

Table of Contents

Query Representation

Explaining Query Modifications: An Alternative Interpretation of Term Addition and Removal	1
<i>Vera Hollink, Jiyin He, and Arjen de Vries</i>	
Modeling Transactional Queries via Templates	13
<i>Edward Bortnikov, Pinar Donmez, Amit Kagan, and Ronny Lempel</i>	
Exploring Query Patterns in Email Search	25
<i>Morgan Harvey and David Elswailer</i>	
Interactive Search Support for Difficult Web Queries	37
<i>Abdigani Diriye, Giridhar Kumaran, and Jeff Huang</i>	

Blog and Online-Community Search

Predicting the Future Impact of News Events	50
<i>Julien Gaugaz, Patrick Siehndel, Gianluca Demartini, Tereza Iofciu, Mihai Georgescu, and Nicola Henze</i>	
Detection of News Feeds Items Appropriate for Children	63
<i>Tamara Polajnar, Richard Glassey, and Leif Azzopardi</i>	
Comparing Tweets and Tags for URLs	73
<i>Morgan Harvey, Mark Carman, and David Elswailer</i>	
Geo-Location Estimation of Flickr Images: Social Web Based Enrichment	85
<i>Claudia Hawff and Geert-Jan Houben</i>	

Semi-structured Retrieval

A Field Relevance Model for Structured Document Retrieval	97
<i>Jin Young Kim and W. Bruce Croft</i>	
Relation Based Term Weighting Regularization	109
<i>Hao Wu and Hui Fang</i>	
A New Approach to Answerer Recommendation in Community Question Answering Services	121
<i>Zhenlei Yan and Jie Zhou</i>	

On the Modeling of Entities for Ad-Hoc Entity Search in the Web of Data	133
<i>Robert Neumayer, Krisztian Balog, and Kjetil Nørkvåg</i>	
Result Disambiguation in Web People Search	146
<i>Richard Berendsen, Bogomil Kovachev, Evangelia-Paraskevi Nastou, Maarten de Rijke, and Wouter Weerkamp</i>	
Evaluation	
On Smoothing Average Precision	158
<i>Stephen Robertson</i>	
New Metrics for Meaningful Evaluation of Informally Structured Speech Retrieval	170
<i>Maria Eskevich, Walid Magdy, and Gareth J.F. Jones</i>	
On Aggregating Labels from Multiple Crowd Workers to Infer Relevance of Documents	182
<i>Mehdi Hosseini, Ingemar J. Cox, Nataša Milić-Frayling, Gabriella Kazai, and Vishwa Vinay</i>	
Applications	
How Random Walks Can Help Tourism	195
<i>Claudio Lucchese, Raffaele Perego, Fabrizio Silvestri, Hossein Vahabi, and Rossano Venturini</i>	
Retrieving Candidate Plagiarised Documents Using Query Expansion . . .	207
<i>Rao Muhammad Adeel Nawab, Mark Stevenson, and Paul Clough</i>	
Reliability Prediction of Webpages in the Medical Domain	219
<i>Parikshit Sondhi, V.G. Vinod Vydiswaran, and ChengXiang Zhai</i>	
Automatic Foldering of Email Messages: A Combination Approach	232
<i>Tony Tam, Artur Ferreira, and André Lourenço</i>	
Retrieval Models	
A Log-Logistic Model-Based Interpretation of TF Normalization of BM25	244
<i>Yuanhua Lu and ChengXiang Zhai</i>	
Score Transformation in Linear Combination for Multi-criteria Relevance Ranking	256
<i>Shima Gerani, ChengXiang Zhai, and Fabio Crestani</i>	

Axiomatic Analysis of Translation Language Model for Information Retrieval	268
<i>Maryam Karimzadehgan and ChengXiang Zhai</i>	
An Information-Based Cross-Language Information Retrieval Model	281
<i>Bo Li and Eric Gaussier</i>	
Extended Expectation Maximization for Inferring Score Distributions . . .	293
<i>Keshi Dai, Virgil Pavlu, Evangelos Kanoulas, and Javed A. Aslam</i>	
Top-k Retrieval Using Facility Location Analysis	305
<i>Guido Zuccon, Leif Azzopardi, Dell Zhang, and Jun Wang</i>	

