

What happened in CLEF 2008

Introduction to the Working Notes

Carol Peters

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

The objective of the Cross Language Evaluation Forum¹ 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 not only to meet but also to anticipate the emerging needs of the R&D community and to encourage the development of next generation multilingual IR systems.

These Working Notes contain descriptions of the experiments conducted within CLEF 2008 – the ninth in a series of annual system evaluation campaigns. The results of the experiments will be presented and discussed in the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. The final papers - revised and extended as a result of the discussions at the Workshop - together with a comparative analysis of the results will appear in the CLEF 2008 Proceedings, to be published by Springer in their Lecture Notes for Computer Science series.

Since CLEF 2005, the Working Notes are published in electronic format only and are distributed to participants at the Workshop together with a printed volume of Extended Abstracts. In previous years, this was in form of a CD-ROM; this year the Working Notes will be distributed on a memory stick. We're moving with the times!

Both Working Notes and Book of Abstracts are divided into ten sections, corresponding to the seven main evaluation tracks, the two pilot tasks, and an additional section describing another evaluation initiative using CLEF data: MorphoChallenge 2008. In addition appendices are included containing run statistics for the Ad Hoc, Domain-Specific, and GeoCLEF tracks, plus a list of all participating groups showing in which track they took part.

The main features of the 2008 campaign are briefly outlined here below in order to provide the necessary background to the experiments reported in the rest of the Working Notes.

1. Tracks and Tasks in CLEF 2008

CLEF 2008 offered seven tracks designed to evaluate the performance of systems for:

- multilingual textual document retrieval (Ad Hoc)
- mono- and cross-language information retrieval on structured scientific data (Domain-Specific)
- interactive cross-language retrieval (iCLEF)
- multiple language question answering (QA@CLEF)
- cross-language retrieval in image collections (ImageCLEF)
- multilingual retrieval of web documents (WebCLEF)
- cross-language geographical information retrieval (GeoCLEF)

Two new tracks were offered as pilot tasks:

- cross-language video retrieval (VideoCLEF)
- multilingual information filtering (INFILE@CLEF)

In addition, Morpho Challenge 2008 was organized in collaboration with CLEF as part of the EU Network of Excellence Pascal Challenge Program². The Morpho Challenge participants will meet separately before the main CLEF workshop on the morning of Wednesday 17 September, to discuss their results.

¹ Since the beginning of 2008, CLEF is included in the activities of the TrebleCLEF Coordination Action, funded by the Seventh Framework Programme of the European Commission. For information on TrebleCLEF, see www.trebleclef.eu.

² See <http://www.cis.hut.fi/morphochallenge2008/>

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

Multilingual Textual Document Retrieval (Ad Hoc): The aim of this track is to promote the development of monolingual and cross-language textual document retrieval systems. From 2000 - 2007, the track exclusively used collections of European newspaper and news agency documents. This year 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). 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 this year used word sense disambiguated (WSD) data. The track was coordinated jointly by ISTI-CNR and Padua University, Italy; Hildesheim University, Germany; and the University of the Basque Country, Spain, with the collaboration of the Database Research Group, University of Tehran, Iran.

Cross-Language Scientific Data Retrieval (Domain-Specific): The focus of the track is research into how the structure of data in collections (i.e. metadata, controlled vocabularies) can be exploited to improve search. Mono- and cross-language domain-specific retrieval is studied in the domain of social sciences using structured data (e.g. bibliographic data, keywords, and abstracts) from scientific reference databases. This year, the target collections provided were: GIRT-4 for German/English, Cambridge Sociological Abstracts for English, and the INION corpus ISISS provided by the Institute of Scientific Information for Social Sciences of the Russian Academy of Science. A multilingual controlled vocabulary (German, English, Russian) suitable for use with GIRT-4 and ISISS together with a bi-directional mapping between this vocabulary and that used for indexing the Sociological Abstracts was provided. Topics were offered in English, German and Russian. The track was coordinated by GESIS-IZ Social Science Information Centre, Bonn, Germany.

