



(12)发明专利申请

(10)申请公布号 CN 106021281 A

(43)申请公布日 2016. 10. 12

(21)申请号 201610281973.6

(22)申请日 2016.04.29

(71)申请人 京东方科技集团股份有限公司
地址 100015 北京市朝阳区酒仙桥路10号

(72)发明人 李慧

(74)专利代理机构 北京同达信恒知识产权代理有限公司 11291

代理人 黄志华

(51) Int. Cl.

G06F 17/30(2006.01)

G06N 5/02(2006.01)

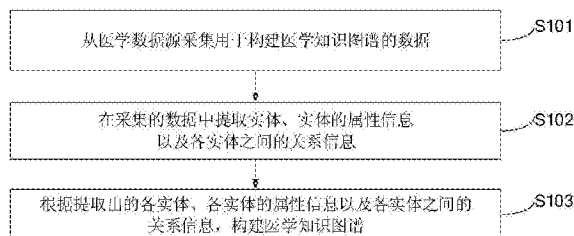
权利要求书2页 说明书7页 附图2页

(54)发明名称

医学知识图谱的构建方法、其装置及其查询方法

(57)摘要

本申请公开了一种医学知识图谱的构建方法、其装置及其查询方法, 通过从医学数据源采集用于构建医学知识图谱的数据; 在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息; 根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息, 构建医学知识图谱。通过上述方式构建的医学知识图谱, 采用非关系型数据存储模式, 更方便医学知识体系的多方向的知识挖掘, 为医护人员提供更直观的参考, 从而降低医疗事故的发生。



1. 一种医学知识图谱的构建方法,其特征在于,包括:
从医学数据源采集用于构建医学知识图谱的数据;
在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息;
根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱。
2. 如权利要求1所述的方法,其特征在于,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱,包括:
根据预设的质量标准对提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息进行筛选;
根据筛选出的符合所述预设的质量标准的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。
3. 如权利要求2所述的方法,其特征在于,还包括:
在提取出的实体、所述实体的属性信息或各所述实体之间的关系信息不符合所述预设的质量标准时,在采集的所述数据中对不符合所述预设的质量标准的数据重新提取实体、所述实体的属性信息以及各所述实体之间的关系信息,直至符合所述预设的质量标准为止。
4. 如权利要求1所述的方法,其特征在于,所述在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,包括:
在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,并将提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息以二维表的形式存储。
5. 如权利要求4所述的方法,其特征在于,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱,包括:
将存储的包括各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息的各二维表转化为图数据;
根据所述图数据构建所述医学知识图谱。
6. 如权利要求1所述的方法,其特征在于,所述从医学数据源采集用于构建医学知识图谱的数据,包括:
从医学指南中采集用于构建所述医学知识图谱的数据。
7. 如权利要求1所述的方法,其特征在于,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱,包括:
根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,以各所述实体或各所述实体的属性信息作为节点,以各所述实体之间的关系信息或各所述实体与属性信息之间的关系信息作为有向线段,构建所述医学知识图谱。
8. 一种利用医学知识图谱的查询方法,其特征在于,包括:
接收用户输入的关键词;
解析出接收的所述关键词中包含的实体信息和/或属性信息;
根据解析出的所述实体信息和/或属性信息在医学知识图谱中查询与所述实体信息

和/或属性信息相关的知识图谱数据,并将查询到的所述知识图谱数据向所述用户显示。

9.如权利要求8所述的方法,其特征在于,所述与所述实体信息和/或属性信息相关的知识图谱数据至少包括以下一种:

与所述实体信息相关的实体和/或属性信息数据;

与所述属性信息相关的实体和/或属性信息数据;

与所述实体信息和/或属性信息相关的互联网数据。

10.一种医学知识图谱的构建装置,其特征在于,包括:

数据采集单元,用于从医学数据源采集用于构建医学知识图谱的数据;

信息提取单元,用于在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息;

医学知识图谱构建单元,用于根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。

11.如权利要求10所述的装置,其特征在于,所述医学知识图谱构建单元,具体用于根据预设的质量标准对提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息进行筛选;根据筛选出的符合所述预设的质量标准的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。

