What Happened in CLEF 2009

Carol Peters

Istituto di Scienza e Tecnologie dell'Informazione (ISTI-CNR), Pisa, Italy
carol.peters@isti.cnr.it

Abstract. The organization of the CLEF 2009 evaluation campaign is described and details are provided concerning the tracks, test collections, evaluation infrastructure, and participation. The aim is to provide the reader of these proceedings with a complete picture of the entire campaign, covering both text and multimedia retrieval experiments. In the final section, the main results achieved by CLEF in the first ten years of activity are discussed and plans for the future of CLEF are presented.

1 Introduction

The objective of the Cross Language Evaluation Forum is to promote research in the field of multilingual system development. This is done through the organisation of annual evaluation campaigns in which a series of tracks designed to test different aspects of mono- and cross-language information retrieval (IR) are offered. The intention is to encourage experimentation with all kinds of multilingual information access – from the development of systems for monolingual retrieval operating on many languages to the implementation of complete multilingual multimedia search services. This has been achieved by offering an increasingly complex and varied set of evaluation tasks over the years. The aim is to meet and anticipate the needs of the multidisciplinary research community working in this area and to encourage the development of next generation multilingual IR systems. CLEF is perhaps one of the few platforms where groups working in many different areas (e.g. Information Retrieval, Natural Language Processing, Image Processing, Speech Recognition, Log Analysis, etc.) have a chance to see what others are doing, and discuss and compare ideas. Figure 1 shows the evolution of CLEF in ten years of activity.

This is the first time that the CLEF post-campaign proceedings have been published in two separate volumes. This decision has been made necessary by the large participation in CLEF 2009 and our desire to provide an exhaustive overview of all the various evaluation activities. We have thus distinguished between papers describing systems and functionality for text retrieval and for multimedia retrieval. This volume reports experiments on various types of textual document collections. It is divided into six main sections presenting the results of the following tracks: Multilingual Document Retrieval (Ad-Hoc), Multiple Language Question Answering (QA@CLEF), Multilingual Information Filtering (INFILE@CLEF), Intellectual Property (CLEF-IP) and Log File Analysis (LogCLEF), plus the activities of the MorphoChallenge program. The papers are mostly extended and revised versions of the initial working notes

C. Peters et al. (Eds.): CLEF 2009 Workshop, Part I, LNCS 6241, pp. 1–12, 2010.

distributed at the workshop. For details on the results of the tracks conducting experiments on multimedia data: Interactive Cross-Language Retrieval (iCLEF), Cross-Language Image Retrieval (ImageCLEF), and Cross-Language Video Retrieval (VideoCLEF), the reader is referred to the companion volume[1].

This Introduction gives a brief overview of entire campaign in order to provide the reader with a complete picture of what happened: Section 2 lists the various tracks and tasks offered in 2009; Sections 3 and 4 describe the participation and the evaluation infrastructure; the final section gives an assessment of the results achieved by CLEF in this first ten years of activity and presents plans for the future.

2 Tracks and Tasks in CLEF 2009

CLEF 2009 offered eight tracks designed to evaluate the performance of systems for:

- multilingual textual document retrieval (Ad Hoc)
- interactive cross-language retrieval (iCLEF)
- multiple language question answering (QA@CLEF)
- cross-language retrieval in image collections (ImageCLEF)
- multilingual information filtering (INFILE@CLEF)
- cross-language video retrieval (VideoCLEF)
- intellectual property (CLEF-IP) – New this year
- log file analysis (LogCLEF) – New this year

