

IL PROGETTO

## Università, lanciata la sfida a Google

di STEFANO VOLTOLINI

A PAGINA 8

**Università** Da ieri tecnici a confronto a Sardinia. Bouquet: l'analisi potrà essere più accurata

# Web, l'ateneo cambia il motore

*Via al progetto «Okkam»: «Ricerca più potente di Google»*

**La fase di studio durerà trenta mesi. Il costo è di oltre otto milioni di euro, di cui cinque finanziati dall'Europa**

TRENTO — «Se affronteremo seriamente il progetto porremo le basi per il semantyc web e per la creazione di un motore di ricerca basato sulle entity invece che sulle parole chiave, che permetterebbe di superare la tecnologia di un motore di ricerca potentissimo come Google». Paolo Bouquet, docente del dipartimento ingegneria e scienza dell'informazione dell'Università di Trento, introduce così l'obiettivo del progetto europeo di cui è coordinatore. «Okkam. Enabling the web of entities», questo il nome, è iniziato il primo gennaio e durerà 30 mesi. Il costo è di 8,2 milioni di euro, di cui 5,1 finanziati dalla Commissione Europea. Ieri la prima giornata dal meeting operativo (fino a domani) svoltosi al Centro congressi di Sardinia. Presenti tutti di tutti i partner: oltre all'Università di Trento, grandi aziende come Sap, Elsevier e l'agenzia di stampa Ansa, piccole-medie imprese come Export System (Italia), National microelectronics application center (Irlanda), Inmark (Spagna), Europe Unlimited (Belgio); università e centri di ricerca come l'Università di Hannover (Germania), il Politecnico di Losanna (Svizzera), l'Università di Malaga (Spagna), la National university of Ireland at Galway (Irlanda).

Sia gli accademici che i funzionari delle aziende sono coscienti dell'importanza a livello economico del progetto. «Noi siamo attenti — nota Bouquet — a proteggere la

proprietà intellettuale sul nostro lavoro e a non comunicare troppo quello che facciamo. Se Google volesse, potrebbe arrivare a dei risultati in metà tempo rispetto a noi».

Il progetto Okkam persegue il superamento del world wide web (www) tradizionale, per arrivare al semantyc web, teorizzato da Tim Berners-Lee (co-inventore del www). «Berners-Lee — spiega Bouquet — disse che il web come ora lo conosciamo è pensato per gli esseri umani. Il passo avanti sarebbe rendere le informazioni accessibili non solo al fruitore umano, ma anche ai programmi informatici». Da qui il concetto di entity, che è il fondamento di Okkam. «L'idea è organizzare la conoscenza attorno agli oggetti reali. Ad ognuno di essi, che possono essere persone, eventi, organizzazioni, città, qualsiasi cosa, corrisponderebbe a livello informatico una ed una sola entity. Il computer vedrebbe quindi le entity e le loro relazioni, permettendo poi la ricerca tramite un motore costruito secondo questo criterio». Le entity a loro volta si basano sugli Uri (Unique resource identifier). «La condizione — prosegue — è riuscire a creare questi identificativi. Oggi ci sono tante uri diverse: è come cambiare ogni giorno il proprio numero di cellulare».

Bouquet porta alcuni esempi di entity: «Attualmente i motori di ricerca si basano sulle parole chiave. Cercando la parola Trento, vengono fuori tutte le pagine con questa parola. Ma alla parola Trento non corrisponde la stessa cosa. Nel semantyc web, Trento sarebbe una entity. Una sola uguale in tutto il mondo, anche in lingue diverse. E cercando due entity, come Trento e Comune, risulterebbero tutti i do-

cumenti in cui queste due sono in relazione». E c'è un altro esempio di possibili applicazioni: «Scrivendo su un programma "voglio fare un viaggio da Trento a New York", mettendo le date, il computer sarebbe in grado di comprendere il testo e di interagire online con i servizi organizzando tutto il viaggio».

Ad Okkam partecipa anche l'Ansa. Ieri in rappresentanza dell'agenzia di stampa nazionale c'era Angelo Marrana, responsabile ricerca e sviluppo: «L'Ansa metterà a disposizione il proprio database per una riclassificazione secondo le entity prodotte. Questo permetterà di effettuare un test significativo sui risultati del progetto». Il settore giornalistico è uno dei campi di applicazione del semantyc web. Oltre alle news, si prevede la riformulazione del materiale scientifico in rete e la creazione di nuovi strumenti per la gestione della conoscenza in ambito aziendale.

**Stefano Voltolini**

