



(12) 发明专利申请

(10) 申请公布号 CN 117573827 A

(43) 申请公布日 2024. 02. 20

(21) 申请号 202311554214.9

(22) 申请日 2023.11.21

(71) 申请人 北京富算科技有限公司

地址 100070 北京市丰台区南四环西路188号十六区18号楼1至15层101内7层701-8

(72) 发明人 尤志强 王兆凯 赵华宇 卞阳 张伟奇

(74) 专利代理机构 北京慧加伦知识产权代理有限公司 16035
专利代理师 赵敏岑

(51) Int. Cl.

G06F 16/332 (2019.01)

G06F 16/335 (2019.01)

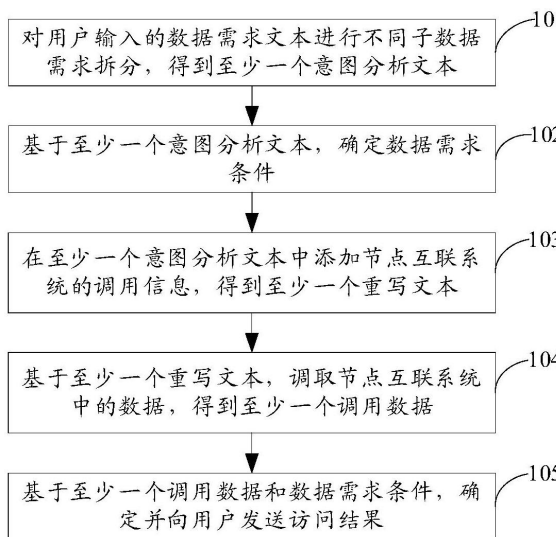
权利要求书2页 说明书12页 附图2页

(54) 发明名称

数据访问方法、装置、计算机设备和介质

(57) 摘要

本公开提供一种数据访问方法、装置、计算机设备和介质,数据访问方法的一具体实施方式包括:对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;基于至少一个意图分析文本,确定数据需求条件;在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据;基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。从而,可以有效地帮助用户寻找自身所需的数据源,提升了用户体验。



1. 一种数据访问方法,其特征在于,包括:
 - 对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;
 - 基于所述至少一个意图分析文本,确定数据需求条件;
 - 在所述至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;
 - 基于所述至少一个重写文本,调取所述节点互联系统中的数据,得到至少一个调用数据;
 - 基于所述至少一个调用数据和所述数据需求条件,确定并向所述用户发送访问结果。
2. 根据权利要求1所述的方法,其特征在于,所述对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本包括:
 - 获取用户输入的数据需求文本;
 - 对所述数据需求文本进行意图识别,确定所述数据需求文本中的至少一个子数据需求;
 - 基于所述至少一个子数据需求中各个子数据需求对应的不同意图,得到至少一个意图分析文本。
3. 根据权利要求1所述的方法,其特征在于,所述在所述至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本包括:
 - 基于所述至少一个意图分析文本中各个意图分析文本对应的节点互联系统的资源模块,确定各个资源模块的访问信息;
 - 将各个访问信息添加到相应的意图分析文本中,得到至少一个重写文本;
 - 所述基于所述至少一个重写文本,调取所述节点互联系统中的数据,得到至少一个调用数据包括:
 - 针对所述至少一个重写文本中各个重新文本,确定该重新文本中的访问信息;
 - 基于所述访问信息,连接所述节点互联系统,并从所述节点互联系统的资源模块获取与该重写文本中意图分析文本相关的数据,得到与该重写文本对应的一个调用数据。
4. 根据权利要求1所述的方法,其特征在于,所述调用信息包括:所述节点互联系统的不同资源的接口调用信息;所述在所述至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本包括:
 - 基于所述至少一个意图分析文本中各个意图分析文本的关键字,确定与各个意图分析文本对应的接口调用信息;
 - 将各个接口调用信息添加到相应的意图分析文本中,得到至少一个重写文本;
 - 所述基于所述至少一个重写文本,调取所述节点互联系统中的数据,得到至少一个调用数据包括:
 - 针对所述至少一个重写文本中各个重写文本,确定该重写文本中的接口调用信息;
 - 基于所述接口调用信息,调取预先设置的与所述节点互联系统连接的功能接口,得到所述功能接口输出的调用数据。
5. 根据权利要求1-4任意一项所述的方法,其特征在于,所述基于所述至少一个意图分析文本,确定数据需求条件包括:
 - 确定所述至少一个意图分析文本中的各个意图分析文本的资源属性;

检测各个意图分析文本中是否具有相应资源属性的属性值；

响应于检测到所述至少一个意图分析文本中具有相应资源属性的属性值，将该资源属性和该资源属性的属性值作为该意图分析文本的数据需求条件。

6. 根据权利要求1-4任意一项所述的方法，其特征在于，所述基于所述至少一个调用数据和所述数据需求条件，确定并向所述用户发送访问结果包括：

基于所述至少一个调用数据，检测是否有满足所述数据需求条件的调用数据；

响应于检测到具有满足所述数据需求条件的调用数据，将满足所述数据需求条件的调用数据作为访问结果，并向所述用户发送所述访问结果。

7. 根据权利要求6所述的方法，其特征在于，所述数据需求条件包括：多个子数据需求条件，在满足所述子数据需求条件的调用数据有多个时，所述方法还包括：

对满足所述子数据需求条件的多个调用数据进行评分，得到多个调用数据的分值；

对多个调用数据的分值进行降序排序，选取分值为前设定位的调用数据作为满足所述子数据需求条件的调用数据。

8. 一种数据访问装置，其特征在于，包括：

拆分模块，用于对用户输入的数据需求文本进行不同子数据需求拆分，得到至少一个意图分析文本；

确定模块，用于基于所述至少一个意图分析文本，确定数据需求条件；

得到模块，用于在所述至少一个意图分析文本中添加节点互联系统的调用信息，得到至少一个重写文本；

调取模块，用于基于所述至少一个重写文本，调取所述节点互联系统中的数据，得到至少一个调用数据；

发送模块，用于基于所述至少一个调用数据和所述数据需求条件，确定并向所述用户发送访问结果。

9. 一种计算机设备，其特征在于，包括存储器和处理器，存储器中存储有计算机程序，处理器执行计算机程序时实现如权利要求1~7中任一所述的数据访问方法。

10. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，计算机程序被处理器执行时实现如权利要求1~7中任一所述的数据访问方法。

