



(12) 发明专利

(10) 授权公告号 CN 113434506 B

(45) 授权公告日 2023.05.16

(21) 申请号 202110724252.9

G06F 16/2455 (2019.01)

(22) 申请日 2021.06.29

G06F 16/901 (2019.01)

(65) 同一申请的已公布的文献号  
申请公布号 CN 113434506 A

(56) 对比文件

CN 110046062 A, 2019.07.23

US 2009100115 A1, 2009.04.16

(43) 申请公布日 2021.09.24

审查员 庞丽金

(73) 专利权人 平安科技(深圳)有限公司  
地址 518000 广东省深圳市福田区福田街  
道福安社区益田路5033号平安金融中  
心23楼

(72) 发明人 梁新宇

(74) 专利代理机构 北京英特普罗知识产权代理  
有限公司 11015  
专利代理师 黄胜波

(51) Int. Cl.

G06F 16/22 (2019.01)

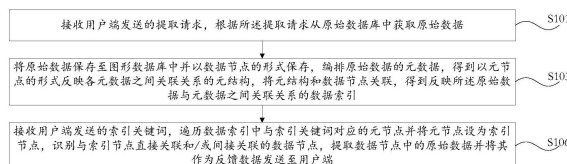
权利要求书3页 说明书15页 附图3页

(54) 发明名称

数据管理及检索方法、装置、计算机设备及  
可读存储介质

(57) 摘要

本发明涉及大数据领域,公开了一种数据管  
理及检索方法、装置、计算机设备及可读存储介  
质,包括:根据提取请求从原始数据库中获取原  
始数据;将原始数据保存至图形数据库中并以数  
据节点的形式保存,编排原始数据的元数据得到  
元结构,将元结构和数据节点关联得到数据索  
引;接收用户端发送的索引关键词,遍历数据索  
引中与索引关键词对应的元节点并将元节点设  
为索引节点,识别与索引节点直接关联和/或间  
接关联的数据节点,提取数据节点中的原始数据  
并将其作为反馈数据。本发明还涉及区块链技  
术,信息可存储于区块链节点中。本发明不仅避  
免了算力消耗过大的问题,提高了反馈数据的获  
取效率,并且保证了反馈数据的全面性进而满足  
检索需求。



1. 一种数据管理及检索方法,其特征在于,包括:

接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据;

识别所述原始数据中的无效值并查询所述无效值所对应的元数据,删除所述原始数据的无效值及其所述元数据;

识别所述原始数据中的缺失值并查询所述缺失值对应的元数据,删除所述原始数据中的缺失值及其元数据;将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库;

接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情况;

向用户端发送搜索索引页面、或可选索引页面、或类别索引页面、或元索引页面,接收所述用户端通过所述搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面发送的索引信息;其中,所述搜索索引页面具有索引输入框,用于所述用户端录入所述索引信息;

通过预置的自然语言模型识别所述索引信息中的主语信息,并将所述主语信息设为所述索引关键词;

通过所述自然语言模型识别所述索引信息中的宾语信息,并将所述宾语信息设为所述索引目的词;

通过所述自然语言模型识别所述索引信息中的谓语信息,并将所述谓语信息设为所述索引目的信息;

根据所述索引关键词识别所述数据索引中的元节点并将所述元节点设为索引节点,根据所述索引目的词识别所述数据索引中的元节点或数据节点,并将所述元节点或所述数据节点设为目的节点,根据所述索引目的信息识别与所述索引节点直接关联和/或间接关联的所述目的节点的数量,并将所述数量发送至所述用户端。

2. 根据权利要求1所述的数据管理及检索方法,其特征在于,所述编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,包括:

建立反映元类别之间逻辑关系的维度数据树,根据所述维度数据树构建以元类别作为分类项的结构表;其中,所述元类别是用于对元数据进行分类的使所述元数据形成层级关系的类别信息;

提取所述图形数据库中原始数据的元数据,将所述元数据及其在所述原始数据中的对应信息,录入所述结构表中与所述元数据对应的类别项之下得到维度表;

根据所述维度表构建表征所述元类别及其元数据的元节点,根据所述元类别之间的逻辑关系构建各所述元节点之间关联关系得到元结构。

3. 根据权利要求2所述的数据管理及检索方法,其特征在于,所述将所述元结构和所述

数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引,包括:

提取所述图形数据库中原始数据的元数据,及所述元数据在所述原始数据中的对应信息,从所述元结构中获得与所述元数据对应的元节点,并将所述元节点设为目标节点;

根据所述维度表中的逻辑关系排列所述目标节点形成节点链,将位于所述节点链末端的目标节点与所述数据节点关联得到所述数据索引。

4. 根据权利要求1所述的数据管理及检索方法,其特征在于,所述接收用户端发送的索引关键词之前,所述方法还包括:

提取所述数据索引中的元节点;

提取所述元节点中的元类别,并对所述元类别设置类别输入框,构建具有所述元类别及其所述类别输入框的类别索引页面,并将用于录入所述索引关键词的所述类别索引页面发送至用户端;或

提取所述元节点中的元数据,并对所述元数据设置元输入框,构建具有所述元数据及其所述元输入框的元索引页面,并将用于录入所述索引关键词的所述元索引页面发送至用户端。

5. 根据权利要求1所述的数据管理及检索方法,其特征在于,所述接收用户端发送的索引关键词之前,所述方法还包括:

提取所述数据索引中的元节点,提取所述元节点中的元类别及其元数据,对所述元数据设置用于启用和禁用所述元数据的选择框,并构建与所述选择框关联的搜索输入框,创制具有所述元类别、所述元数据及其选择框,以及与所述选择框关联的搜索输入框的可选索引页面,将用于录入所述索引关键词的所述可选索引页面发送至用户端。

6. 根据权利要求3所述的数据管理及检索方法,其特征在于,所述将所述数量发送至所述用户端之后,所述方法还包括:

将所述目标节点的数量上传至区块链中。

7. 一种数据管理及检索装置,其特征在于,包括:

数据输入模块,用于接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据;

无效清洗单元,用于识别所述原始数据中的无效值并查询所述无效值所对应的元数据,删除所述原始数据的无效值及其所述元数据;

缺失清洗单元,用于识别所述原始数据中的缺失值并查询所述缺失值对应的元数据,删除所述原始数据中的缺失值及其元数据;

索引构建模块,用于将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库;

检索反馈模块,用于接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,

所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情况;

数量检索模块,用于接收用户端发送的索引关键词、索引目标词和索引目的信息,根据所述索引关键词识别所述数据索引中的元节点并将所述元节点设为索引节点,根据所述索引目的词识别所述数据索引中的元节点或数据节点,并将所述元节点或所述数据节点设为目的节点,根据所述索引目的信息识别与所述索引节点直接关联和/或间接关联的所述目的节点的数量,并将所述数量发送至所述用户端;

所述数量检索模块还包括:

页面发送单元,用于向用户端发送搜索索引页面、或可选索引页面、或类别索引页面、或元索引页面,接收所述用户端通过所述搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面发送的索引信息;其中,所述搜索索引页面具有索引输入框,用于所述用户端录入所述索引信息;

关键识别单元,用于通过预置的自然语言模型识别所述索引信息中的主语信息,并将所述主语信息设为所述索引关键词;

目的识别单元,用于通过所述自然语言模型识别所述索引信息中的宾语信息,并将所述宾语信息设为所述索引目的词;

