Lecture Notes in Computer Science

11437

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board Members

David Hutchison

Lancaster University, Lancaster, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Zurich, Switzerland

John C. Mitchell

Stanford University, Stanford, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

C. Pandu Rangan

Indian Institute of Technology Madras, Chennai, India

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

More information about this series at http://www.springer.com/series/7409

Leif Azzopardi · Benno Stein · Norbert Fuhr · Philipp Mayr · Claudia Hauff · Djoerd Hiemstra (Eds.)

Advances in Information Retrieval

41st European Conference on IR Research, ECIR 2019 Cologne, Germany, April 14–18, 2019 Proceedings, Part I



Editors
Leif Azzopardi
University of Strathclyde
Glasgow, UK

Norbert Fuhr Duisburg-Essen Duisburg, Germany

Claudia Hauff Delft University of Technology Delft, The Netherlands

Benno Stein Bauhaus Universität Weimar Weimar, Germany

Philipp Mayr
GESIS - Leibniz Institute for the Social Sciences Cologne, Germany

Djoerd Hiemstra D University of Twente Enschede, The Netherlands

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-030-15711-1 ISBN 978-3-030-15712-8 (eBook) https://doi.org/10.1007/978-3-030-15712-8

Library of Congress Control Number: 2019934339

LNCS Sublibrary: SL3 - Information Systems and Applications, incl. Internet/Web, and HCI

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The 41st European Conference on Information Retrieval (ECIR) was held in Cologne, Germany, during April 14–18, 2019, and brought together hundreds of researchers from Europe and abroad. The conference was organized by GESIS – Leibniz Institute for the Social Sciences and the University of Duisburg-Essen—in cooperation with the British Computer Society's Information Retrieval Specialist Group (BCS-IRSG).

These proceedings contain the papers, presentations, workshops, and tutorials given during the conference. This year the ECIR 2019 program boasted a variety of novel work from contributors from all around the world and provided new platforms for promoting information retrieval-related (IR) activities from the CLEF Initiative. In total, 365 submissions were fielded across the tracks from 50 different countries.

The final program included 39 full papers (23% acceptance rate), 44 short papers (29% acceptance rate), eight demonstration papers (67% acceptance rate), nine reproducibility full papers (75% acceptance rate), and eight invited CLEF papers. All submissions were peer reviewed by at least three international Program Committee members to ensure that only submissions of the highest quality were included in the final program. As part of the reviewing process we also provided more detailed review forms and guidelines to help reviewers identify common errors in IR experimentation as a way to help ensure consistency and quality across the reviews.

The accepted papers cover the state of the art in IR: evaluation, deep learning, dialogue and conversational approaches, diversity, knowledge graphs, recommender systems, retrieval methods, user behavior, topic modelling, etc., and also included novel application areas beyond traditional text and Web documents such as the processing and retrieval of narrative histories, images, jobs, biodiversity, medical text, and math. The program boasted a high proportion of papers with students as first authors, as well as papers from a variety of universities, research institutes, and commercial organizations.

In addition to the papers, the program also included two keynotes, four tutorials, four workshops, a doctoral consortium, and an industry day. The first keynote was presented by this year's BCS IRSG Karen Sparck Jones Award winner, Prof. Krisztian Balog, On Entities and Evaluation, and the second keynote was presented by Prof. Markus Strohmaier, On Ranking People. The tutorials covered a range of topics from conducting lab-based experiments and statistical analysis to categorization and deep learning, while the workshops brought together participants to discuss algorithm selection (AMIR), narrative extraction (Text2Story), Bibliometrics (BIR), as well as social media personalization and search (SoMePeAS). As part of this year's ECIR we also introduced a new CLEF session to enable CLEF organizers to report on and promote their upcoming tracks. In sum, this added to the success and diversity of ECIR and helped build bridges between communities.

The success of ECIR 2019 would not have been possible without all the help from the team of volunteers and reviewers. We wish to thank all our track chairs for coordinating the different tracks along with the teams of meta-reviewers and reviewers who helped ensure the high quality of the program. We also wish to thank the demo chairs: Christina Lioma and Dagmar Kern; student mentorship chairs: Ahmet Aker and Laura Dietz; doctoral consortium chairs: Ahmet Aker, Dimitar Dimitrov and Zeljko Carevic; workshop chairs: Diane Kelly and Andreas Rauber; tutorial chairs: Guillaume Cabanac and Suzan Verberne; industry chair: Udo Kruschwitz; publicity chair: Ingo Frommholz; and sponsorship chairs: Jochen L. Leidner and Karam Abdulahhad. We would like to thank our webmaster, Sascha Schüller and our local chair, Nina Dietzel along with all the student volunteers who helped to create an excellent online and offline experience for participants and attendees.

ECIR 2019 was sponsored by: DFG (Deutsche Forschungsgemeinschaft), BCS (British Computer Society), SIGIR (Special Interest Group on Information Retrieval), City of Cologne, Signal Media Ltd, Bloomberg, Knowledge Spaces, Polygon Analytics Ltd., Google, Textkernel, MDPI Open Access Journals, and Springer. We thank them all for their support and contributions to the conference.

