

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
07 October 2021 (07.10.2021)



(10) International Publication Number
WO 2021/198807 A1

(51) International Patent Classification:
G06F 17/00 (2019.01)

(21) International Application Number:
PCT/IB2021/051847

(22) International Filing Date:
05 March 2021 (05.03.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/839,200 03 April 2020 (03.04.2020) US

(71) Applicant: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York 10504 (US).

(71) Applicants (for MG only): **IBM (CHINA) INVESTMENT COMPANY LTD.** [CN/CN]; 25/F, Pangu Plaza, No. 27, Central North 4th Ring Road, Chaoyang District,

Beijing 100101 (CN). **IBM DEUTSCHLAND GMBH** [DE/DE]; IBM-Allee 1, 71139 Ehningen (DE).

(72) Inventors: **BODZIONY, Michal**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **FILIP, Marcin**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **LUCZYNSKI, Marcin**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **ZATORSKI, Tomasz**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **LASKAWIEC, Andrzej**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **PIATEK, Monika**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **STUDZIENNY, Lukasz**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL).

(74) Agent: **VETTER, Svenja**; c/o IBM Deutschland GmbH, Patentwesen und Urheberrecht, IBM-Allee 1, 71139 Ehningen (DE).

(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,

(54) Title: MULTI-VALUE PRIMARY KEYS FOR PLURALITY OF UNIQUE IDENTIFIERS OF ENTITIES

Alias/PK	Row-ID	Name	EIN
CNR1#A	CRID1	ABC	
CNR2#B	CRID2	A-B-C	EIN1234
		ABC	EIN1234

FIG. 4a

Alias/PK	Row-ID	Name	EIN
CNR1#A, CNR2#B	CRID1	ABC	EIN1234

FIG. 4b

(57) Abstract: A computer-implemented method for unambiguously identifying entities in a database system may be provided. The method comprises storing data items as records with different attributes in a table of a database, storing naming rules for selected combinations of the attributes of the data items, and prioritizing the naming rules. The method also comprises determining a hash value for each of the selected combinations of the attributes of the data items, and identifying duplicate data items using the determined hash values and the prioritized naming rules.

[Continued on next page]

WO 2021/198807 A1

CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN,
KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE