

(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
2018 年 2 月 15 日 (15.02.2018)



(10) 国际公布号
WO 2018/028443 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2017/094790

(22) 国际申请日: 2017 年 7 月 28 日 (28.07.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201610657498.8 2016 年 8 月 11 日 (11.08.2016) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

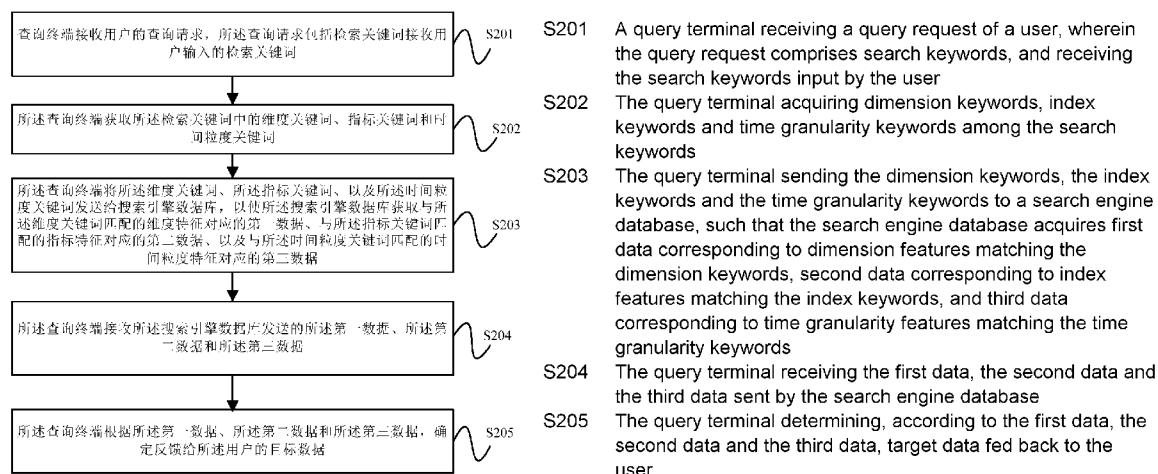
(72) 发明人: 谭纯 (TAN, Chun); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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。

(54) Title: DATA PROCESSING METHOD, DEVICE AND SYSTEM

(54) 发明名称: 数据处理方法、设备及系统



(57) Abstract: Provided are a data processing method, device and system, the method comprising: a query terminal receiving a query request of a user, wherein the query request comprises search keywords; acquiring dimension keywords, index keywords and time granularity keywords among the search keywords, and first data corresponding to dimension features matching the dimension keywords, second data corresponding to index features matching the index keywords, and third data corresponding to time granularity features matching the time granularity keywords; and the query terminal determining target data fed back to the user according to the first data, the second data and the third data. In the present embodiment, the user only need to input the search keywords once without traversing each data outlet to search data, and a search engine database can then find data related to the search keywords in all the data outlets, thereby improving the efficiency of searching data.

(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))。

(57) 摘要: 本公开提供一种数据处理方法、设备及系统, 该方法包括: 查询终端接收用户的查询请求, 查询请求包括检索关键词; 获取检索关键词中的维度关键词、指标关键词和时间粒度关键词, 以及与维度关键词匹配的维度特征对应的第一数据、与指标关键词匹配的指标特征对应的第二数据、以及与时间粒度关键词匹配的时间粒度特征对应的第三数据; 查询终端根据第一数据、第二数据、第三数据, 确定反馈给用户的目标数据。本实施例中, 用户无需遍历每个数据出口进行数据查找, 仅需输入一次检索关键词, 搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据, 从而提高了查找数据的效率。

数据处理方法、设备及系统

本申请要求 2016 年 08 月 11 日递交的申请号为 201610657498.8、发明名称为“数据处理方法、设备及系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本公开涉及互联网技术，尤其涉及一种数据处理方法、设备及系统。

背景技术

10 随着互联网的飞速发展，数据呈爆炸性增长。目前，所有具备大数据资产的公司，其存储数据的数据量均较大。而上述公司一般通过四种数据出口，该数据出口是指存储有数据的存储空间或能够生成数据的软件应用，且该存储空间或软件应用能够为数据库提供数据来源，将其存储的海量数据呈现给公司的所有员工，该四种数据出口分别为数据应用程序出口（比如阿里巴巴公司的淘宝生意经和百度公司的百度指数等）、报表出口
15 口（比如公司的工资报表）、知识库平台出口（比如百度公司的百度百科）和集群物理表出口（比如公司用户的个人信息）。

对于上述公司的非技术员工，一般需依次查找上述四种数据出口，才能获得所需的数据。比如公司的一非技术员工，有获取公司“某天家装产品的成交金额”的需求，那么该非技术人员，需依次查找公司的数据应用程序出口、报表出口、知识库平台出口以及
20 及集群物理表出口，直至查找到公司的“某天家装产品的成交金额”为止。

由于在实际应用中，上述每种数据出口所呈现数据的数据量均较大，那么非技术员工通过依次查找每种数据出口进行数据的查找，势必会造成查询数据的效率低下。

发明内容

25 本公开提供一种数据处理方法、设备及系统，以提高查找数据的效率。

一个方面，本公开提供一种数据处理系统，包括：查询终端和搜索引擎数据库；

所述查询终端，用于接收用户的查询请求，所述查询请求包括检索关键词；所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词，并将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给所述搜索引擎数据库；

30 所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所

述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征；

所述搜索引擎数据库，用于获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据，并将所述第一数据、所述第二数据和所述第三数据
5 发送给所述查询终端；

所述查询终端，还用于根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据，并将所述目标数据显示给所述用户。

另一方面，本公开提供一种数据处理方法，包括：

查询终端接收用户的查询请求，所述查询请求包括检索关键词；

10 所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词；

所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、
15 指标特征和时间粒度特征；

所述查询终端接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述第三数据；

20 所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

另一方面，本公开提供一种数据处理方法，包括：

查询终端接收用户的查询请求，所述查询请求包括检索关键词；

所述查询终端至少获取所述检索关键词中的两类关键词；

25 所述查询终端将至少两类关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述至少两类关键词分别对应的源数据；

所述查询终端接收所述搜索引擎数据库发送的所述源数据；

所述查询终端根据所述源数据，确定反馈给所述用户的目标数据。

另一方面，本公开提供一种数据处理方法，包括：

30 搜索引擎数据库接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词；

词,所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接收用户的查询请求,并从所述查询请求包括的检索关键词中获取的;

所述搜索引擎数据库预先存储有数据出口中的数据,以及所述数据的特征信息,所述特征信息包括下述至少一种:维度特征、指标特征和时间粒度特征;

- 5 所述搜索引擎数据获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据;

所述搜索引擎数据将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端,以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据,确定反馈
10 给所述用户的目标数据。

还一方面,本公开提供一种数据处理方法,包括:

搜索引擎数据库获取数据应用程序中的第一数据,以及所述第一数据的维度特征、指标特征、时间粒度特征;

- 所述搜索引擎数据库分别获取报表、知识库平台、集群物理表中的第二数据,以及
15 所述第二数据的维度特征;

所述搜索引擎数据库存储所述第一数据,以及所述第一数据的维度特征、指标特征、时间粒度特征;

所述搜索引擎数据库存储所述第二数据,以及所述第二数据的维度特征。

- 另一方面,本公开提供一种查询终端,包括:接收单元、处理单元、以及发送单元;
20 所述接收单元,用于接收用户的查询请求,所述查询请求包括检索关键词;

所述处理单元,耦合到所述接收单元,用于获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词;

- 所述发送单元,耦合到所述处理单元,用于将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库,以使所述搜索引擎数据库获取与所述
25 维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据,所述搜索引擎数据库预先存储有数据出口中的数据,以及所述数据的特征信息,所述数据出口包括下述至少一种:数据应用程序、报表、知识库平台以及集群物理表,所述特征信息包括下述至少一种:维度特征、指标特征和时间粒度特征;

- 30 所述接收单元还用于接收所述搜索引擎数据库发送的所述第一数据、所述第二数据

和所述第三数据；

所述处理单元还用于根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

再一方面，本公开提供一种搜索引擎数据库，包括：接收器、存储器、处理器、以及发送器；

所述接收器，用于接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接收用户的查询请求，并从所述查询请求包括的检索关键词中获取的；

所述存储器，用于存储数据出口中的数据，以及所述数据的特征信息，所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征；

所述处理器，耦合到所述接收器和所述存储器，用于获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据；

所述发送器，耦合到所述处理器，用于将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端，以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

在本公开中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还

可以根据这些附图获得其他的附图。

图 1 为本公开的一种可选的应用场景的示意图；

图 2 为本公开实施例提供的数据处理系统的结构示意图；

图 3 为本公开实施例一提供的数据处理方法的流程图；

5 图 4 为本公开实施例二提供的数据处理方法的流程图；

图 5 为本公开实施例三提供的数据处理方法的流程图；

图 6 为本公开实施例四提供的数据处理方法的流程图；

图 7 为本公开实施例五提供的数据处理方法的流程图；

图 8 为本公开实施例六提供的数据处理方法的流程图；

10 图 9 为本公开实施例七提供的数据处理方法的流程图；

图 10 为本公开实施例八提供的数据处理方法的流程图；

图 11 为本公开实施例九提供的数据处理方法的流程图；

图 12 为本公开实施例一提供的查询终端的结构示意图；

图 13 为本公开实施例二提供的查询终端的结构示意图；

15 图 14 为本公开实施例三提供的查询终端的结构示意图；

图 15 为本公开实施例提供的搜索引擎数据库的结构示意图。

具体实施方式

20 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

现有技术中，当公司的一非技术员工，需要获取该公司“某天家装产品的成交金额”时，需要依次查找公司的数据应用程序出口、报表出口、知识库平台出口以及集群物理表出口，直至查找到公司的“某天家装产品的成交金额”为止，如此将导致数据查找效率的下降。针对这个问题，本案提出了一种数据处理方法，现将结合图 1 介绍本案提供的数据处理方法的具体过程。

25

如图 1 所示，用户 10 通过查询终端 11 查询数据，用户 10 可以是公司里的非技术人员，还可以是消费者，查询终端 11 可以是用户 10 所属的公司内的终端设备，还可以是用户 10 的个人电脑、笔记本电脑等设备。查询终端 11 安装有搜索引擎，用户 10 可

30

通过查询终端 11 的键盘在搜索引擎的搜索框中输入搜索关键词，例如，搜索关键词是“最近一天家装类目成交金额”，语义识别模块 12 将该搜索关键词拆分为大数据领域的维度关键词、指标关键词和时间粒度关键词，具体地，维度关键词是“家装类目”、指标关键词是“成交金额”、时间粒度关键词是“最近一天”。语义识别模块 12 将该搜索关键词拆分为维度关键词、指标关键词和时间粒度关键词的方法将在下述实施例中详细描述。

语义识别模块 12 将拆分后的维度关键词“家装类目”、指标关键词“成交金额”、以及时间粒度关键词“最近一天”发送给搜索引擎数据库 13，搜索引擎数据库 13 的数据来源包括数据应用程序 15、报表 16、知识库平台 17 和集群物理表 18，其中，数据应用程序 15 具体可以是数据产品，比如阿里巴巴公司的淘宝生意经和百度公司的百度指数等，数据产品是 web 页面形式的 web 产品，数据产品与普通的 web 产品最大区别在于：数据产品承载有大量数据，且需要频繁与后台数据源交互，该后台数据源具体是存储有该数据应用程序 15 可操作的数据的器件。在本实施例中，数据应用程序 15、报表 16 中的数据可通过语法解析器 19 存储在搜索引擎数据库 13，以数据应用程序 15 为例，由于数据应用程序 15 是通过软件开发工具包（Software Development Kit，简称 SDK）开发的，所以可通过 SDK 将数据应用程序 15 中的数据采集到语法解析器 19 中。语法解析器 19 可解析出一段结构化查询语言（Structured Query Language，简称 SQL）的维度特征、指标特征、时间粒度特征和读取的表名，例如一段 SQL 具体如下：

```
SELECT stat_date AS 日期
, user_type AS 用户类型
, se_lpv_pc_ld_001 AS Pv
, se_uv_pc_ld_001 AS Uv
FROM tbbi.ads_tb_log_ld
WHERE ds = '20151026'
```

语法解析器 19 可解析出该段 SQL 的维度特征是“用户类型”，指标特征是“Pv、Uv”，时间粒度特征是“最近一天”，读取的表名是“tbbi.ads_tb_log_ld”。通过前述方法语法解析器 19 可解析出数据应用程序 15 和报表 16 中每个数据的维度特征、指标特征和时间粒度特征。

语法解析器 19 将解析后的数据发送给搜索引擎数据库 13，搜索引擎数据库 13 中不仅存储有数据本身，同时还存储有数据的维度特征、指标特征和时间粒度特征。另外，

搜索引擎数据库 13 还可以存储有知识库平台 17 和集群物理表 18 中的数据，存储过程具体为：对知识库平台 17 和集群物理表 18 中的每个数据进行拆分，提取出拆分后的每个数据的维度特征，并将知识库平台 17 和集群物理表 18 中的每个数据，以及每个数据的维度特征存储在搜索引擎数据库 13。如此，搜索引擎数据库 13 中存储的每个数据至少具有维度特征。

