

(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 10 月 22 日 (22.10.2020)



(10) 国际公布号
WO 2020/211657 A1

- (51) 国际专利分类号:
G06F 16/22 (2019.01)
- (21) 国际申请号: PCT/CN2020/083092
- (22) 国际申请日: 2020 年 4 月 3 日 (03.04.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910310499.9 2019年4月17日 (17.04.2019) 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: 徐明明 (**XU, Mingming**); 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, 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,

(54) **Title:** METHOD AND DEVICE FOR CREATING DATA TABLE

(54) 发明名称: 一种数据表的创建方法和装置

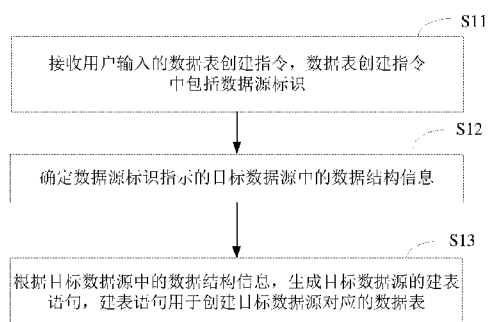


图 1

- S11 Receive a data table creation command input by a user, the data table creation command comprising a data source identifier
- S12 Determine data structure information of a target data source indicated by the data source identifier
- S13 Generate, according to the data structure information of the target data source, table creation statement for the target data source, the table creation statements being used to create a data table corresponding to the target data source

(57) **Abstract:** A method and device for creating a data table. The method comprises: receiving a data table creation command input by a user, the data table creation command comprising a data source identifier (S11); determining data structure information of a target data source indicated by the data source identifier (S12); and generating, according to the data structure information of the target data source, table creation statement for the target data source, the table creation statements being used to create a data table corresponding to the target data source (S13). The method eliminates the need for a user to manually input fields, thereby improving the efficiency of table creation.

(57) **摘要:** 一种数据表的创建方法和装置。该方法包括: 接收用户输入的数据表创建指令, 所述数据表创建指令中包括数据源标识 (S11); 确定所述数据源标识指示的目标数据源中的数据源标识 (S12); 根据所述目标数据源中的数据源标识, 生成所述目标数据源的建表语句, 所述建表语句用于创建所述目标数据源对应的数据表 (S13)。上述方法可以避免用户手动输入字段, 从而可以提高建表效率。

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))。

一种数据表的创建方法和装置

本申请要求 2019 年 04 月 17 日递交的申请号为 201910310499.9、发明名称为“一种数据表的创建方法和装置”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

本公开涉及计算机技术领域，尤其涉及一种数据表的创建方法和装置。

背景技术

10 目前，一般数据库都会提供建表语句，例如：`create table A like B`，利用该建表语句可以根据数据库里面已经有的一个数据表 A 复制出一个同样结构但是不一样表名的新数据表 B。但是，在直接根据数据库中存储的原始数据创建新数据表的情况下，则需要用户手动输入新数据表中的各个字段，导致建表效率较低。

发明内容

有鉴于此，本公开提出了一种数据表的创建方法和装置，使得可以避免用户手动输入字段，从而可以提高建表效率。

20 根据本公开的第一方面，提供了一种数据表的创建方法，包括：接收用户输入的数据表创建指令，所述数据表创建指令中包括数据源标识；确定所述数据源标识指示的目标数据源中的数据结构信息；根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，所述建表语句用于创建所述目标数据源对应的数据表。

在一种可能的实现方式中，确定所述数据源标识指示的目标数据源中的数据结构信息，包括：提取所述目标数据源中存储的至少一条数据记录；根据所述至少一条数据记录，确定所述目标数据源中的数据结构信息。

25 在一种可能的实现方式中，确定所述数据源标识指示的目标数据源中的数据结构信息，包括：确定所述目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为所述数据结构信息。

30 在一种可能的实现方式中，根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，包括：根据所述目标数据源中的各个数据项，确定所述建表语句中的至少一个字段；根据所述任一数据项对应的第一数据类型标识，确定所述数据项对应

的字段的第二数据类型标识，其中，所述第一数据类型标识指示的数据类型和所述第二数据类型标识指示的数据类型相同。

在一种可能的实现方式中，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

- 5 在一种可能的实现方式中，所述目标数据源包括下述至少一种：数据库、具备数据存储功能的存储服务。

在一种可能的实现方式中，所述目标数据源包括下述至少一种：对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MangoDB，和数据库 Redis。

