

Lo scaffale

È noto che “troppa informazione = nessuna informazione” e che questo sembra essere uno dei maggiori problemi di Internet che potenzialmente offre infinite risorse, le quali tuttavia sono difficilmente reperibili per mancanza di strutturazione e classificazione. Molti degli sforzi degli operatori sono volti a classificare, ordinare, indicizzare questo “magma” di informazioni della rete. In particolare, per quanto riguarda le risorse Www sono stati creati e sono molto utilizzati sia da esperti che da meno esperti navigatori i “database di Www” detti anche “motori di ricerca”. Tuttavia anche il numero di questi database disponibili sulla rete è cresciuto così rapidamente che non è tanto semplice scegliere il più appropriato, conoscerne le caratteristiche e sfruttarle al me-

glio allo scopo di individuare le risorse che ci interessano. Normalmente il navigatore si affeziona a uno o due database e utilizza solo questi pur sapendo che nessuno può dirsi completo. Un articolo di M.P. Courtois, W.M. Baer e M. Stark, *Cool Tools for Searching Web - A Performance Evaluation* (“Online”, 1995; 19(6): 15-32) descrive gli elementi comuni della struttura di questi database (attenzione! Sono molto diversi da quelli che siamo soliti utilizzare online) e presenta e valuta sette tra i maggiori di questi motori di ricerca descrivendone le caratteristiche specifiche e paragonandone le performance; l'articolo presenta inoltre, sebbene più concisamente, almeno altri 15 database.

I database valutati sono tra i più noti: Cui W3 Catalog, Harvest, Lycos, Open Text Web Index, WebCrawler, World Wide Web Worm, Yahoo. Rimando all'istruttiva lettura dell'articolo per i dettagli e ri-

porto le conclusioni degli autori che, con le ovvie riserve valide per qualsiasi test di questo tipo, attribuiscono a Open Text la palma del migliore motore di ricerca seguito da Lycos. Al di là dei risultati, comunque, gli autori forniscono un utile check list per l'uso e la valutazione dei database di Www:

- 1) prima di utilizzarli studiare bene le istruzioni e le faq (frequently asked questions);
- 2) prendere dimestichezza con più di un database (nessuno contiene “tutto”);
- 3) valutare sempre molto attentamente i risultati ottenuti;
- 4) quando ci si muove in un campo nuovo, preferire i database che indicizzano il testo completo e/o quelli suddivisi in categorie;
- 5) evitare di utilizzare i database nelle ore di “affollamento” della rete;
- 6) tenersi aggiornati sugli sviluppi di questi “motori” e sulla nascita di nuovi.

Chi vuole approfondire invece

in particolare la propria conoscenza di uno dei più popolari database di Www, Yahoo, trova informazioni nell'articolo di D. Lester (“Database” 1995; 18(6)† :46-50) che ne traccia un profilo completo.

A conclusione di questa panoramica sui database di Www un accenno a quella che si può certamente definire un'ulteriore evoluzione di questo campo. Si tratta del prodotto di Ibm “InfoMarket” che consente di ricercare contemporaneamente con un'unica interfaccia alcuni di questi database (es. Yahoo, Open Text Index, ecc.), il contenuto di oltre 13.000 Usenet news groups ed altre fonti ancora (es. McKinley Inter-net Directory, Cia year book 1995). Il risultato è sorprendente, in pochissimo tempo avete i risultati della ricerca che può essere condotta sia in modo elementare che avanzato. L'indirizzo è <http://www.infomkt.ibm.com>

Mariateresa Pesenti