Finally, we wish to thank all the authors, reviewers, and contributors to the conference.

April 2019

Leif Azzopardi Benno Stein Norbert Fuhr Philipp Mayr Claudia Hauff Djoerd Hiemstra

Organization

General Chairs

Norbert Fuhr Universität Duisburg-Essen, Germany

Philipp Mayr GESIS – Leibniz Institute for the Social Sciences,

Germany

Program Chairs

Leif Azzopardi University of Glasgow, UK

Benno Stein Bauhaus-Universität Weimar, Germany

Short Papers and Poster Chairs

Claudia Hauff Delft University of Technology, The Netherlands

Djoerd Hiemstra University of Twente, The Netherlands

Workshop Chairs

Diane Kelly University of Tennessee, USA

Andreas Rauber Vienna University of Technology, Austria

Tutorial Chairs

Guillaume Cabanac University of Toulouse, France Suzan Verberne Leiden University, The Netherlands

detail verberne Leiden University, The Neuterlands

Demo Chairs

Christina Lioma University of Copenhagen, Denmark

Dagmar Kern GESIS – Leibniz Institute for the Social Sciences,

Germany

Industry Day Chair

Udo Kruschwitz University of Essex, UK

Proceedings Chair

Philipp Mayr GESIS – Leibniz Institute for the Social Sciences,

Germany

Publicity Chair

Ingo Frommholz University of Bedfordshire, UK

Sponsor Chairs

Jochen L. Leidner Thomson Reuters/University of Sheffield, UK
Karam Abdulahhad GESIS – Leibniz Institute for the Social Sciences,

Germany

Student Mentoring Chairs

Ahmet Aker Universität Duisburg-Essen, Germany Laura Dietz University of New Hampshire, USA

Website Infrastructure

Sascha Schüller GESIS – Leibniz Institute for the Social Sciences,

Germany

Local Chair

Nina Dietzel GESIS – Leibniz Institute for the Social Sciences,

Germany

Program Committee

Mohamed Abdel Maksoud Codoma.tech Advanced Technologies, Egypt

Ahmed Abdelali Research Administration

Karam Abdulahhad GESIS - Leibniz institute for the Social Sciences,

Germany

Dirk Ahlers Norwegian University of Science and Technology,

Norway

Qingyao Ai University of Massachusetts Amherst, USA Ahmet Aker University of Duisburg Essen, Germany

Elif Aktolga Apple Inc., USA Dyaa Albakour Signal Media, UK

Giambattista Amati Fondazione Ugo Bordoni, Italy

Linda Andersson Vienna University of Technology, Austria Avi Arampatzis Democritus University of Thrace, Greece

Ioannis Arapakis Telefonica Research, Spain

Jaime Arguello The University of North Carolina at Chapel Hill, USA

Leif Azzopardi University of Strathclyde, UK
Ebrahim Bagheri Ryerson University, USA
Krisztian Balog University of Stavanger, Norway
Alvaro Barreiro University of A Coruña, Spain

Alberto Barrón-Cedeño Qatar Computing Research Institute, Qatar

Srikanta Bedathur IIT Delhi, India

Alejandro Bellogin Universidad Autonoma de Madrid, Spain

Patrice Bellot Aix-Marseille Université - CNRS (LSIS), France Universidad de Castilla-La Mancha, Spain Pablo Bermeio

Catherine Berrut LIG, Université Joseph Fourier Grenoble I, France

Yahoo Prakhar Biyani

Pierre Bonnet CIRAD. France

National Research Council of Italy - CNR, Italy Gloria Bordogna **Dimitrios Bountouridis** Delft University of Technology, The Netherlands

Pavel Braslavski Ural Federal University, Russia

Paul Buitelaar Insight Centre for Data Analytics, National University

of Ireland Galway, Ireland

IRIT - Université Paul Sabatier Toulouse 3, France Guillaume Cabanac

Universidade da Coruña, Spain Fidel Cacheda

Sylvie Calabretto LIRIS. France

Pável Calado Universidade de Lisboa, Portugal

Delft University of Technology, The Netherlands Arthur Camara

Polytechnic Institute of Tomar, Portugal Ricardo Campos

Bilkent University, Turkey Fazli Can

Universidad Autónoma de Madrid, Spain Iván Cantador Cornelia Caragea University of Illinois at Chicago, USA

GESIS, Germany Zeljko Carevic

Claudio Carpineto Fondazione Ugo Bordoni, Italy

Universidad Autónoma de Madrid, Spain Pablo Castells

University of Glasgow, UK Long Chen

Max Chevalier IRIT. France Manoj Chinnakotla Microsoft, India

Nurendra Choudhary International Institute of Information Technology,

Hyderabad, India

Vincent Claveau IRISA - CNRS, France

Fabio Crestani University of Lugano (USI), Switzerland Bruce Croft University of Massachusetts Amherst, USA

Carnegie Mellon University, USA Zhuyun Dai University of Glasgow, UK Jeffery Dalton University of Washington, USA Martine De Cock Pablo de La Fuente Universidad de Valladolid, Spain

