

(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

(43) 国际公布日
2020 年 3 月 5 日 (05.03.2020)



(10) 国际公布号
WO 2020/042530 A1

(51) 国际专利分类号:
G06F 16/9032 (2019.01)

(21) 国际申请号: PCT/CN2019/071606

(22) 国际申请日: 2019 年 1 月 14 日 (14.01.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811021831.1 2018 年 8 月 31 日 (31.08.2018) CN

(71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。

(72) 发明人: 邱寒(QIU, Han); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。 徐国强(XU, Guoqiang); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。 柳明辉(LIU, Minghui); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。 赵云松(ZHAO, Yunsong); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。 江琳(JIANG, Lin); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518052 (CN)。

(74) 代理人: 北京英特普罗知识产权代理有限公司(INTELLECPRO CHINA LIMITED); 中国北

(54) **Title:** NATURAL LANGUAGE-BASED DATA QUERY INTENT DETERMINATION METHOD AND APPARATUS, AND COMPUTER DEVICE

(54) 发明名称: 自然语言的数据查询意图确定方法、装置和计算机设备

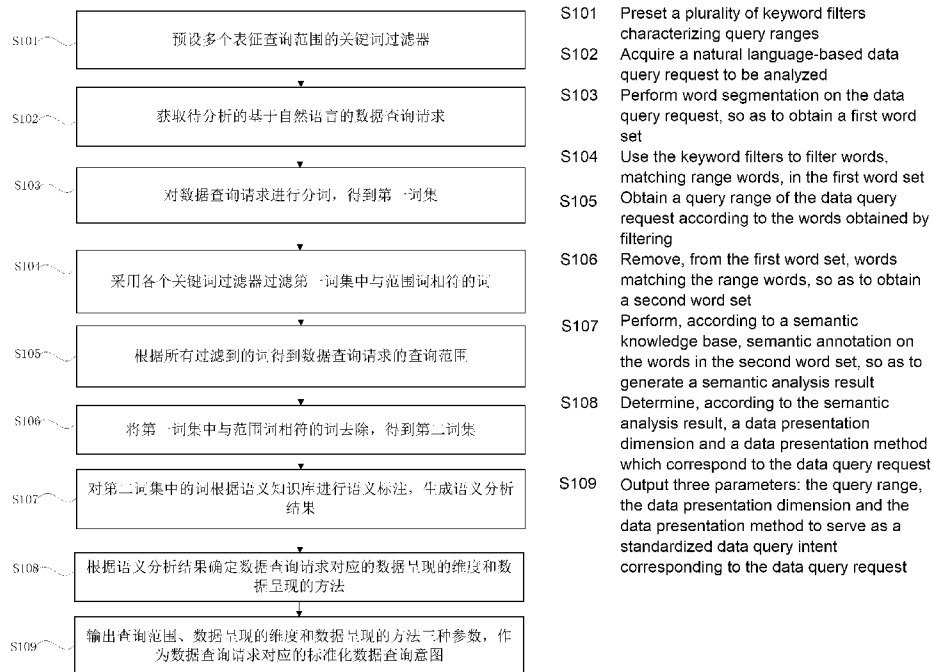


图 1

(57) **Abstract:** A natural language-based data query intent determination method and apparatus, a computer device and a storage medium, which relate to the technical field of natural language processing (NLP). Said method comprises: presetting a plurality of keyword filters (S101); acquiring a natural language-based data query request (S102); performing word segmentation on said request, so as to obtain a first word set (S103); using the keyword filters to filter the first word set (S104); obtaining a query range according to the words obtained by filtering (S105); removing, from the first word set, words matching range words, so as to obtain a second

京市西城区车公庄大街9号五栋大楼C座11层, Beijing 100044 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))。

word set (S106); performing semantic annotation on the words in the second word set, so as to generate a semantic analysis result (S107); determining, according to the semantic analysis result, a data presentation dimension and a data presentation method (S108); and outputting the query range, the data presentation dimension and the data presentation method to serve as a standardized data query intent (S109). Said method can improve natural language-based data query efficiency.

(57) 摘要: 一种自然语言的数据查询意图确定方法、装置、计算机设备及存储介质, 涉及自然语言处理(NLP, Natural Language Processing)技术领域。该方法包括预设多个关键词过滤器(S101); 获取基于自然语言的数据查询请求(S102); 对该请求进行分词得到第一词集(S103); 采用各个关键词过滤器过滤第一词集(S104); 根据过滤到的词得到查询范围(S105); 将第一词集中与范围词相符的词去除得到第二词集(S106); 对第二词集中的词进行语义标注生成语义分析结果(S107); 根据语义分析结果确定数据呈现的维度和数据呈现的方法(S108); 输出查询范围、数据呈现的维度和数据呈现的方法作为标准化数据查询意图(S109)。能够提升基于自然语言的数据查询效率。

自然语言的数据查询意图确定方法、装置和计算机设备

本申请申明享有 2018 年 8 月 31 日递交的申请号为 CN 201811021831.1、名称为“自然语言的数据查询意图确定方法、装置和计算机设备”的中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

技术领域

本申请涉及数据查询技术领域，尤其涉及一种自然语言的数据查询意图确定方法、装置、计算机设备及存储介质。

背景技术

目前，用户进行数据的查询时，通常需要用户按照自己的查询意图，先选定查询界面上设置的选择项，然后服务器一端将用户选择的选择项形成用户的查询内容，基于查询内容得到查询结果。

然而，因为是通过选择查询项的方式进行搜索，这就将导致，用户只能按照选择项中所提供的选择项进行选择，在选择项太少的时候，用户的选择范围会受到限制，在选择项过多的时候，用户在选择选择项的时候，选择的操作比较复杂。为了使用户的搜索和查询更加方便，基于自然语言的查询成为未来的重要的查询方式，其中，识别自然语言的数据查询意图是基于自然语言的查询的基础。

因而，提供一种自然语言的数据查询意图确定方法、装置、计算机设备及存储介质，以准确的从自然语言中捕捉用户数据查询意图，明确的提取出提问范围和分析维度，是本领域需要解决的技术问题。

