

# Introduction to the CLEF2013 Labs and Workshop

Roberto Navigli<sup>1</sup> and Dan Tufis<sup>2</sup>

<sup>1</sup> *Sapienza University of Rome, Italy*

<sup>2</sup> *RACAI, Romania*

*navigli@di.uniroma1.it; tufis@racai.ro*

The CLEF 2013 conference is a new 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. In 2010 CLEF was launched in a new format, as a conference with research presentations, panels, poster and demo sessions and laboratory evaluation workshops. These are proposed and operated by groups of organizers volunteering their time and effort to define, promote, administrate and run an evaluation activity. Labs for CLEF 2013, as in the editions from 2010, 2011 and 2012, are of two types: **laboratories** to conduct evaluation of information access systems, and **workshops** to discuss and pilot innovative evaluation activities. CLEF 2013 is the densest campaign until now as there were accepted 9 laboratories and one workshop. To identify the best proposals, 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 organizing consortium, this year we stressed movement beyond previous years' efforts and connection to real-life usage scenarios. Each Lab, building on previous experience, demonstrated maturity coming with new tasks, new and larger data sets, new ways of evaluation or more languages. They are described by the Lab organizers in details, here we just brief on them.

## **PAN - Uncovering Plagiarism, Authorship, and Author Profiling**

PAN 2013 addressed issues related to Digital Text Forensics and evaluated the participants' submissions along three tasks:

- Plagiarism Detection: Given a document, is it an original?
- Author Identification: Given a document, who wrote it?
- Author Profiling: Given a document, what are the author's demographics?

## **ImageCLEF 2013 - Cross Language Image Annotation and Retrieval**

The main goal of ImageCLEF, which started in 2003, is supporting the development of visual media analysis, indexing, classification, and retrieval by building the infrastructure for the evaluation of visual information retrieval systems operating in monolingual, language-independent and multi-modal contexts. The three challenging tasks of the ImageCLEF 2013 were:

- Photo Annotation and Retrieval: semantic concept detection using private collection data, and large-scale annotation using general Web data;
- Plant Identification: visual classification of leaf images for the identification of plant species;
- Robot Vision: semantic spatial understanding for a mobile robot using multimodal data.

## **INEX - INitiative for the Evaluation of XML retrieval**

INEX was concerned with three different aspects of focused information access:

- The Linked Data Track studies ad-hoc search and faceted search over entities in a strongly structured collection of Linked Data (DBpedia) tied to a large textual corpus (Wikipedia).
- The Social Book Search Track studies the value of user-generated descriptions in addition to formal metadata on a collection of Amazon Books and LibraryThing.com data. In addition, the track studies the challenges of searching full text of scanned books.
- Focused retrieval: the Snippet Retrieval Track studies how to generate informative snippets for search results; the Tweet Contextualization Track studies tweet contextualization, answering questions of the form

"what is this tweet about?" with a synthetic summary of contextual information from Wikipedia and evaluated by both the relevant text retrieved, and the "last point of interest."

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

The goal of QA4MRE is to evaluate Machine Reading abilities through Question Answering and Reading Comprehension Tests.

- The task focuses on the reading of single documents and selection of the answers to a set of questions about information that is stated or implied in the text. Two additional pilots are also proposed:
- Machine Reading of Biomedical Texts about Alzheimer's Disease: aimed at answering questions specific to the biomedical domain, with a special focus on the Alzheimer's disease.
- Entrance Exams: aiming at answering multiple-choice questions of real English Reading Comprehension tests contained in Japanese University Entrance Exams.

### **QALD-3 - Question Answering over Linked Data**

QALD-3 is the third in a series of evaluation campaigns on question answering over linked data, for the first time with a strong emphasis on multilinguality. Two open challenges are offered:

- Question answering: given an RDF dataset and a set of natural language questions of varying complexity and in multiple languages, participating systems are asked to provide correct answers (or SPARQL queries that retrieve those answers).
- Ontology lexicalization: it focuses on lexica that can facilitate multilingual information access. Participants are asked to find lexicalizations of a set of classes and properties from English DBpedia across languages in a given corpus.

### **CHiC - Cultural Heritage in CLEF**

The CHiC 2013 evaluation lab aims at moving towards a systematic and large-scale evaluation of cultural heritage digital libraries and information access systems. Three different tasks were run:

- Multilingual ad-hoc and semantic enrichment, assessing IR in a multilingual collection both for ad-hoc IR and query enrichment;
- Polish ad-hoc, evaluating Polish-language retrieval;
- Interactive, where the evaluation framework is extended to an interactive study observing users during a non-intentional browsing activity.

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

The CLEF-IP lab provides a large collection of XML documents representing patents and patent images. The following tasks were organized on this document collection, but only the first one enjoyed submissions:

- Passage retrieval starting from claims: Starting from a given claim, we ask to retrieve relevant documents in the collection and mark out the relevant passages in these documents;
- Image to text, text to image: Given a patent application document - as an XML file - and the set of images occurring in the application, extract the links between the image labels and the text pointing to the object of the image label.
- Image to structure task: Extract the information in patent images (flowcharts, electrical diagrams) and return it in a predefined textual format.

### **CLEFeHealth 2013**

Discharge summaries describe the course of treatment, the status at release, and care plans. Both nurses and patients are likely to have difficulties in understanding their content, because of their compressed language full of medical jargon, nonstandard abbreviations, and ward-specific idioms. The three tasks of this Lab were:

- Identification (1a) and normalization (1b) of disorders in clinical reports with respect to terminology standards in healthcare,
- Normalizations of abbreviations and acronyms in clinical reports with respect to terminology standards in healthcare,
- IR to address questions patients may have when reading clinical reports.

### **RepLab 2013**

RepLab 2013 is focused on the problem of real-time tracking the reputation of companies or individuals in Twitter. The participants were invited to submit results from a full processing flow or from modules that contribute only partially to the problem. The organizers provided baseline components for all aspects of the task, so that research groups could test systems that address partial problems. Evaluation results were provided for:

- clustering + ranking of the tweets (the main task) and for two subtasks:
- polarity for reputation;
- name ambiguity resolution.

### **CLEF-ER workshop - Entity Recognition Reading Evaluation**

The only workshop at CLEF 2013 was CLEF-ER, organized as a challenge by the EC Mantra project. The workshop is set up to address entity recognition in biomedical text, in different languages and at a large scale. Semantic integration is and will be an important focus. The workshop brings together stakeholders from different domains and researchers who take part in the Mantra challenge. The researchers will explore on the evaluation and results of the Mantra challenge from the first half of 2013 and provide input, such as proposals for novel tasks and evaluations, for future challenges. The current Mantra challenge targets the identification of entity mentions and their concept identifiers (CUIs) from a standard terminological resource in multi-lingual texts. To this end, parallel biomedical corpora have been prepared. These corpora are also exploited to identify entity correspondences and to augment multi-lingual terminologies.

### **Acknowledgements**

We would like to thank previous co-chairs of CLEF-LOC 2012 (the CLEF Lab Organisation Committee) for sharing their experience and ensuring thoughtful and elaborate contributions to assessing the proposals during the selection process:

Jussi Karlgren, Gavagai and SICS (CLEF LOC 2012 co-chair)  
Christa Womser-Hacker, Universität Hildesheim (CLEF LOC 2012 co-chair)

We also thank all the proposal authors, whose effort, enthusiasm and creativity provided new energy to this new edition of CLEF, the lab and workshop organisers, the colleagues and friends involved in running the labs, and all the participants who made the labs and workshops a great success.

Thank you all!