索引目的单元,用于通过所述自然语言模型识别所述索引信息中的谓语信息,并将所述谓语信息设为所述索引目的信息。

8. 一种计算机设备,其包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述计算机设备的处理器执行所述计算机程序时实现权利要求1至6任一项所述数据管理及检索方法的步骤。

9. 一种计算机可读存储介质,所述可读存储介质上存储有计算机程序,其特征在于,所述可读存储介质存储的所述计算机程序被处理器执行时实现权利要求1至6任一项所述数据管理及检索方法的步骤。

## 数据管理及检索方法、装置、计算机设备及可读存储介质

### 技术领域

[0001] 本发明涉及大数据技术领域,尤其涉及一种数据管理及检索方法、装置、计算机设备及可读存储介质。

### 背景技术

[0002] 当前的数据库包括关系型数据库和非关系型数据库,其中,非关系型数据库是一种非关系型的、分布式的、且一般不保证遵循ACID原则的数据存储系统,因此,非关系型数据库以键值对存储,且结构不固定,每一个元组可以有不一样的字段,每个元组可以根据需要增加一些自己的键值对,不局限于固定的结构,可以减少一些时间和空间的开销;关系型数据库是指采用了关系模型来组织数据的数据库,关系模型指的就是二维表格模型,而一个关系型数据库就是由二维表及其之间的联系所组成的一个数据组织。

[0003] 当用户需要查询数据库中的指定数据并获得相应的反馈数据时,发明人意识到,对于非关系型数据库,需要将用户端发送的检索关键字与各主键进行匹配,将匹配的主键所对应的键值作为反馈数据发送至用户端,然而,这种方式对检索关键字的质量要求较高,一旦检索关键字无法准确描述用户所需的反馈数据时,会导致用户端获得反馈数据无法满足用户检索需要;对于关系型数据库,需要根据用户端发送的检索关键字遍历其中的元数据,以获得与所述检索关键字对应的数据并将其作为反馈数据,这种查询方式导致服务器需要消耗大量的算力对所述数据库中的所有数据进行遍历,不仅查询效率低下,而且服务器算力消耗庞大。

### 发明内容

[0004] 本发明的目的是提供一种数据管理及检索方法、装置、计算机设备及可读存储介质,用于解决现有技术存在的因需要对检索关键字的质量具有较高要求,导致用户端获得反馈数据难以满足用户检索需要,以及需要消耗大量的算力对所述数据库中的所有数据进行遍历,不仅查询效率低下,而且服务器算力消耗庞大的问题。

[0005] 为实现上述目的,本发明提供一种数据管理及检索方法,包括:

[0006] 接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据;

[0007] 将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库;

[0008] 接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情

况。

[0009] 上述方案中,所述根据所述提取请求从原始数据库中获取原始数据之后,所述方法还包括:

[0010] 对所述原始数据进行数据清洗,用以删除所述原始数据中的无效数据及其元数据,以及缺失值及其元数据。

[0011] 上述方案中,所述编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,包括:

[0012] 建立反映元类别之间逻辑关系的维度数据树,根据所述维度数据树构建以元类别作为分类项的结构表;其中,所述元类别是用于对元数据进行分类的使所述元数据形成层级关系的类别信息;

[0013] 提取所述图形数据库中原始数据的元数据,将所述元数据及其在所述原始数据中的对应信息,录入所述结构表中与所述元数据对应的类别项之下得到维度表;

[0014] 根据所述维度表构建表征所述元类别及其元数据的元节点,根据所述元类别之间的逻辑关系构建各所述元节点之间关联关系得到元结构。

[0015] 上述方案中,所述将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引,包括:

[0016] 提取所述图形数据库中原始数据的元数据,及所述元数据在所述原始数据中的对应信息,从所述元结构中获得与所述元数据对应的元节点,并将所述元节点设为目标节点;

[0017] 根据所述维度表中的逻辑关系排列所述目标节点形成节点链,将位于所述节点链末端的的目标节点与所述数据节点关联得到所述数据索引。

[0018] 上述方案中,所述接收用户端发送的索引关键词之前,所述方法还包括:

[0019] 提取所述数据索引中的元节点;

[0020] 提取所述元节点中的元类别,并对所述元类别设置类别输入框,构建具有所述元类别及其所述类别输入框的类别索引页面,并将用于录入所述索引关键词的所述类别索引页面发送至用户端;或

[0021] 提取所述元节点中的元数据,并对所述元数据设置元输入框,构建具有所述元数据及其所述元输入框的元索引页面,并将用于录入所述索引关键词的所述元索引页面发送至用户端。

[0022] 上述方案中,所述接收用户端发送的索引关键词之前,所述方法还包括:

[0023] 提取所述数据索引中的元节点,提取所述元节点中的元类别及其元数据,对所述元数据设置用于启用和禁用所述元数据的选择框,并构建与所述选择框关联的搜索输入框,创制具有所述元类别、所述元数据及其选择框,以及与所述选择框关联的搜索输入框的可选索引页面,将用于录入所述索引关键词的所述可选索引页面发送至用户端。

[0024] 上述方案中,所述得到反映所述原始数据与所述元数据之间关联关系的数据索引之后,所述方法还包括:

[0025] 接收用户端发送的索引关键词、索引目标词和索引目的信息,根据所述索引关键词识别所述数据索引中的元节点并将所述元节点设为索引节点,根据所述索引目的词识别所述数据索引中的元节点或数据节点,并将所述元节点或所述数据节点设为目的节点,根据所述索引目的信息识别与所述索引节点直接关联和/或间接关联的所述目的节点的数

量,并将所述数量发送至所述用户端;

[0026] 所述将所述数量发送至所述用户端之后,所述方法还包括:

[0027] 将所述目标节点的数量上传至区块链中。

[0028] 为实现上述目的,本发明还提供一种数据管理及检索装置,包括:

[0029] 数据输入模块,用于接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据;

[0030] 索引构建模块,用于将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库;

[0031] 检索反馈模块,用于接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情况。

[0032] 为实现上述目的,本发明还提供一种计算机设备,其包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序,所述计算机设备的处理器执行所述计算机程序时实现上述数据管理及检索方法的步骤。

[0033] 为实现上述目的,本发明还提供一种计算机可读存储介质,所述可读存储介质上存储有计算机程序,所述可读存储介质存储的所述计算机程序被处理器执行时实现上述数据管理及检索方法的步骤。

[0034] 本发明提供的数据管理及检索方法、装置、计算机设备及可读存储介质,通过根据提取请求从原始数据库中获取相应的原始数据,实现有针对性的提取原始数据库中的原始数据的技术效果;

[0035] 通过编排原始数据的元数据得到元结构,根据所述元结构和数据节点构建反映所述原始数据与所述元数据之间关联关系的数据索引,因此,仅需从数据索引中获取与所述索引关键词对应的节点并将其设为索引节点获取数据节点,并将数据节点中的原始数据设为反馈信息,实现快速识别出原始数据与元数据之间的关联关系,以便于根据用户的需求通过所述元数据检索出与其关联的原始数据的技术效果;

[0036] 通过基于索引关键词识别与其对应的元节点并获得元数据,将与所述元数据对应的原始数据作为反馈数据的方式,不仅避免了便利数据库中所有元数据及其对应的原始数据,导致服务器算力消耗过大的问题发生,提高了反馈数据的获取效率,并且,由于基于索引关键词获得所有与其相关的原始数据的方式,无需对索引关键词的质量进行把控,即可全面反馈与索引关键词相关的反馈数据,保证了反馈数据的全面性,进而保证用户端获得的反馈数据能够满足其检索需求。

