



(12)发明专利申请

(10)申请公布号 CN 106776808 A

(43)申请公布日 2017.05.31

(21)申请号 201611050900.2

(22)申请日 2016.11.23

(71)申请人 百度在线网络技术(北京)有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦三层

(72)发明人 饶俞卉子 李鹏 乔芳 孙晶晶
刘露 张涛

(74)专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201
代理人 宋合成

(51)Int.Cl.
G06F 17/30(2006.01)

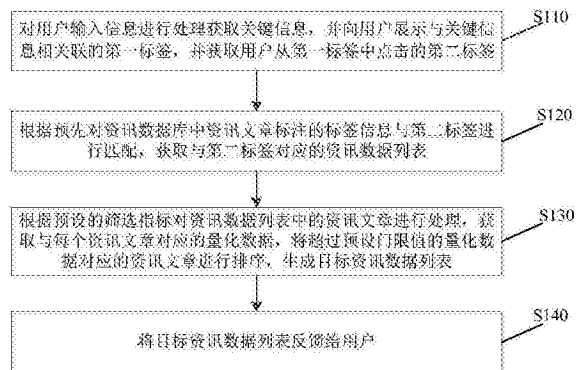
权利要求书3页 说明书13页 附图4页

(54)发明名称

基于人工智能的资讯数据提供方法及装置

(57)摘要

本发明提出一种基于人工智能的资讯数据提供方法及装置,其中,方法包括:对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签;根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表;根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;将目标资讯数据列表反馈给用户。由此,能够为用户提供全面性的资讯,以及针对不同的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。



1. 一种基于人工智能的资讯数据提供方法,其特征在于,包括以下步骤:

对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;

根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;

根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;

将所述目标资讯数据列表反馈给所述用户。

2. 如权利要求1所述的方法,其特征在于,在所述根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配之前,还包括:

将所述资讯数据库中资讯文章进行切词得到多个词语;

根据所述多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重;

对所述多个词语进行去重处理获取目标词语;

根据所述目标词语及其对应的权重为所述资讯文章标注对应的标签信息。

3. 如权利要求1所述的方法,其特征在于,所述根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表,包括:

将所述第二标签与所述资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章;

将预先对所述候选资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表。

4. 如权利要求1所述的方法,其特征在于,所述根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表,包括:

根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;

将超过预设门限值的的第一量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

5. 如权利要求1所述的方法,其特征在于,所述根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将所述量化数据与预设门限值进行比较,生成目标资讯数据列表,包括:

根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;

将超过预设门限值的的第一量化数据对应的资讯文章进行排序,生成第一资讯数据列表;

根据预设的第二筛选指标对所述第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据;

将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

6. 如权利要求5所述的方法,其特征在于,所述预设的第一筛选指标包括:网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的至少一种;

所述预设的第二筛选指标包括:资讯文章的时效性、资讯文章的标签、资讯文章的点击率、所述用户的地理位置和资讯文章的来源中的至少一种。

7. 如权利要求1-6任一所述的方法,其特征在于,在将所述目标资讯数据列表反馈给所述用户之后,还包括:

接收用户对所述目标资讯数据列表中的目标资讯文章的点击操作;

向所述用户显示所述目标资讯文章的内容以及对应的标签信息。

8. 一种基于人工智能的资讯数据提供装置,其特征在于,包括:

第一处理模块,对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;

第二处理模块,用于根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;

第三处理模块,用于根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;

反馈模块,用于将所述目标资讯数据列表反馈给所述用户。

9. 如权利要求8所述的装置,其特征在于,还包括:

切词模块,用于将所述资讯数据库中资讯文章进行切词得到多个词语;

确定模块,用于根据所述多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重;

去重模块,用于对所述多个词语进行去重处理获取目标词语;

标注模块,用于根据所述目标词语及其对应的权重为所述资讯文章标注对应的标签信息。

10. 如权利要求8所述的装置,其特征在于,所述第二处理模块,用于:

将所述第二标签与所述资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章;

将预先对所述候选资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表。

11. 如权利要求8所述的装置,其特征在于,所述第三处理模块,用于:

根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;

将超过预设门限值的第一量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

12. 如权利要求8所述的装置,其特征在于,所述第三处理模块,还用于:

根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;

将超过预设门限值的第一量化数据对应的资讯文章进行排序,生成第一资讯数据列表;

根据预设的第二筛选指标对所述第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据;

将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

13. 如权利要求12所述的装置,其特征在于,所述预设的第一筛选指标包括:网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的至少一种;

所述预设的第二筛选指标包括:资讯文章的时效性、资讯文章的标签、资讯文章的点击率、所述用户的地理位置和资讯文章的来源中的至少一种。

14. 如权利要求8-13任一所述的装置,其特征在于,还包括:

接收模块,用于接收用户对所述目标资讯数据列表中的目标资讯文章的点击操作;

显示模块,用于向所述用户显示所述目标资讯文章的内容以及对应的标签信息。

基于人工智能的资讯数据提供方法及装置

技术领域

[0001] 本发明涉及互联网应用技术领域,尤其涉及一种基于人工智能的资讯数据提供方法及装置。

背景技术

[0002] 人工智能(Artificial Intelligence),英文缩写为AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。人工智能是计算机科学的一个分支,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器,该领域的研究包括机器人、语音识别、图像识别、自然语言处理和专家系统等。其中,人工智能最重要的方面就是语音识别技术。