当搜索引擎数据库 13 接收到语义识别模块 12 发送的维度关键词“家装类目”、指标关键词“成交金额”、以及时间粒度关键词“最近一天”时，分别查找出与维度关键词“家装类目”匹配的数据、与指标关键词“成交金额”匹配的数据、以及与时间粒度关键词“最近一天”匹配的数据，搜索引擎数据库 13 将查找出的匹配数据发送给排序器 14，若搜索引擎数据库 13 查找出的匹配数据只有一个，则排序器 14 将该匹配数据发送给查询终端 11，查询终端 11 显示该匹配数据；若搜索引擎数据库 13 查找出的匹配数据有多个，则排序器 14 按照预设算法对该多个匹配数据进行排序，将排序后的多个匹配数据发送给查询终端 11，查询终端 11 按照排序的先后顺序显示该多个匹配数据。在本实施例中，排序器 14 对该多个匹配数据进行排序的预设算法包括如下至少一种：PageRank 算法、CUS—距离算法、文档主题生成模型（Latent Dirichlet Allocation，简称 LDA）算法、宽度优先搜索（Breadth First Search，简称 BFS）算法等。

本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

25

图 2 为本公开实施例提供的数据处理系统的结构示意图，如图 2 所示，数据处理系统包括查询终端 1 和搜索引擎数据库 2，其中，查询终端 1 用于接收用户的查询请求，所述查询请求包括检索关键词；所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词，并将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给所述搜索引擎数据库。

30

如图 1 所示, 查询终端 11 接收用户 10 的查询请求, 查询请求的方式可以有多种, 例如, 用户 10 在查询终端 11 的搜索引擎上输入文字、语音, 该文字或语音包括用户 10 预检索的关键词。如图 1 所示, 语义识别模块 12 和排序器 14 可以是属于查询终端 11 中的模块, 语义识别模块 12 将该搜索关键词拆分为大数据领域的维度关键词、指标关键词和时间粒度关键词, 具体地, 维度关键词是“家装类目”、指标关键词是“成交金额”、时间粒度关键词是“最近一天”。语义识别模块 12 还将维度关键词“家装类目”、指标关键词“成交金额”、时间粒度关键词“最近一天”发送给搜索引擎数据库 2。

搜索引擎数据库 2 预先存储有数据出口中的数据, 以及所述数据的特征信息, 所述特征信息包括下述至少一种: 维度特征、指标特征和时间粒度特征。

10 可选的, 在本实施例中, 数据出口包括: 数据应用程序、报表、知识库平台以及集群物理表, 搜索引擎数据库 13 存储有数据应用程序、报表、知识库平台以及集群物理表中的数据, 以及每个数据的特征信息, 数据应用程序中的每个数据具有维度特征、指标特征和时间粒度特征, 报表、知识库平台以及集群物理表中的数据均具有维度特征。

15 搜索引擎数据库 2 用于获取与所述维度关键词匹配的维度特征对应的第一数据、与
所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时
间粒度特征对应的第三数据, 并将所述第一数据、所述第二数据和所述第三数据发送给
所述查询终端。

当搜索引擎数据库 13 接收到语义识别模块 12 发送的维度关键词“家装类目”、指
标关键词“成交金额”、以及时间粒度关键词“最近一天”时, 可分别查找出与维度关
20 键词“家装类目”匹配的数据、与指标关键词“成交金额”匹配的数据、以及与时间粒
度关键词“最近一天”匹配的数据。搜索引擎数据库 13 可将语义识别模块 12 识别出的
维度关键词“家装类目”与其存储的数据的维度特征进行匹配, 获得与所述维度关键词
匹配的维度特征对应的第一数据, 该第一数据可以是多个数据, 并且该第一数据可以是
源自于数据应用程序 15、报表 16、知识库平台 17 或集群物理表 18 的数据。

25 另外, 搜索引擎数据库 13 还可将语义识别模块 12 识别出的指标关键词“成交金额”
与其存储的数据的指标特征进行匹配, 获得与所述指标关键词匹配的指标特征对应的第
二数据, 该第二数据可以是源自于数据应用程序 15 的多个数据。

此外, 搜索引擎数据库 13 还可将语义识别模块 12 识别出的时间粒度关键词“最近
一天”与其存储的数据的时间粒度特征进行匹配, 获得与所述时间粒度关键词匹配的时
30 间粒度特征对应的第三数据, 该第三数据可以是源自于数据应用程序 15 的多个数据。

搜索引擎数据库 13 其获得的所述第一数据、所述第二数据和所述第三数据发送给查询终端 11，具体可以发送给查询终端 11 中的排序器 14。

查询终端 1 还用于根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据，并将所述目标数据显示给所述用户。

5 若搜索引擎数据库 13 查找出的匹配数据只有一个，即所述第一数据、所述第二数据和所述第三数据为同一数据，则排序器 14 将该匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器显示该匹配数据。

若搜索引擎数据库 13 查找出的匹配数据有多个，即所述第一数据、所述第二数据和所述第三数据不为同一数据，则排序器 14 按照预设算法对该多个匹配数据进行排序，
10 将排序后的多个匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器按照排序的先后顺序显示该多个匹配数据。

本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词
15 进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

20

图 3 为本公开实施例一提供的数据处理方法的流程图，如图 3 所示，该方法包括如下步骤：

步骤 S201、查询终端接收用户的查询请求，所述查询请求包括检索关键词。

如图 1 所示，查询终端 11 接收用户 10 的查询请求，查询请求的方式可以有多种，
25 例如，用户 10 在查询终端 11 的搜索引擎上输入文字、语音，该文字或语音包括用户 10 预检索的关键词；或者，查询终端 11 的搜索引擎上设置有下拉列表，该列表中预先存储有关键词，用户可以通过选择列表中的关键词并点击的方式输入预检索的关键词；再或者，用户 10 在查询终端 11 上预览文字信息，用户 10 从其预览的文字信息中选择关键词，通过拖动、滑动、点击功能键的方式对该关键词进行检索。

30 用户 10 通过查询终端 11 查询数据，用户 10 可以是公司里的非技术人员，还可以

是消费者，查询终端 11 可以是用户 10 所属的公司内的终端设备，还可以是用户 10 的个人电脑、笔记本电脑等设备。查询终端 11 安装有搜索引擎，用户 10 可通过查询终端 11 的键盘在搜索引擎的搜索框中输入搜索关键词，例如，搜索关键词是“最近一天家装类目成交金额”。

- 5 步骤 S202、所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词。

如图 1 所示，语义识别模块 12 和排序器 14 可以是属于查询终端 11 中的模块，也可以是属于搜索引擎数据库 13 中的模块，查询终端 11 和搜索引擎数据库 13 可以直接连接，也可以通过其他设备间接连接。在本实施例中，以语义识别模块 12 和排序器 14
10 属于查询终端 11、查询终端 11 和搜索引擎数据库 13 直接连接为例。

语义识别模块 12 将该搜索关键词拆分为大数据领域的维度关键词、指标关键词和时间粒度关键词，具体地，维度关键词是“家装类目”、指标关键词是“成交金额”、时间粒度关键词是“最近一天”。

- 步骤 S203、所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度
15 关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

在实施例中，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征。

- 20 所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征，所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表。可选的，在本实施例中，数据出口包括：数据应用程序、报表、知识库平台以及集群物
25 理表，搜索引擎数据库 13 存储有数据应用程序、报表、知识库平台以及集群物理表中的数据。

- 如图 1 所示，搜索引擎数据库 13 的数据来源包括数据应用程序 15、报表 16、知识库平台 17 和集群物理表 18，其中，数据应用程序 15 具体可以是数据产品，比如阿里巴巴公司的淘宝生意经和百度公司的百度指数等，数据产品是 web 页面形式的 web 产品，
30 数据产品与普通的 web 产品最大区别在于：数据产品承载有大量数据，且需要频繁与后

台数据源交互，该后台数据源具体是存储有该数据应用程序 15 可操作的数据的器件。
在本实施例中，数据应用程序 15 中的数据可通过语法解析器 19 存储在搜索引擎数据库 13，具体的，通过 SDK 将数据应用程序 15 中的数据采集到语法解析器 19 中。语法解析器 19 可解析出一段结构化查询语言(Structured Query Language，简称 SQL)的维度特征、指标特征、时间粒度特征和读取的表名，例如一段 SQL 具体如下：

```
SELECT stat_date AS 日期  
    , user_type AS 用户类型  
    , se_lpv_pc_ld_001 AS Pv  
    , se_uv_pc_ld_001 AS Uv  
FROM tbbi.ads_tb_log_ld  
WHERE ds = '20151026'
```

语法解析器 19 可解析出该段 SQL 的维度特征是“用户类型”，指标特征是“Pv、Uv”，时间粒度特征是“最近一天”，读取的表名是“tbbi.ads_tb_log_ld”。通过前述方法，语法解析器 19 可解析出数据应用程序 15 中每个数据的维度特征、指标特征和时间粒度特征。语法解析器 19 将解析后的数据发送给搜索引擎数据库 13，搜索引擎数据库 13 中不仅存储有数据本身，同时还存储有数据的维度特征、指标特征和时间粒度特征。

另外，搜索引擎数据库 13 还可以存储有报表 16、知识库平台 17 和集群物理表 18 中的数据，存储过程具体为：对报表 16、知识库平台 17 和集群物理表 18 中的每个数据进行拆分，提取出拆分后的每个数据的维度特征，并将报表 16、知识库平台 17 和集群物理表 18 中的每个数据，以及每个数据的维度特征存储在搜索引擎数据库 13。如此，搜索引擎数据库 13 中存储的每个数据至少具有维度特征。

当搜索引擎数据库 13 接收到语义识别模块 12 发送的维度关键词“家装类目”、指标关键词“成交金额”、以及时间粒度关键词“最近一天”时，可分别查找出与维度关键词“家装类目”匹配的数据、与指标关键词“成交金额”匹配的数据、以及与时间粒度关键词“最近一天”匹配的数据。

在本实施例中，搜索引擎数据库 13 中存储有数据应用程序 15 中的数据，以及数据应用程序 15 中每个数据的维度特征、指标特征和时间粒度特征。另外，搜索引擎数据库 13 还存储有报表 16、知识库平台 17 和集群物理表 18 中的数据，以及报表 16、知识库平台 17 和集群物理表 18 中每个数据的维度特征。另外，搜索引擎数据库 13 中各数

据的维度特征可能不同，可能相同；各数据的指标特征可能不同，可能相同；各数据的时间粒度特征可能不同，可能相同。

本实施例中的搜索引擎数据库 13 可将语义识别模块 12 识别出的维度关键词“家装类目”与其存储的数据的维度特征进行匹配，获得与所述维度关键词匹配的维度特征对应的第一数据，该第一数据可以是多个数据，并且该第一数据可以是源自于数据应用程序 15、报表 16、知识库平台 17 或集群物理表 18 的数据。

另外，搜索引擎数据库 13 还可将语义识别模块 12 识别出的指标关键词“成交金额”与其存储的数据的指标特征进行匹配，获得与所述指标关键词匹配的指标特征对应的第二数据，该第二数据可以是源自于数据应用程序 15 的多个数据。

此外，搜索引擎数据库 13 还可将语义识别模块 12 识别出的时间粒度关键词“最近一天”与其存储的数据的时间粒度特征进行匹配，获得与所述时间粒度关键词匹配的时间粒度特征对应的第三数据，该第三数据可以是源自于数据应用程序 15 的多个数据。

步骤 S204、所述查询终端接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述第三数据。

搜索引擎数据库 13 其获得的所述第一数据、所述第二数据和所述第三数据发送给查询终端 11，具体可以发送给查询终端 11 中的排序器 14。

步骤 S205、所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

若搜索引擎数据库 13 查找出的匹配数据只有一个，即所述第一数据、所述第二数据和所述第三数据为同一数据，则排序器 14 将该匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器显示该匹配数据。

若搜索引擎数据库 13 查找出的匹配数据有多个，即所述第一数据、所述第二数据和所述第三数据不为同一数据，则排序器 14 按照预设算法对该多个匹配数据进行排序，将排序后的多个匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器按照排序的先后顺序显示该多个匹配数据。

本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数

据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

5 图 4 为本公开实施例二提供的数据处理方法的流程图，如图 4 所示，在图 3 所示实施例的基础上，所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词的方法可以具体包括如下步骤：

步骤 S301、所述查询终端对所述检索关键词进行分词处理获得多个目标分词。

例如步骤 S201 所述，用户输入的搜索关键词是“最近一天家装类目成交金额”。

10 查询终端 11 还可通过 TF-idf 算法对用户输入的检索关键词进行拆分，获得多个目标分词，多个目标分词分别为“最近一天”、“家装类目”、“成交金额”。

步骤 S302、所述查询终端根据各目标分词查询预设的映射表，所述映射表包括维度分词、指标分词和时间粒度分词。

在本实施例中，查询终端 11 预先建立有映射表，该映射表包括维度分词、指标分
15 词和时间粒度分词，维度分词可以是多个具有维度特征的分词，指标分词可以是多个具有指标特征的分词，时间粒度分词可以是多个具有时间粒度特征的分词。根据步骤 S301 拆分后的多个目标分词，查询终端 11 分别查询该映射表，对于每个目标分词，确定该映射表中是否存在与该目标分词匹配的分词。