CLEF 2000	▪ mono-, bi- & multilingual text doc retrieval (Ad Hoc) ▪ mono- and cross-language information on structured scientific data (Domain-Specific)
CLEF 2001 New	▪ interactive cross-language retrieval (iCLEF)
CLEF 2002 New	▪ cross-language spoken document retrieval (CL-SR)
CLEF 2003 New	▪ multiple language question answering (QA@CLEF) ▪ cross-language retrieval in image collections (ImageCLEF)
CLEF 2005 New	▪ multilingual retrieval of Web documents (WebCLEF) ▪ cross-language geographical retrieval (GeoCLEF)
CLEF 2008 New	▪ cross-language video retrieval (VideoCLEF) ▪ multilingual information filtering (INFILE@CLEF)
CLEF 2009 New	▪ intellectual property (CLEF-IP) ▪ log file analysis (LogCLEF)

Fig. 1. Evolution of CLEF Tracks

[1] Multilingual Information Access Evaluation II: Multimedia Experiments, LNCS Vol. 6242, Springer.

An experimental pilot task was also offered:

- Grid Experiments (Grid@CLEF)

In addition, Morpho Challenge 2009 was organized in collaboration with CLEF as part of the EU Network of Excellence Pascal Challenge Program[2].

Here below we give a brief overview of the various activities.

Multilingual Textual Document Retrieval (Ad Hoc): The aim of this track has been to promote the development of monolingual and cross-language textual document retrieval systems. From 2000 - 2007, the track used collections of European newspaper and news agency documents. In CLEF 2008, the focus of the track was considerably widened: we introduced very different document collections, a non-European target language, and an information retrieval (IR) task designed to attract participation from groups interested in natural language processing (NLP). Ad Hoc 2009 was to a large extent a repetition of the previous year's activities, with the same three tasks: Tel@CLEF, Persian@CLEF, and Robust-WSD. An important objective was to create good reusable test collections for each of them The track was thus structured in three distinct streams. The first task offered monolingual and cross-language search on library catalog records and was organized in collaboration with The European Library (TEL)[3]. The second task resembled the ad hoc retrieval tasks of previous years but this time the target collection was a Persian newspaper corpora. The third task was the robust activity which used word sense disambiguated (WSD) data. The track was coordinated jointly by ISTI-CNR and Padua University, Italy; the University of the Basque Country, Spain; with the collaboration of the Database Research Group, University of Tehran, Iran.

Interactive Cross-Language Retrieval (iCLEF): In iCLEF, cross-language search capabilities have been studied from a user-inclusive perspective. A central research question has been how best to assist users when searching information written in unknown languages, rather than how best an algorithm can find information written in languages different from the query language. Since 2006, iCLEF has based its experiments on Flickr, a large-scale, web-based image database where image annotations constitute a naturally multilingual folksonomy. In an attempt to encourage greater participation in user-orientated experiments, a new task was designed for 2008 and continued in 2009. The main novelty has been to focus experiments on a shared analysis of a large search log, generated by iCLEF participants from a single search interface provided by the iCLEF organizers. The focus has been, therefore, on search log analysis rather than on system design. The idea has been to study the behaviour of users in an (almost) naturalistic search scenario, having a much larger data set than in previous iCLEF campaigns. The track was coordinated by UNED, Madrid, Spain; Sheffield University, UK; Swedish Institute of Computer Science, Sweden.

[2] MorphoChallenge is part of the EU Network of Excellence Pascal: `http://www.cis.hut.fi/morphochallenge2009/`

[3] See `http://www.theeuropeanlibrary.org/`

Multilingual Question Answering (QA@CLEF): This track has offered monolingual and cross-language question answering tasks since 2003. QA@CLEF 2009 proposed three exercises: ResPubliQA, QAST and GikiCLEF:

- ResPubliQA: The hypothetical user considered for this exercise is a person close to the law domain interested in making inquiries on European legislation. Given a pool of 500 independent natural language questions, systems must return the passage that answers each question (not the exact answer) from the JRC-Acquis collection of EU parliamentary documentation. Both questions and documents are translated and aligned for a subset of languages. Participating systems could perform the task in Basque, Bulgarian, English, French, German, Italian, Portuguese, Romanian and Spanish.
- QAST: The aim of the third QAST exercise was to evaluate QA technology in a real multilingual speech scenario in which written and oral questions (factual and definitional) in different languages are formulated against a set of manually and automatically transcribed audio recordings related to speech events in those languages. The scenario proposed was the European Parliament sessions in English, Spanish and French.
- GikiCLEF: Following the previous GikiP pilot at GeoCLEF 2008, the task focused on open list questions over Wikipedia that require geographic reasoning, complex information extraction, and cross-lingual processing, for collections in Bulgarian, Dutch, English, German, Italian, Norwegian (both Bokmål and Nynorsk), Portuguese and Romanian or Spanish.

The track was organized by a number of institutions (one for each target language), and jointly coordinated by CELCT, Trento, Italy, and UNED, Madrid, Spain.

Cross-Language Retrieval in Image Collections (ImageCLEF): This track evaluated retrieval from visual collections; both text and visual retrieval techniques were employed. A number of challenging tasks were offered:

- multilingual ad-hoc retrieval from a photo collection concentrating on diversity in the results;
- a photographic annotation task using a simple ontology;
- retrieval from a large scale, heterogeneous collection of Wikipedia images with user-generated textual metadata;
- medical image retrieval (with visual, semantic and mixed topics in several languages);
- medical image annotation form two databases, a database of chest CTs to detect nodules and a database of x-ray images;
- detection of semantic categories from robotic images (non-annotated collection, concepts to be detected).

A large number of organisations have been involved in the complex coordination of these tasks. They include: Sheffield University, UK; University of Applied Sciences Western Switzerland; Oregon Health and Science University, USA; University of Geneva, Switzerland; CWI, The Netherlands; IDIAP, Switzerland; University of Geneva, Switzerland; Fraunhofer Gesellschaft, Germany; Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands.

Multilingual Information Filtering (INFILE@CLEF): INFILE (INformation, FILtering & Evaluation) was a cross-language adaptive filtering evaluation track sponsored by the French National Research Agency. INFILE has extended the last filtering track of TREC 2002 in a multilingual context. It used a corpus of 100,000 Agence France Press comparable newswires for Arabic, English and French; and evaluation was performed using an automatic querying of test systems with a simulated user feedback. Each system can use the feedback at any time to increase performance. The track was coordinated by the Evaluation and Language resources Distribution Agency (ELDA), France; University of Lille, France; and CEA LIST, France.

Cross-Language Video Retrieval (VideoCLEF): VideoCLEF 2009 was dedicated to developing and evaluating tasks involving access to video content in a multilingual environment. Participants were provided with a corpus of video data (Dutch-language television, predominantly documentaries) accompanied by speech recognition transcripts. In 2009, there were three tasks: "Subject Classification", which involved automatically tagging videos with subject labels; "Affect", which involved classifying videos according to characteristics beyond their semantic content; ``Finding Related Resources Across Languages", which involved linking video to material on the same subject in a different language. The track was jointly coordinated by Delft University of Technology, The Netherlands, and Dublin City University, Ireland.

Intellectual Property (CLEF-IP): This was the first year for the CLEF-IP track. The purpose of the track was twofold: to encourage and facilitate research in the area of patent retrieval by providing a large clean data set for experimentation; to create a large test collection of patents in three main European languages for the evaluation of cross-language information access. The track focused on the task of prior art search. A large test collection for evaluation purposes was created by exploiting patent citations. The collection consists of a corpus of 1,9 million patent documents and 10,000 topics with an average of 6 relevance assessments per topic.

Log File Analysis (LogCLEF): LogCLEF was an evaluation initiative for the analysis of queries and other logged activities as expression of user behaviour. The goal was the analysis and classification of queries in order to understand search behaviour in multilingual contexts and ultimately to improve search systems. The track used log data from the files of The European Library.