发明内容

本申请的目的是提供一种自然语言的数据查询意图确定方法、装置、计算机设备及存储介质，用于解决现有技术存在的上述问题。

为实现上述目的，本申请提供一种自然语言的数据查询意图确定方法。

该方法包括：预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；获取待分析的基于自然语言的数据查询请求；对所述数据查询请求进行分词，得到第一词集；采用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；根据所有过滤到的词得到所述数据查询请求

的查询范围；将所述第一词集中与所述范围词相符的词去除，得到第二词集；对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

为实现上述目的，本申请提供一种自然语言的数据查询意图确定装置。

该装置包括：多个关键词过滤器，其中，所述关键词过滤器用于表征查询范围，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；获取模块，用于获取待分析的基于自然语言的数据查询请求；分词模块，用于对所述数据查询请求进行分词，得到第一词集；调用模块，用于调用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；第一确定模块，用于根据所有过滤到的词得到所述数据查询请求的查询范围；筛除模块，用于将所述第一词集中与所述范围词相符的词去除，得到第二词集；标注模块，用于对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；第二确定模块，用于根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及输出模块，用于输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

为实现上述目的，本申请还提供一种计算机设备，包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现自然语言的数据查询意图确定方法的以下步骤：

预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；获取待分析的基于自然语言的数据查询请求；对所述数据查询请求进行分词，得到第一词集；采用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；根据所有过滤到的词得到所述数据查询请求的查询范围；将所述第一词集中与所述范围词相符的词去除，得到第二词集；对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

为实现上述目的，本申请还提供计算机可读存储介质，其上存储有计算机程序，所述程序被处理器执行时实现自然语言的数据查询意图确定方法的以下步骤：

预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；获取待分析的基于自然语言的数据查询请求；对所述数据查询请求进行分词，得到第一词集；采用各个所述关键词过滤器过滤所述第一

词集中与所述范围词相符的词；根据所有过滤到的词得到所述数据查询请求的查询范围；将所述第一词集中与所述范围词相符的词去除，得到第二词集；对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

本申请提供的自然语言的数据查询意图确定方法、装置、计算机设备及存储介质，预设关键词过滤器，通过该关键词过滤器能够将关键词过滤器对应的范围词集合中的范围词过滤出来，因而，根据需要设置范围词结合，就可将预定的范围词进行过滤。在获取到基于自然语言的数据查询请求之后，首先对数据查询请求进行分词，得到一个词集，然后通过预设的关键词过滤器对该词集进行过滤，将该词集中与范围词相符的词过滤出来，根据过滤出来的词能够形成查询范围。得到查询范围后，将词集中过滤出来的词去除掉，对剩余的词进行语义标注，按照语义标注的结果确定出数据呈现的维度和数据呈现的方法。最后，以查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图，从而在进行数据查询时，能够利用统一的查询逻辑按照标准化的查询意图进行查询，提高查询效率。

附图说明

图 1 为本申请实施例 1 提供的自然语言的数据查询意图确定方法的流程图；

图 2 为本申请实施例 2 提供的自然语言的数据查询意图确定装置的框图。

图 3 为本申请实施例 3 提供的计算机设备的硬件结构图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

实施例 1

为了使用户能够基于自然语言表达查询内容，提高查询效率，对于服务器而言直接根据用户的自然语言生成查询结果，该实施例 1 提供一种自然语言的数据查询意图确定方法，该方法将基于自然语言的数据查询请求转化为标准化数据查询意图，其中，数据查询意图包括查询范围、数据呈现的维度和数据呈现的方法三种参数，具体地，图 1 为本申请实施

例 1 提供的自然语言的数据查询意图确定方法的流程图，如图 1 所示，该自然语言的数据查询意图确定方法包括如下的步骤 S101 至步骤 S109：

步骤 S101：预设多个表征查询范围的关键词过滤器。

其中，每个关键词过滤器对应一个范围词集合，范围词集合包括多个范围词，例如，
5 设置一个关键词过滤器对应的范围词集合包括的范围词有上海、天津、北京、南昌和沈阳等，又如，设置另一个关键词过滤器对应的范围词集合包括的范围词有老年、婴幼儿、青少年和中年等，具体怎样设置一个关键词过滤器对应的范围词集合，可根据数据内容进行设定。

步骤 S102：获取待分析的基于自然语言的数据查询请求。

10 提供接收用户查询请求的接口，以接收用户输入的基于自然语言的数据查询请求，具体接收方式可以为文字输入，也可以为语音输入，对于语音输入，可在后台识别为文字，无论以何种输入方式，基于自然语言的数据查询请求直接是通过自然语言体现的，例如，基于自然语言的数据查询请求为“上海地区客户的整体逾期情况怎么样”，又如基于自然语言的数据查询请求可以为“三月份第一销售小组的销售总额是多少”等。

15 步骤 S103：对数据查询请求进行分词，得到第一词集。

例如，对于数据查询请求“上海地区的男性用户的学历是如何分布的”进行分词，得到的第一词集为“上海”、“地区”、“的”、“男性”、“用户”、“的”、“学历”、“是”、“如何”、“分布”、“的”，优选地，可将分词得到的第一词集进行过滤，过滤掉无用的词，例如“的”、“是”等。

20 步骤 S104：采用各个关键词过滤器过滤第一词集中与范围词相符的词。

在该步骤中，通过预先设置的各个关键词过滤器，对第一词集中的词进行过滤，以得到第一词集中与范围词相符的词。

步骤 S105：根据所有过滤到的词得到数据查询请求的查询范围。

在上述步骤 S104 中，可以通过一个过滤器得到一个与范围词相符的词，也可以通过多个过滤器得到多个与范围词相符的词，当通过多个过滤器得到多个与范围词相符的词时，
25 各个与范围词相符的词组合得到数据查询请求的查询范围。例如，采用关键词过滤器对“上海地区的男性用户的学历是如何分布的”所对应的第一词集进行过滤时，得到的与范围词相符的词包括“上海”和“男性”，因而，得到的数据查询请求的查询范围为“上海男性”。