步骤 S303、所述查询终端将所述多个目标分词中与所述维度分词匹配的目标分词确
20 定为所述维度关键词。

例如，上述多个目标分词中的“家装类目”与映射表中的维度分词匹配，则将“家装类目”作为检索关键词中的维度关键词。

步骤 S304、所述查询终端将所述多个目标分词中与所述指标分词匹配的目标分词确
定为所述指标关键词。

25 例如，上述多个目标分词中的“成交金额”与映射表中的指标分词匹配，则将“成交金额”作为检索关键词中的指标关键词。

步骤 S305、所述查询终端将所述多个目标分词中与所述时间粒度分词匹配的目标分词确定为所述时间粒度关键词。

例如，上述多个目标分词中的“最近一天”与映射表中的时间粒度分词匹配，则将
30 “最近一天”作为检索关键词中的时间粒度关键词。

本实施例中，通过对检索关键词进行分词处理获得多个目标分词，根据预先建立的映射表查询该多个目标分词中的维度关键词、指标关键词以及时间粒度关键词，提高了确定检索关键词中维度关键词、指标关键词以及时间粒度关键词的效率。

5 图 5 为本公开实施例三提供的数据处理方法的流程图，如图 5 所示，在上述任一实施例的基础上，以实施例二为基础，本实施例提供的数据处理方法的具体步骤如下：

步骤 S401、查询终端接收用户的查询请求，所述查询请求包括检索关键词。

步骤 S402、所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词。

10 步骤 S403、所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

15 步骤 S404、所述查询终端接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述第三数据。

步骤 S401-步骤 S404 分别与步骤 S201-S204 一致，具体方法此处不再赘述。

步骤 S405、所述查询终端确定所述第一数据、所述第二数据和所述第三数据是否为同一数据，若是，则执行步骤 S406，否则，执行步骤 S407。

步骤 S406、所述查询终端将所述同一数据确定为反馈给所述用户的目标数据。

20 如图 1 所示，若搜索引擎数据库 13 查找出的匹配数据只有一个，即所述第一数据、所述第二数据和所述第三数据为同一数据，则排序器 14 将该匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器显示该匹配数据。

步骤 S407、所述查询终端对所述第一数据、所述第二数据和所述第三数据进行排序，将排序后的数据确定为反馈给所述用户的目标数据。

25 若搜索引擎数据库 13 查找出的匹配数据有多个，即所述第一数据、所述第二数据和所述第三数据不为同一数据，则排序器 14 按照预设算法对该多个匹配数据进行排序，将排序后的多个匹配数据发送给查询终端 11 的显示器，查询终端 11 的显示器按照排序的先后顺序显示该多个匹配数据。

30 在步骤 S407 中，所述查询终端对所述第一数据、所述第二数据和所述第三数据进行排序的方法具体可以包括如下步骤：

步骤 S51、所述查询终端计算所述第一数据、所述第二数据和所述第三数据中每个数据的权重值。

具体可通过 Pagerank 算法计算每个数据的权重值。

步骤 S52、所述查询终端计算所述第一数据、所述第二数据和所述第三数据中每个数据与所述检索关键词的相似度。

具体可利用 CUS—距离算法，计算每个数据与用户输入的检索关键词的相似度。

步骤 S53、所述查询终端根据所述每个数据的权重值和相似度，计算所述每个数据的排序值。

具体的，可将每个数据的权重值和相似度相加得到的值作为该数据的排序值。

步骤 S54、所述查询终端根据所述每个数据的排序值，对所述第一数据、所述第二数据和所述第三数据中的每个数据进行排序。

具体的，可根据每个数据的排序值，按照从大到小的顺序对所述第一数据、所述第二数据和所述第三数据中的每个数据进行排序。

可选的，所述查询终端根据所述每个数据的排序值，确定所述第一数据、所述第二数据和所述第三数据中排序值大于第一阈值的数据；所述查询终端对所述排序值大于第一阈值的数据，按照所述排序值的大小进行排序。

另外，计算出所述第一数据、所述第二数据和所述第三数据中每个数据的排序值后，可确定出所述第一数据、所述第二数据和所述第三数据中排序值大于第一阈值的数据，并对排序值大于第一阈值的数据，按照所述排序值的大小进行排序。

本实施例中，对搜索引擎数据库查找出的多个与检索关键词匹配的数据进行排序，排序的依据是每个数据的排序值，该排序值与每个数据的权重值和该数据与检索关键词的相似度有关，则排序值越大，表示该数据与检索关键词的关联性越强，将排序后的多个数据反馈给用户，用户可方便的查看到与检索关键词关联性最强的数据，提高了用户体验。

图 6 为本公开实施例四提供的数据处理方法的流程图，如图 6 所示，在上述任一实施例的基础上，以实施例二为基础，本实施例提供的数据处理方法的具体步骤如下：

步骤 S601、查询终端接收用户的查询请求，所述查询请求包括检索关键词。

步骤 S602、所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词。

步骤 S603、所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

5 步骤 S604、所述查询终端接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述第三数据。

步骤 S605、所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

步骤 S601-步骤 S605 分别与步骤 S201-步骤 S205 一致，具体方法此处不再赘述。

10 步骤 S606、所述查询终端接收所述用户对所述目标数据的点击操作。

在步骤 S407 之后，可将排序后的多个目标数据显示在查询终端，用户通过查询终端可点击查看到该多个目标数据。当用户点击某个目标数据时，查询终端可接收到该用户对该目标数据的点击操作。

15 步骤 S607、所述查询终端根据所述点击操作建立所述用户与所述目标数据的关联关系。

所述关联关系包括关联度，所述关联度标识所述用户与所述目标数据的关联程度。

在本实施例中，查询终端根据用户点击某个目标数据产生的点击操作建立所述用户与所述目标数据的关联关系，另外，还可根据关联规则和协同过滤规则计算用户与其点击的目标数据的关联度，该用户点击的目标数据的个数可以是多个。

20 步骤 S608、当用户未输入所述检索关键词时，所述查询终端根据所述关联关系显示所述目标数据。

当用户在查询终端 11 未输入检索关键词时，查询终端 11 可根据用户与其点击过的目标数据之间的关联关系显示该目标数据，即查询终端 11 可将用户点击过的目标数据显示给用户。

25 具体的，所述关联关系包括关联度，所述关联度标识所述用户与所述目标数据的关联程度。所述查询终端根据所述关联关系显示所述目标数据，包括：所述查询终端显示关联度大于第二阈值的所述目标数据。

30 可选的，查询终端显示关联度大于第二阈值的所述目标数据。用户与其点击过的每个目标数据的关联关系还包括用户与该目标数据的关联度，查询终端 11 还可以显示用户点击过的关联度大于第二阈值的目标数据。

本实施例中，通过建立用户与其点击过的目标数据之间的关联关系，当用户未输入检索关键词时，可根据用户与目标数据之间的关联关系，显示用户点击过的目标数据，提高了用户查询数据的便捷性。

5 图 7 为本公开实施例五提供的数据处理方法的流程图，如图 7 所示，本实施例提供的数据处理方法的具体步骤如下：

步骤 S501、查询终端接收用户的查询请求，所述查询请求包括检索关键词。

10 如图 1 所示，查询终端 11 接收用户 10 的查询请求，查询请求的方式可以有多种，例如，用户 10 在查询终端 11 的搜索引擎上输入文字、语音，该文字或语音包括用户 10 预检索的关键词。

步骤 S502、所述查询终端至少获取所述检索关键词中的两类关键词。

在本实施例中，查询终端对用户请求查询的检索关键词分类时，可以不局限于维度关键词、指标关键词和时间粒度关键词这三类关键词，因为，并不是用户请求查询的每个检索关键词都包括维度关键词、指标关键词和时间粒度关键词这三类关键词，因此，
15 如图 1 所示的查询终端 11 对应的语义识别模块 12 还可以将用户请求查询的检索关键词拆分为至少两类关键词，例如，用户是卖家，卖家请求查询的检索关键词是“有客户评价我的商品吗”，可拆分出动词“评价”、名词“商品”。

步骤 S503、所述查询终端将至少两类关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述至少两类关键词分别对应的源数据。

20 查询终端将动词“评价”、名词“商品”发送给搜索引擎数据库，搜索引擎数据库存储有卖家的所有商品的商品信息，以及每件商品的评价信息。搜索引擎数据库根据“商品”获得该卖家的所有商品的商品信息，该商品信息具体包括名称、产地、材料等，根据“评价”获得所有商品的评价信息。

步骤 S504、所述查询终端接收所述搜索引擎数据库发送的所述源数据。

25 搜索引擎数据库将商品信息和评价信息发送给查询终端，由于此处的商品信息可以是多个，评价信息也可以是多个。

步骤 S505、所述查询终端根据所述源数据，确定反馈给所述用户的目标数据。

30 查询终端可以根据每个商品的评价信息的个数，确定反馈给所述用户评价信息最多的商品的商品信息，也可以将每个商品的前几条评价信息反馈给所述用户，本实施例不限定查询终端确定反馈给所述用户的目标数据的具体实现方式。

本实施例中，通过对检索关键词进行分类，分类的结果并不局限于维度关键词、指标关键词以及时间粒度关键词，提高了对检索关键词分类的灵活度，增加了对检索关键词进行检索的灵活度，同时也扩大了检索范围。

5 图 8 为本公开实施例六提供的数据处理方法的流程图，如图 8 所示，本实施例提供的数据处理方法的具体步骤如下：

步骤 S701、搜索引擎数据库接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词。

10 其中，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接收用户的查询请求，并从所述查询请求包括的检索关键词中获取的。

在本实施例中，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征。

所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表。

15 步骤 S702、所述搜索引擎数据获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

20 步骤 S703、所述搜索引擎数据将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端，以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

本实施例所述方法的原理与图 3 所示实施例方法的原理一致，具体过程此处不再赘述。

25 本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

30

图 9 为本公开实施例七提供的数据处理方法的流程图，如图 9 所示，本实施例提供的数据处理方法的具体步骤如下：

步骤 S801、所述搜索引擎数据库存储所述数据应用程序、所述报表、所述知识库平台以及所述集群物理表中的数据。

在图 3 所示实施例的基础上，在接收用户输入的检索关键词之前，搜索引擎数据库 13 预先存储有数据应用程序、报表、知识库平台以及集群物理表中的数据。

具体的，数据应用程序 15 中的数据可通过语法解析器 19 存储在搜索引擎数据库 13，具体的，通过 SDK 将数据应用程序 15 中的数据采集到语法解析器 19 中。语法解析器 19 可解析出一段结构化查询语言 (Structured Query Language，简称 SQL) 的维度特征、指标特征、时间粒度特征和读取的表名，例如一段 SQL 具体如下：

```
SELECT stat_date AS 日期
      , user_type AS 用户类型
      , se_lpv_pc_ld_001 AS Pv
15      , se_uv_pc_ld_001 AS Uv
FROM tbbi.ads_tb_log_ld
WHERE ds = '20151026'
```

语法解析器 19 可解析出该段 SQL 的维度特征是“用户类型”，指标特征是“Pv、Uv”，时间粒度特征是“最近一天”，读取的表名是“tbbi.ads_tb_log_ld”。通过前述方法，语法解析器 19 可解析出数据应用程序 15 中每个数据的维度特征、指标特征和时间粒度特征。语法解析器 19 将解析后的数据发送给搜索引擎数据库 13，搜索引擎数据库 13 中不仅存储有数据本身，同时还存储有数据的维度特征、指标特征和时间粒度特征。

另外，搜索引擎数据库 13 还可以存储有报表 16、知识库平台 17 和集群物理表 18 中的数据，存储过程具体为：对报表 16、知识库平台 17 和集群物理表 18 中的每个数据进行拆分，提取出拆分后的每个数据的维度特征，并将报表 16、知识库平台 17 和集群物理表 18 中的每个数据，以及每个数据的维度特征存储在搜索引擎数据库 13。如此，搜索引擎数据库 13 中存储的每个数据至少具有维度特征。

步骤 S802、搜索引擎数据库接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查

询终端接收用户的查询请求，并从所述查询请求包括的检索关键词中获取的。

在本实施例中，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征。

所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表。

步骤 S803、所述搜索引擎数据获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

步骤 S804、所述搜索引擎数据将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端，以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

步骤 S802-步骤 S804 所述的方法原理与步骤 S701-步骤 S703 所述的方法原理一致，此处不再赘述。

本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

图 10 为本公开实施例八提供的数据处理方法的流程图，如图 10 所示，所述搜索引擎数据库存储所述数据应用程序、所述报表、所述知识库平台以及所述集群物理表中的数据具体可以包括如下步骤 S901 和 S902：

