

Joemon M. Jose · Claudia Hauff
Ismail Sengor Altingovde · Dawei Song
Dyaa Albakour · Stuart Watt
John Tait (Eds.)

Advances in Information Retrieval

39th European Conference on IR Research, ECIR 2017
Aberdeen, UK, April 8–13, 2017
Proceedings

Contents

Entity Linking to One Thousand Knowledge Bases	1
<i>Ning Gao and Silviu Cucerzan</i>	
Where Do All These Search Terms Come From? – Two Experiments in Domain-Specific Search	15
<i>Daniel Hienert and Maria Lusky</i>	
Using Query Performance Predictors to Reduce Spoken Queries	27
<i>Jaime Arguello, Sandeep Avula, and Fernando Diaz</i>	
Entity Linking in Queries: Efficiency vs. Effectiveness	40
<i>Faegheh Hasibi, Krisztian Balog, and Svein Erik Bratsberg</i>	
Cross-Lingual Sentiment Relation Capturing for Cross-Lingual Sentiment Analysis	54
<i>Qiang Chen, Wenjie Li, Yu Lei, Xule Liu, Chuwei Luo, and Yanxiang He</i>	
Hierarchical Re-estimation of Topic Models for Measuring Topical Diversity	68
<i>Hosein Azarbonyad, Mostafa Dehghani, Tom Kenter, Maarten Marx, Jaap Kamps, and Maarten de Rijke</i>	
Collective Entity Linking in Tweets Over Space and Time	82
<i>Wen-Haw Chong, Ee-Peng Lim, and William Cohen</i>	
Simple Personalized Search Based on Long-Term Behavioral Signals	95
<i>Anna Sepliarskaia, Filip Radlinski, and Maarten de Rijke</i>	
The Effects of Search Task Determinability on Search Behavior	108
<i>Rob Capra, Jaime Arguello, and Yinglong Zhang</i>	
Inferring User Interests for Passive Users on Twitter by Leveraging Follower Biographies	122
<i>Guangyuan Piao and John G. Breslin</i>	
Fusion of Bag-of-Words Models for Image Classification in the Medical Domain	134
<i>Leonidas Valavanis, Spyridon Stathopoulos, and Theodore Kalambovikis</i>	
Do Topic Shift and Query Reformulation Patterns Correlate in Academic Search?	146
<i>Xinyi Li and Maarten de Rijke</i>	

Learning to Re-rank Medical Images Using a Bayesian Network-Based Thesaurus	160
<i>Hajer Ayadi, Mouna Torjmen Khemakhem, Jimmy Xiangji Huang, Mariam Daoud, and Maher Ben Jemaa</i>	
A Part-of-Speech Enhanced Neural Conversation Model	173
<i>Chuwei Luo, Wenjie Li, Qiang Chen, and Yanxiang He</i>	
“Are Machines Better Than Humans in Image Tagging?” - A User Study Adds to the Puzzle	186
<i>Ralph Ewerth, Matthias Springstein, Lo An Phan-Vogtmann, and Juliane Schütze</i>	
Utterance Retrieval Based on Recurrent Surface Text Patterns	199
<i>Guillaume Dubuisson Duplessis, Franck Charras, Vincent Letard, Anne-Laure Ligozat, and Sophie Rosset</i>	
How Do Order and Proximity Impact the Readability of Event Summaries? . . .	212
<i>Arunav Mishra and Klaus Berberich</i>	
Sentiment Propagation for Predicting Reputation Polarity	226
<i>Anastasia Giachanou, Julio Gonzalo, Ida Mele, and Fabio Crestani</i>	
Transitivity, Time Consumption, and Quality of Preference Judgments in Crowdsourcing	239
<i>Kai Hui and Klaus Berberich</i>	
Exploring Time-Sensitive Variational Bayesian Inference LDA for Social Media Data	252
<i>Anjie Fang, Craig Macdonald, Iadh Ounis, Philip Habel, and Xiao Yang</i>	
E-Government and the Digital Divide: A Study of English-as-a-Second-Language Users’ Information Behaviour	266
<i>David Brazier and Morgan Harvey</i>	
A Task Completion Engine to Enhance Search Session Support for Air Traffic Work Tasks	278
<i>Yashar Moshfeghi, Raoul Rothfeld, Leif Azzopardi, and Peter Triantafillou</i>	
Personalized Keyword Boosting for Venue Suggestion Based on Multiple LBSNs	291
<i>Mohammad Aliannejadi, Dimitrios Rafailidis, and Fabio Crestani</i>	
Feature-Oriented Analysis of User Profile Completion Problem.	304
<i>Morteza Haghir Chehrehgani</i>	

