

# Dig *Italia*

Anno VI, Numero 1 - 2011

Rivista del digitale nei beni culturali

ICCU-ROMA

# La rete di eccellenza europea PROMISE

**Nicola Ferro**

*Dipartimento di Ingegneria dell'Informazione – Università degli Studi di Padova*

## Motivazioni

Con una popolazione di più di 500 milioni di abitanti nei suoi 27 stati membri in cui vengono parlate più di 80 lingue in 3 alfabeti, senza contare quelle derivanti dall'immigrazione, i cittadini e le aziende dell'Unione Europea necessitano di sistemi informativi in grado di consentire loro di interagire con contenuti culturalmente, politicamente e linguisticamente diversi e spesso disponibili in svariati media<sup>1</sup>.

In questo contesto, i sistemi di gestione e accesso all'informazione, tra cui i motori di ricerca e le biblioteche digitali, divengono ogni giorno più complessi in quanto si trovano a dover soddisfare tipologie di utente e supportare attività sempre più variegata ed articolate. Infatti, gli utenti non percepiscono più come "inviolabili" i confini tra lingue e media diversi ma considerano l'accesso all'informazione come intrinseco al processo di comunicazione, rivestendo non solo il ruolo tradizionale di consumatori di informazione ma anche e sempre più spesso di produttori di informazione.



Ad esempio, nel caso del patrimonio culturale, gli utenti desiderano sistemi di biblioteche digitali che permettano loro di interrogare simultaneamente cataloghi di biblioteche, archivi e musei per accedere a libri, scansioni di manoscritti, riproduzioni di quadri, opere musicali, documenti audio-visivi, e così via. Tali sistemi dovrebbero fornire agli utenti la possibilità di "contestualizzare" l'informazione recuperata, collegandola con al-

<sup>1</sup> Commission of the European communities. *Multilingualism. An asset for Europe and a shared commitment*. COM (2008) 566, 2008.

tre risorse pertinenti disponibili su Web, arricchendola con annotazioni e recensioni fatte da altri utenti, e personalizzandola in base alle preferenze dell'utente.

Come ulteriore esempio, nel caso della proprietà intellettuale e dei brevetti, i sistemi si trovano a dover gestire collezioni di documenti intrinsecamente multilingue, in quanto

lo stesso brevetto può essere stato registrato in nazioni diverse, e multimediali, in quanto immagini, grafici e diagrammi fanno parte integrante dei brevetti. Gli utenti hanno, ad esempio, l'esigenza di recuperare tutti i brevetti che potrebbero invalidare un nuovo brevetto e tralasciare anche un solo brevetto pertinente può avere delle notevoli ricadute economiche.

La progettazione e sviluppo dei sistemi di gestione e accesso all'informazione multilingua e multimediale diviene di conseguenza un'attività sempre più complessa e che richiede una costante verifica per assicurare che essi siano conformi ai requisiti specificati e garantiscano le prestazioni attese dagli utenti.

In questo processo di progettazione e sviluppo svolge un ruolo centrale la *valutazione sperimentale* di tali sistemi, cioè l'attività volta a verificare e quantificare le effettive prestazioni di tali sistemi rispetto a dei "campioni" di riferimento: è in quest'area che si colloca il progetto *Participative Research labORatory for Multimedia and Multilingual Information Systems Evaluation* (PROMISE)<sup>2</sup>,

### Obbiettivi e strumento

PROMISE mira a far progredire lo stato dell'arte nella valutazione sperimentale dei sistemi di gestione e accesso all'informazione multilingue e multimediale al fine di supportare individui e organizzazioni di ricerca e commerciali, ma anche comunità che progettano, sviluppano, adottano e migliorano tali sistemi.

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.

PROMISE è una rete di eccellenza finanziata nell'ambito del settimo programma quadro della Commissione Europea, cui partecipano 10 partner scientifici e industriali a livello Europeo tutti leader del settore:



<sup>2</sup> PROMISE, <http://www.promise-noe.eu/>.