步骤 S901、所述搜索引擎数据库存储所述数据应用程序中的数据。

步骤 S901 的可以通过如下步骤 S11-S13 来实现：

步骤 S11、所述搜索引擎数据库获取所述数据应用程序访问数据源的访问逻辑。

所述访问逻辑包括所述数据应用程序中的数据，所述数据源存储有所述数据的产出逻辑。

本实施例介绍将所述数据应用程序中的数据存储在所述搜索引擎数据库的方法，且本实施例所述的方法不同于上述实施例所述的通过语法解析器 19 将数据应用程序 15 中的数据存储在搜索引擎数据库 13 的方法。

在本实施例中，所述数据应用程序可具体为 Web 页面形式，需频繁与后台数据源进行交互；所述后台数据源可具体为存储所述数据应用程序操作数据的器件。由于数据应用程序是根据 SDK 所开发的，SDK 对数据应用程序具有最大的操作权限，因此可通过 SDK 捕获数据应用程序对后台数据源的第一访问逻辑，所述第一访问逻辑中包括数据应用程序访问后台数据源的时间、用户对数据应用程序的第二访问逻辑等字段。因此，通过现有的解析方式，即可获取用户对数据应用程序的第二访问逻辑。用户对数据应用程序的第二访问逻辑中包括用户访问数据应用程序的时间、用户当前所访问的数据应用程序中的数据等字段信息。因此，通过现有的解析方式，即可获取用户当前所访问的数据应用程序中的数据。

假设用户对数据应用程序的第二访问逻辑为：

```
SELECT stat_date AS 日期
, user_type AS 用户类型
, se_lpv_pc_ld_001 AS Pv
, se_uv_pc_ld_001 AS Uv
FROM tbbi.ads_tb_log_ld
WHERE ds = '20151026'
```

通过对上述第二访问逻辑进行解析，即可获取用户当前所访问的数据应用程序中的数据为“tbbi.ads_tb_log_ld”，即 FROM 字段后的信息。

另外，后台数据源中存储有每个数据的产出逻辑，因此，在后台数据源中，可直接查找用户当前所访问的数据应用程序中的数据的产出逻辑。

步骤 S12、所述搜索引擎数据库根据所述产出逻辑，确定所述数据应用程序中的数据的特征信息。

具体的，所述产出逻辑包括所述数据的聚合对象信息、聚合过程中参与运算的指标信息以及指标运算的时间信息。

步骤 S12 的实现方式具体为：所述搜索引擎数据库确定所述数据的聚合对象信息为所述数据的维度特征；所述搜索引擎数据库确定所述数据在聚合过程中参与运算的指标信息为所述数据的指标特征；所述搜索引擎数据库根据所述指标运算的时间信息，确定

所述数据的时间粒度特征。

对用户当前所访问的数据应用程序中的数据的产出逻辑进行解析，获取当前数据的聚合对象信息、聚合过程中参与运算的指标信息以及指标运算的时间信息。

在本实施例中，假设用户当前所访问的数据应用程序中的数据的产出逻辑，如下：

```
5      Select stat_date ,user_type,count(1) se_lpv_pc_ld_001,count(distinct uid)
      se_uv_pc_ld_001
      From tbcdm. dwd_tb_log_ld where ds=' 20160119'
      Group by user_type, stat_date
```

通过对上述产出逻辑进行解析，可获得用户当前所访问的数据应用程序中的数据的聚合对象信息为 stat_date , user_type, 即 Group by 字段后的信息；聚合过程中参与运算的指标信息为 se_lpv_pc_ld_001 , se_uv_pc_ld_001 , 即 count(1) 和 count(distinct uid) 字段后的信息；指标运算的时间信息为 ' 20160119' , 即 where ds 字段后的分数区。

确定当前数据的聚合对象信息为当前数据的维度特征，确定当前数据聚合过程中参与运算的指标信息为当前数据的指标特征，以及，根据当前数据指标运算的时间信息，确定当前数据的时间粒度特征。

另外，还可将上述指标运算的时间信息所代表的时间区间，作为当前数据的时间粒度特征，比如，当前数据指标运算的时间信息，即 where 字段后的分数区为 “ds=' 20160119' ”，则当前数据的时间粒度特征为 1，再如，当前数据指标运算的时间信息，即 where 字段后的分数区为 “ds>=' 20160101' and ds<=' 20160107' ”，则当前数据的时间粒度特征为 7。

步骤 S13、所述搜索引擎数据库存储所述数据应用程序中的数据，以及所述数据的特征信息。

最后，为用户当前所访问的数据应用程序中的数据添加维度特征、指标特征以及时间粒度特征，且将添加特征后的数据存储到搜索引擎数据库中。

在本实施例中，由于用户每访问一次数据应用程序，即可获取一次用户当前所访问的数据应用程序中的数据的维度特征、指标特征以及时间粒度特征，且为用户当前所访问的数据应用程序中的数据添加上述维度特征、指标特征以及时间粒度特征，最后，将添加上述特征后的数据，存储到搜索引擎数据库中。当用户访问尽数据应用程序中的所有数据时，即可将数据应用程序中的所有数据存储到搜索引擎数据库内，则搜索引擎数据

库内的每条数据均有维度特征、指标特征和时间粒度特征。

步骤 S902、所述搜索引擎数据库存储所述报表、所述知识库平台以及所述集群物理表中的数据。

步骤 S902 的可以通过如下步骤 S21-S23 来实现：

- 5 步骤 S21、所述搜索引擎数据库分别获取所述报表、所述知识库平台以及所述集群物理表中的数据。

本实施例可通过 TF-IDF 算法拆分所述报表、所述知识库平台以及所述集群物理表中的每个数据。

- 10 步骤 S22、所述搜索引擎数据库根据预设算法，确定所述报表、所述知识库平台以及所述集群物理表中每个数据的维度特征。

利用 LDA 算法和 TOPIC MODEL 算法对拆分后的数据进行特征提取，并将提取的特征作为对应数据的维度特征。

步骤 S23、所述搜索引擎数据库存储所述报表、所述知识库平台以及所述集群物理表中每个数据，以及所述数据的维度特征。

- 15 为所述报表、所述知识库平台以及所述集群物理表中的每个数据添加维度特征，且将添加维度特征后的数据，存储到搜索引擎数据库中。

- 本实施例中，搜索引擎数据库中存储有数据应用程序中的所有数据，且从数据应用程序存储到搜索引擎数据库中的每个数据关联有维度特征、指标特征和时间粒度特征；另外，搜索引擎数据库中存储有报表、知识库平台以及集群物理表中的所有数据，且从
20 报表、知识库平台以及集群物理表存储到搜索引擎数据库中的每个数据关联有维度特征。

图 11 为本公开实施例九提供的数据处理方法的流程图，如图 11 所示，本实施例提供的数据处理方法可以包括如下步骤：

- 25 步骤 S1001、搜索引擎数据库获取数据应用程序中的第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征。

在本实施例中，步骤 S1001 的实现方式可以包括以下两种：

- 第一种：所述搜索引擎数据库接收语法解析器发送的所述第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征，所述语法解析器用于采集所述数据应用程序中的第一数据，以及解析所述第一数据的维度特征、指标特征、时间粒度特征。
30

具体的,数据应用程序 15 中的数据可通过语法解析器 19 存储在搜索引擎数据库 13,具体的,通过 SDK 将数据应用程序 15 中的数据采集到语法解析器 19 中。语法解析器 19 可解析出一段结构化查询语言 (Structured Query Language, 简称 SQL) 的维度特征、指标特征、时间粒度特征和读取的表名,例如一段 SQL 具体如下:

```
5      SELECT stat_date AS 日期
      , user_type AS 用户类型
      , se_lpv_pc_ld_001 AS Pv
      , se_uv_pc_ld_001 AS Uv
      FROM tbbi.ads_tb_log_ld
10     WHERE ds = '20151026'
```

语法解析器 19 可解析出该段 SQL 的维度特征是“用户类型”,指标特征是“Pv、Uv”,时间粒度特征是“最近一天”,读取的表名是“tbbi.ads_tb_log_ld”。通过前述方法,语法解析器 19 可解析出数据应用程序 15 中每个数据的维度特征、指标特征和时间粒度特征。语法解析器 19 将解析后的数据发送给搜索引擎数据库 13,搜索引擎数据库 13 中不仅存储有数据本身,同时还存储有数据的维度特征、指标特征和时间粒度特征。

第二种包括如下步骤 S31-S32:

步骤 S31、所述搜索引擎数据库获取所述数据应用程序访问数据源的访问逻辑,所述访问逻辑包括所述数据应用程序中的第一数据,所述数据源存储有所述第一数据的产出逻辑。

在本实施例中,所述数据应用程序可具体为 Web 页面形式,需频繁与后台数据源进行交互;所述后台数据源可具体为存储所述数据应用程序操作数据的器件。由于数据应用程序是根据 SDK 所开发的,SDK 对数据应用程序具有最大的操作权限,因此可通过 SDK 捕获数据应用程序对后台数据源的第一访问逻辑,所述第一访问逻辑中包括数据应用程序访问后台数据源的时间、用户对数据应用程序的第二访问逻辑等字段。因此,通过现有的解析方式,即可获取用户对数据应用程序的第二访问逻辑。用户对数据应用程序的第二访问逻辑中包括用户访问数据应用程序的时间、用户当前所访问的数据应用程序中的数据等字段信息。因此,通过现有的解析方式,即可获取用户当前所访问的数据应用程序中的数据。

假设用户对数据应用程序的第二访问逻辑为:

```

SELECT stat_date AS 日期
, user_type AS 用户类型
, se_lpv_pc_ld_001 AS Pv
, se_uv_pc_ld_001 AS Uv

```

```

5 FROM tbbi.ads_tb_log_ld
WHERE ds = '20151026'

```

通过对上述第二访问逻辑进行解析,即可获取用户当前所访问的数据应用程序中的数据为“tbbi.ads_tb_log_ld”,即 FROM 字段后的信息。

另外,后台数据源中存储有每个数据的产出逻辑,因此,在后台数据源中,可直接
10 查找用户当前所访问的数据应用程序中的数据的产出逻辑。

所述产出逻辑包括所述第一数据的聚合对象信息、聚合过程中参与运算的指标信息以及指标运算的时间信息。具体的,所述搜索引擎数据库确定所述第一数据的聚合对象信息为所述第一数据的维度特征;所述搜索引擎数据库确定所述第一数据在聚合过程中参与运算的指标信息为所述第一数据的指标特征;所述搜索引擎数据库根据所述指标运
15 算的时间信息,确定所述第一数据的时间粒度特征。

步骤 S32、所述搜索引擎数据库根据所述产出逻辑,确定所述数据应用程序中的第一数据的特征信息,所述特征信息包括维度特征、指标特征、时间粒度特征。

在本实施例中,假设用户当前所访问的数据应用程序中的数据的产出逻辑,如下:

```

Select stat_date ,user_type,count(1) se_lpv_pc_ld_001,count(distinct uid)
20 se_uv_pc_ld_001
From tbcdm. dwd_tb_log_ld where ds=' 20160119'
Group by user_type, stat_date

```

通过对上述产出逻辑进行解析,可获得用户当前所访问的数据应用程序中的数据的聚合对象信息为 stat_date , user_type, 即 Group by 字段后的信息;聚合过程中参与
25 运算的指标信息为 se_lpv_pc_ld_001 , se_uv_pc_ld_001 , 即 count(1) 和 count(distinct uid) 字段后的信息;指标运算的时间信息为' 20160119' , 即 where ds 字段后的分数区。

确定当前数据的聚合对象信息为当前数据的维度特征,确定当前数据聚合过程中参与运算的指标信息为当前数据的指标特征,以及,根据当前数据指标运算的时间信息,
30 确定当前数据的时间粒度特征。

另外，还可将上述指标运算的时间信息所代表的时间区间，作为当前数据的时间粒度特征，比如，当前数据指标运算的时间信息，即 where 字段后的分数区为“ds=' 20160119' ”，则当前数据的时间粒度特征为 1，再如，当前数据指标运算的时间信息，即 where 字段后的分数区为“ds>=' 20160101' and ds<=' 20160107' ”，
5 则当前数据的时间粒度特征为 7。

步骤 S1002、所述搜索引擎数据库分别获取报表、知识库平台、集群物理表中的第二数据，以及所述第二数据的维度特征。

具体的，所述搜索引擎数据库分别获取所述报表、所述知识库平台以及所述集群物理表中的第二数据；所述搜索引擎数据库根据预设算法，确定所述报表、所述知识库平台以及所述集群物理表中每个第二数据的维度特征。
10

本实施例可通过 TF-IDF 算法拆分所述报表、所述知识库平台以及所述集群物理表中的每个数据。利用 LDA 算法和 TOPIC MODEL 算法对拆分后的数据进行特征提取，并将提取的特征作为对应数据的维度特征。

步骤 S1003、所述搜索引擎数据库存储所述第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征。
15

步骤 S1004、所述搜索引擎数据库存储所述第二数据，以及所述第二数据的维度特征。

为所述报表、所述知识库平台以及所述集群物理表中的每个数据添加维度特征，且将添加维度特征后的数据，存储到搜索引擎数据库中。

20 本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入
25 一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