[0003] 随之智能终端设备的发展,用户可以很方便地使用智能终端设备查看各种各样的内容,尤其是资讯文章。因此,如何利用已有的资讯数据池,满足搜索用户的资讯需求变得越来越重要。

[0004] 相关技术中,通过个别数据资源方定向提交对接的方式、或者通过时效性产品直接抓取新闻还是直接给用户其他网站的地址,让用户去第三方网站进行阅读等。

[0005] 然而,上述方式中让原本在搜索的用户进入产品设计样式不一样,内容质量无保证的网站,给用户的体验非常不一致,增加了搜索内容阅读的难度。另外,不能够为用户提供全面性的资讯,以及无法针对不用的用户提供个性化的资讯,不能满足各个用户对于资讯的个性化需求。

发明内容

[0006] 本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

[0007] 为此,本发明的第一个目的在于提出一种基于人工智能的资讯数据提供方法,该方法能够为用户提供全面性的资讯,以及针对不用的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。

[0008] 本发明的第二个目的在于提出一种基于人工智能的资讯数据提供装置。

[0009] 本发明的第三个目的在于提出另一种基于人工智能的资讯数据提供装置。

[0010] 本发明的第四个目的在于提出一种非临时性计算机可读存储介质。

[0011] 本发明的第五个目的在于提出一种计算机程序产品。

[0012] 为达上述目的,本发明第一方面实施例提出了一种基于人工智能的资讯数据提供方法,包括:对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;将所述目标资讯数据列表反馈给所述用户。

[0013] 本发明实施例的基于人工智能的资讯数据提供方法,通过对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,以及获取用户从第一标签中点击的第二标签,再根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表,最后根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表反馈给用户。由此,能够为用户提供全面性的资讯,以及针对不用的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。

[0014] 另外,根据本发明上述实施例的基于人工智能的资讯数据提供方法还可以具有如下附加的技术特征:

[0015] 在本发明的一个实施例中,在所述根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配之前,还包括:将所述资讯数据库中资讯文章进行切词得到多个词语;根据所述多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重;对所述多个词语进行去重处理获取目标词语;根据所述目标词语及其对应的权重为所述资讯文章标注对应的标签信息。

[0016] 在本发明的一个实施例中,所述根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表,包括:将所述第二标签与所述资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章;将预先对所述候选资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表。

[0017] 在本发明的一个实施例中,所述根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表,包括:根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的所述第一量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0018] 在本发明的一个实施例中,所述根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将所述量化数据与预设门限值进行比较,生成目标资讯数据列表,包括:根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的所述第一量化数据对应的资讯文章进行排序,生成第一资讯数据列表;根据预设的第二筛选指标对所述第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据;将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0019] 在本发明的一个实施例中,所述预设的第一筛选指标包括:网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的至少一种;所述预设的第二筛选指标包括:资讯文章的时效性、资讯文章的标签、资讯文章的点击率、所述用户的地理位置和资讯文章的来源中的至少一种。

[0020] 在本发明的一个实施例中,在将所述目标资讯数据列表反馈给所述用户之后,还包括:接收用户对所述目标资讯数据列表中的目标资讯文章的点击操作;向所述用户显示所述目标资讯文章的内容以及对应的标签信息。

[0021] 为达上述目的,本发明第二方面实施例提出了一种基于人工智能的资讯数据提供装置,包括:第一处理模块,对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;第二处理模块,用于根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;第三处理模块,用于根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;反馈模块,用于将所述目标资讯数据列表反馈给所述用户。

[0022] 本发明实施例的基于人工智能的资讯数据提供装置,通过对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,以及获取用户从第一标签中点击的第二标签,再根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表,最后根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表反馈给用户。由此,能够为用户提供全面性的资讯,以及针对不同的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。

[0023] 另外,根据本发明上述实施例的资讯数据提供装置还可以具有如下附加的技术特征:

[0024] 在本发明的一个实施例中,所述的装置还包括:切词模块,用于将所述资讯数据库中资讯文章进行切词得到多个词语;确定模块,用于根据所述多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重;去重模块,用于对所述多个词语进行去重处理获取目标词语;标注模块,用于根据所述目标词语及其对应的权重为所述资讯文章标注对应的标签信息。

[0025] 在本发明的一个实施例中,所述第二处理模块,用于:将所述第二标签与所述资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章;将预先对所述候选资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表。

[0026] 在本发明的一个实施例中,所述第三处理模块,用于:根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的所述第一量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0027] 在本发明的一个实施例中,所述第三处理模块,还用于:根据预设的第一筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的所述第一量化数据对应的资讯文章进行排序,生成第一资讯数据列表;根据预设的第二筛选指标对所述第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据;将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0028] 在本发明的一个实施例中,所述预设的第一筛选指标包括:网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的至少一种;所述预设的第二筛选指标包括:资讯文章的时效性、资讯文章的标签、资讯文章的点击率、所述用户的地