## 附图说明

- [0037] 图1为本发明数据管理及检索方法实施例一的流程图；
- [0038] 图2为本发明数据管理及检索方法实施例二中数据管理及检索方法的环境应用示意图；
- [0039] 图3是本发明数据管理及检索方法实施例二中数据管理及检索方法的具体方法流程图；
- [0040] 图4为本发明数据管理及检索装置实施例三的程序模块示意图；
- [0041] 图5为本发明计算机设备实施例四中计算机设备的硬件结构示意图。

## 具体实施方式

[0042] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本发明，并不用于限定本发明。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0043] 本发明提供的数据管理及检索方法、装置、计算机设备及可读存储介质，适用于大数据的数据处理技术领域，为提供一种基于数据输入模块、索引构建模块、数据检索模块的数据管理及检索方法。本发明通过根据接收到的提取请求从原始数据库中获取原始数据，将原始数据保存至预置的图形数据库中并以数据节点的形式保存，编排原始数据的元数据，得到以元节点的形式反映各元数据之间关联关系的元结构，将元结构和数据节点关联，得到反映原始数据与元数据之间关联关系的数据索引；接收用户端发送的索引关键词，遍历数据索引中与索引关键词对应的元节点并将元节点设为索引节点，识别与索引节点直接关联和/或间接关联的数据节点，提取数据节点中的原始数据并将其作为反馈数据发送至用户端。

[0044] 实施例一：

[0045] 请参阅图1，本实施例的一种数据管理及检索方法，包括：

[0046] S101：接收用户端发送的提取请求，根据所述提取请求从原始数据库中获取原始数据。

[0047] S103：将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存，编排所述原始数据的元数据，得到以元节点的形式反映各所述元数据之间关联关系的元结构，将所述元结构和所述数据节点关联，得到反映所述原始数据与所述元数据之间关联关系的数据索引；其中，图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库。

[0048] S106：接收用户端发送的索引关键词，遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点，识别与所述索引节点直接关联和/或间接关联的数据节点，提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端；其中，所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况，所述间接关联是指所述元节点和所述数据节点之间的关联关系，是经过其他元节点或数据节点所形成的情况。

[0049] 在示例性的实施例中，通过根据用户端发送的提取请求，从所述原始数据库中获



取相应的原始数据,其中,所述提取请求中具有数据名称和数据类型,用于提取所述原始数据库中与所述数据名称和所述数据类型对应的原始数据,实现有针对性的提取原始数据库中的原始数据的技术效果。

[0050] 通过将所述原始数据保存至预置的图形数据库中,编排所述图形数据库中原始数据的元数据得到元结构,根据所述元结构和数据节点构建反映所述原始数据与所述元数据之间关联关系的数据索引,其中,获得的数据索引反映了各原始数据之间的关联关系,因此,仅需从数据索引中获取与所述索引关键词对应的节点并将其设为索引节点,以便于获取与所述索引节点直接关联和/或间接关联的数据节点,并将所述数据节点中的原始数据设为反馈信息;实现快速识别出原始数据与元数据之间的关联关系,以便于根据用户的需求通过所述元数据检索出与其关联的原始数据的技术效果。

[0051] 通过基于索引关键词识别与其对应的元节点并获得元数据,将与所述元数据对应的原始数据作为反馈数据的方式,不仅避免了便利数据库中所有元数据及其对应的原始数据,导致服务器算力消耗过大的问题发生,提高了反馈数据的获取效率,并且,由于基于索引关键词获得所有与其相关的原始数据的方式,无需对索引关键词的质量进行把控,即可全面反馈与索引关键词相关的反馈数据,保证了反馈数据的全面性,进而保证用户端获得的反馈数据能够满足其检索需求。

[0052] 实施例二:

[0053] 本实施例为上述实施例一的一种具体应用场景,通过本实施例,能够更加清楚、具体地阐述本发明所提供的方法。

[0054] 下面,以在运行有数据管理及检索方法的服务器中,将原始数据以数据节点的形式保存并编排原始数据的元数据,得到反映原始数据与元数据之间关联关系的数据索引,以及识别与索引节点直接关联和/或间接关联的数据节点,提取数据节点中的原始数据并将其作为反馈数据为例,来对本实施例提供的方法进行具体说明。需要说明的是,本实施例只是示例性的,并不限制本发明实施例所保护的范围。

[0055] 图2示意性示出了根据本申请实施例二的数据管理及检索方法的环境应用示意图。

[0056] 在示例性的实施例中,数据管理及检索方法所在的服务器2通过网络3连接用户端4;所述服务器2可以通过一个或多个网络3提供服务,网络3可以包括各种网络设备,例如路由器,交换机,多路复用器,集线器,调制解调器,网桥,中继器,防火墙,代理设备和/或等等。网络3可以包括物理链路,例如同轴电缆链路,双绞线电缆链路,光纤链路,它们的组合和/或类似物。网络3可以包括无线链路,例如蜂窝链路,卫星链路,Wi-Fi链路和/或类似物;所述用户端4可为智能手机、平板电脑、笔记本电脑、台式电脑等计算机设备。

[0057] 图3是本发明一个实施例提供的一种数据管理及检索方法的具体方法流程图,该方法具体包括步骤S201至S207。

[0058] S201:接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据。

[0059] 为实现有针对性的提取原始数据库中的原始数据,本步骤根据用户端发送的提取请求,从所述原始数据库中获取相应的原始数据,其中,所述提取请求中具有数据名称和数据类型,用于提取所述原始数据库中与所述数据名称和所述数据类型对应的原始数据。

[0060] 示例性地,所述原始数据是描述实体的数据信息,不同的系统或平台对原始数据的描述通常会采用不同维度的元数据,其中,所述维度可包括:子系统维度、应用运行环境维度、部署逻辑实体维度、集群维度、实例维度、主机维度、存储维度、DB实体维度、Oracle维度、Redis维度、mysql/postgresql维度。

[0061] 元类别为子系统对应的元数据包括:主键(uuid)、子系统ID(idItmisSubSystem)、英文名简称(enNameAbbr)、英文名称(enName)、中文名称(cnName)、系统状态(status)、重要等级(importantGrade)、是否外购(isOutsourcing)、供应商(supplier)、是否外网可访问(isInternet)、系统架构(architecture)、开发语言(devLanguage)、服务窗口(serviceWindow)、子系统简介(subsysDesc)。

[0062] 元类别为部署逻辑实体对应的元数据包括:主键(uuid)、子系统主键(subsysId)、IAMS LE主键(1eId)、部署子系统(dssId)、子系统编码(subsysCode)、部署实体编码(1eCode)、部署实体名称(1eName)、安全区名(zoneName)、LE类型(typeName)、数据库类型(dbType)、调用类型(invokeType)、状态(stateName)、子系统状态(subsysStatus)、类型(1eType)。

[0063] 元类别为应用运行环境对应的元数据包括:主键(uuid)、应用英文名(enName)、应用中文名(areNameCn)、环境类型(environment)、部署名(ARE名)(areName)、状态(status)、是否标准应用(isStandard)、不可标准化(nonStandardization)。

[0064] 元类别为集群对应的元数据包括:主键(uuid)、集群名称(clusterName)、部署单元名称(deployName)、集群软件(clusterSoft)、HA类型(haType)、状态(status)、可用区(availabilityZone)、网络区域(area)、集群VIP(VIP)、VIP端口(vipPort)、IP类型(ipType)、DNS(dns)、DNS类型(dnsType)、Https(Https)、Config(Config)、使用者类型(ownerType)。

