

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Mohand Boughanem Catherine Berrut
Josiane Mothe Chantal Soule-Dupuy (Eds.)

Advances in Information Retrieval

31th European Conference on IR Research, ECIR 2009
Toulouse, France, April 6-9, 2009
Proceedings

Volume Editors

Mohand Boughanem
Université de Toulouse - IRIT
118 Route de Narbonne, 31062, Toulouse Cedex 4, France
E-mail: Mohand.Boughanem@irit.fr

Catherine Berrut
Université Joseph Fourier, Laboratoire d'Informatique de Grenoble
BP 53, 38041 Grenoble Cedex 9, France
E-mail: catherine.berrut@imag.fr

Josiane Mothe
Université de Toulouse - IRIT
118 Route de Narbonne, 31062 Toulouse Cedex 4, France
E-mail: Josiane.Mothe@irit.fr

Chantal Soule-Dupuy
Université de Toulouse - IRIT
118 Route de Narbonne, 31062 Toulouse Cedex 4, France
E-mail: Chantal.Soule-Dupuy@irit.fr

Library of Congress Control Number: Applied for

CR Subject Classification (1998): H.3, H.2, I.2.6-7, H.4, H.2.8, I.7

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-642-00957-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-00957-0 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

springer.com

© Springer-Verlag Berlin Heidelberg 2009
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12642687 06/3180 5 4 3 2 1 0

Preface

These proceedings contain the papers presented at ECIR 2009, the 31st European Conference on Information Retrieval. The conference was organized by the Institut de Recherche en Informatique de Toulouse (IRIT), in cooperation with the Information Retrieval Specialist Group of the British Computer Society (BCS-IRSG) and was supported by the Special Interest Group on Information Retrieval (ACM SIGIR), the French Association of Information Retrieval (ARIA) and E-IRSG. It was held during April 6–9, 2007 in Toulouse, hosted by the Pierre Baudis congress center.

ECIR 2009 confirmed the growth trend of submissions, with 188 full-paper submissions.

Many of these submissions came from outside Europe, and 25 countries were represented in the final ECIR 2009 program, reflecting the international popularity and reputation of the conference series. Furthermore the traditional strong focus on students was well preserved, with their participation being great.

All submitted papers were reviewed by at least three members of the international Program Committee. Out of the 188 papers 42 were selected as full research papers and a further 18 were accepted as short research papers. Short papers were given 8 pages in the proceedings (instead of 12). Out of these 60 selected papers, 30 have a student as the primary author. The accepted papers themselves come from a mixture of universities, research institutes and commercial organizations.

There was also a separate call for posters. Each of the 60 poster submissions was reviewed by three Program Committee members, as with paper submissions. Twenty-five posters were accepted for presentation.

Taken together, these papers cover the state of the art in information retrieval, including topics such as retrieval models, evaluation and Web search, and emerging topics such as learning to rank, expert search and opinion detection.

The success of such a conference is not possible without the huge effort of several people and organizations. We owe a great thanks to our various committees—Program and Poster Committees, Tutorial and Workshop Committee and Award Committees—for their hard work in ensuring the quality of the ECIR 2009 program. We really appreciate the support, the expertise and the effort given so freely. We thank the researchers who submitted their results, and the members of the local Organizing Committee for their hard work over many months.

We also thank our sponsoring institutions, Google, Matrixware Information Services, Microsoft Research, Yahoo! Research, Exalead, GDR I3, Université Paul Sabatier, ARIA, Inforsid and Région Midi-Pyrénées for their support of ECIR 2009.

We are grateful to our keynote speakers, Bruce W. Croft, Gregory Grefenstette and Maarten de Rijke, for their stimulating contributions to the conference.

Our final thanks go to the IRIT communication group, for dealing with all local arrangements and for sorting out the registration process and the preparation of the conference venues with efficiency and tact, and the many local volunteers for their huge contribution to the smooth running of ECIR 2009.