理位置和资讯文章的来源中的至少一种。

[0029] 在本发明的一个实施例中,所述的装置,还包括:接收模块,用于接收用户对所述目标资讯数据列表中的目标资讯文章的点击操作;显示模块,用于向所述用户显示所述目标资讯文章的内容以及对应的标签信息。

[0030] 为达上述目的,本发明第三方面实施例提出了另一种基于人工智能的资讯数据提供装置,包括:处理器;用于存储所述处理器可执行指令的存储器;其中,所述处理器被配置为:

[0031] 对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;

[0032] 根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;

[0033] 根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;

[0034] 将所述目标资讯数据列表反馈给所述用户。

[0035] 为了实现上述目的,本发明第四方面实施例提出了一种非临时性计算机可读存储介质,当所述存储介质中的指令由服务器端的处理器被执行时,使得服务器端能够执行一种基于人工智能的资讯数据提供方法,所述方法包括:

[0036] 对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;

[0037] 根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;

[0038] 根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;

[0039] 将所述目标资讯数据列表反馈给所述用户。

[0040] 为了实现上述目的,本发明第五方面实施例提出了一种计算机程序产品,当所述计算机程序产品中的指令处理器执行时,执行一种基于人工智能的资讯数据提供方法,所述方法包括:

[0041] 对用户输入信息进行处理获取关键信息,并向所述用户展示与所述关键信息相关联的第一标签,并获取所述用户从所述第一标签中点击的第二标签;

[0042] 根据预先对资讯数据库中资讯文章标注的标签信息与所述第二标签进行匹配,获取与所述第二标签对应的资讯数据列表;

[0043] 根据预设的筛选指标对所述资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;

[0044] 将所述目标资讯数据列表反馈给所述用户。

[0045] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0046] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0047] 图1为本发明一个实施例所提供的基于人工智能的资讯数据提供方法的流程示意图;

[0048] 图2为本发明一个实施例所提供的展示与关键信息相关联的第一标签的示意图;

[0049] 图3为本发明另一种实施例所提供的基于人工智能的资讯数据提供方法的流程示意图;

[0050] 图4为本发明又一种实施例所提供的基于人工智能的资讯数据提供方法的流程示意图;

[0051] 图5为本发明一个实施例所提供的显示目标资讯文章的内容以及对应的标签信息的示意图

[0052] 图6为本发明一个实施例所提供的基于人工智能的资讯数据提供装置的结构示意图;

[0053] 图7为本发明另一个实施例所提供的基于人工智能的资讯数据提供装置的结构示意图。

具体实施方式

[0054] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,旨在用于解释本发明,而不能理解为对本发明的限制。

[0055] 下面参考附图描述本发明实施例的基于人工智能的资讯数据提供方法及装置。

[0056] 通常,搜索引擎可以接收用户输入信息,并在互联网中获取与输入信息相关的资讯数据提供给用户,以使用户阅读到需求的资讯文章。

[0057] 然而,相关技术中通过个别数据资源方定向提交对接的方式、或者通过时效性产品直接抓取新闻还是直接给用户其他网站的地址,让用户去第三方网站进行阅读等。

[0058] 由此,上述方式无法满足各个用户的资讯需求、以及不能提供个性化的资讯服务、不能方便用户阅读资讯等。

[0059] 为了解决上述问题,本发明提出了一种基于人工智能的资讯数据提供方法。具体如下:

[0060] 图1为本发明一个实施例所提供的基于人工智能的资讯数据提供方法的流程示意图。

[0061] 如图1所示,该基于人工智能的资讯数据提供包括以下步骤:

[0062] 步骤110,对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签。

[0063] 具体地,用户可以通过键盘、语音等方式输入信息,通过语义分析方法、预设算法等方式对输入信息进行处理获取关键信息。

[0064] 进一步地,向用户展示与关键信息相关联的第一标签。其中,第一标签的有很多

个,可以根据实际需要,选择部分展示或者全部展示。

[0065] 为了本领域人员更加清楚如何向用户展示与关键信息相关联的第一标签,下面结合图2举例说明如下:

[0066] 图2为本发明一个实施例所提供的展示与关键信息相关联的第一标签的示意图。

[0067] 首先,如图2所示,用户输入信息为“赵薇”,经过对“赵薇”进行处理获取关键信息为“赵薇”。

[0068] 进一步地,向用户展示与关键信息相关联的第一标签“赵薇女儿”、“赵薇艺考青涩记录”、“赵薇林心如关系”、“赵薇陈坤友情破裂”、“赵薇童年旧照”和“赵薇老公”等六个第一标签。

[0069] 可以理解的是,用户可以自己需求或者兴趣点击上述第一标签中的任意一个即第二标签。

[0070] 需要说明的是,用户输入信息有很多种,不限于上述例子,比如输入信息为“驴得水这部电影好看吗”,经过对“驴得水这部电影好看吗”进行处理获取关键信息为“驴得水、好看”。