12.如权利要求11所述的装置,其特征在于,所述信息提取单元,还用于在提取出的实体、所述实体的属性信息或各所述实体之间的关系信息不符合所述预设的质量标准时,对不符合所述预设的质量标准的数据重新提取实体、所述实体的属性信息以及各所述实体之间的关系信息,直至符合所述预设的质量标准为止。

13.如权利要求10所述的装置,其特征在于,所述信息提取单元,具体用于在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,并将提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息以二维表的形式存储。

14.如权利要求13所述的装置,其特征在于,所述医学知识图谱构建单元,具体用于将存储的包括各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息的各二维表转化为图数据;根据所述图数据构建所述医学知识图谱。

15.如权利要求10所述的装置,其特征在于,所述数据采集单元,具体用于从医学指南中采集用于构建所述医学知识图谱的数据。

16.如权利要求10所述的装置,其特征在于,所述医学知识图谱构建单元,具体用于根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,以各所述实体或各所述实体的属性信息作为节点,以各所述实体之间的关系信息或各所述实体与属性信息之间的关系信息作为有向线段,构建所述医学知识图谱。

医学知识图谱的构建方法、其装置及其查询方法

技术领域

[0001] 本申请涉及大数据技术领域,尤指一种医学知识图谱的构建方法、其装置及其查询方法。

背景技术

[0002] 知识图谱是一种图结构的知识库,属于知识工程的范畴。不同于普通知识库,知识图谱融合所有学科,将不同来源、不同类型、不同结构的知识单元通过链接关联成图,基于各学科的元数据,为用户提供更广度、更深度的知识体系并不断扩充。其本质上是将领域知识数据体系化、关系化,并以图的方式将知识可视化。简单来说,知识图谱是基于信息系统建立的知识体系,通过数据采集、数据挖掘、信息处理、知识计量和图形绘制等技术把复杂的知识领域系统地显示出来,揭示知识领域的动态发展规律。

[0003] 知识图谱在应用,扩展了原科学知识图谱的内涵,使其应用场景得到延伸。但是目前知识图谱的应用仍局限于搜索引擎及问答系统等方面,其他方面应用较少。医学领域中的病症、疾病与诊疗手段之间通常存在着错综复杂的关系,而现有的关系模型的数据存储模式不便于医学知识体系内容的扩充,也不能为医护人员提供直观的参考。

发明内容

[0004] 本申请实施例提供一种医学知识图谱的构建方法、其装置及其查询方法,方便医学知识体系的多方向的知识挖掘,为医护人员提供更直观的参考,从而降低医疗事故的发生。

[0005] 第一方面,本申请实施例提供一种医学知识图谱的构建方法,包括:

[0006] 从医学数据源采集用于构建医学知识图谱的数据;

[0007] 在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息;

[0008] 根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱。

[0009] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱,包括:

[0010] 根据预设的质量标准对提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息进行筛选;

[0011] 根据筛选出的符合所述预设的质量标准的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。

[0012] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,还包括:

[0013] 在提取出的实体、所述实体的属性信息或各所述实体之间的关系信息不符合所述预设的质量标准时,在采集的所述数据中对不符合所述预设的质量标准的数据重新提取实

体、所述实体的属性信息以及各所述实体之间的关系信息,直至符合所述预设的质量标准为止。

[0014] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,包括:

[0015] 在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,并将提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息以二维表的形式存储。

[0016] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱,包括:

[0017] 将存储的包括各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息的各二维表转化为图数据;

[0018] 根据所述图数据构建所述医学知识图谱。

[0019] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述从医学数据源采集用于构建医学知识图谱的数据,包括:

[0020] 从医学指南中采集用于构建所述医学知识图谱的数据。

[0021] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建医学知识图谱,包括:

[0022] 根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,以各所述实体或各所述实体的属性信息作为节点,以各所述实体之间的关系信息或各所述实体与属性信息之间的关系信息作为有向线段,构建所述医学知识图谱。

[0023] 第二方面,本申请实施例提供一种利用医学知识图谱的查询方法,包括:

[0024] 接收用户输入的关键词;

[0025] 解析出接收的所述关键词中包含的实体信息和/或属性信息;

