

(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 2024/077906 A1

(43) 国际公布日
2024 年 4 月 18 日 (18.04.2024)

- (51) 国际专利分类号:
G10L 15/26 (2006.01) *G10L 15/18* (2013.01)
G10L 15/22 (2006.01) *G10L 15/06* (2013.01)
- (21) 国际申请号: PCT/CN2023/087793
- (22) 国际申请日: 2023 年 4 月 12 日 (12.04.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202211231004.1 2022 年 10 月 9 日 (09.10.2022) CN
- (71) 申请人: 京 东 科 技 信 息 有 限 公 司 (JINGDONG TECHNOLOGY INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市大

- 兴区北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。
- (72) 发明人: 冯明超 (FENG, Mingchao); 中国北京市大兴区北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。陈蒙 (CHEN, Meng); 中国北京市大兴区北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。覃杰 (QIN, Jie); 中国北京市大兴区北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。
- (74) 代理人: 中科专利商标代理有限责任公司 (CHINA SCIENCE PATENT & TRADEMARK AGENT LTD.); 中国北京市海淀区西三环北路 87 号 4-312 室, Beijing 100089 (CN)。

(54) Title: SPEECH TEXT GENERATION METHOD AND APPARATUS, AND TRAINING METHOD AND APPARATUS FOR SPEECH TEXT GENERATION MODEL

(54) 发明名称: 语音文本生成方法、语音文本生成模型的训练方法、装置

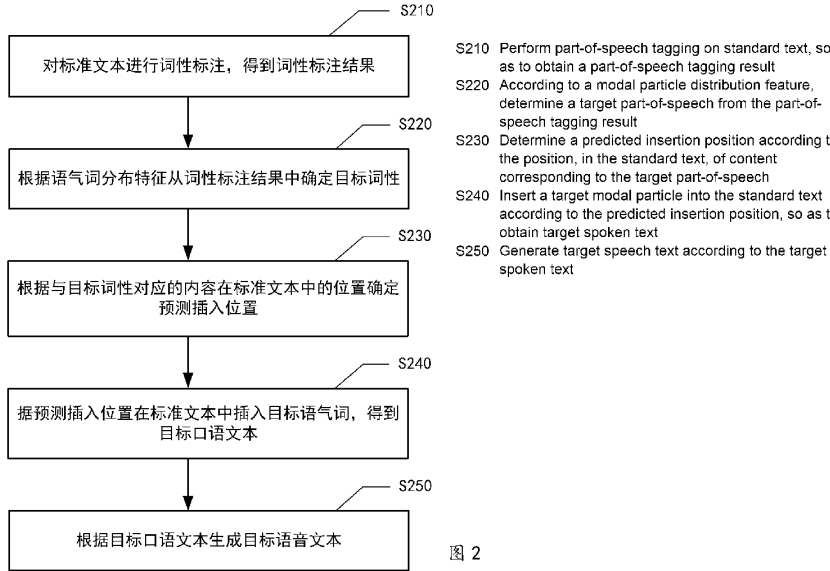


图 2

(57) Abstract: Provided in the present disclosure are a speech text generation method, which can be applied to the technical field of artificial intelligence and the field of intelligent customer service. The speech text generation method comprises: performing part-of-speech tagging on standard text, so as to obtain a part-of-speech tagging result; according to a modal particle distribution feature, determining a target part-of-speech from the part-of-speech tagging result; determining a predicted insertion position according to the position, in

[见续页]

(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))。

the standard text, of content corresponding to the target part-of-speech; inserting a target modal particle into the standard text according to the predicted insertion position, so as to obtain target spoken text; and generating target speech text according to the target spoken text. Further provided in the present disclosure are a training method for a speech text generation model, and a speech text generation apparatus, a training apparatus for a speech text generation model, and a device, a medium and a program product.

(57) 摘要: 本公开提供了一种语音文本生成方法, 可以应用于人工智能技术领域和智能客服领域。该语音文本生成方法包括: 对标准文本进行词性标注, 得到词性标注结果; 根据语气词分布特征从词性标注结果中确定目标词性; 根据与目标词性对应的内容在标准文本中的位置确定预测插入位置; 根据预测插入位置在标准文本中插入目标语气词, 得到目标口语文本; 以及根据目标口语文本生成目标语音文本。本公开还提供了语音文本生成模型的训练方法、语音文本生成装置、语音文本生成模型的训练装置、设备、介质及程序产品。