



Lecția MOODLE 1. Căutarea resurselor pe Internet

Motoare de căutare

Cele mai utilizate, în rândul instrumentelor de căutare, sunt motoarele de căutare (search engines). Aceste programe vizitează pagini Web, analizează textul și cuvintele cheie și le stochează în baza de date a motorului de căutare. Când un utilizator trimite o cerere de căutare, motorul consultă baza de date proprie și extrage adresele care conțin cuvintele cheie specificate de interogator (utilizator), creând un catalog. Catalogul va fi transmis spre vizualizare, pagină cu pagină, către acel browser care a transmis căutarea. Dintre multele motoare de căutare, cele mai cunoscute sunt:

- Google (www.google.com);
- Excite (www.excite.com);
- Bing (www.bing.com);
- Lycos (www.lycos.com/).

Orice motor de căutare are trei elemente majore:

- **robotul de căutare:** (engleză - robot, spider sau web crawler) este un program care vizitează paginile web, le citește și apoi urmărește legăturile către alte pagini.
- **index sau catalog:** unde sunt trecute informațiile despre fiecare pagină web pe care le găsește robotul (titlul și adresa paginii, cuvinte cheie, porțiuni de text din paginile vizitate, legături către alte pagini etc.); pentru aceasta, robotul utilizează un fișier text (robots.txt) aflat în rădăcina site-ului care poate permite sau restricționa roboții motoarelor de căutare să acceseze paginile site-ului.
- **software-ul de căutare:** care permite căutarea printre multitudinea de pagini înregistrate în index precum și afișarea rezultatului sub forma unei liste de legături, în ordinea relevanței

Motorul de căutare Google și algoritmul de lucru al robotului motorului de căutare Google

Principial, există două modalități mai importante de indexare a informației în timpul generării bazelor de date web: indexarea full-text, respectiv indexarea “manuală” (human indexing).

Indexarea full-text. Așa cum îi spune și numele, indexarea full-text se caracterizează prin includerea tuturor cuvintelor dintr-o pagină în baza de date pentru căutare, cu ajutorul unor programe speciale de calculator numite “păianjeni sau roboți” (spiders sau robots). Altavista și Google folosesc pe scară largă această tehnică de indexare. Indexarea full-text ne permite astfel să regăsim toate referirile la un anumit termen din documentul indexat. Pe de altă parte însă, căutarea după termeni generici în baze de date web ce folosesc indexarea de tip full-text va duce la returnarea a numeroase pagini cu informație „sterilă”, cu alte cuvinte nerelevantă pentru autorul căutării.

Indexare “manuală”. Spre deosebire de tehnica de indexare de mai sus, în cazul indexării manuale (human indexing) o persoană examinează paaginile ce urmează a fi indexate și decide asupra câtorva cuvinte (fraze) cheie ce descriu cel mai bine informația conținută în respectiva pagină. Aceasta permite utilizatorului să regăsească mai multe legături utile în urma căutării, tocmai pentru că un om și o nu o mașină a ales cuvintele cheie ce au fost incluse în indexul bazei de date.

Motoare de căutare - directoare, metamotoare de căutare, motoare de căutare semantice.

Un serviciu de directoare web (numit uneori și subject directory - director de subiecte) este o colecție de pagini Web selecționate și organizate ierarhic în categorii de subiecte de către un editor uman. Mai mult, un concept și mai selectiv este cel de bibliotecă virtuală, care este un director web ce include legături spre pagini cu informație de înaltă specializare, pe domenii alese de editor (bibliotecarul virtual). Este evident faptul că serviciile de directoare acoperă și indexează o porțiune mult mai mică din paginile WEB existente, comparativ cu motoarele de căutare. Dar folosirea lor poate duce la regăsirea unor rezultate ale căutării mult mai relevante pentru utilizator.

Trebuie menționat însă că o serie de motoare de căutare sunt de fapt unelte hibride, fiind în același timp atât motoare de căutare cât și servicii de directoare (Google™, de exemplu, unul



dintre cele mai cunoscute motoare de căutare, are și un serviciu de directoare, bazat pe soluția Open Directory Project).