- Università degli Studi di Padova (coordinatore), Italia;
- Sapienza Università di Roma, Italia;
- University of Amsterdam, Olanda;
- Humboldt-Universität zu Berlin, Germania;
- Swedish Institute of Computer Science (SICS), Svezia;
- University of Applied Sciences Western Switzerland, Svizzera;
- Zurch University of Applied Sciences, Svizzera;
- Vienna University of Technology, Austria;
- Centre for the Evaluation of Language Communication Technology (CELCT), Trento, Italia;
- Evaluations and Language resources Distribution Agency (ELDA), Francia.

Il progetto ha una durata triennale a partire da settembre 2010 ad agosto 2013. Lo strumento della rete ha l'obiettivo specifico di unire e amalgamare competenze e linee di ricerca, precedentemente separate, al fine di affrontare una problematica emergente e per la quale non sono già disponibili né soluzioni né lo spettro di competenze richiesto. La rete si propone di sviluppare le metodologie, infrastrutture, e protocolli necessari per far progredire la valutazione sperimentale dei sistemi di gestione e accesso all'informazione multilingue e multimediale così come svolgere attività di formazione e scambio di giovani ricercatori e organizzare eventi di disseminazione e trasferimento dei risultati raggiunti alle comunità applicative interessate e all'industria.

### Attività

PROMISE offrirà un laboratorio virtuale ed aperto per condurre attività di ricerca e sperimentazione collaborative, facilitando l'accesso ai dati sperimentali prodotti, tenendo in considerazione le problematiche di conservazione e preservazione di tali dati, favorendo la loro analisi e riutilizzo, offrendo strumenti innovativi per visualizzare tali informazioni e interagire attivamente con esse.

In particolare PROMISE si focalizza su quattro attività:

- **organizzazione di regolari attività di valutazione:** PROMISE porterà avanti le attività del *Cross-Language Evaluation Forum (CLEF)*<sup>3</sup>, una delle tre maggiori campagne di valutazione a livello internazionale, focalizzando maggiormente le sue attività su casi d'uso innovativi e incentrati sulle effettive esigenze dell'utenza;
- **incremento dell'automazione nella valutazione:** PROMISE svilupperà le metodologie e l'infrastruttura software necessaria per trasformare la valutazione da un processo "artigianale" ad un processo "industriale", aumentando il numero e la dimensione degli esperimenti condotti e diminuendo lo sforzo necessario per condurli;

<sup>3</sup> CLEF, <http://www.clef-campaign.org/>



Figura 1: Attività principali della rete di eccellenza PROMISE

- **collaborazione e ri-uso dei dati sperimentali prodotti:** l’infrastruttura sviluppata consentirà un’efficiente gestione ed accesso ai dati sperimentali prodotti, sia fornendo strumenti per arricchirli, come le annotazioni degli utenti, e tracciarne la storia e provenienza, sia sviluppando tecniche di rappresentazione e visualizzazione di tali dati al fine di consentirne una migliore interpretazione e confronto;
- **trasferimento di conoscenza e formazione:** PROMISE favorirà lo sviluppo di competenze e la formazione di giovani ricercatori nel settore, disseminerà conoscenza, strumenti, e soluzioni ai problemi, faciliterà la partecipazione di organizzazioni commerciali e industriali e darà origine ad una comunità con esperienze e competenze multidisciplinari.

### **Infrastruttura per la valutazione sperimentale**

La valutazione sperimentale viene usualmente condotta nel contesto di campagne internazionali di valutazione su larga scala al fine di favorire la condivisione delle risorse, la comparabilità e il confronto tra i risultati sperimentali, le tecniche e gli algoritmi, e la formazione di una comunità con le competenze multidisciplinari necessarie per sviluppare i sistemi di accesso all’informazione multilingua e multime-

diale. Le maggiori campagne di valutazione sperimentale a livello internazionale sono la *Text REtrieval Conference* (TREC)<sup>4</sup> organizzata dal *National Institute for Standards and Technology* (NIST) negli Stati Uniti, il *Cross-Language Evaluation Forum* (CLEF) in Europa, e il *NII-NACSIS Test Collection for IR Systems* (NTCIR)<sup>5</sup> organizzato dal *National Institute of Informatics* (NII) in Giappone e Asia. Inoltre, stanno sorgendo nuove iniziative di questo genere per dare voce e supporto a nuove comunità e per indirizzare esigenze specifiche, come ad esempio il *Forum for Information Retrieval Evaluation* (FIRE)<sup>6</sup> in India.