步骤 S106：将第一词集中与范围词相符的词去除，得到第二词集。

30 当通过步骤 S104 得到的与范围词相符的词包括“上海”和“男性”时，去除这两个词之后得到的第二词集包括“地区”、“用户”、“学历”、“如何”和“分布”。

步骤 S107：对第二词集中的词根据语义知识库进行语义标注，生成语义分析结果。

例如，语义分析结果包括多个词-义对，词-义对包括第二词集中的一个词和词的语义。

其中，词的语义包括词性和词义，在进行语义标注时，根据语义知识库，来确定各个词所对应的参数，例如，第二词集中的词包括“地区”、“用户”、“学历”、“如何”和“分布”，根据语义知识库，可确定“地区是一个名词，行政划分单位”，“用户是一个名词，人的类别”，“学历是一个名词，一种描述受教育程度的方式”、“如何是一个疑问词”和“分布是一个动词”，则语义分析结果可通过如下方式展示：

<地区，名词，行政划分单位>;

<用户，名词，人的类别>;

<学历，名词，描述受教育程度的方式>;

<如何，代词，表示疑问>;

<分布，动词，表示某种客体在一定范围散布>。

步骤 S108：根据语义分析结果确定数据查询请求对应的数据呈现的维度和数据呈现的方法。

对于数据呈现的维度的确定，预设数据内容中各个数据呈现的维度所对应的语义，例如，数据内容为针对全国 25 至 30 岁之间青年的普查数据，数据呈现的维度包括“学历”、“年收入”、“年消费”和“生育情况”等，分别预设“学历”、“年收入”、“年消费”和“生育情况”所对应的语义。

在该步骤中，根据语义分析结果确定数据查询请求对应的数据呈现的维度具体包括：

匹配词-义对中的语义和数据内容中各个数据呈现的维度所对应的语义；

当词-义对中的语义与数据内容中的第一数据呈现的维度所对应的语义相同时，将第一数据呈现的维度作为数据查询请求对应的数据呈现的维度。

其中，数据呈现的方法是指对数据进行表达的一种方式，例如针对某学校学生的年龄数据，数据呈现的方法包括通过所有年龄的平均值来呈现数据、通过各年龄段所占的比值来呈现数据、通过每个年龄值的分布情况来呈现数据等等。

对于数据呈现的方法的确定，预设数据内容中各个数据呈现的方法所对应的语义，例如，数据内容为针对全国 25 至 30 岁之间青年的普查数据，数据呈现的方法包括“分布”、“平均”等，分别预设“分布”、“平均”所对应的语义。在该步骤中，根据语义分析结果确定数据查询请求对应的数据呈现的方法具体包括：

匹配词-义对中的语义和数据内容中各个数据呈现的方法所对应的语义；

当词-义对中的语义与数据内容中的第一数据呈现的方法所对应的语义相同时，将第一数据呈现的方法作为数据查询请求对应的方法呈现的维度。

步骤 S109：输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查

询请求对应的标准化数据查询意图。

输出的每个数据查询意图均包括查询范围、数据呈现的维度和数据呈现的方法三种参数，成为标准化的、统一的数据查询意图。

5 采用该实施例提供的自然语言的数据查询意图确定方法，预设关键词过滤器，通过该关键词过滤器能够将关键词过滤器对应的范围词集合中的范围词过滤出来，因而，根据需要设置范围词结合，就可将预定的范围词进行过滤。在获取到基于自然语言的数据查询请求之后，首先对数据查询请求进行分词，得到一个词集，然后通过预设的关键词过滤器对该词集进行过滤，将该词集中与范围词相符的词过滤出来，根据过滤出来的词能够形成查询范围。得到查询范围后，将词集中过滤出来的词去除掉，对剩余的词进行语义标注，按照语义标注的结果确定出数据呈现的维度和数据呈现的方法。最后，以查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图，从而在进行数据查询时，能够利用统一的查询逻辑按照标准化的查询意图进行查询，提高查询效率。

15 通常情况下，对于一个过滤器，至多可得到一个与范围词相符的词，在一种具体的实施例中，如果得到两个或两个以上的与范围词相符的词时，通过该数据查询请求可得到相应个数的数据查询意图，其中，每个数据查询意图的查询范围不同，数据呈现的维度和数据呈现的方法相同，例如，数据查询请求为“上海地区和北京地区的男性用户的学历是如何分布的”，得到的查询范围包括“上海地区的男性”和“北京地区的男性”。

20 可选地，本申请在确定数据查询意图时，可对不完整的数据查询请求进行确定，也即，可对不包括查询范围、数据呈现的维度和数据呈现的方法的全部内容的数据查询请求进行数据查询意图的确定。对于不完整的数据查询请求，本申请提供的实施例首先进行补充，然后再进行数据查询意图的确定，以使不完整的数据查询请求也能输出标准化的、统一的数据查询意图。

25 具体地，在一种实施例中，当采用各个关键词过滤器均过滤不到第一词集中与范围词相符的词时，获取历史数据查询意图；确定与数据呈现的维度和数据呈现的方法匹配度最高的历史数据查询意图，具体地，在多条历史数据查询意图中查找包括数据呈现的维度和数据呈现的方法的历史数据查询意图，如果查询到一条，则该条历史数据查询意图即为与数据呈现的维度和数据呈现的方法匹配度最高的历史数据查询意图，如果查询到多条且多条历史数据查询意图中包括内容相同的历史数据查询意图，则将内容相同的历史数据查询意图中条数最多的历史数据查询意图作为与数据呈现的维度和数据呈现的方法匹配度最高的历史数据查询意图；获取匹配度最高的历史数据查询意图中的历史查询范围作为数据查询请求的查询范围，例如，待分析的数据查询请求为“平均年收入”，可得到数据呈现的维

度为“年收入”，数据呈现的方法为“平均”，此时，在多条历史数据查询意图中找到包括“年收入”和“平均”最多的历史数据查询意图作为匹配度最高的历史数据查询意图，该匹配度最高的历史数据查询意图中的历史查询范围为“北京青年”，则将“北京青年”作为数据查询请求“平均年收入”所对应的查询范围。其中，在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征查询范围的文字字体颜色设置为灰色，也即对于上述数据查询请求“平均年收入”，在输出标准化数据查询意图时，将“北京青年”的文字字体颜色设置为灰色，以向用户特别提醒，以使用户对查询范围进行确定。

