PROMISE

Participative Research labOratory for Multimedia and
Multilingual Information Systems Evaluation

FP7 ICT 2009.4.3, Intelligent Information Management

Deliverable 7.10
Third Annual Public Report

Version 1.0, 30 November 2012

Document Information

Deliverable number:	7.10
Deliverable title:	Third Annual Public Report
Delivery date:	30/11/2012
Lead contractor for this deliverable	UNIPD
Author(s):	Nicola Ferro and Debora Leoncini, UNIPD
Participant(s):	All
Workpackage:	7
Workpackage title:	Dissemination, IPR, and Resources
Workpackage leader:	ELDA
Dissemination Level:	PU – Public
Version:	1.0
Keywords:	annual public report

History of Versions

Version	Date	Status	Author (Partner)	Description/Approval Level
0.1	05/11/2012	Draft	UNIPD	First draft circulated to all partners
0.2	25/11/2012	Draft	UNIPD	Second draft sent to internal revision after partners comments
1.0	30/11/2012	Final	UNIPD	Final version after internal revision

Abstract

This document reports on the activities conducted in the year 2012 by the PROMISE project which achieved several goals:

- It organized the CLEF 2012 Conference on Multilingual and Multimedia Information Access Evaluation, which attracted more than 190 participants.
- It released a new prototype of its evaluation infrastructure, based on the DIRECT system.
- It focused on use case based evaluation activities and proposed alternative evaluation methodologies for ground-truth creation.as well as best practices of MLIA systems.
- It organized and contributed to the organization of several knowledge transfer and dissemination events, such as the PROMISE Winter School 2012 and the PROMISE Retreat.

The document will also outlook future work and activities planned for 2013.

Table of Contents

PROMISE Annual Report 2012

http://www.promise-noe.eu/

With a population of over 500 million in its 27 states in which there are more than 80 indigenous and many more immigrant languages, the citizens and companies of the EU demand information systems that allow them to interact with the culturally and politically diverse content that surrounds them in multiple media.

The technology and research behind multilingual and multimedia information systems are, today, in the position of intercepting these emerging trends but their design and development is becoming increasingly complex and needs proper means for ensuring that they meet the expected user requirements and provide the desired effectiveness.

Measuring is a key to scientific progress. This is particularly true for research and development concerning complex systems, whether natural or human-built. PROMISE will provide a virtual laboratory for conducting participative research and experimentation to carry out, advance and bring automation into the evaluation and benchmarking of complex multilingual and multimedia information systems.

Summary of Activities

PROMISE is a Network of Excellence, started in September 2010 and scheduled until August 2013, with 13 academic and industrial partners, who are leaders in the field. PROMISE considers experimental evaluation – both laboratory and interactive – a key means for supporting and fostering the development of multilingual and multimedia information systems which are more adherent to the new user needs in order to ensure that they meet the expected user requirements, provide the desired effectiveness and efficiency, guarantee the required robustness and reliability, and operate with the necessary scalability.

PROMISE aims at advancing the experimental evaluation of complex multimedia and multilingual information systems in order to support individuals, commercial entities, and communities, who design, develop, employ, and improve such complex systems. The overall goal of PROMISE is to deliver a unified environment where to collect data, knowledge, tools, methodologies, and bringing together the user communities, which are involved in the experimental evaluation.

Important Work Areas
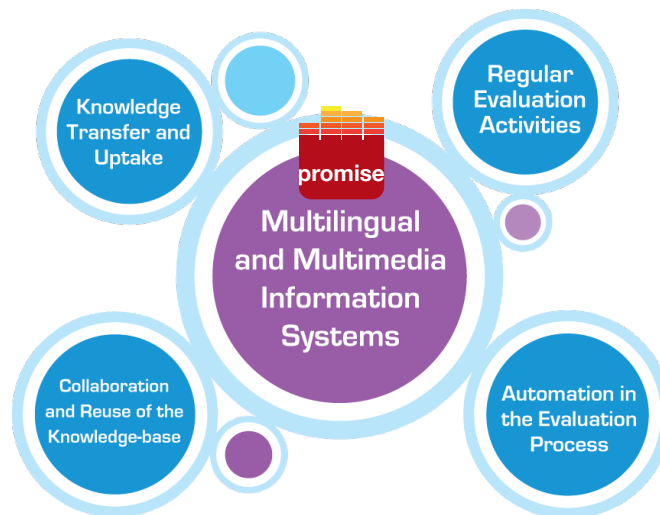
To achieve its goals, PROMISE pursues the following main activites:

