CLEF 2012 Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics

1 Summary

This report describes the activities of the CLEF 2012 Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics conferences and the use of the founds granted by ELIAS Science Meeting 4479 for supporting such an event.

The CLEF 2012 Conference on Multilingual and Multimodal Information Access Evaluation was held at Rome, 17-20 September 2012, and was organized by the Department of Computer, Control and Management Engineering Antonio Ruberti of "La Sapienza" Università di Roma. In particular, Prof. Tiziana Catarci and Prof. Giuseppe were directly involved in the organization. Details about the conference are on Section 2 and Section 3.

Founds have been used for covering travel expenses, as reported on Section 4.

2 Description of the scientific content of and discussion at the event

Since 2000 the Cross-Language Evaluation Forum (CLEF) has played a leading role in stimulating research and innovation in a wide range of key areas in the domain of multimodal and multilingual information access. Through the years, CLEF has promoted the study and implementation of evaluation methodologies for diverse tasks, resulting in the creation of a broad, strong and multidisciplinary research community.

Until 2010, the outcomes of experiments carried out under the CLEF umbrella were presented and discussed at annual workshops in conjunction with the European Conference for Digital Libraries. CLEF 2010 represented a radical departure from this "classic" CLEF format. While preserving CLEF's traditional core goals, namely, benchmarking activities carried in various tracks, we complemented these activities with a peer-reviewed conference component aimed at advancing research in the evaluation of complex information access systems in different languages and modalities.

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.

This year, the papers accepted for the conference included research on Information Access and evaluation initiatives, methodologies and infrastructures. Two keynote speakers highlighted important issues related to our field.

Peter Clark (Vulcan Inc., USA) 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. Here are details of his talk:

Title:
From Information Retrieval to Knowledgeable Machines

Abstract:
Ultimately we would like our machines to 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. In this talk, I will present our work in Project Halo to create an (iPad hosted) "knowledgeable biology textbook", called Inquire. Inquire includes a formal, hand-crafted knowledge base encoding some of the book's content, being augmented (this year) with capabilities for textual entailment and question-answering directly from the book text itself. Inquire allows the user to not only read and browse the 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. In this talk I will describe the project, in particular the textual question-answering component and its use of natural language processing, paraphrasing, textual entailment, and its exploitation of the formal knowledge base. I will also discuss the interplay being developed between the hand-built knowledge and automatic text-extracted knowledge, how each offers complementary strengths, and how each can leverage the other. Finally I will discuss the value of this approach, and argue for the importance of creating a deeper understanding of textual

material, and ultimately more knowledgeable machines.

Tobias Schreck (University of Konstanz, Germany), on the other hand, showed current approaches, applications and challenges for the application of visual analytics in document repositories. Here are details of his talk:

Title:
Visual Search and Analysis in Textual and Non-Textual Document Repositories - Approaches, Applications, and Research Challenges.

Abstract:
Information retrieval and analysis are key tasks in dealing with the information overload problem characteristic for today's networked digital environments. Advances in data acquisition, transmission and storage, and emergence of social media, lead to an abundance of textual and non-textual information items available to everyone at any time. Advances in visual-interactive data analysis can provide for effective visual interfaces for query formulation, navigation, and result exploration in complex information spaces.

In this presentation, we will discuss selected approaches for visual analysis in large textual and non-textual document collections. First, recent techniques for visual analysis of readability, sentiment and opinion properties in large amounts of textual documents, including promising application possibilities, will be discussed. Then, we will focus on visual analysis support for information retrieval in non-textual documents, in particular multimedia and time-oriented research data. We argue 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. We will conclude the presentation by discussing research opportunities at the intersection of visual data analysis, information retrieval, and evaluation.

CLEF 2012 featured seven benchmarking activities:

Seven labs will follow a "campaign-style" evaluation practice for specific information access problems in the tradition of past CLEF campaign tracks:

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, offering four tasks:
    1. 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;
    2. Matching claim to description in a single document (Pilot): Starting from the claims of a patent application, we ask to indicate the paragraphs in the application's description section (same document) that best explain the contents of the given claim.
    3. Flowchart Recognition Task: Extract the information in flowchart images and return it in a predefined textual format.
    4. Chemical Structure Recognition Task: Starting from TIFF images containing patent scans, we ask 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.
3. ImageCLEF a benchmarking activity on the experimental evaluation of image classification and retrieval, focusing on the combination of textual and visual evidence. ImageCLEF offered four tasks:

1. 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;
2. Photo annotation and retrieval: semantic concept detection and concept-based retrieval using Flickr data, and large-scale annotation using general Web data;
3. Plant identification: visual classification of leaf images for the
4. identification of plant species;
5. Robot vision: semantic localisation of a mobile robot using multimodal place classification, with special focus on generalization.