Maarten de Rijke University of Amsterdam, The Netherlands Radboud University, The Netherlands Arjen de Vries Polytechnic University of Milan, Italy Yashar Deldjoo

Kuntal Dev IBM India Research Lab, India Emanuele Di Buccio University of Padua, Italy Giorgio Maria Di Nunzio University of Padua, Italy

University of New Hampshire, USA Laura Dietz

GESIS, Germany Dimitar Dimitrov

University of Strathclyde, UK Mateusz Dubiel

Carsten Eickhoff Brown University, USA Tamer Elsayed Qatar University, Qatar

Liana Ermakova UBO, France

Cristina España-Bonet UdS and DFKI, Germany

Jose Alberto Esquivel Signal Media, UK

Hui Fang University of Delaware, USA

Hossein Fani University of New Brunswick, Canada

Paulo Fernandes PUCRS, Brazil

Nicola Ferro University of Padua, Italy Ingo Frommholz University of Bedfordshire, UK

Norbert Fuhr University of Duisburg-Essen, Germany
Michael Färber University of Freiburg, Germany
Patrick Gallinari LIP6 - University of Paris 6, France

Shreyansh Gandhi Walmart Labs, USA

Debasis Ganguly IBM Ireland Research Lab, Ireland

Wei Gao Victoria University of Wellington, New Zealand

Dario Garigliotti University of Stavanger, Norway
Anastasia Giachanou University of Lugano, Switzerland
Giorgos Giannopoulos Imis Institute, Athena R.C., Greece

Lorraine Goeuriot University Grenoble Alpes, CNRS, France

Julio Gonzalo UNED, Spain

Pawan Goyal IIT Kharagpur, India

Michael Granitzer University of Passau, Germany

Guillaume Gravier CNRS, IRISA, France

Shashank Gupta International Institute of Information Technology,

Hyderabad, India

Cathal Gurrin Dublin City University, Ireland

Matthias Hagen Martin-Luther-Universität Halle-Wittenberg, Germany

Shuguang Han Google, USA

Allan Hanbury Vienna University of Technology, Austria

Preben Hansen Stockholm University, Sweden

Donna Harman NIST, USA

Morgan Harvey Northumbria University, UK

Faegheh Hasibi Norwegian University of Science and Technology,

Norway

Claudia Hauff Delft University of Technology, The Netherlands

Jer Hayes Accenture, Ireland

Ben He University of Chinese Academy of Sciences, China

Nathalie Hernandez IRIT, France

Djoerd Hiemstra University of Twente, The Netherlands
Frank Hopfgartner The University of Sheffield, UK
Andreas Hotho University of Würzburg, Germany

Gilles Hubert IRIT, France

Ali Hürriyetoğlu Koc University, Turkey

Dmitry Ignatov National Research University Higher School

of Economics, Russia

Bogdan Ionescu University Politehnica of Bucharest, Romania

Radu Tudor Ionescu University of Bucharest, Romania

Shoaib Jameel Kent University, UK Adam Jatowt Kyoto University, Japan

Shen Jialie Queen's University, Belfast, UK

Jiepu Jiang Virginia Tech, USA Xiaorui Jiang Aston University, UK

Alexis Joly Inria, France

Gareth Jones Dublin City University, Ireland

Jaap Kamps University of Amsterdam, The Netherlands

Nattiya Kanhabua NTENT, Spain

Jaana Kekäläinen University of Tampere, Finland Diane Kelly University of Tennessee, USA Liadh Kelly Maynooth University, Ireland

Dagmar Kern GESIS, Germany

Roman Kern Graz University of Technology, Austria

Dhruv Khattar International Institute of Information Technology

Hyderabad, India

Julia Kiseleva Microsoft Research AI, USA Dietrich Klakow Saarland University, Germany

Yiannis Kompatsiaris CERTH - ITI, Greece

Kriste Krstovski University of Massachusetts Amherst, USA

Udo Kruschwitz University of Essex, UK

Vaibhav Kumar Carnegie Mellon University, USA

Oren Kurland Technion, Israel

Chiraz Latiri University of Tunis, Tunisia

Wang-Chien Lee The Pennsylvania State University, USA

Teerapong Leelanupab King Mongkut's Institute of Technology Ladkrabang

Mark Levene Birkbeck, University of London, UK Liz Liddy Center for Natural Language Processing,

Syracuse University, USA

Nut Limsopatham Amazon

Chunbin Lin Amazon AWS, USA

Christina Lioma University of Copenhagen, Denmark Aldo Lipani University College London, UK

Nedim Lipka Adobe, USA

Elena Lloret University of Alicante, Spain Fernando Loizides Cardiff University, UK

David Losada University of Santiago de Compostela, Spain Natalia Loukachevitch Research Computing Center of Moscow

State University, Russia

Bernd Ludwig University of Regensburg, Germany

Mihai Lupu Research Studios, Austria
Craig Macdonald University of Glasgow, UK
Andrew Macfarlane City University London, UK

Joao Magalhaes Universidade NOVA de Lisboa, Portugal

Walid Magdy The University of Edinburgh, UK

Marco Maggini University of Siena, Italy Maria Maistro University of Padua, Italy