Interactive Cross-Language Retrieval (iCLEF): In iCLEF, cross-language search capabilities are studied from a user-inclusive perspective. A central research question is 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 moved from news collections (a standard for text retrieval experiments) in order to explore user behaviour in a collection where the cross-language search necessity arises more naturally for average users. The choice fell on Flickr⁴, a large-scale, online image database based on a large social network of WWW users, with the potential for offering both challenging and realistic multilingual search tasks for interactive experiments. The search interface provided by the iCLEF organizers was a basic cross-language retrieval system for the Flickr image database presented as an online game: the user is given an image, and must find it again without any a priori knowledge of the language(s) in which the image is annotated. The game was publicized on the CLEF mailing list and prizes were offered for the best results in order to encourage participation.

The main novelty of the iCLEF 2008 experiments was the shared analysis of a search log from a single search interface provided by the organizers (i.e. the focus was on log analysis, rather than on system design). Search logs were harvested from the search interface described above and iCLEF participants could essentially do two things:

- Search log analysis: participants had access to the search logs, and could freely perform data mining studies on them, such as looking for differences in search behaviour according to language skills, or looking for correlations between search success and search strategies, etc.
- Interactive experiments: participants could recruit their own users and conduct their own experiments with the interface. For instance, they could recruit a set of users with passive abilities and another with active abilities in certain languages and, besides studying the search logs, could perform observational studies on how they search, conduct interviews, etc.

The track was coordinated by UNED, Madrid, Spain; Sheffield University, UK; Swedish Institute of Computer Science, Sweden.

Multilingual Question Answering (QA@CLEF): This track has been offering monolingual and cross-language question answering tasks since 2003. QA@CLEF 2008 proposed both main and pilot tasks. The main scenario was event-targeted QA on a heterogeneous document collection (news articles and Wikipedia). A large number of questions were topic-related, i.e. clusters of related questions possibly containing anaphoric references. Besides the usual news collections, articles from Wikipedia were also considered as sources of answers. Many monolingual and cross-language sub-tasks were offered: Basque, Bulgarian, Dutch, English, French, German,

³ See <http://www.theeuropeanlibrary.org/>

⁴ See <http://www.flickr.com/>

Italian, Portuguese, Romanian and Spanish were proposed as both query and target languages; not all were used in the end. The additional exercises were the following:

- The Answer Validation Exercise (AVE) in its third edition was aimed at evaluating answer validation systems based on recognizing textual entailment.
- QAST was focused on Question Answering over Speech Transcriptions of seminars. In this 2nd year pilot task, answers to factual and definitional questions in English were to be extracted from spontaneous speech transcriptions related to separate scenarios in English, French and Spanish.
- QA-WSD provided questions and collections with already disambiguated Word Senses in order to study their contribution to QA performance.

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 of images from multilingual collections; both text and visual retrieval techniques were exploitable. Five challenging tasks were offered in 2008:

- A photo retrieval task: a good image search engine ensures that duplicate or near duplicate documents retrieved in response to a query are hidden from the user. Ideally the top results of a ranked list will contain diverse items representing different sub-topics within the results. This task focused on the study of successful clustering to provide diversity in the top-ranked results. The target collection contained images with captions in English and German; queries were in English.
- A medical image retrieval task: this is a domain-specific retrieval task in a domain where many ontologies exist; the target collection was a subset of the Goldminer collection containing images from English articles published in Radiology and Radiographics with captions and html links to the full text articles. Queries were provided in English, French and German.
- A visual concept deception task: the objective was to identify language-independent visual concepts that would help in solving the photo retrieval task. A training database was released with approximately 1,800 images classified according to a concept hierarchy. This data was used to train concept detection/annotation techniques. For each of the 1,000 images in the test database, participating groups were required to determine the presence/absence of the concepts.
- An automatic medical image annotation task: automatic image annotation or image classification can be an important step when searching for images from a database of radiographs. The aim of the task was to find out how well current language-independent techniques can identify image modality, body orientation, body region, and biological system on the basis of the visual information provided by the images.
- A Wikipedia image retrieval task: this was an ad hoc image search task where the information structure can be exploited for retrieval. The aim was to investigate retrieval approaches in the context of a larger scale and heterogeneous collection of images (similar to those encountered on the Web) that are searched for by users with diverse information needs.