[0065] 元类别为实例对应的元数据包括:主键(uuid)、实例名称(instanceName)、实例端口(port)、协议(protocol)、中间件类型(type)。

[0066] 元类别为主机对应的元数据包括:主键(uuid)、子系统标准名(hostSubsystemName)、OS名称(hostName)、NAT IP(natIp)、IP(ipAddress)、NAS IP(nasIp)、OS类别(hostType)、主机区域(中间件专用)(hostArea)、本地磁盘(GB)(localdiskGb)、对外IP(hostOutIp)。

[0067] 元类别为存储对应的元数据包括:主键(uuid)、卷名称(volumeName)、格式(format)、路径(path)、容量(capacity)、存储名称(storageName)、存储类型(storageType)、存储路径(storagePath)、是否挂载(isMounted)。

[0068] 元类别为DB实体对应的元数据包括:主键(uuid)、数据库实体名(entityName)、实体英文名(englishDesc)、实体中文名(chineseDesc)、数据库块大小(blockSize)、是否有备份(isBackup)、是否有GG(isGg)、数据库类型(databaseType)、字符集(characterSet)、数据版本(dataVersion)。

[0069] 元类别为Oracle对应的元数据包括:主键(uuid)、数据库名(databaseName)、环境类型(environment)、数据库唯一名(uniqueName)、状态(Status)、架构类型(architectureName)、存储类型(storageType)、数据库块大小(blockSize)、db\_Domain(dbDomain)、db id(dbid)。

[0070] 元类别为Redis对应的元数据包括：主键(uuid)、数据库名(databaseName)、实例名(instanceName)、实例角色(defaultRole)、实例端口(instancePort)、实例状态(status)、环境类型(enviroment)、HA类型(haType)、实例字符集(characterSet)、VIP(serverVip)、高可用架构(isKeepAlived)。

[0071] 元类别为mysql/postgresql对应的元数据包括：主键(uuid)、数据库名(databaseName)、实例名(instanceName)、实例角色(defaultRole)、实例端口(instancePort)、实例状态(status)、环境类型(enviroment)、HA类型(haType)、实例字符集(characterSet)、VIP(serverVip)。

[0072] 于本实施例中,所述原始数据库是统一CMDB及应用映射(UCMDB)模块(UCMDB全称Universal CMDB),是业务可用性中心底层的支撑模块,主要用于管理存放所有的配置管理信息。这样各个资源项的配置信息均可以永久保存在CMDB里,并能够通过工具的检查保证其数据的准确性,比如发现系统配置的变更等。

[0073] S202:对所述原始数据进行数据清洗,用以删除所述原始数据中的无效数据及其元数据,以及缺失值及其元数据。

[0074] 为避免无效数据和缺失值对数据索引作业造成影响,本步骤通过对原始数据进行数据清洗,以删除无效数据和所述缺失值及其元数据,避免用户端输入的元数据对应的索引数据为无效数据或空值时,获得具有无效数据或缺失值的原始数据导致索引搜索准确度下降的情况发生。

[0075] 在一个优选的实施例中,所述对所述原始数据进行数据清洗的步骤,包括:

[0076] S21:识别所述原始数据中的无效值并查询所述无效值所对应的元数据,删除所述原始数据的无效值及其所述元数据。

[0077] 本步骤中,获取无效标准值,将所述无效标准值与所述原始数据中的数据值进行比对,将与所述无效标准值一致的数据值设为无效值,获取与所述无效值对应的元数据,并删除所述无效值及其元数据,以避免无效值对数据查询作业产生影响。

[0078] S22:识别所述原始数据中的缺失值并查询所述缺失值对应的元数据,删除所述原始数据中的缺失值及其元数据。

[0079] 本步骤中,识别原始数据中值为空值的数据值并将其设为缺失值,获取与所述缺失值对应的元数据,并删除所述缺失值及其元数据,以避免缺失值对数据查询作业产生影响。

[0080] S203:将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库。

[0081] 为识别出原始数据与元数据之间的关联关系,以便于根据用户端的需求通过所述元数据检索出与其关联的原始数据,本步骤通过将所述原始数据保存至预置的图形数据库中,编排所述图形数据库中原始数据的元数据得到元结构,根据所述元结构和数据节点构建反映所述原始数据与所述元数据之间关联关系的数据索引,其中,获得的数据索引反映了各原始数据之间的关联关系,因此,仅需从数据索引中获取与所述索引关键词对应的节

点并将其设为索引节点,以便于获取与所述索引节点直接关联和/或间接关联的数据节点,并将所述数据节点中的原始数据设为反馈信息;其中,所述直接关联是指所述两个节点之间直接具有关联关系,所述间接关联关系是指两个节点之间通过其他节点方可产生关联关系。此时,所述反馈节点中的原始数据将作为反馈数据发送至用户端。

[0082] 需要说明的是,所述图形数据库是NoSQL数据库的一种类型,它应用图形理论存储实体之间的关系信息。图形数据库是一种非关系型数据库,它应用图形理论存储实体之间的关系信息。

[0083] 在一个优选的实施例中,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,包括:

[0084] S31:建立反映元类别之间逻辑关系的维度数据树,根据所述维度数据树构建以元类别作为分类项的结构表;其中,所述元类别是用于对元数据进行分类的使所述元数据形成层级关系的类别信息。

[0085] 本步骤中,所述维度数据树是由用户端开发人员构建的,用于反映各元类别之间逻辑关系的树形数据结构,将所述维度数据树录入预置的关系型数据库中,得到以表格形式显示的结构表。同时,根据需求构建元类别,并构建反映所述元类别之间逻辑关系的维度数据树,使得各元数据可按照用户需求进行编排,以便于构建符合用户端使用场景的数据索引。

[0086] 示例性地:所述维度数据树包括第一层级、第二层级、第三层级和第四层级;所述第一层级的元类别为子系统类别,所述第二层级的元类别为逻辑实体类别,所述第三层级的元类别包括数据库类别、中间件类别和负载类别,所述第四层级的元类别为实例类别。

[0087] 其中,所述子系统类别所对应的元数据包括:英文名简称、中文名称、分摊比例、应用状态、中心分类;

[0088] 所述逻辑实体类别所对应的元数据包括:leCode、le描述、网络区域

[0089] 所述数据库类别所对应的元数据包括:数据库实体名、实体中文名、数据库类型、数据库名、详细版本、数据库环境、数据库实例名、数据库域名、数据库实例端口、数据库VIP、主管DA;

[0090] 所述中间件类别对应的元数据包括:中间件类型(type)、主机区域(中间件专用);

[0091] 所述负载类别对应的元数据包括:集群名称、部署环境、环境类型;

[0092] 所述实例类别对应的元数据包括:应用实例名、应用实例状态、应用主机名、应用主机IP、主机环境。

[0093] S32:提取所述图形数据库中原始数据的元数据,将所述元数据及其在所述原始数据中的对应信息,录入所述结构表中与所述元数据对应的类别项之下得到维度表。

[0094] 本步骤中,从所述图形数据库中各数据节点中获取原始数据,提取所述原始数据中的元数据,将所述元数据录入所述结构表中与所述类别对应的元类别项中。

[0095] 示例性地,将原始数据中与子系统对应的元数据及其所述对应信息,录入所述结构表中的子系统项中;将原始数据中与逻辑实体对应的元数据及其所述对应信息,录入所述结构表中的逻辑实体项中;将原始数据中与数据库对应的元数据及其所述对应信息,录入所述结构表中的数据库项中;将原始数据中与中间件对应的元数据及其所述对应信息,录入所述结构表中的中间项中;将原始数据中与负载对应的元数据及其所述对应信息,录

