

Founding Editors

Gerhard Goos

Karlsruhe Institute of Technology, Karlsruhe, Germany

Juris Hartmanis

Cornell University, Ithaca, NY, USA

Editorial Board Members

Elisa Bertino

Purdue University, West Lafayette, IN, USA

Wen Gao

Peking University, Beijing, China

Bernhard Steffen 

TU Dortmund University, Dortmund, Germany

Gerhard Woeginger 

RWTH Aachen, Aachen, Germany

Moti Yung

Columbia University, New York, NY, USA


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

Avi Arampatzis · Evangelos Kanoulas ·
Theodora Tsikrika · Stefanos Vrochidis ·
Hideo Joho · Christina Lioma ·
Carsten Eickhoff · Aurélie Névéol ·
Linda Cappellato · Nicola Ferro (Eds.)


Experimental IR Meets Multilinguality, Multimodality, and Interaction


11th International Conference of the CLEF Association, CLEF 2020
Thessaloniki, Greece, September 22–25, 2020
Proceedings


Editors


Avi Arampatzis 
Department of Electrical
and Computer Engineering
Democritus University of Thrace
Xanthi, Greece


Theodora Tsikrika
Information Technologies Institute
Centre for Research and Technology Hellas
Thessaloniki, Greece


Hideo Joho 
Faculty of Library, Information
and Media Science
University of Tsukuba
Ibaraki, Japan


Carsten Eickhoff 
Brown University
Providence, RI, USA

Linda Cappellato 
Department of Information Engineering
University of Padova
Padua, Italy

Evangelos Kanoulas 
University of Amsterdam
Amsterdam, The Netherlands

Stefanos Vrochidis 
Information Technologies Institute
Centre for Research and Technology Hellas
Thessaloniki, Greece

Christina Lioma 
Department of Computer Science
University of Copenhagen
Copenhagen, Denmark

Aur lie N v  ol 
LIMSI-CNRS
Orsay, France

University Paris-Saclay
Saint-Aubin, France

Nicola Ferro 
Department of Information Engineering
University of Padova
Padua, Italy

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Computer Science

ISBN 978-3-030-58218-0

ISBN 978-3-030-58219-7 (eBook)

<https://doi.org/10.1007/978-3-030-58219-7>

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

  Springer Nature Switzerland AG 2020

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

Since 2000, the Conference and Labs of the Evaluation Forum (CLEF) has played a leading role in stimulating research and innovation in the domain of multimodal and multilingual information access. Initially founded as the Cross-Language Evaluation Forum and running in conjunction with the European Conference on Digital Libraries (ECDL/TPDL), CLEF became a stand-alone event in 2010 combining a peer-reviewed conference with a multi-track evaluation forum. The combination of the scientific program and the track-based evaluations at the CLEF conference creates a unique platform to explore information access from different perspectives, in any modality and language.

The CLEF conference has a clear focus on experimental information retrieval (IR) as seen in evaluation forums (like CLEF Labs, TREC, NTCIR, FIRE, MediaEval, RomIP, TAC, etc.) with special attention to the challenges of multimodality, multilinguality, and interactive search ranging from unstructured, to semi-structured, and structured data. CLEF invites submissions on new insights demonstrated by the use of innovative IR evaluation tasks or in the analysis of IR test collections and evaluation measures, as well as on concrete proposals to push the boundaries of the Cranfield/TREC/CLEF paradigm.

CLEF 2020¹ was jointly organized by the Center for Research and Technology Hellas (CERTH), the University of Amsterdam, and the Democritus University of Thrace, and it was expected to be hosted by CERTH, and in particular by the Multimedia Knowledge and Social Media Analytics Laboratory of its Information Technologies Institute, at the premises of CERTH, in Thessaloniki, Greece, during September 22–25, 2020. The outbreak of the COVID-19 pandemic in early 2020 affected the organization of CLEF 2020. The CLEF Steering Committee, along with the organizers of CLEF 2020, after detailed discussions, decided to run the conference fully virtually. The conference format remained the same as in past years, and consisted of keynotes, contributed papers, lab sessions, and poster sessions, including reports from other benchmarking initiatives from around the world. All sessions were organized and run online.