The University and University Hospitals of Geneva, Switzerland; RWTH Aachen, Germany; Oregon Health and Science University, USA; Victoria University, Australia; Sheffield University, UK; Vienna University of Technology, Austria; CWI, The Netherlands, collaborated in the track organization.

Multilingual Web Retrieval (WebCLEF): In the past three years this track has focused on evaluation of systems providing multi- and cross-lingual access to web data. WebCLEF 2008 repeated the track setup of the 2007 edition. In 2008, a multilingual information synthesis task was offered, where, for a given topic, participating systems were asked to extract important snippets from web pages (fetched from the live web and provided by the task organizers). The systems had to focus on extracting, summarizing, filtering and presenting information relevant to the topic, rather than on large scale web search and retrieval per se. The focus was on refining the assessment procedure and evaluation measures. WebCLEF 2008 had lots of similarities with (topic-oriented) multi-document summarization and with answering complex questions. An important difference is that at WebCLEF, topics could come with extensive descriptions and with many thousands of documents from which important facts have to be mined. In addition, WebCLEF worked with web documents, that may be very noisy and redundant. The track was coordinated by the University of Amsterdam, The Netherlands.

Cross-Language Geographical Retrieval (GeoCLEF): The purpose of GeoCLEF is to test and evaluate cross-language geographic information retrieval for topics with a geographic specification. How best to transform into a machine readable format the imprecise description of a geographic area found in many user queries is still an open research problem. As in previous years, GeoCLEF 2008 examined geographic search of a text corpus. Some topics simulated the situation of a user who poses a query when looking at a map on the screen. For these topics, the system received the content part and a rectangular shape which defines the geographic context.

75 training topics were made available and 25 test topics in three languages were developed (English, German and Portuguese) A new pilot task on Wikipedia, GikiP, was also offered. The track was organized by Hildesheim University; Germany; Linguatca, Norway and Portugal; Sheffield University, UK.

Cross-Language Video Retrieval (VideoCLEF): The VideoCLEF Vid2RSS feed task was a classification task performed on a video corpus containing episodes of a dual language television program in Dutch and English. Participants were provided with speech recognition transcripts, metadata and keyframes for the video data. It is important to note that the languages occur side by side in the program and are not translations of the same content. The task was to group videos into topic categories and generate an RSS-feed for each category. The videos were classified (i.e. assigned to the topic categories) using the speech recognition transcripts for both languages. Keyframes and metadata could support the generation of the RSS-feeds, but could also be used to support classification, if participants so chose. The dual language programming of Dutch TV offered a unique scientific opportunity, presenting the challenge of how to exploit speech features from both languages. The track was coordinated by the University of Amsterdam; data was provided by The Netherlands Institute of Sound and Vision; University of Twente, The Netherlands, provided the speech transcripts; Dublin City University, Ireland, provided the shot segmentations and the key frames.

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 extended the last filtering track of TREC 2002 in the following ways:

- Monolingual and cross-language tasks were offered using a corpus of 100,000 Agence France Press (AFP) comparable newswire stories for Arabic, English and French;
- Evaluation was performed by an automatic interrogation of test systems with a simulated user feedback. A curve of the evolution of efficiency was computed along with more classical measures tested in TREC. The track was coordinated by the Evaluation and Language resources Distribution Agency (ELDA), France.

Unsupervised Morpheme Analysis (Morpho Challenge): The objective of Morpho Challenge is to design a statistical machine learning algorithm that discovers which morphemes (smallest individually meaningful units of language) form words. The scientific goals are:

- to understand the phenomena underlying word construction in natural languages
- to discover approaches suitable for a wide range of languages
- to advance machine learning methodology

The aim of Morpho Challenge 2008 was similar to that of Morpho Challenge 2007, where the goal was to find the morpheme analysis of the word forms in the data. Two tasks were offered. CLEF data for English, Finnish and German was used in the second task in which information retrieval experiments were performed where the words in the documents and queries were replaced by their proposed morpheme representations. The search was then based on morphemes instead of words. The activity was coordinated by Helsinki University of Technology, Finland.