入所述结构表中的负载项中。

[0096] S33:根据所述维度表构建表征所述元类别及其元数据的元节点,根据所述元类别之间的逻辑关系构建各所述元节点之间关联关系得到元结构。

[0097] 本步骤中,获取所述维度表中元类别项,构建表征所述元类别项下元数据的元节点,根据所述维度表中各所述元类别之间的逻辑关系,构建各所述元节点之间的逻辑关系并形成元结构。

[0098] 示例性地,子系统类别、逻辑实体类别、数据库类别、中间件类别、负载类别、实例类别的分类,分别是按照各元数据在所述原始数据中的对应信息进行分类的,为便于理解,将所述对应信息分别按照1、2、3的序号分类进行举例,例如:如果维度表中,子系统类别的元数据为“中文名称”,与其相应的对应信息为系统1、系统2和系统3,那么,就分别构建子系统类别:中文名称:系统1,子系统类别:中文名称:系统2,子系统类别:中文名称:系统3的元节点。进一步地,在构建元结构时,将元类别作为一级节点、将元类别下的元数据作为二级节点、将元数据的所述对应信息作为所述二级节点的关联连线,用于连接下一层级的一级节点或数据节点,如:将子系统类别为一级节点,将中文名称为二级节点,将系统1、系统2和系统3分别作为所述二级节点的关联连线,用于连接逻辑实体类别的一级节点;以此类推,得到的元结构为:{子系统类别:中文名称:系统1、系统2、系统3}—{逻辑实体类别:leCode:实体1、实体2、实体3}—{数据库类别:数据库实体名:数据库1、数据库2、数据库3,中间件类别:中间件类型:中间件1,中间件2,负载类别:集群名称:负载1,负载2}—{实例类别:应用实例名:实例1,实例2}。

[0099] 在一个优选的实施例中,所述将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引,包括:

[0100] S34:提取所述图形数据库中原始数据的元数据,及所述元数据在所述原始数据中的对应信息,从所述元结构中获得与所述元数据对应的元节点,并将所述元节点设为目标节点。

[0101] 本步骤中,获取所述原始数据的元数据,从所述图形数据库中依次获取与所述元数据对应的元节点。

[0102] 示例性地,假设所述原始数据在子系统类别的中文名称元数据的对应信息为:系统1,在逻辑实体类别的leCode元数据的对应信息为:实体2,在数据库类别的数据库实体名的对应信息为:数据库1,在中间件类别的中间件类型元数据的对应信息为:中间件2,在负载类别的集群名称元数据的对应信息为:负载1,在实例类别的应用实例名元数据的对应信息为:实例2。

[0103] 那么,将获得第一层级目标节点,即:一级节点-子系统,二级节点-中文名称,关联连线为数据库1的元节点1;

[0104] 第二层级目标节点,即:一级节点-逻辑实体类别,二级节点-leCode,关联连线为的实体2的元节点2;

[0105] 第三层级目标节点,即:一级节点-数据库类别,二级节点-数据库实体名,关联连线为数据库1;及

[0106] 一级节点-中间件类别,二级节点-中间件类型,关联连线为中间件2;及

[0107] 一级节点-负载类别,二级节点-集群名称,关联连线为负载1的元节点3;

[0108] 第四层级目标节点,即:一级节点-实例类别,二级节点-应用实例名,关联连线为实例2的元节点4。

[0109] S35:根据所述维度表中的逻辑关系排列所述目标节点形成节点链,将位于所述节点链末端的目标节点与所述数据节点关联得到所述数据索引。

[0110] 示例性地,基于上述举例,将所述元节点1、元节点2、元节点3和元节点4依次通过其中的关联连线连接,将元节点4的关联连线与所述数据节点连接,使所述目标节点与所述数据节点关联,形成所述数据节点的数据索引。

[0111] S204:提取所述数据索引中的元节点;

[0112] 提取所述元节点中的元类别,并对所述元类别设置类别输入框,构建具有所述元类别及其所述类别输入框的类别索引页面,并将用于录入所述索引关键词的所述类别索引页面发送至用户端;或

[0113] 提取所述元节点中的元数据,并对所述元数据设置元输入框,构建具有所述元数据及其所述元输入框的元索引页面,并将用于录入所述索引关键词的所述元索引页面发送至用户端。

[0114] 本步骤中,通过向用户端提供类别索引页面,使用户端可通过在类别输入框中录入索引关键词,即可根据所述索引关键词和与所述类别输入框对应的元类别(例如:子系统类别),获取与所述索引关键词对应的元节点,其中,所述索引关键词与所述元节点中的元类别一致或匹配,提高了索引节点的识别效率。

[0115] 同时,通过向用户端提供元索引页面,使用户端可通过在元输入框中录入索引关键词,即可根据所述索引关键词与所述元输入框对应的元数据(例如:中文名称),获取与所述索引关键词对应的元节点,其中,所述索引关键词与所述元节点中的元数据一致或匹配,提高了索引节点的识别效率。

[0116] S205:提取所述数据索引中的元节点,提取所述元节点中的元类别及其元数据,对所述元数据设置用于启用和禁用所述元数据的选择框,并构建与所述选择框关联的搜索输入框,创制具有所述元类别、所述元数据及其选择框,以及与所述选择框关联的搜索输入框的可选索引页面,将用于录入所述索引关键词的所述可选索引页面发送至用户端。

[0117] 本步骤中,通过向用户端提供可选索引页面,使用户端在搜索输入框内录入索引关键词,并在启用部分元数据且禁用另一部分元数据之后,将根据索引关键词遍历与启用的部分元数据对应的元节点,以获得与所述索引关键词对应的索引节点,提高了索引接点的识别效率。

[0118] S206:接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情况。

[0119] 本步骤中,基于索引关键词识别与其对应的元节点并获得元数据,将与所述元数据对应的原始数据作为反馈数据的方式,不仅避免了便利数据库中所有元数据及其对应的原始数据,导致服务器算力消耗过大的问题发生,提高了反馈数据的获取效率,并且,由于

基于索引关键词获得所有与其相关的原始数据的方式,无需对索引关键词的质量进行把控,即可全面反馈与索引关键词相关的反馈数据,保证了反馈数据的全面性,进而保证用户端获得的反馈数据能够满足其检索需求。

[0120] S207:接收用户端发送的索引关键词、索引目标词和索引目的信息,根据所述索引关键词识别所述数据索引中的元节点并将所述元节点设为索引节点,根据所述索引目的词识别所述数据索引中的元节点或数据节点,并将所述元节点或所述数据节点设为目的节点,根据所述索引目的信息识别与所述索引节点直接关联和/或间接关联的所述目的节点的数量,并将所述数量发送至所述用户端。

[0121] 本步骤中,所述索引关键词是用户端要检索的对象,所述索引目标词是识别与所述对象关联的元节点或数据节点,及所述元节点或所述数据节点的数量,以实现用户端可根据其需求,直观获得图形数据库中各元类别的元数据分布、各元类别的原始数据分布、各元数据下的原始数据分布、各元类别的原始数据分布等情况,扩展了图形数据库及其数据索引的适用范围。