数据访问方法、装置、计算机设备和介质

技术领域

[0001] 本公开的实施例涉及数据处理技术领域,具体地,涉及适用于一种数据访问方法、装置、计算机设备和介质。

背景技术

[0002] 随着互联网的发展,各类政务主体、行业主体、公司主体、机构主体可通过互联网关联起来。每一个主体被看作一个节点,一个节点常常需要与其他节点进行数据交易或者挑选其他节点一起执行任务。

[0003] 但是,对于上述互联网中各个节点的数据资源的有效查询,现有技术中没有较好的实现方法。

发明内容

[0004] 本文中描述的实施例提供了一种数据访问方法、数据访问装置、计算机设备和计算机可读存储介质,克服了上述问题。

[0005] 第一方面,根据本公开的内容,提供了一种数据访问方法,包括:对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;基于至少一个意图分析文本,确定数据需求条件;在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据;基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。

[0006] 第二方面,根据本公开的内容,提供了一种数据访问装置,包括:拆分模块,用于对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;确定模块,用于基于至少一个意图分析文本,确定数据需求条件;得到模块,用于在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;调取模块,用于基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据;发送模块,用于基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。

[0007] 第三方面,提供了一种计算机设备,包括存储器和处理器,存储器中存储有计算机程序,处理器执行计算机程序时实现如以上任意一个实施例中数据访问方法的步骤。

[0008] 第四方面,提供了一种计算机可读存储介质,计算机可读存储介质上存储有计算机程序,计算机程序被处理器执行时实现如以上任意一个实施例中数据访问方法的步骤。

[0009] 本公开实施例提供的数据访问方法和装置,首先,对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;然后,基于至少一个意图分析文本,确定数据需求条件;再次,在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;从次,基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据;最后,基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。如此,将用户的数据需求文本拆分为至少一个满足子数据需求的意图分析文本,基于意图分析文本,确定重写文本和数据需求条件,为数据访问提供了可靠的条件支持;通过重写

文本调取调用数据,可以与节点互联系统的数据相关功能执行强相关;将满足数据需求条件的调用数据作为反馈用户的数据需求文本的访问数据,提升了用户数据访问体验。

[0010] 上述说明仅是本公开实施例技术方案的概述,为了能够更清楚了解本公开实施例的技术手段,而可依照说明书的内容予以实施,并且为了让本公开实施例的上述和其它目的、特征和优点能够更明显易懂,以下特举本公开的具体实施方式。

附图说明

[0011] 为了更清楚地说明本公开的实施例的技术方案,下面将对实施例的附图进行简要说明,应当知道,以下描述的附图仅仅涉及本公开的一些实施例,而非对本公开的限制,其中:

[0012] 图1是本公开提供的一种数据访问方法的流程示意图。

[0013] 图2是本公开提供的一种系统架构图。

[0014] 图3是本公开提供的一种数据访问装置的结构示意图。

[0015] 图4是本公开提供的一种计算机设备的结构示意图。

[0016] 需要注意的是,附图中的元素是示意性的,没有按比例绘制。

具体实施方式

[0017] 为了使本公开的实施例的目的、技术方案和优点更加清楚,下面将结合附图,对本公开的实施例的技术方案进行清楚、完整的描述。显然,所描述的实施例是本公开的一部分实施例,而不是全部的实施例。基于所描述的本公开的实施例,本领域技术人员在无需创造性劳动的前提下所获得的所有其它实施例,也都属于本公开保护的范围。

[0018] 除非另外定义,否则在此使用的所有术语(包括技术和科学术语)具有与本公开主题所属领域的技术人员所通常理解的相同含义。进一步将理解的是,诸如在通常使用的词典中定义的那些的术语应解释为具有与说明书上下文和相关技术中它们的含义一致的含义,并且将不以理想化或过于正式的形式来解释,除非在此另外明确定义。如在此所使用的,将两个或更多部分“连接”或“耦接”到一起的陈述应指这些部分直接结合到一起或通过一个或多个中间部件结合。

[0019] 在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本公开的至少一个实施例中。在说明书中的各个位置出现该短语“实施例”并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

[0020] 本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:存在A,同时存在A和B,存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。诸如“第一”和“第二”的术语仅用于将一个部件(或部件的一部分)与另一个部件(或部件的另一部分)区分开。

[0021] 在本公开的描述中,除非另有说明,“多个”的含义是指两个以上(包括两个),同理,“多组”指的是两组以上(包括两组)。

[0022] 为了使本技术领域的人员更好地理解本公开方案,下面将结合附图,对本公开实施例中的技术方案进行清楚、完整地描述。

[0023] 随着ChatGPT(Chat Generative Pre-trained Transformer,生成型预训练变换模型)的出现,基于大语言模型的能力,可以支持对于复杂语义的理解,且基于思维链,可以实现逐步的意图准确定位,进而支持对节点互联系统中用户所需数据资源的准确高效发现及定位。

[0024] 对于大语言模型可以准确识别出每一条细致的要求,非常准确,但是其只能做分析,无法执行动作。

[0025] 针对上述技术问题,本实施例提供的数据访问方法,可以对节点互联系统中的数据进行分析以及执行数据访问动作,帮助节点互联系统的用户能够快速寻找自身所需的数据源,降低数据发现的难度,提升节点互联系统的用户体验。图1是本公开实施例提供的一种数据访问方法的流程示意图。如图1所示,数据访问方法的具体步骤包括:

[0026] 步骤S110,对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本。

[0027] 本实施例中,数据需求文本是由用户在前端输入的对节点互联系统中数据进行访问的需求文本,通过数据需求文本,可以具体确定用户意图访问的多方面的信息主体以及每种信息主体需要满足的要求,而各种信息以及各种信息主体需求满足的要求共同组成子数据需求。

[0028] 本实施例中,节点互联系统可以是传统的数联网,也可以是面向数据保护的数联网;其中,传统的数联网仅仅是“数数相连的互联网”,仅涉及数据在数联网中透明传输,应用端可以直接获取数据,产生各种基于数据的业务应用。面向数据保护的数联网是一种用隐私计算构建数据要素安全互通新范式,面向数据保护的数联网更多关注数据安全和数据处理的隐私保护,面向数据保护的数联网具有一下三个关键特征:(1) 权益分离,保障双方权益,针对数据可复制性,实现与价值的部分分离;(2) 数据安全,安全性可验证,从根本上解决数据泄漏风险问题,同时大幅降低合规成本;(3) 数据流动,经济模型可持续,数据使用权转让合理有序的同时,实现数据可计算价值的充分挖掘与跨主体使用。

[0029] 可选地,由于节点互联系统是多种资源的网络,对应于数据需求文本中不同的资源以及各个资源需要满足的要求,也可以作为一种子数据需求。

[0030] 例如,一种数据需求文本为:“帮我找一个移动运营商,数据条数超过100万条,包含用户的夜晚通信次数分析的特征数据”;在该数据需求文本中,具有“移动运营商”、“数据条”、“通信次数”三种信息主体,对于该三种信息主体中各个信息主体均有各自需要满足的数据子要求,为此,与“移动运营商”对应的子数据需求为:移动运营商:需要找到一个移动运营商,这将是数据的来源。与“数据条”对应的子数据需求为:数据条数超过100万条:需要获取的数据量必须超过100万条。与“通信次数”对应的子数据需求为:包含用户的夜晚通信次数分析的特征数据:所需数据必须包含关于用户夜晚通信次数的特征数据,以便进行分析。

[0031] 本实施例中,上述对用户输入的数据需求文本进行不同子数据需求拆分包括:对数据需求文本中的信息主体进行不同子数据需求分析,确定各个信息主体以及各个信息主体的子数据需求;基于各个信息主体的子数据需求,通过意图分析,得到对应各个子数据需求的意图分析文本。例如,上述“需要找到一个移动运营商,这将是数据的来源”进行意图分析之后,得到的意图分析文本为:在节点互联系统中找到一个移动运行商的数据源提供方。

[0032] 可选地,上述步骤S110还可以包括:确定数据需求文本中各个信息主体;对各个信息主体的子数据需求进行分析,确定各个信息主体的意图分析文本。

[0033] 步骤S120,基于至少一个意图分析文本,确定数据需求条件。

[0034] 本实施例中,数据需求条件是数据需求文本中各个节点互联系统的数据资源需求满足的条件。针对多个子数据需求,每个子数据需求的数据需求条件可以不同,为此,至少一个意图分析文本中每个意图分析文本均可以具有一个子数据需求条件,至少一个意图分析文本对应的多个子数据需求条件组成数据需求条件,例如,针对上述“在节点互联系统中找到一个移动运行商的数据源提供方”,其对应的数据需求条件为:数据源提供方为移动运营商。

[0035] 步骤S130,在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本。

[0036] 本实施例中,调用信息是调取节点互联系统中数据的信息,通过调用信息可以连接节点互联系统,并且从节点互联系统中获取到与意图分析文本相关的数据资源。

[0037] 本实施例中,至少一个重写文本是添加了调用信息的至少一个意图分析文本,至少一个重写文本中各个重写文本中可以是均添加了调用信息的拆分文本,至少一个重写文本也可以是有部分重写文本添加了调用信息的拆分文本,至少一个重写文本中其它重写也可以是拆分文本。

[0038] 上述步骤S130包括:基于至少一个意图分析文本,确定至少一个节点互联系统数据资源;基于预先设置的节点互联系统中数据资源的调用信息表格以及至少一个节点互联系统数据资源,得到调用信息;将调用信息添加到至少一个意图分析文本中,得到至少一个重写文本。

[0039] 步骤S140,基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据。

[0040] 本实施例中,基于至少一个重写文本中各个重写文本中记载的调用信息,可以直接调取节点互联系统中的数据,达到访问节点互联系统中数据资源的目的。调用数据是从节点互联系统中获取到的与至少一个重写文本相关的数据资源的数据。

[0041] 具体地,上述步骤S140包括:将至少一个重写文本中各个重写文本输入到动作执行程序生成智能体,由动作执行程序生成智能体生成可执行代码块,将可执行代码块输出到动作程序执行模块进行程序执行,在程序执行过程,动作程序执行模块访问节点互联系统,得到与各个重写文本对应的调用数据,从而达到访问节点互联系统中与各个重写文本对应的数据资源的目的。其中,上述动作执行程序生成智能体与动作程序执行模块连接,动作程序执行模块与节点互联系统连接,动作执行程序生成智能体与动作程序执行模块均可以通过大语言模型实现。

[0042] 步骤S150,基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。

[0043] 本实施例中,数据需求条件包括多个子数据需求条件,只有至少一个调用数据均满足数据需求条件中所有子数据需求条件时,该满足数据需求条件的调用数据才能够作为访问结果发送给用户。如上述数据需求文本为:“帮我找一个移动运营商,数据条数超过100万条,包含用户的夜晚通信次数分析的特征数据”的数据需求条件包括:a.数据源提供方为移动运营商(子数据需求条件一);b.数据规模在100万条以上(子数据需求条件二);c.数据

字段需包含用户夜晚通信次数(子数据需求条件三)。

[0044] 随着大语言模型训练对话模型ChatGPT的发展,将其作为智能体,能够实现对输入命令的自动思考,并实现任务编排与调度工具的逻辑,进而完成用户的需求,其中,智能体可以针对用户的需求,自动进行方案的设计与优化,并针对优化的方案进行提供工具的编排与实际调用,返回给用户需要的访问结果。