- 10 根据本公开的第二方面，提供了一种数据表的创建装置，包括：接收模块，用于接收用户输入的数据表创建指令，所述数据表创建指令中包括数据源标识；确定模块，用于确定所述数据源标识指示的目标数据源中的数据结构信息；创建模块，用于根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，所述建表语句用于创建所述目标数据源对应的数据表。

- 15 在一种可能的实现方式中，所述确定模块包括：提取子模块，用于提取所述目标数据源中存储的至少一条数据记录；第一确定子模块，用于根据所述至少一条数据记录，确定所述目标数据源中的数据结构信息。

在一种可能的实现方式中，所述确定模块具体用于：确定所述目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为所述数据结构信息。

- 20 在一种可能的实现方式中，所述创建模块包括：第二确定子模块，用于根据所述目标数据源中的各个数据项，确定所述建表语句中的至少一个字段；第三确定子模块，用于根据所述任一数据项对应的第一数据类型标识，确定所述数据项对应的字段的第二数据类型标识，其中，所述第一数据类型标识指示的数据类型和所述第二数据类型标识指示的数据类型相同。

- 25 在一种可能的实现方式中，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

在一种可能的实现方式中，所述目标数据源包括下述至少一种：数据库、具备数据存储功能的存储服务。

- 30 在一种可能的实现方式中，所述目标数据源包括下述至少一种：对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据

库 PolarDB、分布式文档存储数据库 MangoDB，和数据库 Redis。

根据本公开的第三方面，提供了一种数据表的创建装置，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为执行上述第一方面所述的数据表的创建方法。

- 5 根据本公开的第四方面，提供了一种非易失性计算机可读存储介质，其上存储有计算机程序指令，其中，所述计算机程序指令被处理器执行时实现上述第一方面所述的数据表的创建方法。

- 通过接收用户输入的包括数据源标识的数据表创建指令，确定该数据源标识指示的目标数据源中的数据结构信息，进而根据目标数据源中的数据结构信息，自动生成用于
10 创建目标数据源对应的数据表的建表语句，使得可以避免用户手动输入字段，从而可以提高建表效率。

根据下面参考附图对示例性实施例的详细说明，本公开的其它特征及方面将变得清楚。

15 附图说明

包含在说明书中并且构成说明书的一部分的附图与说明书一起示出了本公开的示例性实施例、特征和方面，并且用于解释本公开的原理。

图 1 示出本公开一实施例的数据表的创建方法的流程示意图；

图 2 示出本公开一实施例的数据表的创建方法的示意图；

- 20 图 3 示出本公开一实施例的数据表的创建装置的结构示意图；

图 4 示出本公开一实施例的电子设备的结构示意图。

具体实施方式

- 以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的
25 附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面，但是除非特别指出，不必按比例绘制附图。

在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

另外，为了更好的说明本公开，在下文的具体实施方式中给出了众多的具体细节。

本领域技术人员应当理解，没有某些具体细节，本公开同样可以实施。在一些实例中，对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述，以便于凸显本公开的主旨。

数据湖是一种新型的数据仓库解决方案，数据湖本身不存储数据，而是在用户需要时，从底层数据存储中读取数据。例如，数据湖从企业的多个数据源获取原始数据，并且针对不同的目的，同一份原始数据还可能有多种满足特定内部模型格式的数据副本，因此，数据湖中被处理的数据可能是结构化数据，也可能是非结构化数据。

目前，在数据湖领域，正如背景技术部分陈述的，创建数据源的数据表需要用户手动输入数据表中的各个字段，但是，由于数据源（例如，企业的业务数据库）中包括的字段较多，手动输入字段导致建表效率较低。

本公开提供的数据表的创建方法可以应用于数据湖领域的数据表创建场景，以实现自动生成建表语句，避免用户手动输入字段，提高建表效率。下面以数据湖领域为例详细介绍本公开提供的数据表的创建方法。本领域技术人员应当理解，数据湖领域仅是本公开的一个应用场景示例，并不构成对本公开的限定，本公开提供的数据表的创建方法还可以应用于其它的应用场景。

图 1 示出本公开一实施例的数据表的创建方法的流程示意图。如图 1 所示，该方法可以包括：

步骤 S11，接收用户输入的数据表创建指令，数据表创建指令中包括数据源标识。

步骤 S12，确定数据源标识指示的目标数据源中的数据结构信息。

步骤 S13，根据目标数据源中的数据结构信息，生成目标数据源的建表语句，建表语句用于创建目标数据源对应的数据表。

用户希望针对目标数据源创建数据表时，向数据湖输入包括数据源标识的数据表创建指令，进而数据湖确定数据源标识指示的目标数据源中的数据结构，从而可以根据该数据结构信息，在数据湖中自动生成用于创建目标数据源对应的数据表的建表语句，以避免用户手动输入数据表中的字段，提高建表效率。

