



# PROMISE

Participative Research labOratory for  
Multimedia and Multilingual Information Systems

FP7 ICT 2009.4.3, Intelligent Information Management  
[www.promise-noe.eu](http://www.promise-noe.eu)

## Deliverable 7.9

### Third PROMISE Annual Conference and Proceedings

CLEF 2012 Conference and Labs of the Evaluation Forum –  
Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics  
Rome, Italy, 17-20 September 2012

Version 1.0, 30<sup>th</sup> September 2012







## Document Information

<b>Deliverable number:</b>	7.9
<b>Deliverable title:</b>	Third PROMISE Annual Conference and Proceedings. CLEF 2012 Conference and Labs of the Evaluation Forum – Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics. Rome, Italy, 17-20 September 2012
<b>Delivery date:</b>	30/09/2012
<b>Lead contractor for this deliverable</b>	CELCT
<b>Author(s):</b>	Pamela Forner
<b>Participant(s):</b>	CELCT
<b>Workpackage:</b>	7
<b>Workpackage title:</b>	Dissemination, IPR, and Resources
<b>Workpackage leader:</b>	ELDA
<b>Dissemination Level:</b>	PU – Public
<b>Version:</b>	1.0
<b>Keywords:</b>	CLEF 2012 conference, labs and workshop, Community sessions

## History of Versions

Version	Date	Status	Author (Partner)	Description/Approval Level
1.0	30/09/2012	Draft	CELCT	Circulated to reviewers
1.1	03/10/2012	Draft	UNIPD and ELDA	Revised Version
1.2	05/10/2012	Final	CELCT	Final Version

## Abstract

Deliverable 7.9 consists of two main parts: the organization of the CLEF 2012 conference and labs, and the publication of the proceedings. The conference and the labs were held at "Roma Eventi - Fontana di Trevi" conference center in Rome from 17<sup>th</sup> to 20<sup>th</sup> September 2012.

This documentation provides an overview of the CLEF 2012 conference together with a description of the labs which were proposed, details on this year's participation, and a short report on the workshops. In addition also a description of the website developed for the event and its contents is provided.

In appendix, we provide:

- the list of the people involved in the organization of CLEF 2012
- the list of the institutions involved in the coordination of CLEF 2012 Labs and Workshop
- the Conference Programme
- The list of accepted papers at the Conference
- the Labs and Workshop Programme
- the list of people attending the event
- the list of the institutions which submitted results in the CLEF 2012 evaluation campaigns (Labs), showing also in which Lab they took part
- CLEF 2012 Sponsors
- Statistics about CLEF 2012 website

## Table of Contents

Document Information .....	4
Abstract.....	4
Table of Contents.....	6
Executive Summary .....	7
1 Introduction.....	9
2 CLEF 2012 Conference.....	11
3 CLEF 2012 Community Session .....	13
4 CLEF 2012 Labs and Workshop.....	14
5 Participation in CLEF 2012 .....	21
6 CLEF 2012 Web Site and Mobile Version.....	25
6.1 Advertising.....	26
7 CLEF 2013 and Beyond.....	27
8 Acknowledgements .....	29
References .....	29
Appendix A: CLEF 2012 Organization .....	30
Appendix B: CLEF 2012 Evaluation Labs and Workshop Coordination .....	31
Appendix C: CLEF 2012 Programme at a Glance .....	34
Appendix D: CLEF 2012 Accepted papers at the Conference .....	37
Appendix E: CLEF 2012 Labs and Workshop Programme.....	39
Appendix F: List of people attending the Event.....	52
Appendix G: List of Institutions participating to the Labs.....	59
Appendix H: CLEF 2012 Sponsors .....	69
Appendix I: CLEF 2012 Website Statistics .....	70

## Executive Summary

Deliverable 7.9 comprises two main parts: the organization of the CLEF 2012 conference and labs, and the publication of the proceedings.

Evaluation activities have been widely credited with contributing to the advancement of information retrieval by (i) providing access to infrastructure and evaluation resources that support researchers in the development of new approaches, and (ii) encouraging collaboration and interaction between researchers both from academia and industry.

The CLEF Initiative is the major forum focusing on information access evaluation, meeting the needs of a large and diverse constituency, which includes researchers, professionals, research students and users. The CLEF Initiative is structured in two main parts:

1. a series of **Evaluation Labs**, i.e. laboratories to conduct evaluation of information access systems and workshops to discuss and pilot innovative evaluation activities;
2. a peer-reviewed **Conference** on a broad range of issues, including
  - investigation continuing the activities of the Evaluation Labs;
  - experiments using multilingual and multimodal data, such as the CLEF data;
  - research in evaluation methodologies and challenges.

The Conference and the annual meetings of the Evaluation Labs will be held as a joint event.

Building on the format first introduced in 2010 and continued in 2011, CLEF 2012 consists of :

1. 14 peer-reviewed research papers and 3 research posters, which are being presented in the conference sessions.
2. the more "classic" CLEF benchmarking activities are presented during the 7 lab sessions: CHiC, CLEF-IP, ImageCLEF, INEX, PAN, QA4MRE, and RepLab.

In addition to these research activities, the program contains a workshop (CLEFeHealth) and 2 Community Sessions, focusing on Other Evaluation Initiatives and evaluations in CLINIC. The combination of this diverse set of activities makes for a rich and interesting CLEF 2012 program.

All sessions (conference, labs, workshop and community sessions) were held at "Roma Eventi - Fontana di Trevi" conference center from 17<sup>th</sup> to 20<sup>th</sup> September 2012. As experimented last year, the conference presentations and the laboratories were interleaved over a three and a half day period.

The outcomes of the Conference and the Labs are collected in two different and separated publication forms:

- conference proceedings are published by Springer in their Lecture Notes for Computer Science (LNCS 7488) series with the title *Information Access Evaluation- Multilinguality, Multimodality and Visual Analytics*

- all the experiments carried out by systems during the CLEF 2012 evaluation campaigns are described in the Working Notes, which are published in digital form only and are available at the CLEF 2012 website (<http://www.clef2012.org/index.php?page=Pages/proceedings.php>). Printed abstracts are collected in the Book of Abstracts. Both Working Notes and Book of Abstracts are divided into eight main sections, corresponding to the eighth Labs activated this year; they are then further subdivided in many sub-sections corresponding to the different tasks planned. Both online Working Notes and Book of Abstracts have been assigned an ISBN and ISSN number, and have also been indexed by DBLP.

Post-conference publications in suitable venues, such as journals, are welcome and under the responsibility of Lab organizers.

The CLEF 2012 Conference Proceedings, together with the Labs Book of Abstracts have been distributed to CLEF 2012 attendees, and are thus also to be considered as part of this deliverable as appendices.

This documentation provides an overview of the CLEF 2012 conference with a description of the labs that have been offered, details on this year's participation, and a short report on the workshops. Moreover, we describe the website developed for the event and the contents available on it.

In appendix, we provide:

- the list of the people involved in the organization of CLEF 2012
- the list of the institutions involved in the coordination of CLEF 2012 Labs and Workshop
- the Conference Programme
- the list of accepted papers at the Conference
- the Labs and Workshop Programme
- the list of the people attending the event
- the list of the institutions which submitted results in the CLEF 2012 evaluation campaigns (Labs), showing also in which Lab they took part
- the CLEF 2012Sponsors
- The CLEF 2012 website statistics



## 1 Introduction

PROMISE is working on providing a virtual laboratory for conducting participative research and experimentation to carry out, advance and bring automation into the evaluation and benchmarking of such complex information systems, by facilitating management and offering access, curation, preservation, re-use, analysis, visualization, and mining of the collected experimental data. To this end, PROMISE organizes CLEF, an experimental evaluation activity for multilingual and multimedia information systems at an international level and on an annual basis.

The goal is to advance the traditional way of conducting evaluation campaigns by relying on large data sets, tackling realistic use cases and evaluation tasks designed for compelling user and industrial needs, advancing and automating the evaluation process to support better the envisioned tasks and use cases, providing a proper evaluation infrastructure, producing information and knowledge resources from the collected experimental data, and involving Europe-wide large researcher and developer communities with multidisciplinary competencies.

CLEF 2012 - Conference and Labs of the Evaluation Forum, (formerly known as Cross-Language Evaluation Forum) is next year's edition of the popular CLEF campaign and workshop series which has run since 2000 contributing to the systematic evaluation of information access systems, primarily through experimentation on shared tasks.

The CLEF Initiative is a self-organized body whose main mission is to promote research, innovation, and development of information access systems with an emphasis on multilingual and multimodal information with various levels of structure. CLEF promotes research and development by providing an infrastructure for:

- multilingual and multimodal system testing, tuning and evaluation;
- investigation of the use of unstructured, semi-structured, highly structured, and semantically enriched data in information access;
- creation of reusable test collections for benchmarking;
- exploration of new evaluation methodologies and innovative ways of using experimental data;
- discussion of results, comparison of approaches, exchange of ideas, and transfer of knowledge.

In 2010 CLEF was launched in a new format, consisting of an independent peer-reviewed annual conference on a broad range of topics in the fields of multilingual and multimodal information access evaluation, with research presentations, panels, poster and demo sessions; along with laboratory evaluation workshops designed to test various aspects of mono and cross-language information retrieval systems that continued the CLEF tradition of community-based evaluation interleaved during three and a half days of intense and stimulating research activities.

CLEF 2011 and CLEF 2012 continued to implement this format, with keynotes, contributed papers, lab sessions, poster sessions, reporting of other benchmarking initiatives and, for the first time this year, an evaluation clinic session where people with retrieval evaluation problems of some kind would be able to talk to evaluation experts and get methodological advice, new ideas, pointers to related problems, available solutions, etc.

CLEF 2012 featured seven benchmarking activities: RepLab, INEX, QA4MRE, CLEF-IP, ImageCLEF, PAN, and CHiC. In parallel, also the CLEFeHealth workshop was hosted, dealing with cross-language evaluation of methods, applications, and resources for eHealth document analysis with a focus on written and spoken natural-language processing.

#### BENCHMARKING ACTIVITIES:

**1. CHiC Cultural Heritage in CLEF**

a benchmarking activity to investigate systematic and large-scale evaluation of cultural heritage digital libraries and information access systems.

**2. CLEF-IP**

a benchmarking activity to investigate IR techniques in the patent domain

**3. ImageCLEF**

a benchmarking activity on the experimental evaluation of image classification and retrieval, focusing on the combination of textual and visual evidence

**4. INEX**

a benchmarking activity on the evaluation of XML retrieval

**5. PAN**

a benchmarking activity on uncovering plagiarism, authorship and social software misuse

**6. QA4MRE**

a benchmarking activity on the evaluation of Machine Reading systems through Question Answering and Reading Comprehension Tests

**7. RepLab**

a benchmarking activity on reputation management technologies

#### WORKSHOP:

1. **CLEFeHealth 2012** on Cross-Language Evaluation of Methods, Applications, and Resources for eHealth Document Analysis

## 2 CLEF 2012 Conference

The CLEF 2012 - Conference and Labs of the Evaluation Forum, – an event organized by PROMISE – took place in Rome, Italy on September 17<sup>th</sup> -20<sup>th</sup> , 2012.

This year the focus of the conference was on Information Access Evaluation with special attention to multilinguality, multimodality and visual analytics. Submission of papers were solicited on all aspects of Information Access in any modality and language. Relevant topics included:

**Information Access in any language or modality:** Information Retrieval, Image Retrieval, Question Answering, Search Interfaces and design, infrastructures, etc.

**Visual Analytics for Information Retrieval:** theoretical and practical results in the Visual Analytics field that are specifically targeted for Information Access data analysis.

**Evaluation initiatives:** conclusions, lessons learned, impact and projection of any evaluation initiative after completing their cycle.

**Evaluation:** methodologies, metrics, statistical and analytical tools, component based, user groups and use cases, ground-truth creation, impact of multilingual/multicultural/multimodal differences, etc.

**Technology transfer:** economic impact/sustainability of information access approaches, deployment and exploitation of systems, market studies, etc.

Submissions were gathered using the Springer OCS system, which allowed Program Chairs to manage the submission and review processes, as well as the collection of camera-ready papers.

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 report. 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.



Figure 1: One session at CLEF 2012 – room Loyola, at Roma Eventi Fontana di Trevi Congress Center  
Two keynote speakers highlighted important issues related to our field:

- **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 claims 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 Challenges* he claims 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.

All accepted papers and the abstract of the two invited talks, are published by Springer in their Lecture Notes for Computer Science (7488) series, which are attached to the paper version of this deliverable.

As conference proceedings, the LNCS publication was ready by the conference and was distributed to all the attendees. This innovation, introduced in 2010, is a novelty with respect to the CLEF tradition where CLEF LNCS post-proceedings containing revised version of the papers of the working notes were published almost one year after the workshop.

