

# (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  
2017年1月19日 (19.01.2017)



(10) 国际公布号  
WO 2017/008598 A1

- (51) 国际专利分类号:  
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2016/084954
- (22) 国际申请日: 2016年6月6日 (06.06.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201510405965.3 2015年7月10日 (10.07.2015) CN
- (71) 申请人: 中兴通讯股份有限公司 (ZTE CORPORATION) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。
- (72) 发明人: 王东 (WANG, Dong); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。
- (74) 代理人: 北京安信方达知识产权代理有限公司 (AFD CHINA INTELLECTUAL PROPERTY LAW OFFICE); 中国北京市海淀区学清路8号B座1601A, Beijing 100192 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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)。

### 根据细则 4.17 的声明:

- 关于申请人有权申请并被授予专利(细则 4.17(ii))
- 发明人资格(细则 4.17(iv))

### 本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

(54) Title: BIG DATA EXCHANGE METHOD AND DEVICE

(54) 发明名称: 大数据交换方法及装置

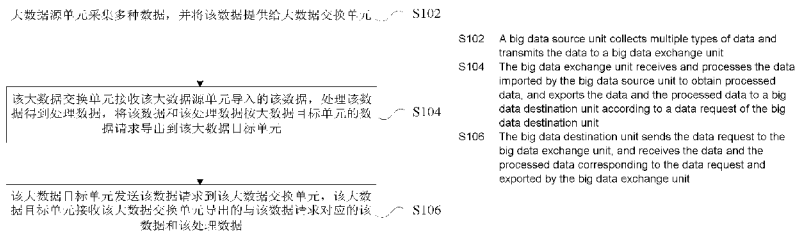


图 1

(57) Abstract: Provided are a big data exchange method and device. The method comprises: (S102) collecting, by a big data source unit, multiple types of data, and provide the data to a big data exchange unit; (104) receiving and processing, by the big data exchange unit, the data imported by the big data source unit to obtain processed data, and exporting the data and the processed data to a big data destination unit according to a data request of the big data destination unit; and (106) sending, by the big data destination unit, the data request to the big data exchange unit, and receiving, by the big data destination unit, the data and the processed data corresponding to the data request and exported by the big data exchange unit.

(57) 摘要: 一种大数据交换方法及装置, 其中, 该方法包括: 大数据源单元采集多种数据, 并将该数据提供给大数据交换单元 (S102); 该大数据交换单元接收该大数据源单元导入的该数据, 处理该数据得到处理数据, 将该数据和该处理数据按大数据目标单元的数据请求导出到该大数据目标单元 (S104); 该大数据目标单元发送该数据请求到该大数据交换单元, 该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该数据和该处理数据 (S106)。



WO 2017/008598 A1

## 大数据交换方法及装置

### 技术领域

5 本发明实施例涉及但不限于通信领域，尤其涉及一种大数据交换方法及装置。

### 背景技术

目前，大数据成为当前业界的技术热点，特别是随着云计算业务的部署，大数据作为其最重要的业务应用被寄予厚望。

10 大数据的主要特点是体量大（Volume）、种类多（Variety）、速度快（Velocity）以及价值高（Value），并同时伴随着多源、异构以及海量的数据特征。

15 大数据所涉及的数据类型包括结构化（Structured）、半结构化（Semi-Structured）以及非结构化（Un-Structured）等类型，这要求大数据交换系统可以针对这些不同数据类型进行区分和处理，比如对于结构化的数据库数据的采集、异常数据清理以及 ETL（数据提取、数据转换以及数据加载）等，对于半结构化数据以及非结构化数据的增量实时采集、解析、注释、异常数据清理、数据分割、特征提取、存储与缓存以及批量非实时处理等。

20 大数据所涉及的数据对象包括文字、语音、视频、图像、网页、传感器数据等，这也就要求大数据交换系统可以有针对性地区分和处理不同数据对象所对应的数据承载格式，比如：文本文件、音频文件、适配文件、图像文件、网页文件、传感器数据文件以及相应的实时或准实时流媒体数据等不同格式。

25 在相关技术中，大数据正在逐步应用于各个行业，但是随着数据体量的逐步增加，所涉及数据种类的日益繁杂，所生成速度的迅速提升以及对于其潜在价值利用的需求增强，当前大数据正遭遇一些新的困境以及面临新的挑战，包括以下几点：

1) 大数据的孤岛性导致大数据无法开放、共享以及互联，从而降低了大

数据用户共享不同产业甚至相同产业不同部门数据的可能性；

2)对于具有巨大潜在价值的大体量的大数据缺乏价值衡量手段和货币化手段，从而削弱了数据拥有者开放数据的积极性；

3)已有公共数据的开放不足阻碍了公共社会资源对于可能的社会生活以及经济活动的贡献性；

4)缺乏有效的商业模式，特别是高效的数据交换机制，阻止了数据拥有者和数据用户之间安全有效地进行数据交换；

5)互连互通标准的缺乏也无法保障大数据提供商和大数据业务提供商之间进行互连。

10 为从大体量的、种类繁多的、高速变化的大数据中充分挖掘出大数据巨大的潜在价值，亟需解决以上所遭遇的各种问题以及相关挑战。而相关技术中，针对大数据无法开放、互联以及共享的问题，还缺乏有效的解决方案。

## 发明内容

15 以下是对本文详细描述的主题的概述。本概述并非是为了限制权利要求的保护范围。

本发明实施例提供了一种大数据交换方法及装置，以至少解决相关技术中大数据无法开放、互联以及共享的问题。

20 本发明实施例提供了一种大数据交换系统，包括：大数据源单元，大数据交换单元，大数据目标单元，其中：

所述大数据源单元，设置为采集多种数据，并将所述数据提供给所述大数据交换单元；

25 所述大数据交换单元，设置为接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到所述大数据目标单元；

所述大数据目标单元，设置为发送所述数据请求到所述大数据交换单元，以及设置为接收所述大数据交换单元导出的与所述数据请求对应的所述数据和所述处理数据。

可选地，所述大数据源单元包括以下至少之一：数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块，其中：

5 所述数据生产模块，设置为生产数据或者将几种数据进行合并为新的数据；

所述数据拥有模块，设置为持有数据并授权进行数据访问；

所述数据中介模块，设置为不直接生产数据，与所述数据生产模块或者所述数据拥有模块通过授权协议收购数据；

10 所述数据代理模块，设置为不直接生产数据，代替所述数据生产模块持有数据；

所述数据摄取适配模块，设置为接收来自于外部数据源的数据并根据所述大数据源单元处理需求，对所摄取的所述外部数据源的数据进行适配；

15 所述数据采集模块，设置为从所述数据生产模块、所述数据代理模块、所述数据拥有模块、所述数据中介模块或者所述数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。