April 2009

Mohand Boughanem
Catherine Berrut
Josiane Mothe
Chantal Soulé-Dupuy

Organization

ECIR 2009 was organized by the IRIT Laboratory and the Université de Toulouse, France.

Organizing Committee

Conference and Program Chair	Mohand Boughanem (Université de Toulouse, IRIT, France)
Poster Chair	Catherine Berrut (Université Joseph Fourier, LIG, France)
Tutorials and Workshop Chair	Josiane Mothe (Université de Toulouse, IRIT, France), Chantal Soulé-Dupuy (Université de Toulouse, IRIT, France)
Website	Max Chevalier (Université de Toulouse, IRIT, France) Cécile Laffaire (IRIT, France)
Local Organization	Guillaume Cabanac (Université de Toulouse, IRIT, France) Marie-Franoise Canut (Université de Toulouse, IRIT, France) Ludovic Chacun (IRIT, France) Max Chevalier (Université de Toulouse, IRIT, France) Véronique Debats (IRIT, France) Corinne Doumerc (IRIT, France) Gilles Hubert (Université de Toulouse, IRIT, France) Christine Julien (Université de Toulouse, IRIT, France) Arlind Kopliku (Université de Toulouse, IRIT, France) Cécile Laffaire (IRIT, France) Sabyne Lartigue (IRIT, France) Karen Pinel-Sauvagnat (Université de Toulouse, IRIT, France) Lynda Tamine-Lechani (Université de Toulouse, IRIT, France)

Full Papers and Posters Program Committee

Agichtein Eugene	Emory University (USA)
Agosti Maristella	University of Padova (Italy)
Allan James	University of Massachusetts (USA)
Amati Giambattista	Fondazione Ugo Bordoni (Italy)
Amini Massih-Reza	Université Pierre et Marie Curie (France)
Amitay Einat	IBM Research, Haifa Lab (Israel)
Azzopardi Leif	University of Glasgow (UK)
Baeza-Yates Ricardo	Yahoo! Research (Spain)
Bailey Alex	Google Switzerland GmbH (Switzerland)
Barreiro Álvaro	University of A Coruña (Spain)
Barry Smyth Barry	University College Dublin (Ireland)
Basili Roberto	University of Rome, Tor Vergata (Italy)
Bast Holger	MPI Informatik (Germany)
Beaulieu Micheline	University of Sheffield (UK)
Beigbeder Michel	Ecole des Mines de Saint-Etienne (France)
Belkin Nicholas	Rutgers University (USA)
Bennett Paul N.	Microsoft Research (USA)
Bilenko Misha	Microsoft Research (USA)
Bordogna Gloria	CNR (Italy)
Bothma Theo	University of Pretoria (South Africa)
Brajnik Giorgio	University of Udine (Italy)
Bruza Peter	Queensland University of Technology (Australia)
Buettcher Stefan	Google Inc. (USA)
Buntine Wray	NICTA (Australia)
Cacheda Fidel	University of A Coruña (Spain)
Callan Jamie	Carnegie Mellon University (USA)
Caracciolo Caterina	Food and Agriculture Organization of the UN (Italy)
Carmel David	IBM Research, Haifa Lab (Israel)
Carpineto Claudio	Fondazione Ugo Bordoni (Italy)
Castillo Ocaranza Carlos Alberto Alejandro	Yahoo! Research (Spain)
Chevalier Max	Université Paul Sabatier - IRIT (France)
Chiarabella Yves	CLIPS-IMAG équipe MRIM (France)
Chirita Paul-Alexandru	Adobe Systems Inc. (Romania)
Clarke Charles L.A.	University of Waterloo (Canada)
Clinchant Stephane	Xerox Research Center Europe (France)
Crestani Fabio	University of Lugano (Switzerland)
Croft Bruce	University of Massachusetts, Amherst (USA)

