



(12) 发明专利申请

(10) 申请公布号 CN 113609833 A

(43) 申请公布日 2021.11.05

(21) 申请号 202110925856.X

G06F 40/35 (2020.01)

(22) 申请日 2021.08.12

G06F 40/253 (2020.01)

(71) 申请人 深圳平安智汇企业信息管理有限公司

G06F 16/35 (2019.01)

G06F 16/903 (2019.01)

地址 518000 广东省深圳市福田区沙头街
道天安社区车公庙泰然七路1号博今
商务广场B座二十一层的整层

(72) 发明人 廖俊瑶 董垒 黎有耕 杨坚
李书君 孙朝辉

(74) 专利代理机构 深圳市世联合知识产权代理
有限公司 44385

代理人 汪琳琳

(51) Int. Cl.

G06F 40/186 (2020.01)

G06F 40/289 (2020.01)

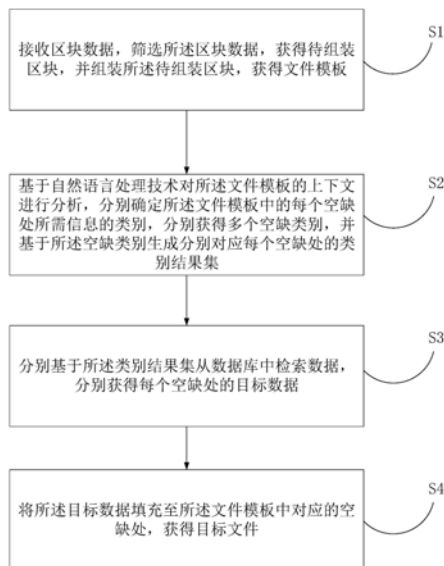
权利要求书2页 说明书11页 附图4页

(54) 发明名称

文件的动态生成方法、装置、计算机设备及
存储介质

(57) 摘要

本申请实施例属于大数据领域,应用于智慧
政务领域中,涉及一种文件的动态生成方法,包
括接收区块数据,筛选区块数据,获得待组装
区块,并组装待组装区块,获得文件模板;基于自然
语言处理技术对文件模板的上下文进行分析,确
定文件模板中的每个空缺处所需信息的类别,获
得多个空缺类别,基于空缺类别生成分别对应每
个空缺处的类别结果集;分别基于类别结果集从
数据库中检索数据,分别获得每个空缺处的目标
数据,数据库中的数据携带有分类标签;将目标
数据填充至文件模板中对应的空缺处,获得目标
文件。本申请还提供一种文件的动态生成装置、
计算机设备及存储介质。其中,目标文件可存储
于区块链中。本申请实现文件的动态生成,减少
人员参与。



1. 一种文件的动态生成方法,其特征在于,包括下述步骤:

接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板;

基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系;

分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据;

将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

2. 根据权利要求1所述的文件的动态生成方法,其特征在于,所述基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值的步骤包括:

分别对所述文件模板中的每个空缺处的上下文进行分词操作,获得词组;

将所述词组与预设的词语库中携带有意图标签的词语进行比较;

在所述词组与所述词语相同时,将所述词语携带的意图标签赋予给所述词组,作为所述词组的词组标签;

确定与所述空缺处距离小于阈值距离的词组所对应的词组标签,作为所述空缺处的空缺类别,距离作为所述空缺类别的分值。

3. 根据权利要求1所述的文件的动态生成方法,其特征在于,所述分别基于所述类别结果集从数据库中检索数据,获得中间数据的步骤包括:

分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致;

在所述空缺类别与所述分类标签一致时,将所述分类标签所关联的数据作为对应空缺处的中间数据。

4. 根据权利要求3所述的文件的动态生成方法,其特征在于,所述分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致的步骤包括:

将所述空缺类别和所述分类标签分别转换为空缺类别向量和分类标签向量;

计算所述空缺类别向量和所述分类标签向量之间的余弦相似度,获得向量相似度;

确定所述向量相似度是否小于相似度阈值,在所述向量相似度大于所述相似度阈值时,确定所述空缺类别与所述分类标签不一致;

在所述向量相似度小于或等于所述相似度阈值时,确定所述空缺类别与所述分类标签一致。

5. 根据权利要求1所述的文件的动态生成方法,其特征在于,所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

分别将每个空缺处的空缺类别的分值赋予对应的中间数据,分别将每个空缺处的分值最高的中间数据作为准确结果,剩余的中间数据作为候选结果,将所述准确结果和所述候选结果展示在前端页面中,以供用户选择;

在检测到用户选择所述准确结果时,将所述准确结果作为当前空缺处的目标数据;

