

Preface

The CLEF 2020 conference is the twentyfirst edition of the popular CLEF campaign and workshop series that has run since 2000 contributing to the systematic evaluation of multilingual and multimodal information access systems, primarily through experimentation on shared tasks. In 2010 CLEF was launched in a new format, as a conference with research presentations, panels, poster and demo sessions and laboratory evaluation workshops. These are proposed and operated by groups of organizers volunteering their time and effort to define, promote, administrate and run an evaluation activity.

CLEF 2020¹ was jointed 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 from 22 to 25 September 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.

15 lab proposals were received and evaluated in peer review based on their innovation potential and the quality of the resources created. To identify the best proposals, besides the well-established criteria from the editions of previous years of CLEF such as topical relevance, novelty, potential impact on future world affairs, likely number of participants, and the quality of the organizing consortium, this year we further stressed the connection to real-life usage scenarios and we tried to avoid as much as possible overlaps among labs in order to promote synergies and integration.

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.

We continued the mentorship program to support the preparation of lab proposals for newcomers to CLEF. The CLEF newcomers mentoring program offered help, guidance, and feedback on the writing of draft lab proposals by assigning a mentor to proponents, who helped them in preparing and maturing the lab proposal for submission. If the lab proposal fell into the scope of an

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

already existing CLEF lab, the mentor helped proponents to get in touch with those lab organizers and team up forces.

Building on previous experience, the Labs at CLEF 2020 demonstrate the maturity of the CLEF evaluation environment by creating new tasks, new and larger data sets, new ways of evaluation or more languages. Details of the individual Labs are described by the Lab organizers in these proceedings. Below is a short summary of them.

ARQMath: Answer Retrieval for Mathematical Questions² 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 information retrieval based on both text and formula analysis.

BioASQ³ challenges researchers with large-scale biomedical semantic indexing and question answering (QA). The challenges include tasks relevant to hierarchical text classification, machine learning, information retrieval, 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 Patents⁴ 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 Claims⁵ 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 eHealth⁶ 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.

² <https://www.cs.rit.edu/~dprl/ARQMath/>

³ <http://www.bioasq.org/workshop2020>

⁴ <http://chemu.eng.unimelb.edu.au/>

⁵ <https://sites.google.com/view/clef2020-checkthat>

⁶ <http://clef-ehealth.org/>

eRisk: Early Risk Prediction on the Internet⁷ 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 in social media for the early detection of signs of self-harm and depression.

HIPE: Named Entity Processing on Historical Newspapers⁸ 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; 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 Retrieval⁹ 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 Prediction¹⁰ 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 information retrieval, machine learning and knowledge engineering.

Lilas: Living Labs for Academic Search¹¹ 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 Stylometry¹² 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.

⁷ <http://erisk.irlab.org/>

⁸ <https://impresso.github.io/CLEF-HIPE-2020/>

⁹ <https://www.imageclef.org/2019>

¹⁰ <http://www.lifeclef.org/>

¹¹ <https://clef-lilas.github.io/>

¹² <http://pan.webis.de/>

Touché: Argument retrieval¹³ 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.

CLEF has always been backed by European projects that complement the incredible amount of volunteering work performed by Lab Organizers and the CLEF community with the resources needed for its necessary central coordination, in a similar manner to the other major international evaluation initiatives such as TREC, NTCIR, FIRE and MediaEval. Since 2014, the organisation of CLEF no longer has direct support from European projects and are working to transform itself into a self-sustainable activity. This is being made possible thanks to the establishment of the CLEF Association¹⁴, a non-profit legal entity in late 2013, which, through the support of its members, ensures the resources needed to smoothly run and coordinate CLEF.

Acknowledgments

We would like to thank the mentors who helped in shepherding the preparation of lab proposals by newcomers:

Lorraine Goeriot, Université Grenoble Alpes, France Frank Hopfgartner, University of Sheffield, UK;

Jaap Kamps, University of Amsterdam, The Netherlands;

Josiane Mothe, IRIT, Université de Toulouse, France;

Henning Müller, University of Applied Sciences Western Switzerland (HES-SO), Switzerland.

We would like to thank the members of CLEF-LOC (the CLEF Lab Organization Committee) for their thoughtful and elaborate contributions to assessing the proposals during the selection process:

Marianna Apidianaki, University of Helsinki, Finland;

Martin Braschler, Zurich University of Applied Sciences, Switzerland;

Ingo Frommholz, University of Bedfordshire, UK;

Donna Harman, National Institute of Standards and Technology (NIST), USA;

Morgan Harvey, University of Sheffield, UK;

Jiyin He, Signal AI, UK

¹³ <https://events.webis.de/touche-20/>

¹⁴ <http://www.clef-initiative.eu/association>

Rezarta Islamaj Doğan, National Library of Medicine, USA;
Yue Ma, University Paris Sud, France;
Henning Müller, University of Applied Sciences Western Switzerland (HES-SO),
Switzerland;
Maarten de Rijke, University of Amsterdam, The Netherlands.

Last but not least, without the important and tireless effort of the enthusiastic and creative proposal authors, the organizers of the selected labs and workshops, the colleagues and friends involved in running them, and the participants who contribute their time to making the labs and workshops a success, the CLEF labs would not be possible.

Thank you all very much!

September, 2020

Linda Cappellato
Carsten Eickhoff
Nicola Ferro
Aurélie Névéol

Organization

CLEF 2020, *Conference and Labs of the Evaluation Forum – Experimental IR meets Multilinguality, 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, Univ. of Amsterdam, the Netherlands

Theodora Tsirikla, Information Technologies Institute, CERTH, Greece

Stefanos Vrochidis, Information Technologies Institute, CERTH, Greece

Avi Arampatzis, Democritus University of Thrace, Greece

Program Chairs

Hideo Joho, University of Tsukuba, Japan

Christina Lioma, University of Copenhagen, Denmark

Lab Chairs

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

Carsten Eickhoff, Brown University, USA

Lab Mentorship Chair

Lorraine Goeuriot, Université Grenoble Alpes, France

Proceedings Chairs

Linda Cappellato, University of Padua, Italy

Nicola Ferro, 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, United Kingdom

Fabio Crestani, Università della Svizzera italiana, Switzerland

Norbert Fuhr, University of Duisburg-Essen, Germany

Lorraine Goeriot, 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 UvA, The Netherlands

Alan Smeaton, Dublin City University, Ireland