Preface

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 organized as an independent four-day event consisting of two main parts, a peer-reviewed conference followed by a series of laboratories and workshops.

CLEF 2011 continued to implement this new format, with keynotes, contributed papers and lab sessions, but with a few additional refinements. First, in this year’s event we interleaved the conference presentations and the laboratories over a three-and-a-half-day period. Second, we added a “community session” aimed at creating awareness of funding opportunities, offering networking opportunities, and demonstrating emerging infrastructures that help support cross-language retrieval experiments.

This year, the papers accepted for the conference included research on evaluation methods and settings, natural language processing within different domains and languages, multimedia and reflections on CLEF. Two keynote speakers highlighted important developments in the field of evaluation. Elaine Toms (University of Sheffield) focused on the role of users in evaluation, whereas Omar Alonso (Microsoft Corporation) presented a framework for the use of crowdsourcing experiments in the setting of retrieval evaluation.

CLEF 2011 featured six benchmarking activities: ImageCLEF, PAN, CLEF-IP, QA4MRE, LogCLEF and, new this year, MusiCLEF. In parallel, there was a workshop dedicated to the evaluation of information access technologies in the setting of cultural heritage. All the experiments carried out by systems during the evaluation campaigns are described in a separate publication, namely, the Working Notes, distributed during CLEF 2011 and available on-line.
The community session at CLEF 2011 was organized around a strategic EU meeting, a networking session organized by the Chorus Network of Excellence, an Evaluation Initiatives session with overviews from other benchmarking fora, and an infrastructure session dedicated to the DIRECT system for handling data resulting from retrieval experiments.

The success of CLEF 2011 would not have been possible without the invaluable contributions of all the members of the Program Committee, Organizing Committee, students and volunteers that supported the conference in its various stages. Thank you all!

Finally, we would like to express our gratitude to the sponsoring organizations for their significant and timely support.

These proceedings were prepared with the assistance of the Center for the Evaluation of Language and Communication Technologies (CELCT), Trento, Italy.

July 2011

Pamela Forner
Julio Gonzalo
Jaana Kekäläinen
Mounia Lalmas
Maarten de Rijke
CLEF 2011 was organized by the Intelligent Systems Lab Amsterdam (ISLA) of the University of Amsterdam.

**General Chairs**

Julio Gonzalo  
National University of Distance Education (UNED), Spain

Maarten de Rijke  
University of Amsterdam, The Netherlands

**Program Chairs**

Jaana Kekäläinen  
University of Tampere, Finland

Mounia Lalmas  
Yahoo! Research Barcelona, Spain

**Labs Chairs**

Paul Clough  
University of Sheffield, UK

Vivien Petras  
Humboldt University of Berlin, Germany

**Resource and Publicity Chair**

Khalid Choukri  
Evaluations and Language resources Distribution Agency (ELDA), France

**Organization Chair**

Emanuele Pianta  
Center for the Evaluation of Language and Communication Technologies (CELCT), Italy

**Organizing Committee**

*Intelligent Systems Lab Amsterdam (ISLA), University of Amsterdam, The Netherlands:*

Richard Berendsen
Bouke Huurnink
Edgar Meij
Maarten de Rijke
Wouter Weerkamp

*Center for the Evaluation of Language and Communication Technologies (CELCT), Italy:*

Pamela Forner
Giovanni Moretti
Program Committee

Pia Borlund Royal School of Library and Information Science, Denmark
Pavel Braslavski Yandex, Russia
Ben Carterette University of Delaware, USA
Nicola Ferro University of Padua, Italy
Norbert Fuhr University of Duisburg, Germany
Kal Järvelin University of Tampere, Finland
Gareth Jones Dublin City University, Ireland
Marko Junkkari University of Tampere, Finland
Noriko Kando National Institute of Informatics, Japan
Heikki Keskustalo University of Tampere, Finland
Birger Larsen Royal School of Library and Information Science, Denmark
Prasenjit Majumder Dhirubhai Ambani Institute of Information and Communication Technology, India
Paul McNamee Johns Hopkins University, USA
Stefano Mizzarro University of Udine, Italy
 Jian-Yun Nie University of Montreal, Canada
Douglas Oard University of Maryland, USA
Carol Peters ISTI CNR Pisa, Italy
Vivien Petras Humboldt University, Germany
Ian Ruthven University of Strathclyde, UK
Tetsuya Sakai Microsoft Research Asia
Christa Womser-Hacker University of Hildesheim, Germany