[0122] 示例性地,接收到的索引关键词、索引目标词和索引目的信息是:系统1中有多少个原始数据,其中,索引关键词是“系统1”,索引目标词是“原始数据”,索引目的信息是“有多少个”,那么,将数据索引中“子系统类别-中文名称-系统1”的元节点并将其设为索引节点,根据所述索引目的信息计算与所述索引节点之间直接关联和间接关联的数据节点数量,并将该数量发送至所述用户端。

[0123] 如果接收到的索引关键词和索引目标词是:系统1具有多少个中间件2,因此,索引关键词是“系统1”,索引目标词是“中间件2”索引目的信息是“有多少个”,那么,将数据索引中“子系统类别-中文名称-系统1”的元节点并将其设为索引节点,根据所述索引目的信息计算与所述索引节点之间直接关联和间接关联的,“中间件类别:中间件类型:中间件2”的元节点的数量,并将该数量发送至所述用户端。

[0124] 于本实施例中,所述接收用户端发送的索引关键词和索引目标词,包括:

[0125] 向用户端发送搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面,接收所述用户端通过所述搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面发送的索引关键词;其中,所述搜索索引页面具有索引输入框,用于所述用户端录入所述索引关键词;

[0126] 向所述用户端发送索引目标页面,接收所述用户端通过所述索引目标页面发送的索引目标词;其中,所述索引目标页面中具有目标输入框,用于所述用户端录入所述索引目标词。

[0127] 向所述用户端发送索引目的页面,接收所述用户端通过所述索引目的页面发送的索引目的信息;其中,所述索引目的页面中具有目的输入框,用于所述用户端录入所述索引目的信息。

[0128] 在一个优选的实施例中,所述接收用户端发送的索引关键词和索引目标词,包括:

[0129] S71:向用户端发送搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面,接收所述用户端通过所述搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面发送的索引信息;其中,所述搜索索引页面具有索引输入框,用于所述用户端录入所述索引信息;

[0130] S72:通过预置的自然语言模型识别所述索引信息中的主语信息,并将所述主语信息设为所述索引关键词;

[0131] S73:通过所述自然语言模型识别所述索引信息中的宾语信息,并将所述宾语信息设为所述索引目的词;

[0132] S74:通过所述自然语言模型识别所述索引信息中的谓语信息,并将所述谓语信息设为所述索引目的信息。

[0133] 需要说明的是,所述自然语言模型是具有句法分析算法的自然语言处理算法(NLP),所述句法分析算法用于确定句子的语法结构或句子中词汇之间的依存关系,其中,句法分析算法分为句法结构分析和依存关系分析两种,以获取整个句子的句法结构为目的的称为完全句法分析,而以获得局部成分为目的的语法分析称为局部分析,依存关系分析简称依存分析。通过所述句法分析算法识别索引信息的句法结构,以及其中各词汇之间的依存关系,识别主语信息、谓语信息和宾语信息。于本实施例中,可采用PCFG(结合上下文无关文法)、Lexical PCFG(基于词典的PCFG)、或Transition-based parsing(基于贪心决策动作的拼接句法树)作为所述句法分析算法。

[0134] 优选的,所述将所述数量发送至所述用户端之后,所述方法还包括:

[0135] 将所述目标节点的数量上传至区块链中。

[0136] 需要说明的是,基于目标节点的数量得到对应的摘要信息,具体来说,摘要信息由目标节点的数量进行散列处理得到,比如利用sha256s算法处理得到。将摘要信息上传至区块链可保证其安全性和对用户的公正透明性。用户设备可以从区块链中下载得该摘要信息,以便查证目标节点的数量是否被篡改。本示例所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了一批网络交易的信息,用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0137] 实施例三:

[0138] 请参阅图4,本实施例的一种数据管理及检索装置1,包括:

[0139] 数据输入模块11,用于接收用户端发送的提取请求,根据所述提取请求从原始数据库中获取原始数据;

[0140] 索引构建模块13,用于将所述原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排所述原始数据的元数据,得到以元节点的形式反映各所述元数据之间关联关系的元结构,将所述元结构和所述数据节点关联,得到反映所述原始数据与所述元数据之间关联关系的数据索引;其中,图形数据库是一种通过图形理论存储实体之间的关系信息的非关系型数据库;

[0141] 数据检索模块16,用于接收用户端发送的索引关键词,遍历所述数据索引中与所述索引关键词对应的元节点并将所述元节点设为索引节点,识别与所述索引节点直接关联和/或间接关联的数据节点,提取所述数据节点中的原始数据并将其作为反馈数据发送至所述用户端;其中,所述直接关联是指所述元节点和所述数据节点之间直接具有关联关系的情况,所述间接关联是指所述元节点和所述数据节点之间的关联关系,是经过其他元节点或数据节点所形成的情况。



[0142] 可选的,所述数据管理及检索装置1还包括:

[0143] 数据清洗模块12,用于对所述原始数据进行数据清洗,用以删除所述原始数据中的无效数据及其元数据,以及缺失值及其元数据。

[0144] 可选的,所述数据清洗模块12还包括:

[0145] 无效清洗单元121,用于识别所述原始数据中的无效值并查询所述无效值所对应的元数据,删除所述原始数据的无效值及其所述元数据;

[0146] 缺失清洗单元122,用于识别所述原始数据中的缺失值并查询所述缺失值对应的元数据,删除所述原始数据中的缺失值及其元数据。

[0147] 可选的,所述索引构建模块13还包括:

[0148] 结构构件单元131,用于建立反映元类别之间逻辑关系的维度数据树,根据所述维度数据树构建以元类别作为分类项的结构表;其中,所述元类别是用于对元数据进行分类的使所述元数据形成层级关系的类别信息;

[0149] 维度构建单元132,用于提取所述图形数据库中原始数据的元数据,将所述元数据及其在所述原始数据中的对应信息,录入所述结构表中与所述元数据对应的类别项之下得到维度表;

[0150] 元结构构建单元133,用于根据所述维度表构建表征所述元类别及其元数据的元节点,根据所述元类别之间的逻辑关系构建各所述元节点之间关联关系得到元结构;

[0151] 节点识别单元134,用于提取所述图形数据库中原始数据的元数据,及所述元数据在所述原始数据中的对应信息,从所述元结构中获得与所述元数据对应的元节点,并将所述元节点设为目标节点;

[0152] 索引构建单元135,用于根据所述维度表中的逻辑关系排列所述目标节点形成节点链,将位于所述节点链末端的目标节点与所述数据节点关联得到所述数据索引。

[0153] 可选的,所述数据管理及检索装置1还包括:

[0154] 节点页面构建模块14,用于提取所述数据索引中的元节点;

[0155] 提取所述元节点中的元类别,并对所述元类别设置类别输入框,构建具有所述元类别及其所述类别输入框的类别索引页面,并将用于录入所述索引关键词的所述类别索引页面发送至用户端;或

[0156] 提取所述元节点中的元数据,并对所述元数据设置元输入框,构建具有所述元数据及其所述元输入框的元索引页面,并将用于录入所述索引关键词的所述元索引页面发送至用户端。

[0157] 可选的,所述数据管理及检索装置1还包括:

[0158] 可选页面构建模块15,用于提取所述数据索引中的元节点,提取所述元节点中的元类别及其元数据,对所述元数据设置用于启用和禁用所述元数据的选择框,并构建与所述选择框关联的搜索输入框,创制具有所述元类别、所述元数据及其选择框,以及与所述选择框关联的搜索输入框的可选索引页面,将用于录入所述索引关键词的所述可选索引页面发送至用户端。

