

Preface

The CLEF 2018 conference is the nineteenth edition of the popular CLEF campaign and workshop series which 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 2018¹ was jointly organized by Avignon, Marseille, and Toulon Universities and was hosted by the University of Avignon, France, 10-14 September 2018.

Ten laboratories were selected and run during CLEF 2018. To identify the best proposals, besides the well-established criteria from previous years' editions 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.

Building on previous experience, the Labs at CLEF 2018 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.

CENTRE@CLEF 2018 -CLEF/NTCIR/TREC Reproducibility² is a joint CLEF/NTCIR/TREC task on challenging problems: 1) to reproduce best results of best/most interesting systems submitted to previous editions of CLEF/NTCIR/TREC by using standard open source IR systems; 2) to contribute back to the community with additional components and resources developed to reproduce the results and to improve the existing open source systems.

CheckThat!³ aims to foster the development of technology capable of both spotting and verifying check-worthy claims in political debates in English and Arabic.

Dynamic Search for Complex Tasks⁴ The lab strives to answer one key question: how can we evaluate, and consequently build, dynamic search algorithms? The 2018 Lab focuses on the development of an evaluation framework,

¹ <http://clef2018.clef-initiative.eu/>

² <http://www.centre-eval.org/clef2018/>

³ <http://alt.qcri.org/clef2018-factcheck/>

⁴ <https://ekanou.github.io/dynamicsearch/>

where participants submit “querying agents” that generate queries to be submitted to a static retrieval system. Effective “querying agents” can then simulate users towards developing dynamic search systems.

CLEFeHealth⁵ provides scenarios which aim to ease patients and nurses understanding and accessing eHealth information. 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. The tasks include: multilingual information extraction; technologically assisted reviews in empirical medicine; and, patient-centred information retrieval.

ImageCLEF⁶ organizes three main tasks and a pilot task: (i) a caption prediction task that aims at predicting the caption of a figure from the biomedical literature based only on the figure image; (ii) a tuberculosis task that aims at detecting the tuberculosis type, severity and drug resistance from CT (Computed Tomography) volumes of the lung; (iii) a lifelog task (videos, images and other sources) about daily activities understanding and moment retrieval, and (iv) a pilot task on visual question answering where systems are asked to answer medical questions.

LifeCLEF⁷ aims at boosting research on the identification of living organisms and on the production of biodiversity data in general. 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.

MC2⁸ mainly focuses on developing processing methods and resources to mine the social media (SM) sphere surrounding cultural events such as festivals, music, books, movies and museums. Following previous editions (CMC 2016 and MC2 2017), the 2018 edition focused on argumentative mining and multilingual cross SM search.

PAN⁹ 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 offers three tasks at CLEF 2018 with new evaluation resources consisting of large-scale corpora, performance measures, and web services that allow for meaningful evaluations. The main goal is to provide sustainable and reproducible evaluations so as to get a clear view of the capabilities of state-of-the-art algorithms. The tasks are: author identification; author profiling; and author obfuscation.

Early risk prediction on the Internet (eRisk)¹⁰ explores issues of evaluation methodology, effectiveness metrics and other processes related to early

⁵ <https://sites.google.com/view/clef-ehealth-2018/>

⁶ <http://www.imageclef.org/2018>

⁷ <http://www.lifeclef.org/>

⁸ <https://mc2.talne.eu/>

⁹ <http://pan.webis.de/>

¹⁰ <http://early.irlab.org/>

risk detection. Early detection technologies can be employed in different areas, particularly those related to health and safety. For instance, early alerts could be sent when a predator starts interacting with a child for sexual purposes, or when a potential offender starts publishing antisocial threats on a blog, forum or social network. Our main goal is to pioneer a new interdisciplinary research area that would be potentially applicable to a wide variety of situations and to many different personal profiles. eRisk 2018 had two campaign-style tasks: early detection of signs of depression and early detection of signs of anorexia.

Personalised Information Retrieval at CLEF (PIR-CLEF)¹¹ provides a framework for the evaluation of Personalised Information Retrieval (PIR). Current approaches to the evaluation of PIR are user-centric, mostly based on user studies, i.e., they rely on experiments that involve real users in a supervised environment. PIR-CLEF aims to develop and test a methodology for the evaluation of personalised search that enables repeatable experiments. The main aim is to enable research groups working on PIR to both experiment with and provide feedback on the proposed PIR evaluation methodology.

CLEF has been always backed by European projects which 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 in late 2013 of the CLEF Association¹², a non-profit legal entity, which, through the support of its members, ensures the resources needed to smoothly run and coordinate CLEF.

Acknowledgments

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:

Martin Braschler, Zurich University of Applied Sciences, Switzerland
Adrian Chifu, Aix-Marseille University, France
Sebastien Fournier, Aix-Marseille Université - CNRS LIS, France
Lorraine Goeuriot, Université Grenoble Alpes, France
Donna Harman, National Institute of Standards and Technology (NIST), USA
Wessel Kraaij, Leiden University, The Netherlands
Léa Laporte, INSA Lyon - LIRIS, France
Jacques Savoy, University of Neuchatel, Switzerland

¹¹ <http://www.ir.disco.unimib.it/pir-clef2018/>

¹² <http://www.clef-initiative.eu/association>

Lynda Said l'hadj, ESI, Alger, Algeria
Lynda Tamine, Paul Sabatier University, France

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!

July, 2018

Linda Cappellato
Nicola Ferro
Jian-Yun Nie
Laure Soulier

Organization

CLEF 2018, *Conference and Labs of the Evaluation Forum – Experimental IR meets Multilinguality, Multimodality, and Interaction*, was hosted by the University of Avignon and jointly co-organized by Avignon, Marseille and Toulon Universities, France.

General Chairs

Patrice Bellot, Aix-Marseille Université - CNRS LSIS, France
Chiraz Trabelsi, University of Tunis El Manar, Tunisia

Program Chairs

Josiane Mothe, SIG, IRIT, France
Fionn Murtagh, University of Huddersfield, United-Kingdom

Lab Chairs

Jian Yun Nie, DIRO, Université de Montréal, Canada
Laure Soulier, LIP6, UPMC, France

Proceedings Chairs

Linda Cappellato, University of Padua, Italy
Nicola Ferro, University of Padua, Italy

Publicity Chair

Adrian Chifu, Aix-Marseille Université - CNRS LSIS, France

Science Outreach Program Chairs

Aurelia Barriere, UAPV, France
Mathieu FERYN, UAPV, France

Sponsoring Chair

Malek Hajjem, UAPV, France

Local Organization

Eric SanJuan, LIA, UAPV, France – **chair**

Tania Jimenez, LIA, UAPV, France – **co-chair**

Sebastien Fournier, Aix-Marseille Université - CNRS LIS, France

Hervé Glotin, Université de Toulon - CNRS LIS, France

Vincent Labatut, LIA, UAPV, France

Elisabeth Muriasco, Université de Toulon - CNRS LIS, France

Magalie Ochs, Aix-Marseille Université - CNRS LIS, France

Juan-Manuel Torres-Moreno, LIA, UAPV, France

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

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

Séamus Lawless, Trinity College Dublin, Ireland

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

Maarten de Rijke, University of Amsterdam UvA, The Netherlands

Giuseppe Santucci, Sapienza University of Rome, Italy

Jacques Savoy, University of Neuchâtel, Switzerland

Christa Womser-Hacker, University of Hildesheim, Germany

Past Members

Jaana Kekäläinen, University of Tampere, Finland

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

Alan Smeaton, Dublin City University, Ireland