图 12 为本公开实施例一提供的查询终端的结构示意图，如图 12 所示，该查询终端
30 包括：接收单元、处理单元、以及发送单元。

所述接收单元，用于接收用户的查询请求，所述查询请求包括检索关键词。

所述处理单元，耦合到所述接收单元，用于获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词。

所述发送单元，耦合到所述处理单元，用于将所述维度关键词、所述指标关键词、
5 以及所述时间粒度关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述
维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的
第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据，所述搜索
引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述数据出口包
括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表，所述特征信息包
10 括下述至少一种：维度特征、指标特征和时间粒度特征。

所述接收单元还用于接收所述搜索引擎数据库发送的所述第一数据、所述第二数据
和所述第三数据。

所述处理单元还用于根据所述第一数据、所述第二数据和所述第三数据，确定反馈
给所述用户的目标数据。

15 本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的
数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒
度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词
进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜
索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数
20 据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入
一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数
据，从而提高了查找数据的效率。

在图 12 所示实施例的基础上，所述处理单元具体用于对所述检索关键词进行分词
25 处理获得多个目标分词；根据各目标分词查询预设的映射表，所述映射表包括维度分词、
指标分词和时间粒度分词；将所述多个目标分词中与所述维度分词匹配的目标分词确定
为所述维度关键词；将所述多个目标分词中与所述指标分词匹配的目标分词确定为所述
指标关键词；将所述多个目标分词中与所述时间粒度分词匹配的目标分词确定为所述时
间粒度关键词。

30 进一步的，所述处理单元具体用于确定所述第一数据、所述第二数据和所述第三数

据是否为同一数据；若所述第一数据、所述第二数据和所述第三数据是同一数据，则所述处理单元将所述同一数据确定为反馈给所述用户的目标数据；若所述第一数据、所述第二数据和所述第三数据不是同一数据，则所述处理单元对所述第一数据、所述第二数据和所述第三数据进行排序，将排序后的数据确定为反馈给所述用户的目标数据。

- 5 本实施例中，通过对检索关键词进行分词处理获得多个目标分词，根据预先建立的映射表查询该多个目标分词中的维度关键词、指标关键词以及时间粒度关键词，提高了确定检索关键词中维度关键词、指标关键词以及时间粒度关键词的效率。

图 13 为本公开实施例二提供的查询终端的结构示意图，如图 13 所示，查询终端还
10 包括：显示器。

所述接收单元还用于接收所述用户对所述目标数据的点击操作。

所述处理单元还用于根据所述点击操作建立所述用户与所述目标数据的关联关系。

所述显示器，耦合到所述处理单元，当用户未输入所述检索关键词时，所述显示器显示所述关联关系关联的所述目标数据。

- 15 本实施例中，通过建立用户与其点击过的目标数据之间的关联关系，当用户未输入检索关键词时，可根据用户与目标数据之间的关联关系，显示用户点击过的目标数据，提高了用户查询数据的便捷性。

图 14 为本公开实施例三提供的查询终端的结构示意图，参照图 14，查询终端 1900
20 包括处理组件 1922，其进一步包括一个或多个处理器，以及由存储器 1932 所代表的存储器资源，用于存储可由处理组件 1922 的执行的指令，例如应用程序。存储器 1932 中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件 1922 被配置为执行指令，以执行上述步骤 S201-S1004 的方法。

- 25 装置 1900 还可以包括一个电源组件 1926 被配置为执行装置 1900 的电源管理，一个有线或无线网络接口 1950 被配置为将装置 1900 连接到网络，和一个输入输出（I/O）接口 1958。装置 1900 可以操作基于存储在存储器 1932 的操作系统，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 或类似。

图 15 为本公开实施例提供的搜索引擎数据库的结构示意图，如图 15 所示，该搜索引擎数据库包括：接收器、存储器、处理器、以及发送器。
30

所述接收器，用于接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接收用户的查询请求，并从所述查询请求包括的检索关键词中获取的。

所述存储器，用于存储数据出口中的数据，以及所述数据的特征信息，所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征。

所述处理器，耦合到所述接收器和所述存储器，用于获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据。

所述发送器，耦合到所述处理器，用于将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端，以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

本实施例中，通过预先采集数据应用程序、报表、知识库平台以及集群物理表中的数据至搜索引擎数据库内，且为所采集的每一条数据添加维度特征、指标特征和时间粒度特征中的至少一个；当搜索引擎接收到用户输入的检索关键词时，首先对检索关键词进行拆分，获得维度关键词、指标关键词以及时间粒度关键词；然后，在预先建立的搜索引擎数据库中，分别查找与维度关键词、指标关键词以及时间粒度关键词相匹配的数据，并将匹配的数据显示给用户；用户无需遍历每个数据出口进行数据查找，仅需输入一次检索关键词，搜索引擎数据库即可查找出所有数据出口中与该检索关键词相关的数据，从而提高了查找数据的效率。

在图 15 所示实施例基础上，所述处理器具体用于获取所述数据应用程序访问数据源的访问逻辑，所述访问逻辑包括所述数据应用程序中的数据，所述数据源存储有所述数据的产出逻辑；根据所述产出逻辑，确定所述数据应用程序中的数据的特征信息；将所述数据应用程序中的数据，以及所述数据的特征信息存储到所述存储器。

或者，在图 15 所示实施例基础上，所述接收器还用于接收语法解析器发送的数据，以及所述数据的维度特征、指标特征、时间粒度特征，所述语法解析器用于采集所述数据应用程序中的数据，以及解析所述数据的维度特征、指标特征、时间粒度特征；所述处理器还用于将所述数据应用程序中的数据，以及所述数据的维度特征、指标特征、时间粒度特征存储到所述存储器。

或者，在图 15 所示实施例基础上，所述处理器具体用于分别获取所述报表、所述知识库平台以及所述集群物理表中的数据；根据预设算法，确定所述报表、所述知识库平台以及所述集群物理表中每个数据的维度特征；将所述报表、所述知识库平台以及所述集群物理表中每个数据，以及所述数据的维度特征存储到所述存储器。

5 本实施例中，搜索引擎数据库中存储有数据应用程序中的所有数据，且从数据应用程序存储到搜索引擎数据库中的每个数据关联有维度特征、指标特征和时间粒度特征；另外，搜索引擎数据库中存储有报表、知识库平台以及集群物理表中的所有数据，且从报表、知识库平台以及集群物理表存储到搜索引擎数据库中的每个数据关联有维度特征。

10 最后应说明的是：以上各实施例仅用以说明本公开的技术方案，而非对其限制；尽管参照前述各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，或者对其中部分或者全部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范围。

权利要求书

1. 一种数据处理系统，其特征在于，包括：查询终端和搜索引擎数据库；

所述查询终端，用于接收用户的查询请求，所述查询请求包括检索关键词；所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词，并将所述
5 维度关键词、所述指标关键词、以及所述时间粒度关键词发送给所述搜索引擎数据库；

所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征；

所述搜索引擎数据库，用于获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹
10 配的时间粒度特征对应的第三数据，并将所述第一数据、所述第二数据和所述第三数据发送给所述查询终端；

所述查询终端，还用于根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据，并将所述目标数据显示给所述用户。

2. 一种数据处理方法，其特征在于，包括：

15 查询终端接收用户的查询请求，所述查询请求包括检索关键词；

所述查询终端获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词；

所述查询终端将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述维度关键词匹配的维度特征对
20 应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据，所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征；

所述查询终端接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述
25 第三数据；

所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

3. 根据权利要求 2 所述的方法，其特征在于，所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表。

30 4. 根据权利要求 3 所述的方法，其特征在于，所述查询终端获取所述检索关键词

中的维度关键词、指标关键词和时间粒度关键词，包括：

所述查询终端对所述检索关键词进行分词处理获得多个目标分词；

所述查询终端根据各目标分词查询预设的映射表，所述映射表包括维度分词、指标分词和时间粒度分词；

5 所述查询终端将所述多个目标分词中与所述维度分词匹配的目标分词确定为所述维度关键词；

所述查询终端将所述多个目标分词中与所述指标分词匹配的目标分词确定为所述指标关键词；

10 所述查询终端将所述多个目标分词中与所述时间粒度分词匹配的目标分词确定为所述时间粒度关键词。

5. 根据权利要求 4 所述的方法，其特征在于，所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据，包括：

所述查询终端确定所述第一数据、所述第二数据和所述第三数据是否为同一数据；

15 若所述第一数据、所述第二数据和所述第三数据是同一数据，则所述查询终端将所述同一数据确定为反馈给所述用户的目标数据；

若所述第一数据、所述第二数据和所述第三数据不是同一数据，则所述查询终端对所述第一数据、所述第二数据和所述第三数据进行排序，将排序后的数据确定为反馈给所述用户的目标数据。

20 6. 根据权利要求 5 所述的方法，其特征在于，所述查询终端对所述第一数据、所述第二数据和所述第三数据进行排序，包括：

所述查询终端计算所述第一数据、所述第二数据和所述第三数据中每个数据的权重值；

所述查询终端计算所述第一数据、所述第二数据和所述第三数据中每个数据与所述检索关键词的相似度；

25 所述查询终端根据所述每个数据的权重值和相似度，计算所述每个数据的排序值；

所述查询终端根据所述每个数据的排序值，对所述第一数据、所述第二数据和所述第三数据中的每个数据进行排序。

7. 根据权利要求 6 所述的方法，其特征在于，所述查询终端根据所述每个数据的排序值，对所述第一数据、所述第二数据和所述第三数据中的每个数据进行排序，包括：

30 所述查询终端根据所述每个数据的排序值，确定所述第一数据、所述第二数据和所

述第三数据中排序值大于第一阈值的数据；

所述查询终端对所述排序值大于第一阈值的数据，按照所述排序值的大小进行排序。

8. 根据权利要求 2-7 任一项所述的方法，其特征在于，所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据之后，还包括：

所述查询终端接收所述用户对所述目标数据的点击操作；

所述查询终端根据所述点击操作建立所述用户与所述目标数据的关联关系；

当用户未输入所述检索关键词时，所述查询终端根据所述关联关系显示所述目标数据。

9. 根据权利要求 8 所述的方法，其特征在于，所述关联关系包括关联度，所述关联度标识所述用户与所述目标数据的关联程度；

所述查询终端根据所述关联关系显示所述目标数据，包括：

所述查询终端显示关联度大于第二阈值的所述目标数据。

10. 一种数据处理方法，其特征在于，包括：

- 查询终端接收用户的查询请求，所述查询请求包括检索关键词；

所述查询终端至少获取所述检索关键词中的两类关键词；

所述查询终端将至少两类关键词发送给搜索引擎数据库，以使所述搜索引擎数据库获取与所述至少两类关键词分别对应的源数据；

所述查询终端接收所述搜索引擎数据库发送的所述源数据；

- 所述查询终端根据所述源数据，确定反馈给所述用户的目标数据。

11. 一种数据处理方法，其特征在于，包括：

搜索引擎数据库接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接收用户的查询请求，并从所述查询请求包括的检索关键词中获取的；

- 所述搜索引擎数据库预先存储有数据出口中的数据，以及所述数据的特征信息，所述特征信息包括下述至少一种：维度特征、指标特征和时间粒度特征；

所述搜索引擎数据获取与所述维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据；

- 所述搜索引擎数据将所述第一数据、所述第二数据和所述第三数据发送给所述查询

终端，以使所述查询终端根据所述第一数据、所述第二数据和所述第三数据，确定反馈给所述用户的目标数据。

12. 根据权利要求 11 所述的方法，其特征在于，所述数据出口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表。

5 13. 根据权利要求 12 所述的方法，其特征在于，所述搜索引擎数据库接收查询终端发送的维度关键词、指标关键词、以及时间粒度关键词之前，还包括：

所述搜索引擎数据库存储所述数据应用程序、所述报表、所述知识库平台以及所述集群物理表中的数据。

10 14. 根据权利要求 13 所述的方法，其特征在于，所述搜索引擎数据库存储所述数据应用程序中的数据，包括：

所述搜索引擎数据库获取所述数据应用程序访问数据源的访问逻辑，所述访问逻辑包括所述数据应用程序中的数据，所述数据源存储有所述数据的产出逻辑；

所述搜索引擎数据库根据所述产出逻辑，确定所述数据应用程序中的数据的特征信息；

15 所述搜索引擎数据库存储所述数据应用程序中的数据，以及所述数据的特征信息。

15. 根据权利要求 14 所述的方法，其特征在于，所述产出逻辑包括所述数据的聚合对象信息、聚合过程中参与运算的指标信息以及指标运算的时间信息；

所述搜索引擎数据库根据所述产出逻辑，确定所述数据应用程序中的数据的特征信息，包括：

20 所述搜索引擎数据库确定所述数据的聚合对象信息为所述数据的维度特征；

所述搜索引擎数据库确定所述数据在聚合过程中参与运算的指标信息为所述数据的指标特征；

所述搜索引擎数据库根据所述指标运算的时间信息，确定所述数据的时间粒度特征。

25 16. 根据权利要求 13 所述的方法，其特征在于，所述搜索引擎数据库存储所述报表、所述知识库平台以及所述集群物理表中的数据，包括：

所述搜索引擎数据库分别获取所述报表、所述知识库平台以及所述集群物理表中的数据；

30 所述搜索引擎数据库根据预设算法，确定所述报表、所述知识库平台以及所述集群物理表中每个数据的维度特征；

所述搜索引擎数据库存储所述报表、所述知识库平台以及所述集群物理表中每个数据，以及所述数据的维度特征。

17. 一种数据处理方法，其特征在于，包括：

搜索引擎数据库获取数据应用程序中的第一数据，以及所述第一数据的维度特征、

5 指标特征、时间粒度特征；

所述搜索引擎数据库分别获取报表、知识库平台、集群物理表中的第二数据，以及所述第二数据的维度特征；

所述搜索引擎数据库存储所述第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征；

10 所述搜索引擎数据库存储所述第二数据，以及所述第二数据的维度特征。

18. 根据权利要求 17 所述的方法，其特征在于，所述搜索引擎数据库获取数据应用程序中的第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征，包括：