Grid Experiments (Grid@CLEF): This experimental pilot has been planned as a long term activity with the aim of: looking at differences across a wide set of languages; identifying best practices for each language; helping other countries to develop their expertise in the IR field and create IR groups. Participants had to conduct experiments according to the CIRCO (Coordinated Information Retrieval Components Orchestration) protocol, an XML-based framework which allows for a distributed, loosely-coupled, and asynchronous experimental evaluation of Information Retrieval (IR) systems. The track was coordinated jointly by University of Padua, Italy, and the National Institute of Standards and Technology, USA.

Unsupervised Morpheme Analysis (Morpho Challenge): Morpheme analysis is particularly useful in speech recognition, information retrieval and machine translation for morphologically rich languages where the amount of different word forms is very large. In Morpho Challenge 2009 unsupervised algorithms that provide morpheme analyses for words in different languages were evaluated in various practical applications. The evaluations consisted of: 1) a comparison to grammatical morphemes, 2) using morphemes instead of words in information retrieval tasks, and 3) combining morpheme and word based systems in statistical machine translation tasks. The evaluation languages in 2009 were: Finnish, Turkish, German, English and Arabic. The track was coordinated by Helsinki University of Technology and Cambridge University Engineering Department.

Details on the technical infrastructure and the organisation of all these tracks can be found in the track overview reports in this volume, collocated at the beginning of the relevant sections.

3 Test Collections

The CLEF test collections are made up of documents, topics and relevance assessments. The topics are created to simulate particular information needs from which the systems derive the queries to search the document collections. System performance is evaluated by judging the results retrieved in response to a topic with respect to their relevance, and computing the relevant measures, depending on the methodology adopted by the track. The document sets that have been used to build the test collections in CLEF 2009 included:

- A subset of the CLEF multilingual corpus of news documents in 14 European languages (Ad Hoc WSD-Robust task, MorphoChallenge)
- Hamshahri Persian newspaper corpus (Ad Hoc Persian task)
- Library catalog records in English, French, German plus log files provided by The European Library (Ad Hoc TEL task and LogCLEF)
- Log files from the Tumba search engine: http://www.tumba.pt/ (LogCLEF)
- Flickr web-based image database (iCLEF)
- ResPubliQA document collection, a subset of the JRC Acquis corpus of European legislation (QAatCLEF: ResPubliQA)
- Transcripts of European parliamentary sessions in English and Spanish, and French news broadcasts (QAatCLEF: QAST)
- BELGAPICTURE image collection (ImageCLEFPhoto)
- A collection of Wikipedia images and their user-generated textual metadata (ImageCLEFwiki)
- Articles and images from the Radiology and Radiography journals of the RSNA (Radiological Society of North America) (ImageCLEFmed); IRMA collection for medical image annotation (ImageCLEFmedAnnotation); a collection from the Lung Image Database Consortium (LIDC) (ImageCLEFmedAnnotation)
- A collection of FlickR images (ImageCLEFanno)
- A collection of robotics images created from KTH, Sweden (ImageCLEFrobot Vision)
- Dutch and English documentary television programs (VideoCLEF)
- Agence France Press (AFP) comparable newswire stories in Arabic, French and English (INFILE)

- Patent documents in English, French and German from the European Patent Office (CLEF-IP)
- Acknowledgements of the valuable contribution of the data providers is given at the end of this paper.

4 CLEF and TrebleCLEF

CLEF is organized mainly through the voluntary efforts of many different institutions and research groups. However, the central coordination has always received some support from the EU IST programme under the unit for Digital Libraries and Technology Enhanced Learning, mainly within the framework of the DELOS Network of Excellence. CLEF 2008 and 2009 were organized under the auspices of TrebleCLEF, a Coordination Action of the Seventh Framework Programme.

TrebleCLEF has built on the results achieved by CLEF, supporting the development of expertise in the multidisciplinary research area of multilingual information access and promoting dissemination actions in the relevant application communities. The aim has been to:

- Provide applications that need multilingual search solutions with the possibility to identify the technology which is most appropriate
- Assist technology providers to develop competitive multilingual search solutions.