Generating Descriptions of Entity Relationships	317
<i>Nikos Voskarides, Edgar Meij, and Maarten de Rijke</i>	
Language Influences on Tweeter Geolocation	331
<i>Ahmed Mourad, Falk Scholer, and Mark Sanderson</i>	
Human-Based Query Difficulty Prediction	343
<i>Adrian-Gabriel Chifu, Sébastien Déjean, Stefano Mizzaro, and Josiane Mothe</i>	
Fixed-Cost Pooling Strategies Based on IR Evaluation Measures.	357
<i>Aldo Lipani, Joao Palotti, Mihai Lupu, Florina Piroi, Guido Zuccon, and Allan Hanbury</i>	
A Formal and Empirical Study of Unsupervised Signal Combination for Textual Similarity Tasks	369
<i>Enrique Amigó, Fernando Giner, Julio Gonzalo, and Felisa Verdejo</i>	
Paper2vec: Combining Graph and Text Information for Scientific Paper Representation	383
<i>Soumyajit Ganguly and Vikram Pudi</i>	
Exploration of a Threshold for Similarity Based on Uncertainty in Word Embedding	396
<i>Navid Rekabsaz, Mihai Lupu, and Allan Hanbury</i>	
Ancient Roman Coin Retrieval: A Systematic Examination of the Effects of Coin Grade	410
<i>Callum Fare and Ognjen Arandjelović</i>	
A Systematic Analysis of Sentence Update Detection for Temporal Summarization	424
<i>Cristina Gârbacea and Evangelos Kanoulas</i>	
A Multiple-Instance Learning Approach to Sentence Selection for Question Ranking	437
<i>Salvatore Romeo, Giovanni Da San Martino, Alberto Barrón-Cedeño, and Alessandro Moschitti</i>	
Enhancing Sensitivity Classification with Semantic Features Using Word Embeddings	450
<i>Graham McDonald, Craig Macdonald, and Iadh Ounis</i>	
Predicting Users' Future Interests on Twitter	464
<i>Fattane Zarrinkalam, Hossein Fani, Ebrahim Bagheri, and Mohsen Kahani</i>	

A New Scheme for Scoring Phrases in Unsupervised Keyphrase Extraction . . .	477
<i>Corina Florescu and Cornelia Caragea</i>	
JustEvents: A Crowdsourced Corpus for Event Validation with Strict Temporal Constraints.	484
<i>Andrea Ceroni, Ujwal Gadiraju, and Marco Fisichella</i>	
Dimension Projection Among Languages Based on Pseudo-Relevant Documents for Query Translation	493
<i>Javid Dadashkarimi, Mahsa S. Shahshahani, Amirhossein Tebbifakhr, Heshaam Faili, and Azadeh Shakery</i>	
Labeling Topics with Images Using a Neural Network.	500
<i>Nikolaos Aletras and Arpit Mittal</i>	
Leveraging Site Search Logs to Identify Missing Content on Enterprise Webpages.	506
<i>Harsh Jhamtani, Rishiraj Saha Roy, Niyati Chhaya, and Eric Nyberg</i>	
LTRo: Learning to Route Queries in Clustered P2P IR	513
<i>Rami S. Alkhalwaldeh, Deepak P., Joemon M. Jose, and Fajie Yuan</i>	
Faster K-Means Cluster Estimation	520
<i>Siddhesh Khandelwal and Amit Awekar</i>	
Predicting Emotional Reaction in Social Networks	527
<i>J�r�mie Clos, Anil Bandhakavi, Nirmalie Wiratunga, and Guillaume Cabanac</i>	
Irony Detection with Attentive Recurrent Neural Networks.	534
<i>Yu-Hsiang Huang, Hen-Hsen Huang, and Hsin-Hsi Chen</i>	
We Used Neural Networks to Detect Clickbaits: You Won't Believe What Happened Next!	541
<i>Ankesh Anand, Tanmoy Chakraborty, and Noseong Park</i>	
Learning to Classify Inappropriate Query-Completions.	548
<i>Parth Gupta and Jose Santos</i>	
Counteracting Novelty Decay in First Story Detection	555
<i>Yumeng Qin, Dominik Wurzer, Victor Lavrenko, and Cunchen Tang</i>	
Predicting Genre Preferences from Cultural and Socio-Economic Factors for Music Retrieval	561
<i>Marcin Skowron, Florian Lemmerich, Bruce Ferwerda, and Markus Schedl</i>	

Batch Incremental Shared Nearest Neighbor Density Based Clustering Algorithm for Dynamic Datasets	568
<i>Panthadeep Bhattacharjee and Amit Awekar</i>	
Improving Tweet Representations Using Temporal and User Context	575
<i>J. Ganesh, Manish Gupta, and Vasudeva Varma</i>	
Personalized Parsimonious Language Models for User Modeling in Social Bookmaking Systems	582
<i>Nawal Ould Amer, Philippe Mulhem, and Mathias Géry</i>	
A Novel Word Embedding Based Stemming Approach for Microblog Retrieval During Disasters	589
<i>Moumita Basu, Anurag Roy, Kripabandhu Ghosh, Somprakash Bandyopadhyay, and Saptarshi Ghosh</i>	
Search Personalization with Embeddings	598
<i>Thanh Vu, Dat Quoc Nguyen, Mark Johnson, Dawei Song, and Alistair Willis</i>	
Do Easy Topics Predict Effectiveness Better Than Difficult Topics?	605
<i>Kevin Roitero, Eddy Maddalena, and Stefano Mizzaro</i>	
Named Entity Classification Using Search Engine’s Query Suggestions	612
<i>Jayendra Barua and Dhaval Patel</i>	
“When Was This Picture Taken?” – Image Date Estimation in the Wild	619
<i>Eric Müller, Matthias Springstein, and Ralph Ewerth</i>	
Low-Cost Preference Judgment via Ties.	626
<i>Kai Hui and Klaus Berberich</i>	
Using Section Headings to Compute Cross-Lingual Similarity of Wikipedia Articles.	633
<i>Monica Lestari Paramita, Paul Clough, and Robert Gaizauskas</i>	
Learning to Rank for Consumer Health Search: A Semantic Approach	640
<i>Luca Soldaini and Nazli Goharian</i>	
Matrix Factorisation with Word Embeddings for Rating Prediction on Location-Based Social Networks.	647
<i>Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis</i>	
Word Similarity Based Model for Tweet Stream Prospective Notification.	655
<i>Abdelhamid Chellal, Mohand Boughanem, and Bernard Dousset</i>	