所述搜索引擎数据库接收语法解析器发送的所述第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征，所述语法解析器用于采集所述数据应用程序中的第一数据，以及解析所述第一数据的维度特征、指标特征、时间粒度特征。

15

19. 根据权利要求 17 所述的方法，其特征在于，所述搜索引擎数据库获取数据应用程序中的第一数据，以及所述第一数据的维度特征、指标特征、时间粒度特征，包括：

所述搜索引擎数据库获取所述数据应用程序访问数据源的访问逻辑，所述访问逻辑包括所述数据应用程序中的第一数据，所述数据源存储有所述第一数据的产出逻辑；

20 所述搜索引擎数据库根据所述产出逻辑，确定所述数据应用程序中的第一数据的特征信息，所述特征信息包括维度特征、指标特征、时间粒度特征。

20. 根据权利要求 19 所述的方法，其特征在于，所述产出逻辑包括所述第一数据的聚合对象信息、聚合过程中参与运算的指标信息以及指标运算的时间信息；

所述搜索引擎数据库根据所述产出逻辑，确定所述数据应用程序中的第一数据的特征信息，包括：

25

所述搜索引擎数据库确定所述第一数据的聚合对象信息为所述第一数据的维度特征；

所述搜索引擎数据库确定所述第一数据在聚合过程中参与运算的指标信息为所述第一数据的指标特征；

30 所述搜索引擎数据库根据所述指标运算的时间信息，确定所述第一数据的时间粒度

特征。

21. 根据权利要求 17-20 任一项所述的方法, 其特征在于, 所述搜索引擎数据库分别获取报表、知识库平台、集群物理表中的第二数据, 以及所述第二数据的维度特征, 包括:

5 所述搜索引擎数据库分别获取所述报表、所述知识库平台以及所述集群物理表中的第二数据;

所述搜索引擎数据库根据预设算法, 确定所述报表、所述知识库平台以及所述集群物理表中每个第二数据的维度特征。

22. 一种查询终端, 其特征在于, 包括: 接收单元、处理单元、以及发送单元;

10 所述接收单元, 用于接收用户的查询请求, 所述查询请求包括检索关键词;

所述处理单元, 耦合到所述接收单元, 用于获取所述检索关键词中的维度关键词、指标关键词和时间粒度关键词;

所述发送单元, 耦合到所述处理单元, 用于将所述维度关键词、所述指标关键词、以及所述时间粒度关键词发送给搜索引擎数据库, 以使所述搜索引擎数据库获取与所述
15 维度关键词匹配的维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及与所述时间粒度关键词匹配的时间粒度特征对应的第三数据, 所述搜索引擎数据库预先存储有数据出口中的数据, 以及所述数据的特征信息, 所述数据出口包括下述至少一种: 数据应用程序、报表、知识库平台以及集群物理表, 所述特征信息包括下述至少一种: 维度特征、指标特征和时间粒度特征;

20 所述接收单元还用于接收所述搜索引擎数据库发送的所述第一数据、所述第二数据和所述第三数据;

所述处理单元还用于根据所述第一数据、所述第二数据和所述第三数据, 确定反馈给所述用户的目标数据。

23. 根据权利要求 22 所述的查询终端, 其特征在于, 所述处理单元具体用于对所述
25 检索关键词进行分词处理获得多个目标分词; 根据各目标分词查询预设的映射表, 所述映射表包括维度分词、指标分词和时间粒度分词; 将所述多个目标分词中与所述维度分词匹配的目标分词确定为所述维度关键词; 将所述多个目标分词中与所述指标分词匹配的目标分词确定为所述指标关键词; 将所述多个目标分词中与所述时间粒度分词匹配的目标分词确定为所述时间粒度关键词。

30 24. 根据权利要求 23 所述的查询终端, 其特征在于, 所述处理单元具体用于确定

所述第一数据、所述第二数据和所述第三数据是否为同一数据；

若所述第一数据、所述第二数据和所述第三数据是同一数据，则所述处理单元将所述同一数据确定为反馈给所述用户的目标数据；

若所述第一数据、所述第二数据和所述第三数据不是同一数据，则所述处理单元对
5 所述第一数据、所述第二数据和所述第三数据进行排序，将排序后的数据确定为反馈给
所述用户的目标数据。

25. 根据权利要求 24 所述的查询终端，其特征在于，还包括：显示器；

所述接收单元还用于接收所述用户对所述目标数据的点击操作；

所述处理单元还用于根据所述点击操作建立所述用户与所述目标数据的关联关系；

10 所述显示器，耦合到所述处理单元，当用户未输入所述检索关键词时，所述显示器
显示所述关联关系关联的所述目标数据。

26. 一种搜索引擎数据库，其特征在于，包括：接收器、存储器、处理器、以及发
送器；

所述接收器，用于接收查询终端发送的维度关键词、指标关键词、以及时间粒度关
15 键词，所述维度关键词、所述指标关键词、以及所述时间粒度关键词是所述查询终端接
收用户的查询请求，并从所述查询请求包括的检索关键词中获取的；

所述存储器，用于存储数据出口中的数据，以及所述数据的特征信息，所述数据出
口包括下述至少一种：数据应用程序、报表、知识库平台以及集群物理表，所述特征信
息包括下述至少一种：维度特征、指标特征和时间粒度特征；

20 所述处理器，耦合到所述接收器和所述存储器，用于获取与所述维度关键词匹配的
维度特征对应的第一数据、与所述指标关键词匹配的指标特征对应的第二数据、以及
与所述时间粒度关键词匹配的时间粒度特征对应的第三数据；

所述发送器，耦合到所述处理器，用于将所述第一数据、所述第二数据和所述第三
数据发送给所述查询终端，以使所述查询终端根据所述第一数据、所述第二数据和所述
25 第三数据，确定反馈给所述用户的目标数据。

27. 根据权利要求 26 所述的搜索引擎数据库，其特征在于，所述处理器具体用于
获取所述数据应用程序访问数据源的访问逻辑，所述访问逻辑包括所述数据应用程序中
的数据，所述数据源存储有所述数据的产出逻辑；根据所述产出逻辑，确定所述数据应
用程序中的数据的特征信息；将所述数据应用程序中的数据，以及所述数据的特征信息
30 存储到所述存储器。

28. 根据权利要求 26 所述的搜索引擎数据库，其特征在于，所述接收器还用于接收语法解析器发送的数据，以及所述数据的维度特征、指标特征、时间粒度特征，所述语法解析器用于采集所述数据应用程序中的数据，以及解析所述数据的维度特征、指标特征、时间粒度特征；

5 所述处理器还用于将所述数据应用程序中的数据，以及所述数据的维度特征、指标特征、时间粒度特征存储到所述存储器。

29. 根据权利要求 26 所述的搜索引擎数据库，其特征在于，所述处理器具体用于分别获取所述报表、所述知识库平台以及所述集群物理表中的数据；根据预设算法，确定所述报表、所述知识库平台以及所述集群物理表中每个数据的维度特征；将所述报表、
10 所述知识库平台以及所述集群物理表中每个数据，以及所述数据的维度特征存储到所述存储器。