In 2009, the TrebleCLEF activities included the organization of a Summer School on Multilingual Information Access (MLIA) and a MLIA Technology Transfer Day, and the publication of three Best Practices studies:

- Best Practices in Language Resources for Multilingual Information Access
- Best Practices in System and User-oriented Multilingual Information Access
- Best Practices for Test Collection Creation, Evaluation Methodologies and Language Processing Technologies

Information on the activities of TrebleCLEF can be found on the project website[4].

5 Technical Infrastructure

TrebleCLEF has supported a data curation approach within CLEF as an extension to the traditional methodology in order to better manage, preserve, interpret and enrich the scientific data produced, and to effectively promote the transfer of knowledge. The current approach to experimental evaluation is mainly focused on creating comparable experiments and evaluating their performance whereas researchers would also greatly benefit from an integrated vision of the scientific data produced, together with analyses and interpretations, and from the possibility of keeping, re-using, and enriching them with further information. The way in which experimental results are managed, made accessible, exchanged, visualized, interpreted, enriched and referenced is an integral part of the process of knowledge transfer and sharing towards relevant application communities.

[4] `http://www.trebleclef.eu/`

The University of Padua has thus developed DIRECT: Distributed Information Retrieval Evaluation Campaign Tool[5], a digital library system for managing the scientific data and information resources produced during an evaluation campaign. A preliminary version of DIRECT was introduced into CLEF in 2005 and subsequently tested and developed in the CLEF 2006 and 2007 campaigns. It has been further developed under TrebleCLEF. In 2009, DIRECT managed the technical infrastructure for several of the CLEF tracks and tasks: Ad Hoc, ImageCLEFphoto, GridCLEF, managing:

- the track set-up, harvesting of documents, management of the registration of participants to tracks;
- the submission of experiments, collection of metadata about experiments, and their validation;
- the creation of document pools and the management of relevance assessment;
- the provision of common statistical analysis tools for both organizers and participants in order to allow the comparison of the experiments;
- the provision of common tools for summarizing, producing reports and graphs on the measured performances and conducted analyses.

6 Participation

Researchers from 117 different academic and industrial institutions submitted runs in CLEF 2009: 81 from Europe, 18 from N.America; 16 from Asia, 1 from S.America and 1 from Africa. Figure 2 shows the trend in participation over the years and Figure 3