4. [INEX](#) a benchmarking activity on the evaluation of XML retrieval, offering five tracks
   1. Social Book Search Track: studying the value of user-generated descriptions in addition to formal metadata on a collection of Amazon Books and LibraryThing.com data.
   2. Data Centric Track: studying adhoc search and facetted search on a collection of Linked Data (DBpedia) tied to a large corpus (Wikipedia).
   3. Snippet Retrieval Track: studying the generation of informative snippets with sufficient information to determine the relevancy of search results.
   4. Show Me Your Code Track: asking participants to submit system components (in particular feedback) rather than results.
   5. Tweet Contextualization Track: retrieving synthetic contextual information from Wikipedia in response to a tweet with a URL on a small terminal like a phone.

5. [PAN](#) a benchmarking activity on uncovering plagiarism, authorship and social software misuse. In particular, PAN offered three tasks:
   1. Plagiarism Detection: This task 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.
   2. Author Identification: This task focuses on identifying sexual predators in chat logs and on authorship verification. Moreover, it features for the first time real cases of disputed authorship.
   3. Quality Flaw Prediction in Wikipedia. This task is newly introduced, and it is about identifying Wikipedia articles which contain certain information quality flaws. It generalizes the vandalism detection task of last year.

6. [QA4MRE](#) a benchmarking activity on the evaluation of Machine Reading systems through Question Answering and Reading Comprehension Tests. Two pilots were proposed:
   1. Processing Modality and Negation for Machine Reading: aimed at evaluating whether systems are able to understand extra-propositional aspects of meaning like modality and negation;
   2. Machine Reading of Biomedical Texts about Alzheimer: aimed at setting questions in the biomedical domain with a special focus on the Alzheimer disease.

7. [RepLab](#) a benchmarking activity on reputation management technologies, offering two shared tasks on Twitter data:
   1. a monitoring task, where the goal is to thematically cluster tweets including a company's name as a step towards early alerting on issues that may damage the company's reputation.
   2. a profiling task, where the goal is annotating tweets according to their polarity for reputation (i.e. as to whether their content has positive/negative implications for the company's reputation).

Moreover, a lab has been organized as a workshop, grouping speaking and discussion sessions to explore issues of evaluation methodology, metrics, and processes in information access and closely related fields:

1. [CLEFeHealth 2012](#) workshop on Cross-Language Evaluation of Methods, Applications, and Resources for eHealth Document Analysis, that required abstracts on:
   1. evaluation of mono- and multilingual methods, applications and resources for eHealth document analysis; and
   2. development of statistical and user-feedback based evaluation protocols, settings, methods and measures for cross-language evaluation of methods, applications, and resources for eHealth document analysis.
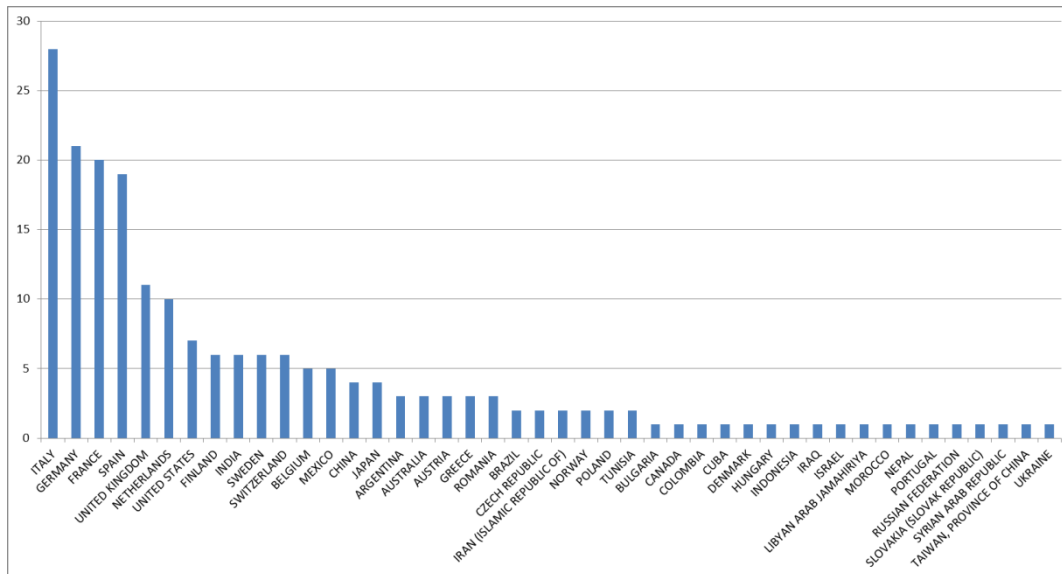
All the experiments carried out by systems during the evaluation campaigns are described in a separate publication, namely, the Working Notes, distributed during CLEF 2013 and available on-line.

3 Assessment of the results and impact of the event on the future direction of the field

A plenary Clef 2013 session

Results and impact of the conference are higher than previous years: overall, CLEF 2012 was attended by more than 200 people from different academic and industrial institutions and the number of submitted papers increased of 40% with respect to 2011. Although the majority of participants came from Europe, we note more interest in CLEF from all around the world, e.g., India, China, Usa.