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.

2. 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.

A number of different document collections were used in CLEF 2008 to build the test collections:

- CLEF multilingual corpus of more than 3 million news documents in 14 European languages. This corpus is divided into two comparable collections: 1994-1995 - Dutch, English, Finnish, French, German, Italian, Portuguese, Russian, Spanish, Swedish; 2000-2002 - Basque, Bulgarian, Czech, English, Hungarian. The Basque data was new this year. Parts of this collections were used in the AdHoc, QuestionAnswering, GeoCLEF and Morpho Challenge tracks.
- Data from The European Library /TEL): approximately 3 million library catalog records in English, French and German, used in the Ad Hoc track.
- Hamshahri Persian newspaper corpus; nearly 170,000 documents used in the Ad Hoc track;
- The GIRT-4 social science database in English and German (over 300,000 documents) and two Russian databases: the Russian Social Science Corpus (approx. 95,000 documents) and the Russian ISISS collection

for sociology and economics (approx. 150,000 docs). The RSSC corpus was not used this year. Cambridge Sociological Abstracts in English (20,000 docs). These collections were used in the domain-specific track.

- Online Flickr database, used in the iCLEF track
- The ImageCLEF track used collections for both general photographic and medical image retrieval:
 - » IAPR TC-12 photo database of 20,000 still natural images (plus 20,000 corresponding thumbnails) with captions in English, and German;
 - » ARRS Goldminer database – nearly 200,000 images published in 249 selected peer-reviewed radiology journals
 - » IRMA collection in English and German of 12,000 classified images for automatic medical image annotation
 - » INEX Wikipedia image collection, approximately 150,000 images associated with unstructured and noisy textual annotations in English
- Videos in Dutch and English of documentary television programs, approximately 30 hours, used in the VideoCLEF track.
- Agence France Press (AFP) comparable newswire stories in Arabic, French and English for the INFILE track

3. CLEF & TrebleCLEF

CLEF is organized mainly through the voluntary efforts of many different institutions and research groups. Section 1 lists the groups responsible for the coordination of this year's tracks. A full list of the people and groups involved in the organization of CLEF2008 is given at the end of this paper. 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 Excellence⁵. CLEF 2008 and 2009 are organized under the auspices of TrebleCLEF, a Coordination Action of the Seventh Framework Programme, Theme ICT 1-4-1.

TrebleCLEF intends to build on and extend the results already achieved by CLEF. The objective is to support the development and consolidation of expertise in the multidisciplinary research area of multilingual information access and to promote a dissemination action in the relevant application communities.

TrebleCLEF thus intends to promote research, development, implementation and industrial take-up of multilingual, multimodal information access functionality in the following ways:

- by supporting the annual system evaluation campaigns of the Cross-Language Evaluation Forum with tracks and tasks designed to stimulate R&D to meet the requirements of the user and application communities, with particular focus on the following key areas:
 - user modeling, e.g. what are the requirements of different classes of users when querying multilingual information sources;
 - language-specific experimentation, e.g. looking at differences across languages in order to derive best practices for each language, best practices for the development of system components and best practices for MLIA systems as a whole;
 - results presentation, e.g. how can results be presented in the most useful and comprehensible way to the user.
- by constituting a scientific forum for the MLIA community of researchers enabling them to meet and discuss results, emerging trends, new directions:
 - providing a scientific digital library to make accessible the scientific data and experiments produced during the course of an evaluation campaign. This library will also provide tools for analyzing, comparing, and citing the scientific data of an evaluation campaign, as well as curating, preserving, annotating, enriching, and promoting the re-use of them;
- by acting as a virtual centre of competence providing a central reference point for anyone interested in studying or implementing MLIA functionality and encouraging the dissemination of information:
 - making publicly available sets of guidelines on best practices in MLIA (e.g. what stemmer to use, what stop list, what translation resources, how best to evaluate, etc., depending on the application requirements);
 - making tools and resources used in the evaluation campaigns freely available to a wider public whenever possible; otherwise providing links to where they can be acquired;
 - organising workshops, and/or tutorials and training sessions.