在一种可能的实现方式中，目标数据源包括下述至少一种：数据库、具备数据存储功能的存储服务。

目标数据源可以是用于存储结构化数据的数据库，也可以是用于存储非结构化数据的具备数据存储功能的存储服务，还可以是其它数据存储，本公开对此不作具体限定。

在一种可能的实现方式中，目标数据源包括下述至少一种：对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MongoDB，和数据库 Redis。

图 2 示出本公开一实施例的数据表的创建方法的流程示意图。如图 2 所示，底层数据源包括对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MongoDB，和数据库 Redis，数据湖可以为底层数据源创建数据表，从而可以基于数据表对底层数据源中的数据进行查询读取。

在一示例中，数据表创建指令为结构化查询语言（SQL，Structured Query Language）。

在一示例中，如图 2 所示，用户希望针对分布式文档存储数据库 MongoDB（目标数据源）创建数据表时，用户向数据湖中输入数据表创建指令：

```
create external table data_lake_table like mapping ('mongodb_collection')。
```

其中，'mongodb_collection' 为数据源标识，用于指示目标数据源 MongoDB。

在一示例中，用户希望针对对象存储服务 OSS（目标数据源）创建数据表时，用户向数据湖中输入数据表创建指令：

```
create external table data_lake_table like mapping ( 'oss :  
//test-bucket/my-file ' ) 。
```

其中，'oss: //test-bucket/my-file' 为数据源标识，用于指示目标数据源 OSS。

在一示例中，用户希望针对数据库 MySQL（目标数据源）创建数据表时，用户向数据湖中输入数据表创建指令：

```
create external table data_lake_table like mapping ('mysql_table ' )。
```

其中，'mysql_table' 为数据源标识，用于指示目标数据源 MySQL。

数据表创建指令的语法格式可以根据实际情况进行改变，本公开对此不作具体限定。

例如，用户希望针对分布式文档存储数据库 MongoDB 创建数据表时，数据表创建指令除了可以为上述记载之外，还可以为：create external table data_lake_table like mapping mongodb_collection；或 create external table data_lake_table like underlying mongodb_collection。

在一种可能的实现方式中，确定数据源标识指示的目标数据源中的数据结构信息，包括：确定目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为数据结构信息。

数据湖接收到用户输入的包括数据源标识的数据表创建指令之后，数据湖的引擎可以自动对数据源标识指示的目标数据源中的数据结构进行分析推导，确定目标数据源中的数据结构信息：各个数据项，以及任一数据项对应的第一数据类型标识。

在一种可能的实现方式中，确定数据源标识指示的目标数据源中的数据结构信息，包括：提取目标数据源中存储的至少一条数据记录；根据该至少一条数据记录，确定目标数据源中的数据结构信息。

10 仍以上述图 2 为例，数据湖接收到用户输入的数据表创建指令：`create external table data_lake_table like mapping ('mongodb_collection')` 后，数据湖从分布式文档存储数据库 MongoDB 中提取一条数据记录进行分析推导，确定分布式文档存储数据库 MongoDB 中的数据结构信息。

例如，`{ "_id": "ObjectId("5c134c3f36d9cf6ad7077043")" "id": 1 "name": "james" "age": 10 "create_time": "ISODate("2018-12-14T06:22:54.369Z")" }` 为提取的数据记录，根据该数据记录，数据湖分析推导确定分布式文档存储数据库 MongoDB 中包括四个数据项：年龄 `age`、创建时间 `creat_time`、标识 `id` 和姓名 `name`；其中，年龄 `age` 对应的第一数据类型标识为 `double`，用于指示的数据类型为双精度浮点类型；创建时间 `creat_time` 对应的第一数据类型标识为 `timestamp`，用于指示的数据类型为时间戳
20 字符序列类型；标识 `id` 对应的第一数据类型标识为 `double`，用于指示的数据类型为双精度浮点类型；姓名 `name` 对应的第一数据类型标识为 `varchar`，用于指示的数据类型为字符串类型。

在一种可能的实现方式中，根据目标数据源中的数据结构信息，生成数据源的建表语句，包括：根据目标数据源中的各个数据项，确定建表语句中的至少一个字段；根据
25 任一数据项对应的第一数据类型标识，确定该数据项对应的字段的第二数据类型标识，其中，第一数据类型标识指示的数据类型和第二数据类型标识指示的数据类型相同。