Sponsoring Institutions

CLEF 2011 benefited from the support of the following organizations:

Center for Creation, Content and Technology, University of Amsterdam, The Netherlands
City of Amsterdam, The Netherlands
ELIAS Research Network Programme
European Language Resources Association, France
Microsoft Research
Netherlands Institute for Sound and Vision, Hilversum, The Netherlands
Netherlands Organization for Scientific Research, The Netherlands
PROMISE Network of Excellence
University of Amsterdam, The Netherlands
Xerox Research Centre Europe, France
# Table of Contents

## Keynote Addresses

“Would You Trust Your IR System to Choose Your Date?” Re-thinking IR Evaluation in the 21st Century ........................................... 1
*Elaine G. Toms*

Crowdsourcing for Information Retrieval Experimentation and Evaluation ............................................. 2
*Omar Alonso*

## Methodologies and Lessons

Building a Cross-Language Entity Linking Collection in Twenty-One Languages ........................................ 3
*James Mayfield, Dawn Lawrie, Paul McNamee, and Douglas W. Oard*

Search Snippet Evaluation at Yandex: Lessons Learned and Future Directions ............................................. 14
*Denis Savenkov, Pavel Braslavski, and Mikhail Lebedev*

Towards a Living Lab for Information Retrieval Research and Development: A Proposal for a Living Lab for Product Search Tasks ... 26
*Leif Azzopardi and Krisztian Balog*

A Comparison of Evaluation Metrics for Document Filtering .......... 38
*Enrique Amigó, Julio Gonzalo, and Felisa Verdejo*

## Language and Processing

Filter Keywords and Majority Class Strategies for Company Name Disambiguation in Twitter ............................................. 50
*Damiano Spina, Enrique Amigó, and Julio Gonzalo*

Automatic Annotation of Bibliographical References for Descriptive Language Materials ............................................. 62
*Harald Hammarström*

A Language-Independent Approach to Identify the Named Entities in Under-Resourced Languages and Clustering Multilingual Documents ... 74
*N. Kiran Kumar, G.S.K. Santosh, and Vasudeva Varma*
Multilingual Question-Answering System in Biomedical Domain on the Web: An Evaluation .............................................. 83
  María-Dolores Olvera-Lobo and Juncal Gutiérrez-Artacho

Simulation of Within-Session Query Variations Using a Text Segmentation Approach ........................................... 89
  Debasis Ganguly, Johannes Leveling, and Gareth J.F. Jones

Visual and Context

Assessing the Scholarly Impact of ImageCLEF ......................... 95
  Theodora Tsikrika, Alba G. Seco de Herrera, and Henning Müller

The Importance of Visual Context Clues in Multimedia Translation .... 107
  Christopher G. Harris and Tao Xu

To Re-rank or to Re-query: Can Visual Analytics Solve This Dilemma? ................................................................. 119
  Emanuele Di Buccio, Marco Dussin, Nicola Ferro, Ivano Masiero, Giuseppe Santucci, and Giuseppe Tino

Evaluation Methods for Rankings of Facetvalues for Faceted Search .... 131
  Anne Schuth and Maarten Marx

Improving Query Expansion for Image Retrieval via Saliency and Picturability ...................................................... 137
  Chee Wee Leong, Samer Hassan, Miguel Enrique Ruiz, and Rada Mihalcea

Author Index ...................................................................... 143