⁵ The DELOS Network closed at the end of 2007 and a self-sustaining DELOS Association was launched.

The aim is 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.

4. Technical Infrastructure

As mentioned in the previous section, TrebleCLEF supports 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.

The University of Padua has thus developed DIRECT: Distributed Information Retrieval Evaluation Campaign Tool⁶, 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 is now being further developed under TrebleCLEF.

DIRECT currently manages the technical infrastructure for several of the CLEF tracks: Ad Hoc, Domain-Specific, GeoCLEF, providing procedures to handle:

- 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.

An extension to also manage the technical infrastructure for ImageCLEF is now under discussion.

In the CLEF 2008 campaign, DIRECT has been used by over 130 participants from 20 countries, who have submitted 490 experiments. Within the DIRECT framework, 80 assessors have created over 200 topics in seven different languages and have assessed about 250,000 documents, including documents in languages like Russian, which uses the Cyrillic alphabet, and Persian, which is written from right to left.

DIRECT is designed and implemented by Giorgio Di Nunzio and Nicola Ferro.

5. Participation

A total of 105 groups submitted runs in CLEF 2008, a big increase on the 81 groups of CLEF 2007: 71(51) from Europe, 13(14) from N.America; 17(14) from Asia, 3(1) from S.America and 1(0) from Africa. The breakdown of participation of groups per track is as follows: Ad Hoc 31(22); Domain-Specific 6(5); iCLEF 6(na); QAatCLEF 29(28); ImageCLEF 42(35); WebCLEF 4(4); GeoCLEF 10(13); VideoCLEF 6 (na); INFILE 1 (na); Morpho Challenge 6(6).⁷ A list of groups and indications of the tracks in which they participated is given in the Appendix to these Working Notes. Figure 1 shows the variation in participation over the years and Figure 2 shows the shift in focus as new tracks are added.

It can be seen that the increase in participation in CLEF this year is almost entirely due to a massive rise in the participation from Europe – the other continents remained more or less stable. It was great to have our first participation from an African country: Uganda, but we missed out on Oceania this year.. It is interesting to note once again that the most popular track at CLEF, ImageCLEF, is also probably the least multilingual track as much of the work is done in a language-independent context. The participation in the WebCLEF and INFILE tracks was very disappointing and it is not expected that these two tracks will be continued next year.

⁶ <http://direct.dei.unipd.it/>

⁷ Last year's figures are between brackets where applicable; we did not register Morpho Challenge as a CLEF activity in 2007.

