



(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/US2024/032887

(22) International Filing Date:
06 June 2024 (06.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/471,420 06 June 2023 (06.06.2023) US

(71) Applicant: **CARNEGIE MELLON UNIVERSITY**
[US/US]; 5000 Forbes Avenue, Pittsburgh, Pennsylvania
15213 (US).

(72) Inventors: **JAIN, Akshath**; 5000 Forbes Avenue, Pitts-
burgh, Pennsylvania 15213 (US). **SADEH, Norman**; 5000
Forbes Avenue, Pittsburgh, Pennsylvania 15213 (US).

(74) Agent: **KNEDEISEN, Mark G.** et al.; K&L Gates LLP,
210 Sixth Avenue, Pittsburgh, Pennsylvania 15222-2613
(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, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, 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,

(54) Title: AUTOMATICALLY DETECTING DISCREPANCIES BETWEEN PRIVACY POLICIES AND PRIVACY LABELS

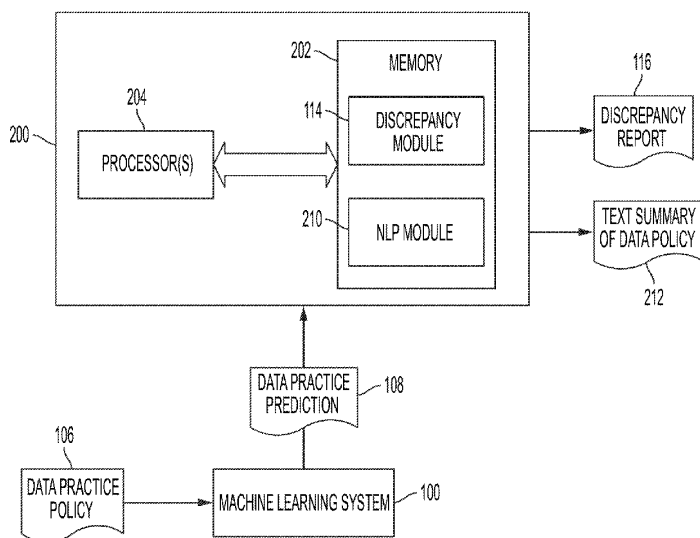


FIGURE 6

(57) Abstract: Computer systems and computer-implemented methods identify potential discrepancies between data practice disclosures made in a data practice policy and a data practice label for a data system. The data system collects or captures data about or for a user of the data system or any other person. The method can also comprise the step of comparing, by the computer system, the prediction for the given data practice with an entry in the data practice label for the data system to identify a possible discrepancy between the prediction and the data practice label. The classifier can be trained via machine learning to make the predictions. The classifier could comprise a classifier for each data practice covered by the data label or a foundation model, such as large language model, that is adapted to make the predictions about a data system's data practice label.

[Continued on next page]

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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*