

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 September 2007 (07.09.2007)

PCT

(10) International Publication Number
WO 2007/100848 A3

(51) International Patent Classification:

G06F 17/30 (2006.01)

(21) International Application Number:

PCT/US2007/005149

(22) International Filing Date:

27 February 2007 (27.02.2007)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

11/364,807 27 February 2006 (27.02.2006) US

(71) Applicant (for all designated States except US): MICROSOFT CORPORATION [US/US]; One Microsoft Way, Redmond, WA 98052-6399 (US).

(72) Inventors: WANG, Jue; One Microsoft Way, Redmond, WA 98052-6399 (US). LI, Mingjing; One Microsoft Way, Redmond, WA 98052-6399 (US). MA, Wei-Ying; One Microsoft Way, Redmond, WA 98052-6399 (US). LI, Zhiwei; One Microsoft Way, Redmond, WA 98052-6399 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, 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, LV, LY, MA, MD, 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, 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).

Declarations under Rule 4.17:

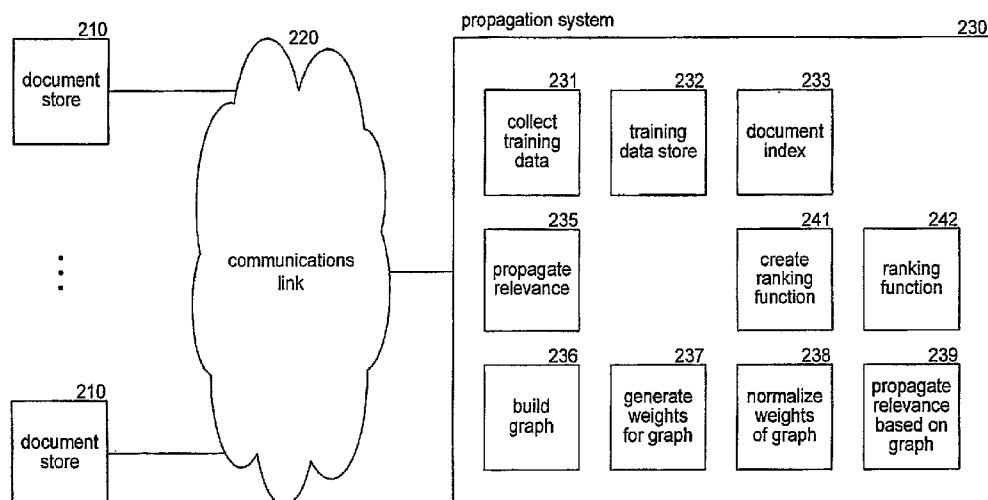
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

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

[Continued on next page]

(54) Title: PROPAGATING RELEVANCE FROM LABELED DOCUMENTS TO UNLABELED DOCUMENTS



(57) Abstract: A method and system for propagating the relevance of labeled documents to a query to unlabeled documents is provided. The propagation system provides training data that includes queries, documents labeled with their relevance to the queries, and unlabeled documents. The propagation system then calculates the similarity between pairs of documents in the training data. The propagation system then propagates the relevance of the labeled documents to similar, but unlabeled, documents. The propagation system may iteratively propagate labels of the documents until the labels converge on a solution. The training data with the propagated relevances can then be used to train a ranking function.

WO 2007/100848 A3



(88) Date of publication of the international search report:

25 October 2007

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.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2007/005149**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/30(2006.01)i**

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

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

Korean Utility models and applications for Utility models since 1975

Japanese Utility models and applications for Utility models since 1975

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

PAJ, FPD, USPAT, eKIPASS, IEEE, YAHOO, GOOGLE, Keyword: "relevance, propagation, score"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DENGYOUNG ZHOU ET AL "Ranking on Data Manifolds" MPI Technical Report (113), Max Planck Institute for Biological Cybernetics, Tübingen, Germany, June 2003, (See Abstract, pages 1-2 See following website: http://www.kyb.mpg.de/publications/pdfs/pdf2290.pdf)	1-20
Y	DENGYOUNG ZHOU ET AL "Learning with Local and Global Consistency", MPI Technical Report(112), Max Planck Institute for Biological Cybernetics, 2003. (See Abstract, pages 1-3 other publication: Advances in Neural Information Processing Systems 16, 321-328. MIT Press, Cambridge, MA, USA 2004 See following webpage: http://www.kyb.tuebingen.mpg.de/bs/people/weston/localglobal.pdf)	1-20
A	US 20040098362 A1 (ULLAS GARGI) 20 May 2004 (See Abstract, paragraph [0009]-[0010], page 2-4)	1-20



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 application or patent 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 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

14 AUGUST 2007 (14.08.2007)

Date of mailing of the international search report

14 AUGUST 2007 (14.08.2007)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

SUN Dong Guk

Telephone No. 82-42-481-8248



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2007/005149

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TAO QIN ET AL. "A study of relevance propagation for web search" Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval Pages: 408 - 415, 2005 (See Abstract, page 408 right column - page 409 left column) See following webpage: http://research.microsoft.com/asia/dload_files/group/wsm/SigIR2005/A%20Study%20of%20Relevance%20Propagation%20for%20Web%20Search.pdf	1-20
A	AZADEH SHAKERY ET AL. "Relevance Propagation for Topic Distillation UIUC TREC 2003 Web Track Experiments" NIST Special Publication 500-255: The Twelfth Text REtrieval Conference (TREC-12). Gaithersburg, Maryland, USA, 2003 (See Abstract, page2 left column - page3 left column) See following webpage : http://trec.nist.gov/pubs/trec12/papers/uillinois-uc.web.pdf	1-20
A	XIN-JING WANG ET AL. "Multi-model similarity propagation and its application for web image retrieval" Proceedings of the 12th annual ACM international conference on Multimedia Pages: 944 - 951, 2004 (See Abstract, page 947 right column- 948 right column) See following webpage : http://research.microsoft.com/asia/dload_files/group/ims/13.pdf	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2007/005149

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US20040098362A1	20.05.2004	US7107520BB	12.09.2006