CLEF 2020 continued the initiative introduced in the 2019 edition during which, the European Conference for Information Retrieval (ECIR) and CLEF joined forces: ECIR 2020 hosted a special session dedicated to CLEF labs where lab organizers presented the major outcomes of their labs and their plans for ongoing activities, followed by a poster session to favor discussion during the conference. This was reflected in the ECIR 2020 proceedings, where CLEF lab activities and results were reported as short papers. The goal was not only to engage the ECIR community in CLEF activities but also to disseminate the research results achieved during CLEF evaluation cycles as submission of papers to ECIR.

¹ <http://clef2020.clef-initiative.eu/>

The following scholars were invited to give a keynote talk at CLEF 2020. Ellen Voorhess (NIST, USA) delivered a talk entitled “Building Reusable Test Collections” which focused on reviewing various approaches for building fair, reusable test collections with large documents sets. Yiannis Kompasiaris (CERTH-ITI, Greece) gave a speech on “Social media mining for sensing and responding to real-world trends and events”, presenting the unique opportunity social media offer to discover, collect, and extract relevant information that provides useful insights in areas ranging from news to environmental and security topics, while addressing key challenges and issues, such as fighting misinformation and analyzing multimodal and multilingual information.

CLEF 2020 received a total of nine submissions, of which a total of seven papers (five long, two short) were accepted. Each submission was reviewed by three Program Committee members, and the program chairs oversaw the reviewing and follow-up discussions. Seven countries are represented in the accepted papers where many of them were a product of international collaboration. This year, researchers addressed the following important challenges in the community: a large-scale evaluation of translation effects in academic search, advancement of assessor-driven aggregation methods for efficient relevance assessments, development of a new test collection or dataset for 1) missing data detection methods in knowledge-base, 2) Russian reading comprehension, and 3) under-resourced languages such as Amharic (Ethiopia), revisiting the concept of session boundaries with fresh eyes, and development of argumentative document retrieval methods.

Like in previous editions since 2015, CLEF 2020 continued inviting CLEF lab organizers to nominate a “best of the labs” paper that was reviewed as a full paper submission to the CLEF 2020 conference according to the same review criteria and PC. Seven full papers were accepted for this “best of the labs” section.

The conference integrated a series of workshops presenting the results of lab-based comparative evaluations. CLEF 2020 was the 11th year of the CLEF conference and the 21st year of the CLEF initiative as a forum for IR evaluation. 15 lab proposals were received and evaluated in peer review based on their innovation potential and the quality of the resources created. The 12 selected labs represented scientific challenges based on new data sets and real-world problems in multimodal and multilingual information access. These data sets provide unique opportunities for scientists to explore collections, to develop solutions for these problems, to receive feedback on the performance of their solutions, and to discuss the issues with peers at the workshops.

In addition to these workshops, the labs reported results of their year-long activities in overview talks and lab sessions. Overview papers describing each of the labs are provided in this volume. The full details for each lab are contained in a separate publication, the Working Notes².

The 12 labs running as part of CLEF 2020 comprised new labs (ARQMath, CheMU, HIPE, Lilas, and Touché) as well as seasoned labs that offered previous editions at CLEF (CheckThat!, CLEF eHealth, eRisk, ImageCLEF, LifeCLEF, and PAN) or in other platforms (BioASQ). The following labs were offered:

² Cappellato, L., Eickhoff, C., Ferro, N., and Névoul, A., editors (2020). *CLEF 2020 Working Notes*. CEUR Workshop Proceedings (CEUR-WS.org), ISSN 1613-0073.

ARQMath: Answer Retrieval for Mathematical Questions³ considers the problem of finding answers to new mathematical questions among posted answers on the community question answering site Math Stack Exchange. The goals of the lab are to develop methods for mathematical IR based on both text and formula analysis.