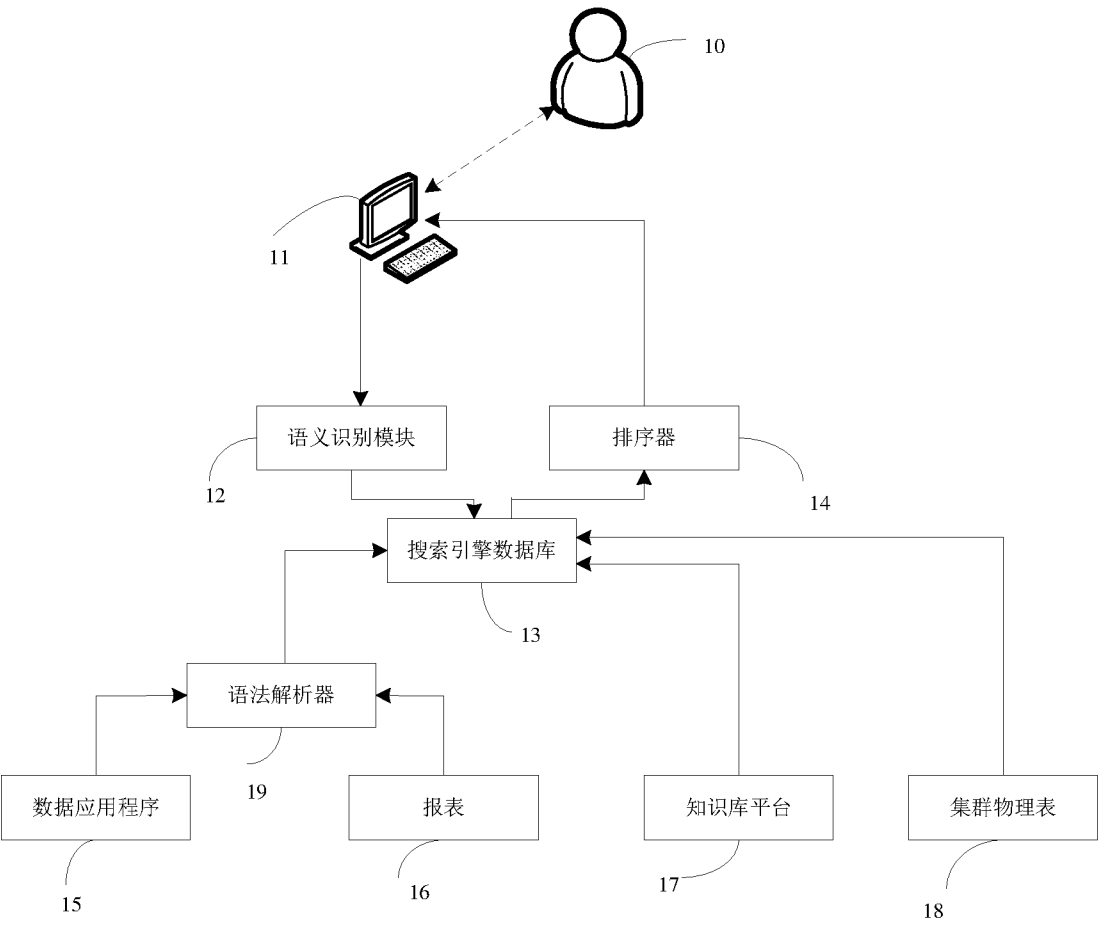


图 1

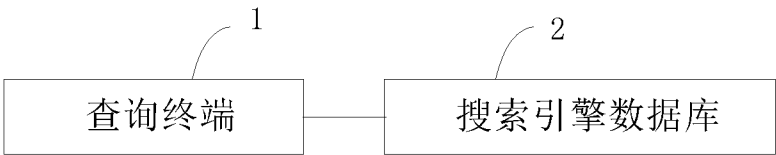


图 2

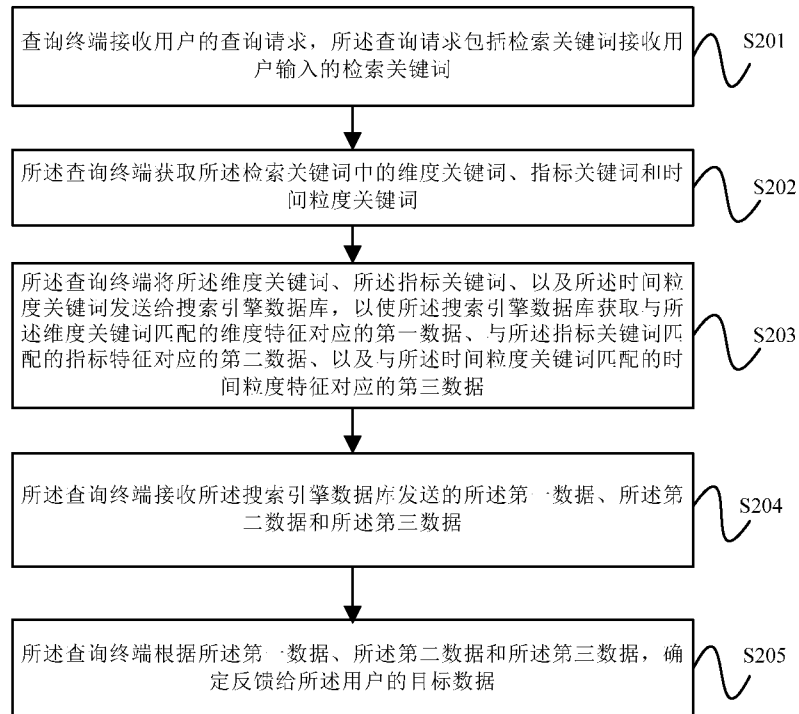


图 3

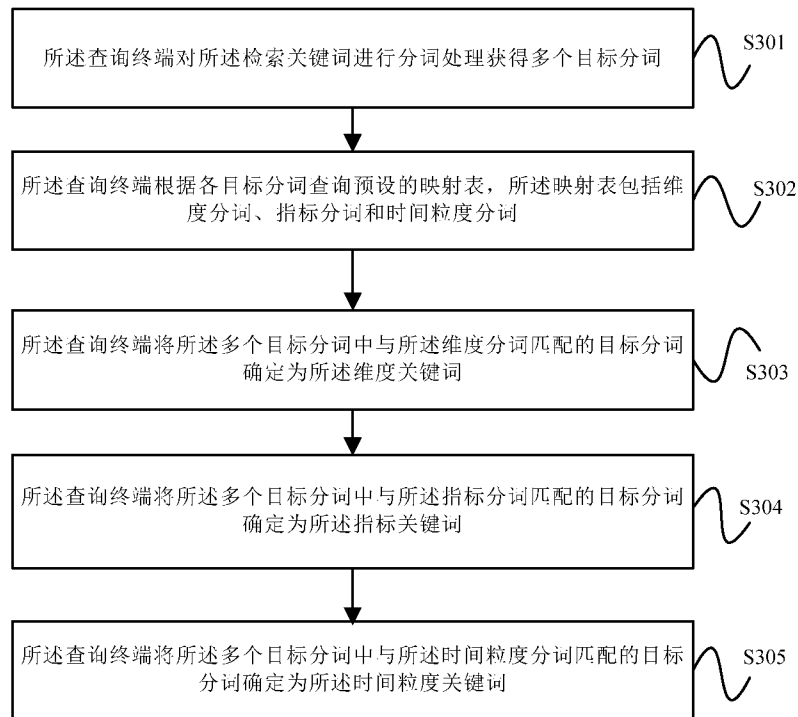


图 4

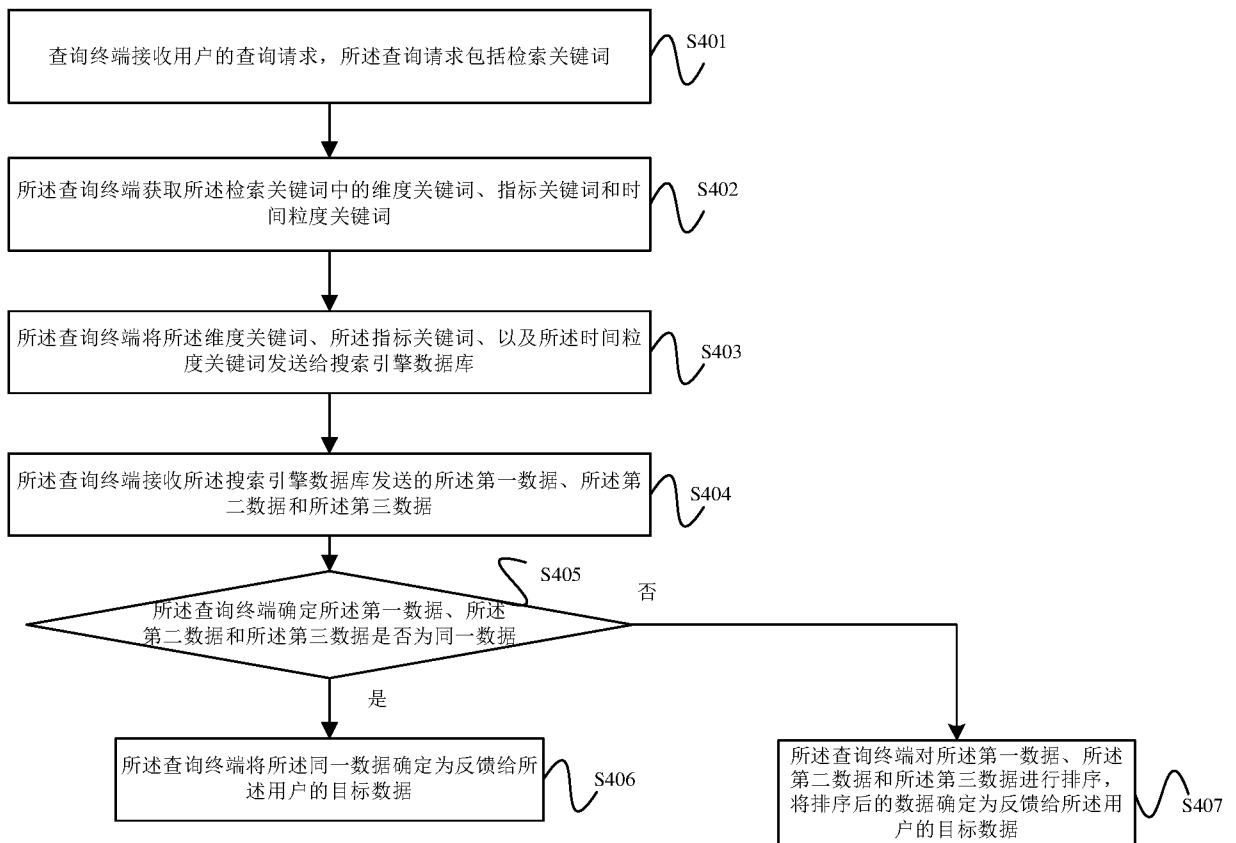


图 5

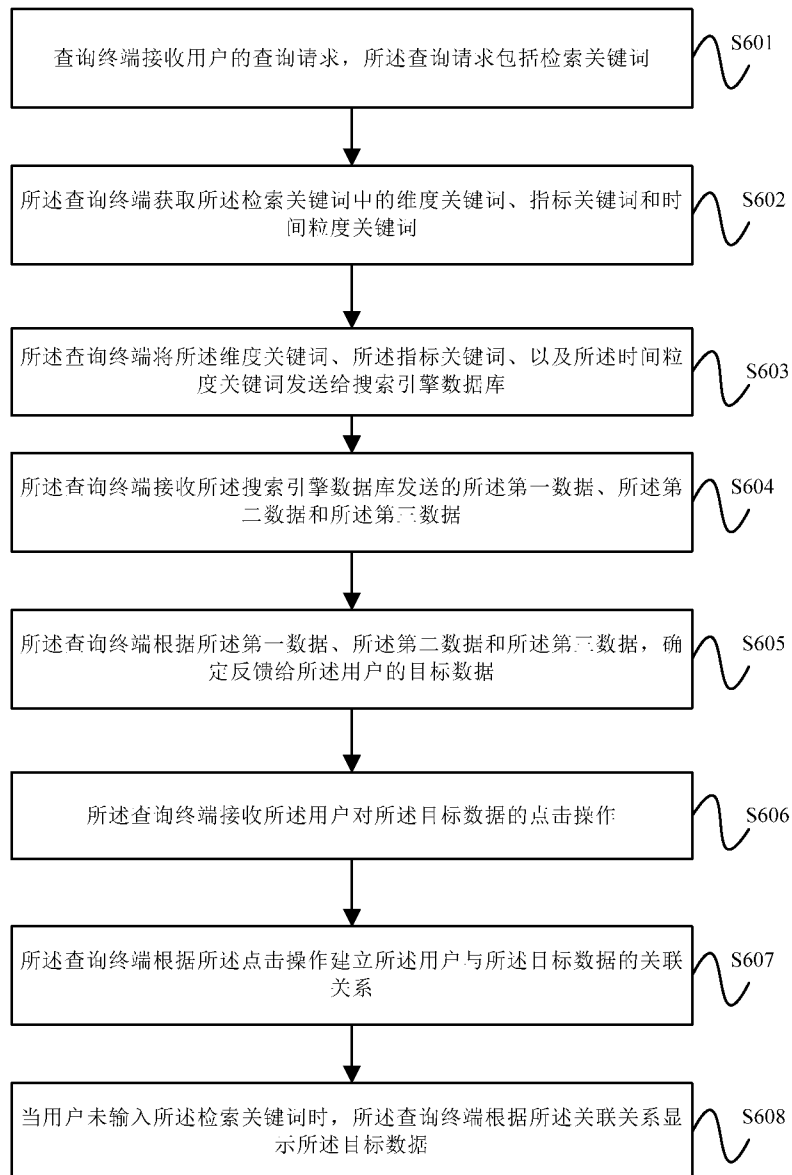


图 6

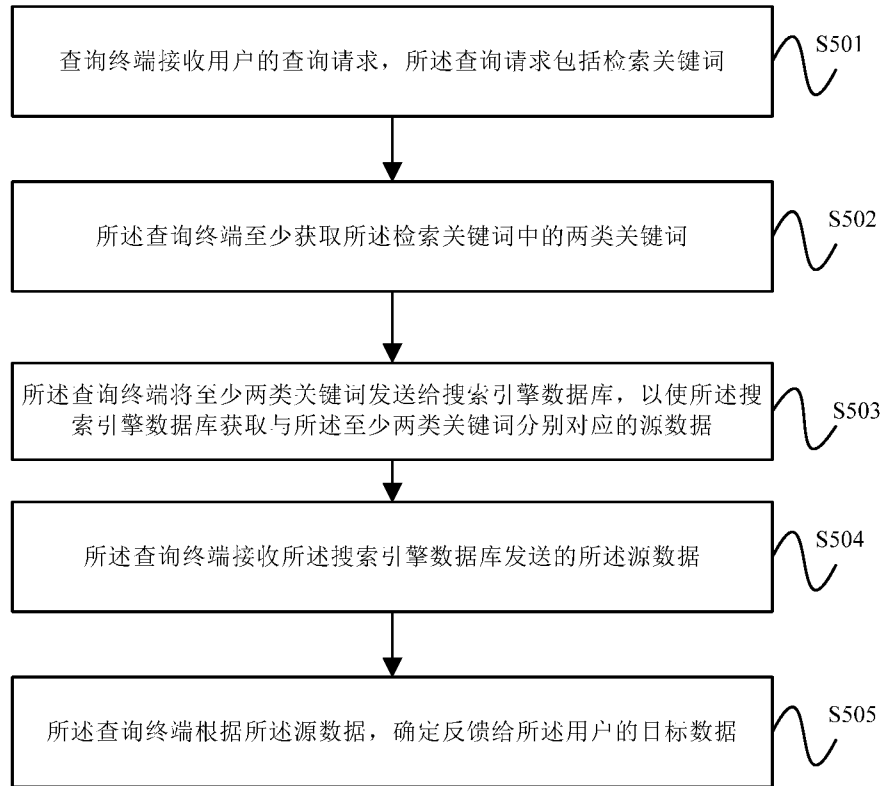


图 7

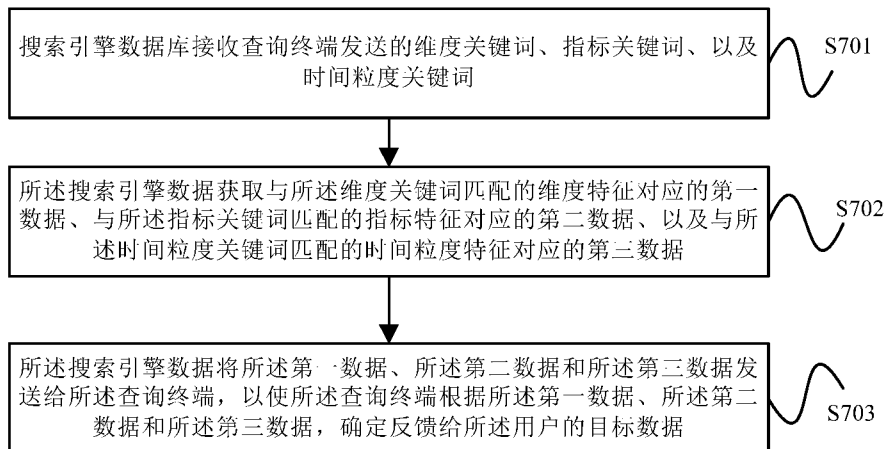


图 8

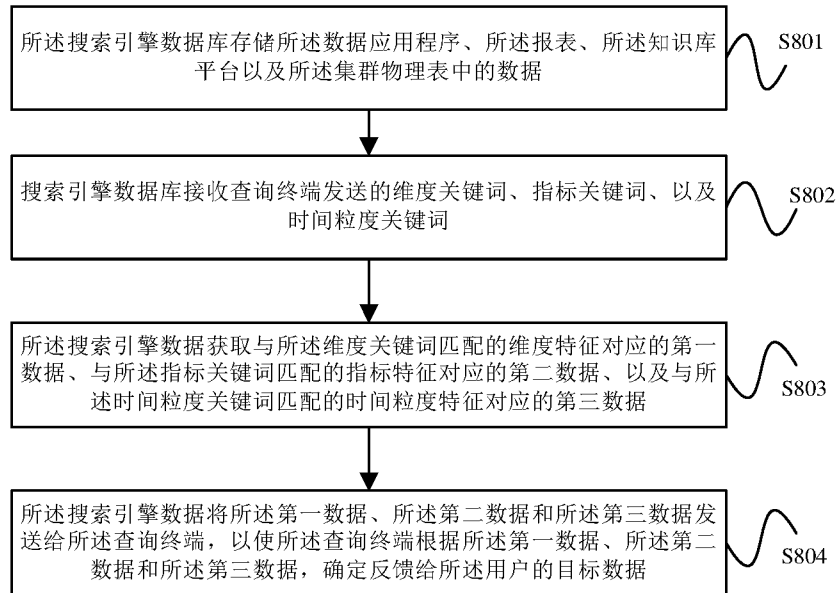


图 9

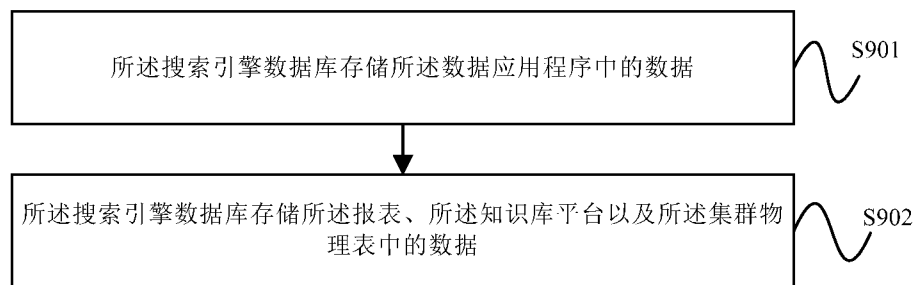


图 10

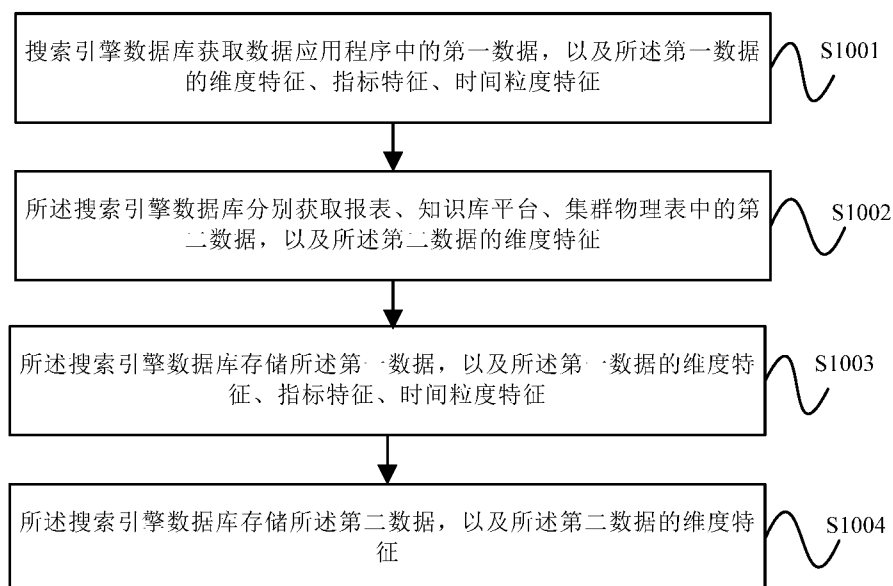


图 11

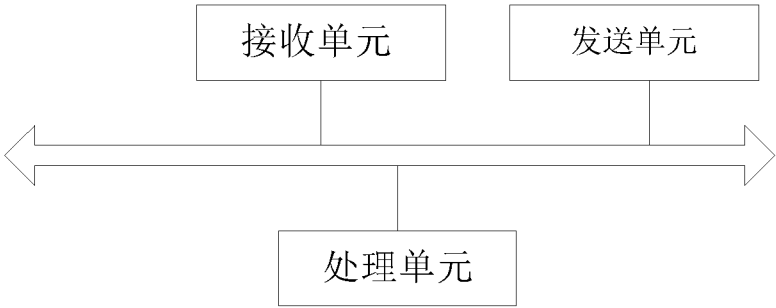


图 12

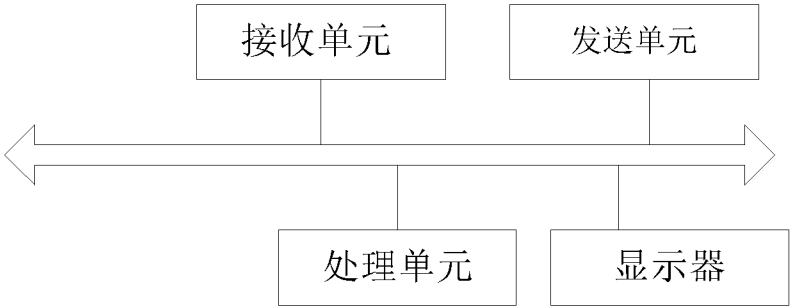


图 13

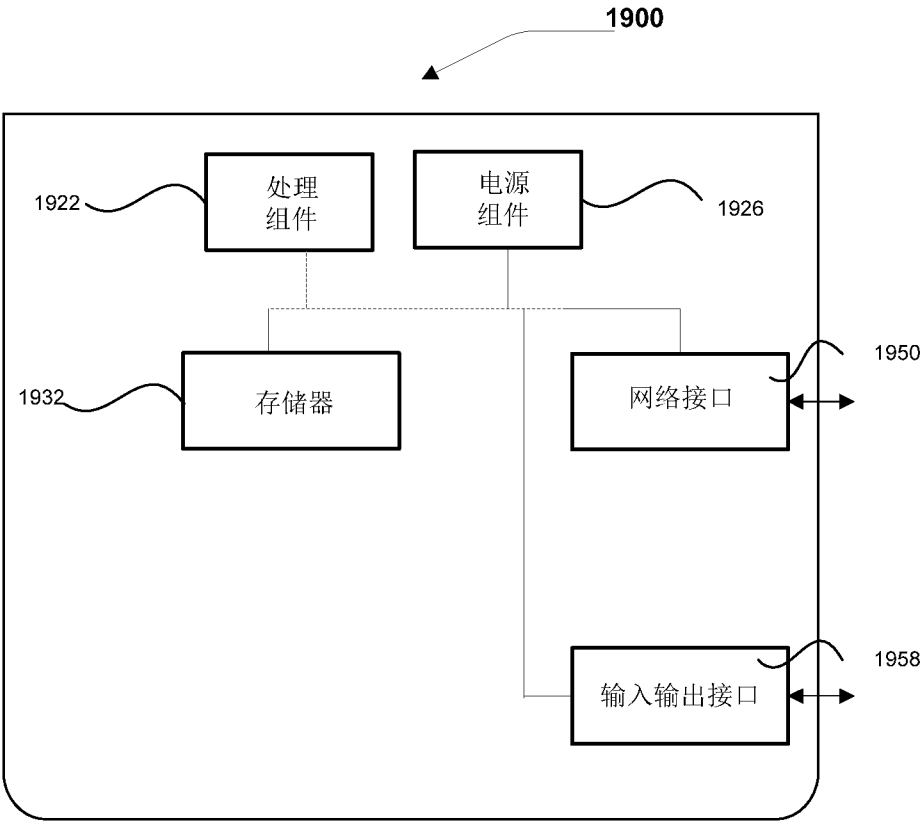


图 14

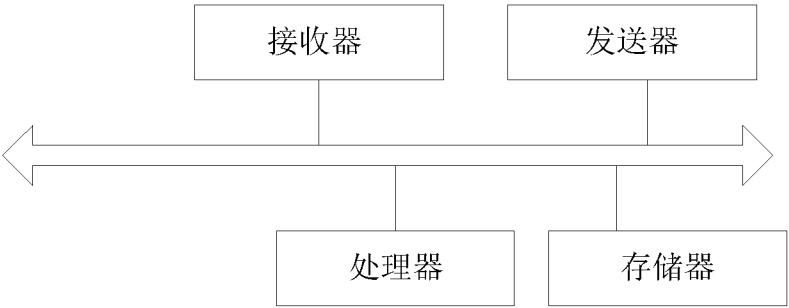


图 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/094790

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) 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)

CNXTX; CNABS; DWPI; CNKI: 数据库, 查询, 检索, 请求, 关键字, 关键词, 分词, 切分, 分割; database, search, retrieve, enquire, request, query, keyword, word, term, partition, split, segment, separate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2014127500 A1 (GOOGLE INC.), 28 August 2014 (28.08.2014), entire document	1-29
A	US 2008033920 A1 (COLCLASURE et al.), 07 February 2008 (07.02.2008), entire document	1-29
A	CN 102521223 A (DAORIGIN TECHNOLOGIES CO., LTD.), 27 June 2012 (27.06.2012), entire document	1-29
A	KR 20080058634 A (NHN CORP.), 26 June 2008 (26.06.2008), entire document	1-29
A	CN 101620605 A (HUAWEI TECHNOLOGIES CO., LTD.), 06 January 2010 (06.01.2010), entire document	1-29

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 22 September 2017	Date of mailing of the international search report 01 November 2017
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer HAO, Xiaoli Telephone No.: (86-10) 62089281

International application No.
PCT/CN2017/094790

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/094790

A. 主题的分类 G06F 17/30 (2006.01) 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)) CNTXT; CNABS; DWPI; CNKI; 数据库, 查询, 检索, 请求, 关键字, 关键词, 分词, 切分, 分割; database, search, retrieve, enquire, request, query, keyword, word, term, partition, split, segment, separate																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>WO 2014127500 A1 (GOOGLE INC) 2014年 8月 28日 (2014 - 08 - 28) 