[0045] 如图2所示,本公开提供的数据访问方法运行于其上的执行主体可以采用多个智能体,为此,可以采用大语言模型定义多种智能体的角色,包括:用户指令智能体、意图分析智能体、动作执行程序生成智能体、动作程序执行智能体、状态分析检测智能体。还可以采用大语言定义多种模块,包括:重写模块、模型评分模块、动作程序执行模块。

[0046] 本实施例中,“状态分析检测智能体”与“意图分析智能体”会进行多次交互,循环检测细分要求,逐项判断是否子数据需求条件是否满足。最终会反馈给“意图分析智能体”满足数据需求条件中所有子数据需求条件的结果列表信息。

[0047] 本公开实施例提供的数据访问方法,首先,对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本;然后,基于至少一个意图分析文本,确定数据需求条件;再次,在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本;从次,基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据;最后,基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。如此,将用户的数据需求文本拆分为至少一个满足子数据需求的意图分析文本,基于意图分析文本,确定重写文本和数据需求条件,为数据访问提供了可靠的条件支持;通过重写文本调取调用数据,可以与节点互联系统的数据相关功能执行强相关;将满足数据需求条件的调用数据作为反馈用户的数据需求文本的访问数据,提升了用户数据访问体验。

[0048] 在本公开的一些可选实现方式中,上述对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本包括:获取用户输入的数据需求文本;对数据需求文本进行意图识别,确定数据需求文本中的至少一个子数据需求;基于至少一个子数据需求中各个子数据需求对应的不同意图,得到至少一个意图分析文本。

[0049] 本可选实现方式中,上述获取用户输入的数据需求文本可以包括:获取用户输入的数据需求语音,对数据需求语音进行语音文本转换得到数据需求文本;或者,上述获取用户输入的数据需求文本还可以包括:获取用户输入的数据需求图像,识别数据需求图像中的图像内容,得到数据需求文本。

[0050] 本可选实现方式中,对数据需求文本进行意图识别可以包括:将数据需求文本输入意图识别模型,得到意图识别模型输出的至少一个子数据需求,各个子数据需求与一个信息主体对应,用于反映对信息主体(例如,节点互联系统中数据资源)的操作需求。需要说明的是,子数据需求可以以多种形式(文本、图像、语音)进行表现。意图识别模型是成熟的模型,此处不再赘述。

[0051] 本可选实现方式中,意图可以分为多种种类,例如,数据提取意图,数据筛选意图,数据优化意图,在得到数据需求文本中至少一个子数据需求之后,对各个子数据需求进行语义分析,确定各个子数据需求属于上述多种意图中的哪一种或多种,得到数据需求文本中所有子数据需求的意图,从而基于意图拆分各个子数据需求,得到包括一个或多个子数据需求的意图分析文本。

[0052] 本可选实现方式提供的得到至少一个意图分析文本的方法,首先对数据需求文本进行意图识别,确定至少一个子数据需求,基于至少一个子数据需求对应的意图,划分意图分析文本,可以使得到的意图分析文本具有明确的意图类型,提高了意图分析文本得到的可靠性和准确性。

[0053] 在本公开的一些实施例中,上述在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本包括:基于至少一个意图分析文本中各个意图分析文本对应的节点互联系统的资源模块,确定各个资源模块的访问信息;将各个访问信息添加到相应的意图分析文本中,得到至少一个重写文本。

[0054] 在本实施例中,上述基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据包括:针对至少一个重写文本中各个重新文本,确定该重新文本中的访问信息;基于访问信息,连接节点互联系统,并从节点互联系统的资源模块获取与该重写文本中意图分析文本相关的数据,得到与该重写文本对应的一个调用数据。

[0055] 在本实施例中,基于节点互联系统中各种数据资源的功能不同,可以将节点互联系统中各个数据资源划分到不同的资源模块中,每个资源模块具有其相应的功能,例如,对于节点互联系统中需要参与到数据交易的数据资源,将该些数据资源划分到数据交易资源模块;对于节点互联系统中需要参与到数据加工与交付的数据资源,将该些数据资源划分到数据加工与交付资源模块。为此,节点互联系统中的资源模块可以包括:数据交易资源模块、数据加工与交付资源模块、数据生产资源模块、数据算力资源模块。

[0056] 在本实施例中,访问信息是访问各个资源模块的信息,通过访问信息可以得到各个资源模块的数据资源,为了实现对各个资源模块的访问,可以与节点互联系统进行通信,通过通信得到各个资源模块的访问信息。进一步地,访问信息可以包括:节点互联系统的地址、访问端口、资源模块的地址、资源模块的访问端口。可选地,为了实现对节点互联系统访问的可靠性,访问信息还可以包括:节点互联系统的访问密钥以及各个资源模块的访问密钥,通过节点互联系统的访问密钥和资源模块的访问密钥可以对加密之后的模块进行解密操作。

[0057] 在本实施例中,将访问信息添加到意图分析文本中,可以使各个拆分文本具有相应的访问节点互联系统的信息,从而通过至少一个重写文本可以随时访问相应的资源模块。

[0058] 本公开实施例提供的数据访问方法,在至少一个意图分析文本的各个意图分析文本中添加访问信息,得到重写文本,可以使执行主体通过重写文本可以直接访问节点互联系统中的资源模块,实现了对节点互联系统数据资源的落地访问,提高了资源模块访问的便利性。

[0059] 在本公开的另一些实施例中,调用信息包括:节点互联系统不同资源的接口调用信息;在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本包括:基于至少一个意图分析文本中各个意图分析文本的关键词,确定与各个意图分析文本对应的接口调用信息;将各个接口调用信息添加到相应的意图分析文本中,得到至少一个重写文本。

[0060] 在本实施例中,上述基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据包括:针对至少一个重写文本中各个重写文本,确定该重写文本中的接口