BioASQ⁴ challenges researchers with large-scale biomedical semantic indexing and question answering (QA). The challenges include tasks relevant to hierarchical text classification, machine learning, IR, QA from texts and structured data, multi-document summarization, and many other areas. The aim of the BioASQ workshop is to push the research frontier towards systems that use the diverse and voluminous information available online to respond directly to the information needs of biomedical scientists.

ChEMU: Information Extraction from Chemical Patents⁵ proposes two key information extraction tasks over chemical reactions from patents. Task 1 aims to identify chemical compounds and their specific types, i.e. to assign the label of a chemical compound according to the role which it plays within a chemical reaction. Task 2 requires identification of event trigger words (e.g. “added” and “stirred”) which all have the same type of “EVENT_TRIGGER”, and then determination of the chemical entity arguments of these events.

CheckThat!: Identification and Verification of Political Claims⁶ aims to foster the development of technology capable of both spotting and verifying check-worthy claims in political debates in English, Arabic, and Italian. The concrete tasks were to assess the check worthiness of a claim in a tweet, check if a (similar) claim has been previously verified, retrieve evidence to fact-check a claim, and verify the factuality of a claim.

CLEF eHealth⁷ aims to support the development of techniques to aid laypeople, clinicians, and policy-makers in easily retrieving and making sense of medical content to support their decision making. The goals of the lab are to develop processing methods and resources in a multilingual setting to enrich difficult-to-understand eHealth texts and provide valuable documentation.

eRisk: Early Risk Prediction on the Internet⁸ explores challenges of evaluation methodology, effectiveness metrics, and other processes related to early risk detection. Early detection technologies can be employed in different areas, particularly those related to health and safety. The 2020 edition of the lab focused on texts written on social media for the early detection of signs of self-harm and depression.

HIPE: Named Entity Processing on Historical Newspapers⁹ aims at fostering named entity recognition on heterogeneous, historical, and noisy inputs. The goals of the lab are to strengthen the robustness of existing approaches on non-standard input; to enable performance comparison of named entity processing on historical texts;

³ <https://www.cs.rit.edu/dprl/ARQMath/>.

⁴ <http://www.bioasq.org/workshop2020>.

⁵ <http://chemu.eng.unimelb.edu.au/>.

⁶ <https://sites.google.com/view/clef2020-checkthat>.

⁷ <http://clef-ehealth.org/>.

⁸ <http://erisk.irlab.org/>.

⁹ <https://impresso.github.io/CLEF-HIPE-2020/>.

and, in the long run, to foster efficient semantic indexing of historical documents in order to support scholarship on digital cultural heritage collections.

ImageCLEF: Multimedia Retrieval¹⁰ provides an evaluation forum for visual media analysis, indexing, classification/learning, and retrieval in medical, nature, security, and lifelogging applications with a focus on multimodal data, so data from a variety of sources and media.

LifeCLEF: Biodiversity Identification and Prediction¹¹ aims at boosting research on the identification and prediction of living organisms in order to solve the taxonomic gap and improve our knowledge of biodiversity. Through its biodiversity informatics related challenges, LifeCLEF is intended to push the boundaries of the state of the art in several research directions at the frontier of multimedia IR, machine learning, and knowledge engineering.

Lilas: Living Labs for Academic Search¹² aims to bring together researchers interested in the online evaluation of academic search systems. The long term goal is to foster knowledge on improving the search for academic resources like literature, research data, and the interlinking between these resources in fields from the Life Sciences and the Social Sciences. The immediate goal of this lab is to develop ideas, best practices, and guidelines for a full online evaluation campaign at CLEF 2021.

PAN: Digital Text Forensics and Stylometry¹³ is a networking initiative for the digital text forensics, where researchers and practitioners study technologies that analyze texts with regard to originality, authorship, and trustworthiness. PAN provides evaluation resources consisting of large-scale corpora, performance measures, and web services that allow for meaningful evaluations. The main goal is to provide for sustainable and reproducible evaluations, to get a clear view of the capabilities of state-of-the-art algorithms.

