

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

(19) World Intellectual Property  
Organization  
International Bureau

(43) International Publication Date  
10 September 2021 (10.09.2021)



(10) International Publication Number  
**WO 2021/178665 A1**

- (51) **International Patent Classification:**  
*G06F 16/25* (2019.01) *G06F 8/34* (2018.01)
- (21) **International Application Number:**  
PCT/US2021/020871
- (22) **International Filing Date:**  
04 March 2021 (04.03.2021)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
62/986,374 06 March 2020 (06.03.2020) US  
17/025,751 18 September 2020 (18.09.2020) US
- (71) **Applicant:** AB INITIO TECHNOLOGY LLC [US/US];  
201 Spring Street, Lexington, Massachusetts 02421 (US).
- (72) **Inventors:** EGENOLF, Jonah; 15 Arbor Lane, Winchester, Massachusetts 01890 (US). ISMAN, Marshall A.; 11 Valley Spring Road, Newton, Massachusetts 02458 (US). SCHECHTER, Ian; 94 Brook Road, Sharon, Massachusetts 02067 (US).
- (74) **Agent:** MCDONOUGH, Christina V.; 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, 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,

(54) **Title:** GENERATION OF OPTIMIZED LOGIC FROM A SCHEMA

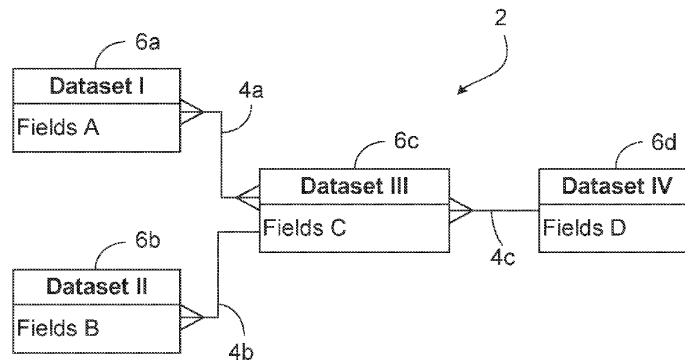


FIG. 1

(57) **Abstract:** A method includes accessing a schema that specifies relationships among datasets, computations on the datasets, or transformations of the datasets, selecting a dataset from among the datasets, and identifying, from the schema, other datasets that are related to the selected dataset. Attributes of the datasets are identified, and logical data representing the identified attributes and relationships among the attributes is generated. The logical data is provided to a development environment, which provides access to portions of the logical data representing the identified attributes. A specification that specifies at least one of the identified attributes in performing an operation is received from the development environment. Based on the specification and the relationships among the identified attributes represented by the logical data, a computer program is generated to perform the operation by accessing, from storage, at least one dataset having the at least one of the attributes specified in the specification.

[Continued on next page]

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))

---