



- (51) **International Patent Classification:**
G06F 11/07 (2006.01) *G06F 16/215* (2019.01)
- (21) **International Application Number:**
PCT/US2022/018310
- (22) **International Filing Date:**
01 March 2022 (01.03.2022)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
63/155,148 01 March 2021 (01.03.2021) US
- (71) **Applicant:** **AB INITIO TECHNOLOGY LLC** [US/US];
201 Spring Street, Lexington, Massachusetts 02421 (US).
- (72) **Inventors:** **MARTIN, Jonathan**; 201 Spring Street, Lexington, Massachusetts 02421 (US). **WEISS, Adam**; 201 Spring Street, Lexington, Massachusetts 02421 (US).
- (74) **Agent:** **DEVRIES, Gretchen A.** et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (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, 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, JM, 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,

(54) **Title:** GENERATION AND EXECUTION OF PROCESSING WORKFLOWS FOR CORRECTING DATA QUALITY ISSUES IN DATA SETS

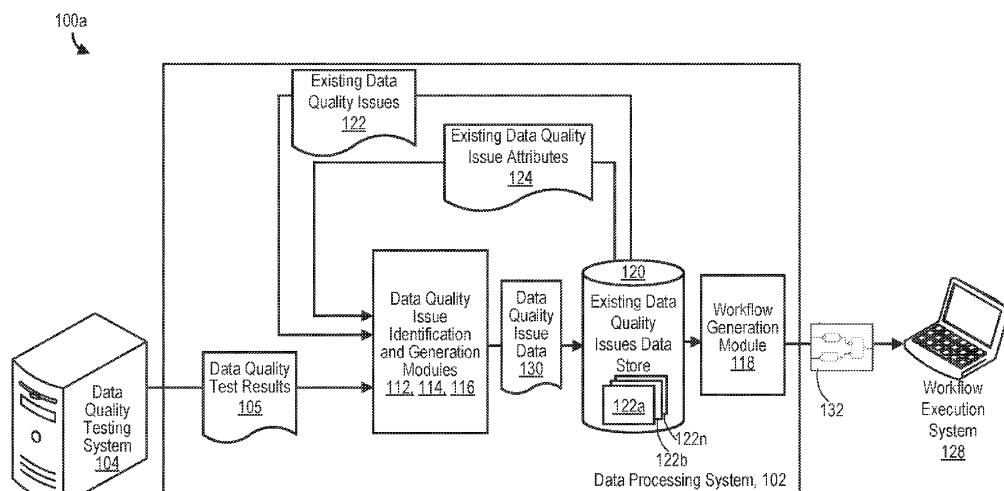


FIG. 1A

(57) **Abstract:** Systems and methods are for executing, by a data processing system, a workflow to process results data indicating an output of a data quality test on data records by generating, responsive to receiving the results data and metadata describing the results data, a data quality issue associated with a state and one or more processing steps of the workflow to resolve a data quality error associated with the data quality test. Operations include generating a workflow for processing results data based a state specified by a data quality issue. Generating the workflow includes: assigning, based on the results data and the state of the data quality issue, an entity responsible for resolving the data quality error; determining, based on the metadata, one or more actions for satisfying the data quality condition specified in the data quality test; and updating the state associated with the data quality issue.

[Continued on next page]

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).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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