[0071] 步骤120,根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表。

[0072] 具体地,首先预先需要对资讯数据库中的每篇资讯文章标注对应的标签信息。以便与第二标签进行匹配以获取第二标签对应的资讯数据列表。

[0073] 具体而言,预先对资讯数据库中的每篇资讯文章标注对应的标签信息可通过图3所示的步骤实现。如图3所示,可具体包括步骤S210-S240。

[0074] 步骤210,将资讯数据库中资讯文章进行切词得到多个词语。

[0075] 步骤220,根据多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重。

[0076] 步骤230,对多个词语进行去重处理获取目标词语。

[0077] 步骤240,根据目标词语及其对应的权重为资讯文章标注对应的标签信息。

[0078] 为了方便描述理解,本实施例中以资讯数据库中的任一篇资讯文章为例进行说明如下:

[0079] 具体地,对资讯文章的内容进行切词处理,可以得到资讯文章对应的多个词语。然后根据每一个词语在资讯文章的正文、标题等中出现的频率赋予不同的权重。

[0080] 举例而言,词语A仅出现在正文中10次赋予权重0.3,词语B在正文和标题中一共出现20次赋予权重0.8。另外还可以根据词语在语句中不同位置赋予一定权重。

[0081] 为了进一步提高标签信息的准确性,以满足用户需求。需要对多个词语进行去重处理。

[0082] 可以理解的是,本示例中的去重处理首先是分析各个词语的语义,将相似含义的词语进行去重处理得到目标词语。

[0083] 进一步地,根据目标词语和其对应的权重为资讯文章标注对应的标签信息的方式有很多种,可以根据实际应用需要进行选择设置。举例说明如下:

[0084] 第一种示例,根据目标词语和其对应的权重可以确定每一个目标词语的综合分值,可以直接对各个目标词语对应的综合分值进行排序,将排名靠前的目标词语对资讯文

章标注对应的标签信息。

[0085] 第二种示例,根据目标词语和其对应的权重可以确定每一个目标词语的综合分值,设置预设阈值,将大于预设阈值的所有目标词语对资讯文章标注对应的标签信息。

[0086] 需要说明的是,可以通过机器学习、人工等方式对资讯文章标注对应的标签信息。

[0087] 进一步地,将得到的资讯文章标注的标签信息与第二标签进行匹配。可以理解的是,匹配的方式有很多种,可以根据需要进行选择,举例说明如下:

[0088] 第一种示例,将资讯数据库中全部资讯文章标注的标签信息与第二标签进行匹配,两者匹配时,将标签信息对应的资讯文章提取出来,从而获取多篇资讯文章形成资讯数据列表。

[0089] 第二种示例,首先通过第二标签与资讯数据库中全部资讯文章的文本进行匹配,筛选出一定量的资讯文章后,再利用资讯文章标注的标签信息与第二标签进行匹配,两者匹配时,将标签信息对应的资讯文章提取出来,从而获取多篇资讯文章形成资讯数据列表。进一步满足了用户对于特定资讯的需求。

[0090] 需要说明的是,资讯数据库可以是公共资讯池数据,可以聚合全网数据,还可以通过机器自动抓取,可以实时更新,保证资讯文章的时效性。

[0091] 步骤130,根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0092] 步骤140,将目标资讯数据列表反馈给用户

[0093] 具体地,预设的筛选指标有很多种,可以根据实际应用需要进行选择设置。举例说明如下:

[0094] 第一种示例,预设的筛选指标包括预设的第一筛选指标。

[0095] 其中,预设的第一筛选指标可以包括网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的一种或者多种,可以根据应用需要进行选择。

[0096] 可以理解的是,根据预设的第一筛选指标比如网页地址的点击率对资讯数据列表中的资讯文章进行处理,即根据用户点击进去资讯数据列表中的资讯文章的网页地址次数与资讯数据列表中的资讯文章的网页地址所有搜索出来的次数的比例,获取与每个资讯文章对应的量化数据。

[0097] 需要说明的是,如果包含网页地址的点击率、输入信息网页地址的点击率两种第一筛选指标,需要根据需要调整两种第一筛选指标的权重,例如网页地址的点击率0.4,输入信息网页地址的点击率0.6,然后两种第一筛选指标进行综合,获取与每个资讯文章对应的量化数据。

[0098] 进一步地,根据需要可以设置预设门限值,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0099] 第二种示例,预设的筛选指标包括预设的第一筛选指标和预设的第二筛选指标。

[0100] 其中,预设的第二筛选指标可以包括资讯文章的时效性、资讯文章的标签、资讯文章的点击率、用户的地理位置和资讯文章的来源的一种或者多种,可以根据应用需要进行选择。

[0101] 可以理解的是,在经过上述第一种示例方式可以获取第一资讯数据列表,经过第

二筛选指标对第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据。

[0102] 进一步地,根据需要可以设置预设门限值,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0103] 进一步地,将目标资讯数据列表反馈给用户。

[0104] 可以理解的是,提供给用户的是统一规范的聚合页面,设计风格一致,给用户的浏览体验一致。通过统一页面以超帧形式承接,可以缩短用户打开新页面的时间,减少用户学习成本。

[0105] 综上所述,本发明实施例的基于人工智能的资讯数据提供方法,通过对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,以及获取用户从第一标签中点击的第二标签,再根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表,最后根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表反馈给用户。由此,能够为用户提供全面性的资讯,以及针对不同的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。