可选地，所述大数据交换单元包括以下至少之一：数据导入模块、数据收集模块、数据准备模块、数据分析模块、数据可视化模块以及数据导出模块，其中：

20 所述数据导入模块，设置为接收所述大数据源单元的数据，并存储和/或缓存所述数据，所述数据导入模块接收数据的方式包括以下之一：数据查询、数据请求、数据抓取、数据提供、数据发现、元数据发布、元数据通知；

所述数据收集模块，设置为从所导入的数据中收集数据并进行初步处理，并提供相应的初步处理结果给所述数据准备模块以及所述数据导出模块；

25 所述数据准备模块，设置为对所收集的数据进行准备处理并提供相应的准备处理的结果给所述数据分析模块以及所述数据导出模块，所述准备处理包括以下之一：数据清洗、数据转换、数据分解、数据验证、数据融合、数据聚合、数据聚类、数据排序、数据摘要、数据分类、数据加密、数据集成；

所述数据分析模块，设置为对经过数据准备处理的数据依据大数据科学，

使用预设算法进行分析处理并从中提取知识数据，并提供相应的分析处理的结果给所述数据可视化模块以及所述数据导出模块；

所述数据可视化模块，设置为对经过数据分析处理的数据进行可视化处理，并提供相应的可视化处理结果给所述数据导出模块；

5 所述数据导出模块，设置为从所述大数据交换单元导出所导入的数据以及所处理的数据到所述大数据目标单元。

本发明实施例还提供了一种大数据交换方法，包括：

大数据源单元采集多种数据，并将所述数据提供给大数据交换单元；

10 所述大数据交换单元接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元；

所述大数据目标单元发送所述数据请求到所述大数据交换单元，所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所述数  
15 据和所述处理数据。

可选地，其中，所述大数据源单元采集多种数据，并将所述数据提供给大数据交换单元，包括：

所述大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种所述数据，  
20 并将所述数据提供给大数据交换单元；

其中，所述数据生产模块，设置为生产数据或者将几种数据进行合并为新的数据；

所述数据拥有模块，设置为持有数据并授权进行数据访问；

所述数据中介模块，设置为不直接生产数据，与所述数据生产模块或者  
25 所述数据拥有模块通过授权协议收购数据；

所述数据代理模块，设置为不直接生产数据，代替所述数据生产模块持有数据；

所述数据摄取适配模块，设置为接收来自于外部数据源的数据并根据所述大数据源单元处理需求，对所摄取的所述外部数据源的数据进行适配；

所述数据采集模块，设置为从数据生产模块、数据代理模块、数据拥有模块、数据中介模块或者数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。

可选地，其中，所述大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种所述数据包括以下至少之一：

所述大数据源单元从所述数据生产模块、所述数据代理模块和所述数据拥有模块获取原始数据；

所述大数据源单元从所述数据拥有模块和所述数据中介模块收购原始数据；

所述大数据源单元从所述数据摄取适配模块摄取来自于外部数据源的数据。

可选地，其中，所述大数据交换单元接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元包括：

所述大数据交换单元通过数据导入从所述大数据源单元导入的数据，通过数据收集、数据准备、数据分析或者数据可视化将这所述导入的数据进行相应处理，输出处理后的数据，所述处理后的数据包括以下之一：所收集的数据、所准备的数据、所分析的数据、可行动的数据，将所述处理后的数据和所述导入的数据按所述大数据目标单元的数据请求导出给所述大数据目标单元。

可选地，其中，所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所述数据和所述处理数据包括：

所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所导入的数据、所收集的数据、所准备的数据、所分析的数据或者可行动的数据。

可选地，其中，所述所导入的数据，设置为经过数据导入所接收、存储和/或缓存的数据以及使用元数据标注的数据；

所述所收集的数据，设置为经过数据收集处理所输出的数据；

5 所述所准备的数据，设置为经过数据准备处理所输出的数据，所述数据包括以下之一：被清洗的数据、被转换的数据、被分解的数据、被验证的数据、被融合的数据、被聚合的数据、被聚类的数据、被注释的数据、被排序的数据、被摘要的数据、被分类的数据、被加密的数据以及被集成的数据；

所述所分析的数据，设置为经过数据分析处理所输出的数据；

所述可行动的数据，设置为经过可视化处理所输出的数据。

10 可选地，其中，所述将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元包括：

判断是对所述数据和所述处理数据进行处理还是直接导出，在确定是直接导出时，按所述大数据目标单元的数据请求导出对所述数据和所述处理数据到大数据目标单元；

15 确定是对所述数据和所述处理数据进行处理时，继续处理所述数据得到所述处理数据以及继续处理所述处理数据得到继续处理的数据。

本发明实施例还提供一种计算机可读存储介质，存储有计算机可执行指令，所述计算机可执行指令用于实现上述方法。

20 通过本发明实施例，大数据源单元采集多种数据，并将该数据提供给大数据交换单元；该大数据交换单元接收该大数据源单元导入的该数据，处理该数据得到处理数据，将该数据和该处理数据按大数据目标单元的数据请求导出到该大数据目标单元；该大数据目标单元发送该数据请求到该大数据交换单元，该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该数据和该处理数据。该方法解决了相关技术中大数据无法开放、互联以及共享的问题，实现了大数据的开放性、互联性以及共享性。

25 在阅读并理解了附图和详细描述后，可以明白其他方面。

## 附图概述

图 1 是根据本发明实施例的一种大数据交换方法的流程图；

图 2 是根据本发明实施例的一种大数据交换系统的结构框图；

图 3 是根据本发明可选实施例的大数据交换系统的结构框图一；

5 图 4 是根据本发明可选实施例的大数据交换系统的结构框图二；

图 5 是根据本发明可选实施例的大数据交换系统的结构框图三。

## 本发明的实施方式

下文中将参考附图来详细说明本发明实施例。需要说明的是，在不冲突  
10 的情况下，本申请中的实施例及实施例中的特征可以相互组合。

需要说明的是，本文说明书、权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。

在本实施例中提供了一种大数据交换方法，图 1 是根据本发明实施例的一种大数据交换方法的流程图，如图 1 所示，该流程包括如下步骤：

15 步骤 S102，大数据源单元采集多种数据，并将该数据提供给大数据交换单元；

步骤 S104，该大数据交换单元接收该大数据源单元导入的该数据，处理该数据得到处理数据，将该数据和该处理数据按该大数据目标单元的数据请求导出到大数据目标单元；

20 步骤 S106，该大数据目标单元发送该数据请求到该大数据交换单元，该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该数据和该处理数据。

通过上述步骤，大数据源单元采集多种数据，并将该数据提供给大数据交换单元，该大数据交换单元接收该大数据源单元导入的该数据，处理该数据  
25 得到处理数据，将该数据和该处理数据按该大数据目标单元的数据请求导出到大数据目标单元，该大数据目标单元发送该数据请求到该大数据交换单元，该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该

数据和该处理数据，解决了相关技术中大数据无法开放、互联以及共享的问题，实现了大数据的开放性、互联性以及共享性。

在本实施例中，该大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种该数据，并将该数据提供给大数据交换单元；