Nationality of Clef 2013 attendee

Several events allowed for increasing the scientific exchange among participants:

- The Community session "Other evaluation activities" allowed the community to better understand similar activities in the world. As a practical result, people from the European community have been invited to give invited talks at the Forum for Information Retrieval Evaluation (FIRE), in Calcutta on 17 - 19 December 2012, strengthening the relationship between Indian and European evaluation activities.

- The conference hosted a Clef Steering Committee meeting that made important decisions on how to proceed in the feature. Main issues regard the lab organization and timing, with the clear indication that lab organizers have to reduce the number of tasks in each lab as much as possible; and that 8 should be the maximum number of labs each year. Moreover the continuation of CLEF has been discussed and the 2013 edition will be on 23 - 26 September 2013, Valencia – Spain.

- Also this edition confirmed that the presence of peer-reviewed conference papers, experimented on 2010 and 2011, allows for focusing on novel scientific issues and acts as an "interlab glue" event, attracting people coming from different, specific, IR fields.

4 Final programme of the meeting

The conference programme was the following:

Monday 17

TIME	
8.00-9.00	Registration
9.00-9.30	Opening - Welcome *(Room Loyola)*
9.30-10.30	**Keynote: Peter Clark, Vulcan Inc., USA** *(Room Loyola)* *From Information Retrieval to Knowledgeable Machines*
10.30-11.00	Break
11.00-12.20	**Evaluation Methodologies and Infrastructure - 4 long papers** *(Room Loyola)*
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	**Benchmarking and Evaluation Initiatives** **2 long papers; 3 short papers** *(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	**Keynote: Tobias Schreck, Konstanz Univesity, Germany** *(Room Loyola)* *Visual Search and Analysis in Textual and Non-Textual Document Repositories - Approaches, Applications, and Research Challenges.*			
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 **Other Evaluation Initiatives**		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)*			

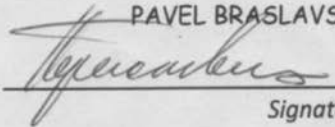The ESF founding has been used as travel grants for seven people (out of 21 applicants). However, because one of

the applicant, Pavel Braslavsky (see the following list), asked only 500 euros we planned to distribute the difference among four people coming from further locations, giving them (1125 euros instead of 1000). Unfortunately, because two of the selected people for this extra founding did not attend the conference (on the last minute the China Government refused them the VISA) we gave the grants to other two excluded European applicants but we were not able to spend all the money (to be fear with other European people). As a conclusion we used 6750, as detailed below for travel grants and 250 euros as contribution to local administrative costs.

Pavel Braslavsky	Ural Federal University - Russia	500 euros
Thomas Wilhelm	Chemnitz University of Technology, Germany	1000 euros
Emanuela Boros	Faculty of Computer Science , University "Alexandru Ioan Cuza" Romanian	1000 euros
Lluís-Pere de las Heras Caballero	DAG Universitat Autonoma de Barcelona, Edifici O, 08193, Bellaterra (Barcelona) Catalonia, Spain	1000 euros
Diego Antonio Rodriguez Torrejón	"Universidad de Huelva", Spain	1000 euros
Sandra Avila	NPDI/DCC/UFMG, Brazil	1125 euros
Prasenjit Majumder	DAIICT,India (FIRE organizer)	1125 euros

In the following we attach the signed receipts of these seven people.

Clef ROME 2012

I receive from ESF Science Meeting – ELIAS the amount of € 500,00 as travel grant for the participation to CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome from September 17 to 20, 2012.

PAVEL BRASLAVSKI

Signature

I receive from ESF Science Meeting – ELIAS the amount of € 1.000,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome** from September 17 to 20, 2012.

WILHELM THOMAS

Signature

CROME 2012 le

I receive from ESF. Science Meeting - ELIAS the amount of € 1.000,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum**, held in Rome from September 17 to 20, 2012.

EMANUELA BOROS

Signature

I receive from ESF Science Meeting – ELIAS the amount of € 1.000,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome** from September 17 to 20, 2012.
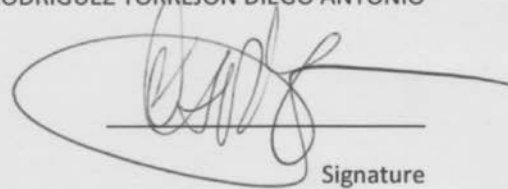
DE LAS HERAS LLUIS PERE

Signature

I receive from ESF Science Meeting – ELIAS the amount of € 1.000,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome** from September 17 to 20, 2012.

RODRIGUEZ TORREJON DIEGO ANTONIO

Signature

C ROME 2012 **le**⌐⌐

I receive from ESF Science Meeting – ELIAS the amount of € 1.125,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome** from September 17 to 20, 2012.

SANDRA AVILA

Signature

I receive from ESF Science Meeting – ELIAS the amount of € 1.125,00 as travel grant for the participation to **CLEF 2012 Conference and Labs of the Evaluation Forum, held in Rome** from September 17 to 20, 2012.

PRASENJIT MAJUMDER

_____

Signature