在检测到用户选择任一所述候选结果时,将用户选择的候选结果作为当前空缺处的目

标数据,并向预先配置的人员发送算法调整提醒。

6. 根据权利要求1所述的文件的动态生成方法,其特征在于,所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

分别将每个空缺处的分值最高的空缺类别所对应的中间数据,作为当前空缺处的目标数据。

7. 根据权利要求1所述的文件的动态生成方法,其特征在于,在所述将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件的步骤之后,还包括:

当识别到用户输入对象名称时,获取所述对象名称,并根据所述对象名称查找数据库中存储的对象信息;

识别所述对象信息中的个人健康信息;

在所述个人健康信息中包括视力障碍时,根据所述视力障碍中的视力度数,根据所述视力度数调整所述目标文件中的文字大小,或者基于音频转换技术将所述目标文件中的文字转化为音频数据并播放;

在所述个人健康信息中的民族为少数民族时,基于所述少数民族,将所述目标文件中的文字转化为对应的民族文字。

8. 一种文件的动态生成装置,其特征在于,包括:

选择模块,用于接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板;

分析模块,用于基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系;

检索模块,用于分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据;

填充模块,用于将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

9. 一种计算机设备,其特征在于,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述处理器执行所述计算机可读指令时实现如权利要求1至7中任一项所述的文件的动态生成方法的步骤。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有计算机可读指令,所述计算机可读指令被处理器执行时实现如权利要求1至7中任一项所述的文件的动态生成方法的步骤。

文件的动态生成方法、装置、计算机设备及存储介质

技术领域

[0001] 本申请涉及大数据技术领域,尤其涉及文件的动态生成方法、装置、计算机设备及存储介质。

背景技术

[0002] 目前,合同或者协议等比较正式的文件起草、生成和处理都是一个比较繁琐的过程,在传统的信息系统中,需要先定好模板,将需要填充信息的地方空缺出来,可以是人工手动填入信息,也可以是从系统自动获取,因为文件结构和需要填充的信息空缺是已知的,所以获取数据的源是已知的。

[0003] 但是,配置一个文件模板后还需要将空缺字段与系统中的数据源字段进行绑定,进而让系统能够对信息进行自动填充。在这种情境下,由于需要配置模板的人懂得专业的计算机知识来给模板绑定数据源,或者需要跨部门合作,配置文件模板成为了一个比较专业化且繁琐的工作,需要大量人员参与,效率较低。并且这种模板是预先定义的固定形式的非动态的模板,如果需要个性化定制,则需要改动模板文本和背后绑定的数据源,繁琐且不够灵活。

发明内容

[0004] 本申请实施例的目的在于提出一种文件的动态生成方法、装置、计算机设备及存储介质,实现文件的动态生成,减少人员参与。

[0005] 为了解决上述技术问题,本申请实施例提供一种文件的动态生成方法,采用了如下所述的技术方案:

[0006] 一种文件的动态生成方法,包括下述步骤:

[0007] 接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板;

[0008] 基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系;

[0009] 分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据;

[0010] 将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

[0011] 进一步的,所述基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值的步骤包括:

[0012] 分别对所述文件模板中的每个空缺处的上下文进行分词操作,获得词组;

[0013] 将所述词组与预设的词语库中携带有意图标签的词语进行比较;

[0014] 在所述词组与所述词语相同时,将所述词语携带的意图标签赋予给所述词组,作为所述词组的词组标签;

[0015] 确定与所述空缺处距离小于阈值距离的词组所对应的词组标签,作为所述空缺处的空缺类别,距离作为所述空缺类别的分值。

[0016] 进一步的,所述分别基于所述类别结果集从数据库中检索数据,获得中间数据的步骤包括:

[0017] 分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致;

[0018] 在所述空缺类别与所述分类标签一致时,将所述分类标签所关联的数据作为对应空缺处的中间数据。

[0019] 进一步的,所述分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致的步骤包括:

[0020] 将所述空缺类别和所述分类标签分别转换为空缺类别向量和分类标签向量;

[0021] 计算所述空缺类别向量和所述分类标签向量之间的余弦相似度,获得向量相似度;

[0022] 确定所述向量相似度是否小于相似度阈值,在所述向量相似度大于所述相似度阈值时,确定所述空缺类别与所述分类标签不一致;

[0023] 在所述向量相似度小于或等于所述相似度阈值时,确定所述空缺类别与所述分类标签一致。

[0024] 进一步的,所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

[0025] 分别将每个空缺处的空缺类别的分值赋予对应的中间数据,分别将每个空缺处的分值最高的中间数据作为准确结果,剩余的中间数据作为候选结果,将所述准确结果和所述候选结果展示在前端页面中,以供用户选择;