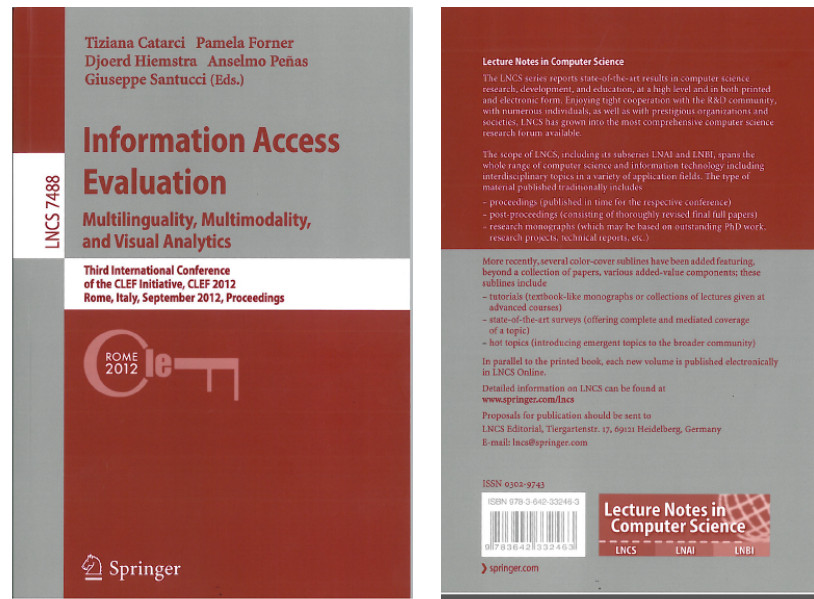


Figure 2: Front and back cover of the CLEF 2012 LNCS proceedings

The conference Programme is given in Appendix C.

### 3 CLEF 2012 Community Session

CLEF 2012 hosted two Community Sessions:

- Evaluation CLINIC: people with evaluation retrieval problem were able to talk to evaluation experts and get methodological advice, new ideas, pointers to related problems, available solutions, etc.
- 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.

#### Other Evaluation Initiatives Speakers and abstracts:

- *TREC at 20 years*, Donna Harman (NIST, USA) reviewed TREC's 20th meeting (2011), with comments on what has been learned (and not learned) about evaluation over those years.



- *TAC at 2011*, Donna Harman (NIST, USA) reviewed the TAC 2011 tasks, as well as looking ahead to 2012 and 2013.
- *What is happening in NTCIR*, Noriko Kando (National Institute of Informatics (NII), Japan) briefly introduced what happened in NTCIR-9 (Sept 2010 - December 2011), and what is happening in NTCIR-10. From NTCIR-9, the structure was changed to have multiple general co-chairs, program co-chairs, and EVIA co-chairs to make the organization stronger. The repertory of the tasks were changed and had CrossLink, GeoTime, Intent, 1Click, PatentMT, SpokenDoc, and VisEx in NTCIR-9.
- *ROMIP: One Step Forward, One Step Aside*, Pavel Braslavski (Kontur labs, Russia) briefly surveyed the history and presented state of ROMIP - a Russian TREC-like information retrieval evaluation campaign started in 2002. In particular, he reported on new tracks - sentiment analysis track introduced in 2011 and machine translation evaluation track scheduled for the next year (2013).
- *FIRE: A community development exercise*, Prasenjit Majumder reported that the FIRE initiative in India is holding its 4th evaluation campaign this year. He outlined the impact of FIRE in terms of community development in the Indian subcontinent. Attracting participants, particularly from the IR community in India, is one of the biggest challenges facing FIRE.
- *MediaEval in 2012*, Gareth Jones explained that MediaEval is a benchmarking initiative dedicated to evaluating new algorithms for multimedia access and retrieval. MediaEval attracts participants who are interested in multimodal approaches to multimedia involving, e.g., speech recognition, multimedia content analysis, user-contributed information (tags, tweets), viewer affective response, social networks, temporal and geo-coordinates. MediaEval 2012 features a range of evaluation tasks exploring all of these dimensions of multimedia search.

## 4 CLEF 2012 Labs and Workshop

CLEF 2012 had, as one of its most central feature, a broad palette of empirical evaluation activities for information access systems of various types. These were proposed and operated by groups of organisers volunteering their time and effort to define, promote, administrate and run an evaluation activity. In 2012, as in 2011 and 2010, two participation types were offered:

1. Evaluation labs that follow a “campaign-style” evaluation practice for specific information access problems (during the year preceding the conference) in the tradition of

past CLEF campaign tracks. Topics covered by evaluation labs can be inspired by any information access-related domain or task.

2. Lab workshops organised as speaking and discussion sessions to explore issues of evaluation methodology, metrics, and processes in information access and closely related fields, such as natural language processing, machine translation, and human-computer interaction. Lab workshops can be a first step towards an evaluation lab. One of the 2010 lab workshops (LogCLEF) evolved into a 2011 evaluation lab. This progression from a lab workshop to an evaluation lab is a development track we wish to encourage - but lab workshops do not need to be associated with an evaluation lab: theoretical and methodological issues are numerous in our field and can well be addressed in focused workshop sessions without shared tasks. If the organisers of the proposal are new to CLEF or to other shared task evaluation campaigns, we highly recommend that a lab workshop first be organised to discuss the format, the problem space, and the practicalities of the shared task.

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.

A call for participation was distributed after CLEF 2011 to solicit proposals for CLEF activities. Twelve creative proposals of very varying scope and domain were submitted by the end of November 2011. A Lab selection committee was formed including:

Jussi Karlgren, Gavagai and SICS (CLEF LOC 2012 co-chair)  
Christa Womser-Hacker, Universität Hildesheim (CLEF LOC 2012 co-chair)  
Paul Clough, The University of Sheffield (CLEF LOC 2011 co-chair)  
Hideo Joho, University of Tsukuba, (NTCIR-10 programme committee chair)  
Jaana Kekäläinen, University of Tampere  
Vanessa Murdock, Yahoo! Research  
Doug Oard, University of Maryland

Eight proposals were accepted as activities into the CLEF 2012 program. In some cases, proposals were judged to be similar to each other, in which case only one of them was accepted.

Seven of the accepted proposals were evaluation labs:

**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. Three tasks were organized:

- Ad-hoc Retrieval Task: to measure information retrieval effectiveness with respect to user input in the form of queries.
- Variability Task: to return diverse objects and resemble the diversity tasks of the Interactive TREC track or the CLEF Image photo tracks.
- Semantic Enrichment Task: retrieval for a query to semantically enrich the query and/or guess the user's information need or original query intent.

**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: starting from TIFF images containing patent scans, to identify the location of the chemical structures depicted on these pages and, for each of them, return the corresponding structure in a chemical structure file format.
- Flowchart Recognition Task: extraction of the information in flowchart images in a predefined textual format.
- Passage retrieval starting from claims: starting from a given claim, retrieval of relevant documents in the collection and mark out the relevant passages in these documents.
- Matching claim to description in a single document (Pilot): starting from the claims of a patent application, indication of the paragraphs in the application's description section that best explain the contents of the given claim.

**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: image modality classification and image retrieval with visual, semantic and mixed topics in several languages, using a data collection from the biomedical literature.
- Photo annotation and retrieval: semantic concept detection and concept-based retrieval using Flickr data, and large-scale annotation using general Web data.
- Personal Photo Annotation
- Plant identification: visual classification of leaf images for the identification of plant species.
- Robot vision: semantic localisation of a mobile robot using multimodal place classification, with special focus on generalization.



**INEX Initiative for the Evaluation of XML Retrieval** has been pioneering structured retrieval since 2002, and joined forces with CLEF running five tasks:

- Social Book Search: study of the value of user-generated descriptions in addition to formal metadata on a collection of Amazon Books and LibraryThing.com data.
- Linked Data: study of ad-hoc search and faceted search on a collection of Linked Data (DBpedia) tied to a large corpus
- Snippet Retrieval: study of the generation of informative snippets with sufficient information to determine the relevancy of search results.
- Relevance Feedback: participants will submit system components (in particular feedback) rather than results.
- Tweet Contextualization: retrieve of synthetic contextual information from Wikipedia in response to a tweet with a URL on a small terminal like a phone.

**PAN Uncovering Plagiarism , Authorship and Social Software Misuse** offered three tasks:

- Plagiarism Detection: features a new plagiarism corpus based on the ClueWeb09, the new search engine ChatNoir which indexes the corpus, the cloud-based algorithm evaluation architecture TIRA and for the first time, real plagiarism cases.
- Author Identification: identification of sexual predators in chat logs and on authorship verification. Moreover, it features for the first time real cases of disputed authorship.
- Quality Flaw Prediction in Wikipedia: newly introduced, and it is about identification of Wikipedia articles that contain certain information quality flaws. It generalizes the vandalism detection task of CLEF2011.

**QA4MRE Question Answering for Machine Reading** evaluates Machine Reading abilities through Question Answering and Reading Comprehension Tests. This lab offers three tasks:

- QA4MRE: reading of single documents and identification of the answers to a set of questions about information that is stated or implied in the text.
- Processing Modality and Negation for Machine Reading (Pilot): evaluation whether systems are able to understand extra-propositional aspects of meaning like modality and negation.
- Machine Reading of Biomedical Texts about Alzheimer (Pilot): setting questions in the biomedical domain with a special focus on the Alzheimer disease.

**RebLab Online Reputation Management** deals 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. Two shared tasks on Twitter data are offered:

- Monitoring task: thematical clusterization of tweets including a company's name as a step towards early alerting on issues that may damage the company's reputation.

- Profiling task: annotation of tweets according to their polarity for reputation (i.e. as to whether their content has positive/negative implications for the company's reputation).

One activity was organised as a lab workshop:

**CLEFeHealth** - Cross-Language Methods, Applications, and Resources for eHealth Document Analysis is a one-day workshop on cross-language evaluation of methods, applications, and resources for eHealth document analysis with a focus on written and spoken NLP.

Besides well-established criteria from previous years' editions of CLEF such as topical relevance, novelty, potential impact on future world affairs, likely number of participants, and the quality of the organising consortium, this year, movement beyond previous years' efforts and connection to real-life usage scenarios was stressed. This is in keeping with work performed in the PROMISE project, which stresses the connection and necessary linkage between the two sides of evaluating information systems: (i) quantitative benchmarking information access systems on the one hand, and (ii) validating hypotheses of usage and application on the other.

Benchmarking has traditionally been the main focus of CLEF and other related evaluation campaigns, but the purposes are to move towards evaluation activities which are comparable across tasks, and to enable validation of the evaluation efforts.

Therefore, this year, the lab proposals were required to address the issue of validation through explicitly stated hypotheses of usage. An evaluation lab should be concrete with respect to situation, context, platform and user preferences for which the suggested evaluation benchmark is valid; a lab workshop should discuss how participants with domain and usage experience and expertise can be recruited to the workshop to provide a grounding of evaluation methodology in application to real-world task. This principle should hopefully lead to wider contact surfaces between evaluation campaigns in CLEF and industrial stakeholders.

Most of the organization of the labs lies in the hands of the individual lab organizers, but the scheduling is determined by lab size and in accordance with the room and time restrictions imposed by the local and lab chairs.

The results of the experiments conducted within CLEF 2012 evaluation campaigns that are run in a yearly circle were presented and discussed as sessions of half a day, one full day or one and a half days at CLEF 2012 on the workshop. 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 Labs workshops were run in parallel sessions distributed over the entire three and a half days of the event interleaving their schedules with the one of the conference.

In addition, a general **poster** session was arranged at the end of the first day (Monday, 17), where participants from all the different Labs had the opportunity to present their works.

This resulted in a lively and interactive discussion among participants and the people attending the event.

Each Lab had its own schedule and programme, but they were almost structured in two main parts: presentation of the main results obtained in the task and reports from the participants discussing the approaches and the techniques used by the systems.

A couple of Labs had also some important invited talks, e.g. PAN hosted two great invited speeches:

- *Babelplagiarism: What can BabelNet do for Cross-language Plagiarism Detection?* given by Roberto Navigli (Univeristy La Sapienza, Rome, Italy). He presented BabelNet, a very large, wide-coverage multilingual semantic network. The resource is automatically constructed by means of a methodology that integrates lexicographic and encyclopedic knowledge from WordNet and Wikipedia. In addition Machine Translation is also applied to enrich the knowledge resource with lexical information for all languages. He presented experiments on new and existing gold-standard datasets to show the high quality and coverage of the resource. In a second set of experiments, he showed that, when provided with a vast amount of high-quality semantic relations, knowledge-rich word sense disambiguation algorithms compete with state-of-the-art supervised WSD systems in a coarse-grained all-words setting and outperform them on gold-standard domain-specific datasets. The second part of the talk was devoted to analyzing cases in which BabelNet could be of help in cross-language plagiarism detection.
- *Cross-lingual Similarity Calculation for Plagiarism Detection and More - Tools and Resources*, given by Ralf Steinberger (European Commission - Joint Research Centre (JRC), Italy) He argues that a system that recognises cross-lingual plagiarism needs to establish - among other things - whether two pieces of text written in different languages are equivalent to each other. There are some cross-lingual functionalities that may well be useful for the plagiarism detection task, i.e. (a) cross-lingual document similarity calculation, (b) subject domain profiling of documents in many different languages according to the same multilingual subject domain categorisation scheme, and (c) the recognition of name spelling variants for the same entity, both within the same language and across different languages and scripts. He presented a number of freely available language resources that can be used to develop software with cross-lingual functionality.

Another exciting talk was held in the QA4MRE workshop:

- *Recognizing Textual Entailment, QA4MRE, and Machine Reading*, Peter Clark (Vulcan Inc. Seattle, USA) declared that Machine Reading remains one of the Grand Challenges of Artificial Intelligence, and also one of the most difficult. Machine Reading requires more than just parsing a text; it requires constructing a coherent

internal model of the world which that text is describing, including inferring facts that are implicit in the text. One basic operation required for this is to infer or recognize Textual Entailment (TE), i.e., recognize plausible implications of a fragment of text. He discussed the overall goal of Machine Reading, and the role of Recognizing Textual Entailment (RTE) within it. He turned to the larger task of Machine Reading in the context of QA4MRE, and discuss two major challenges it elicits, namely the requirement for knowledge, and the requirement for evidential reasoning. Finally he offered some reflections on how these might be addressed, with the aim of moving from sentence-matching strategies towards model-building strategies, and ultimately towards machines that can read.