调用信息;基于接口调用信息,调取预先设置的与节点互联系统连接的功能接口,得到功能接口输出的调用数据。

[0061] 在本实施例中,基于节点互联系统,需要预先定义和开发多种功能接口,比如数据资源搜索接口、数据资源描述信息获取接口、数据处理脚本执行接口、联邦数据质量评估功能调用接口、模型评分推理接口。通过功能接口可以直接访问节点互联系统的数据资源。需要说明的是,节点互联系统的各个数字资源的访问以及各个功能的功能接口的定义与开发采用本领域的常规方式,此处不再赘述。

[0062] 本实施例中,接口调用信息是调用上述预先定义的功能接口的信息,通过接口调用信息可以直接调用接口,从而达到访问节点互联系统的数据资源的目的。本实施例中,当数据访问方法运行于其上的执行主体是大语言模型时,大语言模型会参与决定调用哪些接口,并且生成调用接口的接口调用信息。动作程序执行模块会执行这个决定的动作。

[0063] 在本公开提供的节点互联系统访问方法的整个流程的不同阶段,会用到不同的功能接口。可预期的是,上述预先定义或开发的所有的功能接口至少被使用一次。

[0064] 在本实施例中,基于各个意图分析文本的关键字,确定各个意图分析文本对应的接口调用信息包括:将各个意图分析文本的关键字与预先定义和开发的功能接口的接口调用信息进行相似度比较;响应于关键字与接口调用信息的相似度大于相似度阈值,确定关键字与该接口调用信息相对应,在意图分析文本的各个意图分析文本中添加接口调用信息,得到重写文本;需要说明的是,每个重写文本是意图分析文本与接口调用信息之和,通过重写文本可以得到相应的接口调用信息,从而调用相应的功能接口。

[0065] 下面以图2为具体示例,对本实施例的数据访问方法进行具体说明:1) 用户通过输入数据需求文本提出数据数据源需求,在数据需求文本中给出文本信息的描述,作为提示语输入到“用户指令智能体”。2) “用户指令智能体”接收用户的数据源需求后,会进行意图的分析,将数据需求文本复杂的数据需求拆分为多个小的子数据要求,针对各个拆分后的子数据要求生成意图分析文本,意图分析文本会作为“提示语需被重写的文本”输入到“重写模块”,该重写模块,会通过获取功能接口等系统信息,相应加入节点互联系统的“功能接口的信息”。3) 重写模块,在意图分析文本中加入功能接口后,得到重写文本。重写文本(提示语中融合系统提供的功能接口调用地址),进入到“意图分析智能体”。“意图分析智能体”会将子数字要求拆分成独立的多个子数字要求条件,形成包括多个子数据要求条件的数据需求条件。然后将每一条意图分析文本(细分意图的提示语文本)输入到“动作执行程序生成智能体”,由“动作执行程序生成智能体”生成可执行的代码块(如python、java等)。代码块输入到“动作程序执行模块”进行执行,也就是执行节点互联系统的功能接口获得满足子数据要求条件的调用数据结果,返回结果到“意图分析智能体”,“意图分析智能体”调用“状态分析检测智能体”进行子数据要求条件是否满足的检测。

[0066] “状态分析检测智能体”会记录需要满足的所有用户数据需求条件。返回的满足要求的候选数据集列表会进一步做模型评分,完成质量评估,对数据进行按照分值从高到低排序,得到访问结果。

[0067] 本公开的实施例提供的数据访问方法,基于意图分析文本的关键字,确定意图分析文本对应的接口调用信息,将接口调用信息添加到相应的意图分析文本中,得到至少一个重写文本;在调取节点互联系统中的数据时,基于重写文本中的接口调用信息,调取与节

点互联系统连接的功能接口,可以使执行主体通过重写文本可以直接访问节点互联系统的功能接口,提高了节点互联系统的数据资源访问的便利性。

[0068] 在本公开的一些可选实现方式中,上述基于至少一个意图分析文本,确定数据需求条件包括:确定至少一个意图分析文本中的各个意图分析文本的资源属性;检测各个意图分析文本中是否具有相应资源属性的属性值;响应于检测到至少一个意图分析文本中具有相应资源属性的属性值,将该资源属性和该资源属性的属性值作为该意图分析文本的数据需求条件。

[0069] 本可选实现方式中,资源属性是表征节点互联系统中数据资源的属性,由于数据资源的种类不同,资源属性也不同。意图分析文本中具有与节点互联系统的数据资源对应的信息主体,而信息主体具有相应的资源属性,例如,意图分析文本的信息主体为移动运营商,而移动运营商在节点互联系统的属性是节点互联系统的数据源的来源。

[0070] 本可选实现方式中,意图分析文本中的信息主体具有相应的资源属性,对于意图分析文本中其它文本可以用于反映信息主体的资源属性的属性值,也可以不反映信息主体的资源属性的属性值;针对信息主体的资源属性的具体特征,资源属性的属性值可以具体数值,也可以是枚举值,如,在资源属性为数据规模时,资源属性的属性值可以100万条;在资源属性为数据源的提供方时,资源属性的属性值为移动运行商。

[0071] 需要说明的是,意图分析文本中可以具有反映信息主体的资源属性的属性值,意图分析文本中也可以不具有反映信息主体的资源属性的属性值,当检测到一个意图分析文本中具有相应资源属性的属性值时,可以直接将资源属性和属性值作为数据需求条件中的子数据需求条件,当至少一个意图分析文本中的所有意图分析文本均检测完成之后,将得到的所有的子数据需求条件作为数据需求条件。

[0072] 可选地,当至少一个意图分析文本中的各个意图分析文本中均不具有资源属性的属性值时,可以直接将各个意图分析文本中的资源属性作为数据需求条件的子数据需求条件。意图分析文本意图分析文本

[0073] 本可选实现方式提供的确定数据需求条件的方法,确定意图分析文本的资源属性,在检测到意图分析文本中具有相应资源属性的属性值时,将至少一个意图分析文本中资源属性和资源属性的属性值作为数据需求条件,为数据需求条件的得到提供了一种可靠的实现方式,提高了数据需求条件得到的可靠性。