Image and Video Retrieval

An Interactive Paper and Digital Pen Interface for Query-by-Sketch Image Retrieval	317
<i>Roman Kreuzer, Michael Springmann, Ihab Al Kabary, and Heiko Schuldt</i>	
Image Abstraction in Crossmedia Retrieval for Text Illustration	329
<i>Filipe Coelho and Cristina Ribeiro</i>	
A Latent Variable Ranking Model for Content-Based Retrieval	340
<i>Ariadna Quattoni, Xavier Carreras, and Antonio Torralba</i>	

Text and Content Classification, Categorisation, Clustering

Language Modelling of Constraints for Text Clustering	352
<i>Javier Parapar and Álvaro Barreiro</i>	
A Framework for Unsupervised Spam Detection in Social Networking Sites	364
<i>Maarten Bosma, Edgar Meij, and Wouter Weerkamp</i>	
Classification of Short Texts by Deploying Topical Annotations	376
<i>Daniele Vitale, Paolo Ferragina, and Ugo Scaiella</i>	
Cluster Labeling for Multilingual Scatter/Gather Using Comparable Corpora	388
<i>Goutham Tholpadi, Mrinal Kanti Das, Chiranjib Bhattacharyya, and Shirish Shevade</i>	

Systems Efficiency

Adaptive Time-to-Live Strategies for Query Result Caching in Web Search Engines	401
<i>Sadiye Alici, Ismail Sengor Altingovde, Rifat Ozcan, B. Barla Cambazoglu, and Özgür Ulusoy</i>	
Intra-query Concurrent Pipelined Processing for Distributed Full-Text Retrieval	413
<i>Simon Jonassen and Svein Erik Bratsberg</i>	

Industry Track

Usefulness of Sentiment Analysis	426
<i>Jussi Karlgren, Magnus Sahlgren, Fredrik Olsson, Fredrik Espinoza, and Ola Hamfors</i>	
Modeling Static Caching in Web Search Engines	436
<i>Ricardo Baeza-Yates and Simon Jonassen</i>	

Posters

Integrating Interactive Visualizations in the Search Process of Digital Libraries and IR Systems	447
<i>Daniel Hienert, Frank Sawitzki, Philipp Schaer, and Philipp Mayr</i>	
On Theoretically Valid Score Distributions in Information Retrieval	451
<i>Ronan Cummins and Colm O’Riordan</i>	
Adaptive Temporal Query Modeling	455
<i>Maria-Hendrike Peetz, Edgar Meij, Maarten de Rijke, and Wouter Weerkamp</i>	
The Design of a Visual History Tool to Help Users Refind Information within a Website	459
<i>Trien V. Do and Roy A. Ruddle</i>	
Analyzing the Polarity of Opinionated Queries	463
<i>Sergiu Chelaru, Ismail Sengor Altingovde, and Stefan Siersdorfer</i>	
Semi-automatic Document Classification: Exploiting Document Difficulty	468
<i>Miguel Martinez-Alvarez, Sirvan Yahyaei, and Thomas Roelleke</i>	
Investigating Summarization Techniques for Geo-Tagged Image Indexing	472
<i>Ahmet Aker, Xin Fan, Mark Sanderson, and Robert Gaizauskas</i>	