The ImageCLEF Lab also organized a very interesting **DEMO** session which was much appreciated. The idea is propose it again next year, maybe together with the poster session

Each lab sessions at the workshop contained ample time for general discussion in order to give space to more in-depth exam of the results and engagement by all participants - not just those presenting campaign results and papers. Overall, the workshops provided an ample panorama of the current state-of-the-art and the latest research directions in the multilingual and multimodal information retrieval area.

More information on the CLEF 2012 Labs can be found in the electronic Working Notes which are available online on the CLEF website (<http://clef2012.org/>).

Moreover, a book of abstract with ISBN (978-88-905389-6-4) and ISSN (2038-4726) numbers were prepared and distributed to all the event attendees, together with USB flash drives containing all the full papers of the Labs.

In the last session of the event, lab organizers were asked to briefly report about the main outcomes of the workshop and express intentions and plans for the future.

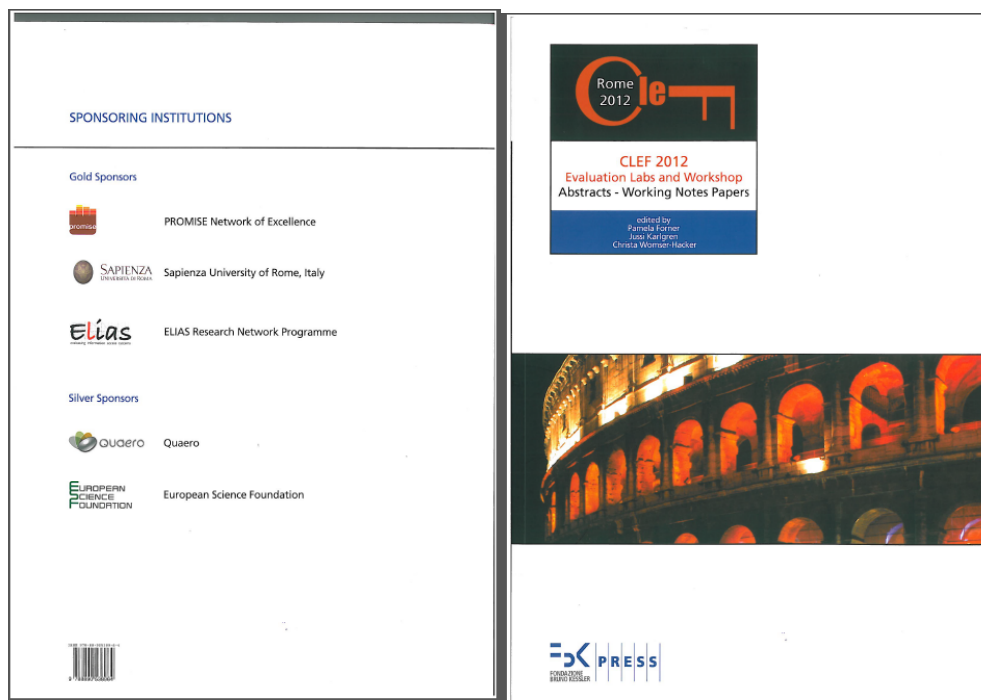


Figure 2: Front and back cover of the CLEF 2012 “Book of Abstracts”

The Labs programmes are given in Appendix E.

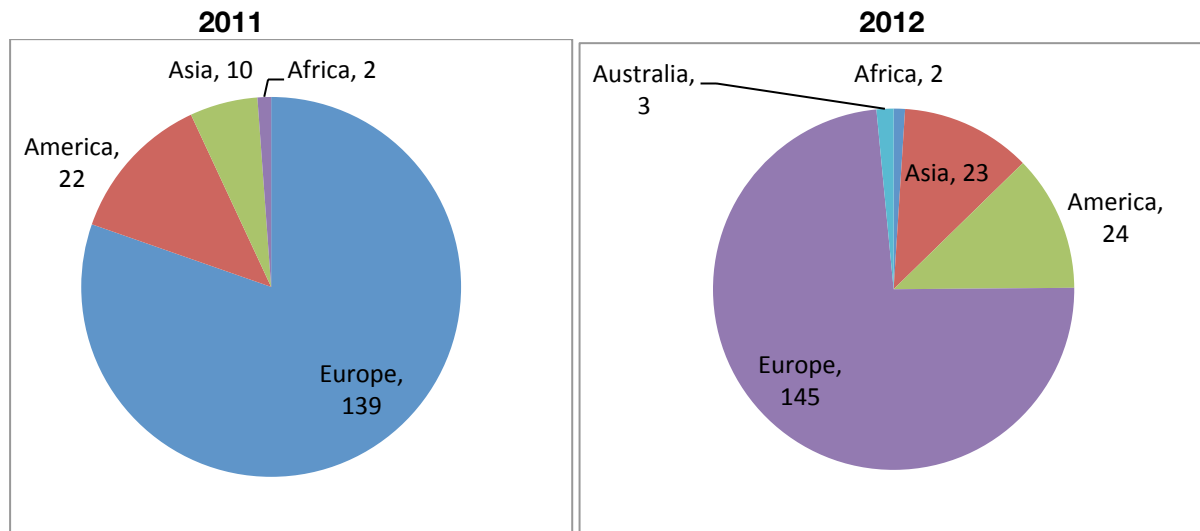
## 5 Participation in CLEF 2012

Overall, CLEF 2012 was attended by almost 200 people from different academic and industrial institutions; the attendance has increased with respect to last year confirming the positive trend started since 2010. The complete list of people attending the event is given in Appendix F. As Figure 3 shows, massive participation is still from Europe, while the other continents remained more or less stable with respect to last years. Although most of the people are still coming from Europe, an increasing interest in CLEF all around the world can be registered, above all from the United States. Important to notice is that, this year, also some participants from Australia were present and, as last year, a small participation from Africa is confirmed, while participants from Asia have almost doubled.

As regards the geographical distribution of the participants according to the country, (Figure 3) most of them came from Germany, France, Spain and Italy and 11 from the United States.

As for the participation in the Labs, more than 280 groups initially registered showing interest in the benchmarking activities proposed this year at CLEF 2012. This proves, once again that CLEF has achieved a high visibility and shows that the labs are still drawing a large participation from the scientific communities in the fields of information retrieval,

natural language processing and multimedia analysis. Multilingual information access is a truly multidisciplinary domain and one of the main objectives of CLEF is to create a forum where groups working in many different areas have a chance to gather together, see what others are doing, and discuss exchanging ideas and comparing results.



**Figure 3: Participation according to the continents**

The registrations to the Labs were gathered via the CELCT infrastructure and a form was made available on the CLEF 2012 website. Registration opened on March 1<sup>st</sup> and was closed on May 20<sup>th</sup>. All registrations were posted on a webpage whose link was made available only to Labs Organizers. In this way, each organizer could have access at any time to the updated status of registrations in real time.

Registrations to the labs were distributed as it is shown in Table 1.

One participant may have registered to more than one Lab. ImageCLEF is still the most popular Lab able to attract many participants not only from Europe but also from the United States and a few from other continents; this year also the PAN lab experienced great popularity and success and was able to reach wide research communities growing considerably with respect to last year.

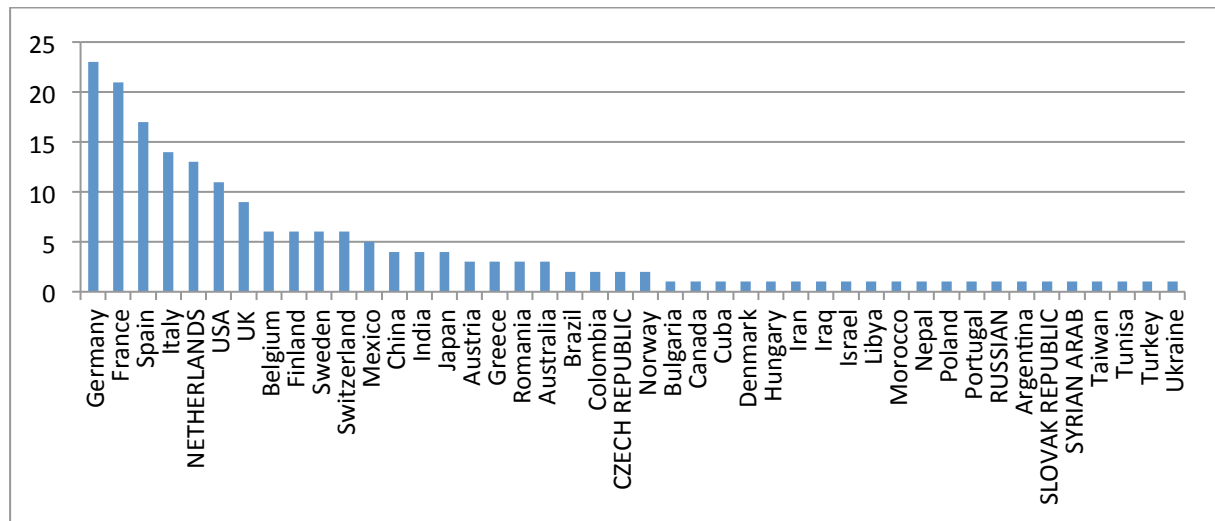


Figure 4: Participation according to the countries

Also the Question Answering for Machine Reading Evaluation attracted a considerable number of registrations and participations: it is a well-established task that continues a long history of question answering tasks at CLEF, but this year two stimulating pilots were also proposed: the Biomedical text about Alzheimer disease task in particular seems to be very promising.

CLEF 2011 LAB	REGISTERED GROUPS	GROUPS SUBMITTING RUNS	N° of Working Notes
IMAGE-CLEF	73	44	44
PAN	45	23	22
CLEF-IP	25	8	8
QA4MRE	25	12	9
LogCLEF	17	4	4
MusiCLEF	4	3	2
CHiC	4		3
<b>TOTAL</b>		<b>94</b>	<b>92</b>

CLEF 2012 LAB	REGISTERED GROUPS	GROUPS SUBMITTING RUNS	N° of Working Notes
IMAGE-CLEF	76	52	50
PAN	74	45	31
CLEF-IP	31	10	10
QA4MRE	38	21	19
CHiC	24	8	5
RepLab	39	11	13
INEX	73	20	20
CLEFeHealth	16	11	16
<b>TOTAL</b>		<b>178</b>	<b>164</b>

Table 1: Registrations, runs and Working Notes in CLEF 2011 and 2012

The CLEFeHealth workshop, newly proposed this year, managed to draw the attention of a considerable number of researches. Its focus is on written and spoken natural-language processing with the use scenario of people using ICT tools to analyse and integrate eHealth



documents across languages, genres, and jargons. The significance of the workshop lies in taking steps toward systematic evaluation and continuous improvement of eHealth ICT. Also to be mentioned this year is the inclusion of the well-known Initiative for Evaluation of XML retrieval which joined efforts with CLEF to target new synergies between multilingual, multimodal and semi-structured information access. Unfortunately, due to the reduced time available for running this year campaign (last year cycle ended on December 2011), the evaluation was still on-going at the time of the conference taking place. So participation data about INEX Lab are still partial.

On the whole, about 180 groups, both researchers and system developers, submitted runs to the different Labs, despite the number of registered participants to the benchmarking activities was much higher than that, as Table 1 shows. It is anyway a very good result also if we compare it to last year: the number of participants taking part to the labs has almost doubled!

A list of groups and indications of the Labs in which they participated is given in the Appendix F to this document.



## 6 CLEF 2012 Web Site and Mobile Version



Figure 5: Home page of the CLEF 2012 Web site.

There is a new website for the CLEF Initiatives<sup>1</sup> which collects all the information regarding the evaluation activities that have been conducted over the first ten years of CLEF from 2000 to 2009 and all the other editions of CLEF – Conference and labs of the Evaluation Forum.

Each year a new website is set up whose purpose is to provide all the necessary information about the event, namely, including information about the CLEF Conference and the CLEF Labs, as well as electronic versions of the Labs Notebook papers. The CLEF websites are used not only to support and promote the visibility of the evaluation activities of the CLEF campaigns, but also to archive their outcomes, and thus act as a vehicle for the dissemination of information and knowledge on multilingual and multimedia information systems.

<sup>1</sup> <http://www.clef-initiative.eu/>

The Web site also contains the electronic version of all the papers submitted to the labs accessible under the “Labs → Working Notes” menu.

Moreover, the slides of the presentations given during the conference are accessible and available for download on the CLEF 2012 Web site.

The publications, both of the conference and of the Labs are indexed by DBPL and linked to the bibliographic social networks BibSonomy.

Like in 2011, a mobile version of the CLEF 2012 website was developed also this year by CELCT containing all useful information for the event like for example the programme, the people giving a presentation, the Labs the sponsor and the map.

CLEF 2012 Conference				Speakers	
17/09	18/09	19/09	20/09	E	
08:00am				Liana Ermakova IRIT, Toulouse	
Registration				F	
09:00am				Edgardo Ferretti Universidad Nacional de San Luis	
Welcome			Loyola	Nicola Ferro University of Padova	
09:30am				Oliver Ferschke Technische Universität Darmstadt	
From Information Retrieval to Knowledgeable Machines			Loyola	Pamela Forner CELCT	
10:30am				Karin Friberg Heppin University of Gothenburg	
Break				G	
11:00am				Alba García Seco de Herrera HES-SO	
Evaluation Methodologies and Infrastructure			Loyola	Shlomo Geva QUT	
12:20am				Lee Gillam University of Surrey	
Lunch				François-Marie Giraud Université Pierre et Marie Curie, UPMC	
2:00pm				Julien Gobeill University of Applied Sciences, Information Studies	
Labs Results Overview			Loyola		
4:00pm					
Break					
4:20pm					
Poster Booster Session					
5:00pm					
Poster Session					
Sessions Speakers Labs Location Road				Sessions Speakers Labs Location Road	

Figure 6: Mobile Version of the CLEF 2012 Web site.

