

Gareth J.F. Jones · Séamus Lawless
Julio Gonzalo · Liadh Kelly
Lorraine Goeriot · Thomas Mandl
Linda Cappellato · Nicola Ferro (Eds.)

Experimental IR Meets Multilinguality, Multimodality, and Interaction

8th International Conference
of the CLEF Association, CLEF 2017
Dublin, Ireland, September 11–14, 2017
Proceedings

Contents

Full Papers

A Pinch of Humor for Short-Text Conversation: An Information Retrieval Approach	3
<i>Vladislav Blinov, Kirill Mishchenko, Valeria Bolotova, and Pavel Braslavski</i>	
A Component-Level Analysis of an Academic Search Test Collection. Part I: System and Collection Configurations	16
<i>Florian Dietz and Vivien Petras</i>	
A Component-Level Analysis of an Academic Search Test Collection. Part II: Query Analysis	29
<i>Florian Dietz and Vivien Petras</i>	
Improving the Reliability of Query Expansion for User-Generated Speech Retrieval Using Query Performance Prediction	43
<i>Ahmad Khwileh, Andy Way, and Gareth J.F. Jones</i>	
Optimized Convolutional Neural Network Ensembles for Medical Subfigure Classification	57
<i>Sven Koitka and Christoph M. Friedrich</i>	
IRIT-QFR: IRIT Query Feature Resource.	69
<i>Serge Molina, Josiane Mothe, Dorian Roques, Ludovic Tanguy, and Md Zia Ullah</i>	
Evaluating and Improving the Extraction of Mathematical Identifier Definitions	82
<i>Moritz Schubotz, Leonard Krämer, Norman Meuschke, Felix Hamborg, and Bela Gipp</i>	

Short Papers

Query Expansion for Sentence Retrieval Using Pseudo Relevance Feedback and Word Embedding	97
<i>Piyush Arora, Jennifer Foster, and Gareth J.F. Jones</i>	
<i>WebShodh</i> : A Code Mixed Factoid Question Answering System for Web . . .	104
<i>Khyathi Raghavi Chandu, Manoj Chinnakotla, Alan W. Black, and Manish Shrivastava</i>	

Extracting Graph Topological Information and Users' Opinion	112
<i>Mirko Lai, Marcella Tambuscio, Viviana Patti, Giancarlo Ruffo, and Paolo Rosso</i>	
Estimating the Similarities Between Texts of Right-Handed and Left-Handed Males and Females	119
<i>Tatiana Litvinova, Pavel Seredin, Olga Litvinova, and Ekaterina Ryzhkova</i>	
Evaluation of Hierarchical Clustering via Markov Decision Processes for Efficient Navigation and Search.	125
<i>Raul Moreno, Wěipéng Huáng, Arjumand Younus, Michael O'Mahony, and Neil J. Hurley</i>	
Plausibility Testing for Lexical Resources	132
<i>Magdalena Parks, Jussi Karlgren, and Sara Stymne</i>	
An Improved <i>Impostors</i> Method for Authorship Verification	138
<i>Nektaria Potha and Efstathios Stamatatos</i>	
Comparison of Character <i>n</i> -grams and Lexical Features on Author, Gender, and Language Variety Identification on the Same Spanish News Corpus	145
<i>Miguel A. Sanchez-Perez, Ilia Markov, Helena Gómez-Adorno, and Grigori Sidorov</i>	
Enriching Existing Test Collections with OXPath	152
<i>Philipp Schaer and Mandy Neumann</i>	
Best of the Labs	
A Highly Available Real-Time News Recommender Based on Apache Spark	161
<i>Jaschar Domann and Andreas Lommatzsch</i>	
The Case for Being Average: A Mediocrity Approach to Style Masking and Author Obfuscation (Best of the Labs Track at CLEF-2017).	173
<i>Georgi Karadzhov, Tsvetomila Mihaylova, Yasen Kiprov, Georgi Georgiev, Ivan Koychev, and Preslav Nakov</i>	
Author Clustering with an Adaptive Threshold	186
<i>Mirco Kocher and Jacques Savoy</i>	
Segmenting Compound Biomedical Figures into Their Constituent Panels . . .	199
<i>Pengyuan Li, Xiangying Jiang, Chandra Kambhamettu, and Hagit Shatkay</i>	

An Analysis of Cross-Genre and In-Genre Performance for Author Profiling in Social Media	211
<i>Maria Medvedeva, Hessel Haagsma, and Malvina Nissim</i>	
TimeLine Illustration Based on Microblogs: When Diversification Meets Metadata Re-ranking	224
<i>Philippe Mulhem, Lorraine Goeuriot, Nayanika Dogra, and Nawal Ould Amer</i>	
Labs Overviews	
CLEF 2017 NewsREEL Overview: A Stream-Based Recommender Task for Evaluation and Education	239
<i>Andreas Lommatzsch, Benjamin Kille, Frank Hopfgartner, Martha Larson, Torben Brodt, Jonas Seiler, and Özlem Özgöbek</i>	
LifeCLEF 2017 Lab Overview: Multimedia Species Identification Challenges	255
<i>Alexis Joly, Hervé Goëau, Hervé Glotin, Concetto Spampinato, Pierre Bonnet, Willem-Pier Vellinga, Jean-Christophe Lombardo, Robert Planqué, Simone Palazzo, and Henning Müller</i>	
Overview of PAN'17: Author Identification, Author Profiling, and Author Obfuscation	275
<i>Martin Potthast, Francisco Rangel, Michael Tschuggnall, Efstathios Stamatatos, Paolo Rosso, and Benno Stein</i>	
CLEF 2017 eHealth Evaluation Lab Overview	291
<i>Lorraine Goeuriot, Liadh Kelly, Hanna Suominen, Aurélie Névéol, Aude Robert, Evangelos Kanoulas, Rene Spijker, João Palotti, and Guido Zuccon</i>	
CLEF 2017 Microblog Cultural Contextualization Lab Overview	304
<i>Liana Ermakova, Lorraine Goeuriot, Josiane Mothe, Philippe Mulhem, Jian-Yun Nie, and Eric SanJuan</i>	
Overview of ImageCLEF 2017: Information Extraction from Images	315
<i>Bogdan Ionescu, Henning Müller, Mauricio Villegas, Helbert Arenas, Giulia Boato, Duc-Tien Dang-Nguyen, Yashin Dicente Cid, Carsten Eickhoff, Alba G. Seco de Herrera, Cathal Gurrin, Bayzidul Islam, Vassili Kovalev, Vitali Liauchuk, Josiane Mothe, Luca Piras, Michael Riegler, and Immanuel Schwall</i>	
Overview of the CLEF 2017 Personalised Information Retrieval Pilot Lab (PIR-CLEF 2017)	338
<i>Gabriella Pasi, Gareth J.F. Jones, Stefania Marrara, Camilla Sanvitto, Debasis Ganguly, and Procheta Sen</i>	

eRISK 2017: CLEF Lab on Early Risk Prediction on the Internet:
Experimental Foundations 346
David E. Losada, Fabio Crestani, and Javier Parapar

CLEF 2017 Dynamic Search Evaluation Lab Overview 361
Evangelos Kanoulas and Leif Azzopardi

CLEF 2017: Multimodal Spatial Role Labeling (mSpRL) Task Overview ... 367
*Parisa Kordjamshidi, Taher Rahgooy, Marie-Francine Moens,
James Pustejovsky, Umar Manzoor, and Kirk Roberts*

Author Index 377