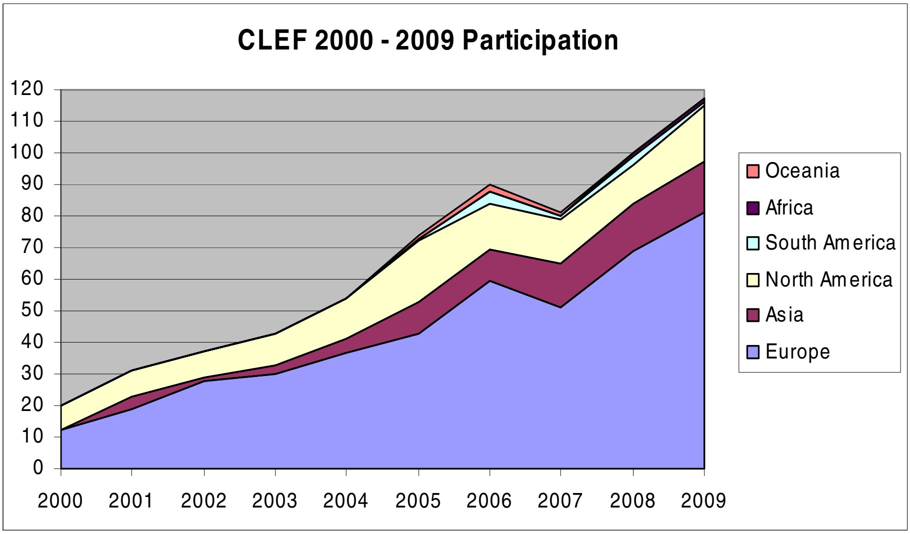

Fig. 2. CLEF 2000 – 2009: Participation

[5] http//direct.dei.unipd.it/

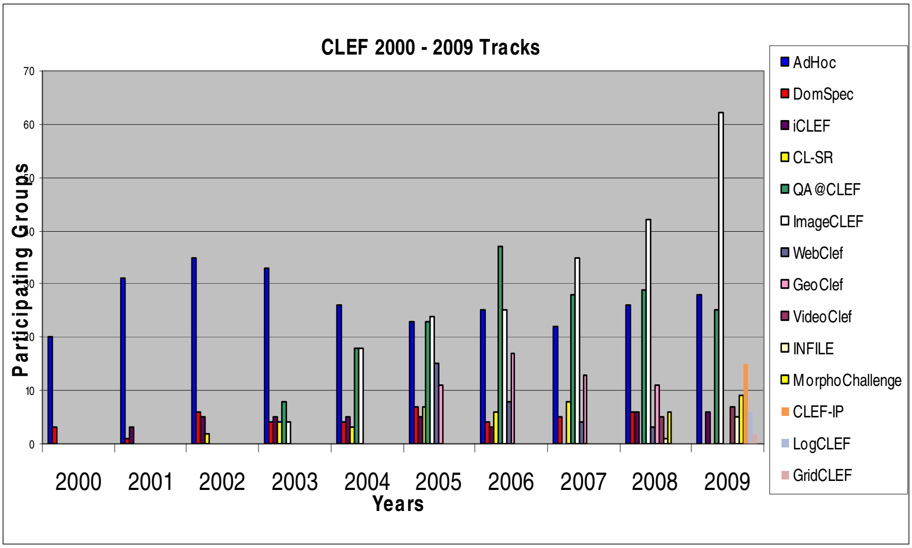

Fig. 3. CLEF 2000 – 2009: Participation per Track

shows the shift in focus as new tracks have been added. As can be seen, the number of groups participating in the Ad Hoc, iCLEF, QA and VideoCLEF tracks is almost the same as last year, there has been a rise of interest in INFILE and participation in the two new tracks (LogCLEF and CLEF-IP) is encouraging.

The most popular track is without doubt ImageCLEF which, with a notable increase from the previous year, tended to dominate the scene in 2009. This gives some cause for reflection as ImageCLEF is the track least concerned with multilinguality. A list of groups and indications of the tracks in which they participated can be found in the CLEF2009 Working Notes on the CLEF website.

7 The Future of CLEF

The main goal of CLEF in this first ten years of activity has been to sustain the growth of excellence in language processing and multilingual information access (MLIA) across language boundaries. A strong motivation has been the desire to promote the study and utilisation of languages other than English on the Internet. In this period, the CLEF activities have produced the following significant results:

- Creation of a very active multidisciplinary international research community, with strong interactions with the other main international initiatives for the evaluation of IR systems: TREC[6], NTCIR[7], and now FIRE[8];

[6] Text REtrieval Conferences, http://trec.nist.gov/

[7] NTCIR (NII Test Collection for IR Systems) Project, http://research.nii.ac.jp/ntcir/

[8] Forum for Information Retrieval Evaluation, http://www.isical.ac.in/~clia/

- Investigation of core issues in MLIA which enable effective transfer over language boundaries, including the development of multiple language processing tools (e.g. stemmers, word decompounders, part-of-speech taggers); creation of linguistic resources (e.g. multilingual dictionaries and corpora); implementation of appropriate cross-language retrieval models and algorithms for different tasks and languages;
- Creation of important reusable test collections and resources in diverse media for a large number of European languages, representative of the major European language typologies;
- Significant and quantifiable improvements in the performance of MLIA systems.