The high visibility of the CLEF websites is evident in the number of visits they attract. CLEF 2010 and 2011 websites had an average of 14,100 visits and 6,200 visitors. These data remain more or less stable also in 2012 Also, much of the visits are registered from Italy but there more and more people from many different countries who are interested in CLEF activities. More details can be found in Appendix I.

## 6.1 Advertising

The CLEF 2010 Conference and Labs (call for labs, call for papers, call for participation) have been announced and advertised on several mailing lists:

- Evaluation Initiatives
  - CLEF (Cross-Language Evaluation Forum community, Europe)
  - TREC (Text Retrieval Conference, USA)
  - TRECvid (TREC Video Retrieval Evaluation, USA)
  - NTCIR (NII Test Collection for IR Systems, Japan)
  - FIRE (Forum for Information Retrieval Evaluation, India)
- Research field news letters
  - SIG-IRList (ACM Special Interest Group on Information Retrieval (SIGIR) moderate newsletter)
  - DBWorld (ACM Special Interest Group on Management of Data (SIGMOD) moderated newsletter)
  - BCS-IRSG list (Information Retrieval Special Interest Group of the BCS, The Chartered Institute for IT, UK)
  - ASIS-L (American Society for Information Science, General List)
  - SIGDL-L (American Society for Information Science, Digital Libraries Interest Group)
  - DBItaly (Italian information management systems – databases, information retrieval systems, digital libraries, - moderated newsletter)
- European Projects
  - Europeana (<http://www.europeana.eu/>)
  - DL.org (<http://www.dlorg.eu/>)
  - DELOS (<http://www.delos.info/>)

A brief report on CLEF 2012 is available in ERCIM News, issue of October 2012, in print at the time of this deliverable. Also, a report will be submitted to SIGIR Forum December issue.

## 7 CLEF 2013 and Beyond

CLEF 2013 will follow a similar format as CLEF 2010, 2011 and 2012, namely

- a series of Evaluation Labs, i.e. *laboratories* to conduct evaluation of information access systems and *workshops* to discuss and pilot innovative evaluation activities;
- a peer-reviewed Conference on a broad range of issues, including
  - investigation continuing the activities of the Evaluation Labs;
  - experiments using multilingual and multimodal data, such as the CLEF data;
  - research in evaluation methodologies and challenges.

The Conference and the annual meetings of the Evaluation Labs will be held as a joint event at the Technical University of Valencia, Spain, 23 - 26 September 2013.

CLEF 2014 will be held at the University of Sheffield, United Kingdom, 15 - 19 September 2014.

During CLEF 2012 in Rome, Italy, it has been decided to re-schedule the overall yearly cycle of CLEF in order to enable having one single publication reporting both the conference papers and the labs overview ready by and delivered at the annual CLEF meeting. The overall schedule for CLEF 2013 will be:

1 October 2012: Call for Lab proposals  
15 October 2012: Deadline for presentation of proposals  
29 October 2012: Notification of acceptance/rejection  
5 November 2012: Release of first lab flyer and call for participation  
1 December 2012: Labs registration opens  
15 December 2012: Labs begin to release data  
15 December 2012 - 1 May 2013: Evaluation Campaign  
1 May 2013: Result submission  
1 June 2013: Results release  
15 June 2013: Working notes papers submitted  
30 June 2013: Lab overview papers submitted in Springer format  
30 June 2013 - 7 July 2013: Lab Chairs review lab overviews  
15 July 2013: Springer proceedings done

For CLEF 2014 the schedule will be:

15 June 2013: Call for Lab Proposals  
1 July 2013: Deadline for presentation of proposals  
20 July 2013: Notification of acceptance/rejection  
CLEF 2013: Release of first lab flyer and call for participation  
1 November 2013: Labs registration opens  
15 November 2013: Labs begin to release data  
15 November 2013 - 1 May 2014: Evaluation Campaign  
1 May 2014: Result submission

1 June 2014: Results release

15 June 2014: Working notes papers submitted

30 June 2014: Lab overview papers submitted in Springer format

30 June 2014 - 7 July 2014: Lab Chairs review lab overviews

15 July 2014: Springer proceedings done

A call for bids to host next editions of CLEF (Conference and labs of the Evaluation Forum) is already open: the Steering Committee of the CLEF Initiative invites interested parties to submit bids to host the CLEF 2015 Conference and Evaluation Labs, held as a joint event of 3-4 days.

Parties interested in hosting the CLEF Evaluation Labs and Conference are invited to submit proposals, in PDF format, by email to the CLEF Steering Committee Chair (chair@clef-initiative.eu) by Friday, April 5, 2013.

## 8 Acknowledgements

We would like to warmly thank all the people who worked hard to make CLEF 2011 possible. We would also like to thank the members of the CLEF Steering Committee who have assisted us with their advice and suggestions.

## References

- [Catarci et al. 2012] T. Catarci, P. Forner, D. Hiemstra, A. Peñas, G. Santucci: Information Access Evaluation. Multilinguality, Multimodality, and Visual Analytics - Third International Conference of the CLEF Initiative, CLEF 2012, Rome, Italy, September 17-20, 2012. (LNCS 7488) Springer 2012
- [Forner et al. 2012] P. Forner, J. Karlgren, C. Womser-Hacker: CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012 CLEF 2012  
<http://www.clef2012.org/index.php?page=Pages/proceedings.php>  
ISBN 978-88-904810-3-1  
ISSN 2038-4963

## Appendix A: CLEF 2012 Organization

CLEF 2012 was organized by Sapienza University of Rome.

### General Chairs

Tiziana Catarci, Sapienza University of Rome, Italy  
Djoerd Hiemstra, University of Twente, The Netherlands

### Program Chairs

Anselmo Peñas, National Distance Learning University (UNED), Spain  
Giuseppe Santucci, Sapienza University of Rome, Italy

### Evaluation Labs Chairs

Jussi Karlgren, Swedish Institute of Computer Science, Sweden  
Christa Womser-Hacker, University of Hildesheim, Germany

### Lab Organising Committee

Jussi Karlgren, Gavagai and SICS (CLEF LOC 2012 co-chair)  
Christa Womser-Hacker, Universität Hildesheim (CLEF LOC 2012 co-chair)  
Paul Clough, The University of Sheffield (CLEF LOC 2011 co-chair)  
Hideo Joho, University of Tsukuba, (NTCIR-10 programme committee chair)  
Jaana Kekäläinen, University of Tampere  
Vanessa Murdock, Yahoo! Research  
Doug Oard, University of Maryland

### Resource 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

*Sapienza University of Rome, Italy:*

Carola Aiello and Giuseppe Santucci

*Consulta Umbria Congressi*

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

Pamela Forner and Giovanni Moretti

## Appendix B: CLEF 2012 Evaluation Labs and Workshop Coordination

It would be impossible to run the CLEF evaluation labs 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 Labs. The following institutions have contributed to the organization of the different Labs and Workshop at CLEF 2012 (alphabetical order):

**Scientific coordination:** Jussi Karlgren (Gavagai, Sweden), Christa Womser-Hacker (University of Hildesheim, Germany)

**Organization:** Emanuele Pianta (CELCT, Trento)

**Web site and Proceedings:** Pamela Forner and Giovanni Moretti (CELCT, Trento)

### Cultural Heritage in CLEF - CHiC

- Berlin School of Library and Information Science, Humboldt-Universität zu Berlin, Germany
- Department of Information Engineering, University of Padova, Italy
- Europeana, The Hague, Netherlands
- Royal School of Library and Information Science, Copenhagen, Denmark
- The Information School, University of Sheffield, United Kingdom

### Information Retrieval in the Intellectual Property Domain - CLEF-IP

- Fraunhofer SCAI, Germany
- Information and Software Engineering Group (IFS), Vienna University of Technology, Austria
- Qatar Computing research Institute (QCRI), Qatar Foundation, Doha, Qatar
- SAIC-Frederick Inc., USA
- School of Computer Science, University of Birmingham, UK

### Image Retrieval in CLEF – ImageCLEF

- CEA LIST, Fontenay-aux-Roses, France
- Harvard Medical School, USA
- Idiap Research Institute, Martigny, Switzerland
- INRA – AMAP, Montpellier, France
- INRIA-IMEDIA, Paris, France
- INRIA-ZENITH, Montpellier, France
- Leiden University, Netherlands
- National Library of Medicine, Bethesda, Maryland, USA
- Nuance Communications, USA
- Universidad Politécnica de Valencia, Spain
- University of Applied Sciences Western Switzerland, Switzerland
- Yahoo! Research, Spain



### **Initiative for the Evaluation of XML Retrieval - INEX**

- IRIT, University of Toulouse, France
- LIA, University of Avignon, France
- LIMSI-CNRS, University Paris-Sud 11, France
- LSIS Aix-Marseille University, France
- Max-Planck-Institut für Informatik, Germany
- Microsoft Research Cambridge, United Kingdom
- Oslo and Akershus University College of Applied Sciences, Norway
- Queensland University of Technology (QUT), Australia
- Renmin University of China, China
- Royal Melbourne Institute of Technology (RMIT), Australia
- Universitat Pompeu Fabra, Spain
- University of Amsterdam, Netherlands
- University of Caen, France
- University of Lugano, Switzerland
- University of Otago, New Zealand

### **Uncovering Plagiarism, Authorship and Social Software Misuse - PAN**

- Bar-Ilan University, Israel
- Duquesne University, Pennsylvania, USA
- Illinois Institute of Technology, Chicago USA
- IRGroup @ Università della Svizzera italiana (University of Lugano), Switzerland
- NLEL @ Universitat Politècnica de València, Spain
- University of the Aegean, Greece
- Webis @ Bauhaus-Universität Weimar, Germany

### **Question Answering for Machine Reading Evaluation - QA4MRE**

- Alexandru Ioan Cuza University, Romania
- Bulgarian Academy of Sciences, Bulgaria
- Center for the Evaluation of Language and Communication Technologies (CELCT), Italy
- Information Sciences Institute of the University of Southern California, USA
- NLP & IR Group, UNED, Spain
- Philips Research North America, USA
- Structural and Computational Biology Group, Spanish National Cancer Research Centre (CNIO), Spain
- University of Antwerp, Belgium
- University of Limerick, Ireland
- University of Saarland, Germany

### **Online Reputation Management Systems - RepLab**

- Llorente & Cuenca, Spain
- NLP & IR Group, UNED, Spain
- University of Amsterdam, Netherlands



### **CLEFeHealth 2012 Workshop**

- Health Business Team (eHealth) and Machine Learning Research Group, NICTA, National ICT Australia

## Appendix C: CLEF 2012 Programme at a Glance

### Monday 17

TIME	
8.00-9.00	Registration
9.00-9.30	Opening - Welcome (Room Loyola)
9.30-10.30	<b>Keynote: Peter Clark, Vulcan Inc., USA (Room Loyola)</b> <i>From Information Retrieval to Knowledgeable Machines</i>
10.30-11.00	Break
11.00-12.20	<b>Evaluation Methodologies and Infrastructure - 4 long papers (Room Loyola)</b>
12.20-14.00	Lunch
14.00-16.00	Labs Results Overview (Room Loyola)
16.00-16.30	Break
16.30-17.00	Poster Booster Session
17.00-18.30	Poster Session (Room Galleria A)
	Welcome Reception

### Tuesday 18

TIME	
9.00-10.10	<b>Benchmarking and Evaluation Initiatives</b> <b>2 long papers; 3 short papers</b> (Room Loyola)
	CLEFeHealth (Room Trilussa)

10.10-10.40	Break Conference/Labs Poster Session			
10.40-12.40	RepLab (Room Leopardi)	INEX (Room Loyola)	QA4MRE (Room Marinetti)	CLEFeHealth (Room Trilussa)
12.40-14.10	Lunch			
14.10-16.10	RepLab (Room Leopardi)	INEX (Room Loyola)	QA4MRE (Room Marinetti)	CLEFeHealth (Room Trilussa)
16.10-16.40	Break Conference/Labs Poster Session			
16.40-18.10	Replab (Room Leopardi)	INEX (Room Loyola)	QA4MRE (Room Marinetti)	CLEFeHealth (Room Trilussa)

## Wednesday 19

### TIME

9.00-10.00	<b>Keynote: Tobias Schreck, Konstanz Univesity, Germany</b> (Room Loyola) <i>Visual Search and Analysis in Textual and Non-Textual Document Repositories - Approaches, Applications, and Research Challenges.</i>			
10.00-10.30	Break Conference/Labs Poster Session			
10.30-12.30	ImageCLEF (Room Loyola)	INEX (Room Leopardi)	PAN (Room Trilussa)	CLEF-IP (Room Marinetti)
12.30-14.00	Lunch			
14.00-16.00	ImageCLEF (Room Loyola)	INEX (Room Leopardi)	PAN (Room Trilussa)	CLEF-IP (Room Marinetti)
16.00-16.30	Break Conference/Labs Poster Session			

16.30-17.50

**Information Access - 3 long papers; 2 short papers**  
(Room Loyola)

Social Dinner

## Thursday 20

TIME				
8.30-9.30	Community Session <b>Other Evaluation Initiatives</b>	Community Session Evaluation CLINIC		
9.30-10.30	ImageCLEF (Room Loyola)	CHiC (Room Trilussa)	PAN (Room Leopardi)	CLEF-IP (Room Marinetti)
10.30-11.00	Break Conference/Labs Poster Session			
11.00-13.00	ImageCLEF (Room Loyola)	CHiC (Room Trilussa)	PAN (Room Leopardi)	CLEF-IP (Room Marinetti)
13.00-14.00	Lunch			
14.00-15.00	ImageCLEF (Room Loyola)	CHiC (Room Trilussa)	PAN (Room Leopardi)	
15.00-15.15	CLEF 2013 Lab Proposals (Room Loyola)			
15.15-15.30	Closing (Room Loyola)			

