

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2025/129400 A1

(43) 国际公布日
2025 年 6 月 26 日 (26.06.2025)

(51) 国际专利分类号:
G06F 16/9032 (2019.01)

(21) 国际申请号: PCT/CN2023/139569

(22) 国际申请日: 2023 年 12 月 18 日 (18.12.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 新华三技术有限公司 (NEW H3C TECHNOLOGIES CO., LTD.) [CN/CN]; 中国浙江省杭州市滨江区长河路466号 310052 (CN)。

(72) 发明人: 王明辉 (WANG, Minghui); 中国北京市朝阳区广顺南大街8号院1号楼利星行中心A座640室 100102 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街9号嘉华大厦B座409 100085 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): 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) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (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)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(54) **Title:** DATA PROCESSING METHOD AND APPARATUS BASED ON LARGE LANGUAGE MODEL, AND DEVICE

(54) 发明名称: 一种基于大语言模型的数据处理方法、装置及设备

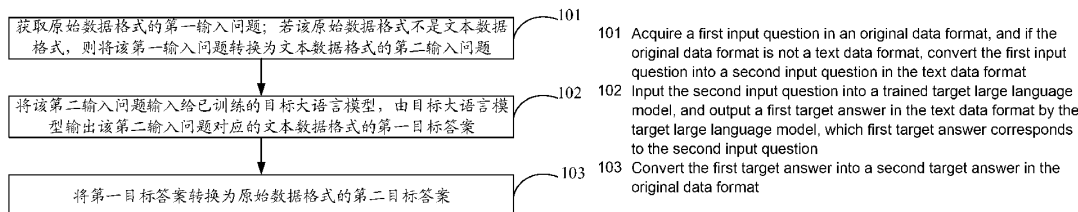


图 1

(57) **Abstract:** Provided in the present application are a data processing method and apparatus based on a large language model, and a device. The method comprises: acquiring a first input question in an original data format, and if the original data format is not a text data format, converting the first input question into a second input question in the text data format; inputting the second input question into a trained target large language model, and outputting a first target answer in the text data format by the target large language model, which first target answer corresponds to the second input question; and converting the first target answer into a second target answer in the original data format. By means of the technical solution of the present application, computing resources of a server can be saved on, time overheads are saved on, and a target answer corresponding to an input question can be quickly determined.

(57) **摘要:** 本申请提供一种基于大语言模型的数据处理方法、装置及设备, 该方法包括: 获取原始数据格式的第一输入问题; 若所述原始数据格式不是文本数据格式, 则将所述第一输入问题转换为文本数据格式的第二输入问题; 将所述第二输入问题输入给已训练的目标大语言模型, 由所述目标大语言模型输出所述第二输入问题对应的文本数据格式的第一目标答案; 将所述第一目标答案转换为所述原始数据格式的第二目标答案。通过本申请的技术方案, 能够节省服务器的计算资源, 并节省时间开销, 能够快速确定出输入问题对应的目标答案。

WO 2025/129400 A1