其中，该数据生产模块，设置为生产数据或者将几种数据进行合并为新的数据；

该数据拥有模块，设置为持有数据并授权进行数据访问；

该数据中介模块，设置为不直接生产数据，与该数据生产模块或者该数据拥有模块通过授权协议收购数据；

该数据代理模块，设置为不直接生产数据，代替该数据生产模块持有数据；

该数据摄取适配模块，设置为接收来自于外部数据源的数据并根据该大数据源单元处理需求，对所摄取的该外部数据源的数据进行适配；

该数据采集模块，设置为从数据生产模块、数据代理模块、数据拥有模块、数据中介模块或者数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。所述适配、配置和管理不同的采集方法、采集接口以及数据处理方式是指：适配不同的采集方法、适配不同的采集接口、适配不同的数据处理方式，以及配置采集方法、采集接口和数据处理方式，以及管理采集方法、采集接口和数据处理方式。

在本实施例中，该大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种该数据包括以下至少之一：

该大数据源单元从该数据生产模块、该数据代理模块和该数据拥有模块获取原始数据；

该大数据源单元从该数据拥有模块和该数据中介模块收购原始数据；

该大数据源单元从该数据摄取适配模块摄取来自于外部数据源的数据。

在本实施例中，该大数据交换单元接收该大数据源单元导入的该数据，

处理该数据得到处理数据，将该数据和该处理数据按该大数据目标单元的数据请求导出到大数据目标单元包括：

- 5 该大数据交换单元通过数据导入从该大数据源单元导入的数据，通过数据收集、数据准备、数据分析或者数据可视化将这该导入的数据进行相应处理，输出处理后的数据，该处理后的数据包括以下之一：所收集的数据、所准备的数据、所分析的数据、可行动的数据，将该处理后的数据和该导入的数据按该大数据目标单元的数据请求导出给该大数据目标单元。

在本实施例中，该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该数据和该处理数据包括：

- 10 该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的所导入的数据、所收集的数据、所准备的数据、所分析的数据或者可行动的数据。

在本实施例中，该所导入的数据，设置为经过数据导入所接收、存储和/或缓存的数据以及使用元数据标注的数据；

- 15 该所收集的数据，设置为经过数据收集处理所输出的数据；

该所准备的数据，设置为经过数据准备处理所输出的数据，该数据包括以下之一：被清洗的数据、被转换的数据、被分解的数据、被验证的数据、被融合的数据、被聚合的数据、被聚类的数据、被注释的数据、被排序的数据、被摘要的数据、被分类的数据、被加密的数据以及被集成的数据；

- 20 该所分析的数据，设置为经过数据分析处理所输出的数据；

该可行动的数据，设置为经过可视化处理所输出的数据。

在本实施例中，可以判断是对该数据和该处理数据进行处理还是直接导出，在确定是直接导出时，按该大数据目标单元的数据请求导出对该数据和该处理数据到大数据目标单元；

- 25 若确定是对该数据和该处理数据进行处理时，继续处理该数据得到处理数据以及继续处理该处理数据得到继续处理的数据。

在本发明的实施例中还提供了大数据交换系统，图 2 是根据本发明实施

例的一种大数据交换系统的结构框图,如图 2 所示,包括:大数据源单元 22,大数据交换单元 24,大数据目标单元 26;

该大数据源单元 22,设置为采集多种数据,并将该数据提供给该大数据交换单元 24;

- 5 该大数据交换单元 24,设置为接收该大数据源单元 22 导入的该数据,处理该数据得到处理数据,将该数据和该处理数据按该大数据目标单元 26 的数据请求导出到该大数据目标单元 26;

10 该大数据目标单元 26,设置为发送该数据请求到该大数据交换单元 24,以及设置为接收该大数据交换单元 24 导出的与该数据请求对应的该数据和该处理数据。

通过上述系统,大数据源单元 22 采集多种数据,并将该数据提供给大数据交换单元 24,该大数据交换单元 24 接收该大数据源单元 22 导入的该数据,处理该数据得到处理数据,将该数据和该处理数据按该大数据目标单元 26 的数据请求导出到大数据目标单元 26,该大数据目标单元 26 发送该数据请求到该大数据交换单元 24,该大数据目标单元 26 接收该大数据交换单元 24 导出的与该数据请求对应的该数据和该处理数据,解决了相关技术中大数据无法开放、互联以及共享的问题,实现了大数据的开放性、互联性以及共享性。

20 在本实施例中,该大数据源单元 22 包括以下至少之一:数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块,其中:

该数据生产模块,设置为生产数据或者将几种数据进行合并为新的数据;

该数据拥有模块,设置为持有数据并授权进行数据访问;

25 该数据中介模块,设置为不直接生产数据,与该数据生产模块或者该数据拥有模块通过授权协议收购数据;

该数据代理模块,设置为不直接生产数据,代替该数据生产模块持有数据;

该数据摄取适配模块,设置为接收来自于外部数据源的数据并根据该大

数据源单元 22 处理需求，对所摄取的该外部数据源的数据进行适配；

该数据采集模块，设置为从该数据生产模块、该数据代理模块、该数据拥有模块、该数据中介模块或者该数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。

5 在本实施例中，该大数据交换单元 24 包括以下至少之一：数据导入模块、数据收集模块、数据准备模块、数据分析模块、数据可视化模块以及数据导出模块，其中：

10 该数据导入模块，设置为接收该大数据源单元 22 的数据，并存储和/或缓存该数据，该数据导入模块接收数据的方式包括以下之一（即数据导入模块采用以下方式之一接收大数据源单元 22 的数据）：数据查询、数据请求、数据抓取、数据提供、数据发现、元数据发布、元数据通知；

该数据收集模块，设置为从所导入的数据中收集数据并进行初步处理，并提供相应的初步处理结果给该数据准备模块以及该数据导出模块；

15 该数据准备模块，设置为对所收集的数据进行准备处理并提供相应的准备处理的结果给该数据分析模块以及该数据导出模块，该准备处理包括以下之一：数据清洗、数据转换、数据分解、数据验证、数据融合、数据聚合、数据聚类、数据排序、数据摘要、数据分类、数据加密、数据集成；

20 该数据分析模块，设置为对经过数据准备处理的数据依据大数据科学，使用预设算法进行分析处理并从中提取知识数据，并提供相应的分析处理的结果给该数据可视化模块以及该数据导出模块；

该数据可视化模块，设置为对经过数据分析处理的数据进行可视化处理，并提供相应的可视化处理结果给该数据导出模块；

25 该数据导出模块，设置为从该大数据交换单元 24 导出所导入的数据以及所处理的数据到该大数据目标单元 26。所述所导入的数据是指大数据源单元 22 导入到大数据交换单元 24 的数据，即数据导入模块接收的数据。所述所处理的数据是指前述数据收集模块、数据准备模块、数据分析模块、数据可视化模块的处理结果。