## Appendix D: CLEF 2012 Accepted papers at the Conference

### Benchmarking and Evaluation Initiatives

Analysis and Refinement of Cross-lingual Entity Linking, Taylor Cassidy, Heng Ji, Hongbo Deng, Jing Zheng, Jiawei Ha

Seven years of INEX interactive retrieval experiments – lessons and challenges, Ragnar Nordlie, and Nils Pharo

Bringing the algorithms to the data: cloud-based benchmarking for medical image analysis, Allan Hanbury, Henning Müller, Georg Langs, Marc Andre Weber, Bjoern H. Menze, and Tomas Salas Fernandez (Short paper)

Going beyond CLEF-IP: The ‘Reality’ for Patent Searchers?, Julia J. Juergens, Preben Hansen, Christa Womser-Hacker (Short paper)

MusiClef: Multimodal Music Tagging Task, Nicola Orio, Cynthia C. S. Liem, Geoffroy Peeters, Markus Schedl (Short paper)

### Information Access

Generating Pseudo Test Collections for Learning to Rank Scientific Articles, Richard Berendsen, Manos Tsagias, Maarten de Rijke, and Edgar Meij

Effects of Language and Topic Size in Patent IR: An Empirical Study, Florina Piroi, Mihai Lupu, Allan Hanbury

Cross-language High Similarity Search using a Conceptual Thesaurus, Parth Gupta, Alberto Barrón-Cedeño, Paolo Rosso

The Appearance of the Giant Component in Descriptor Graphs and Its Application for Descriptor Selection, Anita Keszler, Levente Kovács, Tamas Sziranyi (Short paper)

Hidden Markov Model for Term Weighting in Verbose Queries, Xueliang Yan, Guanglai Gao, Xiangdong Su, Hongxi Wei, Xueliang Zhang, and Qianqian Lu (Short paper)

### Evaluation Methodologies and Infrastructure

DIRECTIONS: Design and Specification of an IR Evaluation Infrastructure, Maristella Agosti, Emanuele Di Buccio, Nicola Ferro, Ivano Masiero, Simone Peruzzo, Gianmaria Silvello

Penalty Functions for Evaluation Measures of Unsegmented Speech Retrieval, Petra Galuscakova, Pavel Pecina, and Jan Hajic  
Cumulated Relative Position: A Metric for Ranking Evaluation, Marco Angelini, Nicola Ferro, Kalervo Jarvelin, Heikki Keskustalo, Ari Pirkola, Giuseppe Santucci, Gianmaria Silvello

Better Than Their Reputation? On the Reliability of Relevance Assessments with Students, Philipp Schaer

### **Posters**

Comparing IR System Components using Beanplots, Jens Kürsten and Maximilian Eibl

Language Independent Query Focused Snippet Generation, Pinaki Bhaskar, and Sivaji Bandyopadhyay

A Test Collection to Evaluate Plagiarism by Missing or Incorrect References, Solange de L. Pertile and Viviane P. Moreira

## Appendix E: CLEF 2012 Labs and Workshop Programme

### CHiC – Thur, September, 20, 2012

09:30 - 10:00 Introduction, General Overview, Results Analysis - Vivien Petras, Humboldt-Universität zu Berlin, Germany

10:00 - 10:30 Creating an Xtrieval Module for Semantic Enrichment - Jens Kürsten, Chemnitz University of Technology, Germany

BREAK

11:00 - 11:30 UniNE at CLEF 2012 - Nada Naji & Mitra Akasereh, University of Neuchatel, Switzerland

11:30 - 12:00 Using Random Walks and Similarity to access Cultural Heritage - Paul Clough, Sheffield University, UK

12:00 - 12:30 Dealing with Sparse Document and Topic Representations - Philipp Schaer, GESIS, Germany

12:30 - 13:00 Query Expansion using Wikipedia and DBpedia - Nitish Aggarwal, DERI, Ireland

BREAK

14:00 - 14:30 Adding interaction to CHiC: Plans for the next Round - Elaine Toms, Sheffield University, UK / Birger Larsen Royal School of Library and Information Science, Copenhagen, Denmark

14:30 - 15:00 Feedback, further Planning - Nicola Ferro, University of Padova, Italy

---

### CLEF IP: Wed, September 19

10:30 - 12:30 Flow-chart and Chemical Structure Recognition Chair: Mihai Lupu

Visual Structure Analysis of Flow Charts in Patent Images, Roland Mörzinger, Joanneum Research Forschungsgesellschaft mbH, Inst. f. Information and Communication Technologies

Textual summarisation of flowcharts in patent drawings for CLEF-IP 2012, Patrice Lopez, Inria

CVC-UAB's participation in the Flowchart Recognition Task of CLEF-IP 2012. Lluís-Pere de las Heras, Computer Vision Center, Universitat Autònoma de Barcelona

Performance of MolRec at CLEF 2012: Overview and Analysis of Results. Nouredin Sadawi. University of Birmingham

14:00 - 16:00 Discussions on the Flow-chart Recognition and on the Chemical Structure Recognition tasks, future image-related tasks

### **CLEF IP: Thu, September 20**

9:30 - 10:30 Passage Retrieval Starting from Claims Chair: Florina Piroi

What makes a baseline a baseline? Tomas Wilhelm, Chemnitz University of Technology, Department of Computer Science

BiTeM site report for the Claims to Passage task in CLEF-IP 2012. Julien Gobeill. University of Applied Sciences, Information Studies, Geneva

Report on the CLEF-IP 2012 Experiments: Search of Topically Organized Patents, Mike Salampasis, Vienna University of Technology, Inst. f. Software Technology and Interactive Systems

11:00 - 13:00

Report on the CLEF-IP 2012 Experiments: Exploring Passage Retrieval with the PIPEXtractor, Linda Andersson, Vienna University of Technology, Inst. f. Software Technology and Interactive Systems

Discussion session on measures, obtaining relevance judgements, etc.

Discussion on future plans, possibilities, etc.



Closing the workshop

---

### **ImageCLEF: Wed, September 19**

10h30-12h30 session chair: Alexis Joly

Overview of the ImageCLEF Flickr Photo Annotation and Retrieval Task , Bart Thomee, Yahoo, Spain

Overview of the ImageCLEF 2012 Scalable Web Image Annotation Task , Mauricio Villegas, University of Valencia, Spain

Overview of the ImageCLEF 2012 Personal Photo Retrieval Task, David Zellhöfer, University of Cottbus, Germany

Overview of the ImageCLEF 2012 Plant Identification Task, Hervé Goëau, INRIA, France

The ImageCLEF 2012 Robot Vision Task: overview and lessons learned, Barbara Caputo, Idiap Research Institute, Switzerland

Overview of the ImageCLEF 2012 medical retrieval and classification tasks, Alba García Seco de Herrera, HES-SO, Switzerland

14h00-16h00 session chair: Bart Thomee

IBM T.J. Watson Research Center Multimedia Analytics: Modality Classification and Case-Based Retrieval tasks of ImageCLEF2012, Michele Merler, IBM, USA

Bioingenium at ImageCLEF 2012: Text and Visual Indexing for Medical Images, Raul Ramos, Universidad Nacional, Colombia

The medGIFT Group in ImageCLEFmed 2012, Alba García Seco de Herrera, HES-SO, Switzerland

Sabanci-Okan at ImageCLEF2012 Plant Identification Competition: Combining Features and Classifiers, Berrin Yanikoglu, Sabanci and Okan Universities, Istanbul, Turkey

A Plant Identification System using Shape and Morphological Features on Segmented Leaflets, Shashwat Mishra, Indian Institute of Technology Kanpur, India

Participation of LSIS/DYNI to ImageCLEF 2012 plant images classification, Sébastien Paris, Aix-Marseille and South Toulon-Var Universities, France

Leveraging robust signatures for mobile robot semantic localization, Jorge Sanchez, University of Cordoba, Argentina

UAIC participation at Robot Vision @ 2012 - An updated vision, Emanuela Boros, Al. I. Cuza University, Romania

### **ImageCLEF: Thur, September 20**

9h30-10h30 session chair: Roberto Paredes

Visual Concepts Detection and Annotation using Histograms of Textual Concepts and Selective Weighted Late Fusion Scheme, Liu Ningning, LIRIS laboratory, Ecole Centrale de Lyon, France

CEA LIST's participation to the Concept Annotation Task of ImageCLEF 2012, Amel Znaidia, Aymen Shabou, Adrian Popescu and Hervé Le Borgne, CEA LIST, Fontenay-aux-Roses, France

IMU @ ImageCLEF 2012, Xueliang Yan, College of Computer Science, Inner Mongolia University Hohhot, China

ISI at ImageCLEF 2012: Scalable System for Image Annotation, Yoshitaka Ushiku, Intelligent Systems and Informatics Lab., the University of Tokyo, Japan

11h00-12h00 session chair: Henning Müller

Presentations of new tasks and discussions on changes requested by participants

Ideas for liver tissue retrieval, Burak Acar, Technical University of Munich, Germany

12h00-13h00 plenary demo session of live systems: session chair: David Zellhöfer

The session will contain an oral presentation of each system.

Group name: LIRIS - ReVeS

Demo name: Folia - A Mobile Application for Tree Species Identification

Person presenting: Guillaume Cerutti - Laure Tougne

Group name: Inria & PlantNet, France

Demo name: PlantNet Identify

Person presenting: Hervé Goëau, Pierre Bonnet

Group name: DBIS, Germany  
Demo name: Pythia MIR system  
Person presenting: David Zellhöfer

Group name: DEMIR, Turkey  
Demo name: The DEMIR system in ImageCLEF 2012  
Person presenting: Adil Alpkocak

14h00-15h00 session chair: Barbara Caputo

ImageCLEF 2013, decisions and directions

---

### **INEX: Wed Tue, September, 18**

10.40-12.40: INEX Social Book Search Session:

Overview of the INEX 2012 Social Book Search Track, Marijn Koolen, University of Amsterdam, The Netherlands

RSLIS at INEX 2012: Social Book Search Track, Toine Bogers, RSLIS Copenhagen, Denmark

Do Social Information Help Book Search?, Romain Deveaud, University of Avignon, France

Using Collaborative Filtering in Social Book Search, Hugo Huurdeman, University of Amsterdam, The Netherlands

OUC's Participation in the 2012 INEX Book and Linked-Data Tracks, Michael Preminger, Oslo University College, Norway

14.10-16.10: INEX Linked Data Session:

Overview of the INEX 2012 Linked Data Track, Qiuyue Wang, Renmin University of China

ENSM-SE at INEX 2012: Basic Experiments, Michel Beigbeder, EMSE, Saint Etienne, France

Running SPARQL-Fulltext Queries Inside a Relational DBMS, Sairam Gurajada, MPI, Saarbruecken, Germany

Integrated Retrieval over Structured and Unstructured Data, Qiuyue Wang, Renmin University of China

16.40-18.40: INEX Relevance Feedback and Snippet Retrieval Session:

Overview of the INEX 2012 Relevance Feedback Track, Shlomo Geva, QUT, Australia

UAM at INEX 2012 Relevance Feedback Track: Using a Probabilistic Method for Ranking Refinement, Esau Villatoro-Tello, UAM Mexico

Overview of the INEX 2012 Snippet Retrieval Track, Shlomo Geva, QUT, Australia

### **INEX Wed, September 19**

10.30-12.30: INEX Tweet Contextualization Session

Overview of the INEX 2012 Tweet Contextualization Track, Eric Sanjuan, University of Avignon, France

A Hybrid Tweet Contextualization System using IR and Summarization, Pinaki Bhaskar, Jadavpur University, Kolkata, India

LIA/LINA at the INEX 2012 Tweet Contextualization track, Romain Deveaud, University of Avignon, France

IRIT at INEX 2012: Tweet Contextualization, Liana Ermakova, IRIT, Toulouse, France

DCU@INEX-2012: Exploring Sentence Retrieval for Tweet Contextualization, Johannes Leveling, DCU, Ireland

INEX 2012 Benchmark A Semantic Space for Tweets Contextualization, Mohamed Morchid, University of Avignon, France

14.00-16.00: INEX Planning Session

---

### **PAN – Wed September, 19**

Quality Flaw Prediction in Wikipedia, Chair: Matthias Hagen

Overview of the 1st International Competition on Quality Flaw Prediction in Wikipedia, Maik Anderka, Benno Stein

On the Use of PU Learning for Quality Flaw Prediction in Wikipedia, Edgardo Ferretti, Donato Hernández Fusilier, Rafael Guzmán Cabrera, Manuel Montes-y-Gómez, Marcelo Errecalde, Paolo Rosso

FlawFinder: A Modular System for Predicting Quality Flaws in Wikipedia, Oliver Ferschke, Iryna Gurevych, Marc Rittberger

Plagiarism Detection, Chair: Matthias Hagen

Overview of the 4th International Competition on Plagiarism Detection, Martin Potthast, Tim Gollub, Matthias Hagen, Jan Graßegger, Johannes Kiesel, Maximilian Michel, Arnd Oberländer, Martin Tippmann, Alberto Barrón-Cedeño, Parth Gupta, Paolo Rosso, Benno Stein

