

1. Record Nr.	UNISA996466168003316
Titolo	Experimental IR Meets Multilinguality, Multimodality, and Interaction [[electronic resource]] : 8th International Conference of the CLEF Association, CLEF 2017, Dublin, Ireland, September 11–14, 2017, Proceedings / / edited by Gareth J.F. Jones, Séamus Lawless, Julio Gonzalo, Liadh Kelly, Lorraine Goeuriot, Thomas Mandl, Linda Cappellato, Nicola Ferro
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2017
ISBN	3-319-65813-1
Edizione	[1st ed. 2017.]
Descrizione fisica	1 online resource (XVIII, 378 p. 59 illus.)
Collana	Information Systems and Applications, incl. Internet/Web, and HCI ; ; 10456
Disciplina	025.04
Soggetti	Natural language processing (Computer science) Information storage and retrieval systems Data mining Application software Artificial intelligence Natural Language Processing (NLP) Information Storage and Retrieval Data Mining and Knowledge Discovery Computer and Information Systems Applications Artificial Intelligence
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- CLEF Steering Committee -- Contents -- Full Papers -- A Pinch of Humor for Short-Text Conversation: An Information Retrieval Approach -- 1 Introduction -- 2 Related Work -- 3 Data -- 3.1 Joke Collection -- 3.2 Yahoo!Answers -- 4 Retrieval Models -- 4.1 BM25 -- 4.2 Query Term Reweighting Model (QTR) -- 4.3 doc2vec -- 5 Evaluation -- 5.1 Yahoo!Answers -- 5.2 Lab Evaluation -- 6 Conclusion and Future Work -- References -- A Component-Level Analysis of an Academic Search Test Collection. -- 1

Introduction -- 2 Component-Level Evaluation in Academic Search -- 3
 Study Configurations -- 3.1 Test Collection and Test IR System -- 3.2
 Component-Level Configurations -- 3.3 Analysis Steps -- 4 Analyzing
 Component Performance -- 4.1 Document and Topic Structure -- 4.2
 IR System Components -- 5 Conclusion -- References -- A
 Component-Level Analysis of an Academic Search Test Collection. -- 1
 Introduction -- 2 Query Performance Prediction -- 3 Study
 Configurations -- 3.1 Test Collection and Test IR System -- 3.2
 Analysis Steps -- 4 Analyzing Query Performance -- 4.1 Recall Base --
 4.2 Topic Specificity -- 4.3 Semantic Heterogeneity of Topic Terms --
 5 Conclusion -- References -- Improving the Reliability of Query
 Expansion for User-Generated Speech Retrieval Using Query
 Performance Prediction -- 1 Introduction -- 2 Previous Work on QE for
 Spoken Content -- 3 UGS Test Collection -- 3.1 UGS Document Set --
 3.2 Query and Relevance Sets -- 4 Initial Investigation of QE for UGS
 Retrieval -- 4.1 Retrieval Model -- 4.2 Query Expansion -- 4.3
 Passage-Based QE -- 4.4 Experimental Results -- 5 Query Performance
 Prediction for QE -- 5.1 Query Performance Prediction Methods -- 5.2
 Evaluation -- 6 Adaptive Passage-Based QE for UGS Retrieval -- 7
 Conclusions and Further Studies -- References.
 Optimized Convolutional Neural Network Ensembles for Medical
 Subfigure Classification -- 1 Introduction -- 2 Revisiting Participation
 at ImageCLEF 2016 -- 3 Dataset Analysis -- 3.1 Data Generation
 Process -- 3.2 Problems of the Dataset -- 3.3 Suggestions for Future
 Datasets -- 4 Methodology -- 5 Evaluation -- 5.1 Results Without
 Additional Images from ImageCLEF 2013 -- 5.2 Self-training Using a
 PubMedCentral Database Snapshot -- 6 Conclusion -- References --
 IRIT-QFR: IRIT Query Feature Resource -- 1 Introduction -- 2 Query
 Features -- 2.1 Document Collection Independent Pre-retrieval
 Features: WordNet-Based and Other Linguistic Features -- 2.2
 Document Collection Dependent Pre-retrieval Features -- 2.3 Post-
 retrieval Features -- 2.4 Collections for Which the Features Have Been
 Estimated -- 3 Analysis of the Resource -- 3.1 Descriptive Analysis --
 4 Applications -- 5 Conclusion -- References -- Evaluating and
 Improving the Extraction of Mathematical Identifier Definitions -- 1
 Introduction -- 2 Related Work -- 2.1 Statistical Feature Analysis (ST)
 -- 2.2 Pattern Matching (PM) -- 2.3 Machine Learning -- 3
 Methodology -- 3.1 Evaluation Framework -- 3.2 Evaluating Existing
 Approaches -- 3.3 New Machine Learning Approach (ML) -- 3.4
 Evaluating the Influence of Namespaces -- 4 Results -- 5 Conclusion
 and Outlook -- References -- Short Papers -- Query Expansion for
 Sentence Retrieval Using Pseudo Relevance Feedback and Word
 Embedding -- 1 Introduction -- 2 Related Work -- 3 Experimental
 Methodology -- 4 Experimental Setup -- 5 Results and Analysis -- 6
 Conclusions and Further Investigation -- References -- WebShodh: A
 Code Mixed Factoid Question Answering System for Web -- 1
 Introduction -- 2 Related Work -- 3 Web Based Code-Mixed QA System
 -- 4 Evaluation Dataset and Results -- 5 Discussion and Conclusion --
 References.
 Extracting Graph Topological Information and Users' Opinion -- 1
 Introduction -- 2 Dataset -- 3 Content and Network Analysis -- 4
 Discussion -- References -- Estimating the Similarities Between Texts
 of Right-Handed and Left-Handed Males and Females -- Abstract -- 1
 Problem Statement -- 2 Materials and Methods -- 3 Results and
 Discussion -- 4 Conclusion and Future Work -- Acknowledgments --
 References -- Evaluation of Hierarchical Clustering via Markov Decision
 Processes for Efficient Navigation and Search -- 1 Introduction -- 2
 Markov Decision Processes -- 2.1 Navigating a Hierarchy as an MDP --

