CLEF 2012
Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics

Tiziana Catarci
Sapienza University of Rome, Italy
catarci@dis.uniroma1.it

Nicola Ferro
University of Padua, Italy
ferro@dei.unipd.it

Pamela Forner
CELCT, Italy
forner@celct.it

Djoerd Hiemstra
University of Twente, The Netherlands
jhiemstra@cs.utwente.nl

Jussi Karlgren
Gavagai, Sweden
jussi@gavagai.se

Anselmo Peñas
UNED, Spain
anselmo@lsi.uned.es

Giuseppe Santucci
Sapienza University of Rome, Italy
santucci@dis.uniroma1.it

Christa Womser-Hacker
University of Hildesheim, Germany
womser@uni-hildesheim.de

1 Introduction

The CLEF 2012 Conference on Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics was held at the "Sapienza" University of Rome, Italy, September 17–20, 2011. CLEF 2012 was organized by the Department of Computer, Control, and Management Enigneering "Antonio Ruberti".

Since 2000 the Cross-Language Evaluation Forum (CLEF) has played a leading role in stimulating research and innovation in a wide range of key areas in the domain of information retrieval. It has become a landmark in the annual research calendar of the international Information Retrieval and Search community. Through the years, CLEF has promoted the study and implementation of evaluation methodologies for diverse types of retrieval task and search scenario. As a result, a broad, strong and multidisciplinary research community has been created, which covers and spans the different areas of expertise needed to deal with the evaluation of solutions to challenging information tasks.

Until 2010, the outcomes of experiments carried out under the CLEF umbrella were presented and discussed at annual workshops in conjunction with the European Conference for Digital Libraries. CLEF 2010 represented a radical departure from this "classic" CLEF format. While preserving CLEF's traditional core business and goals, namely benchmarking activities carried in various tracks, we complemented these activities with a peer-reviewed conference component aimed at advancing research in the evaluation of complex information systems for cross-language tasks and scenarios. CLEF 2010 was thus organised as an independent four-day event consisting of two main parts, a peer-reviewed conference followed by a series of laboratories and workshops.

CLEF 2012 continued to implement this new format, with keynotes, contributed papers and lab sessions, but we added a small number of refinements. As in CLEF 2011, we interleaved the conference presentations and the laboratories over a three and a half day period.

Overall, CLEF 2012 was attended by more than 190 people from different academic and industrial institutions; this is encouraging as the participation increased with respect to last year. Although the majority of participants came from Europe, we note more interest in CLEF from all around the world, above all from the United States with 24 participants, then Asia with 23 participants, Australia with 3 participants and Africa with 2 participants.

2 CLEF 2012: The Conference

The conference aimed at advancing research into multilingual and multimodal information access evaluation and deployment.

On the whole 33 papers were received. This number is a great improvement with respect to previous years: in 2011, 23 papers were submitted and almost the same number in 2010. On average, 4 papers were assigned to each reviewer, and each paper received 3 evaluation reports. In total, 17 papers (9 full, 5 short, and 3 posters) were finally accepted with an overall acceptance rate of 51.52%.

This year, the papers accepted for the conference included research on information access, evaluation initiatives, and methodologies and infrastructures.

Two keynote speakers highlighted important developments in the field of evaluation.

Peter Clark (Vulcan Inc., USA) argued for the importance of creating a deeper understanding of textual material, and ultimately more knowledgeable machines. In his talk "From Information Retrieval to Knowledgeable Machines", he claimed that our machines should not only search and retrieve information, but also have some "understanding" of the material that they are manipulating so that they can better meet the user's needs. He presented a case of innovation turned into a company product that allows users to not only read and browse a textbook, but also to ask questions and get reasoned or retrieved answers back, explore the material through semantic connections, and receive suggestions of useful questions to ask.

Tobias Schreck (University of Konstanz, Germany) showed current approaches, applications and challenges for the application of visual analytics in document repositories. In his talk "Visual Search and Analysis in Textual and Non-Textual Document Repositories – Approaches, Applications, and Research Challenge" he claimed that advances in visual-interactive data analysis can provide for effective visual interfaces for query formulation, navigation, and result exploration in complex information spaces. He discussed selected approaches for visual analysis in large textual and non-textual document collections. He argued that new visual-interactive approaches can

provide for effective user access to large document corpora, including discovering of interesting relationships between data items, and understanding the space of similarity notions for a given document repository.

CLEF 2012 hosted two Community Sessions: (i) "Evaluation CLINIC" where people with evaluation retrieval problem were able to talk to evaluation experts and get methodological advice, new ideas, pointers to related problems, available solutions, and so on; (ii) "Other Evaluation Initiatives": the goal of the session was to present the current focal points and lessons learned in the other experimental evaluation settings around the world and to promote discussion with respect to what the future objectives in this field should be.

3   CLEF 2012: The Lab Sessions

The CLEF Labs are a continuation of tracks from previous CLEF workshops. In 2010, CLEF went from being a workshop collocated with an existing conference to a conference in its own right. The CLEF Labs are an integral part of the conference and two different types of lab are offered: (1) benchmarking or "campaign-style" and (2) workshop-style. The benchmarking labs on the whole follow the traditional ("campaign-style") cycle of activities in a large- scale information retrieval evaluation experiment set-up. Workshop-style labs are offered as a way of exploring possible benchmarking activities and to provide a means to discuss information retrieval evaluation issues from different perspectives.

