

(51) International Patent Classification:
G06F 17/30 (2006.01)School of Computing, Department of Computer Science,
21 Lower Kent Ridge Road, Singapore 119077 (SG).(21) International Application Number:
PCT/SG2011/000139(74) Agent: **AMICA LAW LLC**; 30 Raffles Place, #18-03/04
Chevron House, Singapore 048622 (SG).(22) International Filing Date:
6 April 2011 (06.04.2011)(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, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/321,133 6 April 2010 (06.04.2010) US(71) Applicant (for all designated States except US): **NATIONAL UNIVERSITY OF SINGAPORE** [SG/SG];
21 Lower Kent Ridge Road, Singapore 119077 (SG).

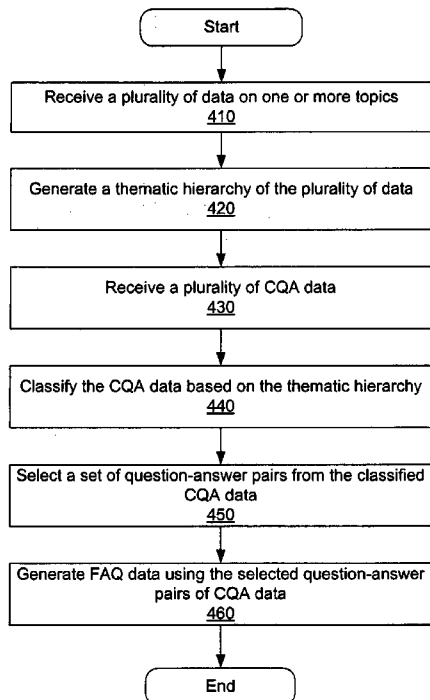
(72) Inventors; and

(75) Inventors/Applicants (for US only): **CHUA, Tat Seng**
[SG/SG]; c/o National University of Singapore, School of
Computing, Department of Computer Science, 21 Lower
Kent Ridge Road, Singapore 119077 (SG). **MING, Zhao**
Yan [CN/SG]; c/o National University of Singapore,(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,
ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), European (AL, 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, RS, SE, SI, SK,

[Continued on next page]

(54) Title: AUTOMATIC FREQUENTLY ASKED QUESTION COMPILATION FROM COMMUNITY-BASED QUESTION
ANSWERING ARCHIVE

FIG. 4

(57) Abstract: Frequently Asked Questions (FAQ) data
are generated using Community-based Question Answer-
ing (CQA) data. A thematic hierarchy generation module
receives multiple data sources and generates a thematic hi-
erarchy of the data source, where a data source has one or
more topics and a topic has one or more themes. A feature
classifier classifies multiple CQA data into one or more
themes based on the thematic hierarchy, where a CQA
data contains multiple question-answer pairs. A selection
module selects multiple question-answer pairs from the
CQA data based on the classification, measures the quality
of the selected question-answer pairs and generates FAQ
data using the selected question-answer pairs of the CQA
data.



SM, 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))*

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*