CLEF 2009 has represented an important milestone for the MLIA community. After ten years of activity focused on stimulating the development MLIA systems and functionality through the organisation of increasingly complex evaluation tasks and presenting the results at an annual workshop, we have decided to widen the format. CLEF 2010 will thus take the form of an independent Conference soliciting the submission of papers that propose new retrieval tasks, new evaluation tools, new measures, and new types of operational evaluation, organised in conjunction with a set of Evaluation Labs, which will continue the CLEF tradition of community-based evaluation and discussion on evaluation issues. Two different forms of labs are offered: "campaign-style" labs running evaluation tasks and experiments during the nine month period preceding the conference, and "workshop-style” labs exploring issues of information access evaluation and related fields.

The Conference will be held in Padua, Italy, September 2010, as a four day event: The first two days will consist of plenary sessions in which keynote speeches and peer-reviewed papers will be presented. The goals will be to explore current needs and practices for information access and discuss new directions for future activities in the European multilingual /multimodal IR system evaluation context. In Days 3 and 4, the results of the Labs will be presented in full and half-day workshops. Information on CLEF 2010 is available online[9].

Acknowledgements

It would be impossible to run the CLEF evaluation initiative and organize the annual workshops without considerable assistance from many groups. CLEF is organized on a distributed basis, with different research groups being responsible for the running of the various tracks. My gratitude goes to all those who have been involved in the coordination of the 2009 campaigns. A list of the main institutions involved is given at the beginning of this volume. Here below, let me thank just some of the people responsible for the coordination of the different tracks. My apologies to all those I have not managed to mention:

- Abolfazl AleAhmad, Hadi Amiri, Eneko Agirre, Giorgio Di Nunzio, Nicola Ferro, Nicolas Moreau, Arantxa Otegi and Vivien Petras for the Ad Hoc Track
- Paul Clough, Julio Gonzalo and Jussi Karlgren for iCLEF

[9] http://clef2010.org/

- Iñaki Alegria, Davide Buscaldi, Luís Miguel Cabral, Pere R. Comas, Corina Forascu, Pamela Forner, Olivier Galibert, Danilo Giampiccolo, Nicolas Moreau, Djamel Mostefa, Petya Osenova, Anselmo Peñas, Álvaro Rodrigo, Sophie Rosset, Paolo Rosso, Diana Santos, Richard Sutcliff and Jordi Turmo for QA@CLEF
- Brian Bakke, Steven Bedrick, Barbara Caputo, Paul Clough, Peter Dunker, Thomas Deselaers, Thomas Deserno, Ivan Eggel, Mark Oliver Güld, William Hersh, Patric Jensfelt, Charles E. Kahn Jr., Jana Kludas, Jayashree Kalpathy–Cramer, Henning Müller, Stefanie Nowak, Monica Lestari Paramita, Andrzej Pronobis, Saïd Radhouani, Mark Sanderson, Tatiana Tommasi, Theodora Tsikrika and Petra Welter for ImageCLEF
- Romaric Besançon, Stéphane Chaudiron, Khalid Choukri, Meriama Laïb, Djamel Mostefa and Ismaïl Timimi for INFILE
- Gareth J.F. Jones, Martha Larson and Eamonn Newman for VideoCLEF
- Giovanna Roda, John Tait, Florina Piroi, and Veronika Zenz for CLEF-IP
- Maristella Agosti, Giorgio Di Nunzio, Christine Doran, Inderjeet Mani, Thomas Mandl, Julia Maria Schulz and Alexander Yeh for LogCLEF
- Nicola Ferro and Donna Harman for GridCLEF
- Graeme W. Blackwood, William Byrne Mikko Kurimo, Ville T. Turunen and Sami Virpioja for MorphoChallenge at CLEF
- Marco Duissin, Giorgio Di Nunzio and Nicola Ferro for developing and managing the DIRECT infrastructure.

