

基于知识图谱的中医养生知识智能问答 APP 构建研究与实践^{*}

李琳菊 李亚春 徐宏宁 韦昌法

(湖南中医药大学信息科学与工程学院 长沙 410208)

[摘要] 分析公众对中医养生知识智能问答 APP 的需求, 详细阐述基于知识图谱的中医养生知识智能问答 APP 设计方案和具体功能实现, 指出其有助于为公众提供权威、可靠、便捷的中医养生保健知识服务, 提升公众养生保健能力。

[关键词] 知识图谱; 中医养生; 智能问答; 应用程序

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2022.07.009

Study and Practice on the Construction of Intelligent Question Answering APP for Traditional Chinese Medicine Health Preserving Knowledge Based on Knowledge Graph LI Linju, LI Yachun, XU Hongning, WEI Changfa, School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China

[Abstract] The paper analyzes the public demand for the intelligent question answering APP for Traditional Chinese Medicine (TCM) health preserving knowledge, elaborates the design scheme and specific function realization of the intelligent question answering APP for TCM health preserving knowledge based on knowledge graph, and points out that it is helpful to provide authoritative, reliable and convenient TCM health preserving knowledge service for the public and improve the public health preserving ability.

[Keywords] knowledge graph; health preservation of Traditional Chinese Medicine (TCM); intelligent question answering; APP

1 引言

中医养生是指在中医理论指导下通过中医疗法颐养生命、增强体质、预防疾病, 从而达到延年益寿目标的医事活动^[1]。知识图谱是人工智能领域的一项重要技术, 具有知识语义化、易关联、易扩充和可视

化等优点, 逐渐成为推动智慧医疗的关键技术之一^[2]。知识图谱有助于中医药知识的表达、关联、整合、共享和应用, 进而实现智能化、个性化的中医药知识服务, 在中医药领域具有广阔的应用前景^[3-4]。公众具有了解、学习中医养生知识的需求, 但缺少快捷获取权威可靠养生知识的渠道^[5-6]。因此开发基于知识图谱的中医养生知识智能问答 APP 有助于提供

[修回日期] 2021-10-09

[作者简介] 李琳菊, 本科生; 通信作者: 韦昌法, 博士, 教授, 发表论文 20 余篇, 主编、参编教材 6 部。

[基金项目] 国家级大学生创新创业训练计划项目“基于知识图谱的中医养生知识智能问答系统构建研究与实现”(项目编号: S202010541053); 湖南省自然科学基金资助项目“基于动态不确定因果图的中医智能辅助辨证知识表示与推理研究”(项目编号: 2020JJ4461); 湖南省教育厅资助科研项目“基于知识图谱的中医智能辅助辨证知识表示与推理研究”(项目编号: 20B431)。

权威、可靠、便捷的中医养生知识服务。

2 公众对中医养生知识智能问答 APP 的需求

2.1 政策推动

2021年6月国家中医药管理局、中央宣传部、教育部、国家卫生健康委员会和国家广电总局共同制定并印发《中医药文化传播行动实施方案(2021-2025年)》，该方案的主要目标之一是“中医药文化更广泛融入群众生产生活，中医药养生保健知识和方法更便捷服务群众健康需要”，重点任务之一是“通过中医养生保健体验设备，帮助群众更加经常接触到规范的中医药养生保健知识”^[7]。

2.2 现有服务无法满足公众养生知识需求

中医养生在我国有着悠久的历史 and 深厚的文化积淀。随着我国经济的持续快速发展，人民群众的生活水平大幅度提高，对健康长寿和高质量生活的追求日益提升，对中医养生知识服务的需求进一步增强。目前介绍中医养生保健知识的书籍和报刊数量较多但是质量良莠不齐。大众通过上网搜索、阅读微信公众号文章等方式获得中医养生保健知识，但是无法确认知识可信度。现有养生应用程序无法精准、高效、便捷地满足用户的养生知识需求。因此日益增长的养生知识需求同落后的养生知识服务之间存在矛盾，严重影响公众利用中医养生知识开展自我防病、养生保健的积极性和成效。

2.3 中医养生知识智能问答 APP

中医养生知识智能问答 APP 是积极响应国家政策、服务群众健康需要的中医养生保健体验设备。开发基于知识图谱的中医养生知识智能问答 APP，从权威渠道采集中医养生保健知识进而提供便捷的养生知识服务，对提升广大人民群众养生保健能力、守护百姓健康具有重要作用和意义。