下面结合可选实施例和实施方式对本发明实施例进行详细说明。

本可选实施例提供一种大数据交换的系统及其实现方法，能够有效的解决大数据的孤岛性、封闭性、无法互联以及无法共享的问题。

本可选实施例提供了一种大数据交换的系统，包括：大数据源单元（Big  
5 Data Source）、大数据交换单元（Big Data Exchanger）和大数据目标单元（Big Data Target）；其中，

该大数据源单元，包括数据生产者模块（Data Producer）、数据拥有者  
模块（Data Owner）、数据中介者模块（Data broker 或 Data Agent）、数据  
代理者模块（Data Proxy）、数据摄取适配者模块（Ingest Adaptor）以及数据  
10 采集模块，主要设置为采集多种数据并将这些数据提供给大数据交换单元；  
其中，该数据生产者模块、数据拥有者模块、数据中介者模块、数据代理者  
模块、数据摄取适配者模块相当于上述实施例中的数据生产模块、数据拥有  
模块、数据中介模块、数据代理模块、数据摄取适配模块。

该数据生产者模块，设置为原始数据的生产者或者将几种原始数据进行  
15 合并而提供新的原始数据的加工者；

该数据拥有者模块，设置为持有数据并拥有授权进行数据访问的权利者；

该数据中介者模块，设置为本身不直接生产数据，但却与数据生产者或  
者原始数据拥有者通过授权协议而收购这些数据从而成为该数据的新拥有者；

该数据代理者模块，设置为本身不直接生产数据但却代替数据生产者持  
20 有（比如数字化、转换以及传输）数据（如某些物联网网关收集多种传感器  
所产生的数据从而成为数据代理者）；

该数据摄取适配者模块，设置为接收来自于外部数据源的数据并根据本  
地（即该大数据源单元）处理需求（如存储和/或缓存以及导出等）对所摄取  
自外部数据源的数据进行适配；

25 该数据采集模块，设置为从数据生产者、数据代理者、数据拥有者、数  
据中介者以及数据摄取适配者采集数据，并适配、配置和管理不同的采集方  
法、采集接口以及数据处理方式。

该数据采集，设置为：从数据生产者、数据代理者和数据拥有者获取

( Capture )原始数据;从数据拥有者和数据中介者收购( Acquire )原始数据;通过数据摄取适配器摄取( Ingest )来自于外部数据源的数据。

5 该大数据交换单元,包括数据导入模块、数据收集模块、数据准备模块、数据分析模块、数据可视化模块以及数据导出模块,主要设置为从大数据源单元导入数据,并将这些所输入以及经过相应处理(包括收集、准备、分析和可视化)后所输出的数据(包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据)按需导出到大数据目标单元;其中,

10 该数据导入模块,设置为接收来自于大数据源单元的原始数据并存储和/或缓存这些数据。接收数据的方式包括以下之一:数据查询、数据请求、数据抓取、数据提供、数据发现、元数据发布以及元数据通知。

该数据收集模块,设置为从所导入的数据中收集数据并进行初步处理(如初始相同结构的数据集的合并等)并提供相应的初步处理结果给数据准备模块以及导出模块;

15 该数据准备模块,设置为对所收集的数据进行准备处理并提供相应的准备处理的结果给数据分析模块以及导出模块,准备处理的方式包括以下之一:数据清洗、数据转换、数据分解、数据验证、数据融合、数据聚合、数据聚类、数据排序、数据摘要、数据分类、数据加密以及数据集集成等;

20 该数据分析模块,设置为对经过数据准备处理的数据依据大数据科学,使用指定算法进行分析处理并从中提取知识,并提供相应的分析处理的结果给可视化模块以及导出模块;

该数据可视化模块,设置为对经过数据分析处理的数据进行可视化处理(如以最佳格式和数据展示形式来传递知识和说明含义)并提供相应的可视化处理结果(如统计图型、图表、报告等)给导出模块;

25 该数据导出模块,设置为从大数据交换单元导出所导入的数据以及所处理的数据(包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据)到大数据目标单元。

该所导入的数据,设置为经过数据导入功能所接收并存储和/或缓存的原始数据以及使用元数据标注的原始数据;

该所收集的数据，设置为经过数据收集处理所输出的数据，如被合并的数据等；

5 该所准备的数据，设置为经过数据准备处理所输出的数据，如包括以下之一：被清洗的数据、被转换的数据、被分解的数据、被验证的数据、被融合的数据、被聚合的数据、被聚类的数据、被注释的数据、被排序的数据、被摘要的数据、被分类的数据、被加密的数据以及被集成的数据等；

该所分析的数据，设置为经过数据分析处理所输出的数据，如（依据大数据科学并使用指定算法进行分析处理并从中所提取的）知识数据等；

10 该可行动的数据，设置为经过可视化处理所输出的数据，如可视化数据，透视数据和报告数据等。

该数据导入，设置为接收来自于大数据源单元的原始数据并存储和/或缓存这些数据；

该数据导出，设置为传输所导入的数据和所处理的数据（包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据）到大数据目标单元。

15 该大数据目标单元，主要设置为向大数据交换单元请求并从其接收所请求的多种数据，如所导入的数据、所收集的数据、所准备的数据、所分析的数据以及可行动的数据。

本可选实施例还提供一种大数据交换的方法，该方法包括：

20 大数据源单元从数据生产者（Data Producer）、数据所有者（Data Owner）、数据中介者（Data broker 或 Data Agent）、数据代理者（Data Proxy）以及数据摄取适配者（Ingest Adaptor）通过数据获取、数据收购、数据摄取等收集多种数据并将这些数据提供给大数据交换单元；

25 大数据交换单元通过数据导入从大数据源单元导入（即接收，存储和/或缓存）数据，通过数据收集、数据准备、数据分析、数据可视化等将这些所导入的数据进行相应处理，输出每个对应步骤处理后的数据（包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据）；处理后的数据连同所导入的数据一起，按需导出给大数据目标单元；

大数据目标单元向大数据交换单元请求所需数据（如所导入的数据、所

收集的数据、所准备的数据、所分析的数据以及可行动的数据)，并从大数据交换单元接收所请求的这些数据。

5 本可选实施例所提供的大数据交换系统及其实现方法的技术方案，定义了大数据交换的生态系统（即大数据源、大数据交换者以及大数据目标）及其角色定义（即数据采集、数据导入和数据导出），给出了大数据交换的参考架构和功能界面（即大数据源功能、模块组成及活动,大数据交换者功能、模块组成及活动,大数据目标功能、模块组成及活动）。通过规范大数据交换的生态系统、角色定义、参考架构以及功能界面，从而有效地解决了大数据的孤岛性、封闭性、无法互联以及无法共享的问题。

10

图 3 是根据本发明可选实施例的大数据交换系统的结构框图一。如图 3 所示，大数据交换的系统结构从左到右分成大数据源单元，大数据交换单元以及大数据目标单元。其中，大数据源单元包括本地大数据单元（301）以及外部大数据单元（301a），二者通过数据摄取接口进行数据交换；大数据交换单元（302）通过数据导入接口与大数据源单元（301）进行数据交换；大  
15 大数据目标单元（303）通过数据导出接口与大数据交换单元（302）进行数据交换。