de la Fuente Pablo	Universidad de Valladolid (Spain)
de Rijke Maarten	University of Amsterdam (The Netherlands)
de Vries Arjen P.	CWI (The Netherlands)
Efthimiadis Efthimis N.	University of Washington (USA)
Fernández-Luna Juan M.	University of Granada (Spain)
Ferragina Paolo	University of Pisa (Italy)
Ferro Nicola	University of Padua (Italy)
Fox Edward	Virginia Tech (USA)
Fürnkranz Johannes	TU Darmstadt (Germany)
Fuhr Norbert	Universität Duisburg-Essen (Germany)
Gallinari Patrick	University Pierre et Marie Curie, Paris 6 (France)
Gaussier Eric	Université J. Fourier-LIG (France)
Goker Ayse	City University London (UK)
Gonzalo Julio	UNED (Spain)
Gray Alasdair J.G.	University of Glasgow (UK)
Grefenstette Gregory	EXALEAD (France)
Greiff Warren Richard	The MITRE Corporation (USA)
Grossman David Adam	IIT (USA)
Gulli Antonio	Ask.com, CTO Europe and UK (Italy)
Gurrin Cathal	Dublin City University (Ireland)
Hansen Preben	SICS - Swedish Institute of Computer Science (SE)
Harman Donna	NIST (USA)
He Ben	University of Glasgow (UK)
Hiemstra Djoerd	University of Twente (The Netherlands)
Hotho Andreas	University of Kassel (Germany)
Hubert Gilles	Université Paul Sabatier - IRIT (France)
Huibers Theo	University of Twente (The Netherlands)
Hull David A.	Google (USA)
Ingwersen Peter	Danmarks Biblioteksskole - Royal (Denmark)
Jansen Bernard	Penn State (USA)
Jarvelin Kal	University of Tampere (Finland)
Johnson Frances	Manchester Metropolitan University (UK)
Joho Hideo	University of Glasgow (UK)
Jones Gareth	Dublin City University (Ireland)
Jose Joemon M.	University of Glasgow (UK)
Kamps Jaap	University of Amsterdam (The Netherlands)
Karlgren Jussi	SICS (SE)
Kelly Diane	University of North Carolina (USA)
Koubarakis Manolis	National and Kapodistrian University of Athens (Greece)
Kraaij Wessel	TNO (The Netherlands)
Kumar Ravi	Yahoo! Research (USA)

Lalmas Mounia	University of Glasgow (UK)
Landoni Monica Angela	USI University of Lugano (Switzerland)
Larsen Birger	Royal School of Library and Information Science (Denmark)
Li Xuelong	Birkbeck College, University of London (UK)
Losada David E.	University of Santiago de Compostela (Spain)
Macdonald Craig	University of Glasgow (UK)
MacFarlane Andrew	City University London (UK)
Maggini Marco	University of Siena (Italy)
Melucci Massimo	University of Padua (Italy)
Micarelli Alessandro	University of Rome 3 (Italy)
Mizzaro Stefano	University of Udine (Italy)
Mladenic Dunja	J. Stefan Institute (Slovenia)
Moens Marie-Francine	Katholieke Universiteit Leuven (Belgium)
Moffat Alistair	University of Melbourne (Australia)
Monz Christof	Department of Computer Science Queen Mary, University of London (UK)
Murdock Vanessa	Yahoo! Research (Spain)
Murray G. Craig	University of Maryland (USA)
Nie Jian-Yun	DIRO, Université de Montréal (Canada)
Oakes Michael	Philip University of Sunderland (UK)
Ogilvie Paul	mSpoke (USA)
Osinski Stanislaw	Carrot Search (Poland)
Ounis Iadh	University of Glasgow (UK)
Pasi Gabriella	Università degli Studi di Milano Bicocca (Italy)
Pedersen Jan O.	A9 (USA)
Pirkola Ari	University of Tampere (Finland)
Plachouras Vassilis	Yahoo! Research (Spain)
Rauber Andreas	Department of Software Technology and Interactive Systems Vienna University of Technology (Austria)
Robertson Stephen E	Microsoft Research Cambridge (UK)
Ruthven Ian	University of Strathclyde (UK)
Schenkel Ralf	Max-Planck-Institut Informatik (Germany)
Sebastiani Fabrizio	Consiglio Nazionale delle Ricerche (Italy)
Sèdes Florence	Université Paul Sabatier-IRIT (France)
Semeraro Giovanni	University of Bari "Aldo Moro" (Italy)
Shen Jialie	Singapore Management University (Singapore)
Siersdorfer Stefan	University of Sheffield (UK)
Silva Mário J.	University of Lisbon (Portugal)
Silvestri Fabrizio	Italian National Research Council (CNR) (Italy)
Smeaton Alan	Dublin City University (Ireland)
Smucker Mark D.	University of Waterloo (Canada)

