

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 December 2009 (30.12.2009)

(10) International Publication Number
WO 2009/156987 A3

(51) International Patent Classification:
G06N 7/00 (2006.01) *G06N 1/00* (2006.01)

(21) International Application Number:
PCT/IL2009/000621

(22) International Filing Date:
23 June 2009 (23.06.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/129,408 24 June 2008 (24.06.2008) US
61/136,414 3 September 2008 (03.09.2008) US

(72) Inventors; and
(71) Applicants : BELENZON, Sharon [IL/IL]; 27/19 MaAgal HaShalom Street, 75801 Rishon-LeZion (IL). BE-LINSON, Liat [IL/IL]; 27/19 MaAgal HaShalom Street, 75801 Rishon-LeZion (IL).

(74) Agents: G.E EHRlich (1995) LTD. et al.; 11 Menachem Begin Street, 52521 Ramat Gan (IL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,

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, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, 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 (Art. 21(3))
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(88) Date of publication of the international search report:
15 July 2010

(54) Title: SEARCH ENGINE AND METHODOLOGY, PARTICULARLY APPLICABLE TO PATENT LITERATURE

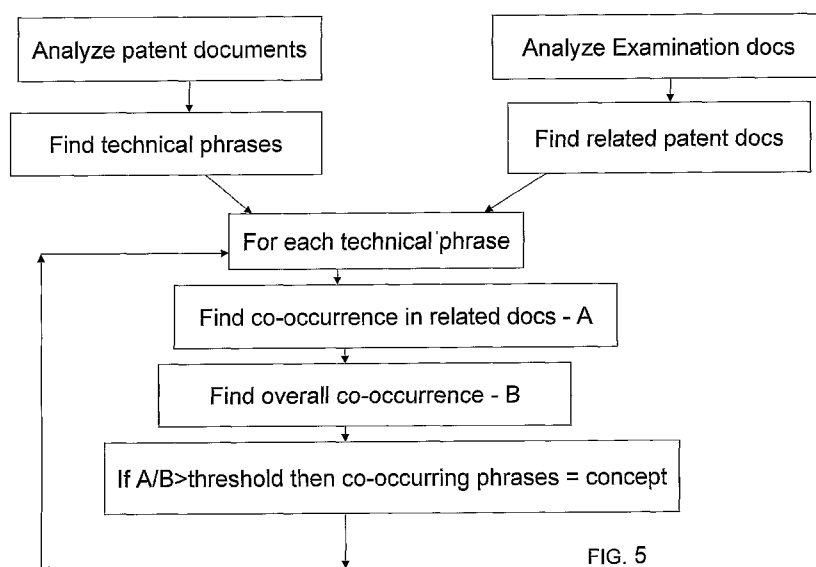


FIG. 5

(57) Abstract: A search engine for searching based on related scientific or technological concepts, comprises: a learning module for learning about relationships between technical phrases based on their rates of occurrence in related documents, therefrom to form concepts from groupings of related phrases, and a search module for searching for related documents to a query document based on occurrence in said related documents of concepts present in said query document, the learning module carrying out said learning based on a training set of documents and inter-document relations.



WO 2009/156987 A3

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2009/000621

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06N7/00 G06N1/00
ADD.

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)
G06N

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GB 2 365 569 A (IBM [US]) 20 February 2002 (2002-02-20)	1-11, 13-19, 21,22
Y	abstract figures 3,5,6,7 page 2, line 7 - page 3, line 6 page 6, line 15 - page 14, line 11	12,20
X	US 2003/026459 A1 (WON JEONG WOOK [KR] ET AL) 6 February 2003 (2003-02-06) abstract figures 3-5	1-3, 13-16
Y	GB 2 417 115 A (HEWLETT PACKARD DEVELOPMENT CO [US]) 15 February 2006 (2006-02-15) abstract	12,20
	----- -/--	

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

17 May 2010

Date of mailing of the international search report

01/06/2010

Name and mailing address of the ISA/

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

Authorized officer

Valencia, Erika

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2009/000621

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>US 5 864 855 A (RUOCCO ANTHONY S [US] ET AL) 26 January 1999 (1999-01-26) abstract column 5 - column 8 -----</p>	5,8,9, 11,12,19
A	<p>L. douglas Baker and Andrew Kachites McCallum: "Distributional Clustering of Words for Text Classification" ACM SIGIR conference on Research and development in information retrieval Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval 31 December 1998 (1998-12-31), pages 96-103, XP002582182 Melbourne, Australia ISBN: 1-58113-015-5 Retrieved from the Internet: URL:http://portal.acm.org/citation.cfm?id=290970> [retrieved on 2010-05-11] the whole document -----</p>	1-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IL2009/000621

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
GB 2365569	A	20-02-2002	CA 2329558 A1 US 6542889 B1	28-07-2001 01-04-2003
US 2003026459	A1	06-02-2003	KR 20030009704 A	05-02-2003
GB 2417115	A	15-02-2006	NONE	
US 5864855	A	26-01-1999	NONE	