Thomas Mandl University of Hildesheim, Germany
Behrooz Mansouri Rochester Institute of Technology, USA
Ilya Markov University of Amsterdam, The Netherlands

Stefania Marrara Consorzio C2T, Italy

Miguel Martinez-Alvarez Signal, UK

Bruno Martins IST and INESC-ID - Instituto Superior Técnico,

University of Lisbon, Portugal

Luis Marujo Snap Inc., USA

Yosi Mass IBM Haifa Research Lab, Israel

Philipp Mayr GESIS, Germany
Edgar Meij Bloomberg L.P., UK
Massimo Melucci University of Padua, Italy

Marcelo Mendoza Universidad Técnica Federico Santa María, Brazil

Alessandro Micarelli Roma Tre University, Italy

Dmitrijs Milajevs Queen Mary University of London, UK Seyed Abolghasem Sharif University of Technology, Iran

Mirroshandel

Stefano Mizzaro University of Udine, Italy

Boughanem Mohand IRIT University Paul Sabatier Toulouse, France Felipe Moraes Delft University of Technology, The Netherlands

Ajinkya More Netflix, USA Jose Moreno IRIT/UPS, France

Yashar Moshfeghi University of Strathclyde, UK

Josiane Mothe Institut de Recherche en Informatique de Toulouse,

France

André Mourão Universidade NOVA de Lisboa, Portugal

Henning Müller HES-SO, Switzerland Franco Maria Nardini ISTI-CNR, Italy

Wolfgang Nejdl
Mahmood Neshati
Dong Nguyen

L3S and University of Hannover, Germany
Sharif University of Technology, Iran
University of Twente, The Netherlands

Massimo Nicosia University of Trento, Italy Jian-Yun Nie University of Montreal, Canada

Qiang Ning University of Illinois at Urbana-Champaign, USA Boris Novikov National Research University Higher School

of Economics, Russia

Andreas Nuernberger Otto-von-Guericke University of Magdeburg, Germany

Aurélie Névéol LIMSI, CNRS, Université Paris-Saclay, France

Neil O'Hare Yahoo Research, USA

Salvatore Orlando Università Ca' Foscari Venezia, Italy

Iadh OunisUniversity of Glasgow, UKMourad OussalahUniversity of Oulu, FinlandDeepak P.Queen's University Belfast, Ireland

Jiaul Paik IIT Kharagpur, India

Joao Palotti Vienna University of Technology/Qatar Computing

Research Institute, Austria/Qatar

Girish Palshikar Tata Research Development and Design Centre, India

Javier Parapar University of A Coruña, Spain

Gabriella Pasi Università degli Studi di Milano Bicocca, Italy

Arian Pasquali University of Porto, Portugal

Bidyut Kr. Patra VTT Technical Research Centre, Finland Pavel Pecina Charles University in Prague, Czech Republic

Gustavo Penha UFMG

Avar Pentel University of Tallinn, Estonia

Raffaele Perego ISTI-CNR, Italy

Vivien Petras Humboldt-Universität zu Berlin, Germany

Jeremy Pickens Catalyst Repository Systems, USA

Karen Pinel-Sauvagnat IRIT, France

Florina Piroi Vienna University of Technology, Austria Benjamin Piwowarski CNRS, Pierre et Marie Curie University, France

Vassilis Plachouras Facebook, UK

Bob Planque Vrije Universiteit Amsterdam, The Netherlands

Senja Pollak University of Ljubljana, Slovenia Martin Potthast Leipzig University, Germany

Georges Quénot Laboratoire d'Informatique de Grenoble, CNRS,

France

Razieh Rahimi University of Tehran, Iran

Nitin Ramrakhiyani TCS Research, Tata Consultancy Services Ltd., India

Jinfeng Rao University of Maryland, USA

Andreas Rauber Vienna University of Technology, Austria
Traian Rebedea University Politehnica of Bucharest, Romania

Navid Rekabsaz Idiap Research Institute, Switzerland Steffen Remus University of Hamburg, Germany

Paolo Rosso Universitat Politècnica de València, Spain

Dmitri Roussinov University of Strathclyde, UK Stefan Rueger Knowledge Media Institute, UK

Tony Russell-Rose UXLabs, UK

Alan Said University of Skövde, Sweden Mark Sanderson RMIT University, Australia

Eric Sanjuan Laboratoire Informatique d'Avignon- Université

d'Avignon, France

Rodrygo Santos Universidade Federal de Minas Gerais, Brazil

Kamal Sarkar Jadavpur University, Kolkata, India
Fabrizio Sebastiani Italian National Council of Research, Italy
Florence Sedes IRIT P. Sabatier University, France

Giovanni Semeraro University of Bari, Italy

Procheta Sen Indian Statistical Institute, India

Armin Seyeditabari UNC Charlotte, USA

Mahsa Shahshahani University of Amsterdam, The Netherlands

Azadeh Shakery University of Tehran, Iran

Manish Shrivastava International Institute of Information Technology,

Hyderabad, India

Ritvik Shrivastava Columbia University, USA

Rajat Singh International Institute of Information Technology,

Hyderabad, India

Eero Sormunen University of Tampere, Finland

Laure Soulier Sorbonne Universités UPMC-LIP6, France

Rene Spijker Cochran, The Netherlands

Efstathios Stamatatos University of the Aegean, Greece Benno Stein Bauhaus-Universität Weimar, Germany

