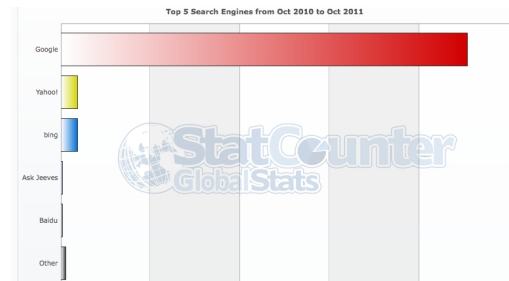


Informatica

Modulo 4 – Ricerca di informazioni nel web

I motori di ricerca



Google - breve storia

- **1998** - Page e Brin acquistano una serie di hard disk di seconda mano, e costruiscono con questi e con dei mattoncini Lego il primo server di Google. Il 7 settembre dello stesso anno, dopo aver cercato a lungo un finanziatore per poter immettere Google sul mercato ufficiale, trovano Andy Bechtolsheim (tra i fondatori di Sun Microsystems), e aprono la prima sede ufficiale di Google Inc. a Menlo Park, California (nel garage di un amico).
- **1999** - Google Inc. si trasferisce a Palo Alto (sempre in California), ha già otto dipendenti, e le ricerche effettuate si aggirano sulle 500.000 al giorno. Viene stipulato un accordo con RedHat e da allora Google ha sempre funzionato su server Linux. In questo stesso anno, Page e Brin incamerano inoltre 25 milioni di dollari grazie al finanziamento di due venture capital companies. Viene creato il Googleplex (un cluster, ovvero un "grappolo" di server, sempre in funzione) a Mountain View, California. Finalmente il motore di ricerca esce dalla fase di test ed entra a tutti gli effetti nel mercato. Il volume di ricerche ha raggiunto i 3.000.000 al giorno ed è in continua crescita.

Interfaccia



Tipi di ricerca

- Testo in pagine web;
- Immagini, sempre su pagine web;
- All'interno di gruppi di discussione Usenet (newsgroups) con la funzione gruppi;
- All'interno di determinate categorie predefinite con la funzione directory;
- All'interno dei servizi di informazione con la funzione news.

Ricerche con selezione

- Di norma, la ricerca verrà effettuata su tutto il web; possiamo però restringerla alle sole pagine in lingua locale, oppure a quelle localizzate geograficamente nel proprio stato di appartenenza (per noi l'Italia) tramite apposita selezione.

Indice di popolarità

- Il PageRank (letteralmente, “classifica delle pagine”, ma più significativamente denominato “indice di popolarità”) è lo strumento che permette a Google di ordinare i risultati delle nostre ricerche.
- Indice determinato da:
 - Quantità di link presenti su altre pagine web;
 - Quantità di link presenti su pagine con PageRank alto;
 - Rilevanza in base ai criteri di ricerca.

Funzioni di ricerca

- L’operatore “AND” permette di selezionare tutte le pagine contenenti tutti i termini indicati nel campo di ricerca. E’ l’opzione di default quindi non c’è comunque bisogno di indicarlo.
- L’operatore “OR” permette di selezionare tutte le pagine contenenti almeno uno dei termini indicati nel campo di ricerca.

Operatori di ricerca

- L’operatore “-” (meno) permette di specificare quali termini associati ad un primo termine generico, vadano ignorati nella ricerca. In pratica, indicando ad esempio nel campo di ricerca “sport –basket” otterremo tutte le pagine in cui sia presente la parola “sport” ad eccezione di quelle in cui sia presente anche la parola “basket”.
- L’operatore “+” (più) permette di rendere la ricerca sensibile ai simboli grafici (accenti, diacritici, ecc.). Ad esempio se vogliamo ricercare precisamente la parola “portò” (passato remoto del verbo portare, terza persona singolare), basterà digitare il simbolo “+” immediatamente prima della parola (“+portò”).
- Nota bene: in entrambi i casi tra l’operatore e il termine da ricercare non ci devono essere spazi!

Altri operatori

- Le virgolette: qualora si voglia ricercare un preciso insieme di parole, o una frase esatta, basterà racchiudere i termini di ricerca tra virgolette.
- L’operatore “*” (asterisco) funge da carattere jolly. Digitando ad esempio nel campo di ricerca la stringa “televi*” Google selezionerà tutte le pagine in cui sia presente una parola che inizi con “televi”. Qualsiasi carattere o serie di caratteri segua, è irrilevante.

Ricerca avanzata

- Se non si vuole scrivere una serie lunga di parole unite ad operatori di ricerca, è possibile utilizzare la “Ricerca Avanzata”, che permette di utilizzare tutte le opzioni già descritte in maniera intuitiva.
- La ricerca avanzata offre inoltre diverse altre opportunità, come la selezione della lingua, del formato del file, la ricerca all’interno di un particolare sito web...

Operatori avanzati

- Operatore “site:” - limita la ricerca ad un particolare indirizzo/dominio (es.: “site:www.rai.it”).
- Operatore “link:” - trova tutte le pagine che hanno un link all’indirizzo/dominio indicato di seguito.
- Operatore “related:” – elenca una serie di siti considerati “simili” all’indirizzo/dominio indicato di seguito.
- Operatore “filetype:” – permette di specificare che tipo di file si sta cercando.