I also thank all those colleagues who have helped us by preparing topic sets in different languages and the members of the CLEF Steering Committee who have assisted me with their advice and suggestions throughout this campaign.

Furthermore, I gratefully acknowledge the support of all the data providers and copyright holders, and in particular:

- The Los Angeles Times, for the American-English newspaper collection.
- SMG Newspapers (The Herald) for the British-English newspaper collection.
- Le Monde S.A. and ELDA: Evaluations and Language resources Distribution Agency, for the French newspaper collection.
- Frankfurter Rundschau, Druck und Verlagshaus Frankfurt am Main; Der Spiegel, Spiegel Verlag, Hamburg, for the German newspaper collections.
- Hypersystems Srl, Torino and La Stampa, for the Italian newspaper data.
- Agencia EFE S.A. for the Spanish news agency data.
- NRC Handelsblad, Algemeen Dagblad and PCM Landelijke dagbladen/Het Parool for the Dutch newspaper data.
- Aamulehti Oyj and Sanoma Osakeyhtiö for the Finnish newspaper data.
- Russika-Izvestia for the Russian newspaper data.
- Hamshahri newspaper and DBRG, Univ. Tehran, for the Persian newspaper data.
- Público, Portugal, and Linguateca for the Portuguese (PT) newspaper collection.
- Folha, Brazil, and Linguateca for the Portuguese (BR) newspaper collection.
- Tidningarnas Telegrambyrå (TT) SE-105 12 Stockholm, Sweden for the Swedish newspaper data.
- Schweizerische Depeschenagentur, Switzerland, for the French, German & Italian Swiss news agency data.

- Ringier Kiadoi Rt. (Ringier Publishing Inc.).and the Research Institute for Linguistics, Hungarian Acad. Sci. for the Hungarian newspaper documents.
- Sega AD, Sofia; Standart Nyuz AD, Sofia, Novinar OD, Sofia and the BulTreeBank Project, Linguistic Modelling Laboratory, IPP, Bulgarian Acad. Sci, for the Bulgarian newspaper documents
- Mafra a.s. and Lidové Noviny a.s. for the Czech newspaper data
- Usurbilgo Udala, Basque Country, Spain, for the Egunkaria, Basque newspaper documents
- The European Commission – Joint Research Centre for the JRC Acquis Parallel corpus of European legislation in many languages.
- AFP Agence France Presse for the English, French and Arabic newswire data used in the INFILE track
- The British Library, Bibliothèque Nationale de France and the Austrian National Library for the library catalog records forming part of The European Library (TEL)
- The European Library (TEL) for use of TEL log files
- Tumba! web search engine of the Faculdade de Ciências da Universidade de Lisboa (FCUL), Portugal, for logfile querying
- Aachen University of Technology (RWTH), Germany, for the IRMA annotated medical images.
- Radiological Society of North America for the images of the Radiology and Radiographics journals.
- Lung Image Database Consortium (LIDC) for their database of lung nodules.
- Belga Press Agency, Belgium, for BELGAPICTURE image collection
- LIACS Medialab, Leiden University, The Netherlands & Fraunhofer IDMT, Ilmenau, Germany for the use of the MIRFLICKR 25000 Image collection
- Wikipedia for the use of the Wikipedia image collection.
- ELDA for the use of the ESTER Corpus: Manual and automatic transcripts of French broadcast news
- ELDA for the use of EPPS 2005/2006 ES & EN Corpora: Manual and automatic transcriptions of European Parliament Plenary Sessions in Spanish and English
- Matrixware Information Services GmbH for the use of a collection of patent documents in English, French and German from the European Patent Office
- The Institute of Sound and Vision, The Netherlands, for the English/Dutch videos, the University of Twente for the speech transcriptions, and Dublin City University for the shot segmentation.

Without their contribution, this evaluation activity would be impossible.