

Conferenza “CLEF 2010: Conference on Multilingual and Multimodal Information Access Evaluation”

Padova, 20-23 Settembre 2010

Maristella Agosti – Nicola Ferro

Università degli studi di Padova

L'Università degli studi di Padova ha organizzato e ospitato dal 20 al 23 settembre 2010 la conferenza internazionale “CLEF 2010: Conference on Multilingual and Multimodal Information Access Evaluation”¹. La conferenza si occupa della *valutazione sperimentale* dei sistemi di reperimento dell'informazione con una particolare attenzione ai problemi del multilinguismo, multimedialità e multi modalità. L'esempio più noto di sistema di reperimento dell'informazione sono i motori di ricerca nel Web ma tali sistemi sono anche usati come componenti di sistemi più ampi, quali i portali aziendali o le biblioteche digitali, con la finalità di fornire un accesso efficace ai contenuti gestiti. La valutazione sperimentale consiste nell'usare collezioni realistiche di documenti quali, ad esempio, pagine HTML prese dal Web, blog e Wikipedia, record catalografici collezionati da cataloghi in linea o biblioteche digitali, brevetti registrati a livello internazionale in varie lingue, oppure immagini mediche corredate dalle corrispondenti diagnosi, e sperimentare i sistemi sviluppati dai partecipanti all'iniziativa su tali documenti al fine di misurarne le prestazioni e comprendere quali siano gli algoritmi e le tecniche più efficaci.

Il CLEF (Cross-Language Evaluation Forum)² è un'importante iniziativa a livello internazionale che si è tenuta con regolarità ogni anno a partire dal 2000 e a cui partecipano i più importanti gruppi di ricerca a livello europeo ed internazionale, circa centocinquanta, così come aziende ed industrie leader nel settore. Assieme a TREC (Text REtrieval Conference)³, organizzato dal National Institute of Standards and Technology (NIST) negli Stati Uniti, e a NTCIR (NII Test Collection for IR Systems)⁴, organizzato dal National Institute of Informatics (NII) in Giappone, CLEF rappresenta una delle tre più importanti iniziative di valutazione sperimentale a livello mondiale. Nel corso degli anni, CLEF ha prodotto un miglioramento significativo e misurabile nelle prestazioni dei motori di ricerca multilingue, che sono quasi raddoppiate, offrendo agli utenti risultati di migliore qualità e riducendo la necessità di scorrere lunghe liste di risultati.

CLEF 2010 rappresenta un'innovazione radicale rispetto al formato classico di CLEF ed un esperimento per comprendere meglio come debbano essere strutturate le campagne di valutazione sperimentale di “prossima ge-

¹ CLEF 2010, <http://www.clef2010.org/>.

² CLEF, <http://www.clef-campaign.org/>.

³ TREC, <http://trec.nist.gov/>.

⁴ NTCIR, <http://research.nii.ac.jp/ntcir/>.

nerazione". Infatti, dopo dieci anni di successi, si è dovuto affrontare il problema di come innovare CLEF pur mantenendo il suo focus tradizionale incentrato sulla valutazione dei sistemi di reperimento dell'informazione. La scelta, dopo vivaci discussioni all'interno della comunità, è stata rendere CLEF un evento indipendente di quattro giorni non più organizzato congiuntamente alla ECDL (European Conference on Digital Libraries)⁵ dove CLEF è sempre stato un workshop della durata di due giorni e mezzo. I primi due giorni sono stati una conferenza vera e propria con articoli selezionati in base ad un rigoroso processo di revisione, relatori invitati e tavole rotonde, il tutto incentrato sull'avanzamento metodologico e scientifico della valutazione sperimentale nel settore del reperimento dell'informazione. I secondi due giorni sono stati costituiti da laboratori e workshop organizzati in sessioni parallele che hanno riguardato la presentazione e discussione delle attività di valutazione condotte con sistemi reali e dei risultati ottenuti.