[0026] 根据解析出的所述实体信息和/或属性信息在医学知识图谱中查询与所述实体信息和/或属性信息相关的知识图谱数据,并将查询到的所述知识图谱数据向所述用户显示。

[0027] 在一种可能的实现方式中,在本申请实施例提供的上述方法中,所述与所述实体信息和/或属性信息相关的知识图谱数据至少包括以下一种:

[0028] 与所述实体信息相关的实体和/或属性信息数据;

[0029] 与所述属性信息相关的实体和/或属性信息数据;

[0030] 与所述实体信息和/或属性信息相关的互联网数据。

[0031] 第三方面,本申请实施例提供一种医学知识图谱的构建装置,包括:

[0032] 数据采集单元,用于从医学数据源采集用于构建医学知识图谱的数据;

[0033] 信息提取单元,用于在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息;

[0034] 医学知识图谱构建单元,用于根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。

[0035] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述医学知识图

谱构建单元,具体用于根据预设的质量标准对提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息进行筛选;根据筛选出的符合所述预设的质量标准的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,构建所述医学知识图谱。

[0036] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述信息提取单元,还用于在提取出的实体、所述实体的属性信息或各所述实体之间的关系信息不符合所述预设的质量标准时,对不符合所述预设的质量标准的数据重新提取实体、所述实体的属性信息以及各所述实体之间的关系信息,直至符合所述预设的质量标准为止。

[0037] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述信息提取单元,具体用于在采集的所述数据中提取实体、所述实体的属性信息以及各所述实体之间的关系信息,并将提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息以二维表的形式存储。

[0038] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述医学知识图谱构建单元,具体用于将存储的包括各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息的各二维表转化为图数据;根据所述图数据构建所述医学知识图谱。

[0039] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述数据采集单元,具体用于从医学指南中采集用于构建所述医学知识图谱的数据。

[0040] 在一种可能的实现方式中,在本申请实施例提供的上述装置中,所述医学知识图谱构建单元,具体用于根据提取出的各所述实体、各所述实体的属性信息以及各所述实体之间的关系信息,以各所述实体或各所述实体的属性信息作为节点,以各所述实体之间的关系信息或各所述实体与属性信息之间的关系信息作为有向线段,构建所述医学知识图谱。

[0041] 本申请有益效果如下:

[0042] 本申请实施例提供的一种医学知识图谱的构建方法、其装置及其查询方法,通过从医学数据源采集用于构建医学知识图谱的数据;在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息;根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。通过上述方式构建的医学知识图谱,采用非关系型数据存储模式,更方便医学知识体系的多方向的知识挖掘,为医护人员提供更直观的参考,从而降低医疗事故的发生。

附图说明

[0043] 图1为本申请实施例中医学知识图谱的构建方法的流程图;

[0044] 图2为本申请实施例中医学知识图谱的示意图;

[0045] 图3为本申请实施例中利用医学知识图谱的查询方法的流程图;

[0046] 图4为本申请实施例中医学知识图谱的构建装置的结构示意图。

具体实施方式

[0047] 针对而现有的关系模型的数据存储模式不便于医学知识体系内容的扩充,也不能为医护人员提供直观的参考,本申请实施例提供一种医学知识图谱的构建方法。

[0048] 具体地,如图1所示,本申请实施例提供的医学知识图谱的构建方法,包括如下步骤:

[0049] S101、从医学数据源采集用于构建医学知识图谱的数据;

[0050] S102、在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息;

[0051] S103、根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。

[0052] 本申请实施例提供的上述医学知识图谱的构建方法,根据采集的数据中的各实体、各实体的属性信息以及各实体之间的关系信息构建医学知识图谱,将医学数据源中大量的文字数据简化,更直观的体现出医学领域的各实体的具有的特征和属性,以及各实体之间的相关性,便于将非知识图谱数据扩充到医学知识图谱中,为医学人员提供直观便利的参考,从而降低医疗事故的发生。

[0053] 如下对上述各步骤进行具体说明。

[0054] 在具体实施时,在上述的步骤S103中,根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱,具体可以包括:

[0055] 根据预设的质量标准对提取出的各实体、各实体的属性信息以及各实体之间的关系信息进行筛选;

