon Sterling, and the members of the executive and of the curriculum and ranking committees.

Justin Zobel
President, CORE
January, 2010
ACSW Conferences and the Australian Computer Science Communications

The Australasian Computer Science Week of conferences has been running in some form continuously since 1978. This makes it one of the longest running conferences in computer science. The proceedings of the week have been published as the *Australian Computer Science Communications* since 1979 (with the 1978 proceedings often referred to as Volume 0). Thus the sequence number of the Australasian Computer Science Conference is always one greater than the volume of the Communications. Below is a list of the conferences, their locations and hosts.

2011. Volume 33. Host and Venue - Curtin University of Technology, Perth, WA.

2010. Volume 32. Host and Venue - Queensland University of Technology, Brisbane, QLD.

2008. Volume 30. Host and Venue - University of Wollongong, NSW.
2007. Volume 29. Host and Venue - University of Ballarat, VIC. First running of HDKM.
2006. Volume 28. Host and Venue - University of Tasmania, TAS.
2001. Volume 23. Hosts - Bond University and Griffith University (Gold Coast), Venue - Gold Coast, QLD.
1998. Volume 20. Hosts - University of Western Australia, Murdoch University, Edith Cowan University and Curtin University, Venue - Perth, WA.
1995. Volume 17. Hosts - Flinders University, University of Adelaide and University of South Australia. Venue - Glenelg, SA.
1990. Volume 12. Host and Venue - Monash University, Melbourne, VIC. Joined by Database and Information Systems Conference which in 1992 became ADC (which stayed with ACSW) and ACIS (which now operates independently).
1989. Volume 11. Host and Venue - University of Wollongong, NSW.
1987. Volume 9. Host and Venue - Deakin University, VIC.
1986. Volume 8. Host and Venue - Australian National University, Canberra, ACT.
1983. Volume 5. Host and Venue - University of Sydney, NSW.
1982. Volume 4. Host and Venue - University of Western Australia, WA.
1981. Volume 3. Host and Venue - University of Queensland, QLD.
1980. Volume 2. Host and Venue - Australian National University, Canberra, ACT.
1979. Volume 1. Host and Venue - University of Tasmania, TAS.
1978. Volume 0. Host and Venue - University of New South Wales, NSW.
Conference Acronyms

ACDC        Australasian Computing Doctoral Consortium
ACE         Australasian Computer Education Conference
ACSC        Australasian Computer Science Conference
ACSW        Australasian Computer Science Week
ADC         Australasian Database Conference
AISC        Australasian Information Security Conference
AUIC        Australasian User Interface Conference
APCCM       Asia-Pacific Conference on Conceptual Modelling
AusPDC      Australasian Symposium on Parallel and Distributed Computing (replaces AusGrid)
CATS        Computing: Australasian Theory Symposium
HIKM        Australasian Workshop on Health Informatics and Knowledge Management

Note that various name changes have occurred, which have been indicated in the Conference Acronyms sections in respective CRPIT volumes.
ACSW and ADC 2010 Sponsors

We wish to thank the following sponsors for their contribution towards this conference.

CORE - Computing Research and Education, www.core.edu.au


Queensland University of Technology, www.qut.edu.au


SAP Research, www.sap.com/about/company/research


University of Queensland, www.uq.edu.au
How Consistent Are Human Judgments of Whether an Open Resource is Educational Material?

Michael C. Harris  
James A. Thom  
Falk Scholer

School of Computer Science and Information Technology, RMIT University, GPO Box 2476, Melbourne, Victoria 3001, Australia.  
Email: {michael.c.harris,james.thom,falk.scholer}@rmit.edu.au

Abstract

Systems that filter web search results to return open educational resources need evaluation. The Cranfield method, which is widely used in information retrieval evaluation, can be used as the basis of a model for evaluating such systems. The Cranfield method requires a collection of resources with associated judgments. In this paper, we describe an experiment to collect judgments of whether open resources are educational material. We show experimentally that judges can agree on what resources are educational material, even in the absence of an educational context, and demonstrate that displaying the query used to retrieve the resource makes a judge less likely to rate a resource as educational.

Keywords: Open Educational Resources, Information Retrieval, Inter-rater Agreement

1 Introduction

The increase in the use of technology to support learners has led to a radical increase in the amount of digital teaching and learning material. Reusing existing material can increase the productivity of teaching staff, provide academics with more time to spend on value-added activities such as interacting with students, increase the return on investment of creating high-quality resources, and lead to the improvement of the quality of resources (Harris & Beiers 2005). However, the benefits of reuse cannot be realised if appropriate resources cannot be found.

This work explores a key issue related to the construction of collections appropriate for the evaluation of systems that filter web resources to identify learning material. The concept “likely to support learning” is not well defined, so complete agreement between judges rating resources according to that concept is unlikely. We investigate if people can broadly agree on what resources are likely to support learning in the face of this ambiguity.

