



(43) International Publication Date  
6 December 2012 (06.12.2012)

- (51) International Patent Classification:  
*G06F 17/00* (2006.01)    *G06F 3/14* (2006.01)
- (21) International Application Number:  
PCT/US2012/039533
- (22) International Filing Date:  
25 May 2012 (25.05.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
61/491,140    27 May 2011 (27.05.2011)    US
- (71) Applicant (for all designated States except US): **CTC TECH CORP.** [US/US]; 4673 York Blvd., Los Angeles, California 90041 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **DOLAN, Brian** [US/US]; 5110 Almaden Drive, Los Angeles, California 90041 (US).
- (74) Agent: **LIZZI, Thomas**; GTC Law Group LLP & Affiliates, c/o CPA Global, P.O. Box 52050, Minneapolis, Minnesota 55402 (US).
- (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, QA, RO, RS, RU, RW, 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.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, 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))
- 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:  
31 January 2013

(54) Title: CREATION, USE AND TRAINING OF COMPUTER-BASED DISCOVERY AVATARS

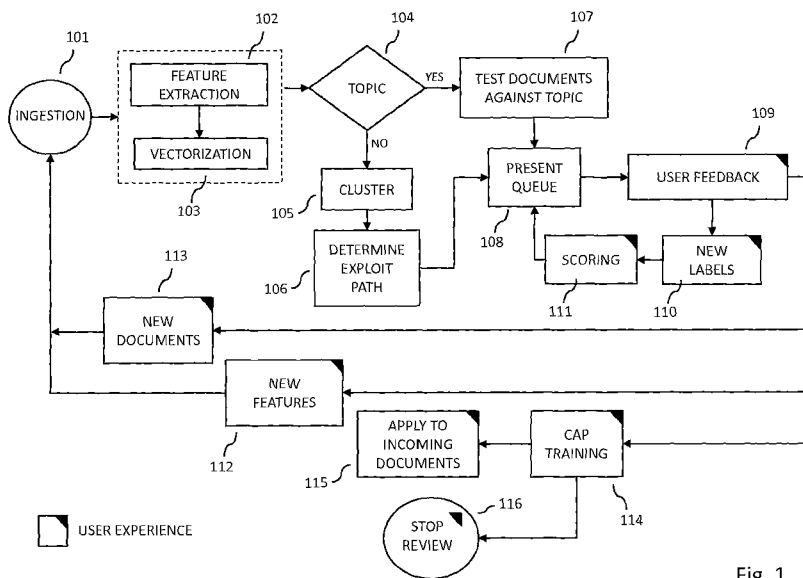


Fig. 1

(57) Abstract: In embodiments of the present invention improved capabilities are described for developing, training, validating and deploying discovery avatars embodying mathematical models that may be used for document and data discovery and deployed within large data repositories.



**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/00(2006.01)i, G06F 3/14(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)

G06F 17/00; G06F 9/44; G06F 17/30; G06F 17/28

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

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) &amp; Keywords: data cluster, score, token, queue

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009-0150370 A1 (CHRISTENSEN LARRY et al.) 11 June 2009 See paragraph 27 - paragraph 28; paragraph 67 - paragraph 75; paragraph 81 - paragraph 86; and figures 2,3,6,7.	1-20
A	US 2005-0222975 A1 (TAPAS NAYAK et al.) 06 October 2005 See paragraph 68 - paragraph 79; and figures 12-15.	1-20
A	US 2008-0262828 A1 (OCH FRANZ JOSEF et al.) 23 October 2008 See paragraph 77 - paragraph 82; and figures 7-9B.	1-20
A	US 2010-0318957 A1 (AKKIRAJU RAMA K. et al.) 16 December 2010 See paragraph 50 - paragraph 67; and figures 4-6.	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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

07 DECEMBER 2012 (07.12.2012)

Date of mailing of the international search report

**10 DECEMBER 2012 (10.12.2012)**

Name and mailing address of the ISA/KR

Korean Intellectual Property Office  
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan  
City, 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

PARK, SANG HYUN

Telephone No. 82-42-481-8263



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2012/039533**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009-0150370 A1	11.06.2009	AU 2007-248585 A1 CA 2651119 A1 EP 2013791 A2 EP 2013791 A4 SG174027A1 WO 2007-130546 A2 WO 2007-130546 A3 WO 2007-130546 A3	15.11.2007 15.11.2007 14.01.2009 20.04.2011 29.09.2011 15.11.2007 02.10.2008 15.11.2007
US 2005-0222975 A1	06.10.2005	US 2009-100022 A1	16.04.2009
US 2008-0262828 A1	23.10.2008	AU 2007-217900 A1 BR P10706404A2 CA 2631815 A1 CN 101361065 A EP 1969493 A2 EP 1969493 A4 EP 1983445 A2 EP 1983445 A3 EP 2511833 A2 JP 2009-527818 A KR 10-2008-0112192 A US 8296123 B2 WO 2007-098055 A2 WO 2007-098055 A3	30.08.2007 29.03.2011 30.08.2007 04.02.2009 17.09.2008 05.05.2010 22.10.2008 18.11.2009 17.10.2012 30.07.2009 24.12.2008 23.10.2012 30.08.2007 22.05.2008
US 2010-0318957 A1	16.12.2010	None	