[0106] 图4为本发明又一个实施例所提供的基于人工智能的资讯数据提供方法的流程示意图。

[0107] 如图4所示,该基于人工智能的资讯数据提供包括以下步骤:

[0108] 步骤310,对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签。

[0109] 需要说明的是,步骤S310的描述与上述步骤S110相对应,因此对步骤S310的描述参考上述步骤S110的描述,在此不再赘述。

[0110] 步骤320,将第二标签与资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章。

[0111] 步骤330,将预先对候选资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表。

[0112] 具体地,首先将第二标签与资讯数据库中每篇资讯文章的文本信息进行匹配。如果第二标签与资讯文章的文本信息匹配,赋予对应的资讯文章相应的分值。由此,资讯文章对应的分值越高,表示其匹配度越高。

[0113] 进而,通过预设的阈值可以删除一些相关性不大的资讯文章,进一步满足用户个性化的资讯需求。由此,得到匹配度大于预设的阈值的候选资讯文章。

[0114] 需要说明的是,本示例中的预先对候选资讯文章标注的标签信息的具体过程可以参见步骤120中预先对资讯数据库中资讯文章标注的标签信息。此处不再详述。

[0115] 进一步地,将预先对候选资讯文章标注的标签信息与第二标签进行匹配。如果匹配资讯文章标注的标签信息与第二标签匹配,则将对应的资讯文章添加入资讯数据列表中。

[0116] 步骤340,根据预设的第一筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据。

[0117] 步骤350,将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成第一资讯数据列表。

[0118] 具体地,根据多个不同权重的第二筛选指标对资讯数据列表中的资讯文章进行处理,可以获取与每个资讯文章对应的第二量化数据。

[0119] 其中,第二筛选指标可以包括网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的一种或者多种,可以根据应用需要进行选择。

[0120] 可以理解的是,网页地址的点击率(url相关的ctr)可以包括图的数量、timeliness时效性、strategy_type来自哪种召回(cb/cf/hot)和is_ent是否娱乐、tag_sum_ctr,url_ctr和domain_ctr。

[0121] 可以理解的是,输入信息网页地址的点击率(query_url相关的ctr)可以包括query_cate_ctr、query_topic_ctr、query_tag_ctr和query_url_ctr。

[0122] 可以理解的是,用户属性特征网页地址的点击率(user_url相关的ctr)可以包括tag_simi、topic_simi、cate_simi和matched_tag_ctr。

[0123] 进一步地,设置预设门限值,将超过第一门限值的第二量化数据对应的资讯文章进行排序生成第一资讯数据列表。

[0124] 步骤360,根据预设的第二筛选指标对第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据。

[0125] 步骤370,将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0126] 步骤380,将目标资讯数据列表反馈给用户。

[0127] 具体地,进一步根据不同领域用户的需求,调整预设的第二筛选指标的权重,将用户需要的内容排在最前面。

[0128] 其中,资讯文章的时效性可以理解为笑话、母婴问答的内容时效性要求低。实时新闻的内容时效性要求高。

[0129] 其中,资讯文章的标签可以理解为人物、体育可根据人名、赛事名、影视作品等实体名词圈定内容范围。

[0130] 其中,资讯文章的点击率可以理解为是否根据用户点击,将热门内容置顶。

[0131] 其中,用户的地理位置可以理解为根据用户IP位置,优先展示和其所在城市相关内容。

[0132] 其中,资讯文章的来源可以理解为根据内容发布者筛选,对于权威的资源方提高置信度。

[0133] 进一步地,根据需要可以设置预设门限值,将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表反馈给用户。

[0134] 步骤390,接收用户对目标资讯数据列表中的目标资讯文章的点击操作。

[0135] 步骤3100,向用户显示目标资讯文章的内容以及对应的标签信息。

[0136] 具体地,用户根据需要点击目标资讯数据列表中的目标资讯文章,从而可以阅读到目标资讯文章的内容以及对应的标签信息。

[0137] 为了本领域人员更加清楚如何向用户显示目标资讯文章的内容以及对应的标签信息,下面结合图5举例说明如下:

[0138] 图5为本发明一个实施例所提供的显示目标资讯文章的内容以及对应的标签信息的示意图。

[0139] 首先,如图5所示,目标资讯文章的内容“对于企业…全面布局”;标签信息“科学地球”、“风趣讲故事”等。

[0140] 其中,标签信息为预先对候选资讯文章标注的标签信息中的标签信息。

[0141] 可以理解的是,根据资讯文章标题、正文分析提取资讯文章目标词语,以及资讯文章所属的类别,还可以为资讯文章打上标签信息。

[0142] 综上所述,本发明实施例的基于人工智能的资讯数据提供方法,通过两次筛选,能够为用户提供更优质的资讯,以及向接收用户对目标资讯数据列表中的目标资讯文章的点击操作,向用户显示目标资讯文章的内容以及对应的标签信息,为用户提供全面性的资讯,方便了用户阅读资讯。