[0159] 可选的,所述数据管理及检索装置1还包括:

[0160] 数量检索模块17,用于接收用户端发送的索引关键词、索引目标词和索引目的信息,根据所述索引关键词识别所述数据索引中的元节点并将所述元节点设为索引节点,根

据所述索引目的词识别所述数据索引中的元节点或数据节点,并将所述元节点或所述数据节点设为目的节点,根据所述索引目的信息识别与所述索引节点直接关联和/或间接关联的所述目的节点的数量,并将所述数量发送至所述用户端。

[0161] 可选的,所述数量检索模块17还包括:

[0162] 页面发送单元171,用于向用户端发送搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面,接收所述用户端通过所述搜索索引页面、或所述可选索引页面、或所述类别索引页面、或所述元索引页面发送的索引信息;其中,所述搜索索引页面具有索引输入框,用于所述用户端录入所述索引信息;

[0163] 关键识别单元172,用于通过预置的自然语言模型识别所述索引信息中的主语信息,并将所述主语信息设为所述索引关键词;

[0164] 目的识别单元173,用于通过所述自然语言模型识别所述索引信息中的宾语信息,并将所述宾语信息设为所述索引目的词;

[0165] 索引目的单元174,用于通过所述自然语言模型识别所述索引信息中的谓语信息,并将所述谓语信息设为所述索引目的信息。

[0166] 本技术方案应用于大数据的数据处理领域,通过根据接收到的提取请求从原始数据库中获取原始数据,将原始数据保存至预置的图形数据库中并以数据节点的形式保存,编排原始数据的元数据,得到以元节点的形式反映各元数据之间关联关系的元结构,将元结构和数据节点关联,得到反映原始数据与元数据之间关联关系的数据索引;接收用户端发送的索引关键词,遍历数据索引中与索引关键词对应的元节点并将元节点设为索引节点,识别与索引节点直接关联和/或间接关联的数据节点,进而实现了数据查询的树表查询;提取数据节点中的原始数据并将其作为反馈数据发送至用户端。

[0167] 实施例四:

[0168] 为实现上述目的,本发明还提供一种计算机设备5,实施例三的数据管理及检索装置的组成部分可分散于不同的计算机设备中,计算机设备5可以是执行程序的智能手机、平板电脑、笔记本电脑、台式计算机、机架式服务器、刀片式服务器、塔式服务器或机柜式服务器(包括独立的服务器,或者多个应用服务器所组成的服务器集群)等。本实施例的计算机设备至少包括但不限于:可通过系统总线相互通信连接的存储器51、处理器52,如图5所示。需要指出的是,图5仅示出了具有组件-的计算机设备,但是应理解的是,并不要求实施所有示出的组件,可以替代的实施更多或者更少的组件。

[0169] 本实施例中,存储器51(即可读存储介质)包括闪存、硬盘、多媒体卡、卡型存储器(例如,SD或DX存储器等)、随机访问存储器(RAM)、静态随机访问存储器(SRAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、可编程只读存储器(PROM)、磁性存储器、磁盘、光盘等。在一些实施例中,存储器51可以是计算机设备的内部存储单元,例如该计算机设备的硬盘或内存。在另一些实施例中,存储器51也可以是计算机设备的外部存储设备,例如该计算机设备上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。当然,存储器51还可以既包括计算机设备的内部存储单元也包括其外部存储设备。本实施例中,存储器51通常用于存储安装于计算机设备的操作系统和各类应用软件,例如实施例三的数据管理及检索装置的程序代码等。此外,存储器51还可以用于暂时地存储已经输出或者将要输出的各类数据。

[0170] 处理器52在一些实施例中可以是中央处理器(Central Processing Unit,CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器52通常用于控制计算机设备的总体操作。本实施例中,处理器52用于运行存储器51中存储的程序代码或者处理数据,例如运行数据管理及检索装置,以实现实施例一和实施例二的数据管理及检索方法。

[0171] 实施例五:

[0172] 为实现上述目的,本发明还提供一种计算机可读存储介质,如闪存、硬盘、多媒体卡、卡型存储器(例如,SD或DX存储器等)、随机访问存储器(RAM)、静态随机访问存储器(SRAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、可编程只读存储器(PROM)、磁性存储器、磁盘、光盘、服务器、App应用商城等等,其上存储有计算机程序,程序被处理器52执行时实现相应功能。本实施例的计算机可读存储介质用于存储实现所述数据管理及检索方法的计算机程序,被处理器52执行时实现实施例一和实施例二的数据管理及检索方法。

[0173] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0174] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。