图 4 是根据本发明可选实施例的大数据交换系统的结构框图二。如图 4 所示，大数据交换的系统结构中：

20 大数据源单元（301）由数据生产者模块、数据拥有者模块、数据中介模块、数据代理模块、数据摄取适配者模块以及数据采集模块等构成，并使用数据获取、数据收购以及数据摄取等不同的数据采集方法从数据生产者、数据代理者、数据拥有者、数据中介者以及数据摄取适配者采集数据，最后将所采集到的数据汇总到数据采集模块，以便导入到大数据交换单元。

25 大数据交换单元（302）由数据导入模块、数据收集模块、数据准备模块、数据分析模块、数据可视化模块以及数据导出模块等构成，其中：

数据导入模块设置为接收来自于大数据源单元的原始数据并存储和/或缓存这些数据，其接收数据的方式包括以下之一：数据查询、数据请求、数据抓取、数据提供、数据发现、元数据发布以及元数据通知；

数据收集模块设置为从所存储和/或缓存的数据中收集数据并进行初步处理（比如对于具有初始相同结构的数据集进行合并等）并提供相应的初步处理结果，以便提供给数据准备模块做处理或者根据需要经由数据导出模块导出所导入的数据给大数据目标单元；

- 5 数据准备模块设置为对经过数据收集处理所收集的数据进行准备处理并提供相应的准备处理的结果，以便提供给数据分析模块做处理或者根据需要经由数据导出模块导出所准备的数据给大数据目标单元。准备处理包括以下之一：数据清洗、数据转换、数据分解、数据验证、数据融合、数据聚合、数据聚类、数据排序、数据摘要、数据分类、数据加密以及数据集成等，相应地，所提供的数据包括以下之一：被清洗的数据、被转换的数据、被分解的数据、被验证的数据、被融合的数据、被聚合的数据、被聚类的数据、被注释的数据、被排序的数据、被摘要的数据、被分类的数据、被加密的数据以及被集成的数据等；
- 10

- 数据分析模块设置为对经过数据准备处理所准备的数据依据大数据科学，使用指定算法进行分析处理并从中提取知识，并提供相应的分析处理的结果，以便提供给可视化模块做处理或者根据需要经由数据导出模块导出所分析的数据给大数据目标单元；
- 15

- 数据可视化模块设置为对经过数据分析处理所分析的数据进行可视化处理（如以最佳格式和数据展示形式来传递知识和说明含义）并提供相应的可视化处理结果（如统计图型、图表、报告等），并需要根据经由数据导出模块导出可视化处理结果给大数据目标单元；
- 20

数据导出模块设置为将所导入的数据以及所处理的数据（包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据）导出到大数据目标单元。

- 25 大数据目标单元（303）设置为向大数据交换单元请求并从其接收所请求的多种数据，如所导入的数据、所收集的数据、所准备的数据、所分析的数据以及可行动的数据。

大数据交换的方法中：

大数据源单元（301）使用数据获取、数据收购以及数据摄取等不同的数

据采集方法从数据生产者、数据代理者、数据拥有者、数据中介者以及数据摄取适配者采集数据，最后将所采集到的数据汇总到数据采集模块，并导入到大数据交换单元。

5 大数据交换单元(302)通过数据导入模块来接收来自于大数据源单元的原始数据并存储和/或缓存这些数据，此时的数据被称为所导入的数据；可选地，根据来自于大数据目标单元(303)的请求来判断针对所导入的数据是做收集处理还是直接导出；如做收集处理，则进入下面的收集处理流程，如不做收集处理，则经由数据导出模块可直接导出给大数据目标单元；

10 大数据交换单元(302)通过数据收集模块来从所导入的数据中收集数据并进行初步处理(比如对于具有初始相同结构的数据集进行合并等)并提供相应的初步处理结果，此时的数据被称为所收集的数据；可选地，根据来自于大数据目标单元(303)的请求来判断针对所收集的数据是做准备处理还是直接导出；如做准备处理，则进入下面的准备处理流程，如不做准备处理，则经由数据导出模块可直接导出给大数据目标单元；

15 大数据交换单元(302)通过数据准备模块来对所收集的数据进行准备处理并提供相应的准备处理的结果，此时的数据被称为所准备的数据；可选地，根据来自于大数据目标单元(303)的请求来判断针对所准备的数据是做分析处理还是直接导出；如做分析处理，则进入下面的分析处理流程，如不做分析处理，则经由数据导出模块可直接导出给大数据目标单元；

20 大数据交换单元(302)通过数据分析模块来对所准备的数据依据大数据科学，使用指定算法进行分析处理并从中提取知识，并提供相应的分析处理的结果，此时的数据被称为所分析的数据；可选地，根据来自于大数据目标单元(303)的请求来判断针对所分析的数据是做可视化处理还是直接导出；如做可视化处理，则进入下面的可视化处理流程，如不做可视化处理，则经  
25 由数据导出模块可直接导出给大数据目标单元；

大数据交换单元(302)通过数据可视化模块来对所分析的数据进行可视化处理(如以最佳格式和数据展示形式来传递知识和说明含义)并提供相应的可视化处理结果(如统计图型、图表、报告等)，此时的数据被称为可行动的数据；根据需要，可行动的数据经由数据导出模块可直接导出给大数据

目标单元;

大数据交换单元(302)通过数据导出模块来将所导入的数据以及所处理的数据(包括所收集的数据、所准备的数据、所分析的数据以及可行动的数据)导出到大数据目标单元。

- 5 大数据目标单元(303)通过向大数据交换单元(302)请求并从其接收所请求的多种数据,如所导入的数据、所收集的数据、所准备的数据、所分析的数据以及可行动的数据。

10 在一个可选实施例中,图5是根据本发明可选实施例的大数据交换系统的结构框图三。如图5所示,包括下面的步骤:

步骤1,大数据目标单元(303)向大数据交换单元(302)发送业务请求(带有目标数据类型等信息),请求交换目标数据(如所导入的数据、所收集的数据、所准备的数据、所分析的数据以及可行动的数据);

15 步骤2,大数据交换单元(302)接收到来自于大数据目标单元(303)的业务请求(带有目标数据类型等信息),则大数据交换单元(302)判断所请求的目标数据是否存在,如果目标数据存在,则直接导出,从而实现数据交换;如果目标数据不存在,则:

步骤2.1,大数据交换单元(302)发送消息给大数据目标单元(303),给出服务失败的响应;或者,

20 步骤2.2,可选地,大数据交换单元(302)启动相应的数据导入流程向大数据源单元(301和301a)请求所需数据并经由大数据交换单元(302)的数据处理流程(即数据收集、收集准备、数据分析以及数据可视化等处理)来尝试满足大数据目标单元(303)的业务请求;