[0143] 为了实现上述实施例,本发明还提出一种基于人工智能的资讯数据提供装置。

[0144] 图6为本发明一个实施例所提供的基于人工智能的资讯数据提供装置的结构示意图。

[0145] 如图6所示,该基于人工智能的资讯数据提供装置包括:第一处理模块61、第二处理模块62、第三处理模块63和反馈模块64。

[0146] 其中,第一处理模块61对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签。

[0147] 第二处理模块62用于根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表。

[0148] 第三处理模块63用于根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0149] 反馈模块64用于将目标资讯数据列表反馈给用户。

[0150] 其中,在本发明的一个实施例中,基于人工智能的资讯数据提供装置还包括:切词模块65、确定模块66、去重模块67和标注模块68。

[0151] 其中,切词模块65用于将资讯数据库中资讯文章进行切词得到多个词语。

[0152] 确定模块66用于根据多个词语中的每个词语在资讯文章中的出现频率确定每个词语的权重。

[0153] 去重模块67用于对多个词语进行去重处理获取目标词语。

[0154] 标注模块68用于根据目标词语及其对应的权重为资讯文章标注对应的标签信息。

[0155] 其中,在本发明的一个实施例中,第二处理模块62用于将第二标签与资讯数据库中每篇资讯文章的文本信息进行匹配,获取匹配度大于预设的阈值的候选资讯文章;将预先对候选资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表。

[0156] 其中,在本发明的一个实施例中,第三处理模块63用于根据预设的第一筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的的第一量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0157] 其中,在本发明的一个实施例中,第三处理模块63还用于根据预设的第一筛选指

标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第一量化数据;将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成第一资讯数据列表;根据预设的第二筛选指标对第一资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的第二量化数据;将超过预设门限值的第二量化数据对应的资讯文章进行排序,生成目标资讯数据列表。

[0158] 其中,预设的第一筛选指标包括:网页地址的点击率、输入信息网页地址的点击率和用户属性特征网页地址的点击率中的至少一种。

[0159] 其中,预设的第二筛选指标包括:资讯文章的时效性、资讯文章的标签、资讯文章的点击率、所述用户的地理位置和资讯文章的来源中的至少一种。

[0160] 需要说明的是,前述对基于人工智能的资讯数据提供方法实施例的解释说明也适用于基于人工智能的资讯数据提供装置,此处不再赘述。

[0161] 综上所述,本发明实施例的基于人工智能的资讯数据提供装置,通过对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,以及获取用户从第一标签中点击的第二标签,再根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表,最后根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表反馈给用户。由此,能够为用户提供全面性的资讯,以及针对不同的用户提供不同的资讯,满足各个用户对于资讯的个性化需求,方便了用户阅读资讯。

[0162] 为了清楚说明上一实施例,本实施例提供了另一种基于人工智能的资讯数据提供装置。

[0163] 图7为本发明另一个实施例所提供的基于人工智能的资讯数据提供装置的结构示意图。

[0164] 如图7所示,在上一实施例的基础上,该装置还包括:接收模块69和显示模块610。

[0165] 其中,接收模块69用于接收用户对目标资讯数据列表中的目标资讯文章的点击操作。

[0166] 显示模块610用于向用户显示目标资讯文章的内容以及对应的标签信息。

[0167] 综上所述,本发明实施例的基于人工智能的资讯数据提供装置,通过两次筛选,能够为用户提供更优质的资讯,以及向接收用户对目标资讯数据列表中的目标资讯文章的点击操作,向用户显示目标资讯文章的内容以及对应的标签信息,为用户提供全面性的资讯,方便了用户阅读资讯。

[0168] 为了实现上述实施例,本发明还提出另一种基于人工智能的资讯数据提供装置,包括:处理器,以及用于存储所述处理器可执行指令的存储器。

[0169] 其中,处理器被配置为:对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签;根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表;根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;将目标资讯数据列表反馈给用户。

[0170] 为了实现上述实施例,本发明还提出一种非临时性计算机可读存储介质,当所述存储介质中的指令由服务器端的处理器被执行时,使得服务器端能够执行一种基于人工智能的资讯数据提供方法,所述方法包括:对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签;根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表;根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;将目标资讯数据列表反馈给用户。

[0171] 为了实现上述实施例,本发明还提出一种计算机程序产品,当所述计算机程序产品中的指令处理器执行时,执行一种基于人工智能的资讯数据提供方法,所述方法包括:对用户输入信息进行处理获取关键信息,并向用户展示与关键信息相关联的第一标签,并获取用户从第一标签中点击的第二标签;根据预先对资讯数据库中资讯文章标注的标签信息与第二标签进行匹配,获取与第二标签对应的资讯数据列表;根据预设的筛选指标对资讯数据列表中的资讯文章进行处理,获取与每个资讯文章对应的量化数据,将超过预设门限值的量化数据对应的资讯文章进行排序,生成目标资讯数据列表;将目标资讯数据列表反馈给用户。

[0172] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不必针对的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0173] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中,“多个”的含义是至少两个,例如两个,三个等,除非另有明确具体的限定。

[0174] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本发明的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本发明的实施例所属技术领域的技术人员所理解。

[0175] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为是在于实现逻辑功能的可执行指令的定序列列表,可以具体实现在任何计算机可读介质中,以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用,或结合这些指令执行系统、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电