[0056] 根据筛选出的符合预设的质量标准的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。

[0057] 由于涉及医学学科,为保证提取的各实体、各实体的属性信息以及各实体之间的关系信息均为可靠数据,需要在根据这些数据构建医学知识图谱之前,对提取出的各实体、各实体的属性信息以及各实体之间的关系信息进行正确性的筛选,在实际应用时,可邀请医学专家或相关领域的权威机构对提取的各实体、各实体的属性信息以及各实体之间的关系信息进行正确性验证。

[0058] 进一步地,在本申请实施例提供的上述方法中,还包括如下步骤:

[0059] 在提取出的实体、实体的属性信息或各实体之间的关系信息不符合预设的质量标准时,在采集的数据中对不符合预设的质量标准的数据重新提取实体、实体的属性信息以及各实体之间的关系信息,直至符合预设的质量标准为止。由于医学领域的数据通常在实体之间存在相互关联的关系,因此,为保证构建的医学知识图谱的完整性,需要对不符合预设的质量标准的数据重新提取实体、实体的属性信息以及各实体之间的关系信息的过程,由此,在采集的数据全部符合上述的质量标准时,执行根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱的步骤。

[0060] 在具体实施时,在上述的步骤S102中,在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息,具体可以包括:

[0061] 在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息,并将提取出的各实体、各实体的属性信息以及各实体之间的关系信息以二维表的形式存储。

[0062] 在本申请实施例提供的上述方法中,需要对采集的数据中提取出实体、实体的属性信息以及各属性之间的关系信息,还需要将提取的信息以二维表的形式存储,例如,可将提取的各实体、各实体的属性信息以及各实体之间的关系信息以excel表的形式进行存储,从而将构建医学知识图谱的数据由繁化简,体现出数据的核心内容,从而可为应用医学图

谱的医护人员提供更直观的参考内容。此外,还可将上述提取的各实体、各实体的属性信息以及各实体之间关系信息存储为其它格式,本申请实施例在此不做限定。

[0063] 进一步地,在上述的步骤S103中,根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱,具体可以包括:

[0064] 将存储的包括各实体、各实体的属性信息以及各实体之间的关系信息的各二维表转化为图数据;

[0065] 根据图数据构建医学知识图谱。

[0066] 在本申请实施例提供的上述方法中,医学领域的实体可为疾病、确诊为某疾病的医疗手段等;实体的属性信息可为疾病的症状、临床表现、确诊信息等;实体与实体之间的关系信息可为两种疾病之间的关联关系、相关性等。将包括各实体、各实体的属性信息以及各实体之间的关系信息的二维表转化为图数据,而图数据为一种非关系型数据存储形式,因此更有利于数据的扩充和挖掘,并且表现形式更为直观,为医护人员提供了便利。在实际应用时,可采用编程的方式将上述的包括各实体、各实体的属性信息以及各实体之间的关系信息的二维表转化进行编程从而转化为图数据。其它转化图数据的方式在此不做限定。

[0067] 如下以一实例说明上述提取实体、实体的属性信息以及各实体之间关系信息的操作过程。

[0068] 由数据源中采集的数据如下:

[0069] “GDM指妊娠期发生的糖代谢异常,妊娠期首次发现且血糖升高已经达到糖尿病标准,应将其诊断为PGDM而非GDM。GDM诊断方法和标准如下:

[0070] 推荐医疗机构对所有尚未被诊断为PGDM或GDM的孕妇,在妊娠24~28周以及28周后首次就诊时行OGTT。

[0071] 75g OGTT方法:OGTT前禁食至少8h,试验前连续3d正常饮食,即每日进食碳水化合物不少于150g,检查期间静坐、禁烟。检查时,5min内口服含75g葡萄糖的液体300ml,分别抽取孕妇服糖前及服糖后1、2h的静脉血(从开始饮用葡萄糖水计算时间),放入含有氟化钠的试管中,采用葡萄糖氧化酶法测定血糖水平。

[0072] 75g OGTT的诊断标准:服糖前及服糖后1h、2h,3项血糖值应分别低于5.1、10.0、8.5mmol/L(92、180、153mg/dl)。任何一项血糖值达到或超过上述标准即诊断为GDM。”