Touché: Argument Retrieval¹⁴ is the first shared task on the topic of argument retrieval. Decision-making processes, be it at the societal or at the personal level, eventually come to a point where one side will challenge the other with a why-question, which is a prompt to justify one's stance. Thus, technologies for argument mining and argumentation processing are maturing at a rapid pace, giving rise for the first time to argument retrieval.

As a group, the 71 lab organizers were based in 14 countries, with Germany and France leading the distribution. Despite CLEF's traditionally Europe-based audience, 18 (25.4%) organizers were affiliated with international institutions outside of Europe. The gender distribution was biased towards 81.3% male organizers.

The success of CLEF 2020 would not have been possible without the huge effort of several people and organizations, including the CLEF Association¹⁵, the Program Committee, the Lab Organizing Committee, the reviewers, and the many students and

¹⁰ <https://www.imageclef.org/2019>.

¹¹ <http://www.lifeclef.org/>.

¹² <https://clef-lilas.github.io/>.

¹³ <http://pan.webis.de/>.

¹⁴ <https://events.webis.de/touche-20/>.

¹⁵ <http://www.clef-initiative.eu/association>.

volunteers who contributed. Finally, we thank the generous support of ACM SIGIR and BCS IRSG, both of which provided general funding support.

July 2020

Avi Arampatzis
Evangelos Kanoulas
Theodora Tsikrika
Stefanos Vrochidis
Hideo Joho
Christina Lioma
Aurélie Névéol
Carsten Eickhoff
Linda Cappellato
Nicola Ferro

Organization

CLEF 2020, Conference and Labs of the Evaluation Forum – Experimental IR meets Multilingualism, Multimodality, and Interaction, was hosted (online) by the Multimedia Knowledge and Social Media Analytics Laboratory (MKLab) of the Information Technologies Institute (ITI) of the Center for Research and Technology Hellas (CERTH), Thessaloniki, Greece.

General Chairs

Evangelos Kanoulas
Theodora Tsikrika
Stefanos Vrochidis
Avi Arampatzis

University of Amsterdam, The Netherlands
Information Technologies Institute, CERTH, Greece
Information Technologies Institute, CERTH, Greece
Democritus University of Thrace, Greece

Program Chairs

Hideo Joho
Christina Lioma

University of Tsukuba, Japan
University of Copenhagen, Denmark

Lab Chairs

Aurélié Névél
Carsten Eickhoff

Université Paris Saclay, CNRS, LIMSI, France
Brown University, USA

Lab Mentorship Chair

Lorraine Goeuriot

Université Grenoble Alpes, France

Proceedings Chairs

Linda Cappellato
Nicola Ferro

University of Padua, Italy
University of Padua, Italy

Local Organization

Vivi Ntrigkogia

Information Technologies Institute, CERTH, Greece

CLEF Steering Committee

Steering Committee Chair

Nicola Ferro University of Padua, Italy

Deputy Steering Committee Chair for the Conference

Paolo Rosso Universitat Politècnica de València, Spain

Deputy Steering Committee Chair for the Evaluation Labs

Martin Braschler Zurich University of Applied Sciences, Switzerland

Members

Khalid Choukri	Evaluations and Language resources Distribution Agency (ELDA), France
Paul Clough	University of Sheffield, UK
Fabio Crestani	Università della Svizzera Italiana, Switzerland
Norbert Fuhr	University of Duisburg-Essen, Germany
Lorraine Goeuriot	Université Grenoble Alpes, France
Julio Gonzalo	National Distance Education University (UNED), Spain
Donna Harman	National Institute for Standards and Technology (NIST), USA
Djoerd Hiemstra	University of Twente, The Netherlands
Evangelos Kanoulas	University of Amsterdam, The Netherlands
Birger Larsen	University of Aalborg, Denmark
David E. Losada	Universidade de Santiago de Compostela, Spain
Mihai Lupu	Vienna University of Technology, Austria
Josiane Mothe	IRIT, Université de Toulouse, France
Henning Müller	University of Applied Sciences Western Switzerland (HES-SO), Switzerland
Jian-Yun Nie	Université de Montréal, Canada
Eric SanJuan	University of Avignon, France
Giuseppe Santucci	Sapienza University of Rome, Italy
Jacques Savoy	University of Neuchâtel, Switzerland
Laure Soulier	Pierre and Marie Curie University (Paris 6), France
Christa Womser-Hacker	University of Hildesheim, Germany