- **Foster the adoption of regular and thorough experimental evaluation activities**: PROMISE carries on the successful and renowned CLEF (*Cross-Language Evaluation Forum*) evaluation campaigns further pushing the evaluation exercises to tackle realistic tasks, use cases, and data sets;

- **Bring automation into the experimental evaluation process**: PROMISE proposes methods and provides a software infrastructure to create larger experimental collections; increase the number and size of the experiments conducted; and develop distributed, asynchronous, and loosely-coupled evaluation protocols, moving experimental evaluation from handicraft process to a mostly automatic one.

- **Promote collaboration and re-use over the acquired knowledge-base**: PROMISE will curate, preserve, and allow for the enrichment of the collected experimental data; it will provide the means for an easy comparison with and a meaningful interpretation and visualisation of the experimental results; and it will facilitate the discussion and collaboration among all the interested stakeholders.

- **Stimulate knowledge transfer and uptake**: PROMISE disseminates know-how, tools, and best practices about multilingual and multimedia information systems; facilitate uptake and participation by commercial entities and industries; and gives rise to multidisciplinary competencies and expertises.

Regular Evaluation Activities: CLEF 2012

CLEF 2012 is the third CLEF conference continuing the popular CLEF campaigns which have run since 2000 contributing to the systematic evaluation of information access systems, primarily through experimentation on shared tasks. Building on the format first introduced in 2010, CLEF 2012 consists of an independent peer-reviewed conference on a broad range of issues in the fields of multilingual and multimodal information access

evaluation, and a set of labs and workshops designed to test different aspects of mono and cross-language Information retrieval systems.

CLEF 2012 has been hosted by the "La Sapienza" University of Rome, Italy, on 17-20 September 2012.

Figure 1 shows various moments of the CLEF 2012 conference, as well as some of its external sponsors.

Figure 1: Different moments of the CLEF 2012 conference.

Eight labs were offered at CLEF 2012:

- **CHiC Cultural Heritage in CLEF** a benchmarking activity to investigate systematic and large-scale evaluation of cultural heritage digital libraries and information access systems.
- **CLEF-IP** a benchmarking activity to investigate IR techniques in the patent domain
- **ImageCLEF** a benchmarking activity on the experimental evaluation of image classification and retrieval, focusing on the combination of textual and visual evidence
- **INEX** a benchmarking activity on the evaluation of XML retrieval
- **PAN** a benchmarking activity on uncovering plagiarism, authorship and social software misuse
- **QA4MRE** a benchmarking activity on the evaluation of Machine Reading systems through Question Answering and Reading Comprehension Tests
- **RepLab** a benchmarking activity on reputation management technologies

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

14 papers and 3 posters were accepted for the Conference and published by Springer in their Lectures Notes for Computer Science (LNCS) series. Two keynote speakers highlighted important developments in the field of evaluation. Peter Clark, Vulcan Inc., USA, focused on how to move from information retrieval to knowledgeable machines. Tobias Schreck, University of Konstanz, Germany, presented research challenges for visual search and analysis in textual and non-textual documents and their evaluation.

The community sessions at CLEF 2012 were organized around the presentation of the activities of the other evaluation initiatives and an "Evaluation Clinic" where participants had the possibility of meeting evaluation experts and discussing with them their evaluation problems.