This paper is organised as follows. We begin in Section 2 by discussing related research. In Section 3 previous work on agreement evaluation is detailed. In Section 4 we describe the user experiment conducted in this research, the results of which are analysed in Section 5. We then discuss what these results suggest about the ability of judges to agree on whether a resource is educational, and what impact this has on an evaluation methodology for systems that filter educational resources. We conclude by outlining possible future work.

2 Related work

Significant time and money has been spent in the private and public sectors developing and maintaining systems designed to manage educational material, often called learning objects. For example, CANARIE1, the Canadian Advanced Network and Research for Industry and Education alone has spent almost $C30 million on online learning projects, with the design and development of e-learning repositories being a major research theme (MacLeod 2005). These repositories often expose incompatible access interfaces and contain few resources (Neven & Duval 2002). Research on the effective retrieval of educational material has assumed that all resources being searched are learning objects. However, many educational resources are released on the World Wide Web, and clearly there is much more on the Web than just learning material.

When publicly released, these resources are known as Open Educational Resources (OERs) (Wiley 2007), which are generally defined as “technology-enabled, open provision of educational resources for consultation, use and adaptation by a community of users for non-commercial purposes” (UNESCO 2002).

Previous work suggests that when people want to find digital material to support learning, they prefer to use a public search engine, such as Google2 (Harris & Beiers 2005, Griffiths & Brophy 2005). These users may be more satisfied with search engine results if only resources likely to support learning were presented. To provide satisfactory result sets, resources that are unlikely to support learning should not be present. A filter to detect material that is likely to support learning would therefore be useful.

To ensure effectiveness, filters should be systematically evaluated. This evaluation can be carried out by assessing filter performance on a dataset where each resource has been labelled according to whether it should be retained or filtered out. This labelled dataset is called a ground truth or gold standard. This paper examines the establishment of a ground truth for evaluating systems that filter web resources for educational material.

The notion of a ground truth is also used in information retrieval (IR) system evaluation, and this suggests a useful starting point for developing an evaluation methodology for learning material filtering. The ground truth for IR systems evaluation is typically constructed by having relevance judgments assigned to resources by human judges. Systems are

1http://www.canarie.ca
2http://www.google.com
measured based on their ability to approximate the human-assigned relevance judgments. This is known as the Cranfield method after experiments carried out at Cranfield University in the 1960s (Cleverdon 1967).

The Cranfield method requires a collection of documents, a set of queries, and a set of relevance judgments linking the documents and the queries (Hildreth 2001). It is the standard approach used in information retrieval evaluation, and is used for evaluation in, for example, the Text RETrieval Conference (TREC) (Buckley & Voorhees 2005) and the Cross Language Evaluation Forum (CLEF) (Braschler & Peters 2002). This evaluation methodology has also been adapted for use in other retrieval evaluation domains, such as for XML retrieval evaluation (Kazai et al. 2003).

We propose that, using the Cranfield method as a model, systems that filter learning material can be evaluated based on their ability to select those resources that have been categorised by human judges as educational.

Experiments based upon the Cranfield method make the assumption that relevance is a property of resources in relation to a query, independent of the user. Under this assumption, the user and the context of retrieval is completely represented by the query (Saracevic 2007). While there has been debate about the validity of this assumption, it has been a useful starting point for IR experiments in general (Buckley & Voorhees 2005). We make a similar simplifying assumption, that a resource can be judged educational or not, independent of the specific educational context. Though context obviously plays an important role in education, we believe making the assumption of context independence is appropriate for the development of an evaluation methodology for retrieving educational resources. Our experiment explores this assumption.

In IR systems, the ground truth is constructed by assessing relevance, a concept which is complex and multi-dimensional (Saracevic 2007). For filtering educational resources, the ground truth should have assigned judgments of whether resources are educational or likely to support learning, which is also a complex concept. We investigate whether people can broadly agree on what resources are likely to support learning in the face of this complexity.

The level of agreement between raters will assist in deciding how many judgments are needed for each resource to establish an accurate ground truth. The use of a single assessor to judge the relevance of each resource has been a criticism of experiments based on the Cranfield method (Harter 1996). However, it is common in IR experiments to use a test collection where each resource has been assigned a relevance assessment by a single assessor, and this methodology has been shown to be adequate on small collections (Burgin 1992). We further examine whether displaying the query used to retrieve a resource influences judgments and affects agreement.

Collaborative recommendation systems also use ratings that users assign to resources (Adomavicius & Tuzhilin 2005). In collaborative recommendation, ratings are collected from users in production systems and used to suggest resources to other users. Also related are information filtering systems, which find or remove resources from an incoming stream of data based on user profiles (Belkin & Croft 1992). However, these systems differ from judgments in IR evaluation experiments, and to what we propose in this paper, in that we collect ratings based on independent assessments of items; that is, the ratings have no relationship to users’ general profiles.

