CLEF 2017

Conference and Labs of the Evaluation Forum

Information Access Evaluation meets Multilinguality, Multimodality, and Visualization

11 - 14 September 2017, Dublin - Ireland

http://clef2017.clef-initiative.eu/

Call for Lab Participation

CLEF 2017 is the 8th CLEF conference continuing the popular CLEF campaigns which have run since 2000 contributing to the systematic evaluation of information access systems. CLEF 2017 consists of an independent peer-reviewed conference, combined with and a set of Labs and Workshops designed to test different aspects of mono and cross-language Information Retrieval systems. CLEF 2017 offers six Labs and four Workshops: **1**. CLEF NewsREEL **2**. LifeCLEF **3**. PAN **4**. CLEF eHealth **5**. MC² Microblog Cultural Contextualization **6**. ImageCLEF **7**. PIR-CLEF **8**. Early risk prediction on the Internet **9**. Dynamic Search for Complex Tasks **10**. Multimodal Spatial Role Labeling

Important Dates

Lab Registration Opens: **4 November 2016**
Registration closes: **21 April 2017**
End Evaluation Cycle: **5 May 2017**
Working Notes papers due: **26 May 2017**
Camera Ready Copy: **3 July 2017**

Organizers

Conference Chairs

Gareth J. F. Jones, Dublin City University, Ireland
Séamus Lawless, Trinity College Dublin, Ireland

Program Chairs

Julio Gonzalo, UNED, Spain
Liadh Kelly, Trinity College Dublin, Ireland

Lab Chairs

Lorraine Goeuriot, Université Joseph Fourier, France
Thomas Mandl, University of Hildesheim, Germany

Proceedings Chairs

Linda Cappellato, University of Padua, Italy
Nicola Ferro, University of Padua, Italy

Participation

Participants must register for the Labs via the CLEF website:
http://clef2017-labs-registration.dei.unipd.it/

Publication

Lab Working Notes will be published in the CEUR-WS Proceedings:
http://ceur-ws.org

Lab Paper Submission via Easychair:
http://easychair.org/conferences/?conf=clef2017

CLEF NewsREEL

The goal of CLEF NewsREEL is to provide a vehicle for the IR/recommender system communities to move from conventional offline evaluation to online evaluation. NewsREEL addresses the following information access challenge: Whenever a visitor of an online news portal reads a news article on their side, the task is to recommend other news articles that the user might be interested in.

Task 1: NewsREEL Live: The first subtask implements the idea of living laboratories, i.e., researchers gain access to the resources of a company to evaluate different recommendation techniques using A/B testing.

Task 2: NewsREEL Replay: This subtask focuses on replaying a previously recorded data stream of activities, hence allowing benchmarking of news recommendations in a simulated environment

Lab Coordination: Frank Hopfgartner (University of Glasgow, UK) Torben Brodt (plista GmbH, Berlin, DE)
Lab website: http://clef-newsreel.org/ @clefnewsreel

LifeCLEF

LifeCLEF lab aims at boosting research on the identification of living organisms and on the production of biodiversity data. Through its biodiversity informatics related challenges, LifeCLEF is intended to push the boundaries of the state-of-the-art in several research directions at the frontier of multimedia information retrieval, machine learning and knowledge engineering. The lab is organized around three tasks (and one pilot task):

Task 1: PlantCLEF: crowdsourced biodiversity monitoring from mobile photo streams
Task 2: BirdCLEF: bird species identification from bird calls and songs
Task 3: SeaCLEF: organisms Identification in Sea-Related Visual Data.
Pilot Task: HabitatCLEF: observation-based habitat classification

Lab Coordination: Alexis Joly (INRIA, Sophia-Antipolis, FR), Henning Müller (University of Applied Sciences Western Switzerland in Sierre, CH), Hervé Goëau (CIRAD, FR)
Lab website: http://lifeclef.org/

PAN

CLEF 2017 is the 14th evaluation lab on uncovering plagiarism, authorship and social software misuse. This year the Lab is organized around three tasks:

Task 1: Author identification: This task will address author clustering and style breach detection. The task will focus on short documents of length. Style breach detection has the goal of identifying breaches of writing style in longer texts, which is a prerequisite to identifying changes of authorship within a document.

Task 2: Author profiling - Gender and language variety identification in Twitter: Demographic traits such as gender and language variety have so far been investigated separately. In this task we will provide participants with a Twitter corpus annotated with authors' gender and their specific variation of their native language.

Task 3: Author obfuscation: This task works against identification and profiling by automatically paraphrasing a text to obfuscate its author's style. The tasks offered are author masking and obfuscation evaluation.