Snasel Vaclav	VSB-Technical University of Ostrava (Czech Republic)
Soboroff Ian	NIST (USA)
Song Dawei	The Open University (UK)
Song Ruihua	Microsoft Research Asia (CN)
Sormunen Eero	University of Tampere (Finland)
Spink Amanda	Queensland University of Technology (Australia)
Stokes Nicola	University College Dublin (Ireland)
Tamine-Lechani Lynda	Université Paul Sabatier - IRIT (France)
Theobald Martin	Stanford University (USA)
Thiel Ulrich	Franceaunhofer (Germany)
Tombros Anastasios	Queen Mary University of London (UK)
Vakkari Pertti	University of Tampere (Finland)
van Rijsbergen Keith	University of Glasgow (UK)
Vechtomova Olga	University of Waterloo (Canada)
Weiss Dawid	Poznan University of Technology (Poland)
Westerveld Thijs	Teezir Search Solutions (The Netherlands)
White Ryen William	Microsoft Corporation (USA)
Yang Tao	ASK.COM/UCSB (USA)
Yom-Tov Elad	IBM Research (Israel)
Zhang Dell	Birkbeck, University of London (UK)
Zobel Justin	University of Melbourne (Australia)

Additional Reviewers

Bai Jing	Yahoo! inc.(USA)
Bashir Shariq	Vienna University of Technology (Austria)
Basile Pierpaolo	University of Bari (Italy)
Biancalana Claudio	University of Rome 3 (Italy)
Burger John	The MITRE Corporation (USA)
Cabanac Guillaume	Université Paul Sabatier - IRIT (France)
Carman Mark	University of Lugano (Switzerland)
Celikik Marjan	Max Planck Institute for Informatics (Germany)
Daoud Mariam	Université de Toulouse, IRIT (France)
de Gemmis Marco	University of Bari (Italy)
Di Buccio Emanuele	University of Padua (Italy)
Di Nunzio Giorgio Maria	University of Padua (Italy)
Gasparetti Fabio	University of Rome 3 (Italy)
Castellano Giovanna	University of Bari (Italy)
Gerani Shima	University of Lugano (Switzerland)
Gurrin Cathal	Dublin City University (Ireland)
Gwadera Robert	University of Lugano (Switzerland)
Hopfgartner Frank	University of Glasgow (UK)

XII Organization

Keikha Mostafa	University of Lugano (Switzerland)
Lops Pasquale	University of Bari (Italy)
Moshfeghi Yashar	University of Glasgow (UK)
Olivares Ximena	Universitat Pompeu Fabra, Barcelona (Spain)
Pinel-Sauvagnat Karen	Université Paul Sabatier - IRIT (France)
Rigutini Leonardo	University of Siena (Italy)
Rusinov Momchil	Max Planck Institute for Informatics (Germany)
Shao Xiaowei	Tokyo University (Japan)
Silvello Gianmaria	University of Padua (Italy)
Sitbon Laurianne	NICTA (Australia)
Surdeanu Mihai	Yahoo! Research, Barcelona (Spain)
Tannebaum Wolfgang	Vienna University of Technology (Austria)
Varlaro Antoni	Exhicon ICT (Italy)
Villa Robert	University of Glasgow (UK)
Weber Ingmar	EPFL Lausanne (Switzerland)