3 Agreement evaluation methodology

In this section we provide some background on measures of agreement. Alongside our discussion of agreement measures, we present a worked example of each method, in the domain of judging educational material.

3.1 Overlap

The usefulness of larger test collections with single assessments was experimentally supported in relation to the TREC collections by Voorhees, who showed that, despite variability of individual relevance assessments, the relative ranking of systems is stable (Voorhees 1998). In this work, TREC collections were reassessed by additional judges, and the level of agreement between all judges was measured using overlap (Voorhees 1996). Overlap is the mean of the size of the intersection of positive ratings divided by the size of the union of positive ratings for each resource. That is, the average of the number of times both judges rated the resource relevant (for our purposes, rated the resource educational) divided by the number of times either judge rated the resource relevant. Voorhees reports the overlap measures between each of the three judges, and overlap across the three judges. However, the value of this overlap calculated across all judges will decrease as the number of judges increases, as a single dissenting judge counts as disagreement. For this reason, mean pairwise overlap is a more useful measure.

We use the example data from Table 1 to illustrate all agreement measures. Let there be J judges and R resources. For our example we use three judges (J = 3) each rating five resources (R = 5). A value of 0 represents a judgment that a resource was not educational, and a value of 1 represents a judgment that a resource was educational.

Table 1: Example ratings of three judges on five resources

<table>
<thead>
<tr>
<th>Judges</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 0 0 1 1</td>
</tr>
<tr>
<td>2</td>
<td>0 0 0 1 1</td>
</tr>
<tr>
<td>3</td>
<td>0 0 1 1 1</td>
</tr>
</tbody>
</table>

For our example, pairwise overlap can be calculated as follows. We have three judges, and three pairwise comparisons: judge 1 with judge 2, judge 1 with judge 3, and judge 2 with judge 3. Consider judges 1 and 2; they agree that resources 4 and 5 are educational, so the intersection is 2. However, judge 1 rated the resource relevant, and overlap across the three judges. However, the value of this overlap calculated across all judges will decrease as the number of judges increases, as a single dissenting judge counts as disagreement. For this reason, mean pairwise overlap is a more useful measure.

Table 2: Example mean pairwise overlap

<table>
<thead>
<tr>
<th>Judges</th>
<th>Resources</th>
<th>intersect</th>
<th>union</th>
<th>pairwise overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>2 3</td>
<td>3</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>2 4</td>
<td>4</td>
<td>0.500</td>
<td></td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>1 4</td>
<td>1</td>
<td>0.250</td>
<td></td>
</tr>
<tr>
<td>mean pairwise overlap</td>
<td>0.472</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2 Raw agreement

A further simple measure commonly used is raw agreement, which is the proportion of observed agreement to possible agreement. Uebersax says, “Much neglected, raw agreement indices are important descriptive statistics. They have unique common-sense value. A study that reports only simple agreement rates can be very useful; a study that omits them but reports complex statistics may fail to inform readers at a practical level.” (Uebersax 2008)

In the context of assessments of whether resources are OERs, if a random resource is selected from a test collection, and we select a random rater who has judged the resource an OER, what is the probability that another random judge will agree? If the proportion of negative and positive judgments differ greatly, overall agreement will be biased towards the dominant judgments (Kundel & Polansky 2003). This often happens in relevance judgments, where there are dominant judgments (Kundel & Polansky 2003). This proportion of negative and positive judgments differs, has judged the resource an OER, what is the proportion of observed agreement to possible agreement, or one fewer than the total number of ratings, \( p_r + n_r - 1 \), and thus positive agreement for a resource is \( p_r(p_r + n_r - 1) \).

\[
A_{\text{pos}}^+ = \sum_{r=1}^{R} p_r(p_r + n_r - 1)
\]

The observed and possible agreement for our example are shown in Table 4.

<table>
<thead>
<tr>
<th>Resources</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_{\text{obs}} )</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>( A_{\text{pos}} )</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>( A_{\text{obs}}^+ )</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>( A_{\text{pos}}^+ )</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>16</td>
</tr>
</tbody>
</table>

Therefore, we can calculate specific agreement for both the positive \( (A^+) \) and negative \( (A^-) \) cases to be the proportion of observed agreement to possible agreement.

\[
A^- = \frac{A_{\text{obs}}^-}{A_{\text{pos}}^-} \quad A^+ = \frac{A_{\text{obs}}^+}{A_{\text{pos}}^+}
\]

For our example data, this means we have \( A^- = \frac{8}{14} = 0.571 \) and \( A^+ = \frac{10}{14} = 0.625 \).