In questo contesto, le campagne di valutazione sperimentale forniscono un contributo significativo all'avanzamento dello stato dell'arte della ricerca nel settore e all'innovazione industriale. Inoltre, come riportato in uno studio condotto dal NIST riguardante l'impatto economico di TREC, per ogni dollaro investito dal NIST e dai suoi associati in TREC, sono stati prodotti da 3,35 a 5,07 dollari in benefici per la comunità di ricerca e il settore in generale<sup>7</sup>.

Durante la loro vita, le campagne di valutazione sperimentale su larga scala hanno prodotto e producono una grande quantità di dati scientifici di estremo valore. Infatti, questi dati forniscono le basi per tutta la produzione scientifica successiva e per lo sviluppo dei sistemi. Inoltre, questi dati hanno valore anche da un punto di vista economico se si considera che, sempre nello stesso studio menzionato in precedenza, il NIST stima in 30 milioni di dollari l'investimento complessivo in TREC e di cui i dati scientifici prodotti sono una parte essenziale.

Tuttavia, minore attenzione è stata posta fino ad oggi alla modellazione, gestione, cura, accesso, preservazione e riuso di tali dati scientifici, anche se l'importanza generale dei dati scientifici è stata evidenziata da diverse istituzioni come, ad esempio, la Commissione Europea<sup>8</sup>, il *National Scientific Board* (NSB)<sup>9</sup> negli Stati Uniti, e il *Australian Working Group on Data for Science*<sup>10</sup> in Australia.

Uno degli obiettivi di PROMISE è quindi sviluppare e fornire una infrastruttura ed un ambiente unificati in cui collezionare i dati scientifici e gli strumenti a sup-

<sup>4</sup> TREC, <http://trec.nist.gov/>.

<sup>5</sup> NTCIR, <http://research.nii.ac.jp/ntcir/>.

<sup>6</sup> FIRE, <http://www.isical.ac.in/~cia/>.

<sup>7</sup> Brent R. Rowe - Dallas W. Wood - Albert L. Link - Diglio A. Simoni, *Economic Impact Assessment of NIST's Text REtrieval Conference (TREC) Program*, RTI Project Number Q211875, RTI International, USA, <http://trec.nist.gov/pubs/2010.economic.impact.pdf>, 2010.

<sup>8</sup> Commission of the European communities. *Communication from the Commission to the European Parliament, the Council and the European economic and social committee on scientific information in the digital age. Access, dissemination and preservation*, COM (2008) 56, 2007.

<sup>9</sup> National Science Board, *Long-Lived digital data collections: Enabling research and education in the 21st century (NSB-05-40)*, National Science Foundation (NSF), <http://www.nsf.gov/pubs/2005/nsb0540/>, 2005.

<sup>10</sup> Working Group on Data for Science, *From data to wisdom. Pathways to successful data management for Australian science*. Report to Prime minister's science, engineering and innovation council (PMSEIC), <http://www.innovation.gov.au/Science/PMSEIC/Documents/FromDataToWisdom.pdf>, 2006.

porto della valutazione sperimentale per le comunità di utenti, individui, e organizzazioni commerciali che si occupano di progettare, realizzare, migliorare, adottare e utilizzare i sistemi di accesso all'informazione multilingue e multimediale.



Figura 2: Esempio di funzionalità supportate dal sistema di biblioteca digitale per dati scientifici DIRECT.

Il primo passo in questa direzione è rappresentato dalla realizzazione del sistema di biblioteca digitale per la gestione dei dati scientifici prodotti dalla valutazione sperimentale chiamato DIRECT (*Distributed Information Retrieval Evaluation Campaign Tool*)<sup>11, 12, 13, 14</sup>, che:

<sup>11</sup> DIRECT, <http://direct.dei.unipd.it/>.

Un video con una demo del sistema è disponibile su YouTube nel canale della rete di eccellenza PROMISE all'indirizzo: <http://www.youtube.com/watch?v=fDsXDCUPkiM>

<sup>12</sup> Maristella Agosti - Nicola Ferro. *Towards an evaluation infrastructure for DL performance evaluation*, EVIA 2010. Proceedings of the 3<sup>rd</sup> international Workshop on evaluating information access, a cura di Tetsuya Sakai, Mark Sanderson e William Webber, Tokio, 2010, p. 16-24.