Handling OOV Words in Indian-language – English CLIR	476
<i>Parin Chheda, Manaal Faruqui, and Pabitra Mitra</i>	
Using a Medical Thesaurus to Predict Query Difficulty	480
<i>Florian Boudin, Jian-Yun Nie, and Martin Dawes</i>	
Studying a Personality Coreference Network in a News Stories Photo Collection	485
<i>José Devezas, Filipe Coelho, Sérgio Nunes, and Cristina Ribeiro</i>	
Phrase Pair Classification for Identifying Subtopics	489
<i>Sujatha Das, Prasenjit Mitra, and C. Lee Giles</i>	
Full and Mini-batch Clustering of News Articles with Star-EM	494
<i>Matthias Gallé and Jean-Michel Renders</i>	
Assessing and Predicting Vertical Intent for Web Queries	499
<i>Ke Zhou, Ronan Cummins, Martin Halvey, Mounia Lalmas, and Joemon M. Jose</i>	
Predicting IMDB Movie Ratings Using Social Media	503
<i>Andrei Oghina, Mathias Breuss, Manos Tsagkias, and Maarten de Rijke</i>	
Squeezing the Ensemble Pruning: Faster and More Accurate Categorization for News Portals	508
<i>Cagri Toraman and Fazli Can</i>	
A General Framework for People Retrieval in Social Media with Multiple Roles	512
<i>Amin Mantrach and Jean-Michel Renders</i>	
Analysis of Query Reformulations in a Search Engine of a Local Web Site	517
<i>M-Dyaa Albakour, Udo Kruschwitz, Nikolaos Nanas, Ibrahim Adeyanju, Dawei Song, Maria Fasli, and Anne De Roeck</i>	
Temporal Pseudo-relevance Feedback in Microblog Retrieval	522
<i>Stewart Whiting, Iraklis A. Klampanos, and Joemon M. Jose</i>	
Learning Adaptive Domain Models from Click Data to Bootstrap Interactive Web Search	527
<i>Deirdre Lungley, Udo Kruschwitz, and Dawei Song</i>	
A Little Interaction Can Go a Long Way: Enriching the Query Formulation Process	531
<i>Abdiganı Diriye, Anastasios Tombros, and Ann Blandford</i>	

Learning to Rank from Relevance Feedback for e-Discovery	535
<i>Peter Lubell-Doughtie and Katja Hofmann</i>	
When Simple is (more than) Good Enough: Effective Semantic Search with (almost) no Semantics	540
<i>Robert Neumayer, Krisztian Balog, and Kjetil Nørkvåg</i>	
Evaluating Personal Information Retrieval	544
<i>Liadh Kelly, Paul Bunbury, and Gareth J.F. Jones</i>	
Applying Power Graph Analysis to Weighted Graphs	548
<i>Niels Bloom</i>	
An Investigation of Term Weighting Approaches for Microblog Retrieval	552
<i>Paul Ferguson, Neil O'Hare, James Lanagan, Owen Phelan, and Kevin McCarthy</i>	
On the Size of Full Element-Indexes for XML Keyword Search	556
<i>Duygu Atilgan, Ismail Sengor Altingovde, and Özgür Ulusoy</i>	
Combining Probabilistic Language Models for Aspect-Based Sentiment Retrieval	561
<i>Lisette García-Moya, Henry Anaya-Sánchez, and Rafael Berlanga-Llavori</i>	
In Praise of Laziness: A Lazy Strategy for Web Information Extraction	565
<i>Rifat Ozcan, Ismail Sengor Altingovde, and Özgür Ulusoy</i>	

Demos

LiveTweet: Monitoring and Predicting Interesting Microblog Posts	569
<i>Arifah Che Alhadi, Thomas Gottron, Jérôme Kunegis, and Nasir Naveed</i>	
A User Interface for Query-by-Sketch Based Image Retrieval with Color Sketches	571
<i>Ivan Giangreco, Michael Springmann, Ihab Al Kabary, and Heiko Schuld</i>	
Crisees: Real-Time Monitoring of Social Media Streams to Support Crisis Management	573
<i>David Maxwell, Stefan Raue, Leif Azzopardi, Chris Johnson, and Sarah Oates</i>	
A Mailbox Search Engine Using Query Multi-modal Expansion and Community-Based Smoothing	576
<i>Amin Mantrach and Jean-Michel Renders</i>	

EmSe: Supporting Children's Information Needs within a Hospital Environment	578
<i>Leif Azzopardi, Doug Dowie, Sergio Duarte, Carsten Eickhoff, Richard Glassey, Karl Gyllstrom, Djoerd Hiemstra, Franciska de Jong, Frea Kruisinga, Kelly Marshall, Sien Moens, Tamara Polajnar, Frans van der Sluis, and Arjen de Vries</i>	
Retro: Time-Based Exploration of Product Reviews	581
<i>Jannik Strötgen, Omar Alonso, and Michael Gertz</i>	
Querium: A Session-Based Collaborative Search System	583
<i>Abdigani Diriye and Gene Golovchinsky</i>	
Author Index	585