Overall agreement can similarly be calculated by dividing the observed agreement from both positive and negative judgments by the number of possible agreements.

\[
A = \frac{A_{\text{obs}}^- + A_{\text{obs}}^+}{\sum_{r=1}^{R} (n_r + p_r)(n_r + p_r - 1)}
\]

Of course, \( n_r + p_r \) is constant, the number of judgments made on a resource, and therefore the number of judges. We defined the number of judges earlier as \( J \), so the overall agreement can simplified to the following.

\[
A = \frac{A_{\text{obs}}^- + A_{\text{obs}}^+}{R \cdot J \cdot (J - 1)}
\]

Using our example data, for overall agreement we have \( A = \frac{8 + 10}{(5)(3)(2)} = \frac{18}{30} = 0.6 \).

3.3 Kappa

The disadvantages of the overlap and raw agreement measures are that they are not corrected for chance, and it is not possible to estimate a confidence interval (Kundel & Polansky 2003). The index \( \kappa \) has been developed to address these issues; Cohen’s \( \kappa \) for two raters (Cohen 1960) and Fleiss’ \( \kappa \) for multiple raters (Fleiss 1971).
In calculating Fleiss' $\kappa$ we ask the question, given that we have some set of observed judgments, what agreement would we expect by chance? The proportion of agreement expected by chance can be represented as $P_e$. If we take this value away from perfect agreement, we have the best agreement possible, $1 - P_e$. If we take the chance agreement away from what was observed, which we can call $P_o$, and divide it by the best possible agreement, we have the proportion of agreement that is not due to chance. Therefore, $\kappa$ can be defined as follows, with a value of 1 indicating complete agreement, and a value less than 0 representing agreement less than would be expected by chance.

$$\kappa = \frac{P_e - P_o}{1 - P_e}$$

It is then possible to calculate the standard error, and a confidence interval. Calculating $\kappa$ for our example data, we have $\kappa = 0.196$ and $p = 0.447$. Therefore, though there is some agreement above chance, our example data does not show statistically significant agreement.

The table developed by Landis and Koch (Landis & Koch 1977) is sometimes used as a way to interpret values of $\kappa$. However, the levels chosen are arbitrary, are not applicable across experiments (Sim & Wright 2003) and can lead to unreliable conclusions (Gwet 2001). For these reasons, we do not report our results in relation to the Landis and Koch table.

While $\kappa$ is used as a measure of agreement, it is not a test of the effect of classifying resources using two methods. We may observe significant agreement both when judging resources with a query visible and with the query not visible, but we want to be able to compare the levels of agreement. For this we use Fisher's exact test (Agresti 1992), which tests the null hypothesis that there is no difference in the proportions that raters assign resources to different categories under each condition.

In the next section, we describe our experiment design. As explained in this section, we take the Cranfield method as a starting point for our work.

4 Experiment design

Evaluation of systems that filter e-learning material differs from evaluation of relevance using the Cranfield method in that it seeks to collect classifications of resources according to a concept (supporting learning) as opposed to drawing a relationship (relevance) between a query and a document. To investigate how such classifications should be collected, we conducted a user experiment. The experiment investigates whether human judges can agree on what resources are likely to support learning, and whether visibility of the query used to retrieve the resource has an effect on judgments.

Eight judges were recruited for the experiment. Participants were acquaintances of the first author, from diverse backgrounds, and all had some experience with using web browsers and web interfaces.

A total of 20 resources were judged by the eight judges under one of two conditions: the query used to retrieve the resource being visible ($q$) or not visible ($q'$). Each judge viewed ten resources under condition $q$ and ten under $q'$. A latin squares design was used to control for ordering effects.

4.1 Resource selection

The Flexible Learning Toolboxes are a collection of OERs managed by e-Works\(^3\) that comply with the Shareable Content Object Reference Model (SCORM). A log of queries submitted to the live repository of the e-Works collection was obtained, containing 21339 queries, 7764 of them unique. Queries were drawn at random from the unique queries. If it was judged improbable that submitting a particular query to a search engine would return educational resources, that query was discarded. For example, the queries “rte2606a” and “a*” were discarded. A total of 20 queries were selected for our experiments, and these are shown in the query terms column of Table 5.

While the queries were originally used for seeking resources from an educational repository, the resources used for our experiment were retrieved from a search across the entire web, with each query being submitted to the Yahoo! Search API\(^4\). Alongside each selected query, Table 5 shows the resources used in our experiment, which were selected as follows.

For the first ten queries, resources returned at rank position one were selected for judging. Call this set of resources $R_A$, resources 1 to 10 from Table 5.

To ensure that the collection contained an adequate proportion of resources for which a positive judgment was probable, resources returned using the second ten queries were judged by one of the experimenters, in rank order, according to the same criteria that would ultimately be used by the participants.

For each query, the highest ranked resource judged adequate was used as the selected query. The judgments were made without reference to the search query used to retrieve the resources. Call these resources $R_B$, resources 11 to 20 from Table 5.