A call for lab proposals was distributed in October 2011 and a lab selection committee formed from leading IR researchers (see Acknowledgements). Lab proposals were requested to include a detailed description of the topics and goals of the lab, the target audience, potential opportunities for future versions of the lab, as well as details about the planning and organisation of the lab (e.g. suggested tasks, the data collections used in the lab, the background of lab organisers and steering committee members, etc.).

In previous years, lab workshops have resulted in a proposal for an evaluation lab for the following year. This was again the case this year: the CHiC lab workshop from 2011 submitted a proposal for an evaluation laboratory in 2012. This progression from a lab workshop to an evaluation lab is a development track the CLEF Lab Organisation Committee wishes to encourage to introduce the organisers of an evaluation campaign to discuss practicalities and the make-up of an evaluation task. However, we do not expect that lab workshops are limited to planning future campaigns: their scope may well be more abstract, more far-reaching or more specific.

Eight lab proposals were accepted for CLEF 2012: 7 benchmarking labs and 1 workshop. By August 2011, 96 research groups had submitted experimental results in a benchmarking activity and 107 participants registered to attend one of the lab sessions at CLEF.

The following benchmarking labs ran in CLEF 2011:

CHiC Cultural Heritage in CLEF is a pilot evaluation lab which aimed at moving towards a systematic and large-scale evaluation of cultural heritage digital libraries and information access systems. Data test collections and queries come from the cultural heritage domain (in 2012 data from Europeana) and tasks contained a mix of conventional system-oriented evaluation scenarios (e.g. ad-hoc retrieval and semantic enrichment) for comparison with other domains and a uniquely customized scenario for the CH domain, i.e., a variability task

to present a particular good overview over the various object types and categories in the collection targeted towards a casual user.

CLEF-IP Information Retrieval in the Intellectual Property Domain provided a large collection of XML documents representing patents and patent images. On this collection the following four tasks were organized: chemical structure recognition task; flowchart Recognition task; passage retrieval starting from claims; matching claim to description in a single document.

ImageCLEF Cross Language Image Retrieval Track evaluates the cross-language annotation and retrieval of images by focusing on the combination of textual and visual evidence. Five challenging tasks are offered: medical task; photo annotation and retrieval; personal photo annotation; plant identification: robot vision.

INEX Initiative for the Evaluation of XML Retrieval has been a stand-alone initiative pioneering structured retrieval since 2002, and joined forces with CLEF running five tasks: social book search; linked data; snippet retrieval; relevance feedback; tweet Contextualization.

PAN a benchmarking activity aimed at uncovering plagiarism, authorship and social software misuse offered three tasks: plagiarism detection; author identification; quality flaw prediction in Wikipedia.

QA4MRE Question Answering for Machine Reading evaluates machine reading abilities through question answering and reading comprehension tests and offered three tasks: reading of single documents and identification of the answers to a set of questions about information that is stated or implied in the text; evaluation whether systems are able to understand extra-propositional aspects of meaning like modality and negation; setting questions in the biomedical domain with a special focus on the Alzheimer disease.

RebLab Online Reputation Management dealt with the image that online media project about individuals and organizations. The aim is to bring together the Information Access research community with representatives from the Online Reputation Management industry, with the goals of (i) establishing a five-year roadmap that includes a description of the language technologies required in terms of resources, algorithms, and applications; (ii) specifying suitable evaluation methodologies and metrics; and (iii) developing of test collections that enable systematic comparison of algorithms and reliable benchmarking of commercial systems.

The following workshop-style lab was also held at CLEF 2012:

CLEFeHealth focused on cross-language methods, applications, and resources for eHealth document analysis a with an emphasis on written and spoken NLP.

4   CLEF 2012 and Beyond

During CLEF 2012 a big effort has been made to further reshape and improve the overall organization of CLEF. In order to allow for more time for planning and organizing the activities the bid for both CLEF 2013 and CLEF 2014 has been carried out and so, it is now possible, to know in advance where the two next CLEFs will happen.

CLEF 2013 will be hosted by the Technical University of Valencia, Spain, 23-26 September 2013 while CLEF 2014 will be hosted by the University of Sheffield, United Kingdom, 15-19 September 2014.

Moreover, in order to have a single publication which reports the whole breadth of CLEF activities, it has been decided the the annual Springer LNCS proceedings will report not only the conference papers but also revised overviews from the labs carried out during the year. In order to be able to deliver this publication at the annual CLEF event in September, the yearly cycle for the evaluation campaign has been rescheduled to run from mid November to early May. Therefore, the Call for Labs proposal for CLEF 2013 will be closing in mid October 2012 and the one for CLEF 2014 will be further anticipated, closing in early July 2013, in order to be able to announce the labs for the next year during the annual CLEF event.

The Call for papers for the CLEF Conference 2013 will be released in late November 2012 and the expected deadline for the submission of papers is late April 2012.

Finally, the bids for hosting CLEF 2015 are open and will close on 5th April 2013. A proposal for hosting CLEF 2015 can be sent to the CLEF Steering Committee Chair at chair@clef-initiative.eu.

Acknowledgments

References

[1] T. Catarci, P. Forner, D. Hiemstra, A. Peñas, and G. Santucci, editors. *Information Access Evaluation. Multilinguality, Multimodality, and Visual Analytics. Proceedings of the Third International Conference of the CLEF Initiative (CLEF 2012).* Lecture Notes in Computer Science (LNCS) 7488, Springer, Heidelberg, Germany, 2012.

[2] P. Forner, J. Karlgren, and C. Womser-Hacker, editors. *CLEF 2012 Labs and Workshops, Notebook Papers.* MINT srl, Trento, Italy. ISBN 978-88-904810-3-1., 2012.