Automation in the Evaluation Process: the DIRECT System

The second prototype of the PROMISE evaluation infrastructure (deliverable 3.2, 3.3, 5.2, and 5.3) allowed use to carry out the PROMISE use cases performed during CLEF 2011, namely "Unlocking Culture) (CHiC 2012 Evaluation Lab) "Search for Innovation" (CLEF-IP 2012 Evaluation Lab) and "Visual Clinical Decision Support" (ImageCLEF 2012 Evalution Lab). An early prototype of it has been demoed during the CLEF 2011 community session entitled "The Promise evaluation infrastructure, where Visual Analytics meets Information Retrieval" on September 21, 2011. Moreover, its data model (see D3.2) and additional feedback on requirements and expected functionalities have been gathered during the DESIRE 2011 Workshop on "Data infrastructurEs for Supporting Information Retrieval Evaluation", co-located with CIKM 2011 on October 28, 2011. The data model has been further revised, aligned, and tested in the light of D4.2 "Tutorial on Evaluation in the Wild" and D5.3 "Collaborative User Interface Prototype with Annotation Functionalities".
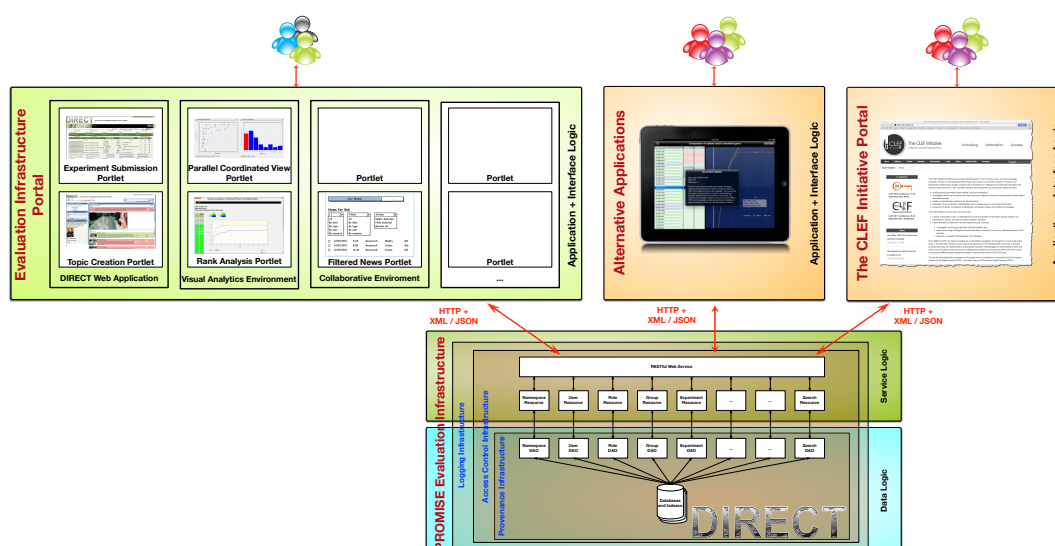

Figure 2: Architecture of the PROMISE evaluation infrastructure.

The PROMISE infrastructure adopts a modular and layered approach, whose architecture is shown in Figure 2, by distinguishing between: (i) data logic layer, where the managed resources are persisted and stored; (ii) service logic layer, where the managed resources are exposed as a RESTful Web service; (iii) application and interface logic layers, where different applications are built, as the one for managing an evaluation campaign or the one for providing advanced visual analytics tools for exploring the experimental data

Collaboration and Reuse of the Knowledge-base

There are two main components in the PROMISE user interfaces: the annotation service, one of the main means that is provided to promote the collaboration among the stakeholders of the evaluation infrastructure and the Visual Analytics component that allows for quick and automated support to the main analysis strategies that are currently adopted by PROMISE's stakeholders. The two components are tightly connected: annotating the visual analysis is one of the most effective means to spread knowledge and share findings. The details presented on the deliverable D3.3 allows for a clear understanding of how the visualizations produced by the Visual Analytics component are managed by the infrastructure: they are treated as "first class objects" and the system saves and retrieves not simple images but all the details behind them: track, run, experimental data, analysis strategy, and image manipulation. Moreover, the FAST annotation service adopts the same architecture of the overall PROMISE infrastructure. Figure 3 shows an example of per topic analysis:

- **table chart**: displays the data table.
- **box-plot chart**: displays an aggregated visualization of values that each topic reached with respect to the chosen metric. In a per topic analysis a box plot chart is used to evaluate the trend of a topic among experiments with respect to a chosen metric. A box plot chart presents a box plot for each topic on x-axis and the chosen metric on y-axis.
- **scatter-plot chart**: is used to compare two experiment on all topic with respect to the chosen metric. In a per topic analysis a bi-dimensional scatter plot presents on x-axis topics and on y-axis the chosen metric. Each metric value is represented by a point. For each topic there are as many points on y-axis as the selected columns. To see the trend of a single experiment (a column) you can unify its points with a polyline. To highlight some point it is possible to use color or markers.
- **bar/point/line/stacked-bar chart**: is used to compare topics. In a per topic analysis a bar chart can be used to compare two or more experiments on all the topics with respect to a chosen metric. Although it is possible to compare more than two experiments, as the number of experiments (topics) increases the chart representation loses clarity. Possible comparisons are two algorithm of the same participant or the best algorithm of two different participants

Figure 3: Example of per topic analysis.

