CROSS-LINGUAL INFORMATION RETRIEVAL

FIG. 1

Abstract: Multi-lingual search and retrieval of digital content. Embodiments are generally directed to methods and systems for creating an English language database that associates non-English terms with English terms in multiple categories of metadata. Language experts use an interface to create equivalencies between non-English terms and English terms, Boolean expressions, synonyms, and other forms of search terms. Language dictionaries and other sources also create equivalencies. The database is used to evaluate non-English search terms submitted by a user, and to determine English search terms that can be used to perform a search for content. The multiple categories of metadata may comprise structured data, such as keywords of a structured vocabulary, and/or unstructured data, such as captions, titles, descriptions, etc. Weighting and/or prioritization can be applied to the search terms, to the process of searching the multiple categories, and/or to the search results, to rank the search results.
INTERNATIONAL SEARCH REPORT

A CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 17/28 (2008.04)
USPC - 704/7

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)
USPC 704/7

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC 704/2.7.8 | 707/4,104 1

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
USPTO WEST (PGPUB USPAT USOCR EPAB JPAB), GOOGLE Scholars, select, choose, language, controlled vocabulary, structured, metadata, content management system, equivalent, keyword, term, query, translation, list

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>
</table>

D Further documents are listed in the continuation of Box C

* Special categories of cited documents
  "A" document defining the general state of the art which is not considered
  "E" earlier application or patent but published on or after the international filing date
  "L" document which may throw doubts on patentability claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  "O" document referring to an oral disclosure, use, exhibition or other means
  "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
30 June 2008 (30 06 2008)

Date of mailing of the international search report
14 JUL 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

Authorized officer
Lee W. Young
PCT Hilpdesk 571-272-4300
PCTSP 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)