L. Venkata Subramaniam
Hanna Suominen
Pascale Sébillot
Lynda Tamine

IBM Research, India
The ANU, Australia
IRISA, France
IRIT, France

Thibaut Thonet University of Grenoble Alpes, France Marko Tkalcic Free University of Bozen-Bolzano, Italy

Nicola Tonellotto ISTI-CNR, Italy

Michael Tschuggnall Institute for computer science, DBIS, Innsbruck,

Austria

Theodora Tsikrika Information Technologies Institute, CERTH, Greece

Denis Turdakov Institute for System Programming RAS

Ferhan Ture Comcast Labs, USA

Yannis Tzitzikas University of Crete and FORTH-ICS, Greece Sumithra Velupillai KTH Royal Institute of Technology, Sweden

Suzan Verberne Leiden University, The Netherlands Vishwa Vinay Adobe Research Bangalore, India

Marco Viviani Università degli Studi di Milano-Bicocca - DISCo,

Italy

Stefanos Vrochidis Information Technologies Institute, Greece Shuohang Wang Singapore Management University, Singapore

Christa Womser-Hacker Universität Hildesheim, Germany

Chenyan Xiong Carnegie Mellon University; Microsoft, USA

Grace Hui Yang Georgetown University, USA

Peilin Yang Twitter Inc., USA

Tao Yang University of California at Santa Barbara, USA Andrew Yates Max Planck Institute for Informatics, Germany

Hai-Tao Yu University of Tsukuba, Japan

Hamed Zamani University of Massachusetts Amherst, USA

Eva Zangerle University of Innsbruck, Austria

Fattane Zarrinkalam Ferdowsi University, Iran

Dan Zhang Facebook, USA Duo Zhang Kunlun Inc.

Shuo Zhang University of Stavanger, Norway
Sicong Zhang Georgetown University, USA
Guoqing Zheng Carnegie Mellon University, USA

Leonardo Zilio Université catholique de Louvain, Belgium Guido Zuccon The University of Queensland, Australia

Sponsors



German Research Foundation



Information Retrieval **Specialist Group**









Bloomberg

Engineering





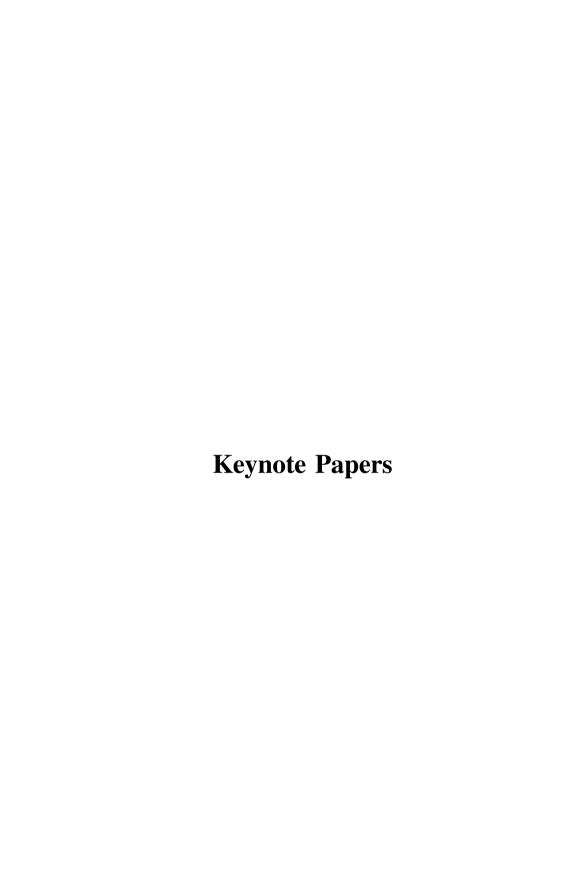






Open Access Journals by MDPI





On Entities and Evaluation

Krisztian Balog

University of Stavanger, Stavanger, Norway krisztian.balog@uis.no

This talk addresses two broad topics, entities and evaluation, which have been the main focus of my research for over ten years. Over the past decade, we have witnessed entities becoming first-class citizens in many information access scenarios [1]. With this has also come an increased reliance on knowledge bases (a.k.a. knowledge graphs), which organize information about entities in a structured and semantically meaningful way. Knowledge bases have enabled significant advancements on specific retrieval tasks, such as entity retrieval and entity linking [2], as well as have contributed to the grand challenge effort of building intelligent personal assistants. The talk provides a brief synthesis of progress thus far, then highlights some open challenges that remain in this space. In particular, the concept of a *personal knowledge graph* is introduced, which is a resource of structured information about entities personally relevant to a given user. A range of tasks associated to personal knowledge graphs are also discussed.