Approaches for Candidate Document Retrieval and Detailed Comparison of Plagiarism Detection, Kong Leilei, Qi Haoliang, Wang Shuai, Du Cuixia, Wang Suhong, Han Yong

Educated guesses and equality judgements: using search engines and pairwise match for external plagiarism detection, Lee Gillam, Neil Newbold, Neil Cooke

Lunch

Plagiarism Detection, Chair: Paolo Rosso

Three way search engine queries with multi-feature document comparison for plagiarism detection, Simon Suchomel, Jan Kasprzak, and Michal Brandejs

Cross-Language Plagiarism Detection (Keynotes and Panel Discussion), Chair: Paolo Rosso

Cross-lingual Similarity Calculation for Plagiarism Detection and More - Tools and Resources, Ralf Steinberger

Babelplagiarism: What can BabelNet do for Cross-language Plagiarism Detection?, Roberto Navigli

Panel discussion

## **PAN, Thur, September 20**

Traditional Authorship Attribution, Chair: Efstathios Stamatatos

An Overview of the Traditional Authorship Attribution Subtask + Mixture of Experts Authorship Attribution, Patrick Juola and Michael Ryan, John Noecker Jr

Authorship attribution: using rich linguistic features when training data is scarcen, Ludovic Tanguy, Franck Sajous, Basilio Calderone, Nabil Hathout

Feature Bagging for Author Attribution, François-Marie Giraud, Thierry Artières

Break

Traditional Authorship Attribution & Sexual Predator Identification, Chair: Patrick Juola

Graph-based and Lexical-Syntactic Approaches for the Authorship Attribution Task, Esteban Castillo, Darnes Vilariño, David Pinto, Iván Olmos, Jesús A. González, Maya Carrillos

Overview of the International Sexual Predator Identification Competition at PAN-2012, Giacomo Inches, Fabio Crestani

Vote/Veto Classification, Ensemble Clustering and Sequence Classification for Author Identification, Roman Kern, Stefan Klampfl and Mario Zechner

Quite Simple Approaches for Authorship Attribution, Intrinsic Plagiarism Detection and Sexual Predator Identification, Anna Vartapetian, Lee Gillam

Kernel Methods and String Kernels for Authorship Analysis, Marius Popescu, Cristian Grozea

Conversation Level Constraints on Pedophile Detection in Chat Rooms, Claudia Peersman, Frederik Vaassen, Vincent Van Asch, Walter Daelemans

Identifying Predators Using ChatCoder 2.0, April Kontostathis, Will West, Andy Garron, Kelly Reynolds, Lynne Edwards

Lunch

Sexual Predator Identification, Chair: Giacomo Inches

A Two-step Approach for Effective Detection of Misbehaving Users in Chats, Esaú Villatoro-Tello, Antonio Juárez-González, Hugo Jair Escalante, Manuel Montes-y-Gómez, and Luis Villaseñor-Pineda

A Learning-Based Approach for the Identification of Sexual Predators in Chat Logs, Javier Parapar, David E. Losada, Alvaro Barreiro

Features for modelling characteristics of conversations, Gunnar Eriksson, Jussi Karlgren

Identifying Sexual Predators by SVM Classification with Lexical and Behavioral Features,  
Colin Morris, Graeme Hirst

---

## **QA4MRE – Tue, September 18**

Welcome Anselmo Peñas, UNED, Spain

Invited talk – Recognizing Textual Entailment, QA4MRE, and Machine Reading, Peter Clark,  
Vulcan Inc, USA

Overview of QA4MRE 2012 Main task, Pamela Forner, CELCT, Italy

Machine Reading of Biomedical Texts about Alzheimer, Roser Morante, University of  
Antwerp, Belgium

Processing Modality and Negation, Roser Morante, University of Antwerp, Belgium

Lunch

Reports from the Participants

Cross-Language Answer Validation, Ivelina Nikolova; Ontotex AD, Bulgaria

Testing Lexical Approaches in QA4MRE, Luísa Coheur, L2F/INESC-ID, Portugal

Bulgarian Question Answering for Machine Reading, Ivelina Nikolova

Index Expansion for Machine Reading and Question Answering, Giuseppe Attardi,  
University of Pisa, Italy

Question Answering for Alzheimer Disease using Information Retrieval, Luca Toldo, Merck  
KGaA, Germany

Adaptation of LIMSI's QALC for QA4MRE, Brigitte Grau, LIMSI, France

A Text Mining Approach as Baseline for QA4MRE' 12, Mathias Verbeke, Katholieke  
Universiteit Leuven, Belgium

Break

Simple similarity-based Question Answering Strategies for Biomedical Text, David Martinez,  
National ICT Australia



An Approach for Detecting Modality and Negation in Texts by using Rule-based Techniques,  
Daniel Sánchez-Cisneros, Universidad Carlos III de Madrid, Spain

Breakout session Chair: Anselmo Peñas, UNED, Spain

- planning for next year
- discussion with Participants

---

### **Replab – Tue, September, 18**

Meet the user, Adolfo Corujo (Llorente & Cuenca)

Profiling overview

José Manuel González Chenlo, FBM-Yahoo! at RepLab 2012

Fredrik Olsson, Gavagai at RepLab 2012

Maria-Hendrike Peetz, UvA at RepLab 2012

Julio Villena Román, Daedalus at RepLab 2012

Lunch

Interactive topic analysis

Monitoring overview

Tamara Martín, UNED at RepLab 2012

Younus Arjumand, DERI at RepLab 2012

RepLab lessons learned for 2012

Break

Plans for 2013

Discussion

---

### **CLEFeHealth – Tue, September 18**

Chair: Senior Researcher, PhD Hanna Suominen (NICTA National ICT Australia and Australian National University)

Keynote of Hanna Suominen (NICTA National ICT Australia and Australian National University): Towards International Privacy-preserving Benchmarks of eHealth Technologies

Poster highlights

Sanna Salanterä (University of Turku): An Ontology to Improve Accessibility and Quality of Patient Instructions

Karin Friberg Heppin and Anni Järvelin (University of Gothenburg): Towards Improving Search Results for Medical Experts and Laypersons

Sumithra Velupillai (Stockholm University): Initial Results in the Development of SCAN: a Swedish Clinical Abbreviation Normalizer

Mattias Kanhov (Stockholm University): Natural Language Generation from SNOMED Specifications

Allan Hanbury (Vienna University of Technology): Considering Subjects and Scenarios in Large-scale User-centered Evaluation of a Multilingual Multimodal Medical Search System

Leif Hanlen (NICTA and The Australian National University): Towards Ease of Building Legos in Assessing eHealth Language Technologies: A RESTful Laboratory for Data and Software

10:40 - 12:40 Session 2

Chair: Senior Researcher, PhD Hanna Suominen (NICTA National ICT Australia and Australian National University)

Poster highlights

Karl Kreiner (AIT, Austrian Institute of Technology GmbH): On the Use of Text Messaging in a Diabetes Telehealth System - Results and Evaluation of a Content Analysis

Riitta Danielsson-Ojala (University of Turku and Southwest Hospital District): Vocabulary in Discharge Documents - The Patient's Perspective

David Martinez (NICTA and The University of Melbourne): Biosurveillance for Invasive Fungal Infections via Text Mining

Hans Moen (Norwegian University of Science and Technology): Towards Retrieving and Ranking Clinical Recommendations Using Cross-lingual Random Indexing

Hanna Suominen (NICTA and The Australian National University): Clinical Speech to Text: Evaluation Setting

Keynote of Wendy W Chapman (University of California): Developing Resources to Assist in Development and Application of NLP to Clinical Texts

Panel with position presentations

Gareth Jones (Dublin City University): Multilingual Evaluation

Jussi Karlgren (Gavagai): Bridging the Gap between Experiences in the Health System and Life at Large

Karl Kreiner (AIT Austrian Institute of Technology): Towards the Next Generation of Collaborative Biomedical Research

Henning Müller (University of Applied Sciences Western Switzerland, HES-SO): The Role of Images for Information Retrieval in the Patient Record

Sanna Salanterä (University of Turku): Clinical Decision-making and Technological Support

14:10 - 16:10 Session 3

Chair: Principal Researcher, Director of Health, PhD Leif Hanlen (NICTA National ICT Australia and Australian National University)

Keynote of Leif Hanlen (NICTA National ICT Australia and Australian National University): National eHealth Living Lab: Removing the “E” from E-Health

Student mentoring: Q&A of students and senior academics in the CLEFeHealth2012

Building partnerships (Trilussa) / Student meeting

16:10-16:40: Posters – Final Chance to Discuss with the Poster Presenters

Sanna Salanterä (University of Turku): An Ontology to Improve Accessibility and Quality of Patient Instructions

Karin Friberg Heppin and Anni Järvelin (University of Gothenburg): Towards Improving Search Results for Medical Experts and Laypersons

Sumithra Velupillai (Stockholm University): Initial Results in the Development of SCAN: a Swedish Clinical Abbreviation Normalizer

Mattias Kanhov (Stockholm University): Natural Language Generation from SNOMED Specifications

Allan Hanbury (Vienna University of Technology): Considering Subjects and Scenarios in Large-scale User-centered Evaluation of a Multilingual Multimodal Medical Search System

Karl Kreiner (AIT, Austrian Institute of Technology GmbH): On the Use of Text Messaging in a Diabetes Telehealth System – Results and Evaluation of a Content Analysis

Riitta Danielsson-Ojala (University of Turku and Southwest Hospital District): Vocabulary in Discharge Documents – The Patient's Perspective

David Martinez (NICTA and The University of Melbourne): Biosurveillance for Invasive Fungal Infections via Text Mining

Hans Moen (Norwegian University of Science and Technology): Towards Retrieving and Ranking Clinical Recommendations Using Cross-lingual Random Indexing

Hanna Suominen (NICTA and The Australian National University): Clinical Speech to Text: Evaluation Setting

Leif Hanlen (NICTA and The Australian National University): Towards Ease of Building Legos in Assessing eHealth Language Technologies: A RESTful Laboratory for Data and Software

16:40-18:10: Session 4

Chair: Senior Researcher, PhD Hanna Suominen (NICTA National ICT Australia and Australian National University)

Workshop: Developing the roadmap for eHealth evaluation

## Appendix F: List of people attending the Event

SURNAME	NAME	AFFILIATION	COUNTRY
ABDULAHHAD	KARAM	CNRS LIG	SYRIAN ARAB REPUBLIC
ADRIAENSSEN	MARC	EUROPEAN PATENT OFFICE	NETHERLANDS
AGGARWAL	NITISH	DERI, NUI GALWAY	INDIA
AKASEREH	MITRA	University of neuchatel	SWITZERLAND
AMIGÓ	ENRIQUE	UNED	SPAIN
ARJUMAND		UNIVERSITY OF MILANO BICOCCA	ITALY
ATIF QURESHI	MUHAMMAD		
ATTARDI	GIUSEPPE	UNIVERSITY OF PISA	ITALY
BARROU	DIALLO	EUROPEAN PATENT OFFICE	FRANCE
BEIGBEDER	MICHEL	ECOLE DES MINES DE SAINT-ÉTIENNE	FRANCE
BERENDSEN	RICHARD	UNIVERSITY OF AMSTERDAM	NETHERLANDS
BERRIN	YANIKOGLU	SABANCI UNIVERSITY	TURKEY
BOGERS	TOINE	ROYAL SCHOOL OF LIBRARY AND INFORMATION SCIENCE	NETHERLANDS
BONNET	PIERRE	CIRAD	FRANCE
BOUZOUBAA	KARIM	MOHAMED 5TH UNIVERSITY	MOROCCO
BRASCHLER	MARTIN	ZURICH UNIVERSITY OF APPLIED SCIENCES	SWITZERLAND
BRIGITTE	GRAU	LIMSI-CNRS	FRANCE
BURAK	ACAR	TU MUNICH	GERMANY
CALDERONE	BASILIO	CLLE-ERSS (UMR 5263) CNRS & UNIVERSITÉ DE TOULOUSE-LE MIRAIL	FRANCE
CAPUTO	BARBARA	IDIAP RESEARCH INSTITUTE	SWITZERLAND
CARRILLO DE ALBORNOZ	JORGE	UNED	SPAIN
CASSIDY	TAYLOR	THE GRADUATE CENTER (CUNY)	UNITED STATES
CASTELLANOS	ANGEL	UNED	SPAIN
CASTILLO JUAVEZ	ESTEBAN	UNIVERSIDAD AUTONOMA DE PUEBLA	MEXICO
CATARCI	TIZIANA	SAPIENZA - UNIVERSITA' DI ROMA	ITALY
CERUTTI	GUILLAUME	UNIVERSITÉ LYON 2	FRANCE
CHAPMAN	WENDY	University of California at San Diego	UNITED STATES
CHIEN	BEEN-CHIAN	NATIONAL UNIVERSITY OF TAINAN	TAIWAN, PROVINCE OF CHINA