L'evento è stato organizzato con la collaborazione di diverse istituzioni a livello nazionale ed internazionale: la prof. Maristella Agosti, Università degli studi di Padova, e il prof. Maarten de Rijke, Università di Amsterdam, hanno svolto il ruolo di *general chair*; il dr. Nicola Ferro, Università degli studi di Padova, e il prof. Alan Smeaton, Università di Dublino, sono stati i *program chair*; la dr. Donna Harman del National Institute of Standards and Technology (NIST), USA, e il dr. Martin Braschler dell'Università di Scienze Applicate di Zurigo hanno svolto il ruolo di *labs co-chair*; infine il dr. Emanuele Pianta del Centre for the Evaluation of Language Communication

Technologies (CELCT) di Trento ha ricoperto il ruolo di *organization chair* mentre il dr. Khalid Choukri della Evaluations and Language resources Distribution Agency (ELDA) di Parigi ha svolto il ruolo di *resource chair*.

Circa 140 ricercatori provenienti da Europa, Stati Uniti, Asia e Sud America hanno partecipato all'evento, la maggior parte per la sua intera durata di quattro giorni, e circa una ventina tra studenti, dottorandi e giovani ricercatori hanno ricevuto una borsa di studio per prendere parte alla conferenza, grazie alle sponsorizzazioni ricevute. Le presentazioni, le foto e i filmati degli interventi tenuti durante l'evento sono disponibili sul sito Web di CLEF 2010.

CLEF 2010: la conferenza

La conferenza mira all'avanzamento della ricerca sulla valutazione sperimentale dei sistemi informativi multilingue e multimediali al fine di supportare individui, organizzazioni e comunità che progettano, sviluppano, utilizzato e migliorano tali sistemi. Sono stati sollecitati articoli che esplorassero le esigenze e le pratiche correnti per l'accesso all'informazione multilingue e multimediale, che studiassero nuove metriche e metodologie di valutazione e che discutessero e proponessero nuove direzioni per le future attività di valutazione. In aggiunta, si è incoraggiata la sottomissione di articoli che analizzassero i risultati ottenuti in dieci anni di attività di CLEF conducendo esperimenti approfonditi sulle collezioni e i dati prodotti da CLEF nel corso del tempo. A questo fine i dati sperimentali degli ultimi dieci anni di CLEF sono stati resi disponibili in linea tramite il sistema DIRECT (Distributed Information Retrieval Evaluation Campaign

⁵ ECDL è stato, per quattordici anni, l'evento scientifico di riferimento in Europa per la comunità delle biblioteche digitali. In occasione del suo quindicesimo anno, la conferenza è stata rinominata *Theory and Practice of Digital Libraries* (TPDL). La prossima edizione di TPDL si terrà a Berlino dal 25 al 29 settembre 2011.

TPDL 2011, <http://www.tpd2011.org/>

Tool)⁶ e accessibili a seguito della registrazione nel sistema. Infine, è stato predisposto un ampio comitato di programma che coprisse non solo le aree tradizionalmente parte di CLEF ma che esplorasse anche nuove aree, come ad esempio la visualizzazione di grandi moli di dati.

Gli articoli accettati per la conferenza hanno riguardato attività di ricerca sulle risorse, gli strumenti e i metodi necessari per la valutazione sperimentale, così come la creazione di collezioni sperimentali realistiche e vicine alle necessità degli utenti.

In aggiunta, due relazioni invitate hanno illustrato prospettive future per la valutazione sperimentale sia dal punto di vista accademico sia dal punto di vista industriale. Il prof. Norbert Fuhr, Università di Duisburg-Essen in Germania, ha esplorato il tema *IR Between Science and Engineering, and the Role of Experimentation*, mentre il dr. Ricardo Baeza-Yates, Yahoo! Research e Universitat Pompeu Fabra in Spagna, ha illustrato le problematiche inerenti *Retrieval Evaluation in Practice*.

Il programma della conferenza ha compreso anche riepiloghi sulle attività condotte nei diversi "laboratori", spiegate poi nel dettaglio nei due giorni successivi, e interventi sui risultati delle altre campagne di valutazione a livello internazionale: TREC (Text REtrieval Conference)⁷ negli Stati Uniti, NII-NACSIS Test Collection for IR Systems (NTCIR)⁸ in Giappone, INitiative for the Evaluation of XML Retrieval (INEX)⁹ in Australia, Russian Information Retrieval Evaluation Seminar (ROMIP)¹⁰ in Russia, e Forum for Information Retrieval Evaluation (FIRE)¹¹ in India.