Dintre serviciile de directoare web, mai cunoscute sunt cele de mai jos:

- Yahoo! (www.yahoo.com);
- Google™ Directory (www.google.com);
- Open Directory Project (ODP) (<http://dmoz.org>);
- Zeal (www.zeal.com) JoeAnt (www.joeant.com);
- Gimpsy (www.gimpsy.com).

Un alt instrument de căutare este meta-motorul de căutare. Metamotoarele de căutare sunt instrumente care utilizează căutarea informațiilor solicitate de utilizatori cu ajutorul mai multor motoare de căutare. Rezultatele acestor căutări se obțin mai lent, dar au șanse de succes mai mari, deoarece centralizează rezultatele obținute, elimină adresele duplicate, apoi le ordonează după categorii.

Din cele menționate mai sus, reiese clar că motoarele de meta-căutare NU au propriile baze de date web, ci le folosesc pe cele ale serviciilor apelate. Tot din acest motiv, motoarele de meta-căutare nu permit înscrierea manuală a unei pagini (site) în baza de date, lucru posibil în cazul motoarelor de căutare clasice.

Unele dintre cele mai cunoscute motoare de meta-căutare sunt:

- Metasearch (www.metasearch.com);
- Ask (www.search.ask.com);
- MetaCrawler (www.metacrawler.com);
- Dogpile (www.dogpile.com);
- Webcrawler (www.webcrawler.com);
- Infonetware (www.infonetware.com);
- iBoogie (iboogie.com.cutestat.com) .

Motorul de căutare Academic Google (Google Scholar)

Google Scholar, o variantă a motorului de căutare care se concentrează asupra materialelor științifice care au fost revizuite de către cercetători sau profesori. Google Scholar este perfect pentru cei care vor să alcătuiască o lucrare științifică și pentru cei care caută materiale pentru o dezbateră.

Navigatorul (browser-ul) Google Chrome

Google Chrome este unul dintre cele mai utilizate și apreciate browsere web la ora actuală. Acesta folosește o tehnologie sofisticată combinat cu o interfață foarte simplă pentru a putea fi realizate căutări rapide, sigure și ușoare. Apărut pe piață în 2008, acest browser a ajuns în topurile motoarelor de căutare.

Unul dintre punctele forte ale acestui browser și care îl diferențiază de celelalte motoare de căutare este viteza extraordinară a acestuia. Cei de la Google au investit mult timp în dezvoltarea acestui browser pentru a-l face mai rapid cu ajutorul motorului de procesare JavaScript. Viteza motorului de căutare Google Chrome împreună cu interfața ușor de utilizat al acestuia, transformă acest browser în unul greu de depășit în ceea ce privește competiția.

Un alt lucru care îl face să fie atât de îndrăgit de utilizatori este însăși simplitatea interfeței sale, acesta fiind alcătuit din 3 rânduri de instrumente.

În ceea ce privește intimitatea pe care o poate acorda Google Chrome datorită modului Incognito prin care utilizatorul poate căuta informații pe internet fără a realiza un istoric al paginilor vizitate, eliminând cookie-urile ce pot urmări apoi calculatorul. Google Chrome are de asemenea inclusă tehnologia de Safe Browsing care anunță utilizatorul atunci când este întâlnit un website suspicios sau care conține activități dăunătoare pentru calculator. Update-urile care au loc regulat asigură securitate împotriva elementelor apărute recent.

Pentru a avea o experiență cât mai plăcută în utilizarea acestui browser, Google Chrome pune la dispoziție o serie de aplicații, teme și setări pentru a realiza un motor de căutare unic, care să se plieze cât mai bine pe preferințele utilizatorului. Utilizatorul se poate loga cu contul Google pentru a avea la îndemână toate contactele, preferințele, istoricul și de asemenea acces la toate instrumentele Google.