CLARK	PETER	VULCAN INC.	UNITED STATES
CLOUGH	PAUL	UNIVERSITY OF SHEFFIELD	UNITED KINGDOM
COHEUR	LUISA	INESC-ID/IST	PORTUGAL
DAELEMANS	WALTER	UNIVERSITY OF ANTWERP	BELGIUM
DANIELSSON-OJALA	RIITTA	UNIVERSITY OF TURKU	FINLAND
DE LAS HERAS	LLUIS PERE	COMPUTER VISION CENTER	SPAIN
DE RIJKE	MAARTEN	UNIVERSITY OF AMSTERDAM	NETHERLANDS
DE VRIES	CHRIS	Queensland University of Technology	AUSTRALIA
DELPORT	BERNARD	EUROPEAN PATENT OFFICE	NETHERLANDS
DEVEAUD	ROMAIN	UNIVERSITY OF AVIGNON	FRANCE
DORR	BONNIE	DARPA	UNITED STATES
EGGEL	IVAN	HES-SO	SWITZERLAND
ERIKSSON	GUNNAR	GAVAGAI AB	SWEDEN
ERMAKOVA	LIANA	PERM STATE UNIVERSITY	RUSSIAN FEDERATION
FERRO	NICOLA	UNIVERSITY OF PADUA	ITALY
FONTES DE AVILA	SANDRA ELIZA	FEDERAL UNIVERSITY OF MINAS GERAIS (UFMG) & PIERRE AND MARIE CURIE UNIVERSITY (UPMC)	BRAZIL
FORNER	PAMELA	CELCT	ITALY
FRANCK	SAJOUS	CLLE-ERSS	FRANCE
FRIBERG HEPPIN	KARIN	UNIVERSITY OF GOTHENBURG	SWEDEN
GAEDE	MARIA	HUMBOLDT-UNIVERSITÄT ZU BERLIN / BERLIN SCHOOL OF LIBRARY AND INFORMATION SCIENCE	GERMANY
GALLO	IGNAZIO	UNIVERSITY OF INSURBIA	ITALY
GALUSCAKOVA	PETRA	INSTITUTE OF FORMAL AND APPLIED LINGUISTICS, CHARLES UNIVERSITY IN PRAGUE	SLOVAKIA (SLOVAK REPUBLIC)
GARCIA SECO DE HERRERA	ALBA	HES-SO	SPAIN
GARCIA-SERRANO	ANA	UNED	SPAIN
GEVA	SHLOMO	QUEENSLAND UNIVERSITY OF TECHNOLOGY	AUSTRALIA
GILBERT	ANDREW	UNIVERSITY OF SURREY	UNITED KINGDOM
GILLAM	LEE	UNIVERSITY OF SURREY	UNITED KINGDOM
GIRAUD	FRANÇOIS-MARIE	UPMC - LIP6	FRANCE
GOBEILL	JULIEN	UNIVERSITY OF APPLIED SCIENCES, GENEVA	FRANCE

GOËAU	HERVÉ	INRIA	FRANCE
GOLLUB	TIM	BAUHAUS-UNIVERSITÄT WEIMAR	GERMANY
GONZÁLEZ CHENLO	JOSE MANUEL	FUNDACIO BARCELONA MEDIA	SPAIN
GROZEA	CRISTIAN	FRAUNHOFER	ROMANIA
GUPTA	PARTH	UPV, SPAIN	INDIA
HAGEN	MATTHIAS	BAUHAUS-UNIVERSITY WEIMAR	GERMANY
HANSEN	PREBEN	SICS	SWEDEN
HARADA	TATSUYA	THE UNIVERSITY OF TOKYO	JAPAN
HARMAN	DONNA	NIST	UNITED STATES
HERNÁNDEZ FUSILIER	DONATO	UNIVERSIDAD DE GUANAJUATO	MEXICO
HIEMSTRA	DJOERD	UNIVERSITY OF TWENTE	NETHERLANDS
HIGUCHI	TAKAYUKI	TOKYO UNIV	JAPAN
HIRST	GRAEME	UNIVERSITY OF TORONTO	AUSTRALIA
HUURDEMAN	HUGO	UNIVERSITY OF AMSTERDAM	NETHERLANDS
INCHES	GIACOMO	University of Lugano (USI)	ITALY
JACQUES	SAVOY	UNIVERSITY OF NEUCHATEL	SWITZERLAND
JARVELIN	KAL	SCHOOL OF INFORMATION SCIENCES	FINLAND
JÄRVELIN	ANNI	SICS	FINLAND
JOLY	ALEXIS	INRIA	FRANCE
JONES	GARETH	DUBLIN CITY UNIVERSITY	UNITED KINGDOM
JUOLA	PATRICK	DUQUESNE UNIVERSITY	UNITED STATES
KALAMBOUKIS	THEODORE	ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS	GREECE
KAMPS	JAAP	UNIVERSITY OF AMSTERDAM	NETHERLANDS
KANDO	NORIKO	NATIONAL INSTITUTE OF INFORMATICS	JAPAN
KANHOV	MATTIAS	STOCKHOLM UNIVERSITY	UNITED KINGDOM
KARLGREN	JUSSI	GAVAGAI	SWEDEN
KASPRZAK	JAN	MASARYK UNIVERSITY	CZECH REPUBLIC
KERN	ROMAN	KNOW-CENTER	AUSTRIA
KLEINEBERG	MICHAEL	BERLIN SCHOOL OF LIBRARY AND INFORMATION SCIENCE (HUMBOLDT-UNIVERSITY)	GERMANY
KONTOSTATHIS	APRIL	URSINUS COLLEGE	UNITED STATES
KOOLEN	MARIJN	UNIVERSITY OF AMSTERDAM	NETHERLANDS
KREINER	KARL	AIT Austrian Institute of Technology GmbH	AUSTRIA
KUEPPERS	ROBIN	UNIVERSITY OF DUSSELDORF	GERMANY



KUERSTEN	JENS	CHEMNITZ UNIVERSITY OF TECHNOLOGY	GERMANY
LANA-SERRANO	SARA	POLYTECHNIC UNIVERSITY OF MADRID	SPAIN
LARSEN	BIRGER	ROYAL SCHOOL OF LIBRARY AND INFORMATION SCIENCE	DENMARK
LEVELING	JOHANNES	DUBLIN CITY UNIVERSITY	GERMANY
LIU	NINGNING	LIRIS	CHINA
LOPEZ	PATRICE	INRIA	FRANCE
LUNDGREN-LAINE	HELJA	UNIVERSITY OF TURKU	FINLAND
MANTZIOU	ELENI	CENTER FOR RESEARCH AND TECHNOLOGY HELLAS	GREECE
MARTIN WANTON	TAMARA	UNED	CUBA
MARTINEZ	DAVID	NICTA	SPAIN
MASSEY	DAVID	OSLO UNIVERSITY COLLEGE	UNITED KINGDOM
MEIJ	EDGAR	UNIVERSITY OF AMSTERDAM	NETHERLANDS
MERLER	MICHELE	IBM TJ WATSON	UNITED STATES
MISHRA	SHASHWAT	INDIAN INSTITUTE OF TECHNOLOGY	INDIA
MISHRA	ARUNAV	MPI	INDIA
MOEN	HANS	NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY	NORWAY
MONTES Y GOMEZ	MANUEL	INSTITUTO NACIONAL DE ASTROFISICA, OPTICA Y ELECTRONICA	MEXICO
MORANTE	ROSER	UNIVERSITEIT ANTWERPEN	BELGIUM
MORCHID	MOHAMED	UNIVERSITY OF AVIGNON	FRANCE
MORETTI	GIOVANNI	CELCT	ITALY
MORRIS	COLIN	UNIVERSITY OF TORONTO	CANADA
MÖRZINGER	ROLAND	JOANNEUM RESEARCH FORSCHUNGSGESELLSCHAFT MBH	AUSTRIA
MOTHE	JOSIANE	IRIT	FRANCE
MÜLLER	HENNING	HES-SO	GERMANY
NAJI	NADA	UNIVERSITY OF NEUCHATEL	IRAQ
NAVIGLI	ROBERTO	SAPIENZA - UNIVERSITA' DI ROMA	ITALY
NIKOLOVA	IVELINA	ONTOTEXT AD	BULGARIA
NORDLIE	RAGNAR	OSLO AND AKERSHUS UNIVERSITY COLLEGE	NORWAY
OLIVER	FERSCHKE	UKP LAB, TECHNISCHE UNIVERSITÄT DARMSTADT	GERMANY
OLSSON	FREDRIK		SWEDEN
ORIO	NICOLA	UNIVERSITY OF PADOVA	ITALY

PALKOVSKII	YURII	ZHYTOMYR STATE UNIVERSITY	UKRAINE
PARAPAR	JAVIER	UNIVERSITY OF A CORUÑA	SPAIN
PEREDES	ROBERTO	DSIC-UPV CAMINO DE VERA S/N	SPAIN
PARIS	SEBASTIEN	AIX-MARSEILLE UNIVERSITY	FRANCE
PATRICE	BELLOT	AIX-MARSEILLE UNIVERSITÉ	FRANCE
PEERSMAN	CLAUDIA	ANTWERP UNIVERSITY	BELGIUM
PEETZ	MARIA-HENDRIKE	UNIVERSITY OF AMSTERDAM	GERMANY
PEÑAS	ANSELMO	UNED	SPAIN
PERTILE	SOLANGE DE LURDES	UFRGS	BRAZIL
PETERS	CAROL ANN	ISTI-CNR	UNITED KINGDOM
PETRAS	VIVIEN	HUMBOLDT-UNIVERSITÄT ZU BERLIN / BERLIN SCHOOL OF LIBRARY AND INFORMATION SCIENCE	GERMANY
PIANTA	EMANUELE	CELCT	SWITZERLAND
PIRAS	LUCA	Dipartimento di Ingegneria Elettrica ed Elettronica dell'Universit di Cagliari	ITALY
PIROI	FLORINA	VIENNA UNIVERSITY OF TECHNOLOGY, ISIS INSTITUTE	ROMANIA
POPESCU	ADRIAN	CEA	ROMANIA
POTTHAST	MARTIN	BAUHAUS-UNIVERSITÄT WEIMAR	GERMANY
PREMINGER	MICHAEL	OSLO AND AKERSHUS UNIVERSITY COLLEGE OF APPLIED SCIENCE	ISRAEL
QURESHI	MUHAMMAD ATIF	UNIVERSITY OF MILANO BICOCCA	ITALY
RAMOS POLLAN	RAUL	UNAL	COLOMBIA
REYNOLDS	KELLY	LEHIGH UNIVERSITY	UNITED STATES
RISCHKA	MAGDALENA	HEINRICH HEINE UNIVERSITY DUESSELDORF	POLAND
ROSSO	PAOLO	UNIVERSIDAD POLITECNICA DE VALENCIA	SPAIN
SADAWI	NOUREDDIN	UNIVERSITY OF BIRMINGHAM	LIBYAN ARAB JAMAHIRIYA
SALAMPASIS	MICHAIL	VIENNA UNIVERSITY OF TECHNOLOGY	GERMANY
SALANTERÄ	SANNA	UNIVRSITY OF TURKU	FINLAND
SANCHEZ	CHRISTIAN	UNIVERSIDAD AUTONOMA METROPOLITANA	MEXICO
SANCHEZ	JORGE ADRIAN	UNIVERSIDAD TECNOLÓGICA	ARGENTINA

		NACIONAL	
SÁNCHEZ CISNEROS	DANIEL	UNIVERSIDAD CARLOS III DE MADRID	SPAIN
SÁNCHEZ-ORO CALVO	JESÚS	UNIVERSIDAD REY JUAN CARLOS	SPAIN
SANJUAN	ERIC	LIA/UNIVERSITÉ D'AVIGNON ET DES PAYS DE VAUCLUSE	FRANCE
SANTUCCI	GIUSEPPE	UNIVERSITY OF ROME LA SAPIENZA	ITALY
SAPKOTA	UPENDRA	UNIVERSITY OF ALABAMA AT BIRMINGHAM	NEPAL
SCHAER	PHILIPP	GESIS	GERMANY
SCHENKEL	RALF	SAARLAND UNIVERSITY	GERMANY
SCHMIDT	CHRISTOPH	BRANDENBURG TECHNICAL UNIVERSITY	GERMANY
SCHRECK	TOBIAS	UNIVERSITÄT KONSTANZ	GERMANY
SCHUTH	ANNE	UNIVERSITY OF AMSTERDAM	NETHERLANDS
SILVELLO	GIANMARIA	UNIVERSITY OF PADUA	ITALY
SPINA	DAMIANO	UNED	ARGENTINA
STAMATATOS	EFSTATHIOS	UNIVERSITY OF THE AEGEAN	GREECE
STEIN	BENNO	BAUHAUS-UNIVERSITÄT WEIMAR	GERMANY
STEINBERGER	RALF	EUROPEAN COMMISSION - JOINT RESEARCH CENTRE	NETHERLANDS
STILLER	JULIANE	BERLIN SCHOOL OF LIBRARY AND INFORMATION SCIENCE	GERMANY
SUCHOMEL	ŠIMON	FI MU	CZECH REPUBLIC
SUOMINEN	HANNA SUOMINEN	NICTA	FINLAND
SZIRÁNYI	TAMÁS	MTA SZTAKI	HUNGARY
TANGUY	LUDOVIC	UNIVERSITY OF TOULOUSE	FRANCE
TANNIER	XAVIER	LIMSI-CNRS, UNIV. PARIS-SUD	FRANCE
THOMEE	BART	YAHOO! RESEARCH	UNITED STATES
TOMS	ELAINE	UNIVERSITY OF SHEFFIELD	UNITED KINGDOM
TOUGNE	LAURE	UNIVERSITÉ LUMIÈRE LYON 2	FRANCE
USHIKU	YOSHITAKA	UNIVERSITY OF TOKYO	JAPAN
VAASSEN	FREDERIK	UNIVERSITY OF ANTWERP	BELGIUM
VAN ASCH	VINCENT	CLIPS - UNIVERSITY OF ANTWERP	BELGIUM
VARTAPETIANCE	ANNA	UNIVERSITY OF SURREY	IRAN (ISLAMIC REPUBLIC OF)
VEENMAN	COR		NETHERLANDS
VELUPILLAI	SUMITHRA	STOCKHOLM UNIVERSITY	SWEDEN
VERBEKE	MATHIAS	KU LEUVEN	BELGIUM