One of the peculiar features of this version of the Collaborative User Interface prototype is the management of the annotations. After the user has made reason about the data that are visualized in the different charts available, he may want to add permanent annotations for future referencing and/or to involve other users in the particular insights that he has discovered.

In order to do so the user will have to simple press the button on the bottom of the screen, representing a pen, and it will led the system to a new tab for annotation shown in Figure 4. In this tab will be present a preview screenshots of the particular set of visualizations under analysis: this preview will take into account all the different manipulations operations done

by the user on the set of visualizations, and will serve as a reference for the textual and non-textual annotations that the user what to share. The second element of this tab is an hyper-textual box where the user can add all the relevant information he want to share with others or simply saving for future uses, and a check boxes system constructed with the title names of the involved visualizations for indexing and referencing to which kind of visualizations are referred the comments.
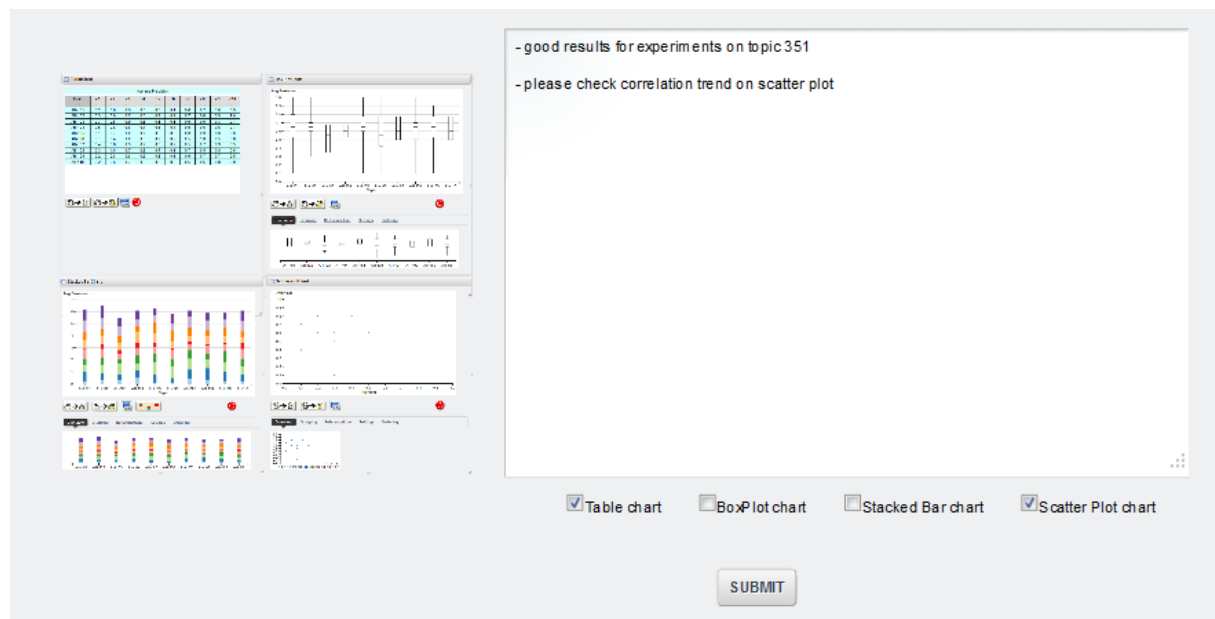

Figure 4: Example of annotation of the per topic analysis.