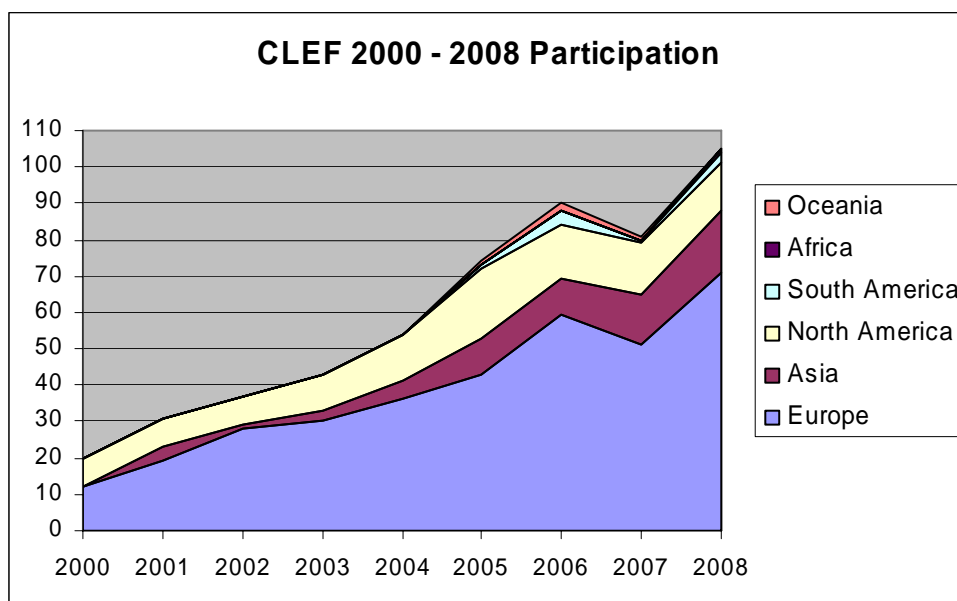


Figure 1. CLEF 2000 – 2008: Variation in Participation

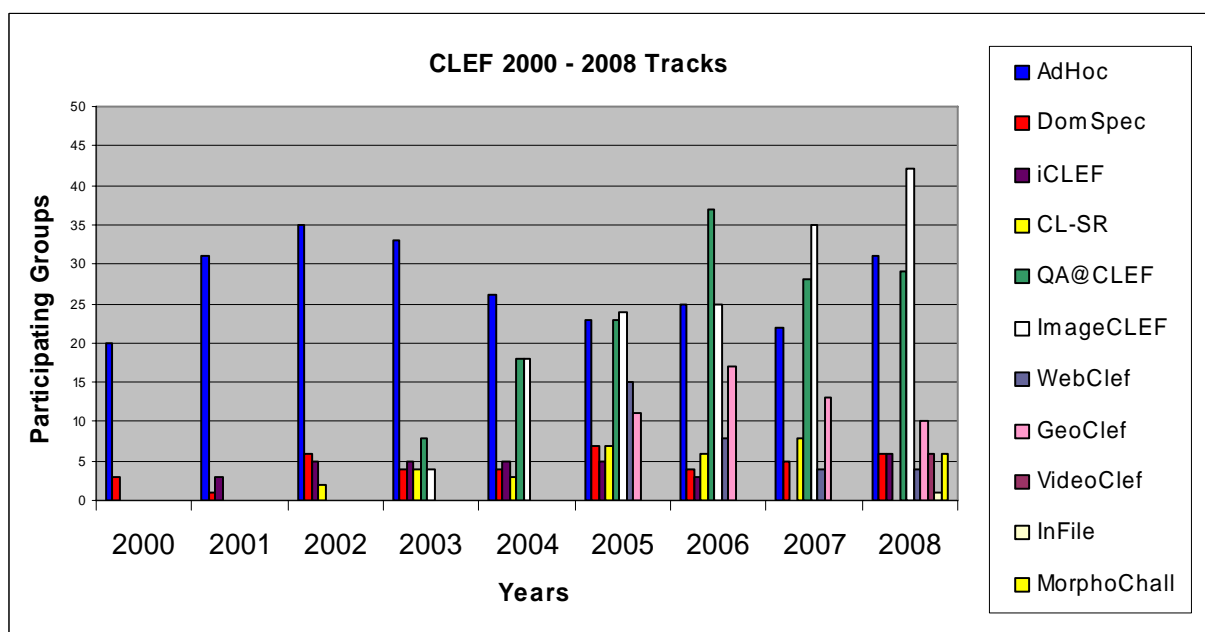


Figure 2. CLEF 2000 – 2008: Participation per Track in Tracks

6. Workshop

CLEF aims at creating a strong CLIR/MLIA research and development community. The Workshop plays an important role by providing the opportunity for all the groups that have participated in the evaluation campaign to get together comparing approaches and exchanging ideas. The work of the groups participating in this year's campaign will be presented in plenary and parallel paper and poster sessions. There will also be break-out sessions for more in-depth discussion of the results of individual tracks and intentions for the future. The final sessions will include discussions on ideas for new tracks in future campaigns. Overall, the Workshop should provide an ample

panorama of the current state-of-the-art and the latest research directions in the multilingual information retrieval area. I very much hope that it will prove an interesting, worthwhile and enjoyable experience to all those who participate.

The final programme and the presentations at the Workshop are posted on the CLEF website at <http://www.clef-campaign.org>.

