



(51) International Patent Classification:
G06F 16/215 (2019.01)

(21) International Application Number:
PCT/IB2021/056095

(22) International Filing Date:
07 July 2021 (07.07.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/944,702 31 July 2020 (31.07.2020) US

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

(71) Applicants (for MG only): **IBM (CHINA) INVEST-
MENT 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: **IYENGAR, Arun**; c/o IBM Thomas J Wat-
son Research Center, 1101 Kitchawan Road, P.O. Box 218,
Yorktown Heights, New York 10598 (US). **BHAMIDI-
PATY, Anuradha**; c/o IBM Thomas J Watson Research
Center, 1101 Kitchawan Road, P.O. Box 218, Yorktown
Heights, New York 10598 (US). **PATEL, Dhavalku-
mar**; c/o IBM Thomas J Watson Research Center, 1101
Kitchawan Road, P.O. Box 218, Yorktown Heights, New
York 10598 (US). **SHRIVASTAVA, Shrey**; c/o IBM
Thomas J Watson Research Center, 1101 Kitchawan Road,
P.O. Box 218, Yorktown Heights, New York 10598 (US).
ZHOU, Nianjun; c/o IBM Thomas J Watson Research
Center, 1101 Kitchawan Road, P.O. Box 218, Yorktown
Heights, New York 10598 (US).

(54) Title: REAL-TIME DATA QUALITY ANALYSIS

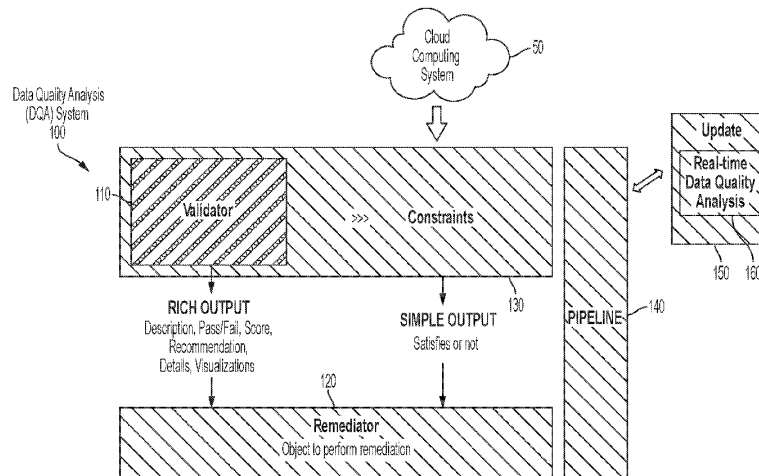


FIG. 1

(57) Abstract: A computer-implemented method for assessing data quality includes using a processor to receive a plurality of updates to data points in a data stream. The processor is further used to compute instances of a data quality metric (DQM) from the data points in the data stream. The instances of the DQM are configured to differentiate the data points in the data stream by time and assign a higher weight to the instances of the DQM computed from more recent data points in the data stream. The instances of the DQM are continuously updated as more of the data points are received by the processor while limiting cycles of the processor consumed by updating the instances of the DQM to a threshold.

[Continued on next page]

(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, 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