[0175] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

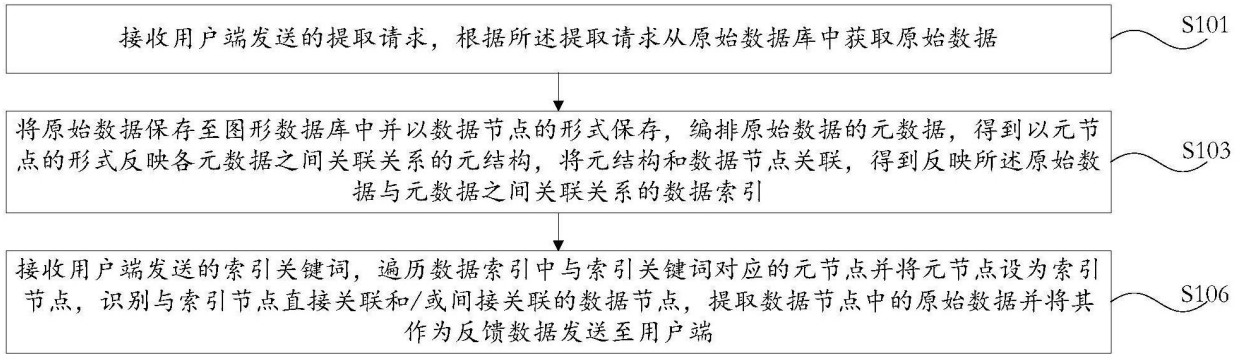


图1

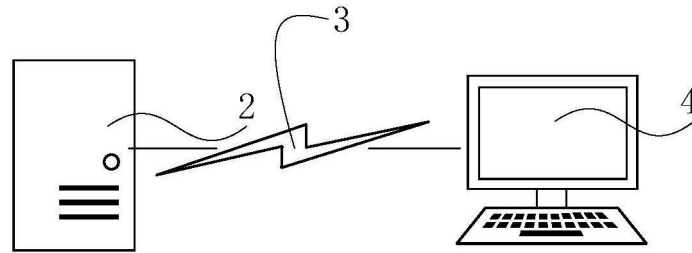


图2

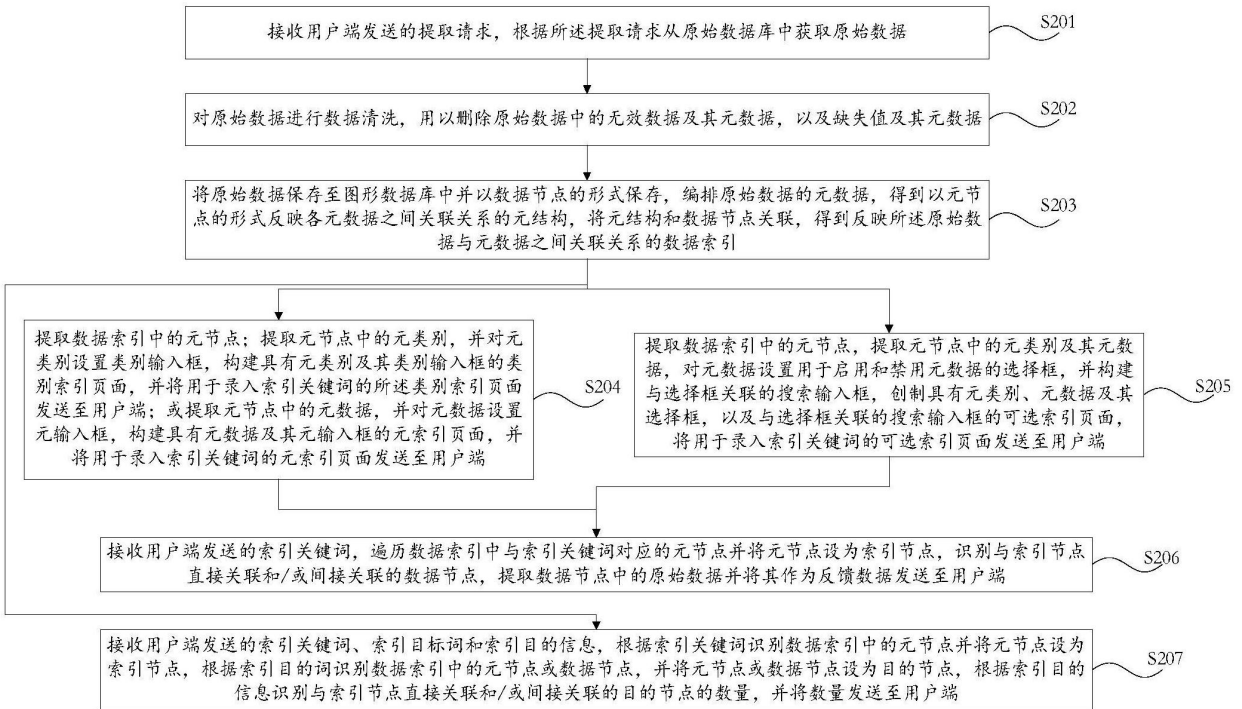


图3

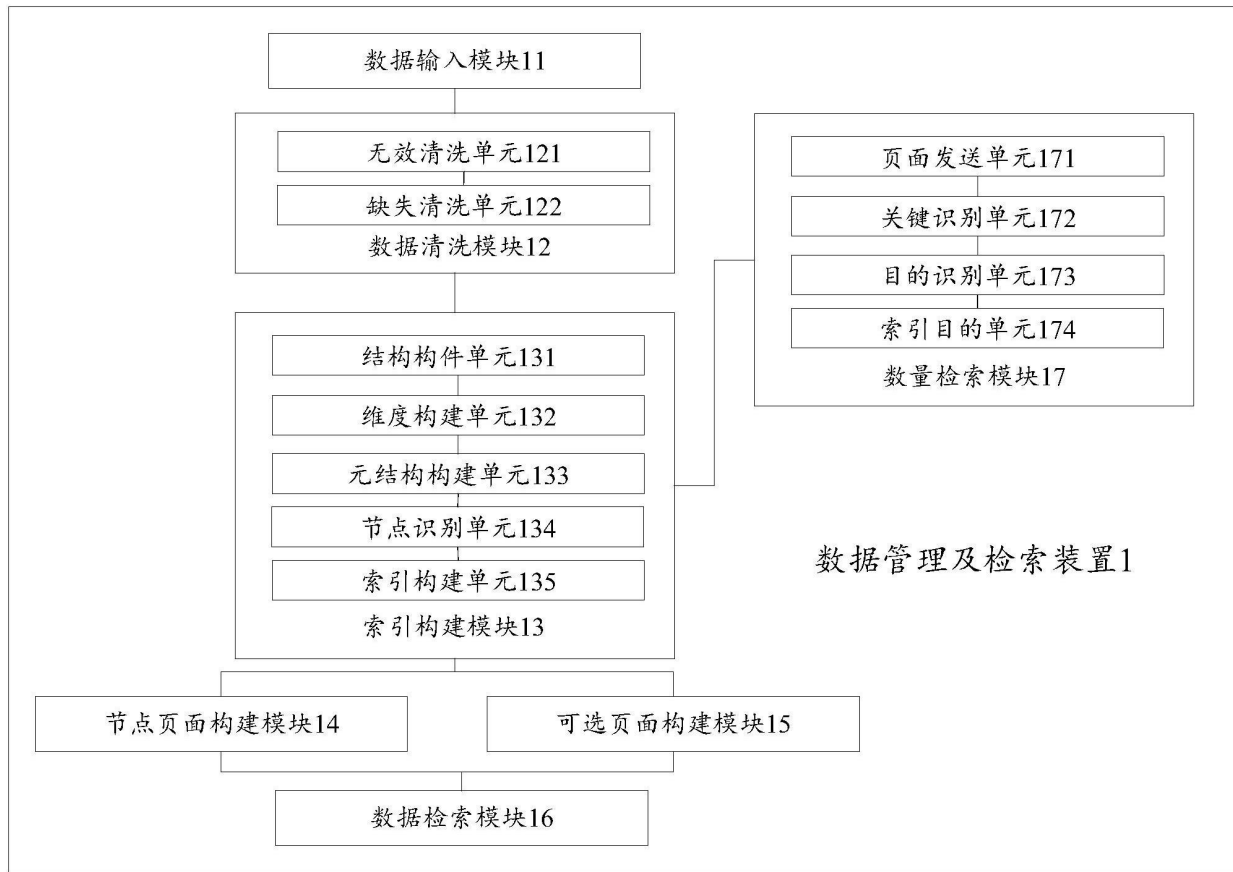


图4

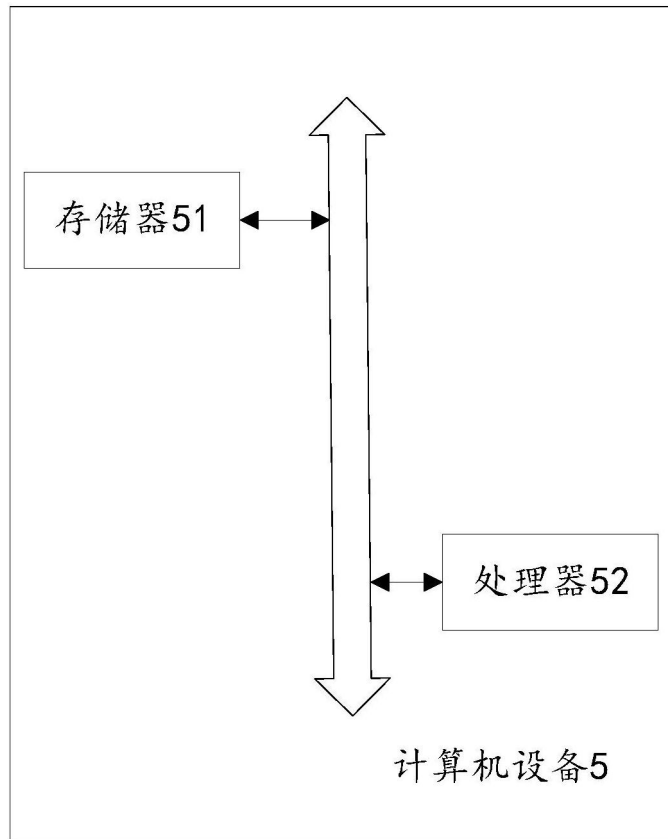


图5