Al fine di stimolare ulteriormente la discussione e coinvolgere i partecipanti, sono state organizzate due tavole rotonde. La prima, *The Four Ladies of Experimental Evaluation*, ha visto coinvolte le quattro ricercatrici che hanno creato e coordinato le principali iniziative di valutazione degli ultimi venti anni: dr. Donna Harman (National Institute of Standards and Technology, NIST, USA), prof. Noriko Kando (National Institute of Informatics, NII, Japan), prof. Mounia Lalmas (University of Glasgow, UK) e dr. Carol Peters (Consiglio Nazionale delle Ricerche, ISTI-CNR, Pisa, Italia). In particolare, si è discusso di quali risultati siano stati effettivamente raggiunti fino ad oggi e quali ancora debbano essere conseguiti.

La seconda tavola rotonda, *A PROMISE for Experimental Evaluation*, è stata presentata dai membri del progetto PROMISE (Participative Research labOratory for Multimedia and Multilingual Information Systems Evaluation), una rete di eccellenza finanziata nel settimo programma quadro della Commissione Europea. PROMISE ambisce a far progredire la valutazione sperimentale dei sistemi di accesso all'informazione multilingue, multimediale e multi modale. L'obiettivo finale del progetto è fornire e rendere disponibili a livello internazionale un ambiente e un'infrastruttura software che unifichino i dati prodotti dalla valutazione sperimentale, la conoscenza da essi estratta, le metodologie di valutazione, gli strumenti per effettuare e automatizzare la valutazione e la comunità di utenti ad essa interessata, accrescendo l'utilizzo di tali dati e strumenti.

I risultati e gli interventi dei due giorni di con-

⁶ DIRECT, <http://direct.dei.unipd.it/>. Maristella Agosti – Giorgio Maria Di Nunzio – Marco Dussin – Nicola Ferro, *10 Years of CLEF Data in DIRECT: Where We Are and Where We Can Go*, in: *Proceedings 3rd International Workshop on Evaluating Information Access (EVIA 2010)*, edited by Tetsuya Sakay, Mark Sanderson, William Webber, Tokyo: National Institute of Informatics, 2010, p. 16-24.

⁷ Vedi nota 3.

⁸ Vedi nota 4.

⁹ INEX, <http://www.inex.otago.ac.nz/>.

¹⁰ ROMIP, <http://romip.ru/en/>.

¹¹ FIRE, <http://www.isical.ac.in/~fire/>.

ferenza sono stati pubblicati negli atti editi da Springer¹².

CLEF 2010: i laboratori e i workshop

I laboratori continuano ed espandono l'approccio tradizionale di CLEF. Sono stati previsti due tipi diversi di laboratori: attività di valutazione ed analisi comparativa delle prestazioni e workshops che esplorino nuove problematiche per la valutazione sperimentale. È stato formato un comitato di selezione dei laboratori per valutare le proposte di laboratorio e scegliere quelle più innovative per CLEF 2010. Sono stati adottati diversi criteri di selezione tra cui: appropriatezza della metodologia adottata, fattibilità dell'attività proposta, riferimento ad esigenze di utenti ben definite, esistenza di un caso d'uso di interesse, numero di potenziali partecipanti, prospettiva di evoluzione della proposta, dimensione e scala degli esperimenti, inter-disciplinarietà, originalità e possibilità di riutilizzare i risultati. I risultati delle attività dei laboratori sono riportate in una pubblicazione separata, le *working notes*, distribuita durante la conferenza e disponibile in linea¹³. È auspicato che i laboratori diano anche origine a pubblicazioni successive all'evento, come ad esempio articoli su rivista.