The second part of the talk concentrates on evaluation, which has been a central theme in information retrieval since the inception of the field. For a long time, system-oriented evaluation has primarily been performed using offline test collections, following the Cranfield paradigm. While this rigorous methodology ensures the repeatability and reproducibility of experiments, it is inherently limited by abstracting the actual user, to a large extent, away. In this talk, an argument is made for the (complementary) need of online evaluation. Specifically, the "living labs" evaluation methodology is presented, along with past and current efforts to implement it as a collaborative research and development scheme [3, 4].

References

- Balog, K.: Entity-Oriented Search. The Information Retrieval Series. Springer, Cham (2018). https://eos-book.org
- Hasibi, F., Balog, K. Bratsberg, S.E.: Exploiting entity linking in queries for entity retrieval. In: Proceedings of the 2nd ACM International Conference on the Theory of Information Retrieval, ICTIR 2016, pp. 209–218 (2016)
- 3. Jagerman, R., Balog, K., Rijke, M.D.: Opensearch: Lessons learned from an online evaluation campaign. J. Data Inf. Qual. **10**(3), 13:1–13:15 (2018)
- 4. Schuth, A., Balog, K., Kelly, L.: Overview of the living labs for information retrieval evaluation (LL4IR) CLEF lab 2015. In: Mothe, J., et al. (eds.) CLEF 2015. LNCS, vol. 9283, pp. 484–496. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24027-5_47

Ranking People

Markus Strohmaier

RWTH Aachen University and GESIS - Leibniz Institute for the Social Sciences markus.strohmaier@cssh.rwth-aachen.de http://cssh.rwth-aachen.de

Abstract. The popularity of search engines on the World Wide Web is a testament to the broad impact of the work done by the information retrieval community over the last decades. The advances achieved by this community have not only made the World Wide Web more accessible, they have also made it appealing to consider the application of ranking algorithms to other domains, beyond the ranking of documents. One of the most interesting examples is the domain of *ranking people*. In this talk, I will first highlight some of the many challenges that come with deploying ranking algorithms to individuals. I will then show how mechanisms that are perfectly fine to utilize when ranking documents can have undesired or even detrimental effects when ranking people. This talk intends to stimulate a discussion on the manifold, interdisciplinary challenges around the increasing adoption of ranking algorithms in computational social systems.

Keywords: Information retrieval · Ranking · Computational social science

Contents - Part I

Modeling Relations

Learning Lexical-Semantic Relations Using Intuitive Cognitive Links Georgios Balikas, Gaël Dias, Rumen Moraliyski, Houssam Akhmouch, and Massih-Reza Amini	3
Relationship Prediction in Dynamic Heterogeneous Information Networks Amin Milani Fard, Ebrahim Bagheri, and Ke Wang	19
Retrieving Relationships from a Knowledge Graph for Question Answering	35
Embedding Geographic Locations for Modelling the Natural Environment Using Flickr Tags and Structured Data	51
Classification and Search	
Recognising Summary Articles	69
Towards Content Expiry Date Determination: Predicting Validity Periods of Sentences	86

PRIN: A Probabilistic Recommender with Item Priors and Neural Models . . .

Alfonso Landin, Daniel Valcarce, Javier Parapar, and Álvaro Barreiro

Structural Similarity Search for Formulas Using Leaf-Root Paths

Information Retrieval Models for Contact Recommendation

Javier Sanz-Cruzado and Pablo Castells

Nektaria Potha and Efstathios Stamatatos

Wei Zhong and Richard Zanibbi

Recommender Systems I

102

116

133

148

Conformative Filtering for Implicit Feedback Data	164
Graphs	
Binarized Knowledge Graph Embeddings	181
A Supervised Keyphrase Extraction System Based on Graph Representation Learning	197
Recommender Systems II	
Comparison of Sentiment Analysis and User Ratings in Venue Recommendation	215
AntRS: Recommending Lists Through a Multi-objective Ant Colony System	229
Automated Early Leaderboard Generation from Comparative Tables	244
Query Analytics	
Predicting the Topic of Your Next Query for Just-In-Time IR	261
Identifying Unclear Questions in Community Question Answering Websites	276
Local and Global Query Expansion for Hierarchical Complex Topics Jeffrey Dalton, Shahrzad Naseri, Laura Dietz, and James Allan	290
Representation	
Word Embeddings for Entity-Annotated Texts	307
Vectors of Pairwise Item Preferences	323

Reproducibility (Systems)	
Compressing Inverted Indexes with Recursive Graph Bisection: A Reproducibility Study	339
An Experimental Study of Index Compression and DAAT Query Processing Methods	353
Reproducing and Generalizing Semantic Term Matching in Axiomatic Information Retrieval	369
Optimizing Ranking Models in an Online Setting	382
Simple Techniques for Cross-Collection Relevance Feedback	397
Reproducibility (Application)	
A Comparative Study of Summarization Algorithms Applied to Legal Case Judgments	413
Replicating Relevance-Ranked Synonym Discovery in a New Language and Domain	429
On Cross-Domain Transfer in Venue Recommendation	443
The Effect of Algorithmic Bias on Recommender Systems for Massive Open Online Courses	457
Neural IR	
Domain Adaptive Neural Sentence Compression by Tree Cutting Litton J. Kurisinkel, Yue Zhang, and Vasudeva Varma	475
An Axiomatic Approach to Diagnosing Neural IR Models	489