3 Choosing a Navigation Policy -- 3.1 Policy: Greedy Descent to Threshold Depth -- 4 Toy Example and Conclusion -- References -- Plausibility Testing for Lexical Resources -- 1 Evaluating Information System Components -- 2 Requirements for a Habitable Evaluation Scheme -- 3 Plausible Utterances -- 4 Method of Fake Sentence Generation -- 5 Swapping Strategies -- 5.1 Word Coconut -- 5.2 Sentence Coconut -- 6 Evaluating the Evaluation -- 7 Conclusion -- References -- An Improved Impostors Method for Authorship Verification -- 1 Introduction -- 2 Previous Work -- 3 The Proposed Method -- 4 Experiments -- 4.1 Setup -- 4.2 Results -- 5 Conclusion -- References -- Comparison of Character n-grams and Lexical Features on Author, Gender, and Language Variety Identification on the Same Spanish News Corpus -- 1 Introduction -- 2 Related Work -- 3 Corpus -- 4 Experimental Settings and Results -- 5 Conclusions -- References -- Enriching Existing Test Collections with OXPath -- 1 Introduction -- 2 Related Work -- 3 Materials and Methods -- 4 Results -- 5 Discussion and Conclusion -- References -- Best of the Labs -- A Highly Available Real-Time News Recommender Based on Apache Spark -- 1 Motivation -- 2 Problem Description -- 3 Related Work -- 4 Approach.