3 系统设计

3.1 系统架构

基于知识图谱的中医养生知识智能问答系统以

Web 系统和 APP 系统两种外在形式呈现给用户。Web 系统通过电脑网页端向用户提供中医养生知识智能问答服务，而 APP 系统则通过手机端向用户提供服务，两种方式相互补充，便捷地为用户提供服务，见图 1。

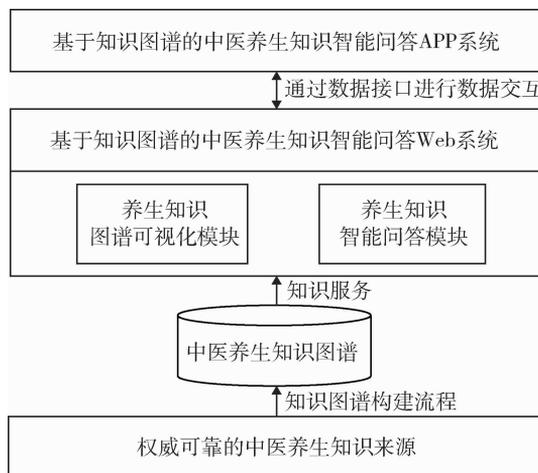


图 1 基于知识图谱的中医养生知识智能问答系统架构

3.2 系统实现

基于知识图谱的中医养生知识智能问答 APP 系统是在 Web 系统基础上实现的。基于权威可靠的中医养生知识来源构建中医养生知识图谱，为 Web 系统的图谱可视化模块和智能问答模块提供数据支持，Web 系统对外提供数据访问接口，APP 系统则通过访问 Web 系统提供的数据接口实现相关功能。

3.3 核心模块

3.3.1 中医养生知识图谱构建模块 通过人机结合梳理书籍知识和智能采集网页知识两种方式获取中医养生知识，进而构建知识图谱。中医养生知识来源较多，其中两个最重要的来源是以全国中医药行业高等教育“十三五”规划教材《中医养生学》^[1]为代表的中医养生权威书籍，以及以名老中医学学术经验国家服务平台中的养生保健精品资源库^[8]为代表的中医养生权威网站。人机结合梳理书籍知识是指通过计算机程序对中医养生权威书籍中的知识进行梳理并辅以人工核验，通过这种方式获

得的养生知识可以直接进行结构化处理；而智能采集网页知识是指通过计算机程序智能采集中医养生权威网站上的养生知识，获取养生知识的速度更快，但是在对获取到的知识进行结构化处理之前需进行筛选。智能采集网页知识主要是使用 lxml 和 Xpath 的语法来解析并获取网页数据。lxml 是 Python 第 3 方库，能够对 html 元素进行精准定位取值，通常用于解析并提取网页数据。XPath 是一种在 html/xml 文档中查找信息的语言，可以遍历 html 或 xml 的元素和属性。智能采集网页知识业务的实现逻辑是通过 Python 第 3 方库 requests 来得到网页信息，使用 lxml 定位工具将网页解析成 html 数据，接着使用 XPath 从 html 数据中将所需数据提取出来。数据经过清洗后，整理成特定格式的知识条目，保存在逗号分隔值（Comma - Separated Values, CSV）格式文档中。

3.3.2 中医养生知识图谱可视化模块 根据用户需求进行养生知识检索并予以可视化展示，用户输入需要检索的中医养生关键词，APP 系统将用户输入的关键词发送给 Web 系统，Web 系统接收后在知识图谱中进行精确检索，再将检索结果通过数据接口传递给 APP 系统，最终由 APP 系统将结果直观地展示给用户。

3.3.3 中医养生知识智能问答模块 根据用户以语音或文字方式输入的问题进行理解分析、抽取关键词，将关键词发送给 Web 系统，Web 系统在知识图谱中进行模板匹配、搜索答案，将答案通过数据接口传递给 APP 系统，APP 系统最终将答案以语音或文字的方式返回给用户，见图 2。智能问答流程主要包括以下步骤：（1）用户通过语音或文字输入要询问的问题。（2）系统后台接收用户问句，对问句采用完全匹配的方式抽取关键词。（3）判断第 2 步是否成功抽取到关键词，是则进入第 5 步，否则进入第 4 步。（4）利用 jieba 分词工具对问句进行分词，采用模糊匹配的方式通过相似度计算抽取关键词，然后进入第 5 步。（5）判断前期抽取到的关键词是否唯一，是则进入第 7 步，否则进入第 6 步。（6）前期抽取到的关键词有多个，存在关键词重复的可能，需进行关键词去重处理，然后进入第 7

步。（7）从用户问句中获取方向词和疑问词。（8）对方向词进行匹配，获取方向词查询语句。（9）在中医养生知识图谱中进行查询，获取与方向词相关的节点和关系数据。（10）结合关键词和疑问词对上一步获取到的数据进行字符串拼接，形成问题答案并返回给用户。