[0073] 在上述数据中提取实体、实体的属性信息以及各实体之间的关系信息后存储的二维表如下所示:

条件		结论
孕周	>24	GDM
空腹血糖	>5.1mmol/L	
餐后1小时血糖	>10.0mmol/L	
餐后2小时血糖	>8.5mmol/L	

[0075] 其中,疾病概念GDM可作为实体,诊断GDM的标准,即上表中的各条件可作为实体,且只有在诊断GDM的各条件同时存在的前提下,诊断结论才为GDM。

[0076] 进一步地,将得到的上述二维表中的有关实体、实体的属性信息以及各实体之间的关系信息的数据进行编程,转化为图数据,从而根据转化的图数据构建医学知识图谱。由于图数据为非关系型数据结构,相比于现有的关系型数据结构,可以更好地体现出数据之间关系类型的多样化,有利于知识的进一步扩充和挖掘。因此,在本申请实施例提供的上述方法中,根据图根据构建医学知识图谱,有利于医学知识体系的多方向的知识挖掘,可为医护人员提供更直观的参考,从而降低医疗事故的发生。

[0077] 在具体实施时,在上述的步骤S101中,从医学数据源采集用于构建医学知识图谱的数据,具体可以包括:

[0078] 从医学指南中采集用于构建医学知识图谱的数据。除此之外,还可将其它具有可靠医学数据的数据源作为本申请实施例中的数据源,在此不做限定。

[0079] 进一步地,在上述的步骤S103中,根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱,具体可以包括:

[0080] 根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,以各实体或各实体的属性信息作为节点,以各实体之间的关系信息或各实体与属性信息之间的关系信息作为有向线段,构建医学知识图谱。例如,根据上述的二维表中,可以表中的结论GDM,以及诊断GDM的条件作为节点,以诊断GDM的条件和GDM之间的关系作为有向线段,构建的医学知识图谱可表示为如图2所示。在最终构建的医学知识图谱中,可以节点表示疾病,以有向线段来表疾病与症状之间的关系、疾病与该疾病的治疗手段之间的关系,以及与该疾病相关疾病之间的关系。根据需要仍可向增加其它概念或主体来作为节点,本申请实施例不对其进行限定。

[0081] 本申请实施例还提供一种利用上述构建出的医学知识图谱的查询方法,如图3所示,本申请实施例提供的利用医学知识图谱的查询方法,包括如下步骤:

[0082] S301、接收用户输入的关键词;

[0083] S302、解析出接收的关键词中包含的实体信息和/或属性信息;

[0084] S303、根据解析出的实体信息和/或属性信息在医学知识图谱中查询与实体信息和/或属性信息相关的知识图谱数据,并将查询到的知识图谱数据向用户显示。

[0085] 具体地,在上述的步骤S303中,与实体信息和/或属性信息相关的知识图谱数据至少包括以下一种:

[0086] 与实体信息相关的实体和/或属性信息数据;

[0087] 与属性信息相关的实体和/或属性信息数据;

[0088] 与实体信息和/或属性信息相关的互联网数据。

[0089] 在实际应用中,用户输入的关键词可为某一疾病名称,或某一症状,那么以该病症作为实体信息的知识图谱的数据、以该症状作为实体信息或属性信息的知识图谱数据将被查询再返回至用户,以供用户参考。

[0090] 基于同一申请构思,本申请实施例提供一种医学知识图谱的构建装置,由于该装置解决问题的原理与前述医学知识图谱的构建方法相似,因此该装置的实施可以参见方法的实施,重复之处不再赘述。

[0091] 本申请实施例提供的一种医学知识图谱的构建装置,如图4所示,包括:

[0092] 数据采集单元41,用于从医学数据源采集用于构建医学知识图谱的数据;

[0093] 信息提取单元42,用于在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息;

[0094] 医学知识图谱构建单元43,用于根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。

[0095] 具体地,医学知识图谱构建单元43,具体用于根据预设的质量标准对提取出的各实体、各实体的属性信息以及各实体之间的关系信息进行筛选;根据筛选出的符合预设的质量标准的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。