4.2 Presentation and user interface

Five resources from $R_A$ and five from $R_B$ were combined and their order randomised to form the first pool, $P_1$. The remaining resources were combined and their order randomised to form $P_2$. The judgment pools contained the following resources:

$P_1 = \{19, 6, 16, 18, 10, 8, 17, 20, 9, 7\}$

$P_2 = \{4, 11, 15, 5, 3, 13, 2, 12, 1, 14\}$

Resources were presented to judges in four ways, as described below.

Group 1) $P_1$ with the query followed by $P_2$ without the query.

Group 2) $P_1$ without the query followed by $P_2$ with the query.

Group 3) $P_2$ with the query followed by $P_1$ without the query.

Group 4) $P_2$ without the query followed by $P_1$ with the query.

For example, Group 1 were presented the resources from $P_1$ with the original search query displayed, and then the resources from $P_2$ were presented without the original search query displayed. The ten resources in $P_1$ and $P_2$ were always presented in the same order.

Two judges were randomly assigned to each group. Resources were presented sequentially via a web interface. Judges were asked to classify resources as likely or unlikely to support learning. Specifically,

\(^3\)http://www.eworks.edu.au

\(^4\)http://developer.yahoo.com/search/web/
## Table 5: Resources in collection.

| resource | query terms | rank | URL |
|----------|-------------|------|-----|---|
| RA1 | Communicate with colleagues and clients in an office environment | 1 | [www.visualdatallc.com/clients.aspx](http://www.visualdatallc.com/clients.aspx) | |
| 4 | Maintain equipment for activities | 1 | [www.govliquidation.com/list/c7484/lna/1.html](http://www.govliquidation.com/list/c7484/lna/1.html) | |
| 5 | Safe beauty | 1 | [www.safeinternetshops.com/beauty.htm](http://www.safeinternetshops.com/beauty.htm) | |
| 6 | Search internet | 1 | [kaftos.com](http://kaftos.com) | |
| 8 | Lifting safely | 1 | [www.smarter.com/--se--qq-Lifting+Safely.html](http://www.smarter.com/--se--qq-Lifting+Safely.html) | |
| 9 | Cash budget | 1 | [www.investopedia.com/terms/c/cashbudget.asp](http://www.investopedia.com/terms/c/cashbudget.asp) | |
| 10 | Process customer complaints | 1 | [www.bcuc.com/Complaint.aspx](http://www.bcuc.com/Complaint.aspx) | |
| RB1 | Costing ingredients | 4 | [www.ces.ncsu.edu/depts/poulsci/tech_manuals/ingredient_sampling.html](http://www.ces.ncsu.edu/depts/poulsci/tech_manuals/ingredient_sampling.html) | |
| 12 | Prepare cook and serve food | 3 | [www.cfsan.fda.gov/~dms/hret2-2.html](http://www.cfsan.fda.gov/~dms/hret2-2.html) | |
| 14 | Treat weeds | 1 | [www.blm.gov/seeds/FAQs/FAQs.htm](http://www.blm.gov/seeds/FAQs/FAQs.htm) | |
| 15 | Library skills | 31 | [www.rock-hill.k12.sc.us/schools/high/sphs/Media/libraryskills.htm#Mod1](http://www.rock-hill.k12.sc.us/schools/high/sphs/Media/libraryskills.htm#Mod1) | |
| 16 | (Plan and conduct and meetings) | 3 | [www.azskillsusa.org/Teachers/meetings.htm](http://www.azskillsusa.org/Teachers/meetings.htm) | |
| 17 | Administer projects | 8 | [sais-jhu.edu/cmtoolkit/issues/evaluation/index.html](http://sais-jhu.edu/cmtoolkit/issues/evaluation/index.html) | |
| 19 | Write simple documents | 1 | [nadoka.vipnet.org:8080/doc/user/08_is1.htm](http://nadoka.vipnet.org:8080/doc/user/08_is1.htm) | |
| 20 | Arts administration | 3 | [en.wikipedia.org/wiki/Arts_administration](http://en.wikipedia.org/wiki/Arts_administration) | |
they were presented with the statement, “The resource is likely/unlikely to support a user to acquire knowledge or a skill,” where the words “likely” and “unlikely” were buttons that recorded the judgment. The judgment interface with the query visible is shown in Figure 1.

An HTML iframe element was used to embed single page resources in the judgment interface. All links within the resources were disabled, and raters evaluated the resources without reference to other web pages.

The resource displayed in Figure 1 is one of the resources used in our experiment, resource 13 from Table 5. The judgment interface without the query visible was identical, save for the removal of the query. Overall, judgments for this resource were split, with four judges believing it was educational material and four believing it was not. When the query used to retrieve the resource was not displayed, three of the four assessors who assessed this resource said it was educational material. However, when judges could see the query, only one of the four assessors judged the resource as educational material.

Each participant answered several questions after each judgment, and after they completed all judgments, including being asked for comments about the resource they had just judged.