VERDEJO	FELISA	UNED	SPAIN
VILLA	ROBERT	UNIVERSITY OF SHEFFIELD	UNITED KINGDOM
VILLATORO-TELLO	ESAU	UNIVERSIDAD AUTONOMA METROPOLITANA	MEXICO
VILLEGAS	MAURICIO	UNIVERSIDAD POLITECNICA DE VALENCIA	COLOMBIA
VILLENA-ROMÁN	JULIO	DAEDALUS	SPAIN
WANG	QIUYUE	RENMIN UNIVERSITY OF CHINA	CHINA
WEST	WILLIAM	LEHIGH UNIVERSITY	UNITED STATES
WILHELM	THOMAS	CHEMNITZ UNIVERSITY OF TECHNOLOGY	GERMANY
WOMSER-HACKER	CHRISTA	UNIVERSITY OF HILDESHEIM	GERMANY
WU	WEI	INNER MONGOLIA UNIVERSITY	CHINA
YAN	XUELIANG	INNER MONGOLIA UNIVERSITY	CHINA
YOUNUS	ARJUMAND		
ZELLHÖFER	DAVID	BTU COTTBUS - ZELLH FER	GERMANY
ZNAIDIA	AMEL	CEA	TUNISIA

## Appendix G: List of Institutions participating to the Labs

AFFILIATION	COUNTRY	CHiCH	CLEF -IP	Image CLEF	INEX	PA N	QA4 MRE	Rep Lab	CLEFe Health
Aristotle University of Thessaloniki	Greece			x					
Athens University of Economics and Business	Greece			x					
Australian National University	Australia								x
Austrian Institute of Technology GmbH	Austria								x
Bar Ilan University	Israel					x			
Barcelona Media	Spain							x	
BeiHang University	China			x					
Benemérita Universidad Autónoma de Puebla	Mexico					x			
Brandenburgis che Technische Universität	Germany			x					
Bulgarian Academy of Sciences	Bulgaria						x		
CEA LIST	France			x					
Chemnitz University of Technology	Germany		x						
CLLE-ERSS: CNRS	France					x			
Concordia University	Canada						x		
Cyril and Methodius University in Skopje	Macedonia , the former Yugoslav Republic of			x					

Daedalus	Spain							x	
Daegu University	South Korea					x			
DAICT Gandhinagar	India				x				
Digital Enterprise Research Institute	Ireland	x							
Dokuz Eylul University	Turkey			x					
Dublin City University	Ireland				x				x
Duquesne University	USA					x			
Ecole Nationale Supérieure des Mines de Saint-Etienne	France				x				
Electronics and Telecommunications Research Institute	South Korea					x			
ETSI Informatica UNED	Spain			x					
FernUniversität in Hagen	Germany						x		
Fraunhofer Institute FIRST	Germany					x			
Fudan University	China						x		
Gavagai AB	Sweden					x		x	
Ge Global Research, Niskayuna, NY	USA			x					
GESIS - Leibniz Institute for the Social Sciences	Germany	x							
Graz University of Technology	Austria					x			
Hefei University of Technology	China			x					

Heinrich Heine University	Germany			x		x			
Higher Institute of Computer Science and Multimedia of Sfax	Tunisia			x					
Humboldt-Universität zu Berlin	Germany	x							
Hungarian Academy of Sciences	Hungary			x					
IBM Hawthorne	USA			x					
Independent University Bagladesh	Bangladesh	x							
Indian Institute of Technology Hyderabad	India			x					
Indian Institute of Technology Kanpur	India			x					
Indian Statistical Institute (Kolkata)	India				x				
Information Technologies Institute Centre for Research and Technology	Greece			x					
Inner Mongolia University	China			x					
INRIA	Germany		x						
INRIA Sophia-Antipolis	France			x					
Institut de Recherche en Informatique de Toulouse	France				x				
Institute of Astrophysics, Optics, and Electronics	Mexico					x			
Institute of Technology	China					x			

Instituto de investigacion en Informatica de Albacete	Spain			x					
ISM Dhanbad	India				x				
Istanbul Technical University	Turkey			x					
Jadavpur University	India				x	x	x		
JOANNEUM RESEARCH Forschungsge sellschaft mbH, DIGITAL	Austria		x						
JRC	Italy							x	
Karolinska University Hospital	Sweden								x
Katholieke Universiteit Leuven	Belgium						x		
Know-Center GmbH	Austria					x			
Kyungsung University	South Korea					x			
L2F INESC-ID Lisboa	Portugal						x		
Laboratoire Informatique d'Avignon	France				x				
Laboratory of Informatics of Grenoble (LIG)	France			x					
Lehigh University	USA					x			
Leiden University	The Netherlands					x			
LIA - University of Avignon	France				x				
LIMSI-CNRS	France						x		
LIRIS ECL	France			x					
LSIS - Univ. Marseille	France				x				
Masaryk University	Czech Republic					x			



Merck KGaA	Germany						x		
Microsoft Advanced Technology Labs Cairo	Egypt			x					
Mohamed Vth University Agdal, Rabat	Morocco						x		
MPI-Inf	Germany				x				
National ICT Australia	Australia						x		x
National Institute of Informatics	Japan			x					
National Library of Medicine	USA			x					
National School of Engineering of Sfax	Tunisia			x					
National Taiwan Normal University	Taiwan, Province of China						x		
National University of Defense Technology	China			x					
Netherlands Forensic Institute	The Netherlands					x			
Norwegian University of Science and Technology	Norway				x				x
Ontotext AD.	Bulgaria						x		
Optenet	Spain					x			
Oslo and Akershus University College of Applied Sciences	Norway				x				
Oxyme	The Netherlands							x	
Politehnica University of	Romania					x			

Bucharest									
Porfiau Inc.	Canada					x			
PSU, IRIT	Russian Federation				x				
Purdue University	USA					x			
Renmin University of China	China				x				
RMIT University GPO	Australia			x					
Royal School of Library & Information Science	Denmark				x				
Sabanci University	Turkey			x					
SAIC-Frederick Inc., Chemical Biology Laboratory at NCI-Frederick	USA		x						
San Francisco State University	USA			x					
SKB Kontur Labs	Russian Federation			x					
SkyLine Inc.	Ukraine								
Stockholm University	Sweden								x
Technische Universität Darmstadt	Germany					x			
TU Chemnitz	Germany	x							
Turku Centre for Computer Science	Finland								x
U.C. Berkeley	USA	x							
UNED	Spain							x	
Univ. Sud Toulon	France			x					
Univ. Tecnológica	Argentina			x					

Nacional, FR Cordoba									
Univerity Paris 6 (LIP6)	France					x			
Universidad Autónoma Metropolitana	Mexico				x	x			
Universidad de Chile	Chile					x			
Universidad de Guanajuato	Mexico					x			
Universidad Nacional Autónoma de México	Mexico				x				
Universidad Nacional de Colombia	Colombia			x					
Universidad Nacional de San Luis	Argentina					x			
Universidad Politécnica de Madrid	Spain					x			
Universidad Politécnica de Valencia	Spain					x			
Universidade de Évora	Portugal						x		
Universidade de Santiago de Compostela	Spain					x			
Universidade Federal de Minas Gerais	Brazil			x					
Universidade de São Paulo	Brazil			x					
Università degli Studi dell'Insubria	Italy			x					
Universitat Autònoma de Barcelona	Spain		x						
Universitat Pompeu Fabra	Spain				x				
Université de Nantes	France				x				
University of Minnesota	USA				x				

Duluth									
University Alexandru Ioan Cuza Iasi	Romania			x		x	x		
University at Buffalo	USA			x					
University Carlos III de Madrid	Spain			x			x		
University College Dublin	Ireland						x		
University of A Coruña	Spain					x			
University of Alabama at Birmingham	USA					x			
University of Amsterdam	The Netherland s				x			x	
University of Antwerp	The Netherland s					x			
University of Applied Sciences, Geneva, Information Studies	Switzerland		x						
University of Banja Luka	Bosnia and Herzegovin a			x					
University of Birmingham, School of Computer Science	United Kingdom		x						
University of Bucharest	Romania					x			
University of Cagliari	Italy			x					
University of California	USA								x
University of Galway	Ireland							x	
University of Gothenburg	Sweden								x
University of Hildesheim	Germany		x						
University of	USA							x	

Iowa									
University of Lugano	Switzerland		x						
University of Maryland	USA					x			
University of Melbourne	Australia								x
University of Michigan Library	USA				x				
University of Neuchatel	Switzerland	x							
University of Pisa	Italy						x		
University of Saint Petersburg	Russia					x			
University of Sfax	Tunisia			x			x		
University of Sheffield	United Kingdom					x		x	
University of Surrey	United Kingdom					x			
University of the Basque Country	Spain	x							
University of Tokyo	Japan			x					
University of Toronto	Canada					x			
University of Toulouse	France					x			
University of Turku	Finland								x
University of Western Sydney	Australia								x
University of Wollongong	Australia								x
Ural Federal University	Russia			x					
Ursinus College	USA					x			
Vanderbilt	USA							x	
Versicherungsanstalt für Eisenbahnen und Bergbau (VAEB)	Austria								x

Vienna University of Technology	Austria		x						
VM Labs	France				x				
Vulcan Inc	USA						x		
Zhytomyr State University	Ukraine					x			
<b>TOTAL Institutions: 166</b>		<b>8</b>	<b>10</b>	<b>46</b>	<b>24</b>	<b>45</b>	<b>20</b>	<b>11</b>	<b>15</b>

## Appendix H: CLEF 2012 Sponsors

CLEF 2012 benefited from the support of the following organizations:



PROMISE Network of Excellence



Sapienza University of Rome, Italy



ELIAS Research Network Programme



Quaero



European Science Foundation, France

## Appendix I: CLEF 2012 Website Statistics

### Visitors Overview

Jan 1, 2012 - Sep 27, 2012

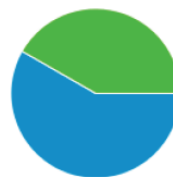
% of visits: 100.00%

#### Overview



6,273 people visited this site

Visits: 14,850  
Unique Visitors: 6,273  
Pageviews: 37,392  
Pages / Visit: 2.52  
Avg. Visit Duration: 00:02:40  
Bounce Rate: 49.91%  
% New Visits: 41.70%



58.27% Returning Visitor  
8,653 Visits  
41.73% New Visitor  
6,197 Visits



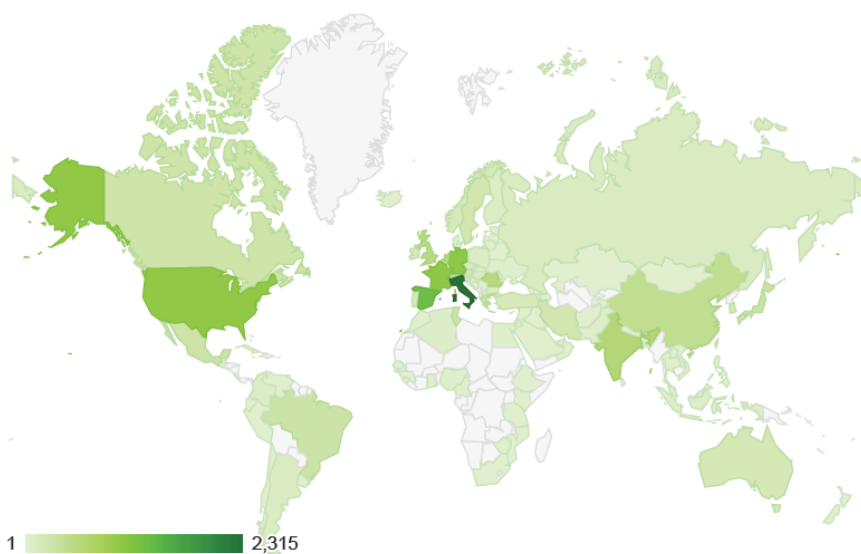
## Location

Jan 1, 2012 - Sep 27, 2012

 % of visits: 100.00%

Map Overlay

Site Usage



Visits  
**14,850**  
% of Total: 100.00% (14,850)

Pages / Visit  
**2.52**  
Site Avg: 2.52 (0.00%)

Avg. Visit Duration  
**00:02:40**  
Site Avg: 00:02:40 (0.00%)

% New Visits  
**41.70%**  
Site Avg: 41.70% (0.00%)

Bounce Rate  
**49.91%**  
Site Avg: 49.91% (0.00%)

Visits		Pages / Visit		Avg. Visit Duration		% New Visits		Bounce Rate	
14,850		2.52		00:02:40		41.70%		49.91%	
% of Total: 100.00% (14,850)		Site Avg: 2.52 (0.00%)		Site Avg: 00:02:40 (0.00%)		Site Avg: 41.70% (0.00%)		Site Avg: 49.91% (0.00%)	
Country / Territory				Visits	Pages / Visit	Avg. Visit Duration	% New Visits	Bounce Rate	
1.	Italy			2,315	2.73	00:03:28	23.46%	50.63%	
2.	Spain			1,294	2.65	00:02:54	34.39%	45.36%	
3.	Germany			1,018	2.80	00:02:40	40.57%	42.24%	
4.	France			1,018	2.76	00:02:48	33.50%	47.35%	
5.	United States			1,001	2.18	00:01:45	67.63%	59.54%	
6.	India			615	2.05	00:01:57	57.72%	54.63%	
7.	Romania			545	2.12	00:01:36	37.61%	61.83%	
8.	United Kingdom			502	2.25	00:02:07	51.99%	53.78%	
9.	Netherlands			493	2.67	00:03:23	41.18%	43.00%	
10.	China			432	2.72	00:03:46	46.99%	46.53%	
Rows 1 - 10 of 109									