连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPROM或闪速存储器),光纤装置,以及便携式光盘只读存储器(CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0176] 应当理解,本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如,如果用硬件来实现和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0177] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0178] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。

[0179] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

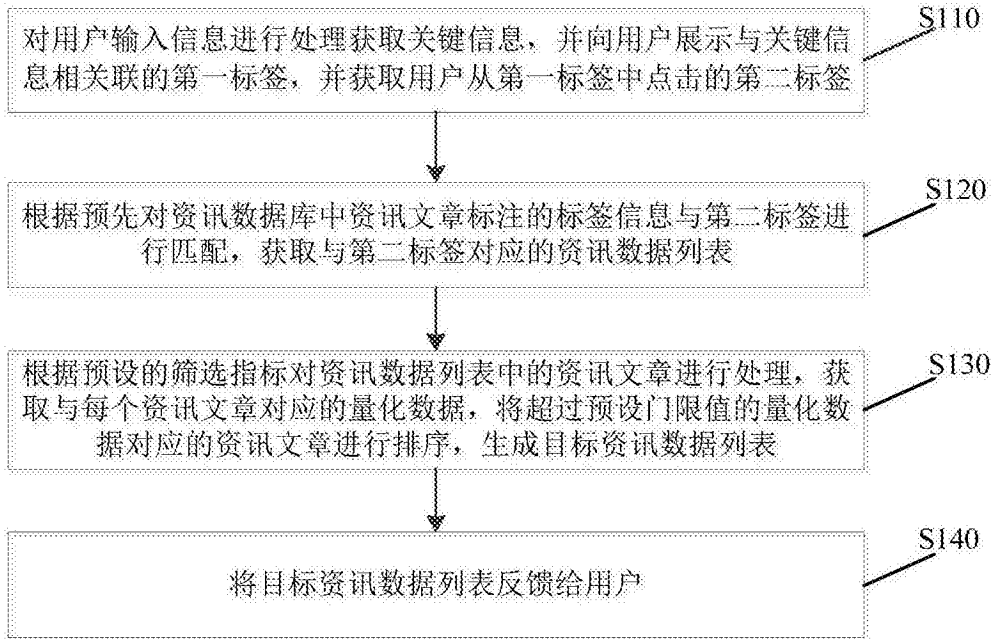


图1



图2

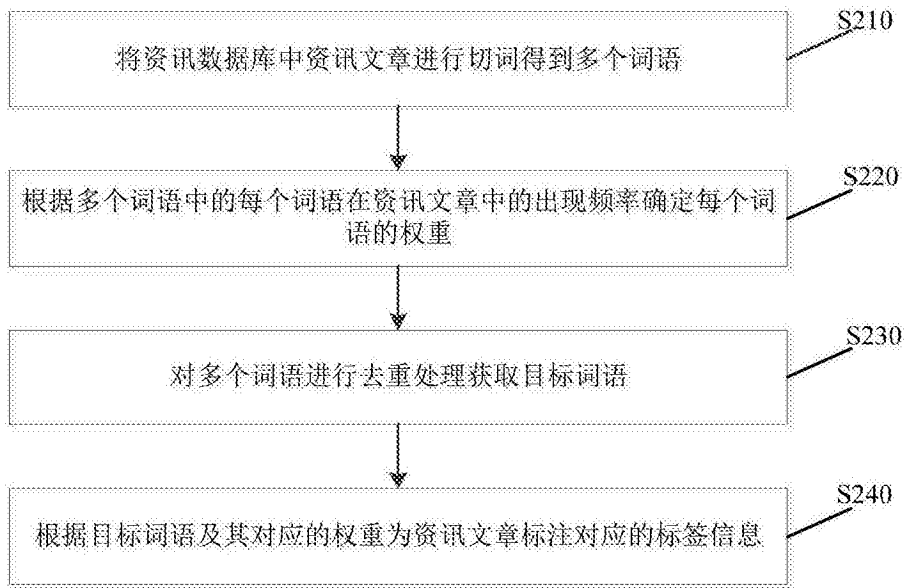


图3

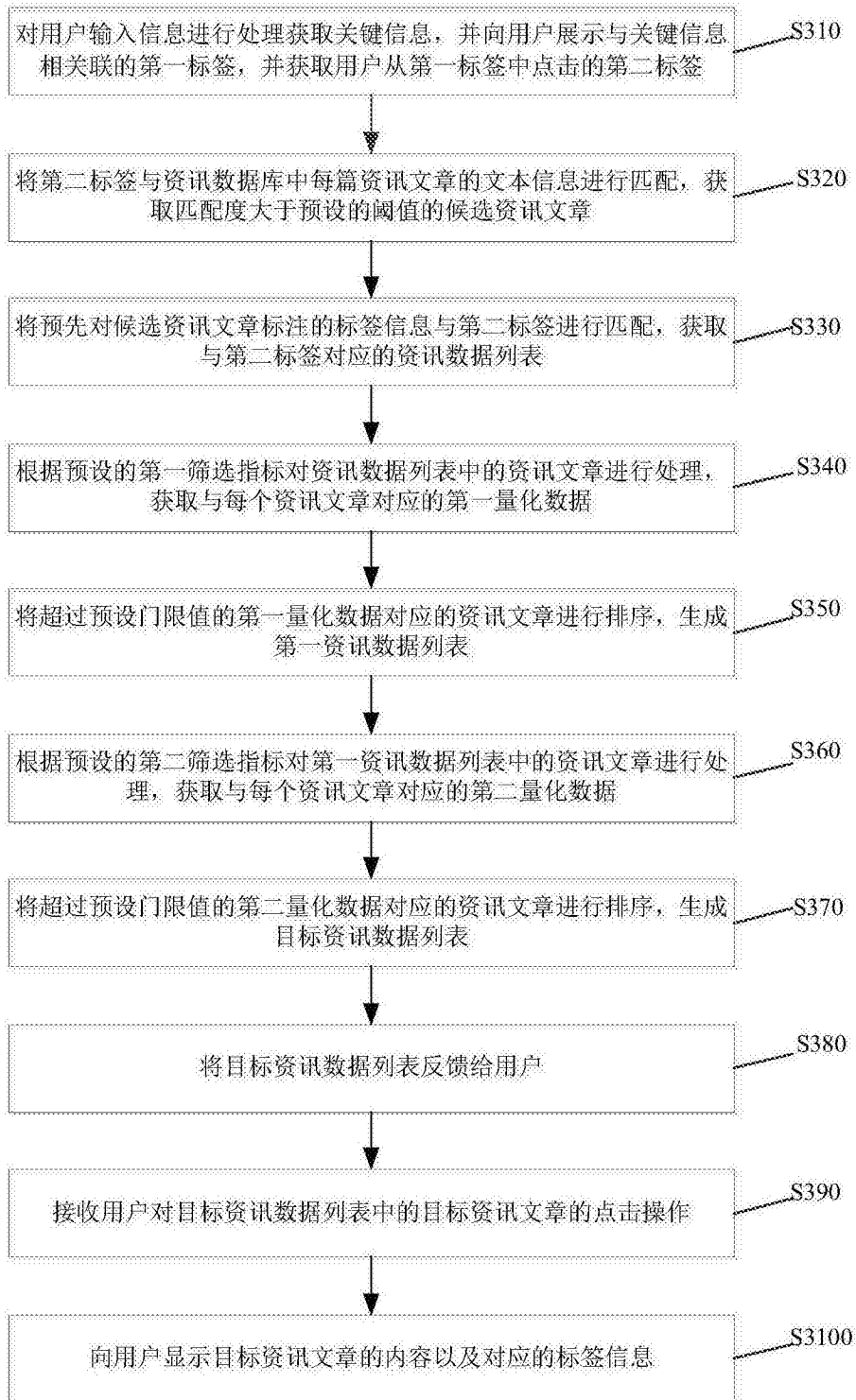


图4



图5

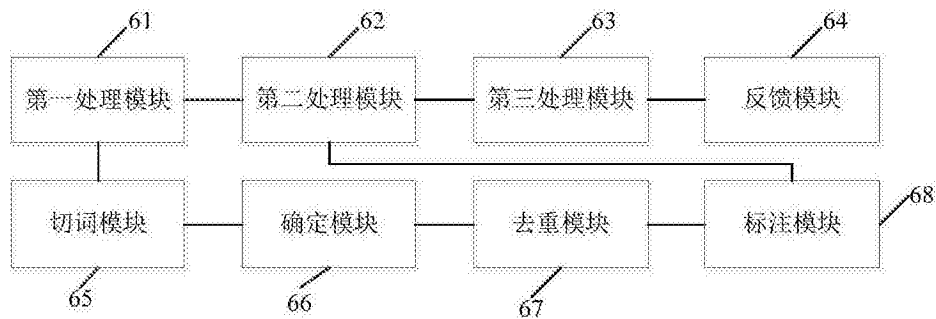


图6

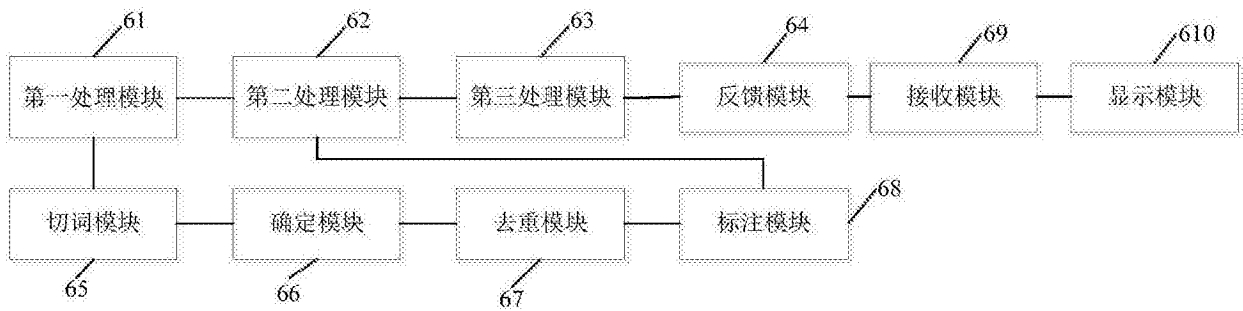


图7