5 Analysis

This section reports on the results of our experimental evaluation of rater agreement, and provides analysis of these results. To give a general picture of the judgments, the number of times a resource was judged educational is shown in Figure 2. For our analysis, we begin by investigating rater agreement across all judgments regardless of other factors. Second, we discuss the impact of query visibility on rater agreement. We then report on the influence of resource type, that is, whether the resource was included in the judgment pool as a first ranked resource, or as a resource that was pre-selected to be a likely educational resource. Finally, we conclude by providing a discussion of comments that raters made after judging each resource.

5.1 General rater agreement

The frequency with which a number of raters judged resources as educational is shown in Figure 3. The leftmost bar represents the number of resources that no rater judged as educational, and the rightmost bar represents the number of times that all raters judged a resource educational.

We see a bimodal distribution, with higher frequencies at the extremes. This is as expected if there is a high level of agreement.

The agreement measures between the eight judges observed across all resources are presented in Table 6. All measures suggest a high level of agreement, and the value of $\kappa$ is highly significant. The calculated mean pairwise overlap measure between the eight judges is 0.595, compared with the mean pairwise overlap measure between three assessors of 0.447 shown in Voorhees’ work that justified the use of a single judge in relevance assessments.

5.2 Query visibility

When building a collection for assessing systems using the Cranfield method, an assessor makes a judgment about the relevance of a document to a query. The query is therefore central to the process,
Resource

Positive ratings

0 2 4 6 8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Figure 2: Positive OER judgments by resource and query visibility.

Figure 4: Frequency of positive OER judgments by query visibility.

Figure 3: Overall frequency of positive OER judgments.

and different queries will cause the resource to be judged differently. However, when judging whether a resource is educational, the judgment criteria is stable, and it is unclear what effect query visibility would have on the judging process. Here we report the results of varying query visibility.

Each resource has an a priori probability of being judged likely to support learning. In this case, we are interested in the conditional probability, that is, the probability a resource will be judged likely to support learning given query visibility.

Figure 4 shows the frequency with which a number of raters judged a resource as educational, separated by query visibility. Each resource was judged by four judges under each condition. The leftmost bar shows that, with the query visible five resources received no positive ratings, while without the query visible three resources received no positive ratings. The rightmost bar indicates that all raters judged a resource educational on three occasions when the query was visible and on eight occasions when the query was not visible.

As with Figure 3, bimodal distributions indicate a high level of agreement. The distributions of frequency with and without the query being visible do appear to be generally bimodal, however, inspection suggests that displaying the query makes it less likely that a resource will be judged educational.

The agreement measures when split by query visibility are presented in Table 7. We can see that on all measures except negative agreement, agreement is noticeably higher when the query is not visible, though $\kappa$ is significant in both cases. There is a very high level of positive agreement when the query is not visible, meaning that when the query is not visible raters very often agree that a resource is educational. It appears that judges use different criteria to rate a resource when the query is visible.

Fisher’s exact test indicates that query visibility has a weakly significant effect on judgments ($p = 0.053$).
Table 7: Agreement by query visibility

<table>
<thead>
<tr>
<th>Overlap</th>
<th>query</th>
<th>no query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>0.516</td>
<td>0.685</td>
</tr>
<tr>
<td>Positive</td>
<td>0.667</td>
<td>0.615</td>
</tr>
<tr>
<td>Overall</td>
<td>0.683</td>
<td>0.815</td>
</tr>
<tr>
<td>κ</td>
<td>0.675</td>
<td>0.750</td>
</tr>
</tbody>
</table>

Table 8: Agreement by resource group

<table>
<thead>
<tr>
<th>Overlap</th>
<th>$R_A$</th>
<th>$R_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>0.622</td>
<td>0.583</td>
</tr>
<tr>
<td>Positive</td>
<td>0.805</td>
<td>0.272</td>
</tr>
<tr>
<td>Overall</td>
<td>0.762</td>
<td>0.741</td>
</tr>
<tr>
<td>κ</td>
<td>0.786</td>
<td>0.618</td>
</tr>
</tbody>
</table>

5.3 Resource rank

Figure 5 shows the positive ratings made on each resource, that is, ratings where raters judged the resource educational, separated by query visibility. As described in Subsection 4.1, resources 1 through 10 were included in the judgment pool because they were returned at rank position one in response to a search for the first 10 queries ($R_A$), and the resources 11 through 20 were included in the judgment pool because they were the highest ranked resource judged educational from the results returned in response to the second 10 queries ($R_B$).

The agreement measures when split by resource group are presented in Table 8. On all measures, agreement is lower for resources in $R_B$, and though we see similar values for overlap, positive agreement and overall agreement, $κ$ does not show significant agreement for $R_B$. Negative agreement is particularly low for $R_B$ when compared with $R_A$.