Workshop and Tutorials Committee

Baeza-Yates Ricardo	Yahoo! Research, Spain
Clarkes Charles	University of Waterloo, Canada
Croft W. Bruce	University of Massachusetts Amherst, USA
Jarvelin Kalervo	University of Tampere, Finland
Liu Tie-Yan	Microsoft Research Asia, China
Pasi Gabriella	University of Milano Bicocca, Italy
Rueger Stefan	The Open University, UK

Best Paper Award Committee

Kraaij Wessel	TNO, The Netherlands
Ounis Iadh	University of Glasgow, UK
Pasi Gabriella	University of Milano Bicocca, Italy

Best Student Paper Award Committee

Azzopardi Leif	University of Glasgow, UK
Mulhem Philippe	Laboratoire d'Informatique de Grenoble, France
Murdock Vanesa	Yahoo! Research, Spain

Best Poster Committee

Crestani Fabio	University of Lugano, Switzerland
Rollecke Thomas	Queen Mary University of London, UK
Tamine-Lechani Lynda	Université Paul Sabatier - IRIT, France

Sponsoring Institutions



Organising Institution



In Co-operation with



Table of Contents

Invited Presentations

Query Evolution	1
<i>W. Bruce Croft</i>	
Searching User Generated Content: What's Next?	2
<i>Maarten de Rijke</i>	
Upcoming Industrial Needs for Search	3
<i>Gregory Grefenstette</i>	

Retrieval Model I

Mean-Variance Analysis: A New Document Ranking Theory in Information Retrieval	4
<i>Jun Wang</i>	
Risk-Aware Information Retrieval	17
<i>Jianhan Zhu, Jun Wang, Michael Taylor, and Ingemar J. Cox</i>	
A Comparative Study of Utilizing Topic Models for Information Retrieval	29
<i>Xing Yi and James Allan</i>	

Collaborative IR/Filtering

Synchronous Collaborative Information Retrieval: Techniques and Evaluation	42
<i>Colum Foley and Alan F. Smeaton</i>	
Movie Recommender: <i>Semantically Enriched</i> Unified Relevance Model for Rating Prediction in Collaborative Filtering	54
<i>Yashar Moshfeghi, Deepak Agarwal, Benjamin Piwowarski, and Joemon M. Jose</i>	
Revisiting IR Techniques for Collaborative Search Strategies	66
<i>Hideo Joho, David Hannah, and Joemon M. Jose</i>	

Learning

Active Sampling for Rank Learning via Optimizing the Area under the ROC Curve	78
<i>Pinar Donmez and Jaime G. Carbonell</i>	

Regression Rank: Learning to Meet the Opportunity of Descriptive Queries 90
Matthew Lease, James Allan, and W. Bruce Croft

Active Learning Strategies for Multi-Label Text Classification 102
Andrea Esuli and Fabrizio Sebastiani

Joint Ranking for Multilingual Web Search 114
Wei Gao, Cheng Niu, Ming Zhou, and Kam-Fai Wong

Multimedia - Metadata

Diversity, Assortment, Dissimilarity, Variety: A Study of Diversity Measures Using Low Level Features for Video Retrieval 126
Martin Halvey, P. Punitha, David Hannah, Robert Villa, Frank Hopfgartner, Anuj Goyal, and Joemon M. Jose

Bayesian Mixture Hierarchies for Automatic Image Annotation 138
Vassilios Stathopoulos and Joemon M. Jose

XML Multimedia Retrieval: From Relevant Textual Information to Relevant Multimedia Fragments 150
Mouna Torjmen, Karen Pinel-Savagnat, and Mohand Boughanem

Effectively Searching Maps in Web Documents 162
Qingzhao Tan, Prasenjit Mitra, and C. Lee Giles

Expert Search - Advertising

Enhancing Expert Finding Using Organizational Hierarchies 177
Maryam Karimzadehgan, Ryen W. White, and Matthew Richardson