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 2008 campaigns. A list of the main institutions involved is given in the following pages. 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, Thomas Mandl, Nicolas Moreau, Alessandro Nardi and Vivien Petras for the Ad Hoc Track
- Vivien Petras and Maximillian Stempfhuber for the Domain-Specific track
- Paul Clough, Julio Gonzalo and Jussi Karlgren for iCLEF
- Danilo Giampiccolo Pamela Forner, Dan Cristea, Corina Forascu, Nicolas Moreau, Petya Osenova, Anselmo Peñas, Iñaki Alegria, Bogdan Sacaleanu, Prokopis Prokopidis, Paulo Rocha and Richard Sutcliffe for QA@CLEF
- Allan Hanbury, Paul Clough, Thomas Arni, Mark Sanderson, Henning Müller, Thomas Deselaers , Thomas Deserno, Michael Grubinger, Jayashree Kalpathy–Cramer, and William Hersh for ImageCLEF
- Valentin Jijkoun and Maarten de Rijke for Web-CLEF
- Thomas Mandl, Fredric Gey, Ray Larson, Mark Sanderson, Diana Santos, Paula Carvalho for GeoCLEF
- Martha Larson and Gareth Jones for VideoCLEF
- Djamel Mostefa for INFILE
- Marco Duissin, Giorgio Di Nunzio and Nicola Ferro for developing and managing the DIRECT infrastructure.

I should also like to thank 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. Without their contribution, this evaluation activity would be impossible.

Finally, I should like to express my gratitude to Francesca Borri and Alessandro Nardi in Pisa and Jette Junge in Aarhus for their assistance in the organisation of the CLEF 2008 Workshop.

Coordination

CLEF is run mainly on a voluntary basis and is coordinated by the Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. The following institutions have contributed to the organisation of the different tracks of the CLEF 2008 campaign:

- Athena Research Center, Greece
- Business Information Systems, Univ. of Applied Sciences Western Switzerland, Sierre, Switzerland
- Centre for the Evaluation of Human Language and Multimodal Communication Technologies (CELCT), Trento, Italy
- Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands
- Computer Science Department, University of the Basque Country, Spain
- Computer Vision and Multimedia Lab, University of Geneva, Switzerland
- Database Research Group, University of Tehran, Iran
- Department of Computer Science, RWTH Aachen University, Germany
- Department of Computer Science and Information Systems, University of Limerick, Ireland
- Department of Information Engineering, University of Padua, Italy
- Department of Information Science, University of Hildesheim, Germany
- Department of Information Studies, University of Sheffield, UK
- Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science Univ., USA
- Department of Medical Informatics, Aachen University of Technology (RWTH), Germany
- Department of Medical Informatics, University Hospitals and University of Geneva, Switzerland
- Evaluations and Language Resources Distribution Agency Sarl, Paris, France
- German Research Centre for Artificial Intelligence, DFKI, Saarbrücken, Germany
- GESIS-IZ Social Science Information Centre, Bonn, Germany
- Information Science, University of Groningen, The Netherlands
- Institute of Computer Aided Automation, Vienna University of Technology, Austria
- Intelligent Systems Lab, University of Amsterdam, Netherlands
- Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), Orsay, France
- Lenguajes y Sistemas Informáticos, Universidad Nacional de Educación a Distancia, Madrid, Spain
- Linguatca, Sintef, Oslo, Norway
- Linguatca, CISUC, Department of Information Engineering, University of Coimbra, Portugal
- Linguatca, XLDB, LasiGE, Department of Information Engineering, University of Lisbon, Portugal
- Linguistic Modelling Laboratory, Bulgarian Academy of Sciences
- Microsoft Research Asia
- National Institute of Standards and Technology, Gaithersburg MD, USA
- Research Computing Center of Moscow State University, Russia
- Romanian Institute for Computer Science, Romania
- School of Computer Science and Mathematics, Victoria University, Australia
- School of Computing, Dublin City University, Ireland
- TALP Research Center, Universitat Politècnica de Catalunya, Barcelona, Spain.
- UC Data Archive and School of Information Management and Systems, UC Berkeley, USA