Fisher’s exact test indicates that how a resource was added to the judgment pool has a significant effect on judgments ($p < 0.001$). This means that the proportions of negative and positive judgments are different depending on whether the resource was a first ranked resource or was the highest ranked resource judged to be educational.

5.4 Rater comments

In the judgment interface raters were invited to make comments about their judgments. In total, raters made 91 comments from the 160 judgments. Seven of the eight raters made at least one comment after judging a resource. Approximately a third of the comments make reference to the query used to retrieve the resource. For example, after judging a resource with the query visible one rater said, “Query asking general question; resource for much more specific request which is likely irrelevant. Therefore, easy to judge,” and another said, “query not specific enough,” and “if it was autism and computing skills this would be a useful resource.”

In some cases, the rater stated that they found the resource difficult to judge because the query was not known, “Specific resource and without search terms, difficult to determine whether relevant to query; therefore, difficult to judge.”

These comments appear to suggest that raters find it more difficult to judge whether a resource is educational in the absence of the context given by a query. It might be expected that a more difficult judgment decision would take longer to make; however timing measurements reveal that there is no significant interaction between judging times and query visibility.

6 Discussion and future work

The methodology presented in this paper produces a ground truth that can be used in the evaluation of systems that filter web search results for educational resources. This methodology is based upon the Cranfield method, which is commonly used in IR experiments. Further, we establish how many judges should rate each resource, and whether the queries used in the retrieval of resources should be presented to assessors as part of the judgment interface.

We present a user experiment in which participants judged whether web resources were educational, in a manner similar to the way relevance assessments are collected when building test collections for use in experiments using the Cranfield method.

Our results show a high level of general agreement. Indeed, our results show a level of agreement higher than that used in the IR literature to justify the use of a single assessor. We conclude that, given this high level of agreement, an appropriate methodology for building a test collection for the evaluation of systems that filter for educational resources would involve having resources categorised by a single judge rather than have multiple judges categorise each resource. In particular, for fixed time and number of judges, it would be more useful to judge a larger number of resources than have multiple judgments on fewer resources.

In relation to query visibility, our results show a high level of agreement both with and without the query visible. The level of agreement is higher when the query is not visible, though this is only weakly significant overall. While we found nothing to suggest that the query needs to be displayed, judges did report that they found the task difficult. The task may be made simpler by asking judges to rate resources in an artificial context and then to make a judgment as to whether the resource is educational in other contexts.

This effect is significant when the resource is not the first ranked result. This result is intuitively reasonable, as the search engine used for retrieving the resource has rated the resource as less relevant than other resources, as reflected in its ranking. It may be that the query distracts raters from the task of judging whether a resource is likely to support learning, and causes them to judge relevance instead.

When people use search engines generally, they issue a query and judge how well the documents returned meet their information need. That is, they judge the relevance of returned resources to their query. Therefore, when presented with a resource to judge, and the query that was used to retrieve it, it is unsurprising that their judgments reflect relevance. As we are interested in filtering educational resources, relevance is handled by the search engine, and therefore should be factored out for our purposes.

Our experiment used a fixed description when asking judges to rate resources. Future work could examine what instructions should be given to judges and what effect this may have on agreement.

The selection of resources appropriate for inclusion in a test collection must also be addressed. We contend that it is appropriate to submit queries to a search engine and select returned resources for the collection. In this work, initial queries came from a log of queries submitted to a repository for e-learning material. This is appropriate in that the users were searching for the type of material in which we are interested. However, a user’s search behaviour may be different when searching a specific repository as opposed to the wider web, and thus the queries may not be representative of the sorts of queries that would be submitted to a filtering system. Equally, queries
selected from a general query log, without knowledge of user intent, are likely to be inappropriate.

Resources must then be selected from the ranked resources returned by a search engine in response to a query. Table 5 shows that most of the resources $R_B$ (those included in the pool because they were judged likely to support learning) were ranked in the top 10 results. The mean rank was 7.7 and the median was 3. Also, half the resources in $R_A$ (those included as the first ranked result) were judged educational by a majority of the judges. This suggests that a reasonable percentage of highly ranked resources will be judged likely to support learning, and that we need not have taken the precaution of pre-judging some resources. Therefore, it is appropriate to include the first $N$ results from the returned ranked results in the collection. The results of this preliminary study suggest that $N = 10$ is sufficient.

Our future work will involve the construction of a test collection appropriate for the evaluation of educational resource filters. We will construct such filters, using some of the features from the resources judged in this work as starting points, and evaluate them using the methodology outlined here. For example, resources not judged educational often include a large proportion of links when compared to content, whereas resources judged educational often have large amounts of text separated by headings and a higher proportion of internal links.

One shortcoming of this work, and an opportunity for future work, is that we have only considered single page resources. However, it is likely that multi-page resources are more interesting from a learning point of view. This is related to the concept of granularity as used in reference to reusable learning objects.