在另一种实施例中，当根据语义分析结果无法确定出数据查询请求对应的数据呈现的维度时，获取历史数据查询意图；确定与查询范围和数据呈现的方法匹配度最高的历史数据查询意图，具体地，在多条历史数据查询意图中查找包括查询范围和数据呈现的方法的历史数据查询意图，如果查询到一条，则该条历史数据查询意图即为与查询范围和数据呈现的方法匹配度最高的历史数据查询意图，如果查询到多条且多条历史数据查询意图中包括内容相同的历史数据查询意图，则将内容相同的历史数据查询意图中条数最多的历史数据查询意图作为与查询范围和数据呈现的方法匹配度最高的历史数据查询意图；获取匹配度最高的历史数据查询意图中的历史数据呈现的维度作为数据查询请求的数据呈现的维度，例如，待分析的数据查询请求为“上海男性平均情况”，可得到查询范围为“上海男性”，数据呈现的方法为“平均”，此时，在多条历史数据查询意图中找到包括“上海男性”和“平均”最多的历史数据查询意图作为匹配度最高的历史数据查询意图，该匹配度最高的历史数据查询意图中的数据呈现的维度为“年收入”，则将“年收入”作为数据查询请求“上海男性平均情况”所对应的数据呈现的维度。其中，在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征数据呈现的维度的文字字体颜色设置为灰色，也即对于上述数据查询请求“上海男性平均情况”，在输出标准化数据查询意图时，将“年收入”的文字字体颜色设置为灰色，以向用户特别提醒，以使用户对数据呈现的维度进行确定。

在又一种实施例中，当根据语义分析结果无法确定出数据查询请求对应的数据呈现的方法时，获取历史数据查询意图；确定与查询范围和数据呈现的维度匹配度最高的历史数据查询意图，具体地，在多条历史数据查询意图中查找包括查询范围和数据呈现的维度的历史数据查询意图，如果查询到一条，则该条历史数据查询意图即为与查询范围和数据呈现的方法匹配度最高的历史数据查询意图，如果查询到多条且多条历史数据查询意图中包括内容相同的历史数据查询意图，则将内容相同的历史数据查询意图中条数最多的历史数据查询意图作为与查询范围和数据呈现的维度匹配度最高的历史数据查询意图；获取匹配

度最高的历史数据查询意图中的历史数据呈现的方法作为数据查询请求的数据呈现的方法，例如，待分析的数据查询请求为“上海男性年收入”，可得到查询范围为“上海男性”，数据呈现的维度为“年收入”，此时，在多条历史数据查询意图中找到包括“上海男性”和“年收入”最多的历史数据查询意图作为匹配度最高的历史数据查询意图，该匹配度最高的历史数据查询意图中的数据呈现的方法为“平均”，则将“平均”作为数据查询请求“上海男性年收入”所对应的数据呈现的方法。其中，在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征数据呈现的方法的文字字体颜色设置为灰色，也即对于上述数据查询请求“上海男性年收入”，在输出标准化数据查询意图时，将“平均”的文字字体颜色设置为灰色，以向用户特别提醒，以使用户对数据呈现的方法进行确定。

实施例 2

对应于上述实施例 1，本申请实施例 2 提供了一种自然语言的数据查询装置，相关部分可与上述实施例 1 相互参考。图 2 为本申请实施例 2 提供的自然语言的数据查询意图确定装置的框图，如图 2 所示，该装置包括多个关键词过滤器 201、获取模块 202、分词模块 203、调用模块 204、第一确定模块 205、筛除模块 206、标注模块 207、第二确定模块 208 和输出模块 209。

其中，关键词过滤器 201 用于表征查询范围，每个关键词过滤器对应一个范围词集合，范围词集合包括多个范围词；获取模块 202 用于获取待分析的基于自然语言的数据查询请求；分词模块 203 用于对数据查询请求进行分词，得到第一词集；调用模块 204 用于调用各个关键词过滤器过滤第一词集中与范围词相符的词；第一确定模块 205 用于根据所有过滤到的词得到数据查询请求的查询范围；筛除模块 206 用于将第一词集中与范围词相符的词去除，得到第二词集；标注模块 207 用于对第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；第二确定模块 208 用于根据语义分析结果确定数据查询请求对应的数据呈现的维度和数据呈现的方法；输出模块 209 用于输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

采用该实施例提供的自然语言的数据查询意图确定装置，预设关键词过滤器，通过该关键词过滤器能够将关键词过滤器对应的范围词集合中的范围词过滤出来，因而，根据需要设置范围词结合，就可将预定的范围词进行过滤。在获取模块获取到基于自然语言的数据查询请求之后，分词模块对数据查询请求进行分词，得到一个词集，调用模块调用预设的关键词过滤器对该词集进行过滤，将该词集中与范围词相符的词过滤出来，第一确定模块根据过滤出来的词能够形成查询范围。得到查询范围后，筛除模块将词集中过滤出来的

词去除掉，标注模块对剩余的词进行语义标注，第二确定模块按照语义标注的结果确定出数据呈现的维度和数据呈现的方法。最后，输出模块以查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图，从而在进行数据查询时，能够利用统一的查询逻辑按照标准化的查询意图进行查询，提高查询效率。

5 优选地，该装置还包括第一补充模块，当采用各个关键词过滤器均过滤不到第一词集中与范围词相符的词时，该第一补充模块用于执行以下步骤：获取历史数据查询意图；确定与数据呈现的维度和数据呈现的方法匹配度最高的历史数据查询意图；获取匹配度最高的历史数据查询意图中的历史查询范围作为数据查询请求的查询范围。其中，输出模块在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标
10 准化数据查询意图的步骤时，将表征查询范围的文字字体颜色设置为灰色。