Cross Lingual IR

Term Selection for Query Expansion in Medical Cross-Lingual Information Retrieval	507
Zero-Shot Language Transfer for Cross-Lingual Sentence Retrieval Using Bidirectional Attention Model	523
QA and Conversational Search	
QRFA: A Data-Driven Model of Information-Seeking Dialogues	541
Iterative Relevance Feedback for Answer Passage Retrieval with Passage-Level Semantic Match	558
Topic Modeling	
LICD: A Language-Independent Approach for Aspect Category Detection Erfan Ghadery, Sajad Movahedi, Masoud Jalili Sabet, Heshaam Faili, and Azadeh Shakery	575
Topic Grouper: An Agglomerative Clustering Approach to Topic Modeling	590
Metrics	
Meta-evaluation of Dynamic Search: How Do Metrics Capture Topical Relevance, Diversity and User Effort?	607
A Markovian Approach to Evaluate Session-Based IR Systems	621
Correlation, Prediction and Ranking of Evaluation Metrics in Information Retrieval	636
Modeling User Actions in Job Search	652

Ima	ge	IR
	-	

Automated Semantic Annotation of Species Names in Handwritten Texts Lise Stork, Andreas Weber, Jaap van den Herik, Aske Plaat, Fons Verbeek, and Katherine Wolstencroft	667
Tangent-V: Math Formula Image Search Using Line-of-Sight Graphs Kenny Davila, Ritvik Joshi, Srirangaraj Setlur, Venu Govindaraju, and Richard Zanibbi	681
Figure Retrieval from Collections of Research Articles	696
"Is This an Example Image?" – Predicting the Relative Abstractness Level of Image and Text	711
Short Papers	
End-to-End Neural Relation Extraction Using Deep Biaffine Attention Dat Quoc Nguyen and Karin Verspoor	729
Social Relation Inference via Label Propagation	739
Wikipedia Text Reuse: Within and Without	747
Stochastic Relevance for Crowdsourcing	755
Exploiting a More Global Context for Event Detection Through Bootstrapping Dorian Kodelja, Romaric Besançon, and Olivier Ferret	763
Faster BlockMax WAND with Longer Skipping	771
A Hybrid Modeling Approach for an Automated Lyrics-Rating System for Adolescents	779
Evaluating Similarity Metrics for Latent Twitter Topics	787

xxvi Contents - Part I

On Interpretability and Feature Representations: An Analysis of the Sentiment Neuron	795
Jonathan Donnelly and Adam Roegiest	1)
Image Tweet Popularity Prediction with Convolutional Neural Network Yihong Zhang and Adam Jatowt	803
Enriching Word Embeddings for Patent Retrieval with Global Context Sebastian Hofstätter, Navid Rekabsaz, Mihai Lupu, Carsten Eickhoff, and Allan Hanbury	810
Incorporating External Knowledge to Boost Machine Comprehension	
Based Question Answering	819
Impact of Training Dataset Size on Neural Answer Selection Models Trond Linjordet and Krisztian Balog	828
Unsupervised Explainable Controversy Detection from Online News Youngwoo Kim and James Allan	836
Extracting Temporal Event Relations Based on Event Networks	844
Dynamic-Keyword Extraction from Social Media	852
Local Popularity and Time in top-N Recommendation	861
Multiple Keyphrase Generation Model with Diversity	869
Author Index	877

Contents – Part II

Short Papers (Continued)

Open-Set Web Genre Identification Using Distributional Features and Nearest Neighbors Distance Ratio	3
Exploiting Global Impact Ordering for Higher Throughput in Selective Search	12
Cross-Domain Recommendation via Deep Domain Adaptation	20
It's only Words and Words Are All I Have	30
Modeling User Return Time Using Inhomogeneous Poisson Process	37
Inductive Transfer Learning for Detection of Well-Formed Natural Language Search Queries	45
Towards Spatial Word Embeddings	53
Asymmetry Sensitive Architecture for Neural Text Matching	62
QGraph: A Quality Assessment Index for Graph Clustering	70
A Neural Approach to Entity Linking on Wikidata	78
Self-attentive Model for Headline Generation	87

Can Image Captioning Help Passage Retrieval in Multimodal Question Answering?	94
A Simple Neural Approach to Spatial Role Labelling	102
Neural Diverse Abstractive Sentence Compression Generation	109
Fully Contextualized Biomedical NER	117
DeepTagRec: A Content-cum-User Based Tag Recommendation Framework for Stack Overflow	125
Document Performance Prediction for Automatic Text Classification Gustavo Penha, Raphael Campos, Sérgio Canuto, Marcos André Gonçalves, and Rodrygo L. T. Santos	132
Misleading Metadata Detection on YouTube	140
A Test Collection for Passage Retrieval Evaluation of Spanish Health-Related Resources Eleni Kamateri, Theodora Tsikrika, Spyridon Symeonidis, Stefanos Vrochidis, Wolfgang Minker, and Yiannis Kompatsiaris	148
On the Impact of Storing Query Frequency History for Search Engine Result Caching	155
AspeRa: Aspect-Based Rating Prediction Model	163
Heterogeneous Edge Embedding for Friend Recommendation Janu Verma, Srishti Gupta, Debdoot Mukherjee, and Tanmoy Chakraborty	172
Public Sphere 2.0: Targeted Commenting in Online News Media Ankan Mullick, Sayan Ghosh, Ritam Dutt, Avijit Ghosh, and Abhijnan Chakraborty	180