Past Members

Jaana Kekäläinen	University of Tampere, Finland
Séamus Lawless	Trinity College Dublin, Ireland
Carol Peters	ISTI, National Council of Research (CNR), Italy (Steering Committee Chair 2000–2009)
Emanuele Pianta	Centre for the Evaluation of Language and Communication Technologies (CELCT), Italy
Maarten de Rijke	University of Amsterdam, The Netherlands
Alan Smeaton	Dublin City University, Ireland

Sponsors



Information Retrieval
Specialist Group

SIGIR
Special Interest Group
on Information Retrieval

Contents

Full Papers

SberQuAD – Russian Reading Comprehension Dataset: Description and Analysis.	3
<i>Pavel Efimov, Andrey Chertok, Leonid Boytsov, and Pavel Braslavski</i>	
s-AWARE: Supervised Measure-Based Methods for Crowd-Assessors Combination.	16
<i>Marco Ferrante, Nicola Ferro, and Luca Piazzon</i>	
Query or Document Translation for Academic Search – What’s the Real Difference?	28
<i>Vivien Petras, Andreas Lüscho, Roland Ramthun, Juliane Stiller, Cristina España-Bonet, and Sophie Henning</i>	
Question Answering When Knowledge Bases are Incomplete	43
<i>Camille Pradel, Damien Sileo, Álvaro Rodrigo, Anselmo Peñas, and Eneko Agirre</i>	
2AIRC: The Amharic Adhoc Information Retrieval Test Collection.	55
<i>Tilahun Yeshambel, Josiane Mothe, and Yaregal Assabie</i>	

Short Papers

The Curious Case of Session Identification.	69
<i>Florian Dietz</i>	
Argument Retrieval from Web	75
<i>Mahsa S. Shahshahani and Jaap Kamps</i>	

Best of CLEF 2019 Labs

File Forgery Detection Using a Weighted Rule-Based System.	85
<i>João Rafael Almeida, Olga Fajarda, and José Luís Oliveira</i>	
Protest Event Detection: When Task-Specific Models Outperform an Event-Driven Method	97
<i>Angelo Basile and Tommaso Caselli</i>	
A Study on a Stopping Strategy for Systematic Reviews Based on a Distributed Effort Approach	112
<i>Giorgio Maria Di Nunzio</i>	

Fact Check-Worthiness Detection with Contrastive Ranking	124
<i>Casper Hansen, Christian Hansen, Jakob Grue Simonsen, and Christina Lioma</i>	
Tuberculosis CT Image Analysis Using Image Features Extracted by 3D Autoencoder.	131
<i>Siarhei Kazlouski</i>	
Twitter User Profiling: Bot and Gender Identification: Notebook for PAN at CLEF 2019	141
<i>Dijana Kosmajac and Vlado Keselj</i>	
Medical Image Tagging by Deep Learning and Retrieval	154
<i>Vasiliki Kougia, John Pavlopoulos, and Ion Androutsopoulos</i>	
CLEF 2020 Lab Overviews	
Overview of ARQMath 2020: CLEF Lab on Answer Retrieval for Questions on Math.	169
<i>Richard Zanibbi, Douglas W. Oard, Anurag Agarwal, and Behrooz Mansouri</i>	
Overview of BioASQ 2020: The Eighth BioASQ Challenge on Large-Scale Biomedical Semantic Indexing and Question Answering	194
<i>Anastasios Nentidis, Anastasia Krithara, Konstantinos Bougiatiotis, Martin Krallinger, Carlos Rodriguez-Penagos, Marta Villegas, and Georgios Paliouras</i>	
Overview of CheckThat! 2020: Automatic Identification and Verification of Claims in Social Media	215
<i>Alberto Barrón-Cedeño, Tamer Elsayed, Preslav Nakov, Giovanni Da San Martino, Maram Hasanain, Reem Suwaileh, Fatima Haouari, Nikolay Babulkov, Bayan Hamdan, Alex Nikolov, Shaden Shaar, and Zien Sheikh Ali</i>	
Overview of ChEMU 2020: Named Entity Recognition and Event Extraction of Chemical Reactions from Patents.	237
<i>Jiayuan He, Dat Quoc Nguyen, Saber A. Akhondi, Christian Druckenbrodt, Camilo Thorne, Ralph Hoessel, Zubair Afzal, Zenan Zhai, Biaoyan Fang, Hiyori Yoshikawa, Ameer Albahem, Lawrence Cavedon, Trevor Cohn, Timothy Baldwin, and Karin Verspoor</i>	
Overview of the CLEF eHealth Evaluation Lab 2020	255
<i>Lorraine Goeuriot, Hanna Suominen, Liadh Kelly, Antonio Miranda-Escalada, Martin Krallinger, Zhengyang Liu, Gabiella Pasi, Gabriela Gonzalez Saez, Marco Viviani, and Chenchen Xu</i>	