25 步骤3,通过步骤2.2,来自于大数据源单元(301和301a)的数据以及经由大数据交换单元(302)所处理后,

步骤3.1,能够满足来自于大数据目标单元(303)的业务请求(带有目标数据类型等信息),则大数据交换单元(302)导出相关数据给大数据目标单元(303),从而实现数据交换;

步骤 3.2, 若不能够满足来自于大数据目标单元 (303) 的业务请求, 则大数据交换单元 (302) 发送消息给大数据目标单元 (303), 给出服务失败的响应。

5 本可选实施例提供了一种大数据交换的系统及其实现方法, 实现了大数据的开放性、互联性以及共享性, 并充分调动孤立且不活跃的大数据能够充分参与到大数据交易活动中, 为生产和生活所用以及创造大数据所应有的价值。

10 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根据上述实施例的方法可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。基于这样的理解, 本发明实施例的技术方案可以以软件产品的形式体现出来, 该计算机软件产品可以存储在一个存储介质(如 ROM/RAM、磁碟、光盘)中, 包括程序指令用以使得一台终端设备(可以是手机, 计算机, 服务器, 或者网络设备等)执行本发明每个实施例所述的方法。

15 本发明的实施例还提供了一种存储介质。可选地, 在本实施例中, 上述存储介质可以被设置为存储用于执行上述实施例的方法步骤的程序代码:

可选地, 在本实施例中, 上述存储介质可以包括但不限于: U 盘、只读存储器 (ROM, Read-Only Memory)、随机存取存储器 (RAM, Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

20 可选地, 在本实施例中, 处理器根据存储介质中已存储的程序代码执行上述实施例的方法步骤。

25 显然, 本领域的技术人员应该明白, 上述的本发明实施例的模块或步骤可以用通用的计算装置来实现, 它们可以集中在单个的计算装置上, 或者分布在多个计算装置所组成的网络上, 可选地, 它们可以用计算装置可执行的程序代码来实现, 从而, 可以将它们存储在存储装置中由计算装置来执行, 并且在某些情况下, 可以以不同于此处的顺序执行所示出或描述的步骤, 或者将它们分别制作成集成电路模块, 或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样, 本发明实施例不限制于任何特定的硬件和软件结合。

以上所述仅为本发明的可选实施例而已, 并不用于限制本发明, 对于本

领域的技术人员来说，本发明实施例可以有多种更改和变化。凡在本申请技术方案的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

## 5 工业实用性

10 通过本发明实施例，大数据源单元采集多种数据，并将该数据提供给大数据交换单元；该大数据交换单元接收该大数据源单元导入的该数据，处理该数据得到处理数据，将该数据和该处理数据按大数据目标单元的数据请求导出到该大数据目标单元；该大数据目标单元发送该数据请求到该大数据交换单元，该大数据目标单元接收该大数据交换单元导出的与该数据请求对应的该数据和该处理数据。该方法解决了相关技术中大数据无法开放、互联以及共享的问题，实现了大数据的开放性、互联性以及共享性。

## 权 利 要 求 书

1、一种大数据交换系统，包括：大数据源单元，大数据交换单元以及大数据目标单元，其中：

5 所述大数据源单元，设置为采集多种数据，并将所述数据提供给所述大数据交换单元；

所述大数据交换单元，设置为接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到所述大数据目标单元；

10 所述大数据目标单元，设置为发送所述数据请求到所述大数据交换单元，以及设置为接收所述大数据交换单元导出的与所述数据请求对应的所述数据和所述处理数据。

2、根据权利要求1所述的系统，其中，所述大数据源单元包括以下至少之一：数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块，其中：

15 所述数据生产模块，设置为生产数据或者将几种数据进行合并为新的数据；

所述数据拥有模块，设置为持有数据并授权进行数据访问；

所述数据中介模块，设置为不直接生产数据，与所述数据生产模块或者所述数据拥有模块通过授权协议收购数据；

20 所述数据代理模块，设置为不直接生产数据，代替所述数据生产模块持有数据；

所述数据摄取适配模块，设置为接收来自于外部数据源的数据并根据所述大数据源单元处理需求，对所摄取的所述外部数据源的数据进行适配；

25 所述数据采集模块，设置为从所述数据生产模块、所述数据代理模块、所述数据拥有模块、所述数据中介模块或者所述数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。

3、根据权利要求1所述的系统，其中，所述大数据交换单元包括以下至

少之一：数据导入模块、数据收集模块、数据准备模块、数据分析模块、数据可视化模块以及数据导出模块，其中：

5 所述数据导入模块，设置为接收所述大数据源单元的数据，并存储和/或缓存所述数据，所述数据导入模块接收数据的方式包括以下之一：数据查询、数据请求、数据抓取、数据提供、数据发现、元数据发布、元数据通知；

所述数据收集模块，设置为从所导入的数据中收集数据并进行初步处理，并提供相应的初步处理结果给所述数据准备模块以及所述数据导出模块；

10 所述数据准备模块，设置为对所收集的数据进行准备处理并提供相应的准备处理的结果给所述数据分析模块以及所述数据导出模块，所述准备处理包括以下之一：数据清洗、数据转换、数据分解、数据验证、数据融合、数据聚合、数据聚类、数据排序、数据摘要、数据分类、数据加密、数据集成；

所述数据分析模块，设置为对经过数据准备处理的数据依据大数据科学，使用预设算法进行分析处理并从中提取知识数据，并提供相应的分析处理的结果给所述数据可视化模块以及所述数据导出模块；

15 所述数据可视化模块，设置为对经过数据分析处理的数据进行可视化处理，并提供相应的可视化处理结果给所述数据导出模块；

所述数据导出模块，设置为从所述大数据交换单元导出所导入的数据以及所处理的数据到所述大数据目标单元。

4、一种大数据交换方法，包括：

20 大数据源单元采集多种数据，并将所述数据提供给大数据交换单元；

所述大数据交换单元接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元；

25 所述大数据目标单元发送所述数据请求到所述大数据交换单元，所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所述数据和所述处理数据。

5、根据权利要求4所述的方法，其中，所述大数据源单元采集多种数据，并将所述数据提供给大数据交换单元，包括：

所述大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种所述数据，并将所述数据提供给大数据交换单元；

5 其中，所述数据生产模块，设置为生产数据或者将几种数据进行合并为新的数据；

所述数据拥有模块，设置为持有数据并授权进行数据访问；

所述数据中介模块，设置为不直接生产数据，与所述数据生产模块或者所述数据拥有模块通过授权协议收购数据；

10 所述数据代理模块，设置为不直接生产数据，代替所述数据生产模块持有数据；

所述数据摄取适配模块，设置为接收来自于外部数据源的数据并根据所述大数据源单元处理需求，对所摄取的所述外部数据源的数据进行适配；

15 所述数据采集模块，设置为从数据生产模块、数据代理模块、数据拥有模块、数据中介模块或者数据摄取适配模块采集数据，并适配、配置和管理不同的采集方法、采集接口以及数据处理方式。

