Abstract

Disclosed information exploration system and method embodiments operate on a document set to determine a document cluster hierarchy. An exclusionary phrase index is determined for each cluster, and representative phrases are selected from the indexes. The selection process may enforce pathwise uniqueness and balanced sub-cluster representation. The representative phrases may be used as cluster labels in an interactive information exploration interface.
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
   IPC: G06F 7/00( 2006.01),17/00( 2006.01),17/30( 2006.01)

   USPC: 707/5,7,102,104.1,E17.089,E17.092
   According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

   Minimum documentation searched (classification system followed by classification symbols)
     U.S. : 707/5,7,102,104.1,E17.089,E17.092

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

   Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
     Google Scholar

C. 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>ZAMIR, Oren et al., Grouper: A Dynamic Clustering Interface to Web Search Results, Computer Netwokrs, Vol. 31, No. 11, May 1999, pages 1-15, especially pages 3-6.</td>
<td>1,4,8,14 and 23</td>
</tr>
<tr>
<td>Y</td>
<td>KUMAR, Sachin, et al, Personalized Profile Based Search Interface With Ranked and Clustered Display, Technical Report 01-023, Univ. of Minnesota, Dept. of Computer Science and Engineering, Minneapolis, MN, June 1, 2001, pages i-iii and 1-18, especially pages 1 and 3.</td>
<td>13</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
17 June 2008 (17.06.2008)

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (571) 273-3201

Date of mailing of the international search report
22 JUL 2008

Authorized officer
[Signature]

Telephone No. (571) 272-4102

Form PCT/ISA/210 (second sheet) (April 2007)
<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>Y</td>
<td>US 2005/0044487 A1 (BELLEGARDA et al) 24 February 2005 (24.02.2005), whole document, especially Fig. 1, Fig. 2B, and paragraphs [0011], [0024] and [0026].</td>
<td>9-12, 17-22 and 28-30</td>
</tr>
<tr>
<td>Y</td>
<td>US 6,847,966 B1 (SOMMER et al) 25 January 2005 (25.01.2005), whole document, especially Fig. 10, col. 1 lines 8-12 and col. 8 lines 46-60.</td>
<td>15</td>
</tr>
</tbody>
</table>