Lab Coordination: Martin Potthast, Benno Stein, Matthias Hagen (Bauhaus-Universitat Weimar, DE), Paolo Rosso (Universitat Politècnica de València, SP), Francisco Rangel (Autoritas Consulting, SP), Efstathios Stamatatos (Univerisity of the Aegean, GR)
Lab website: http://pan.webis.de/

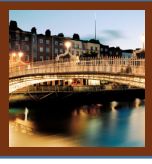
CLEF eHealth

Task 1. Multilingual Information Extraction: Participants will be required to extract the causes of death from death certificates, authored by physicians in European languages. This can be seen as a named entity recognition, normalization, and/or text classification task. The task also has a replication track, where participants are given the option to have not only their automatically codified documents evaluated, but also the system used to produce them.

Task 2. Technologically Assisted Reviews in Empirical Medicine: Participants will be challenged to retrieve medical studies relevant to conducting a systematic review on a given topic. This can be seen as a total recall problem and is addressed by both query generation and document ranking. The task consists of two subtasks: interactive information retrieval; and query construction (optional task).

Task 3. Patient-centred Information Retrieval: Participants must retrieve web pages that fulfil a given patient's personalised information need. This needs to fulfil the following criteria: information reliability, quality, and suitability, at both the individual query level and optimised over an entire search session. The task also has a multilingual track, which challenges participants to fulfil these information needs in a multilingual setting.

Lab Coordination: Lorraine Goeuriot (Univ. J.Fourier, FR), Evangelos Kanoulas (Univ. of Amsterdam, NL), Liadh Kelly (Trinity College, IR), Aurélie Névéol (CNRS-LIMSI, FR), Joao Palotti (Vienna Univ., AU), Aude Robert (INSERM/CepiDC, FR), Rene Spijker (Julius Center, NL), Hanna Suominen (Australian National Univ., AUS), Guido Zuccon (Queensland Univ. of Technology, AUS)
Lab Website: https://sites.google.com/site/clefehealth2017/ @clefehealth



MC² Microblog Cultural Contextualization

This lab deals with how cultural context of a microblog affects its social impact at large. This involves microblog search, classification, filtering, language recognition, localization, entity extraction, linking open data and summarization. Lab participants will have access to the massive multilingual microblog stream of the festival galleries project. The microblog stream and related URLs of festivals are appropriate for experiments into advanced social media search and mining methods. There are three tasks:

1. Content Analysis: Given a stream of microblogs, the content analysis task consists of filtering microblogs dealing with festivals, language(s) identification, event localization, author categorization, recognition of DBpedia entities and automatic summarization of linked Wikipedia pages in four languages.

2. MicroBlog Search: Given a cultural entity as a set of Wikipedia pages, this task will involve two sub-tasks: Task 2a: Retrieval of relevant microblogs for an entity; Task 2b: Summarization of the most informative microblogs.

3. Timeline Illustration: The goal of the Timeline illustration based on Microblogs task is to provide, for each event of a cultural festival, the most interesting tweets.

Lab Coordination: Liana Ermakova, Josiane Mothe (Institut de Recherche en Informatique de Toulouse, FR), Lorraine Goeuriot, Philippe Mulhem (Université Grenoble Alpes, FR), Fionn Murtagh (University of Derby & University of London, UK), JianYun Nie (Université de Montréal, CAN), Eric SanJuan (Université d'Avignon, FR)

Lab website: https://mc2.talne.eu/

ImageCLEF

1. ImageCLEFlifelog: The availability of a large variety of personal devices, e.g., smartphones, video cameras, as well as wearable devices that allow capturing pictures, and videos in every moment of our life, is creating the need for systems that can automatically categorize, summarize and retrieve these data. The task addresses the problems of lifelogging data retrieval and summarization.

2. **ImageCLEFcaption**: Interpreting and summarizing the insights gained from medical images such as radiology output is a time-consuming task. So, there is a considerable need for automatic methods that can approximate this mapping from visual information to condensed textual descriptions. The task addresses the problem of bio-medical image caption prediction from large training data.

3. ImageCLEFtuberculosis: The objective of this task is to determine tuberculosis subtypes and drug resistances, as far as possible automatically, from the volumetric image information in computed tomography (CT) volumes (mainly texture analysis) and based on clinical information (e.g., age, gender, etc).

4. **ImageCLEFremote** (pilot task - FabSpace 2.0 Exploring Sentinel Copernicus Images): The objective of this task is to explore Earth observation data images (Sentinel Copernicus satellite images) in order to discover unknown information.