在一种可能的实现方式中，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

在不同的数据源中，指示相同数据类型的第一数据类型标识可能不相同。例如，在

分布式文档存储数据库 MongoDB 中，指示字符串类型的第一数据类型标识为 string，而在数据库 MySQL 中，指示字符串类型的第一数据类型标识为 varchar。数据湖中对指示相同数据类型的不同第一数据类型标识进行统一映射，即根据任一数据项对应的第一数据类型标识，确定该数据项对应的字段的第二数据类型标识时，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

例如，在数据湖中，指示字符串类型的第二数据类型标识为 varchar。创建分布式文档存储数据库 MongoDB 对应的数据表时，针对分布式文档存储数据库 MongoDB 中第一数据类型标识为 string（用于指示字符串类型）的数据项，确定该数据项对应的字段的第二数据类型标识为 varchar；创建数据库 MySQL 对应的数据表时，针对数据库 MySQL 中第一数据类型标识为 varchar（用于指示字符串类型）的数据项，确定该数据项对应的字段的第二数据类型标识仍为 varchar。

仍以上述图 2 为例，数据湖确定的分布式文档存储数据库 MongoDB 中的数据结构信息后，根据该数据结构信息，自动生成用于创建分布式文档存储数据库 MongoDB 对应的数据表的建表语句。

例如，用于创建分布式文档存储数据库 MongoDB 对应的数据表的建表语句为：

```
CREATE EXTERNAL TABLE 'data_lake_table' , //在数据湖中创建数据表；  
'age' double NULL COMMENT, //创建一个字段 age，数据类型为双精度浮点类型；  
'create_time' timestamp NULL COMMENT, //创建一个字段 create_time，数据类型为时间戳字符序列类型；
```

```
'id' double NULL COMMENT, //创建一个字段 id，数据类型为双精度浮点类型；  
'name' varchar NULL COMMENT, //创建一个字段 name，数据类型为字符串类型；  
TBLPROPERTIES ( TABLE_MAPPING = 'mongo_test.mongo_collection')。
```

数据湖基于建表语句，创建目标数据源对应的数据表，进而用户可以在数据湖中基于该数据表，对目标数据源中的数据进行查询。

仍以上述图 2 为例，数据湖自动生成用于创建分布式文档存储数据库 MongoDB 对应的数据表的建表语句之后，基于该建表语句创建分布式文档存储数据库 MongoDB 对应的数据表。当数据湖接收到用户输入的针对分布式文档存储数据库 MongoDB 的数据查询指令时，数据湖可以基于该数据表，对分布式文档存储数据库 MongoDB 中的数据进行查询。

通过接收用户输入的包括数据源标识的数据表创建指令，确定该数据源标识指示的目标数据源中的数据结构信息，进而根据目标数据源中的数据结构信息，自动生成用于创建目标数据源对应的数据表的建表语句，使得可以避免用户手动输入字段，从而提高建表效率。

5

图3示出本公开一实施例的数据表的创建装置的结构示意图。图3所示的装置30可以用于执行图1所示方法实施例的步骤，装置30包括：

接收模块31，用于接收用户输入的数据表创建指令，数据表创建指令中包括数据源标识；

10 确定模块32，用于确定数据源标识指示的目标数据源中的数据结构信息；

创建模块33，用于根据目标数据源中的数据结构信息，生成目标数据源的建表语句，建表语句用于创建目标数据源对应的数据表。

在一种可能的实现方式中，确定模块32包括：

提取子模块，用于提取目标数据源中存储的至少一条数据记录；

15 第一确定子模块，用于根据该至少一条数据记录，确定目标数据源中的数据结构信息。

在一种可能的实现方式中，确定模块32具体用于：

确定目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为数据结构信息。

20 在一种可能的实现方式中，创建模块33包括：

第二确定子模块，用于根据目标数据源中的各个数据项，确定建表语句中的至少一个字段；

第三确定子模块，用于根据任一数据项对应的第一数据类型标识，确定该数据项对应的字段的第二数据类型标识，其中，第一数据类型标识指示的数据类型和第二数据类型标识指示的数据类型相同。