References


Author Index

Antoine, Elizabeth, 37
Bellahsene, Zohra, 151
Benatallah, Boualem, 141
Bennett, Kerry, 3
Bertino, Elisa, 93
Bezdek, James C., 123
Bjorklund, Truls A., 57
Bouguettaya, Athman, iii

Cao, Jinli, 133
Cheema, Muhammad Aamir, 75
Cheng, Hong, 179
Chi, Yunxiang, 93
Chrysanthis, Panos K., 103

Deng, Ke, 47
Flender, Christian, 67

Grimsmo, Nils, 57

Harris, Michael C., 9
Hoang, Dat Dac, 141

Jiang, Jian, 161

Kabir, Md Enamul, 93
Kalinov, Pavel, 113
Ke, Yiping, 179

Labrinidis, Alexandros, 103
Landberg, Anders H., 85
Lau, Ricky, 161
Lin, Xuemin, 75
Liu, Jun, 151
Liu, Xiaoyan, 161
Lodi, Stefano, 171

Moro, Gianluca, 171

Nguyen, Khanh, 133
Paik, Hye-young, 141
Pan, Qi, 179
Pardede, Eric, 85
Pupunwiwat, Prapassara, 19
Rahayu, J. Wenny, 85
Ramamohanarao, Kotagiri, 5, 37, 123
Roantree, Mark, 151

Sartori, Claudio, 171
Sattar, Abdul, 113
Scholer, Falk, 9
Shang, Shuo, 47
Shao, Jie, 37
Sharaf, Mohamed A., 103
Shen, Heng Tao, ii
Stantic, Bela, 19, 113

Thom, James A., 9
Toyama, Motomichi, 29

Wang, Hua, 93
Wang, Huaiqing, 161
Wang, Liping, 3
Wu, Di, 179

Xu, Guandong, 3
Xu, Jeffrey, 179

Yamada, Hiroyuki, 29

Zhang, Rui, 37, 161
Zhang, Shaoyi, 123
Zhang, Wenjie, 75
Zhang, Yanchun, 3
Zhang, Ying, 75
Zhang, Zhe, 161
Zheng, Kai, 47
Volume 100 - Safety Critical Systems and Software 2007

Volume 99 - Grid Computing and e-Research 2009

Volume 98 - Information Security 2009
Contains the proceedings of the 3rd Australasian Ontology Workshop (AGO 2009), Gold Coast, Queensland, Australia.

Volume 97 - Computer Science 2009
Contains the proceedings of the Thirty-Second Australasian Computer Science Conference (ACSC2009), Wellington, New Zealand, January 2009.

Volume 96 - Conceptual Modelling 2009
Contains the proceedings of the Fifth Asia-Pacific Conference on Conceptual Modelling (APCCM2009), Wollongong, NSW, Australia, January 2009.

Volume 95 - Computing Education 2009
Contains the proceedings of the Third Australasian Workshop on Health Data and Knowledge Management (HDKM 2009), Wellington, New Zealand, January 2009.

Volume 94 - Theory of Computing 2009

Volume 93 - User Interfaces 2009
Contains the proceedings of the Tenth Australian User Interface Conference (AUIC2009), Wellington, New Zealand, January 2009.

Volume 92 - Database Technologies 2009
Contains the proceedings of the Twenty-fifth Australasian Database Conference (ADC2009), Wellington, New Zealand, January 2009.

Volume 91 - Computer Science 2009
Contains the proceedings of the Thirty-Second Australasian Computer Science Conference (ACSC2009), Wellington, New Zealand, January 2009.

Volume 90 - Advances in Ontologies 2009
Contains the proceedings of the Knowledge Representation Ontology Workshop (KROW 2009), Wellington, New Zealand, January 2009.

Volume 89 - Australian Video
Contains the proceedings of the 1st Australian Video Conference.

Volume 88 - Koli Calling 2009
Contains the proceedings of the 7th Baltic Sea Conference on Computing Education Research.

Volume 87 - Data Mining and Analytics 2008
Contains the proceedings of the 7th Australasian Data Mining Conference (AusDM 2008), Adelaide, Australia. December 2008.

Contains the proceedings of the 13th Australian Conference on Safety Critical Systems and Software, Canberra Australia.

Volume 85 - Advances in Ontologies 2007
Contains the proceedings of the 3rd Australasian Ontology Workshop (AGO 2007), Gold Coast, Queensland, Australia.

Volume 84 - Artificial Intelligence and Data Mining 2007
Contains the proceedings of the 2nd International Workshop on Integrating AI and Data Mining (AIDM 2007), Gold Coast, Australia. December 2007.

Recent Volumes in the CRPIT Series

Listed below are some of the latest volumes published in the ACS Series Conferences in Research and Practice in Information Technology. The full text of most papers (in either PDF or Postscript format) is available at the series website http://crpit.com.

Volume 100 - Safety Critical Systems and Software 2007

Volume 99 - Grid Computing and e-Research 2009