Knowledge Transfer and Uptake

Best practices report

Gathering, sorting, evaluating and retrieving information is increasingly becoming crucial for professional (information) workers and also in every-day life. Consequently, information retrieval (IR) technology has been widely adopted in some fields, with some IR applications, such as Web search services becoming immensely popular. Academic research has started to branch into many aspects covering advanced IR issues such as the handling of multilingual and multimedia information, but there is as yet little take-up in operational systems for the proposed solutions. This report presents best practice recommendations for information retrieval (IR) system developers, IR application implementers and IR application maintainers. A focus of the work has been on clear, concise recommendations, which are listed in a structured form, accompanied by a more detailed discussion of associated aspects. To this end, we introduce the following structural elements, where applicable:

1. **Validity and Qualification**: Restrictions of applicability depending on use cases domains or other circumstances.
2. **Dependencies**: Dependencies and effects on other BPs.

3. **Action**: Concrete action plan: e.g., features to be implemented or steps to be taken when following the BP.
4. **Expected Impact**: Quantified (if possible) impact of the implementation of this BP on the application, the system or on individual components.
5. **Support of Claim**: Source material for BP, including references, experiments, etc.
6. **Tested Configurations**: Actually implemented and tested configurations of the BP which performed well. Currently, this is mostly a placeholder. In the future, it is desirable and foreseeable that rich sets of experiments can be linked for each BP. Ultimately; external contributors should be able to add to the element through the DIRECT system (D3.2 and D3.3).

PROMISE Winter School 2012

The first PROMISE Winter school[1] on "Information Retrieval meets Information Visualization" has been organized in Zinal, Valais, Switzerland, on 23 - 27 January 2012. This Winter School corresponds to D7.7 "First PROMISE Summer School" which was scheduled for month 24 but it has anticipated to month 17.

The aim of the Promise Winter School on information retrieval and information visualization is to give participants a grounding in the core topics that constitute the multidisciplinary area of Multilingual Information Retrieval. The school was a week-long event consisting of guest lectures from invited speakers who are recognized experts in the field. The School was intended for PhD students, master students or senior researchers such as post-doctoral researchers form the fields of information visualization and information retrieval and related fields.

The school has been attended by participants who came from one domain or the other and offered them the possibility of starting to acquire cross-disciplinary competencies. Interestingly enough, the school turned out to be a brainstorming and discussion opportunity also for the lecturers, since they had the occasion of meeting colleagues from a quite different field with their own perspectives on a ground of shared topics and issues, such as how to envision models and design systems around user needs, how to consider the user interaction and context, how to conduct evaluation, an so on.

17 high quality lecturers from academia and industry were invited to speak on a large variety of topics from introduction talks to hot topics such as crowd sourcing and social media. 62 participants from 25 countries and all continents but Australia and four organizers followed the courses and helped to create many lively discussions and an open atmosphere with many questions. Also most of the speakers stayed for the entire week and enriched the discussions as well.

---

[1] http://www.promise-noe.eu/events/winter-school-2012/

Figure 5: Participants to the PROMISE Winter School 2012.

PROMISE Retreat

The PROMISE network of excellence organized a two-days brainstorming workshop on 30th and 31st May 2012 in Padua, Italy, to discuss and envisage future directions and perspectives for the evaluation of information access and retrieval systems in multiple languages and multiple media.

25 researchers from 10 different European countries attended the event, covering many different research areas – information retrieval, information extraction, natural language processing, human-computer interaction, semantic technologies, information visualization and visual analytics, system architectures, and so on. The event has been organized as a "retreat" allowing researchers to work back to back and propose hot topics where to focus research in the field in the coming years.

Six promising research lines have been envisaged: search applications; contextual evaluation; challenges in test collection design and exploitation; component-based evaluation; ongoing evaluation; and signal-aware evaluation.

The ultimate goal of the PROMISE retreat is to stimulate and involve the research community along these research lines and to provide funding agencies with effective and scientifically sound ideas for coordinating and supporting information access research.

Figure 6: Group photo of the participants to the retreat and front-page of the printed retreat report.

The outcomes of the PROMISE retreat have been summarized in a report that has been printed (ISBN 978-88-6321-039-2) and distributed to about 200 participant at CLEF 2012 in Rome, September 2012. The PROMISE retreat report has been also advertised on the main mailing lists in the field and, from September 2012 to date, its online version has been downloaded more than 500 times.

A short communication about the PROMISE retreat has been published in the October 2012 issue of the ERCIM News.

Finally, a condensed version of the PROMISE retreat report has been submitted for the December 2012 issue of SIGIR Forum.

PROMISE and Other EU Projects

Khresmoi

The Khresmoi Integrated Project (contract no. 257528, http://khresmoi.eu/) collaborated in the organization of the PROMISE Winter School 2012.