CLEF Steering Committee

- Maristella Agosti, University of Padova, Italy
- Martin Braschler, Zurich University of Applied Sciences Winterthur, Switzerland
- Amedeo Cappelli, ISTI-CNR & CELCT, Italy
- Hsin-Hsi Chen, National Taiwan University, Taipei, Taiwan
- Khalid Choukri, Evaluations and Language resources Distribution Agency, Paris, France
- Paul Clough, University of Sheffield, UK
- Thomas Deselaers, RWTH Aachen University, Germany
- Giorgio Di Nunzio, University of Padova, Italy
- David A. Evans, Clairvoyance Corporation, USA
- Marcello Federico, ITC-irst, Trento, Italy
- Nicola Ferro, University of Padova, Italy
- Christian Fluhr, CEA-LIST, Fontenay-aux-Roses, France
- Norbert Fuhr, University of Duisburg, Germany
- Frederic C. Gey, U.C. Berkeley, USA
- Julio Gonzalo, LSI-UNED, Madrid, Spain
- Donna Harman, National Institute of Standards and Technology, USA
- Gareth Jones, Dublin City University, Ireland
- Franciska de Jong, University of Twente, The Netherlands
- Noriko Kando, National Institute of Informatics, Tokyo, Japan
- Jussi Karlgren, Swedish Institute of Computer Science, Sweden
- Michael Kluck, German Institute for International and Security Affairs, Berlin, Germany
- Natalia Loukachevitch, Moscow State University, Russia
- Bernardo Magnini, ITC-irst, Trento, Italy
- Paul McNamee, Johns Hopkins University, USA
- Henning Müller, University & University Hospitals of Geneva, Switzerland
- Douglas W. Oard, University of Maryland, USA
- Anselmo Peñas, LSI-UNED, Madrid, Spain
- Maarten de Rijke, University of Amsterdam, The Netherlands
- Diana Santos, Linguatca, Sintef, Oslo, Norway
- Jacques Savoy, University of Neuchatel, Switzerland
- Peter Schäuble, Eurospider Information Technologies, Switzerland
- Max Stempfhuber, Informationszentrum Sozialwissenschaften Bonn, Germany
- Richard Sutcliffe, University of Limerick, Ireland
- Hans Uszkoreit, German Research Center for Artificial Intelligence (DFKI), Germany
- Felisa Verdejo, LSI-UNED, Madrid, Spain
- José Luis Vicedo, University of Alicante, Spain
- Ellen Voorhees, National Institute of Standards and Technology, USA
- Christa Womser-Hacker, University of Hildesheim, Germany

Data Providers

The support of all the data providers and copyright holders that have contributed to the creation of the CLEF test collections is gratefully acknowledged, 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 Depeschagentur, 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
- 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)
- InformationsZentrum Sozialwissen-schaften, Bonn, for the GIRT social science database.
- SocioNet system for the Russian Social Science Corpora
- Institute of Scientific Information for Social Sciences of the Russian Academy of Science (ISISS RAS) for the ISISS database
- University and University Hospitals, Geneva, Switzerland and Oregon Health and Science University for the ImageCLEFmed Radiological Medical Database.
- Aachen University of Technology (RWTH), Germany, for the IRMA annotated medical images.
- The Radiology Dept. of the University Hospitals of Geneva for the Casimage database and the PEIR (Pathology Education Image Resource) for the images and the HEAL (Health Education Assets Library) for the Annotation of the Peir dataset.
- Mallinkrodt Institute of Radiology for permission to use their nuclear medicine teaching file.
- University of Basel's Pathopic project for their Pathology teaching file.
- Michael Grubinger, administrator of the IAPR Image Benchmark, Clement Leung who initiated and supervised the IAPR Image Benchmark Project, and André Kiwitez, the Managing Director of Viventura for granting access to the image database and the raw image annotations of the tour guides.
- USC Shoah Foundation Institute, and IBM (English) and The Johns Hopkins University Center for Language and Speech Processing (Czech) for the speech transcriptions.
- 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.