25

在一种可能的实现方式中，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

在一种可能的实现方式中，目标数据源包括下述至少一种：

数据库、具备数据存储功能的存储服务。

在一种可能的实现方式中，目标数据源包括下述至少一种：

对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MangoDB，和数据库 Redis。

本公开提供的装置 30 能够实现图 1 所示方法实施例中的各个步骤，并实现相同的技
5 术效果，为避免重复，这里不再赘述。

图 4 示出本公开一实施例的电子设备的结构示意图。如图 4 所示，在硬件层面，该电子设备包括处理器，可选地还包括内部总线、网络接口、存储器。其中，存储器可能包含内存，例如高速随机存取存储器 (Random-Access Memory, RAM)，也可能还包括非
10 易失性存储器 (non-volatile memory)，例如至少 1 个磁盘存储器等。当然，该电子设备还可能包括其他业务所需要的硬件。

处理器、网络接口和存储器可以通过内部总线相互连接，该内部总线可以是 ISA (Industry Standard Architecture, 工业标准体系结构) 总线、PCI (Peripheral Component Interconnect, 外设部件互连标准) 总线或 EISA (Extended Industry
15 Standard Architecture, 扩展工业标准结构) 总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 4 中仅用一个双向箭头表示，但并不表示仅有一根总线或一种类型的总线。

存储器，存放程序。具体地，程序可以包括程序代码，所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器，并向处理器提供指令和数据。

20 处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成数据表的创建装置。处理器，执行存储器所存放的程序，并具体执行：接收用户输入的数据表创建指令，数据表创建指令中包括数据源标识；确定数据源标识指示的目标数据源中的数据结构信息；根据目标数据源中的数据结构信息，生成目标数据源的建表语句，所述建表语句用于创建所述目标数据源对应的数据表。

25 在一种可能的实现方式中，处理器被具体配置为执行：提取目标数据源中存储的至少一条数据记录；根据该至少一条数据记录，确定目标数据源中的数据结构信息。

在一种可能的实现方式中，处理器被具体配置为执行：确定目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为数据结构信息。

在一种可能的实现方式中，处理器被具体配置为执行：根据目标数据源中的各个数

据项，确定建表语句中的至少一个字段；根据任一数据项对应的第一数据类型标识，确定该数据项对应的字段的第二数据类型标识，其中，第一数据类型标识指示的数据类型和第二数据类型标识指示的数据类型相同。

在一种可能的实现方式中，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

在一种可能的实现方式中，目标数据源包括下述至少一种：数据库、具备数据存储功能的存储服务。

在一种可能的实现方式中，目标数据源包括下述至少一种：对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MongoDB，和数据库 Redis。

处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器（Central Processing Unit，CPU）、网络处理器（Network Processor，NP）等；还可以是数字信号处理器（Digital Signal Processor，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现场可编程门阵列（Field-Programmable Gate Array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本说明书实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本说明书实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

该电子设备可执行图 1 所示方法实施例执行的方法，并实现上述图 1 所示

方法实施例的功能，本说明书实施例在此不再赘述。

本说明书实施例还提出了一种计算机可读存储介质，该计算机可读存储介质存储一个或多个程序，该一个或多个程序包括指令，该指令当被包括多个应用程序的电子设备执行时，能够使该电子设备执行图 1 所示实施例中的数据表的创建方法，并具体执行图 1 所示方法实施例的步骤。

本公开可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于使处理器实现本公开的各个方面的计算机可读程序指令。

计算机可读存储介质可以是保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一—但不限于—电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、静态随机存取存储器（SRAM）、便携式压缩盘只读存储器（CD-ROM）、数字多功能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的计算机可读存储介质中。

用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码，所述编程语言包括面向对象的编程语言—诸如 Smalltalk、C++等，以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络—包括局域网（LAN）或广域网（WAN）—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列（FPGA）或可编程逻辑阵列（PLA），该电子电路可以执行计算机

可读程序指令，从而实现本公开的各个方面。

这里参照根据本公开实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

- 5 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，
- 10 这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

- 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上
- 15 执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

- 附图中的流程图和框图显示了根据本公开的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注
- 20 的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意
- 的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

- 25 以上已经描述了本公开的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术的技术改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

权 利 要 求 书

1、一种数据表的创建方法，其特征在于，包括：

接收用户输入的数据表创建指令，所述数据表创建指令中包括数据源标识；

确定所述数据源标识指示的目标数据源中的数据结构信息；

- 5 根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，所述建表语句用于创建所述目标数据源对应的数据表。