Chorus+

The CHORUS+ Coordination Action (contract no. 249008) organized the "Information Retrieval from Scientific Multimedia Data" community session during CLEF 2011.

CULTURA

CULTURA is a STREP (contract no. 269973, http://www.cultura-strep.eu/) which cooperates with PROMISE along the following action lines:

- Investigation of proper evaluation methodologies for adaptive services for digital cultural collections;

- Cooperation to support dissemination, training, and awareness events which promote the goals of both projects.

CULTURA contributed to the organization of the PROMISE Winter School.

LiMoSINe

LiMoSINe is a STREP (contract no. 288024, http://limosine-project.eu/) which organized the "RepLab" evaluation laboratory on reputation management at CLEF 2012 and contacts are ongoing to preprarea MoU between PROMISE and LiMoSINe on aspects related to evaluation.

Future Work

Evaluation Infrastructure

The PROMISE Evaluation Infrastructure is currently being extended to support the continuous evaluation process by targeting a real system, i.e. Europeana via its search APIs.
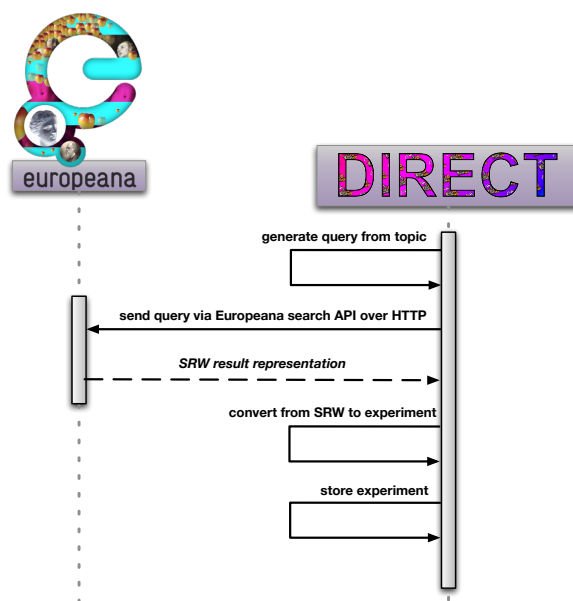
Figure 7: Continuous evaluation of Europeana currently being integrated in the PROMISE Evaluation Infrastructure.

It is also taking into account synergies with new tasks: in particular, the part of the model concerning the contribution will be exploited for Task 3.6 "Semantic Exploitation of Experimental Data" and Task 6.6 "Impact Analysis for the CLEF Initiative" while the part concerning experiments and guerrilla experiments may be exploited for storing the data of Task 4.7 "User-centered Evaluation Methodology".

Finally, other synergies are being explored, such as the possibility of representing the best practices (Task 2.5) as first-class object in the infrastructure and relate them to the relevant experimental data.

PROMISE Winter School 2013

The second PROMISE Winter School[2] on "Bridging between Information Retrieval and Databases" is being organized in Bressanone, Italy, on 4 - 8 February 2013. The idea of the school stems from the observation that nowadays databases are more and more getting into techniques that have traditionally been typical of information retrieval and, viceversa, information retrieval is using more and more database-oriented techniques.

Figure 8: PROMISE Winter School 2013 program.

PROMISE Technology Transfer Day at CeBIT

The Technology Transfer Day" is being organized together with the CeBIT 2013, from 5 to 9 March 2013, Hannover, Germany. It will consist of:

- a shared stand with KRESHMOI (and other 5 projects) in the CeBIT Lab
- a 45 minutes slot (application submitted) in the main program of the CeBIT Global Conference (CGC)
  - Title: Is your Search Engine making you miss Business Opportunities?
  - Subtitle: How Lessons learned in Research about Information Access Evaluation can help Industry.

