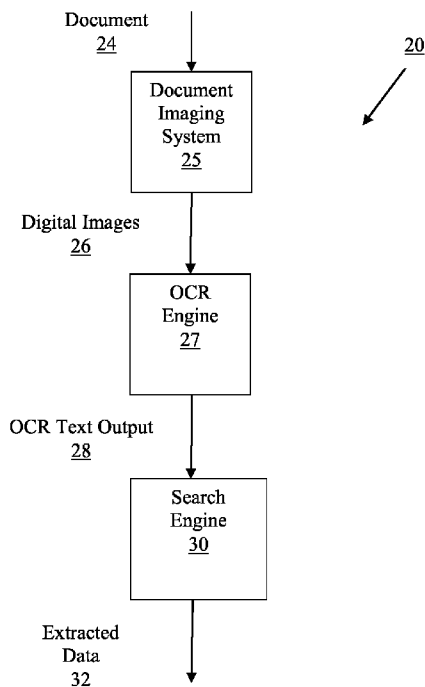




- (51) International Patent Classification: *G06K 9/60* (2006.01)
- (21) International Application Number: PCT/US2014/030904
- (22) International Filing Date: 17 March 2014 (17.03.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 61/799,091 15 March 2013 (15.03.2013) US
- (71) Applicant: MERCURY FILE, LLC [—/US]; 10 Martingale, Ste 400, Schaumburg, IL 60173 (US).
- (72) Inventors; and
- (71) Applicants : SUAREZ, Sergio, David, Jr. [US/US]; 821 N Neva, Addison, IL 60101 (US). MESKE, Joshua, Daniel [US/US]; 5121 N. East River Rd., Apt. 2h, Chicago, IL 60656 (US).
- (74) Agent: PASKY, Jonathan, R.; Pasky IP Law, 320 W Ohio St Ste 300, Chicago, IL 60654 (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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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,

[Continued on next page]

(54) Title: SYSTEM FOR METHOD FOR DATA SWEEPING USING KEYWORDS



(57) Abstract: A method for extracting data relating to a search term from a searchable text document is disclosed. The method includes executing on a processor instructions for extracting the data from the search term, wherein the search term is defined by at least one character pattern. The instructions include: searching the text document for the search term; performing a data sweep within an area around the search term within the text document; identifying at least one string within the area; comparing the at least one string to the at least one character pattern; and extracting data from the at least one string when the data of the at least one string matches the at least one character pattern.

Fig. 1





TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

— *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))*

Published:

— *with international search report (Art. 21(3))*

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

24 December 2014

INTERNATIONAL SEARCH REPORT

~~47030904-10-10-2014~~

International application No.

PCT/US 14/30904

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06K 9/60 (2014.01)

CPC - G06F17/30011

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)
USPC: 382/306; CPC: G06F17/30011; IPC(8): G06K 9/60 (2014.01)Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 707/738, 707/736, 707/722, 707/708, 707/E17.024, 707/999.003; 715/711, 715/254, 715/255; 706/934, 382/306, 382/181; CPC: G06F17/30017, G06F17/30011, G06F17/30265, G06K19/06037, H04N1/2175; IPC(8): G06K 9/60 (2014.01) (keyword limited)Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, Google Patents, IEEE; Search Terms: extracting, pulling, retrieving, sweeping, identifying, segmenting, searching, scanning
text document, database, data base, repository, catalog, string, search term, character pattern, record, data, segment zone, region,
paragraph, page, part, area, OCR, optical character recognition

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011/0047149 A1 (Vaananen) 24 February 2011 (24.02.2011), entire document especially Figure 1, steps 100 - 140, paras [0072], [0074]	1 - 2
A	US 2011/0161144 A1 (Mizuguchi et al.) 30 June 2011 (30.06.2011), entire document	1 - 2
A	US 2002/0184188 A1 (Mandyam et al.) 05 December 2002 (05.12.2002), entire document	1 - 2
A	US 2007/0300295 A1 (Kwok et al.) 27 December 2007 (27.12.2007), entire document	1 - 2

 Further documents are listed in the continuation of Box C.

* 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 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

04 September 2014 (04.09.2014)

Date of mailing of the international search report

10 OCT 2014

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774