Preface

The CLEF 2016 conference is the seventeenth 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.

Seven laboratories and one workshop were selected and run during CLEF 2016. To identify the best proposals, besides 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.

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 as TREC, NTCIR, FIRE and MediaEval. Since 2014, the organisation of CLEF no longer has direct support from European projects and working to transform itself into a self-sustainable activity. This is being made possible thanks to the establishment in late 2013 of the CLEF Association[1], a non-profit legal entity, which, through the support of its members, ensures the resources needed to smoothly run and coordinate CLEF.

The Labs at CLEF 2016, building on previous experience, demonstrate the maturity of the CLEF evaluation environment via the incorporation of 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, here we just provide brief comment on each one.

CLEFeHealth[2] provides scenarios which aim to ease patients and nurses understanding and accessing of 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

---

[1] http://www.clef-initiative.eu/association
[2] https://sites.google.com/site/clefehealth2016/

tasks are: handover information extraction; multilingual Information extraction; and, patient-centred information retrieval.

ImageCLEF[3] organizes three main tasks with a global objective of benchmarking automatic annotation, indexing and retrieval of images. The tasks tackle different aspects of the annotation and retrieval problem and are aimed at supporting and promoting cutting-edge research addressing the key challenges in the field. A wide range of source images and objectives are considered, such as general multi-domain images for object or concept detection, as well as domain-specific tasks such as labelling and separation of compound figures from biomedical literature and scanned pages from historical documents.

LifeCLEF[4] proposes three data-oriented challenges related to this vision, in the continuity of the two previous editions of the lab, but with several consistent novelties intended to push the boundaries of the state-of-the-art in several research directions at the frontier of information retrieval, machine learning and knowledge engineering including: an audio record-based bird identification task (BirdCLEF); an image-based plant identification task (PlantCLEF); and, a fish video surveillance task (FishCLEF).

Living Labs for IR (LL4IR)[5] inprovides a benchmarking platform for researchers to evaluate their ranking systems in a live setting with real users in their natural task environments. The lab acts as a proxy between commercial organizations (live environments) and lab participants (experimental systems), facilitates data exchange, and makes comparison between the participating systems. The task focuses on on-line product search.

News Recommendation Evaluation Lab (NEWSREEL)[6] provides two tasks designed to address the challenge of real-time news recommendation. Participants can: a) develop news recommendation algorithms and b) have them tested by millions of users over the period of a few weeks in a living lab. The tasks are: benchmark news recommendations in a living lab; benchmarking news recommendations in a simulated environment.

Uncovering Plagiarism, Authorship and Social Software Misuse (PAN)[7] provides evaluation of uncovering plagiarism, authorship, and social software misuse. PAN offers three tasks at CLEF 2016 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 for sustainable and reproducible evaluations, to get a clear view of the capabilities of state-of-the-art-algorithms. The tasks are: author identification; author profiling; and, author obfuscation.

Social Book Search (SBS)[8] provides evaluation of real-world information needs which are generally complex, yet almost all research focuses instead on

---

[3] http://www.imageclef.org/2016

[4] http://www.imageclef.org/node/197

[5] http://living-labs.net/clef-ll4ir-2016/

[6] http://www.clef-newsreel.org/

[7] http://pan.webis.de/

[8] http://social-book-search.humanities.uva.nl/

either relatively simple search based on queries or recommendation based on profiles. The goal of the Social Book Search Lab is to investigate techniques to support users in complex book search tasks that involve more than just a query and results list. The tasks are: a user-oriented interactive task investigating systems that support users in each of multiple stages of a complex search tasks; a system-oriented task for systems to suggest books based on rich search requests combining several topical and contextual relevance signals, as well as user profiles and real-world relevance judgements; and, an NLP/Text Mining track focusing on detecting and linking book titles in online book discussion forums, as well as detecting book search research in forum posts for automatic book recommendation.

Cultural Microblog Contextualization (CMC) Workshop[9] aims at developing processing methods for social media mining. The focus is around festivals that are organized or that have a large presence on social media. For its first edition, this workShop gives access to a massive collection of microblogs and urls and allows researchers in IR and NLP to experiment a broad variety of multilingual microblog search techniques (WikiPedia entity search, automatic summarization, and more).

Acknowledgments

July, 2016

Krisztian Balog
Linda Cappellato
Nicola Ferro
Craig Macdonald

---

[9] `https://mc2.talne.eu/~cmc/spip/`