2、根据权利要求 1 所述的方法，其特征在于，确定所述数据源标识指示的目标数据源中的数据结构信息，包括：

- 10 提取所述目标数据源中存储的至少一条数据记录；

根据所述至少一条数据记录，确定所述目标数据源中的数据结构信息。

3、根据权利要求 1 或 2 所述的方法，其特征在于，确定所述数据源标识指示的目标数据源中的数据结构信息，包括：

- 15 确定所述目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为所述数据结构信息。

4、根据权利要求 3 所述的方法，其特征在于，根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，包括：

- 20 根据所述目标数据源中的各个数据项，确定所述建表语句中的至少一个字段；

根据所述任一数据项对应的第一数据类型标识，确定所述数据项对应的字段的第二数据类型标识，其中，所述第一数据类型标识指示的数据类型和所述第二数据类型标识指示的数据类型相同。

- 25 5、根据权利要求 4 所述的方法，其特征在于，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

6、根据权利要求 1 所述的方法，其特征在于，所述目标数据源包括下述至少一种：数据库、具备数据存储功能的存储服务。

- 30 7、根据权利要求 1 所述的方法，其特征在于，所述目标数据源包括下述至少一种：

对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MangoDB，和数据库 Redis。

5 8、一种数据表的创建装置，其特征在于，包括：

接收模块，用于接收用户输入的数据表创建指令，所述数据表创建指令中包括数据源标识；

确定模块，用于确定所述数据源标识指示的目标数据源中的数据结构信息；

10 创建模块，用于根据所述目标数据源中的数据结构信息，生成所述目标数据源的建表语句，所述建表语句用于创建所述目标数据源对应的数据表。

9、根据权利要求 8 所述的装置，其特征在于，所述确定模块包括：

提取子模块，用于提取所述目标数据源中存储的至少一条数据记录；

15 第一确定子模块，用于根据所述至少一条数据记录，确定所述目标数据源中的数据结构信息。

10、根据权利要求 8 或 9 所述的装置，其特征在于，所述确定模块具体用于：

确定所述目标数据源中的各个数据项，以及任一数据项对应的第一数据类型标识，作为所述数据结构信息。

20 11、根据权利要求 10 所述的装置，其特征在于，所述创建模块包括：

第二确定子模块，用于根据所述目标数据源中的各个数据项，确定所述建表语句中的至少一个字段；

25 第三确定子模块，用于根据所述任一数据项对应的第一数据类型标识，确定所述数据项对应的字段的第二数据类型标识，其中，所述第一数据类型标识指示的数据类型和所述第二数据类型标识指示的数据类型相同。

12、根据权利要求 11 所述的装置，其特征在于，根据指示相同数据类型的第一数据类型标识所确定的第二数据类型标识是相同的。

30 13、根据权利要求 8 所述的装置，其特征在于，所述目标数据源包括下述至少一种：

数据库、具备数据存储功能的存储服务。

14、根据权利要求 8 所述的装置，其特征在于，所述目标数据源包括下述至少一种：

对象存储服务 OSS、表格存储服务 OTS、数据库 MySQL、数据库 SQL Server、数据库 Postgres、关系型数据库 PolarDB、分布式文档存储数据库 MangoDB，和数据库 Redis。