A Vector Space Model for Ranking Entities and Its Application to Expert Search 189
Gianluca Demartini, Julien Gaugaz, and Wolfgang Nejdl

Sentiment-Oriented Contextual Advertising 202
Teng-Kai Fan and Chia-Hui Chang

Lexical Graphs for Improved Contextual Ad Recommendation 216
Symeon Papadopoulos, Fotis Menemenis, Yiannis Kompatsiaris, and Ben Bratu

Retrieval Model II

A Probabilistic Retrieval Model for Semistructured Data 228
Jinyoung Kim, Xiaobing Xue, and W. Bruce Croft

Model Fusion in Conceptual Language Modeling	240
<i>Loic Maisonnasse, Eric Gaussier, and Jean-Pierre Chevallet</i>	
Graded-Inclusion-Based Information Retrieval Systems	252
<i>Patrick Bosc, Vincent Claveau, Olivier Pivert, and Laurent Ughetto</i>	
Multidimensional Relevance: A New Aggregation Criterion	264
<i>Célia da Costa Pereira, Mauro Dragoni, and Gabriella Pasi</i>	

Evaluation

Using Multiple Query Aspects to Build Test Collections without Human Relevance Judgments	276
<i>Miles Efron</i>	
If I Had a Million Queries	288
<i>Ben Carterette, Virgil Pavlu, Evangelos Kanoulas, Javed A. Aslam, and James Allan</i>	
The Combination and Evaluation of Query Performance Prediction Methods	301
<i>Claudia Hauff, Leif Azzopardi, and Djoerd Hiemstra</i>	

Opinion Detection

Investigating Learning Approaches for Blog Post Opinion Retrieval	313
<i>Shima Gerani, Mark J. Carman, and Fabio Crestani</i>	
Integrating Proximity to Subjective Sentences for Blog Opinion Retrieval	325
<i>Rodrygo L.T. Santos, Ben He, Craig Macdonald, and Iadh Ounis</i>	
Adapting Naive Bayes to Domain Adaptation for Sentiment Analysis . . .	337
<i>Songbo Tan, Xueqi Cheng, Yuefen Wang, and Hongbo Xu</i>	

Web IR

PathRank: Web Page Retrieval with Navigation Path	350
<i>Jianqiang Li and Yu Zhao</i>	
Query Expansion Using External Evidence	362
<i>Zhijun Yin, Milad Shokouhi, and Nick Craswell</i>	
Selective Application of Query-Independent Features in Web Information Retrieval	375
<i>Jie Peng and Iadh Ounis</i>	

Measuring the Search Effectiveness of a Breadth-First Crawl 388
Dennis Fetterly, Nick Craswell, and Vishwa Vinay

Representation

Using Contextual Information to Improve Search in Email Archives 400
Wouter Weerkamp, Krisztian Balog, and Maarten de Rijke

Part of Speech Based Term Weighting for Information Retrieval 412
Christina Lioma and Roi Blanco

Word Particles Applied to Information Retrieval 424
Evandro B. Gouvêa and Bhiksha Raj

“They Are Out There, If You Know Where to Look”: Mining
 Transliterations of OOV Query Terms for Cross-Language Information
 Retrieval 437
Raghavendra Udupa, Saravanan K, Anton Bakalov, and Abhijit Bhole

Clustering / Categorization

E-Mail Classification for Phishing Defense 449
Wilfried N. Gansterer and David Pölz

Multi-facet Rating of Product Reviews 461
Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani

Exploiting Surface Features for the Prediction of Podcast Preference . . . 473
Manos Tsagkias, Martha Larson, and Maarten de Rijke

Distributed IR

A Topic-Based Measure of Resource Description Quality for Distributed
 Information Retrieval 485
Mark Baillie, Mark J. Carman, and Fabio Crestani

Simple Adaptations of Data Fusion Algorithms for Source Selection 497
Georgios Paltoglou, Michail Salampasis, and Maria Satratzemi