<sup>13</sup> 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: Evaluation of digital libraries. An insight into useful applications and methods, a cura di Giannis Tsakonas e Christos Papatheodorou, Oxford (UK): Chandos Publishing, 2009, pp. 93-120.

<sup>14</sup> Nicola Ferro - Allan Hanbury - Henning Müller - Giuseppe Santucci. *Harnessing the scientific data produced by the experimental evaluation of search engines and information access systems*. «Procedia Computer Science», 4(2011), p. 740-749.

- introduce uno schema concettuale per modellare e rendere chiaro quali siano le entità coinvolte nello spazio informativo della valutazione sperimentale, i loro attributi e le relazioni tra loro;
- fornisce le basi per sviluppare formati di metadati che consentano di descrivere il significato dei dati scientifici che vengono gestiti, di facilitare la loro condivisione, scambio e riuso nel tempo;
- adotta un meccanismo di identificazione univoca che consente la citazione esplicita dei dati scientifici, ad esempio negli articoli dove vengono utilizzati, e ne favorisce un accesso facilitato;
- gestisce i diversi aspetti di una campagna di valutazione sperimentale come, ad esempio, la gestione delle collezioni di documenti, la sottomissione degli esperimenti, il calcolo delle misure di prestazione, e così via.

PROMISE si propone, ogni anno di progetto, di rilasciare una nuova versione del sistema DIRECT con nuove ed innovative funzionalità come, ad esempio, la possibilità di annotare i dati scientifici e collaborare attivamente su di essi oppure l'adozione di tecniche avanzate di *visual analytics* per rappresentare meglio i risultati sperimentali, rendendoli più facilmente fruibili, comprensibili e comunicabili. Questo si traduce anche nell'esplorazione dell'utilizzo di dispositivi alternativi per navigare, manipolare ed interagire con i dati sperimentali, come mostrato nella figura sottostante che riporta un primo prototipo di applicazione per iPad finalizzata a confrontare i risultati sperimentali con quelli ottimali ed ideali.



Figura 3: Esempio di applicazione iPad per l'esplorazione dei risultati sperimentali.



### Attività di disseminazione e formazione

Al fine di favorire l'avanzamento della ricerca nel settore della valutazione sperimentale dei sistemi di accesso all'informazione e di consentire la conduzione di rigorose e regolari attività di valutazione, ogni anno PROMISE organizza una conferenza e una serie di laboratori dedicati a questo tema.

La prima importante iniziativa realizzata in questo ambito è stata la conferenza *CLEF 2010 (Conference on Multilingual and Multimodal Information Access Evaluation)*<sup>15</sup>, svoltasi a Padova dal 20 al 23 settembre 2010 e descritta in nel volume precedente di *Digitalia*<sup>16</sup>. È ora in corso di svolgimento e di organizzazione la conferenza *CLEF 2011*<sup>17</sup>, che si terrà ad Amsterdam dal 19 al 22 settembre 2011.

Per quanto riguarda il problema specifico della gestione dei dati scientifici prodotti dalla valutazione sperimentale, si tratta di una questione particolarmente difficile e spesso non affrontata appieno nel settore del reperimento dell'informazione. Inoltre, questo problema richiede competenze che appartengono anche ad altri settori oltre quello del reperimento dell'informazione come, ad esempio, le basi di dati e la gestione della conoscenza. A questo proposito, PROMISE organizza un workshop intitolato *Data infrastructureS for Supporting Information Retrieval Evaluation (DESIRE 2011)*<sup>18</sup> che si terrà in occasione della *20th ACM Conference on Information and Knowledge Management (CIKM 2011)* il 28 ottobre 2011 a Glasgow.

Infine, per favorire lo sviluppo di competenze multidisciplinari che abbraccino conoscenze specifiche sia per quanto riguarda la valutazione sperimentale sia per quanto riguarda la visualizzazione e rappresentazione dei risultati sperimentali, PROMISE organizza una scuola invernale dal titolo *Information Retrieval meets Information Visualization*<sup>19</sup> che si terrà a Zinal in Svizzera dal 23 al 27 gennaio 2012.