15、一种数据表的创建装置，其特征在于，包括：

处理器；

10 用于存储处理器可执行指令的存储器；

其中，所述处理器被配置为执行权利要求 1-7 任一项所述的数据表的创建方法。

16、一种非易失性计算机可读存储介质，其上存储有计算机程序指令，其特征在于，所述计算机程序指令被处理器执行时实现权利要求 1-7 任一项所述的数据表的创建方

15 法。

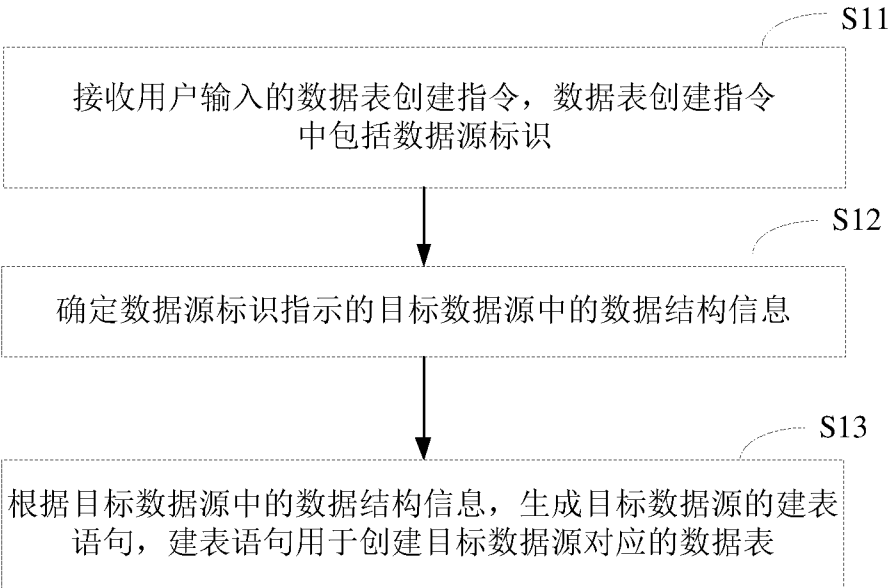


图 1

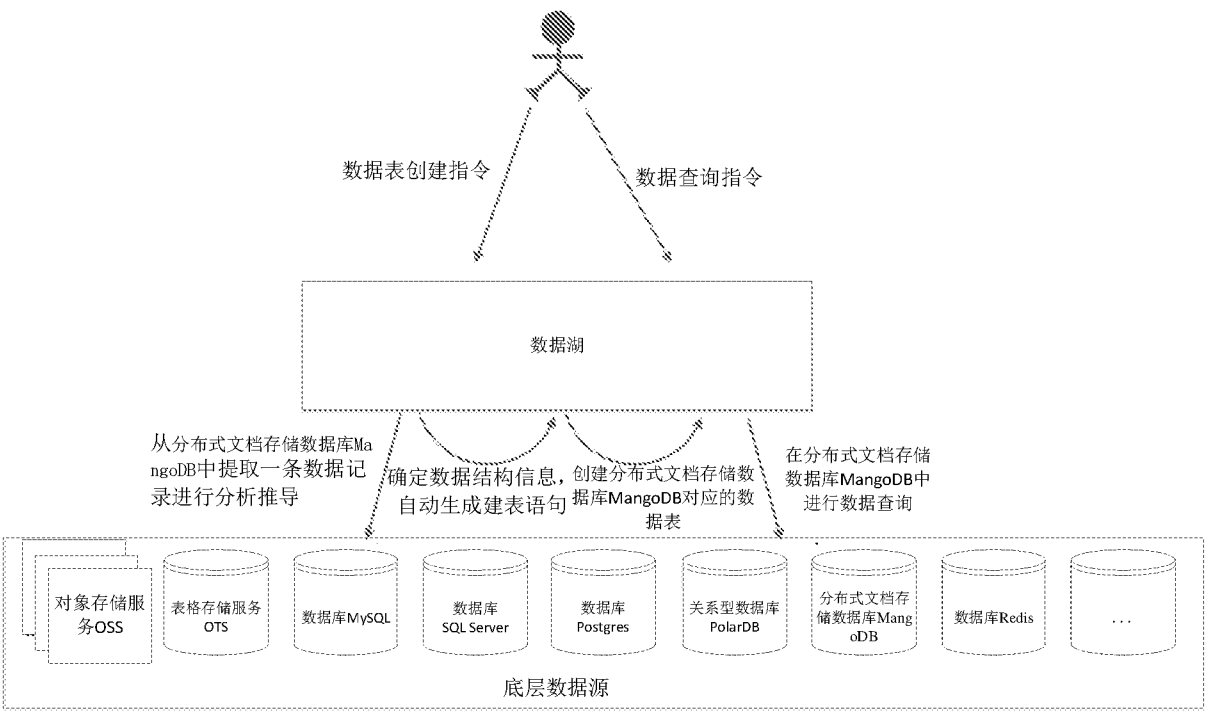


图 2

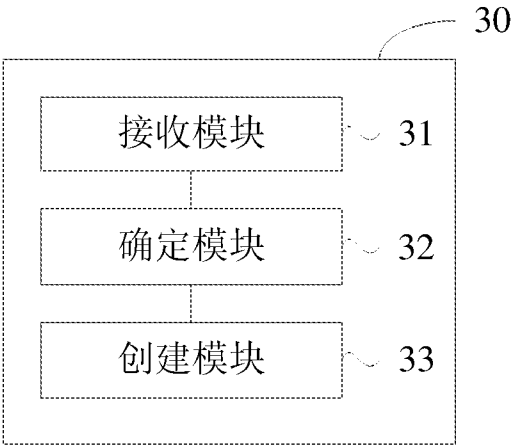


图 3

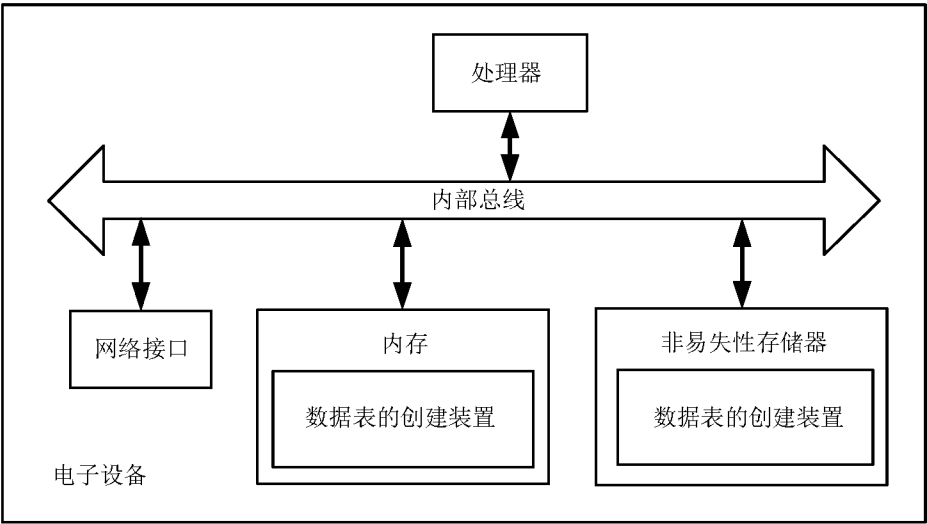


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/083092

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/22(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)

CNPAT, EPODOC, WPI, CNKI, IEEE: 数据, 表, 数据库, 指令, 语句, 创建, 结构, 标识, 字段, data, table, database, SQL, instruction, statement, create, structure, identifier, field

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108170809 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 15 June 2018 (2018-06-15) claims 1-10, description, paragraphs [0002], [0054]-[0085], figures 1-8	1-16
A	CN 101661508 A (KINGDEE SOFTWARE CHINA CO., LTD.) 03 March 2010 (2010-03-03) entire document	1-16
A	CN 104899045 A (PHICOMM (SHANGHAI) CO., LTD.) 09 September 2015 (2015-09-09) entire document	1-16
A	CN 105608088 A (XI'AN AIRCRAFT DESIGN INSTITUTE OF AVIATION INDUSTRY CORPORATION OF CHINA) 25 May 2016 (2016-05-25) entire document	1-16
A	US 2017091234 A1 (ZEKKO INC.) 30 March 2017 (2017-03-30) entire document	1-16

☐ 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

19 June 2020

Date of mailing of the international search report

02 July 2020

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/CN2020/083092