优选地，该装置还包括第二补充模块，当第二确定模块根据语义分析结果无法确定出数据查询请求对应的数据呈现的维度时，该第二补充模块用于执行以下步骤：获取历史数据查询意图；确定与查询范围和数据呈现的方法匹配度最高的历史数据查询意图；获取匹配度最高的历史数据查询意图中的历史数据呈现的维度作为数据查询请求的数据呈现的维
15 度。其中，输出模块在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征数据呈现的维度的文字字体颜色设置为灰色。

优选地，该装置还包括第三补充模块，当第二确定模块根据语义分析结果无法确定出数据查询请求对应的数据呈现的方法时，该第三补充模块用于执行以下步骤：获取历史数据查询意图；确定与查询范围和数据呈现的维度匹配度最高的历史数据查询意图；获取匹
20 配度最高的历史数据查询意图中的历史数据呈现的方法作为数据查询请求的数据呈现的方法，其中，输出模块在输出查询范围、数据呈现的维度和数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征数据呈现的方法的文字字体颜色设置为灰色。

25 优选地，标注模块生成的语义分析结果包括多个词-义对，词-义对包括第二词集中的一个词和词的语义。

该装置还包括第一预设模块，用于预设数据内容中各个数据呈现的维度所对应的语义，第二确定模块在根据语义分析结果确定数据查询请求对应的数据呈现的维度时，具体执行以下步骤：匹配词-义对中的语义和数据内容中各个数据呈现的维度所对应的语义；当词-
30 义对中的语义与数据内容中的第一数据呈现的维度所对应的语义相同时，将第一数据呈现的维度作为数据查询请求对应的数据呈现的维度。

该装置还包括第二预设模块，用于预设数据内容中各个数据呈现的方法所对应的语义，

第二确定模块在根据语义分析结果确定数据查询请求对应的数据呈现的方法时，具体执行以下步骤：匹配词-义对中的语义和数据内容中各个数据呈现的方法所对应的语义；当词-义对中的语义与数据内容中的第一数据呈现的方法所对应的语义相同时，将第一数据呈现的方法作为数据查询请求对应的数据呈现的方法。

5

实施例 3

本实施例还提供一种计算机设备，如可以执行程序的智能手机、平板电脑、笔记本电脑、台式计算机、机架式服务器、刀片式服务器、塔式服务器或机柜式服务器（包括独立的服务器，或者多个服务器所组成的服务器集群）等。如图 3 所示，本实施例的计算机设备 20 至少包括但不限于：可通过系统总线相互通信连接的存储器 21、处理器 22，如图 3 所示。需要指出的是，图 3 仅示出了具有组件 21-22 的计算机设备 20，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。

本实施例中，存储器 21（即可读存储介质）包括闪存、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（RAM）、静态随机访问存储器（SRAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、可编程只读存储器（PROM）、磁性存储器、磁盘、光盘等。在一些实施例中，存储器 21 可以是计算机设备 20 的内部存储单元，例如该计算机设备 20 的硬盘或内存。在另一些实施例中，存储器 21 也可以是计算机设备 20 的外部存储设备，例如该计算机设备 20 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，闪存卡（Flash Card）等。当然，存储器 21 还可以既包括计算机设备 20 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 21 通常用于存储安装于计算机设备 20 的操作系统和各类应用软件，例如实施例 2 的自然语言的数据查询装置的程序代码等。此外，存储器 21 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。

处理器 22 在一些实施例中可以是中央处理器（Central Processing Unit, CPU）、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器 22 通常用于控制计算机设备 20 的总体操作。本实施例中，处理器 22 用于运行存储器 21 中存储的程序代码或者处理数据，例如自然语言的数据查询装置等。

实施例 4

本实施例还提供一种计算机可读存储介质，如闪存、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（RAM）、静态随机访问存储器（SRAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、可编程只读存储器（PROM）、磁性存储器、磁盘、光盘、服务器、App 应用商城等等，其上存储有计算机程序，程序被

处理器执行时实现相应功能。本实施例的计算机可读存储介质用于自然语言的数据查询装置，被处理器执行时实现实施例 1 的自然语言的数据查询方法。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可
5 借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

权利要求书

1.一种自然语言的数据查询意图确定方法，其特征在于，包括：

预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个
5 范围词集合，所述范围词集合包括多个范围词；

获取待分析的基于自然语言的数据查询请求；

对所述数据查询请求进行分词，得到第一词集；

采用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；

根据所有过滤到的词得到所述数据查询请求的查询范围；

10 将所述第一词集中与所述范围词相符的词去除，得到第二词集；

对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；

根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现
的方法；以及

输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为
15 数据查询请求对应的标准化数据查询意图。

2.根据权利要求1所述的自然语言的数据查询意图确定方法，其特征在于，所述方法还包括：

当采用各个所述关键词过滤器均过滤不到所述第一词集中与所述范围词相符的词
时，获取历史数据查询意图；

20 确定与所述数据呈现的维度和所述数据呈现的方法匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史查询范围作为所述数据查询请求的查询范围；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参
25 数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述查询范围的
文字字体颜色设置为灰色。

3.根据权利要求1所述的自然语言的数据查询意图确定方法，其特征在于，所述方法还包括：

当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的维度
30 时，获取历史数据查询意图；

确定与所述查询范围和所述数据呈现的方法匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的维度作为所述数据查询请求的数据呈现的维度；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的维度的文字字体颜色设置为灰色。

4.根据权利要求1所述的自然语言的数据查询意图确定方法，其特征在于，所述方法还包括：

当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的方法时，获取历史数据查询意图；

10 确定与所述查询范围和所述数据呈现的维度匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的方法作为所述数据查询请求的数据呈现的方法；

15 其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的方法的文字字体颜色设置为灰色。

5.根据权利要求1所述的自然语言的数据查询意图确定方法，其特征在于，所述语义分析结果包括多个词-义对，所述词-义对包括所述第二词集中的一个词和所述词的语义。

20 6.根据权利要求5所述的自然语言的数据查询意图确定方法，其特征在于，所述方法还包括：预设数据内容中各个数据呈现的维度所对应的语义；根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度包括：匹配所述词-义对中的语义和所述数据内容中各个数据呈现的维度所对应的语义；当所述词-义对中的语义与所述数据内容中的第一数据呈现的维度所对应的语义相同时，将所述第一数据呈现的维度作为所述数据查询请求对应的数据呈现的维度。