6、根据权利要求5所述的方法，其中，所述大数据源单元通过数据生产模块、数据拥有模块、数据中介模块、数据代理模块、数据摄取适配模块以及数据采集模块摄取收集多种所述数据包括以下至少之一：

20 所述大数据源单元从所述数据生产模块、所述数据代理模块和所述数据拥有模块获取原始数据；

所述大数据源单元从所述数据拥有模块和所述数据中介模块收购原始数据；

所述大数据源单元从所述数据摄取适配模块摄取来自于外部数据源的数据。

25 7、根据权利要求4所述的方法，其中，所述大数据交换单元接收所述大数据源单元导入的所述数据，处理所述数据得到处理数据，将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元包括：

所述大数据交换单元通过数据导入从所述大数据源单元导入的数据，通

过数据收集、数据准备、数据分析或者数据可视化将这所述导入的数据进行相应处理，输出处理后的数据，所述处理后的数据包括以下之一：所收集的数据、所准备的数据、所分析的数据、可行动的数据，将所述处理后的数据和所述导入的数据按所述大数据目标单元的数据请求导出给所述大数据目标单元。

8、根据权利要求4所述的方法，其中，所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所述数据和所述处理数据包括：

所述大数据目标单元接收所述大数据交换单元导出的与所述数据请求对应的所导入的数据、所收集的数据、所准备的数据、所分析的数据或者可行动的数据。

9、根据权利要求8所述的方法，其中，

所述所导入的数据，设置为经过数据导入所接收、存储和/或缓存的数据以及使用元数据标注的数据；

所述所收集的数据，设置为经过数据收集处理所输出的数据；

所述所准备的数据，设置为经过数据准备处理所输出的数据，所述数据包括以下之一：被清洗的数据、被转换的数据、被分解的数据、被验证的数据、被融合的数据、被聚合的数据、被聚类的数据、被注释的数据、被排序的数据、被摘要的数据、被分类的数据、被加密的数据以及被集成的数据；