[0096] 在一种可能实施的方式中,信息提取单元42,还用于在提取出的实体、实体的属性信息或各实体之间的关系信息不符合预设的质量标准时,对不符合预设的质量标准的数据重新提取实体、实体的属性信息以及各实体之间的关系信息,直至符合预设的质量标准为止。

[0097] 具体地,信息提取单元42,具体用于在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息,并将提取出的各实体、各实体的属性信息以及各实体之间的关系信息以二维表的形式存储。

[0098] 进一步地,医学知识图谱构建单元43,具体用于将存储的包括各实体、各实体的属性信息以及各实体之间的关系信息的各二维表转化为图数据;根据图数据构建医学知识图谱。

[0099] 具体地,数据采集单元41,具体用于从医学指南中采集用于构建医学知识图谱的数据。

[0100] 具体地,医学知识图谱构建单元43,具体用于根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,以各实体或各实体的属性信息作为节点,以各实体之间的关系信息或各实体与属性信息之间的关系信息作为有向线段,构建医学知识图谱。

[0101] 本申请实施例提供的一种医学知识图谱的构建方法、其装置及其查询方法,通过从医学数据源采集用于构建医学知识图谱的数据;在采集的数据中提取实体、实体的属性信息以及各实体之间的关系信息;根据提取出的各实体、各实体的属性信息以及各实体之间的关系信息,构建医学知识图谱。通过上述方式构建的医学知识图谱,采用非关系型数据存储模式,更方便医学知识体系的多方向的知识挖掘,为医护人员提供更直观的参考,从而降低医疗事故的发生。