25 7.根据权利要求5所述的自然语言的数据查询意图确定方法，其特征在于，所述方法还包括：预设数据内容中各个数据呈现的方法所对应的语义；根据所述语义分析结果确定所述数据查询请求对应的数据呈现的方法包括：匹配所述词-义对中的语义和所述数据内容中各个数据呈现的方法所对应的语义；当所述词-义对中的语义与所述数据内容中的第一数据呈现的方法所对应的语义相同时，将所述第一数据呈现的方法作为所述数据查询请求对应的方法呈现的维度。

8.一种自然语言的数据查询装置，其特征在于，包括：

多个关键词过滤器，其中，所述关键词过滤器用于表征查询范围，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；

获取模块，用于获取待分析的基于自然语言的数据查询请求；

5 分词模块，用于对所述数据查询请求进行分词，得到第一词集；

调用模块，用于调用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；

第一确定模块，用于根据所有过滤到的词得到所述数据查询请求的查询范围；

筛除模块，用于将所述第一词集中与所述范围词相符的词去除，得到第二词集；

10 标注模块，用于对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；

第二确定模块，用于根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及

15 输出模块，用于输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

9.一种计算机设备，所述计算机设备包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现自然语言的数据查询意图确定方法的以下步骤：

20 预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；

获取待分析的基于自然语言的数据查询请求；

对所述数据查询请求进行分词，得到第一词集；

采用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；

根据所有过滤到的词得到所述数据查询请求的查询范围；

25 将所述第一词集中与所述范围词相符的词去除，得到第二词集；

对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；

根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及

30 输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

10.根据权利要求9所述的计算机设备，其特征在于，所述方法还包括：

当采用各个所述关键词过滤器均过滤不到所述第一词集中与所述范围词相符的词

时，获取历史数据查询意图；

确定与所述数据呈现的维度和所述数据呈现的方法匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史查询范围作为所述数据查询请求的查询范围；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述查询范围的文字字体颜色设置为灰色。

11.根据权利要求 9 所述的计算机设备，其特征在于，所述方法还包括：

当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的维度时，获取历史数据查询意图；

确定与所述查询范围和所述数据呈现的方法匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的维度作为所述数据查询请求的数据呈现的维度；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的维度的文字字体颜色设置为灰色。

12.根据权利要求 9 所述的计算机设备，其特征在于，所述方法还包括：

当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的方法时，获取历史数据查询意图；

确定与所述查询范围和所述数据呈现的维度匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的方法作为所述数据查询请求的数据呈现的方法；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的方法的文字字体颜色设置为灰色。

13.根据权利要求 9 所述的计算机设备，其特征在于，所述语义分析结果包括多个词-义对，所述词-义对包括所述第二词集中的一个词和所述词的语义。

14.根据权利要求 13 所述的计算机设备，其特征在于，

所述方法还包括：预设数据内容中各个数据呈现的维度所对应的语义；

根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度包括：

匹配所述词-义对中的语义和所述数据内容中各个数据呈现的维度所对应的语义；

当所述词-义对中的语义与所述数据内容中的第一数据呈现的维度所对应的语义相同时，将所述第一数据呈现的维度作为所述数据查询请求对应的数据呈现的维度。

5 15.一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：所述程序被处理器执行时实现自然语言的数据查询意图确定方法的以下步骤：

预设多个表征查询范围的关键词过滤器，其中，每个所述关键词过滤器对应一个范围词集合，所述范围词集合包括多个范围词；

获取待分析的基于自然语言的数据查询请求；

10 对所述数据查询请求进行分词，得到第一词集；

采用各个所述关键词过滤器过滤所述第一词集中与所述范围词相符的词；

根据所有过滤到的词得到所述数据查询请求的查询范围；

将所述第一词集中与所述范围词相符的词去除，得到第二词集；

对所述第二词集中的词根据语义知识库进行语义标注，生成语义分析结果；

15 根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度和数据呈现的方法；以及

输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图。

16.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述方法还包括：

20 当采用各个所述关键词过滤器均过滤不到所述第一词集中与所述范围词相符的词时，获取历史数据查询意图；

确定与所述数据呈现的维度和所述数据呈现的方法匹配度最高的所述历史数据查询意图；

25 获取所述匹配度最高的历史数据查询意图中的历史查询范围作为所述数据查询请求的查询范围；

其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述查询范围的文字字体颜色设置为灰色。

17.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述方法还包括：

30 当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的维度时，获取历史数据查询意图；

确定与所述查询范围和所述数据呈现的方法匹配度最高的所述历史数据查询意

图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的维度作为所述数据查询请求的数据呈现的维度；

5 其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的维度的文字字体颜色设置为灰色。

18.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述方法还包括：当根据所述语义分析结果无法确定出所述数据查询请求对应的数据呈现的方法时，获取历史数据查询意图；

10 确定与所述查询范围和所述数据呈现的维度匹配度最高的所述历史数据查询意图；

获取所述匹配度最高的历史数据查询意图中的历史数据呈现的方法作为所述数据查询请求的数据呈现的方法；

15 其中，在输出所述查询范围、所述数据呈现的维度和所述数据呈现的方法三种参数，作为数据查询请求对应的标准化数据查询意图的步骤时，将表征所述数据呈现的方法的文字字体颜色设置为灰色。

19.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述语义分析结果包括多个词-义对，所述词-义对包括所述第二词集中的一个词和所述词的语义。

20 20.根据权利要求 19 所述的计算机可读存储介质，其特征在于，
所述方法还包括：预设数据内容中各个数据呈现的维度所对应的语义；
根据所述语义分析结果确定所述数据查询请求对应的数据呈现的维度包括：
匹配所述词-义对中的语义和所述数据内容中各个数据呈现的维度所对应的语义；
当所述词-义对中的语义与所述数据内容中的第一数据呈现的维度所对应的语义相同时，将所述第一数据呈现的维度作为所述数据查询请求对应的数据呈现的维度。

