

CLEF 2014

Conference and Labs of
the Evaluation Forum

15 to 18th September
2014 Sheffield

Call for Labs Participation

CLEF 2014 is next year's edition of the popular CLEF campaign and workshop series which contributes to the systematic evaluation of information access systems, primarily through experimentation on shared tasks. Labs are of two types: laboratories to conduct evaluation of information access systems, and workshops to discuss and pilot innovative evaluation activities. There are eight labs for CLEF 2014.

NEWSREEL — News Recommendation Evaluation Lab

NEWSREEL offers two tasks:

- **Task 1:** Predict the items a user will click in the next 10 Minutes based on the offline dataset. A sliding window approach is used for the evaluation of the recommender algorithm's quality. The main emphasis is here on the reproducible, deep analysis of the user's behavior.
- **Task 2:** Predict the articles users will click. The prediction algorithms are evaluated in an online scenario based on live user-interactions. The main focus is on providing real-time recommendations for current news articles. Participants will be allowed to fine-tune their algorithms before submission starts in the weeks leading up to CLEF 2014.

Lab Coordination: Technische Universität Berlin, plista GmbH, Berlin.

Lab website: <http://www.clef-newsreel.org/>

QA Track — CLEF Question Answering Track

In the current general scenario for the CLEF QA Track, the starting point is always a Natural Language question. However, answering some questions may need to query Linked Data (especially if aggregations or logical inferences are required); whereas some questions may need textual inferences and querying free-text. Answering some queries may need both. The tasks are:

- **QALD: Question Answering over Linked Data**
- **BioASQ: Biomedical semantic indexing and question answering**
- **Entrance Exams**

Lab Coordination: INRIA, NCSR, Carnegie Mellon, University Leipzig, UNED, University of Limerick, CITEC

Lab website: <http://nlp.uned.es/clef-qa>

CLEF eHealth - ShARe/CLEF eHealth Evaluation Lab

The usage scenario of the CLEF eHealth lab is to ease patients and next-of-kins' ease in understanding eHealth information. The lab contains three tasks:

- **Visual-Interactive Search and Exploration of eHealth Data**
- **Information extraction from clinical text**
- **User-centred health information retrieval**

Lab Coordination: Dublin City University; Universities of Arizona, Konstanz, Canberra, Utah, Pittsburgh, Melbourne, Turku; Australian National University; SICS and Stockholm University; NICTA; DSV Stockholm University; Columbia University; KTH and Gavagai; Karolinska Institutet; Harvard Medical School and Boston Children's Hospital; Vienna University of Technology; HES-SO; Charles University; and the Australian e-Health Research Centre.

Lab website: <http://clefehealth2014.dcu.ie/>

ImageCLEF

ImageCLEF aims at providing benchmarks for the challenging task of image annotation for a wide range of source images and annotation objective, such as general multi-domain images for object or concept detection, as well as domain-specific tasks such as visual-depth images for robot vision and volumetric medical images for automated structured reporting. The tasks address different aspects of the annotation problem and are aimed at supporting and promoting the cutting-edge research addressing the key challenges in the field, such as multi-modal image annotation, domain adaptation and ontology driven image annotation. The Lab tasks are:

- **Robot Vision**
- **Scalable concept Image Annotation**
- **Liver CT Annotation**
- **Domain Adaptation**

Lab Coordination: University of Rome La Sapienza, University of Castila-La Mancha.

Lab website: <http://www.imageclef.org/2014>



15 — 18th September 2014
Sheffield — England

Participation

Registration is via the CLEF website

Publication

Working Notes are published online, with an ISBN.