[0102] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

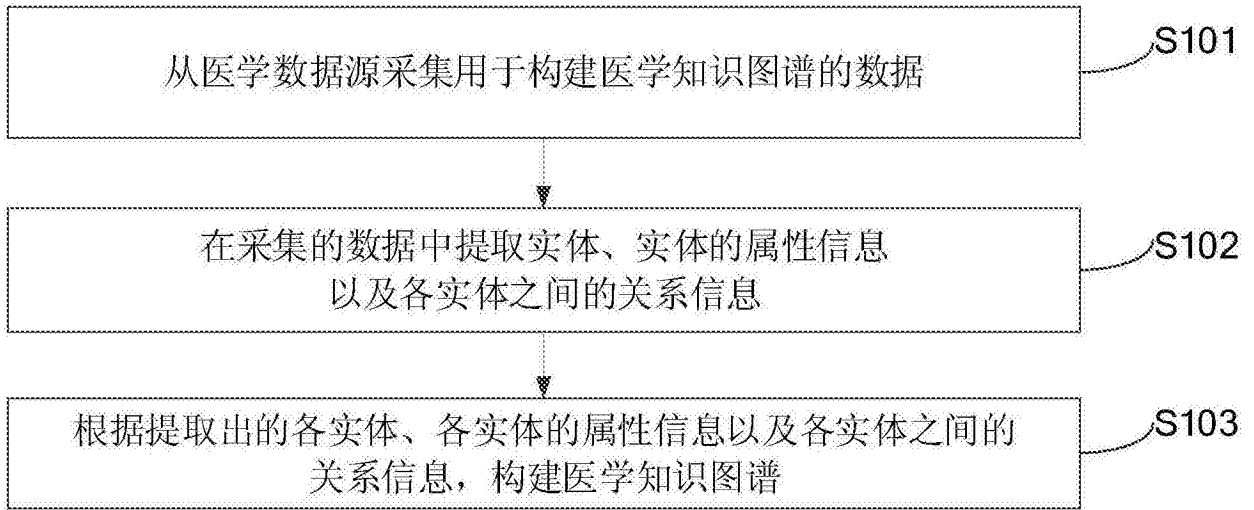


图1

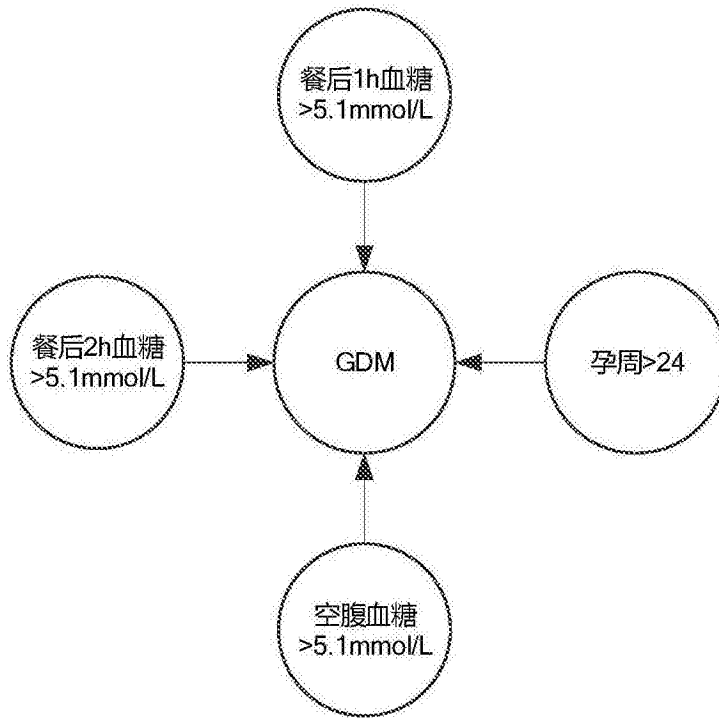


图2

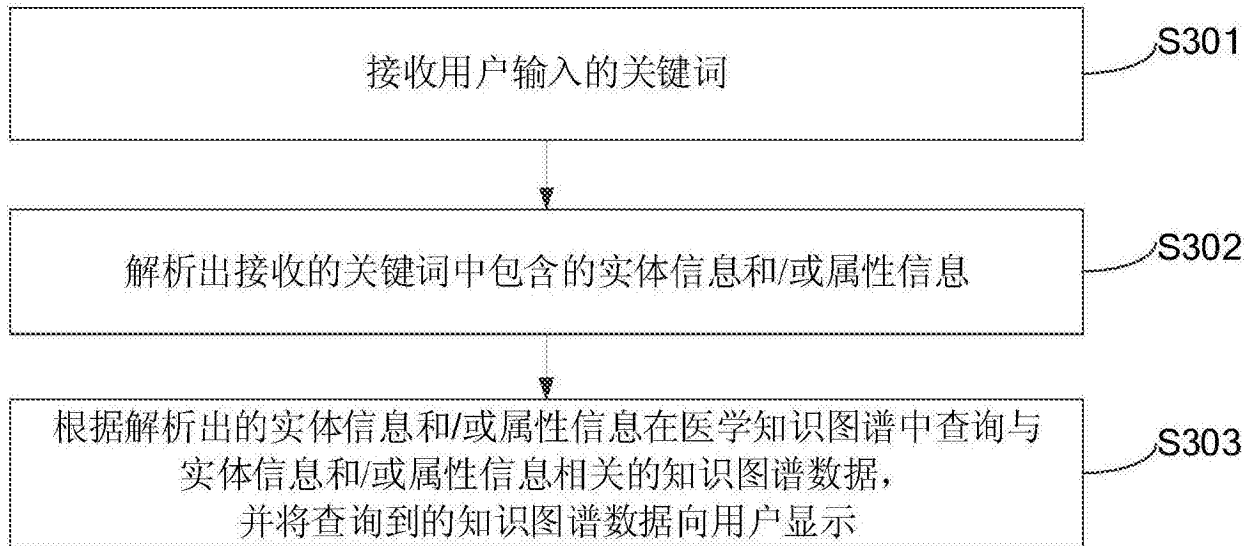


图3

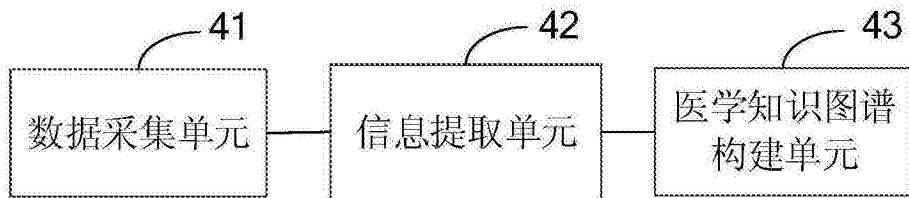


图4