25

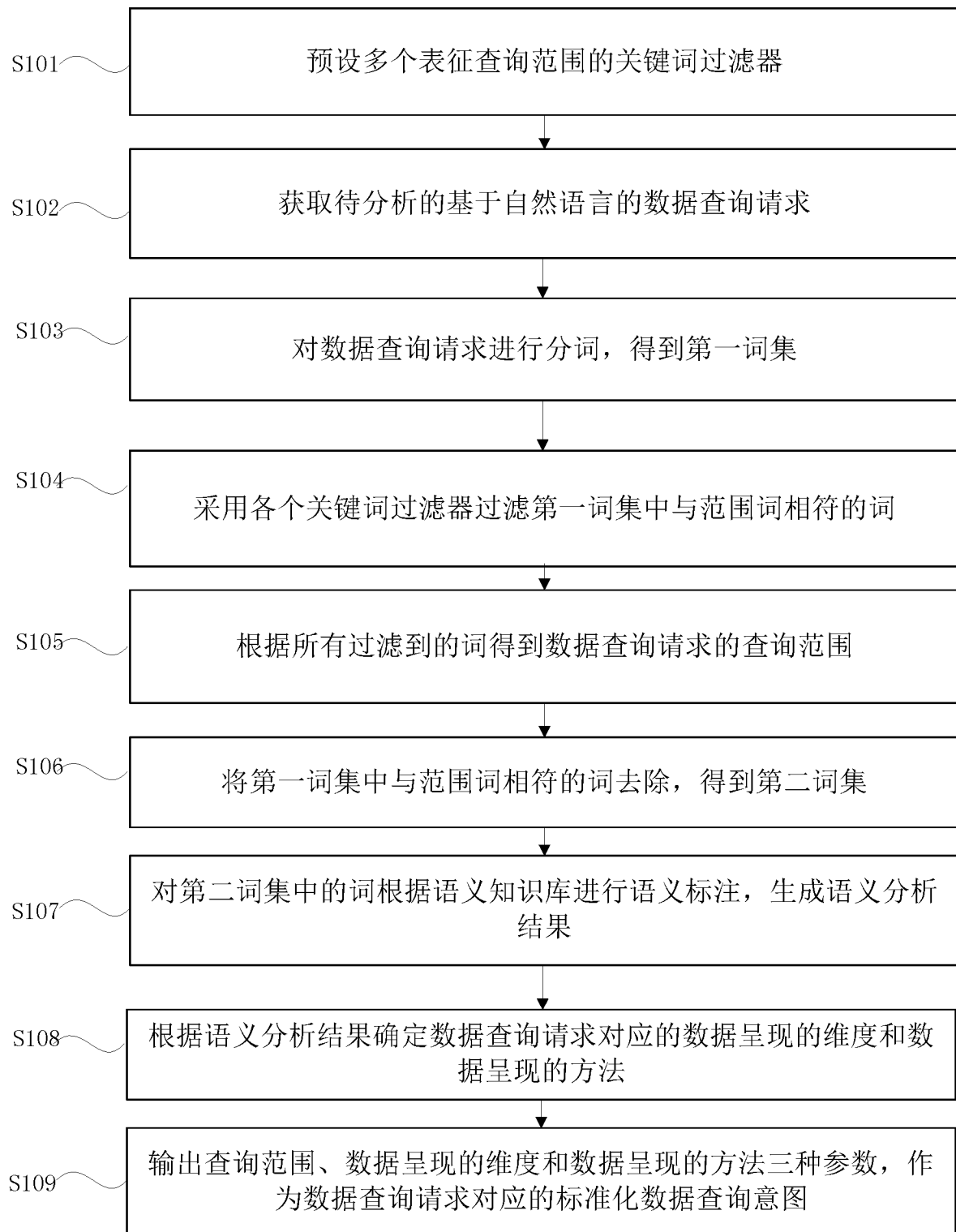


图 1

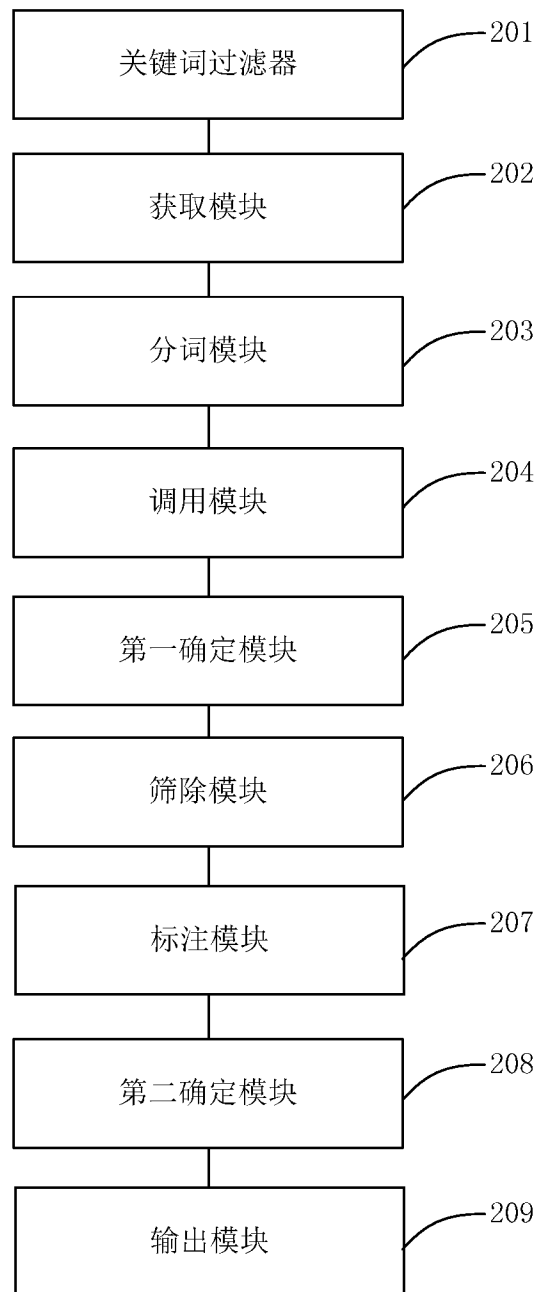


图 2

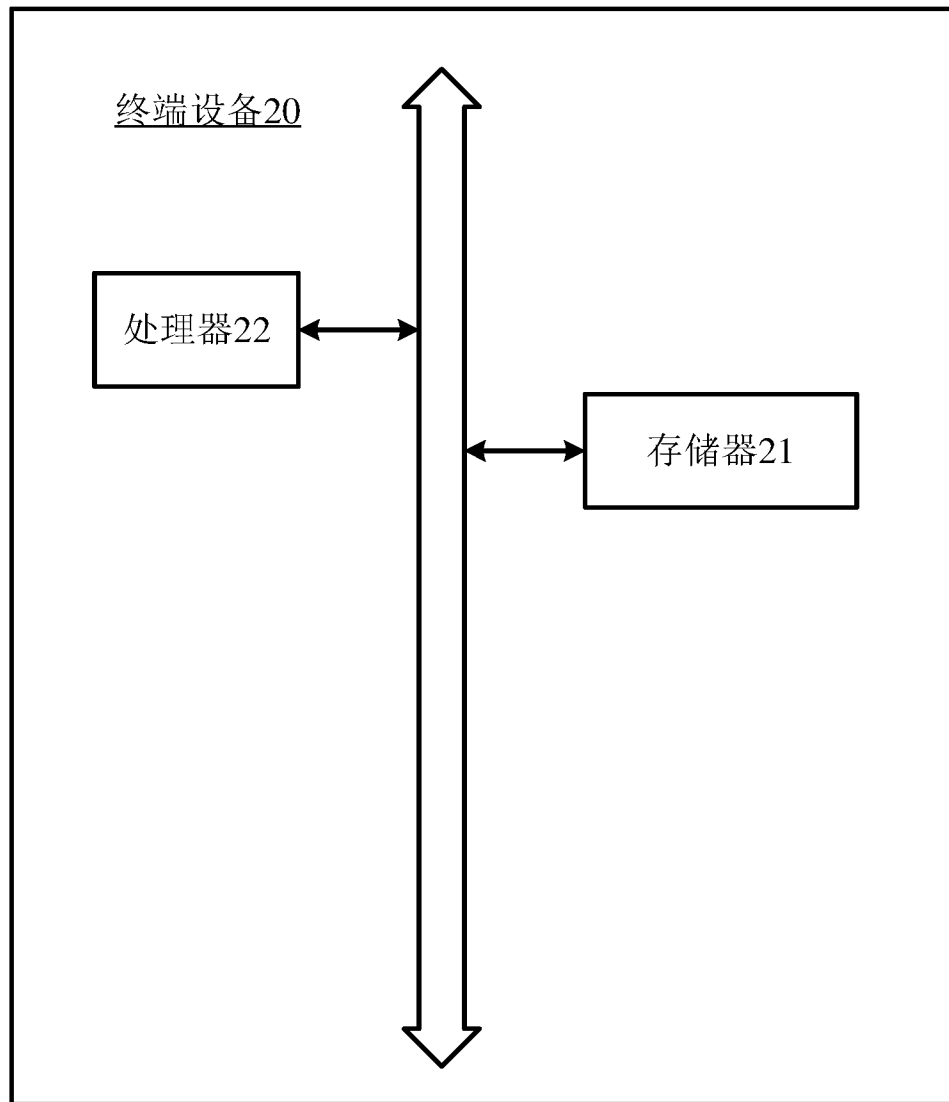


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/071606

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/9032(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, WPI, EPODOC, GOOGLE, IEEE: 查询, 自然语言, 意图, 请求, 过滤, 语义, 标注, 维度, search+, question?, understand+, NLP, natural language processing, filter, divid+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 107729336 A (ALIBABA GROUP HOLDING LIMITED) 23 February 2018 (2018-02-23) description, paragraphs [0119]-[0129], [0183], [0184] and [0304]-[0309], and figures 4 and 8	1-20
Y	CN 102737049 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 17 October 2012 (2012-10-17) claims 1-3	1-20
A	CN 107748784 A (XING, JIAHE) 02 March 2018 (2018-03-02) entire document	1-20
A	US 2009307194 A1 (DELEFEVRE, PATRICK Y.) 10 December 2009 (2009-12-10) entire document	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

10 May 2019

Date of mailing of the international search report

29 May 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/ CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/071606

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107729336	A	23 February 2018	TW	201805839	A	16 February 2018