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108170809	A	15 June 2018	WO	2019127837	A1	04 July 2019
CN	101661508	A	03 March 2010	None			
CN	104899045	A	09 September 2015	None			
CN	105608088	A	25 May 2016	None			
US	2017091234	A1	30 March 2017	JP	2017068475	A	06 April 2017

国际检索报告

国际申请号

PCT/CN2020/083092

A. 主题的分类

G06F 16/22 (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))

CNPAT, EPDOC, WPI, CNKI, IEEE: 数据, 表, 数据库, 指令, 语句, 创建, 结构, 标识, 字段, data, table, database, SQL, instruction, statement, create, structure, identifier, field

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
X	CN 108170809 A (平安科技深圳有限公司) 2018年 6月 15日 (2018 - 06 - 15) 权利要求1-10, 说明书第[0002], [0054]-[0085]段, 图1-8	1-16
A	CN 101661508 A (金蝶软件中国有限公司) 2010年 3月 3日 (2010 - 03 - 03) 全文	1-16
A	CN 104899045 A (上海斐讯数据通信技术有限公司) 2015年 9月 9日 (2015 - 09 - 09) 全文	1-16
A	CN 105608088 A (中国航空工业集团公司西安飞机设计研究所) 2016年 5月 25日 (2016 - 05 - 25) 全文	1-16
A	US 2017091234 A1 (ZEKKO INC.) 2017年 3月 30日 (2017 - 03 - 30) 全文	1-16

☐ 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

2020年 6月 19日

国际检索报告邮寄日期

2020年 7月 2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

受权官员

王莹

电话号码 86-(10)-53961411

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/083092

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	108170809	A	2018年 6月 15日	WO	2019127837	A1	2019年 7月 4日
CN	101661508	A	2010年 3月 3日	无			
CN	104899045	A	2015年 9月 9日	无			
CN	105608088	A	2016年 5月 25日	无			
US	2017091234	A1	2017年 3月 30日	JP	2017068475	A	2017年 4月 6日