所述所分析的数据，设置为经过数据分析处理所输出的数据；

所述可行动的数据，设置为经过可视化处理所输出的数据。

10、根据权利要求4至9任一项所述的方法，其中，所述将所述数据和所述处理数据按所述大数据目标单元的数据请求导出到大数据目标单元包括：

判断是对所述数据和所述处理数据进行处理还是直接导出，在确定是直接导出时，按所述大数据目标单元的数据请求导出对所述数据和所述处理数据到大数据目标单元；

确定是对所述数据和所述处理数据进行处理时，继续处理所述数据得到所述处理数据以及继续处理所述处理数据得到继续处理的数据。

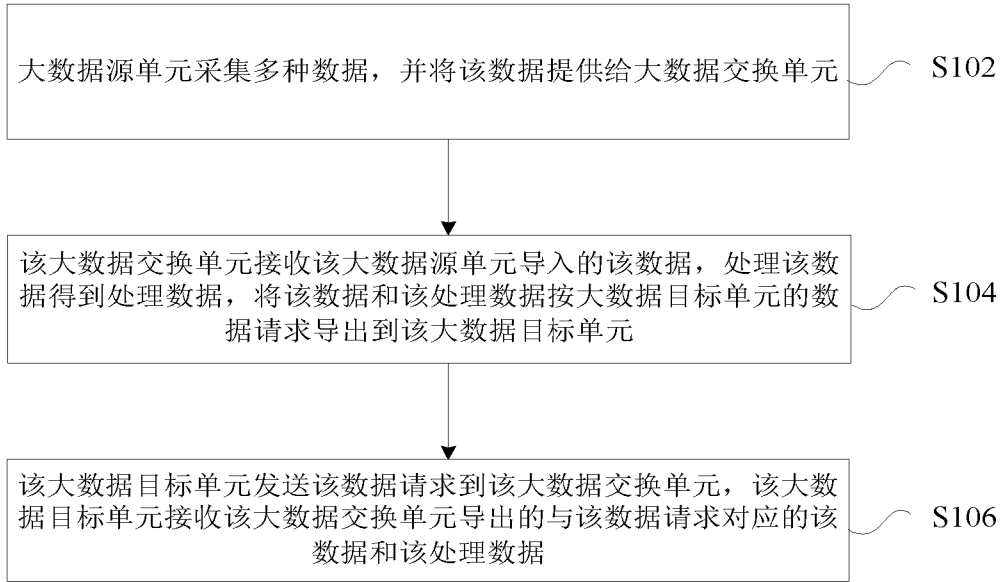


图 1



图 2

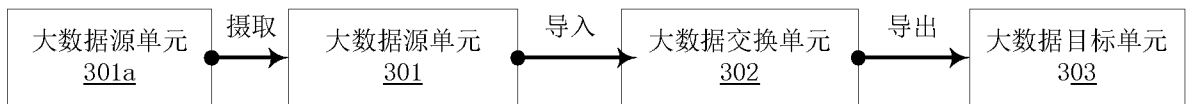


图 3

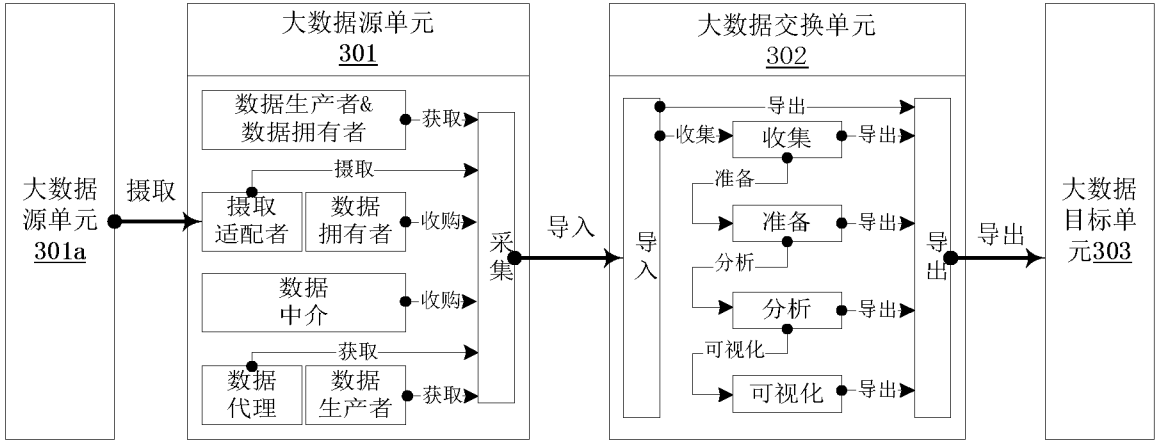


图 4

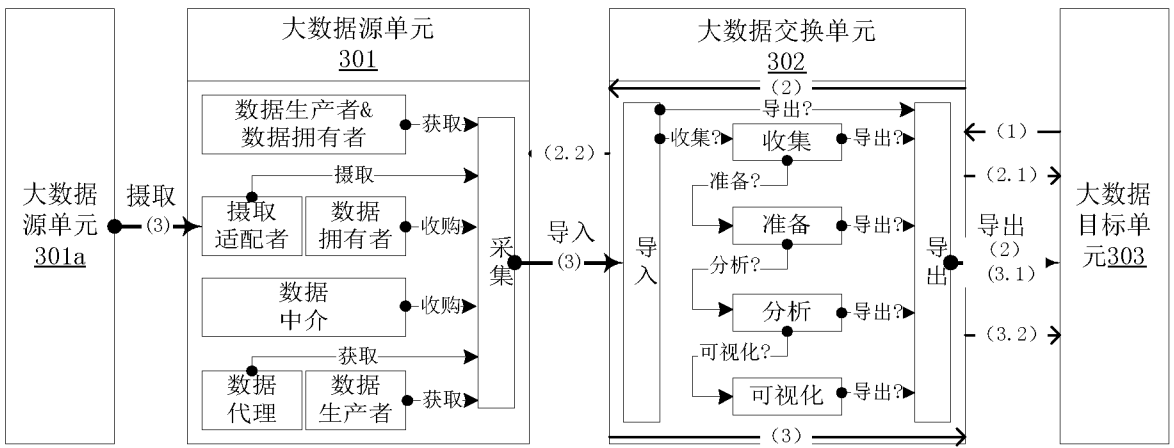


图 5

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2016/084954**

## 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 H04L

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; WPI; EPODOC; CNKI: large data, island, distributed, heterogeneous, acquisition, distribution, forward, large, data, exchang+, extract+, source

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103731298 A (INSTITUTE 706, THE SECOND ACADEMY OF CHINESE AEROSPACE SCIENCE AND INDUSTRY CORP. et al.), 16 April 2014 (16.04.2014), see description, paragraphs [0095]-[0105], and figures 2 and 4	1-10
X	CN 104699855 A (CHENGDU CALABAR INFORMATION TECHNOLOGY CO., LTD.), 10 June 2015 (10.06.2015), description, paragraphs [0003] and [0006]	1-10
A	CN 102546730 A (CHINA MOBILE COMMUNICATIONS CORPORATION), 04 July 2012 (04.07.2012), the whole document	1-10
A	CN 102081656 A (JIANGSU MENGLANSHENCAI TECHNOLOGY DEVELOPMENT CO., LTD.), 01 June 2011 (01.06.2011), the whole document	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” document member of the same patent family</p>
---	---

<p>Date of the actual completion of the international search</p> <p style="text-align: center;">16 August 2016 (16.08.2016)</p>	<p>Date of mailing of the international search report</p> <p style="text-align: center;"><b>26 August 2016 (26.08.2016)</b></p>
<p>Name and mailing address of the ISA/CN:</p> <p>State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451</p>	<p>Authorized officer</p> <p style="text-align: center;"><b>LIU, Man</b></p> <p>Telephone No.: (86-10) <b>62414025</b></p>

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

**PCT/CN2016/084954**

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103731298 A	16 April 2014	None	
CN 104699855 A	10 June 2015	None	
CN 102546730 A	04 July 2012	None	
CN 102081656 A	01 June 2011	None	

国际检索报告

国际申请号

PCT/CN2016/084954

<p>A. 主题的分类</p> <p>G06F 17/30(2006.01)i</p> <p>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</p>																											
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06F H04L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT;WPI;EPODOC;CNKI:大数据, 交换, 孤岛, 分布式, 异构, 采集, 分发, 转发, 抽取, large, data, exchang+, extract+, source</p>																											
<p>C. 相关文件</p> <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>X</td> <td>CN 103731298 A (中国航天科工集团第二研究院七〇六所 等) 2014年 4月 16日 (2014 - 04 - 16) 参见说明书第[0095]-[0105]段、说明书附图2, 4</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>CN 104699855 A (成都卡莱博尔信息技术有限公司) 2015年 6月 10日 (2015 - 06 - 10) 说明书第[0003], [0006]段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 102546730 A (中国移动通信集团公司) 2012年 7月 4日 (2012 - 07 - 04) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 102081656 A (江苏梦兰神彩科技发展有限公司) 2011年 6月 1日 (2011 - 06 - 01) 全文</td> <td>1-10</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 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          “&amp;” 同族专利的文件</p> <table border="1"> <tr> <td>国际检索实际完成的日期</td> <td>国际检索报告邮寄日期</td> </tr> <tr> <td>2016年 8月 16日</td> <td>2016年 8月 26日</td> </tr> <tr> <td>ISA/CN的名称和邮寄地址</td> <td>受权官员</td> </tr> <tr> <td>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</td> <td>刘曼</td> </tr> <tr> <td>传真号 (86-10)62019451</td> <td>电话号码 (86-10)62414025</td> </tr> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103731298 A (中国航天科工集团第二研究院七〇六所 等) 2014年 4月 16日 (2014 - 04 - 16) 参见说明书第[0095]-[0105]段、说明书附图2, 4	1-10	X	CN 104699855 A (成都卡莱博尔信息技术有限公司) 2015年 6月 10日 (2015 - 06 - 10) 说明书第[0003], [0006]段	1-10	A	CN 102546730 A (中国移动通信集团公司) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-10	A	CN 102081656 A (江苏梦兰神彩科技发展有限公司) 2011年 6月 1日 (2011 - 06 - 01) 全文	1-10	国际检索实际完成的日期	国际检索报告邮寄日期	2016年 8月 16日	2016年 8月 26日	ISA/CN的名称和邮寄地址	受权官员	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	刘曼	传真号 (86-10)62019451	电话号码 (86-10)62414025
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																									
X	CN 103731298 A (中国航天科工集团第二研究院七〇六所 等) 2014年 4月 16日 (2014 - 04 - 16) 参见说明书第[0095]-[0105]段、说明书附图2, 4	1-10																									
X	CN 104699855 A (成都卡莱博尔信息技术有限公司) 2015年 6月 10日 (2015 - 06 - 10) 说明书第[0003], [0006]段	1-10																									
A	CN 102546730 A (中国移动通信集团公司) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-10																									
A	CN 102081656 A (江苏梦兰神彩科技发展有限公司) 2011年 6月 1日 (2011 - 06 - 01) 全文	1-10																									
国际检索实际完成的日期	国际检索报告邮寄日期																										
2016年 8月 16日	2016年 8月 26日																										
ISA/CN的名称和邮寄地址	受权官员																										
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	刘曼																										
传真号 (86-10)62019451	电话号码 (86-10)62414025																										

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2016/084954

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	103731298	A	2014年 4月 16日	无	
CN	104699855	A	2015年 6月 10日	无	
CN	102546730	A	2012年 7月 4日	无	
CN	102081656	A	2011年 6月 1日	无	

表 PCT/ISA/210 (同族专利附件) (2009年7月)