Lab Coordination: Bogdan Ionescu (University Politehnica of Bucharest, RO), Mauricio Villegas (Universitat Politècnica de València, SP), Henning Müller (University of Applied Sciences in Sierre, CH)

Lab website: http://imageclef.org/2017 @imageclef #imageclef

PIR-CLEF

The aim of the PIR-CLEF is to provide a framework for evaluation of Personalised Information Retrieval (PIR). Current approaches to the evaluation of PIR are user-centered. The pilot test collection will provide all the traditional components needed in a laboratory-based evaluation experiment plus a set of user-related information for modelling and introducing profiles in the evaluation experiment: user personal information; search logs; user's documents of interest; basic user profile representations in the form of a bag-of-words; user satisfaction. We will use ClueWeb12, which contains over 730 million Web pages, as the basic repository to extract the above mentioned user-related information. Recognising that operating a ClueWeb12 service is a significant undertaking, API access to an existing ClueWeb12 search service will be made available to task participants. The workshop at CLEF 2017 will include a report on the PIR-CLEF pilot task, with short participant presentations, some invited presentations on the themes of benchmarking, personalization and adaptation, and a discussion of potential PIR tasks for CLEF 2018.

Workshop Coord.: Gabriella Pasi, Stefania Marrara (University of Milano Bicocca, IT), Gareth Jones, Debais Ganguly (Dublin City University, IR) Nicola Ferro, Maria Maistro (University of Padova, IT)

Workshop website: http://www.ir.disco.unimib.it/pirclef2017/

Early risk prediction on the Internet - eRisk

The main purpose of eRisk is to explore issues of evaluation methodology, effectiveness metrics and processes related to early risk detection. Early detection technologies can be employed in different areas, particularly related to health and safety. Our main goal is to pioneer a new interdisciplinary research area that would be potentially applicable to a wide variety of situations and to many different personal profiles. Examples include potential paedophiles, stalkers, individuals that could fall into the hands of criminal organizations, people with suicidal inclinations, or people susceptible to depression. eRisk 2017 has two possible ways to participate:

1. Research Papers Submission: Submission of papers describing test collections or data sets suitable for early risk prediction, early risk prediction challenges, tasks and evaluation metrics or specific early risk detection solutions.

2. Pilot Task: Early Detection of Depression: The challenge consists of sequentially processing pieces of evidence and detect early traces of depression as soon as possible

Workshop Coordination: David E. Losada (University of Santiago de Compostela, SP), Fabio Crestani (University of Lugano, CH), Javier Parapar (University of A Coruña, SP)

Workshop website: http://early.irlab.org/ @earlyrisk

Dynamic Search for Complex Tasks

Information Retrieval research has traditionally focused on serving the best results for a single query. However, users typically search iteratively, refining and reformulating their queries. A key challenge in the study of this interaction is the creation of suitable evaluation resources to assess the effectiveness of IR systems over sessions. The goal of the Dynamic Search lab is to propose and standardize an evaluation methodology that can lead to reusable resources and evaluation metrics able to assess retrieval performance over an entire session, keeping the "user" in the loop. The objective of the lab is: to produce the methodology and algorithms that will lead to a dynamic test collection by simulating the users; to understand and quantify in terms of evaluation measures what constitutes a good ranking of documents at different stages of a session, and a good ranking for the overall session; to develop algorithms that can provide an optimal ranking throughout a user's session. There are two possible ways to participate: (1) By submitting a scientific paper (2) By submitting a run to a pilot task

Workshop Coordination: Evangelos Kanoulas (University of Amsterdam, NL)

Workshop website: https://ekanou.github.io/dynamicsearch/ @clef_dynamic

Multimodal Spatial Role Labeling

The main goal of mSpRL is to explore the extraction of spatial information from two information resources that is: image and text. This is important for various applications such as semantic search, question answering, geographical information systems and even in robotics for machine understanding of navigational instructions or instructions for grabbing and manipulating objects. It is also essential for some specific tasks such as text to scene conversion or vice-versa, scene understanding as well as general information retrieval tasks when using a huge amount of available multimodal data from various resources. Moreover, there is an increasing interest in extraction of spatial information from medical images that are accompanied by natural language descriptions. mSpRL 2017 has two possible ways to participate: research paper submission and participation in a pilot task on multimodal Spatial Role Labeling.

Workshop Coordination: Parisa Kordjamshidi (Tulane University, USA), Taher Rahgooy (Bu-Ali Sina University, IR), Marie-Francine Moens (KULeuven, BE), James Pustejovsky (Brandeis University, USA), Kirk Roberts (University of Texas, USA)

Workshop website: http://www.cs.tulane.edu/~pkordjam/mSpRL_CLEF_lab.htm @mSpRL2017