

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
05 June 2025 (05.06.2025)



(10) International Publication Number
WO 2025/113585 A1

(51) International Patent Classification:
G06N 5/04 (2023.01)

(21) International Application Number:
PCT/CN2024/135384

(22) International Filing Date:
28 November 2024 (28.11.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
202311624147.3 30 November 2023 (30.11.2023) CN
18/956,845 22 November 2024 (22.11.2024) US

(71) Applicant: **HANGZHOU ALIBABA INTERNATIONAL INTERNET INDUSTRY CO., LTD.** [CN/CN]; Room 303, 3rd Floor, Building 5, 699 Wangshang Road, Changhe Sub-District, Binjiang District, Hangzhou, Zhejiang 310052 (CN).

(72) Inventors: **LIU, Sen**; 699 Wang Shang Road Binjiang District, Hangzhou, Zhejiang 310052 (CN). **JIN, Linbo**; 699 Wang Shang Road Binjiang District, Hangzhou, Zhejiang 310052 (CN).

(74) Agent: **BEIJING TONGJUN LAW FIRM**; Room 808, 7th Floor Bldg. 1, No. 32 Xizhimen North Street, Haidian District, Beijing 100082 (CN).

(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, 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:

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

(54) Title: METHOD AND SYSTEM FOR DIALOGUE DATA GENERATION AND PROCESSING

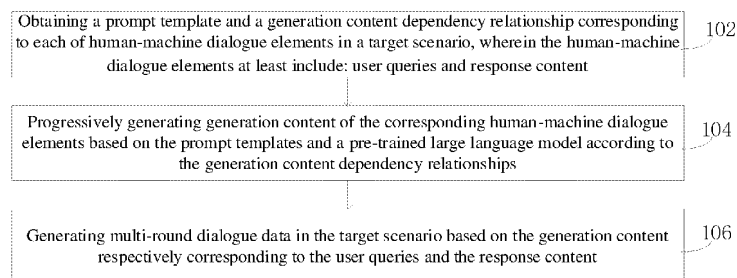


FIG. 1

(57) Abstract: The application provides a method for generating dialogue data, a method for training a model, and a method for processing dialogues. The dialogue data generation method includes: obtaining a prompt template and a generation content dependency relationship corresponding to each of human-machine dialogue elements in a target scenario, wherein the human-machine dialogue elements at least include: user queries and response content; progressively generating generation content of the corresponding human-machine dialogue elements based on the prompt templates and a pre-trained large language model according to the generation content dependency relationships; generating multi-round dialogue data in the target scenario based on the generation content respectively corresponding to the user queries and the response content. This method enables the fully automatic generation of multi-round dialogue data in a target domain, eliminating the need for manual annotation, reducing the cost of acquiring multi-round dialogue data, and improving the efficiency of dialogue data acquisition.

WO 2025/113585 A1