Overview of eRisk 2020: Early Risk Prediction on the Internet.	272
<i>David E. Losada, Fabio Crestani, and Javier Parapar</i>	
Overview of CLEF HIPE 2020: Named Entity Recognition and Linking on Historical Newspapers.	288
<i>Maud Ehrmann, Matteo Romanello, Alex Flückiger, and Simon Clemenide</i>	
Overview of the ImageCLEF 2020: Multimedia Retrieval in Medical, Lifelogging, Nature, and Internet Applications	311
<i>Bogdan Ionescu, Henning Müller, Renaud Péteri, Asma Ben Abacha, Vivek Datla, Sadid A. Hasan, Dina Demner-Fushman, Serge Kozlovski, Vitali Liauchuk, Yashin Dicente Cid, Vassili Kovalev, Obioma Pelka, Christoph M. Friedrich, Alba García Seco de Herrera, Van-Tu Ninh, Tu-Khiem Le, Liting Zhou, Luca Piras, Michael Riegler, Pål Halvorsen, Minh-Triet Tran, Mathias Lux, Cathal Gurrin, Duc-Tien Dang-Nguyen, Jon Chamberlain, Adrian Clark, Antonio Campello, Dimitri Fichou, Raul Berari, Paul Brie, Mihai Dogariu, Liviu Daniel Ștefan, and Mihai Gabriel Constantin</i>	
Overview of LifeCLEF 2020: A System-Oriented Evaluation of Automated Species Identification and Species Distribution Prediction.	342
<i>Alexis Joly, Hervé Goëau, Stefan Kahl, Benjamin Deneu, Maximilien Servajean, Elijah Cole, Lukáš Pícek, Rafael Ruiz de Castañeda, Isabelle Bolon, Andrew Durso, Titouan Lorieul, Christophe Botella, Hervé Glotin, Julien Champ, Ivan Eggel, Willem-Pier Vellinga, Pierre Bonnet, and Henning Müller</i>	
Overview of LiLAS 2020 – Living Labs for Academic Search	364
<i>Philipp Schaer, Johann Schaible, and Leyla Jael Garcia Castro</i>	
Overview of PAN 2020: Authorship Verification, Celebrity Profiling, Profiling Fake News Spreaders on Twitter, and Style Change Detection.	372
<i>Janek Bevendorff, Bilal Ghanem, Anastasia Giachanou, Mike Kestemont, Enrique Manjavacas, Ilia Markov, Maximilian Mayerl, Martin Potthast, Francisco Rangel, Paolo Rosso, Günther Specht, Efsthios Stamatatos, Benno Stein, Matti Wiegmann, and Eva Zangerle</i>	
Overview of Touché 2020: Argument Retrieval: Extended Abstract.	384
<i>Alexander Bondarenko, Maik Fröbe, Meriem Beloucif, Lukas Gienapp, Yamen Ajjour, Alexander Panchenko, Chris Biemann, Benno Stein, Henning Wachsmuth, Martin Potthast, and Matthias Hagen</i>	
Author Index	397