Sono state offerte le seguenti attività di valutazione e workshop:

- CLEF-IP: un'attività di valutazione riguardante la proprietà intellettuale e i brevetti multilingue;
- ImageCLEF: un'attività di valutazione riguardante la ricerca e recupero di immagini e testo, con particolare riferimento ad

immagini di tipo medico, fotografie prese da Flickr e immagini di Wikipedia;

- PAN: un'attività di valutazione sull'identificazione del plagio e dell'abuso dei software sociali, con particolare riferimento a Wikipedia;
- RespubliQA: un'attività di valutazione focalizzata sul *question answering* utilizzando una collezione multilingua di documenti del parlamento europeo;
- WePS: un'attività di valutazione focalizzata sulla ricerca di persone nel Web e sulla gestione della propria reputazione su Web;
- CriES: un workshop finalizzato ad esplorare la valutazione della ricerca di competenze e conoscenze nelle reti sociali;
- LogCLEF: un workshop finalizzato all'analisi e classificazione delle interrogazioni al fine di comprendere il comportamento dell'utente in un contesto multilingua, utilizzando i log di TEL (The European Library)¹⁴ e del motore di ricerca DBS (Deutscher Bildungsserver).

CLEF 2010: l'evento

I primi due giorni, la conferenza, sono stati ospitati al Palazzo del Bo, la sede storica dell'Università degli studi di Padova, nell'aula magna Galileo Galilei. I secondi due giorni, i laboratori, si sono tenuti presso la Facoltà di Scienze Statistiche, nel campus di Santa Caterina, un ex-convento del XIV secolo.

Il programma dei lavori è stato arricchito con una serie di eventi sociali che hanno favorito la conoscenza reciproca e lo scambio di idee in modo informale tra i partecipanti a CLEF 2010. In particolare la cena sociale è stata anche l'occasione per consegnare alla dr. Carol

¹² *Multilingual and Multimodal Information Access Evaluation: Proceedings of the International Conference of the Cross-Language Evaluation Forum (CLEF 2010)*, edited by Maristella Agosti, Nicola Ferro, Carol Peters, Maarten de Rijke, Alan Smeaton, Heidelberg: Springer 2010.

¹³ *CLEF 2010 Labs and Workshops, Notebook Papers*, edited by Martin Braschler, Donna Harman, Emanuele Pianta, Trento: MINT srl, 2010, <http://www.clef2010.org/index.php?page=pages/proceedings.php>.

¹⁴ TEL è un portale progettato per far interoperare le biblioteche nazionali europee; ad oggi sono 48 le biblioteche nazionali accessibili in linea attraverso TEL. <http://www.theeuropeanlibrary.org/>.

Peters il premio Tony Kent Strix Award¹⁵ che ha vinto nel 2009 per il suo lavoro di ideatore e coordinatore del Cross-Language Evaluation Forum negli ultimi dieci anni.

CLEF 2011 e prospettive future

CLEF 2011 seguirà un formato simile a quello introdotto con CLEF 2010 è cioè una conferenza seguita da laboratori e workshop.



Figura 1. Alcuni momenti di CLEF 2010

CLEF 2010 è stato parzialmente supportato dal progetto PROMISE¹⁶, rete di eccellenza finanziata nel settimo programma quadro della Commissione Europea. Inoltre, per la prima volta, ha ricevuto il supporto finanziario di diverse organizzazioni e istituzioni: Google, Svizzera; Xerox Research Centre Europe (XRCE), Francia; Information Retrieval Facility (IRF), Austria; CELI Language Information Technology, Italia; Dipartimento di ingegneria dell'informazione dell'Università degli studi di Padova, Italia; Consorzio per la formazione e la ricerca in ingegneria dell'informazione in Padova (COFRIDIP), Italia.

CLEF 2011 verrà organizzato dall'Università di Amsterdam, in Olanda, nel mese di Settembre 2011.

Per CLEF 2012 e successivi, si progetta di istituire una procedura di selezione in modo tale che le organizzazioni e le istituzioni interessate possano candidarsi ad organizzare l'evento e competano per ospitarlo.

Ringraziamenti

Il lavoro riportato in questo articolo è stato parzialmente supportato dalla rete di eccellenza PROMISE (contratto n. 258191), parte del settimo programma quadro della Commissione Europea.

The work reported has been partially supported by the PROMISE network of excellence (contract n. 258191) project, as part of the 7th Framework Program of the European Commission.

¹⁵ Tony Kent Strix Award, <http://www.ukeig.org.uk/awards/tony-kent-strix>.

¹⁶ PROMISE, <http://www.promise-noe.eu/>