  The slot will take the shape of an interactive event, where a professional journalist will interview researchers and panellists on the above topics.
- a one day event in the middle of CeBIT exhibition grounds in the Convention Center targeting three different kinds of stakeholders in separate sessions during the day:
  - 1st slot: decision makers of producer companies
  - 2nd slot : decision makers of consumer companies
  - 3rd slot: technology people and demos (both consumers and producers)

CLEF 2013

Work has already started to organize the CLEF 2013 conference which will be held in Valencia, Spain, on 23-26 September 2013.
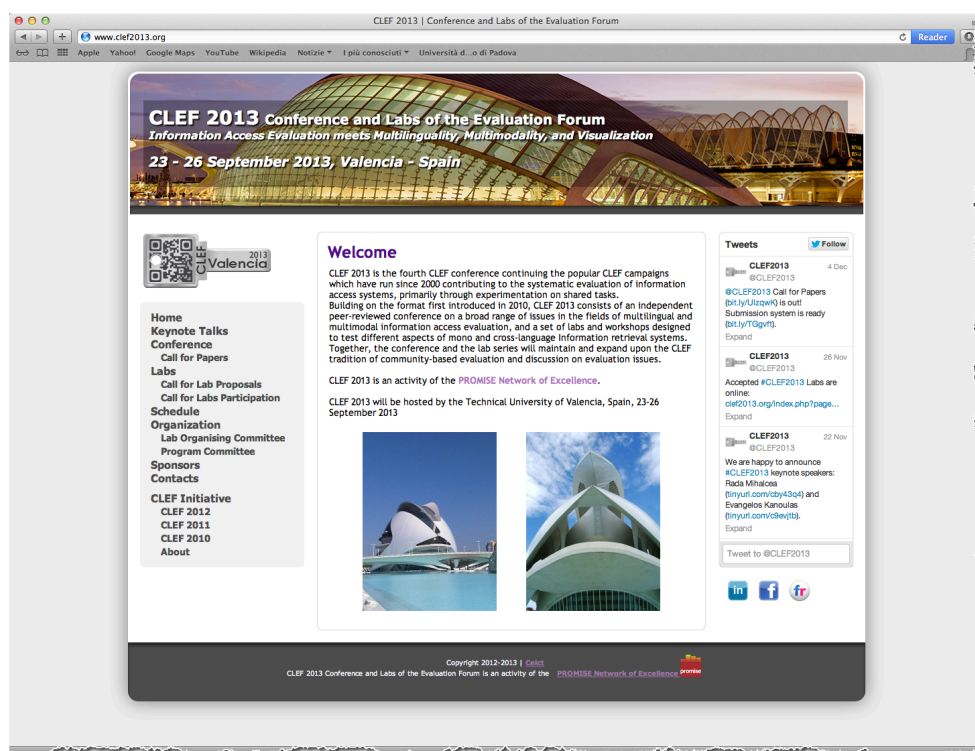

Figure 9: Homepage of the CLEF 2013 conference.

Further Information

Further information about PROMISE, its activities, and related events can be found at:

- CLEF 2012 home page
  `http://www.clef2012.org/`

- CLEF 2013 home page
  `http://www.clef2013.org/`

- CLEF General home page
  `http://www.clef-initiative.eu/`

- PROMISE Winter School 2013
  `http://www.promise-noe.eu/events/winter-school-2013`

PROMISE Consortium

- University of Padua, Italy (coordinator)
  http://www.unipd.it/

- University of Amsterdam, The Netherlands
  http://www.uva.nl/

- Sapienza University of Rome, Italy
  http://www.uniroma1.it/

- University of Applied Sciences Western Switzerland, Switzerland
  http://www.hes-so.ch/

- Zurich University of Applied Sciences, Switzerland
  http://www.zhaw.ch/

- Humboldt-Universität zu Berlin, Germany;
  http://www.hu-berlin.de/

- Evaluations and Language resources Distribution Agency (ELDA), France
  http://www.elda.org/

- Centre for the Evaluation of Language Communication Technologies (CELCT), Italy
  http://www.celct.it/
- Technical University of Vienna, Austria
  http://www.tuwien.ac.at/

- University of Sheffield, United Kingdom

  http://www.shef.ac.uk/

- National University of Ireland, Galway, Ireland

  http://www.nuigalway.ie/

- Royal School of Library and information Science, Denmark

  http://www.iva.dk/

- University of Gothenburg, Sweden

  http://www.gu.se/

Contact Person

Dr. Nicola Ferro

Department of Information Engineering University of Padua

Via Gradenigo 6/b

35131 Padova, Italy

e-mail: ferro@dei.unipd.it

Tel: +39 049 827 7939

Fax: +39 049 827 7799