(54) Title: SCHEMA AND ETL TOOLS FOR STRUCTURED AND UNSTRUCTURED DATA

(57) Abstract: A system and method of making unstructured data available to structured data analysis tools. The system includes middleware software that can be used in combination with structured data tools to perform analysis on both structured and unstructured data. Data can be read from a wide variety of unstructured sources. The data may then be transformed with commercial data transformation products that may, for example, extract individual pieces of data and determine relationships between the extracted data. The transformed data and relationships may then be passed through an extraction/transform/load (ETL) layer and placed in a structured schema. The structured schema may then be made available to commercial or proprietary structured data analysis tools.
INTERNATIONAL SEARCH REPORT

INTERNATIONAL SEARCH REPORT

PCT/US 06/25814

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8): G06F 17/00 (2006.01)

USPC: 707/102

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)

IPC(8): G06F 17/00 (2006.01)

USPC: 707/102

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

USPC: 707/2, 100-102, 104.1; 700/47, 52, 60; 704/243, 251; 715/500, 509, 513, 517

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>Y</td>
<td>US 2005/0108256 A1 (Wat 0691/10 et al.) 19 May 2005 (19.05.2005)</td>
<td>1-17</td>
</tr>
<tr>
<td></td>
<td>[See abstract, figs. 1-4b, Paragraphs 3, 10-13, 23, 28-31, 71, and 94-98]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[See abstract, figs. 2-4, col. 2 (line 18) - col. 5 (line 35), and col. 7 (line 20) - col. 9 (line 64)]</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>US 2003/0206201 A1 (Ly) 06 November 2003 (06.11.2003)</td>
<td>1-18</td>
</tr>
</tbody>
</table>

D. Further documents are listed in the continuation of Box C.

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 actual completion of the international search

21 November 2006 (21.11.2006)

Date of mailing of the international search report

03 JAN 2007

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

PCTHelpdesk: 571-272-4300
PCTOSP: 571-272-7774

FormPCT ASA/210 (second sheet) (April 2005)