[0074] 在本实施例的一些可选实现方式中,上述基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果包括:基于至少一个调用数据,检测是否有满足数据需求条件的调用数据;响应于检测到具有满足数据需求条件的调用数据,将满足数据需求条件的调用数据作为访问结果,并向用户发送访问结果。

[0075] 本可选实现方式中,至少一个调用数据中的各个调用数据可能来自节点互联系统的不同数据资源,上述基于至少一个调用数据,检测是否满足数据需求条件的调用数据包括:在得到至少一个调用数据之后,将至少一个调用数据中的各个调用数据与该调用数据相适配的子数据需求条件进行匹配,响应于该调用数据满足该子数据需求条件,确定有调用数据满足子数据需求条件,存储该满足子数据需求条件的调用数据,再将至少一个调用数据与各个子数据需求条件进行比较,若数据需求条件中所有子数据需求条件均被各自的调用数据满足,则确定具有满足数据需求条件的调用数据。

[0076] 可选地,在数据需求条件中有子数据需求条件未被当前的至少一个调用数据满足时,即当前轮的至少一个调用数据无法数据需求条件中的子数据需求条件时,可以修改重写文本,并重新调取节点互联系统中的数据,得到新的至少一个调用数据,再采用新的至少一个调用数据满足数据需求条件,直至数据需求条件满足为止。

[0077] 本可选实现方式提供的确定并向用户发送访问结果的方法,基于至少一个调用数据,检测是否具有满足数据需求条件的调用数据,响应于检测到具有满足数据需求条件的调用数据,将满足数据需求条件的调用数据作为访问结果,发送给用户,为访问结果的得到提供了一种可靠的实现方式。

[0078] 在本公开的一些可选实现方式中,上述数据需求条件包括:多个子数据需求条件,在满足子数据需求条件的调用数据有多个时,方法还包括:对满足子数据需求条件的多个调用数据进行评分,得到多个调用数据的分值;对多个调用数据的分值进行降序排序,选取分值为前设定位的调用数据作为满足子数据需求条件的调用数据。

[0079] 本可选实现方式中,前设定位可以基于设计需求确定,例如,前设定位位前3位。分值是用于评价调用数据是否满足子数据需求的标准,分值越高,则调用数据越满足子数据需求。

[0080] 本可选实现方式中,上述对满足子数据需求条件的多个调用数据进行评分包括:将满足子数据需求条件的多个调用数据输入模型评分模块,得到模型评分模块输出的对应各个调用数据的分值。其中,模型评分模块是采用大语言模型的模块。

[0081] 本可选实现方式提供的确定并向用户发送访问结果的方法,在数据需求条件包括多个子数据条件,并且满足子数据需求条件的调用数据有多个时,通过对满足子数据需求条件的多个调用数据进行评分和排序,选取分值为前设定位的调用数据作为子数据需求条件的调用数据,提高了访问数据的可靠性和准确性。

[0082] 可选地,上述数据需求条件包括:多个子数据需求条件,在满足子数据需求条件的调用数据有多个时,方法还包括:对满足子数据需求条件的多个调用数据进行评分,得到多个调用数据的分值;对多个调用数据的分值进行升序排序,选取分值为后设定位的调用数据作为满足子数据需求条件的调用数据。

[0083] 图3为本公开的实施例提供的一种数据访问装置的结构示意图。数据访问装置可以包括:拆分模块310、确定模块320、得到模块330、调取模块340和发送模块350。

[0084] 其中,上述拆分模块310,用于对用户输入的数据需求文本进行不同子数据需求拆分,得到至少一个意图分析文本。上述确定模块320,用于基于至少一个意图分析文本,确定数据需求条件。上述得到模块330,用于在至少一个意图分析文本中添加节点互联系统的调用信息,得到至少一个重写文本。上述调取模块340,用于基于至少一个重写文本,调取节点互联系统中的数据,得到至少一个调用数据。上述发送模块350,用于基于至少一个调用数据和数据需求条件,确定并向用户发送访问结果。

[0085] 在本实施例中,数据访问装置中:拆分模块310、确定模块320、得到模块330、调取模块340和发送模块350的具体处理及其所带来的技术效果可分别参考图1对应实施例中的步骤101、步骤102、步骤103、步骤104、步骤105的相关说明,在此不再赘述。

[0086] 在本实施例的一些可选的实现方式中,上述拆分模块310具体用于:获取用户输入的数据需求文本;对数据需求文本进行意图识别,确定数据需求文本中的至少一个子数据

需求;基于至少一个子数据需求中各个子数据需求对应的不同意图,得到至少一个意图分析文本。

[0087] 在本实施例的一些可选的实现方式中,上述得到模块330具体用于:基于至少一个意图分析文本中各个意图分析文本对应的节点互联系统的资源模块,确定各个资源模块的访问信息;将各个访问信息添加到相应的意图分析文本中,得到至少一个重写文本。上述调取模块340具体用于:针对至少一个重写文本中各个重新文本,确定该重新文本中的访问信息;基于访问信息,连接节点互联系统,并从节点互联系统的资源模块获取与该重写文本中意图分析文本相关的数据,得到与该重写文本对应的一个调用数据。

[0088] 在本实施例的一些可选的实现方式中,调用信息包括:节点互联系统不同资源的接口调用信息;上述得到模块330具体用于:基于至少一个意图分析文本中各个意图分析文本的关键字,确定与各个意图分析文本对应的接口调用信息;将各个接口调用信息添加到相应的意图分析文本中,得到至少一个重写文本。上述调取模块340具体用于:针对至少一个重写文本中各个重写文本,确定该重写文本中的接口调用信息;基于接口调用信息,调取预先设置的与节点互联系统连接的功能接口,得到功能接口输出的调用数据。

