

Modele de căutare pentru articole multimedia

Ionuț-Dragoș Roban

Rezumat

Lucrarea de față presupune o aplicație complexă ce urmărește regăsirea informațiilor, pe web, bazate pe un subiect predefinit de către autor, mai exact implementarea unui robot de căutare ce facilitează munca utilizatorului de a căuta informațiile dorite pe web.

Aplicația propusă are prestabilit un set de informații pe baza căruia se extrag de pe web, dintr-un număr restrâns de locații/domenii, date cu privire la tema specificată. În ceea ce privește ansamblul de surse de unde se extrag informațiile, această lucrare își propune restrângerea domeniului de căutare, caracterul documentelor căutate fiind bazat pe subiectul celor de tip multimedia, știri, noutăți, ziare online.

Pentru împlinirea scopului ca regăsirea să fie cât mai riguroasă, implementarea robotului web vine în avantajul regăsirii volumelor mari de informații reprezentate prin documente web. În cadrul lucrării, volumele mari de date extrase de către robotul web, ce navighează automat, sunt procesate și supuse unui algoritm de indexare. Nevoia de un astfel de algoritm vine tocmai din definiția indexării ce reprezintă aplicarea unei relații de ordine asupra unui set de entități pentru a putea fi regăsite cu ușurință.

În ceea ce privește rezultatul lucrării, comportamentul este similar unui portal web ce reprezintă un website având la bază un sistem funcțional ce realizează schimbul de informații între aplicații/sisteme independente din punct de vedere al limbajului de care sunt descrise, presupunând anumite proceduri unificate de regăsire a informațiilor. Lucrarea are la bază modelul ce descrie un vortal. Un vortal este o particularizare a conceptului de „Portal Web” ce simulează un comportament bazat pe un subiect concret, în care se restricționează scopul regăsirii informațiilor în funcție de tema/semnificația celor dorite, în cazul de față fiind multimedia, știri, articole. Astfel, produsul software rezultat oferă următoarele avantaje:

- performanță în căutarea și indexarea documentelor web
- funcționalitate simplă ce folosește doar resurse strict necesare
- posibilități de extindere
- portabilitate
- regăsirea ușoară și rapidă a documentelor

Pentru a elimina ambiguitățile pe care le pot crea unele funcționalități inutile, interacțiunea cu utilizatorul se limitează la accesul unei simple pagini web, prin intermediul căreia acesta are posibilitatea de a regăsi anumite documente, sub formă de pagini web, indexate de către sistemul de față, proces similar unui motor de căutare clasic. Afișarea rezultatelor se face, de asemenea, într-o manieră simplă și într-un cadru ușor de vizualizat și de accesat.