CLEF 2014

Conference and Labs of
the Evaluation Forum

15 to 18th September
2014 Sheffield

PAN Lab on Uncovering Plagiarism, Authorship, and Social Software Misuse

PAN centers around the topics of plagiarism, authorship, and social software misuse. The goal is to foster research on automatic detection and uncovering. People increasingly share their work online, contribute to open projects and engage in web-based social interactions. The ease and anonymity with which this can be done raises concerns about verifiability and trust: Is a given text an original? Is the author the one who she claims to be? Does a piece of information come from a trusted source? Answers to such questions are crucial to deal with and to rely on information obtained online, while the scale at which answers should be given calls for an automatic means. Lab Tasks:

- **Author Identification**
- **Author Profiling**
- **Plagiarism Detection**

Lab Coordination: Bauhaus-Universität Weimar, Universität Politècnica de València, University of the Aegean
Lab website: <http://pan.webis.de>

RepLab

The aim of RepLab is to bring together the Information Access research community with representatives from the Online Reputation Management industry, with the ultimate goals of (i) establishing a roadmap on the topic that includes a description of the language technologies required in terms of resources, algorithms, and applications; (ii) specifying suitable evaluation methodologies and metrics to measure scientific progress; and (iii) developing of test collections that enable systematic comparison of algorithms and reliable benchmarking of commercial systems. Lab Tasks:

- Task 1. Annotating company-related tweets according to the dimension(s) of the company affected by their content (social, financial, etc.)
- Task 2. Generating brief pseudo-summaries of each of the topics (tweet clusters) that may serve as surrogate of a set of tweets for the purposes of reputation management.

Lab contact: Julio Gonzalo, julio@lsi.uned.es

Lab Coordination: UNED, University of Amsterdam, Yahoo! Research Barcelona, Llorente & Cuenca

General Chairs

Paul Clough and Mark Sanderson

Evaluation lab chairs

Martin Halvey and Wessel Kraaij

Programme chairs

Evangelos Kanoulas and
Mihai Lupu

Proceedings chairs

Elaine Toms and Allan Hanbury

Local Organization

Robert Villa and Mark Hall

INEX — Initiative for the Evaluation of XML retrieval

INEX builds evaluation benchmarks for search in the context of rich structure such as document structure, semantic metadata, entities, or genre/topical structure. INEX 2014 runs four tasks studying different aspects of focused information access:

- **Social Book Search Task:** investigates the relative value of authoritative metadata and user-generated content. The test collection is from Amazon and LibraryThing, and user profiles and personal catalogues.
- **Interactive Social Book Search Task:** investigates user information seeking behavior when interacting with various sources of information for realistic task scenarios, and how the user interface impacts search and the search experience.
- **Linked Data Task:** investigates complex questions to be answered by DBpedia/Wikipedia, with the help of SPARQL queries and additional keyword filters, aiming to express natural language search cues more effectively (in collaboration with the QA Lab).
- **Tweet Contextualization Task:** investigates tweet contextualization, helping a user understand a tweet by providing a short background summary generated from relevant Wikipedia passages aggregated into a coherent summary (in collaboration with RepLab).

Lab Coordination: Queensland University of Technology, University of Amsterdam, University of Passau.

Lab website: <https://inex.mmci.uni-saarland.de/>

LifeCLEF

LifeCLEF aims at evaluating multimedia analysis and retrieval techniques on biodiversity data for species identification. Tasks are:

- **BirdCLEF:** a bird songs identification task based on Xeno-Canto audio recordings
- **PlantCLEF:** an image-based plant identification task based on the data of Tela Botanica social network
- **FishCLEF:** a fish video surveillance task based on the data of the Fish4Knowledge network

Lab Coordination: INRIA Sophia-Antipolis, University of Applied Sciences Western Switzerland, University of Toulon, INRIA, TU Wien, Xeno-Canto, Cirad – AMAP, University of Catania, University of Edinburgh

Lab webpage: <http://www.lifeclef.org>

Important Dates

Lab registration opens:	22nd Nov. 2013
Lab registration closes:	7th May 2014
Evaluation cycle:	Dec. 2013 - May 2014
Working notes papers due:	7th June 2014
Lab overview papers due:	30th June 2014