Negative Feedback in the Language Modeling Framework for Text Recommendation	662
<i>Hossein Rahmatizadeh Zagheli, Mozhdeh Ariannezhad, and Azadeh Shakery</i>	
Plagiarism Detection in Texts Obfuscated with Homoglyphs.	669
<i>Faisal Alvi, Mark Stevenson, and Paul Clough</i>	
Iterative Estimation of Document Relevance Score for Pseudo-Relevance Feedback	676
<i>Mozhdeh Ariannezhad, Ali MontazerAlghaem, Hamed Zamani, and Azadeh Shakery</i>	
Design Patterns for Fusion-Based Object Retrieval	684
<i>Shuo Zhang and Krisztian Balog</i>	
On the Long-Tail Entities in News	691
<i>José Esquivel, Dyaa Albakour, Miguel Martinez, David Corney, and Samir Moussa</i>	
Search Costs vs. User Satisfaction on Mobile	698
<i>Manisha Verma and Emine Yilmaz</i>	
On the Efficiency of Selective Search	705
<i>Fatih Hafizoglu, Emre Can Kucukoglu, and Ismail Sengor Altinogvde</i>	
Does Online Evaluation Correspond to Offline Evaluation in Query Auto Completion?	713
<i>Alexandros Bampoulidis, João Palotti, Mihai Lupu, Jon Brassey, and Allan Hanbury</i>	
A Neural Attention Model for Categorizing Patient Safety Events	720
<i>Arman Cohan, Allan Fong, Nazli Goharian, and Raj Ratwani</i>	
Promoting Understandability in Consumer Health Information Search	727
<i>Hua Yang and Teresa Goncalves</i>	
A Social Framework for Set Recommendation in Group Recommender Systems	735
<i>Lorena Recalde</i>	
Pyndri: A Python Interface to the Indri Search Engine	744
<i>Christophe Van Gysel, Evangelos Kanoulas, and Maarten de Rijke</i>	
“Hey, vitivr!” – A Multimodal UI for Video Retrieval	749
<i>Prateek Goel, Ivan Giangreco, Luca Rossetto, Claudiu Tănase, and Heiko Schuldt</i>	

FairScholar: Balancing Relevance and Diversity for Scientific Paper Recommendation	753
<i>Ankesh Anand, Tanmoy Chakraborty, and Amitava Das</i>	
The SENSEI Overview of Newspaper Readers' Comments	758
<i>Adam Funk, Ahmet Aker, Emma Barker, Monica Lestari Paramita, Mark Hepple, and Robert Gaizauskas</i>	
Temporal Semantic Analysis of Conference Proceedings	762
<i>Fedelucio Narducci, Pierpaolo Basile, Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro</i>	
QweetFinder: Real-Time Finding and Filtering of Question Tweets	766
<i>Ameer Albahem, Maram Hasanain, Marwan Torki, and Tamer Elsayed</i>	
Integration of the Scientific Recommender System Mr. DLib into the Reference Manager JabRef	770
<i>Stefan Feyer, Sophie Siebert, Bela Gipp, Akiko Aizawa, and Joeran Beel</i>	
Abstracts of Doctoral Consortium Papers and Workshops	
A Social Framework for Set Recommendation in Group Recommender Systems	777
<i>Lorena Recalde</i>	
User Behavior Analysis and User Modeling for Complex Search	778
<i>Jiaxin Mao</i>	
First International Workshop on Exploitation of Social Media for Emergency Relief and Preparedness (SMERP).	779
<i>Saptarshi Ghosh, Kripabandhu Ghosh, Tanmoy Chakraborty, Debasis Ganguly, Gareth Jones, and Marie-Francine Moens</i>	
Bibliometric-Enhanced Information Retrieval: 5th International BIR Workshop	784
<i>Philipp Mayr, Ingo Frommholz, and Guillaume Cabanac</i>	
OnST'17: The 2nd International Workshop on <u>Online Safety Trust</u> Fraud Prevention.	790
<i>Marco Fisichella</i>	
Workshop on Social Media for Personalization and Search (SoMePeAS)	795
<i>Ludovico Boratto, Andreas Kaltenbrunner, and Giovanni Stilo</i>	
Author Index	797