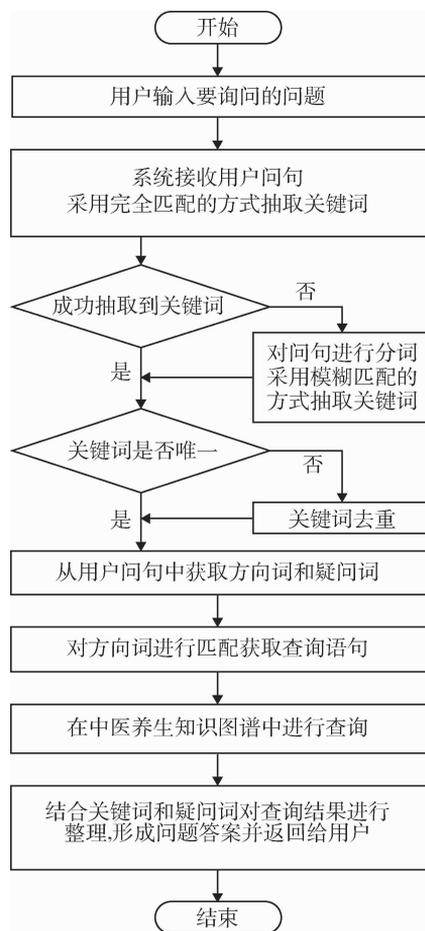


图 2 中医养生知识智能问答模块业务流程

4 系统功能实现

4.1 开发语言及工具

基于知识图谱的中医养生知识智能问答 APP 采用的开发语言是 Java、Python 和 Html，开发工具是 Android studio 和 Pycharm。

4.2 中医养生知识图谱构建模块

该 APP 所使用的养生知识来源于中医养生权威书籍和网站，通过人机结合梳理书籍知识和智能采

集网页知识这两种方式获取到中医养生知识后,需要对其进行结构化处理,抽取中医养生知识关键词(如水浴、温泉浴等),构成中医养生知识节点表,见图 3。再将各中医养生知识关键词通过某种关系(如概念、分类、作用等)联系起来,构成中医养生知识关系,见表 1。例如“水浴”和“温泉浴”这两个中医养生知识关键词通过关系词“分类”联系起来,表达的含义是“温泉浴”是“水浴”的一种类型;“水浴”分别和“化学效应”“机械作用”“温度刺激”“清洁皮肤”这 4 个中医养生知识关键词通过关系词“作用”联系起来,表达的含义是“水浴”的作用包括了“化学效应”“机械作用”“温度刺激”和“清洁皮肤”。将整理好的中医养生知识节点表和关系表存储在 CSV 文件中,最后通过 Cypher 语言批量导入到 Neo4j 数据库中,构建并存储中医养生知识的知识图谱。开发者可以在 Neo4j 数据库平台中对知识图谱进行查看和管理;而普通用户可以通过中医养生知识智能问答 APP 的图谱可视化功能对知识图谱进行检索和查看。

中医养生知识关键词
水浴
以水为介质,利用水温、浮力、压力、冲击力和所含的特殊化学成分对人体产生作用的沐浴方法
温泉浴
蒸气浴
冷水浴
热水浴
化学效应
机械作用
温度刺激
清洁皮肤
水浴后应避风寒,注意保暖
空腹、饱腹、醉酒后和过度疲劳时不宜进行

图 3 中医养生知识节点表示例

表 1 中医养生知识关系表示例

中医养生知识关键词	关系词	中医养生知识关键词
水浴	概念	以水为介质,利用水温、浮力、压力、冲击力和所含的特殊化学成分对人体产生作用的沐浴方法
水浴	分类	温泉浴
水浴	分类	蒸气浴
水浴	分类	冷水浴
水浴	分类	热水浴
水浴	作用	化学效应
水浴	作用	机械作用
水浴	作用	温度刺激
水浴	作用	清洁皮肤
水浴	注意	水浴后应避风寒,注意保暖
水浴	注意	空腹、饱腹、醉酒后和过度疲劳时不宜进行

4.3 中医养生知识图谱可视化模块

中医养生知识图谱可视化模块根据用户输入的检索词在养生知识图谱中进行检索,对检索结果予以可视化展示。例如当用户想了解“水浴”相关知识时,可以在 APP 系统查询页面搜索栏中输入“水浴”,APP 系统将该检索词发送给 Web 系统,Web 系统在中医养生知识图谱中进行检索,将检索结果封装渲染成接口,APP 系统利用 Webview 访问该接口并将结果展示给用户。系统将中医养生知识以图谱和表格形式展示给用户,有助于用户理解和记忆。