				WO	2018028443	A1	15 February 2018
CN	102737049	A	17 October 2012	None			
CN	107748784	A	02 March 2018	None			
US	2009307194	A1	10 December 2009	None			

国际检索报告

国际申请号

PCT/CN2019/071606

A. 主题的分类

G06F 16/9032 (2019.01) i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNKI, CNPAT, WPI, EPDOC, GOOGLE, IEEE: 查询, 自然语言, 意图, 请求, 过滤, 语义, 标注, 维度, search+, question?, understand+, NLP, natural language processing, filter, divid+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 107729336 A (阿里巴巴集团控股有限公司) 2018年 2月 23日 (2018 - 02 - 23) 说明书第[0119]-[0129]、[0183]-[0184]、[0304]-[0309]段, 附图4、8	1-20
Y	CN 102737049 A (腾讯科技深圳有限公司) 2012年 10月 17日 (2012 - 10 - 17) 权利要求1-3	1-20
A	CN 107748784 A (邢加和) 2018年 3月 2日 (2018 - 03 - 02) 全文	1-20
A	US 2009307194 A1 (DELEFEVRE, PATRICK Y.) 2009年 12月 10日 (2009 - 12 - 10) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 5月 10日

国际检索报告邮寄日期

2019年 5月 29日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

吴瑶

电话号码 86-10-53961339

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/071606

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	107729336	A	2018年 2月 23日	TW	201805839	A	2018年 2月 16日
				WO	2018028443	A1	2018年 2月 15日
CN	102737049	A	2012年 10月 17日	无			
CN	107748784	A	2018年 3月 2日	无			
US	2009307194	A1	2009年 12月 10日	无			