[0089] 在本实施例的一些可选的实现方式中,上述确定模块320具体用于:确定至少一个意图分析文本中的各个意图分析文本的资源属性;检测各个意图分析文本中是否具有相应资源属性的属性值;响应于检测到至少一个意图分析文本中具有相应资源属性的属性值,将该资源属性和该资源属性的属性值作为该意图分析文本的数据需求条件。

[0090] 在本实施例的一些可选的实现方式中,上述发送模块350具体用于:基于至少一个调用数据,检测是否有满足数据需求条件的调用数据;响应于检测到具有满足数据需求条件的调用数据,将满足数据需求条件的调用数据作为访问结果,并向用户发送访问结果。

[0091] 在本实施例的另一些可选的实现方式中,数据需求条件包括:多个子数据需求条件,在满足子数据需求条件的调用数据有多个时,上述发送模块350具体还用于:对满足子数据需求条件的多个调用数据进行评分,得到多个调用数据的分值;对多个调用数据的分值进行降序排序,选取分值为前设定位的调用数据作为满足子数据需求条件的调用数据。

[0092] 本公开提供的数据访问装置,可执行上述方法实施例,其具体实现原理和技术效果,可参见上述方法实施例,本公开此处不再赘述。

[0093] 本公开实施例还提供了一种计算机设备。具体请参阅图4,图4为本实施例计算机设备基本结构框图。

[0094] 计算机设备包括通过系统总线相互通信连接存储器410和处理器420。需要指出的是,图中仅示出了具有组件410-420的计算机设备,但是应理解的是,并不要求实施所有示出的组件,可以替代的实施更多或者更少的组件。其中,本技术领域技术人员可以理解,这里的计算机设备是一种能够按照事先设定或存储的指令,自动进行数值计算和/或信息处理的设备,其硬件包括但不限于微处理器、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程门阵列(Field-Programmable Gate Array,FPGA)、数字处理器(Digital Signal Processor,DSP)、嵌入式设备等。

[0095] 计算机设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。计算机设备可以与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互。

[0096] 存储器410至少包括一种类型的可读存储介质,可读存储介质包括非易失性存储

器(non-volatile memory)或易失性存储器,例如,闪存(flash memory)、硬盘、多媒体卡、卡型存储器(例如,SD或DX存储器等)、随机访问存储器(random access memory,RAM)、只读存储器(read-only memory,ROM)、可擦写可编程只读存储器(erasable programmable read-only memory,EPR0M)、电可擦写可编程只读存储器(electrically erasable programmable read-only memory,EEPROM)、可编程只读存储器(programmable read-only memory,PROM)、磁性存储器、磁盘、光盘等,RAM可以包括静态RAM或动态RAM。在一些实施例中,存储器410可以是计算机设备的内部存储单元,例如,该计算机设备的硬盘或内存。在另一些实施例中,存储器410也可以是计算机设备的外部存储设备,例如该计算机设备上配备的插接式硬盘、智能存储卡(Smart Media Card,SMC)、安全数字(Secure Digital,SD)卡或闪存卡(Flash Card)等。当然,存储器410还可以既包括计算机设备的内部存储单元也包括其外部存储设备。本实施例中,存储器410通常用于存储安装于计算机设备的操作系统和各类应用软件,例如上述方法的程序代码等。此外,存储器410还可以用于暂时地存储已经输出或者将要输出的各类数据。

[0097] 处理器420通常用于执行计算机设备的总体操作。本实施例中,存储器410用于存储程序代码或指令,程序代码包括计算机操作指令,处理器420用于执行存储器410存储的程序代码或指令或者处理数据,例如运行上述方法的程序代码。

[0098] 本文中,总线可以是工业标准体系结构(Industry Standard Architecture,ISA)总线、外设部件互连标准(Peripheral Component Interconnect,PCI)总线或扩展工业标准结构(Extended Industry Standard Architecture,EISA)总线等。该总线系统可以分为地址总线、数据总线、控制总线等。为便于表示,图中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0099] 本公开的另一实施例还提供一种计算机可读介质,计算机可读介质可以是计算机可读信号介质或者计算机可读介质。计算机中的处理器读取存储在计算机可读介质中的计算机可读程序代码,使得处理器能够执行在上述数据访问方法中每个步骤、或各步骤的组合中规定的功能动作;生成实施在框图的每一块、或各块的组合中规定的功能动作的装置。

[0100] 计算机可读介质包含但不限于电子、磁性、光学、电磁、红外的存储器或半导体系统、设备或者装置,或者前述的任意适当组合,存储器用于存储程序代码或指令,程序代码包括计算机操作指令,处理器用于执行存储器存储的上述数据访问方法的程序代码或指令。

[0101] 存储器和处理器的定义,可以参考前述计算机设备实施例的描述,在此不再赘述。

[0102] 在本公开所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0103] 在本公开各个实施例中的各功能单元或模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0104] 集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。基于这样的理解,本公开的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(processor)执行本公开各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory, RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0105] 在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。本公开描述的“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本公开可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了装置若干的单元权利要求中,这些装置中的若干个单元可以通过同一个硬件项来具体体现。第一、第二、以及第三等的使用不表示任何顺序,可将这些单词解释为名称。上述实施例中的步骤,除有特殊说明外,不应理解为对执行顺序的限定。

[0106] 以上实施例仅用以说明本公开的技术方案,而非对其限制;尽管参照前述实施例对本公开进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本公开各实施例技术方案的精神和范围。