4.4 中医养生知识智能问答模块

中医养生知识智能问答模块对用户以语音或文字方式输入的问题进行自动分析并返回答案。智能问答实例的内部流程如下:(1)用户通过文字输入要询问的问题,即“我最近经常失眠,听说菊花枕可以治疗失眠,但是我不知道做一个菊花枕需要哪些原材料”。(2)系统后台接收用户问句,对问句采用完全匹配的方式抽取关键词,关键词有“菊花”“菊花枕”。(3)判断第 2 步成功抽取了关键

词, 进入第 4 步。(4) 判断前期抽取到的关键词不唯一, 进入第 5 步。(5) 前期抽取到的关键词有多个, 存在关键词重复的可能, 需进行关键词去重处理, 然后进入第 6 步。(6) 从用户问句中获取方向词和疑问词, 方向词为“原材料”, 疑问词为“哪些”。(7) 对方向词“原材料”进行匹配, 获取方向词查询语句。(8) 在中医养生知识图谱中进行查询, 获取与方向词“原材料”相关的节点和关系数据。(9) 结合关键词“菊花枕”和疑问词“哪些”对上一步获取到的数据进行字符串拼接, 形成问题答案并返回给用户。

5 结语

本文分析公众对中医养生知识智能问答 APP 的需求, 介绍中医养生知识智能问答 APP 设计方案和实现过程, 展示中医养生知识图谱构建、中医养生知识图谱可视化、中医养生知识智能问答模块实现效果。目前基于知识图谱的中医养生知识智能问答 APP 已获得国家版权局计算机软件著作权登记证书。构建中医养生知识智能问答 APP, 从权威渠道采集可靠的中医养生保健知识, 为公众提供可信、可靠、便捷的中医养生保健知识服务, 既积极响应国家相关政策又切实服务于大众健康需求, 对提升公众养生保健能力、守护公众健康具有重要作用和

意义。

参考文献

- 1 马烈光, 蒋力生. 中医养生学 (第 10 版) [M]. 北京: 中国中医药出版社, 2018.
- 2 周海, 彭斐灵, 韦昌法. 中医诊断学辨证知识图谱构建研究与实践 [J]. 医学信息学杂志, 2020, 41 (12): 41-44.
- 3 孙华君, 李海燕, 聂莹, 等. 知识图谱及其在中医药领域应用研究进展 [J]. 世界科学技术-中医药现代化, 2020, 22 (6): 1969-1974.
- 4 李新龙, 刘岩, 何丽云, 等. 知识图谱研究概况及其在中医药领域的应用 [J]. 中国中医药信息杂志, 2017, 24 (7): 129-132.
- 5 吉彬彬, 林信冰, 唐陶然, 等. 社区中老年人中医养生知信行及服务需求的调查与分析 [J]. 湖南中医药大学学报, 2020, 40 (6): 768-771.
- 6 汪洋, 冯梅, 邓沂, 等. 皖南地区居民中医养生保健知识获取途径的调查分析 [J]. 中国中医药现代远程教育, 2018, 16 (12): 55-57.
- 7 国家中医药管理局. 中医药文化传播行动实施方案 (2021-2025 年) [EB/OL]. [2021-07-26]. <http://bgs.satcm.gov.cn/zhengcewenjian/2021-07-07/22232.html>.
- 8 世界中医药学会联合会. 名老中医学术经验国家服务平台养生保健精品资源库 [EB/OL]. [2021-07-26]. <http://www.gjmlzy.com:83/KBJPYSBJ/Main/Index>.

《医学信息学杂志》版权声明

(1) 作者所投稿件无“抄袭”“剽窃”“一稿两投或多投”等学术不端行为, 对于署名无异议, 不涉及保密与知识产权的侵权等问题, 文责自负。对于因上述问题引起的一切法律纠纷, 完全由全体署名作者负责, 无需编辑部承担连带责任。(2) 来稿刊用后, 该稿包括印刷出版和电子出版在内的出版权、复制权、发行权、汇编权、翻译权及信息网络传播权已经转让给《医学信息学杂志》编辑部。除以纸载体形式出版外, 本刊有权以光盘、网络期刊等其他方式刊登文稿, 本刊已加入万方数据“数字化期刊群”、重庆维普“中文科技期刊数据库”、清华同方“中国期刊全文数据库”、中邮阅读网。(3) 作者著作权使用费与本刊稿酬一次性给付, 不再另行发放。作者如不同意文章入编, 投稿时敬请说明。

《医学信息学杂志》编辑部