全文</td> <td>1-29</td> </tr> <tr> <td>A</td> <td>US 2008033920 A1 (COLCLASURE 等) 2008年 2月 7日 (2008 - 02 - 07) 全文</td> <td>1-29</td> </tr> <tr> <td>A</td> <td>CN 102521223 A (天津市道本科技有限公司) 2012年 6月 27日 (2012 - 06 - 27) 全文</td> <td>1-29</td> </tr> <tr> <td>A</td> <td>KR 20080058634 A (NHN CORP) 2008年 6月 26日 (2008 - 06 - 26) 全文</td> <td>1-29</td> </tr> <tr> <td>A</td> <td>CN 101620605 A (华为技术有限公司) 2010年 1月 6日 (2010 - 01 - 06) 全文</td> <td>1-29</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	WO 2014127500 A1 (GOOGLE INC) 2014年 8月 28日 (2014 - 08 - 28) 全文	1-29	A	US 2008033920 A1 (COLCLASURE 等) 2008年 2月 7日 (2008 - 02 - 07) 全文	1-29	A	CN 102521223 A (天津市道本科技有限公司) 2012年 6月 27日 (2012 - 06 - 27) 全文	1-29	A	KR 20080058634 A (NHN CORP) 2008年 6月 26日 (2008 - 06 - 26) 全文	1-29	A	CN 101620605 A (华为技术有限公司) 2010年 1月 6日 (2010 - 01 - 06) 全文	1-29
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
A	WO 2014127500 A1 (GOOGLE INC) 2014年 8月 28日 (2014 - 08 - 28) 全文	1-29																		
A	US 2008033920 A1 (COLCLASURE 等) 2008年 2月 7日 (2008 - 02 - 07) 全文	1-29																		
A	CN 102521223 A (天津市道本科技有限公司) 2012年 6月 27日 (2012 - 06 - 27) 全文	1-29																		
A	KR 20080058634 A (NHN CORP) 2008年 6月 26日 (2008 - 06 - 26) 全文	1-29																		
A	CN 101620605 A (华为技术有限公司) 2010年 1月 6日 (2010 - 01 - 06) 全文	1-29																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td style="vertical-align: top;"> * 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 </td> <td style="vertical-align: top;"> “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 </td> </tr> </table>			* 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																
* 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	国际检索报告邮寄日期																			
2017年 9月 22日	2017年 11月 1日																			
ISA/CN的名称和邮寄地址	授权官员																			
中华人民共和国国家知识产权局 (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	郝晓丽 电话号码 (86-10) 62089281																			

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/094790

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)
WO	2014127500	A1	2014年 8月 28日	CN	105900081	A	2016年 8月 24日
				EP	2959405	A1	2015年 12月 30日
US	2008033920	A1	2008年 2月 7日	US	7536383	B2	2009年 5月 19日
CN	102521223	A	2010年 10月 6日	无			
KR	20080058634	A	2008年 6月 26日	WO	2008078884	A1	2008年 7月 3日
CN	101620605	A	2010年 1月 6日	WO	2010000211	A1	2010年 1月 7日