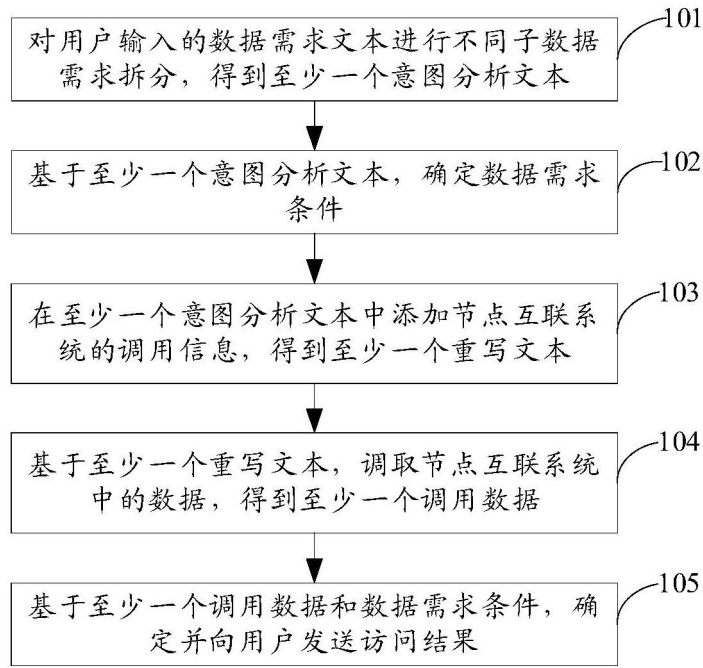


图1

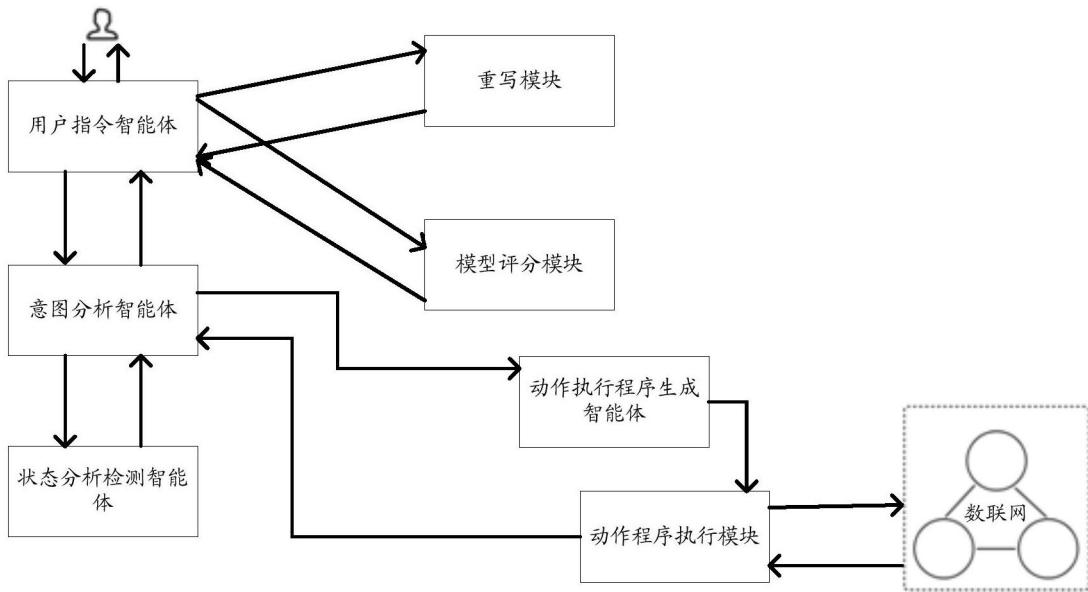


图2

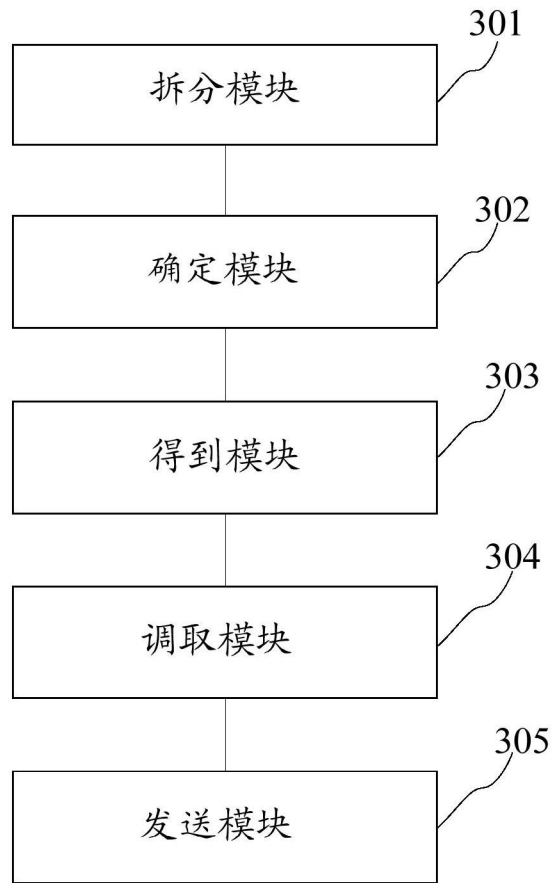


图3

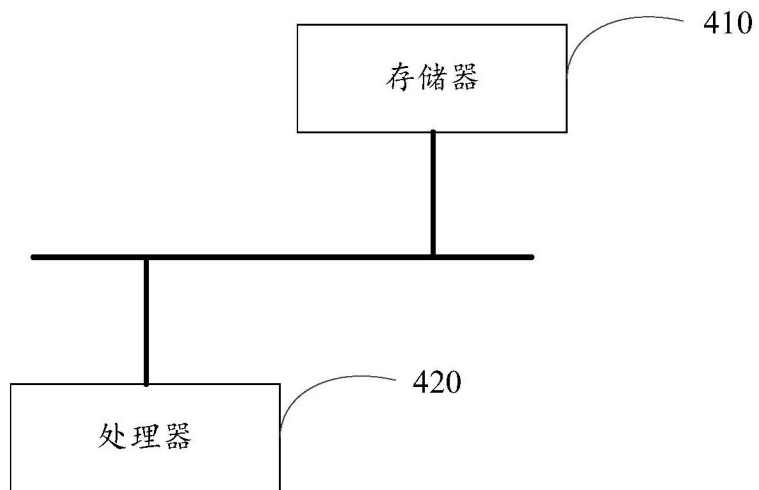


图4