

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 November 2007 (15.11.2007)

PCT

(10) International Publication Number
WO 2007/130675 A3

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/US2007/011022

(22) International Filing Date: 7 May 2007 (07.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/429,928 8 May 2006 (08.05.2006) US

(71) Applicant (for all designated States except US): **BLACK DUCK SOFTWARE, INC.** [US/US]; 265 Winter Street, Waltham, MA 02451 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **PEDERSEN, Palle, M.** [US/US]; 83 Harvard Avenue, #2, Brookline, MA 02446 (US).

(74) Agents: **HEFFAN, Ira V.** et al.; Goodwin Procter LLP, Exchange Place, Boston, MA 02109 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

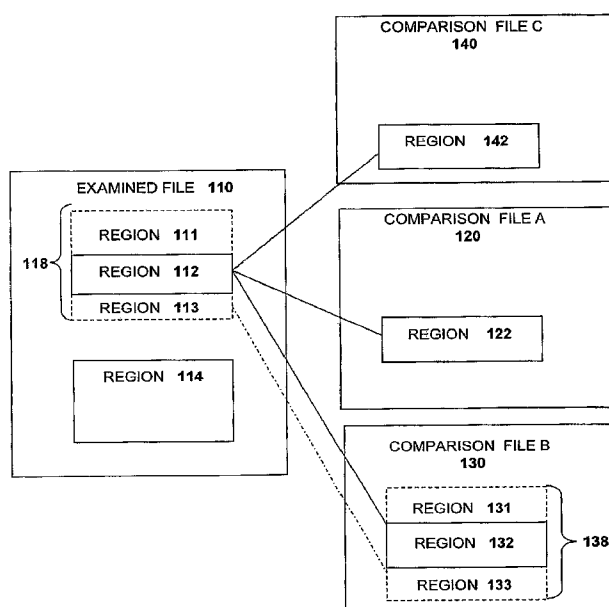
Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
10 January 2008

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.

(54) Title: METHODS AND SYSTEMS FOR REPORTING REGIONS OF INTEREST IN CONTENT FILES



(57) Abstract: A content file is examined and compared against one or more comparison files. An indication is provided that the content file is similar to the one comparison file that is the best match with the examined content file.

WO 2007/130675 A3

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/011022

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06F17/30

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)

G06F

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 practical, search terms used)

EPO-Internal, WPI Data, IBM-TDB, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/065930 A1 (SWAMINATHAN KISHORE [US] ET AL) 24 March 2005 (2005-03-24) page 1, paragraph 12 - page 6, paragraph 56	1-26
A	RICHARD COLE AND RAMESH HARIHARAN: "Verifying Candidate Matches in Sparse and Wildcard Matching" STOC 02, [Online] 19 May 2002 (2002-05-19), - 21 May 2002 (2002-05-21) XP002457005 Montreal, Quebec, Canada Retrieved from the Internet: URL:www.citeseer.org> [retrieved on 2007-10-30] the whole document	1-26

-/--

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority 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 October 2007

Date of mailing of the international search report

12/11/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Warry, Lawrence

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2007/011022

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 893 095 A (JAIN RAMESH [US] ET AL) 6 April 1999 (1999-04-06) the whole document -----	1-26
A	JAAP KAMPS, MAARTEN MARX, MAARTEN DE RIJKE, ORKUR SIGURBJÖRNSSON: "BestMatch Querying from DocumentCentric XML" SEVENTH INTERNATIONAL WORKSHOP ON THE WEB AND DATABASES (WEBDB 2004), [Online] 17 June 2004 (2004-06-17), - 18 June 2004 (2004-06-18) XP002457006 Amsterdam, NL Retrieved from the Internet: URL: http://staff.science.uva.nl/{mdr/Publications/Files/webdb2004.pdf} [retrieved on 2007-10-30] the whole document -----	1-26
A	FARID MEZIANE A, YACINE REZGUI: "A document management methodology based on similarity contents" ELSEVIER, 2003, XP002457007 Salford, UK the whole document -----	1-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2007/011022

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005065930	A1	24-03-2005	
		BR PI0413717 A	17-10-2006
		CA 2538526 A1	24-03-2005
		EP 1665095 A1	07-06-2006
		WO 2005026982 A1	24-03-2005
<hr/>			
US 5893095	A	06-04-1999	NONE
<hr/>			