(51) International Patent Classification: G06F 7/00

(21) International Application Number: PCT/US01/43248

(22) International Filing Date: 20 November 2001 (20.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

(71) Applicant (for all designated States except US):
SINGINGFISH.COM [US/US]; 2401 Fourth Avenue, Suite 400, Seattle, WA 98121 (US).

(72) Inventors; and


(54) Title: A SYSTEM AND PROCESS FOR MEDIATED CRAWLING

(57) Abstract: A system and method for searching networked based content limits searching unnecessary content. The first time a web site is encountered, an exhaustive search is conducted (522), and a site map (300) is generated (524) and the URL of the web site is added to a directory of encountered web sites (526). The next time the web site is encountered, the system utilizes the site map and directory to search only for relevant content (614). Web sites are revisited, in accordance with information derived from previous visits, to conduct subsequent exhaustive searches in order to update the site map and directory. A site map includes a structured data storage format, wherein content of the web site is organized in levels.
Published: with international search report

(88) Date of publication of the international search report:
19 September 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
# INTERNATIONAL SEARCH REPORT

**Classification of Subject Matter**

<table>
<thead>
<tr>
<th>IPC</th>
<th>US CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>C06B 7/00, 17/00</td>
<td>707/104, 707/202; 345/302</td>
</tr>
</tbody>
</table>

According to International Patent Classification (IPC) or to both national classification and IPC

**Fields Searched**

Minimum documentation searched (classification system followed by classification symbols)

- U.S.: 707/1-10, 100-104, 200-204; 711/118-216; 345/302; 379/202

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database consulted during the international search (name of database and where practicable, search terms used)

West DB, PALM

**Documents Considered to be Relevant**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 6,055,543 A (Christensen et al.) 25 Apr. 2000 (25.04.2000), Abstract, Figures 1-13, col. 3, lines 28-41; col. 4, lines 1-7; col. 4, line 51-; col. 5, line 6; col. 6, line 6-18; col. 7, lines 30-34; lines 49-64; col. 7, lines 49-55.</td>
<td>1-24</td>
</tr>
<tr>
<td>Y</td>
<td>US 6,128,627 A (Matis et al.) 03 October 2000 (03.10.2000), Abstract; col. 5, lines 23-37, lines 56-64; col. 6, lines 58-61; Figures 8-12.</td>
<td>25-32</td>
</tr>
<tr>
<td>A</td>
<td>US 6,067,552 A (Tommy Yu) 23 May 2000 (23.05.2000), All</td>
<td>1-32</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. See patent family annex.

Date of the actual completion of the international search: 25 April 2002 (25.04.2002)

Date of mailing of the international search report: 30 May 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Te Y Chen
Telephone No. (703)305-6296

Form PCT/ISA/210 (second sheet) (July 1998)