Document Compaction for Efficient Query Biased Snippet
 Generation 509
Yohannes Tsegay, Simon J. Puglisi, Andrew Turpin, and Justin Zobel

Short Papers

Organizing Suggestions in Autocompletion Interfaces 521
*Alia Amin, Michiel Hildebrand, Jacco van Osssenbruggen,
 Vanessa Evers, and Lynda Hardman*

Building a Graph of Names and Contextual Patterns for Named Entity Classification	530
<i>César de Pablo-Sánchez and Paloma Martínez</i>	
Combination of Documents Features Based on Simulated Click-through Data	538
<i>Ali Mohammad Zareh Bidoki and James A. Thom</i>	
Discovering Association Rules on Experiences from Large-Scale Blog Entries	546
<i>Takeshi Kurashima, Ko Fujimura, and Hidenori Okuda</i>	
Extracting Geographic Context from the Web: GeoReferencing in MyMoSe	554
<i>Álvaro Zubizarreta, Pablo de la Fuente, José M. Cantera, Mario Arias, Jorge Cabrero, Guido García, César Llamas, and Jesús Vegas</i>	
What Else Is There? Search Diversity Examined	562
<i>Mark Sanderson, Jiayu Tang, Thomas Arni, and Paul Clough</i>	
Using Second Order Statistics to Enhance Automated Image Annotation	570
<i>Ainhoa Llorente and Stefan Rürger</i>	
Classifying and Characterizing Query Intent	578
<i>Azin Ashkan, Charles L.A. Clarke, Eugene Agichtein, and Qi Guo</i>	
Design and Evaluation of a University-Wide Expert Search Engine	587
<i>Ruud Liebrechts and Toine Bogers</i>	
A Study of the Impact of Index Updates on Distributed Query Processing for Web Search	595
<i>Charalampos Sarigiannis, Vassilis Plachouras, and Ricardo Baeza-Yates</i>	
Generic and Spatial Approaches to Image Search Results Diversification	603
<i>Monica Lestari Paramita, Jiayu Tang, and Mark Sanderson</i>	
Studying Query Expansion Effectiveness	611
<i>Ben He and Iadh Ounis</i>	
Correlation of Term Count and Document Frequency for Google N-Grams	620
<i>Martin Klein and Michael L. Nelson</i>	
A Cost-Aware Strategy for Query Result Caching in Web Search Engines	628
<i>Ismail Sengor Altingovde, Rifat Ozcan, and Özgür Ulusoy</i>	

Quality-Oriented Search for Depression Portals	637
<i>Thanh Tang, David Hawking, Ramesh Sankaranarayana, Kathleen M. Griffiths, and Nick Craswell</i>	
Evaluation of Text Clustering Algorithms with N-Gram-Based Document Fingerprints	645
<i>Javier Parapar and Álvaro Barreiro</i>	
Exploiting Flickr Tags and Groups for Finding Landmark Photos	654
<i>Rabeh Abbasi, Sergey Chernov, Wolfgang Nejdl, Raluca Paiu, and Steffen Staab</i>	
Refining Keyword Queries for XML Retrieval by Combining Content and Structure	662
<i>Desislava Petkova, W. Bruce Croft, and Yanlei Diao</i>	
Posters	
Cover Coefficient-Based Multi-document Summarization	670
<i>Gonenc Ercan and Fazli Can</i>	
A Practitioner’s Guide for Static Index Pruning	675
<i>Ismail Sengor Altıngöve, Rifat Özcan, and Özgür Ulusoy</i>	
Revisiting N-Gram Based Models for Retrieval in Degraded Large Collections	680
<i>Javier Parapar, Ana Freire, and Álvaro Barreiro</i>	
A Simple Linear Ranking Algorithm Using Query Dependent Intercept Variables	685
<i>Nir Ailon</i>	
Measurement Techniques and Caching Effects	691
<i>Stefan Pohl and Alistair Moffat</i>	
On Automatic Plagiarism Detection Based on <i>n</i> -Grams Comparison	696
<i>Alberto Barrón-Cedeño and Paolo Rosso</i>	
Exploiting Visual Concepts to Improve Text-Based Image Retrieval	701
<i>Sabrina Tollari, Marcin Detyniecki, Christophe Marsala, Ali Fakeri-Tabrizi, Massih-Reza Amini, and Patrick Gallinari</i>	
Choosing the Best MT Programs for CLIR Purposes – Can MT Metrics Be Helpful?	706
<i>Kimmo Kettunen</i>	
Entropy-Based Static Index Pruning	713
<i>Lei Zheng and Ingemar J. Cox</i>	