4.1 Recommender Models -- 4.2 Implementation -- 4.3 Discussion -- 5 Evaluation -- 5.1 Scalability Analysis in the Offline Evaluation Lab -- 5.2 Click-Through Rate Analysis in the Offline Evaluation Lab -- 5.3 Response Rate Analysis in the Living Lab Evaluation -- 5.4 Discussion -- 6 Conclusion -- References -- The Case for Being Average: A Mediocrity Approach to Style Masking and Author Obfuscation -- 1 Introduction -- 2 Related Work -- 3 Our Mediocrity Approach to Style Masking and Author Obfuscation -- 3.1 Calculating Text Metrics -- 3.2 Modulizing the Text -- 3.3 Text Transformations -- 3.4 Noise -- 3.5 General Transformations -- 3.6 Other Methods -- 3.7 Transformation Magnitude -- 4 Evaluation -- 5 Conclusion and Future Work -- References -- Author Clustering with an Adaptive Threshold -- Abstract -- 1 Introduction -- 2 Related Work -- 3 Test Collections and Evaluation Methodology -- 4 Simple Clustering Algorithm -- 5 Evaluation -- 6 Sensibility Assessment -- 7 Adaptive Thresholding -- 8 Conclusion -- Acknowledgments -- References -- Segmenting Compound Biomedical Figures into Their Constituent Panels -- 1 Introduction -- 2 Methods -- 3 Experiments and Results -- 3.1 Experiments -- 3.2 Datasets and Evaluation -- 3.3 Results -- 4 Conclusion -- References -- An Analysis of Cross-Genre and In-Genre Performance for Author Profiling in Social Media -- 1 Introduction and Background -- 2 GronUP -- 2.1 Preprocessing -- 2.2 Classifier -- 2.3 Features -- 3 Experiments -- 3.1 Datasets -- 3.2 Experimental Settings -- 4 Results -- 5 Analysis -- 6 Conclusions -- References -- TimeLine Illustration Based on Microblogs: When Diversification Meets Metadata Re-ranking -- 1 Introduction -- 2 Description of the Task -- 3 Related Work -- 4 Description of the Approach -- 4.1 Pre-processing of the Official Tweet Corpus -- 4.2 Content-Based Matching.

4.3 Diversification -- 4.4 Re-ranking -- 5 Experimental Results -- 5.1 Dataset -- 5.2 Parameters Settings -- 5.3 Runs Submitted -- 6 Evaluation Results -- 6.1 Classical IR Metrics -- 6.2 Informativeness -- 7 Discussion -- 7.1 Evaluation of Tweet Retrieval -- 7.2 Building a Tweet Retrieval Benchmark -- 8 Conclusion -- References -- Labs Overviews -- CLEF 2017 NewsREEL Overview: A Stream-Based Recommender Task for Evaluation and Education -- 1 Introduction -- 2 Related Work -- 2.1 Evaluation of News Recommendation Systems -- 2.2 Education -- 3 News Recommendation Scenario -- 3.1 NewsREEL Live -- 3.2 NewsREEL Replay -- 3.3 Summary -- 4 NewsREEL for

Learning and Teaching -- 4.1 Participant's Perspective -- 4.2 An Instructor's Perspective -- 4.3 Summary -- 5 Conclusion and Outlook -- References -- LifeCLEF 2017 Lab Overview: Multimedia Species Identification Challenges -- 1 LifeCLEF Lab Overview -- 2 Task1: PlantCLEF -- 2.1 Dataset and Evaluation Protocol -- 2.2 Participants and Results -- 3 Task2: BirdCLEF -- 3.1 Data and Task Description -- 3.2 Participants and Results -- 4 Task3: SeaCLEF -- 4.1 Data and Task Description -- 4.2 Participants and Results -- 5 Conclusions and Perspectives -- References -- Overview of PAN'17 -- 1 Introduction -- 2 Author Identification -- 2.1 Style Breach Detection -- 2.2 Author Clustering -- 3 Author Profiling -- 3.1 Evaluation Framework -- 3.2 Results -- 4 Author Obfuscation -- 4.1 Evaluation Datasets -- 4.2 Performance Measures -- 4.3 Results -- 5 Summary -- References -- CLEF 2017 eHealth Evaluation Lab Overview -- 1 Introduction -- 2 Materials and Methods -- 2.1 Text Documents -- 2.2 Human Annotations, Queries, and Relevance Assessments -- 2.3 Evaluation Methods -- 3 Results -- 4 Conclusions -- References -- CLEF 2017 Microblog Cultural Contextualization Lab Overview. 1 Introduction: From Microblog to Cultural Contextualization.

Sommario/riassunto

This book constitutes the refereed proceedings of the 8th International Conference of the CLEF Initiative, CLEF 2017, held in Dublin, Ireland, in September 2017. The 7 full papers and 9 short papers presented together with 6 best of the labs papers were carefully reviewed and selected from 38 submissions. In addition, this volume contains the results of 10 benchmarking labs reporting their year long activities in overview talks and lab sessions. The papers address all aspects of information access in any modality and language and cover a broad range of topics in the field of multilingual and multimodal information access evaluation.