Contents – Part II	xxix
An Extended CLEF eHealth Test Collection for Cross-Lingual Information Retrieval in the Medical Domain	188
An Axiomatic Study of Query Terms Order in Ad-Hoc Retrieval	196
Deep Neural Networks for Query Expansion Using Word Embeddings Ayyoob Imani, Amir Vakili, Ali Montazer, and Azadeh Shakery	203
Demonstration Papers	
Online Evaluations for Everyone: Mr. DLib's Living Lab for Scholarly Recommendations	213
StyleExplorer: A Toolkit for Textual Writing Style Visualization	220
MedSpecSearch: Medical Specialty Search	225
Contender: Leveraging User Opinions for Purchase Decision-Making Tiago de Melo, Altigran S. da Silva, Edleno S. de Moura, and Pável Calado	230
Rethinking 'Advanced Search': A New Approach to Complex Query Formulation	236
node-indri: Moving the Indri Toolkit to the Modern Web Stack Felipe Moraes and Claudia Hauff	241
PaperHunter: A System for Exploring Papers and Citation Contexts Michael Färber, Ashwath Sampath, and Adam Jatowt	246
Interactive System for Automatically Generating Temporal Narratives Arian Pasquali, Vitor Mangaravite, Ricardo Campos, Alípio Mário Jorge, and Adam Jatowt	251
CLEF Organizers Lab Track	
Early Detection of Risks on the Internet: An Exploratory Campaign David E. Losada, Fabio Crestani, and Javier Parapar	259

CLEF eHealth 2019 Evaluation Lab	267
LifeCLEF 2019: Biodiversity Identification and Prediction Challenges Alexis Joly, Hervé Goëau, Christophe Botella, Stefan Kahl, Marion Poupard, Maximillien Servajean, Hervé Glotin, Pierre Bonnet, Willem-Pier Vellinga, Robert Planqué, Jan Schlüter, Fabian-Robert Stöter, and Henning Müller	275
CENTRE@CLEF 2019	283
A Decade of Shared Tasks in Digital Text Forensics at PAN	291
ImageCLEF 2019: Multimedia Retrieval in Lifelogging, Medical, Nature, and Security Applications. Bogdan Ionescu, Henning Müller, Renaud Péteri, Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler, Minh-Triet Tran, Mathias Lux, Cathal Gurrin, Yashin Dicente Cid, Vitali Liauchuk, Vassili Kovalev, Asma Ben Abacha, Sadid A. Hasan, Vivek Datla, Joey Liu, Dina Demner-Fushman, Obioma Pelka, Christoph M. Friedrich, Jon Chamberlain, Adrian Clark, Alba García Seco de Herrera, Narciso Garcia, Ergina Kavallieratou, Carlos Roberto del Blanco, Carlos Cuevas Rodríguez, Nikos Vasillopoulos, and Konstantinos Karampidis	301
CheckThat! at CLEF 2019: Automatic Identification and Verification of Claims	309
A Task Set Proposal for Automatic Protest Information Collection Across Multiple Countries	316
Doctoral Consortium Papers	
Exploring Result Presentation in Conversational IR Using a Wizard-of-Oz Study	327

	Contents – Part II	xxxi
Keyword Search on RDF Datasets		332
Learning User and Item Representations for Recommen Alfonso Landin	nder Systems	337
Improving the Annotation Efficiency and Effectiveness in the Text Domain		343
Logic-Based Models for Social Media Retrieval Firas Sabbah		348
Adapting Models for the Case of Early Risk Prediction Razan Masood	on the Internet	353
Integration of Images into the Patent Retrieval Process Wiebke Thode		359
Dialogue-Based Information Retrieval		364
Dynamic Diversification for Interactive Complex Search Ameer Albahem	h	369
The Influence of Backstories on Queries with Varying in Task-Based Specialised Information Retrieval Manuel Steiner		375
Workshops		
Proposal for the 1 st Interdisciplinary Workshop on Algo and Meta-Learning in Information Retrieval (AMIR) Joeran Beel and Lars Kotthoff		383
The 2 nd International Workshop on Narrative Extraction Text2Story 2019		389
Bibliometric-Enhanced Information Retrieval: 8th Internal BIR Workshop		394
Third Workshop on Social Media for Personalization at (SoMePeAS 2019)		401

xxxii Contents - Part II

Tutorials

Conducting Laboratory Experiments Properly with Statistical Tools: An Easy Hands-On Tutorial	405
Text Categorization with Style	408
Concept to Code: Neural Networks for Sequence Learning	410
A Tutorial on Basics, Applications and Current Developments in PLS Path Modeling for the IIR Community	413
Author Index	417