Representing User Navigation in XML Retrieval with Structural Summaries	719
<i>Mir Sadek Ali, Mariano P. Consens, and Birger Larsen</i>	
ESUM: An Efficient System for Query-Specific Multi-document Summarization	724
<i>C. Ravindranath Chowdary and P. Sreenivasa Kumar</i>	
Using WordNet's Semantic Relations for Opinion Detection in Blogs....	729
<i>Malik Muhammad Saad Missen and Mohand Boughanem</i>	
Improving Opinion Retrieval Based on Query-Specific Sentiment Lexicon	734
<i>Seung-Hoon Na, Yeha Lee, Sang-Hyob Nam, and Jong-Hyeok Lee</i>	
Automatically Maintained Domain Knowledge: Initial Findings	739
<i>Deirdre Lungley and Udo Kruschwitz</i>	
A Framework of Evaluation for Question-Answering Systems	744
<i>Sarra El Ayari and Brigitte Grau</i>	
Combining Content and Context Similarities for Image Retrieval	749
<i>Xiaojun Wan</i>	
Investigating the Global Semantic Impact of Speech Recognition Error on Spoken Content Collections.....	755
<i>Martha Larson, Manos Tsagkias, Jiyin He, and Maarten de Rijke</i>	
Supervised Semantic Indexing	761
<i>Bing Bai, Jason Weston, Ronan Collobert, and David Grangier</i>	
Split and Merge Based Story Segmentation in News Videos.....	766
<i>Anuj Goyal, P. Punitha, Frank Hopfgartner, and Joemon M. Jose</i>	
Encoding Ordinal Features into Binary Features for Text Classification	771
<i>Andrea Esuli and Fabrizio Sebastiani</i>	
Topic and Trend Detection in Text Collections Using Latent Dirichlet Allocation	776
<i>Levent Bolelli, Seyda Ertekin, and C. Lee Giles</i>	
Measuring Similarity of Geographic Regions for Geographic Information Retrieval	781
<i>Andreas Henrich and Volker Lüdecke</i>	
Towards the Selection of Induced Syntactic Relations	786
<i>Nicolas Béchet, Mathieu Roche, and Jacques Chauché</i>	

DiffPost: Filtering Non-relevant Content Based on Content Difference
between Two Consecutive Blog Posts 791
Sang-Hyob Nam, Seung-Hoon Na, Yeha Lee, and Jong-Hyeok Lee

An Unsupervised Approach to Product Attribute Extraction 796
Santosh Raju, Prasad Pingali, and Vasudeva Varma

Workshops

Workshop on Contextual Information Access, Seeking and Retrieval
Evaluation 801
*Bich-Liên Doan, Joemon M. Jose, Massimo Melucci, and
Lynda Tamime-Lechani*

Workshop on Information Retrieval over Social Networks 803
Stephane Marchand-Maillet, Arjen P. de Vries, and Mor Naaman

Workshop on Geographic Information on the Internet Workshop
(GIIW) 804
*Gregory Grefenstette, Pierre-Alain Moëllic, Adrian Popescu, and
Florence Sèdes*

Tutorials

Current Developments in Information Retrieval Evaluation 806
Thomas Mandl

Information Extraction and Linking in a Retrieval Context 810
Marie-Francine Moens and Djoerd Hiemstra

Mining Query Logs 814
Salvatore Orlando and Fabrizio Silvestri

Author Index 819