### Ricadute

Come già accennato, l'obiettivo principale di PROMISE è far avanzare lo stato dell'arte nella valutazione sperimentale dei sistemi di accesso all'informazione multilingue e multimediale.

<sup>15</sup> Clef 2010, <http://www.clef2010.org/>.

<sup>16</sup> Maristella Agosti – Nicola Ferro. Conferenza *CLEF 2010. Conference on multilingual and multimodal information access evaluation*, «*Digitalia*», 5(2010), n.2, p.176–180.

<sup>17</sup> Clef 2011, <http://www.clef2011.org/>.

<sup>18</sup> Desire 2011, <http://www.promise-noe.eu/events/desire-2011/>.

<sup>19</sup> Promise Winter School 2012, <http://www.promise-noe.eu/events/winter-school>.

In particolare, come illustrato in Figura 4, grazie all'aumentato grado di automazione che PROMISE porterà nella valutazione sperimentale, ci si aspetta una decisa riduzione dello sforzo e del lavoro umano necessari per condurre la valutazione e, al contempo, un notevole aumento nella quantità di dati prodotti dovuto al maggior numero

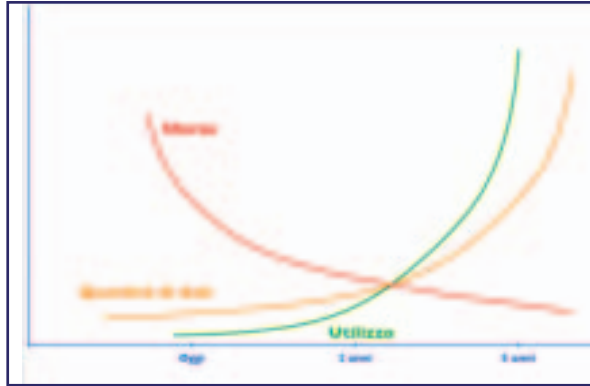


Figura 4: Impatto atteso per PROMISE.

di esperimenti che sarà possibile effettuare. Inoltre, l'infrastruttura messa a disposizione faciliterà l'utilizzo, l'accesso e l'analisi dei dati stessi favorendone una migliore comprensione. L'effetto sarà quindi una maggiore conoscenza e una più dettagliata analisi del comportamento dei sistemi di accesso all'informazione in diversi domini, come ad esempio il patrimonio culturale o la proprietà intellettuale, e, di conseguenza, un miglioramento degli stessi.

PROMISE, infine, costituirà una sorta di istituto virtuale, entità dinamica che agirà come forum in cui i ricercatori, gli sviluppatori, e gli utilizzatori di sistemi informativi multilingue e multimediali avranno la possibilità di incontrarsi, collaborare, condividere idee e conoscenze e condurre le proprie attività di valutazione.

L'istituto virtuale di PROMISE sarà costituito da:

- **basamento:** l'infrastruttura di valutazione aperta, la crescente base di conoscenza che viene acquisita grazie alle regolari attività di valutazione, la possibilità di collaborare sopra di essa, la maggiore automazione nel processo di valutazione, e le metodologie e metriche di valutazione proposte costituiscono il basamento dell'istitu-



Figura 5: Visione dell'istituto virtuale di PROMISE.

to virtuale e forniranno le fondamenta per promuovere gli avanzamenti scientifici e tecnologici attesi.

- **pilastr**i: le regolari attività di valutazione organizzate annualmente tramite le campagne CLEF e i casi d'uso realistici e la valutazione condotta intorno ad essi rappresentano i pilastri dell'istituto virtuale. Essi servono a stimolare la ricerca e lo sviluppo nel settore dei sistemi informativi multilingue e multimediali e contribuiscono a creare una comunità di ricercatori e sviluppatori con competenze multidisciplinari che abbia le conoscenze necessarie per progettare e realizzare tali sistemi complessi.
- **tetto**: il basamento e i pilastri dell'istituto virtuale forniranno il supporto necessario per progettare e sviluppare i sistemi informativi multilingue e multimediali di prossima generazione in grado di soddisfare le esigenze di utenti che stanno emergendo.

Per maggior informazioni o aggiornamenti sulle attività condotte nel progetto, si invita il lettore a visitare il sito di PROMISE all'indirizzo <http://www.promise-noe.eu/>.

### 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.