[0026] 在检测到用户选择所述准确结果时,将所述准确结果作为当前空缺处的目标数据;

[0027] 在检测到用户选择任一所述候选结果时,将用户选择的候选结果作为当前空缺处的目标数据,并向预先配置的人员发送算法调整提醒。

[0028] 进一步的,所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

[0029] 分别将每个空缺处的分值最高的空缺类别所对应的中间数据,作为当前空缺处的目标数据。

[0030] 进一步的,在所述将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件的步骤之后,还包括:

[0031] 当识别到用户输入对象名称时,获取所述对象名称,并根据所述对象名称查找数据库中存储的对象信息;

[0032] 识别所述对象信息中的个人健康信息;

[0033] 在所述个人健康信息中包括视力障碍时,根据所述视力障碍中的视力度数,根据所述视力度数调整所述目标文件中的文字大小,或者基于音频转换技术将所述目标文件中的文字转化为音频数据并播放;

[0034] 在所述个人健康信息中的民族为少数民族时,基于所述少数民族,将所述目标文

件中的文字转化为对应的民族文字。

[0035] 为了解决上述技术问题,本申请实施例还提供一种文件的动态生成装置,采用了如下所述的技术方案:

[0036] 一种文件的动态生成装置,包括:

[0037] 选择模块,用于接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板;

[0038] 分析模块,用于基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系;

[0039] 检索模块,用于分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据;

[0040] 填充模块,用于将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

[0041] 为了解决上述技术问题,本申请实施例还提供一种计算机设备,采用了如下所述的技术方案:

[0042] 一种计算机设备,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述处理器执行所述计算机可读指令时实现上述的文件的动态生成方法的步骤。

[0043] 为了解决上述技术问题,本申请实施例还提供一种计算机可读存储介质,采用了如下所述的技术方案:

[0044] 一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机可读指令,所述计算机可读指令被处理器执行时实现上述的文件的动态生成方法的步骤。

[0045] 与现有技术相比,本申请实施例主要有以下有益效果:

[0046] 通过组装的区块的不同,从而可以实现获得灵活的动态的文件模板。通过自然语言处理技术对文件模板的上下文进行分析,能够有效的确定空缺处的类别,进而便于在数据库中检索到目标数据,填入空缺处,获得完整的目标文件,无需额外配置每个空缺处的数据源字段,更加无需手工填写,实现目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

附图说明

[0047] 为了更清楚地说明本申请中的方案,下面将对本申请实施例描述中所需要使用的附图作一个简单介绍,显而易见地,下面描述中的附图是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0048] 图1是本申请可以应用于其中的示例性系统架构图;

[0049] 图2是根据本申请的文件的动态生成方法的一个实施例的流程图;

[0050] 图3是根据本申请的文件的动态生成装置的一个实施例的结构示意图;

[0051] 图4是根据本申请的计算机设备的一个实施例的结构示意图。

[0052] 附图标记:200、计算机设备;201、存储器;202、处理器;203、网络接口;300、文件的动态生成装置;301、选择模块;302、分析模块;303、检索模块;304、填充模块。

具体实施方式

[0053] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同;本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的,不是旨在于限制本申请;本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排他的包含。本申请的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象,而不是用于描述特定顺序。

[0054] 在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

[0055] 为了使本技术领域的人员更好地理解本申请方案,下面将结合附图,对本申请实施例中的技术方案进行清楚、完整地描述。

[0056] 如图1所示,系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0057] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

[0058] 终端设备101、102、103可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

[0059] 服务器105可以是提供各种服务的服务器,例如对终端设备101、102、103上显示的页面提供支持的后台服务器。

[0060] 需要说明的是,本申请实施例所提供的文件的动态生成方法一般由服务器/终端设备执行,相应地,文件的动态生成装置一般设置于服务器/终端设备中。

[0061] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器。

[0062] 继续参考图2,示出了根据本申请的文件的动态生成方法的一个实施例的流程图。所述的文件的动态生成方法,包括以下步骤:

[0063] S1:接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板。

[0064] 在本实施例中,对于所述区块数据的筛选过程为:本申请通过获取当前登录的用户的用户名称,根据用户名称调取日志,从日志中确定出该用户使用频率大于频率阈值的历史区块数据,作为用户区块;对比接收到的每个区块数据与每个用户区块是否相同,在区块数据与用户区块相同时,将该区块数据作为待组装区块,最终获得至少一个待组装区块。在所有的区块数据与所有的用户区块均不相同时,获取所述区块数据携带的权重值,基于

权重值对所述区块数据进行降序排序,获得区块数据列表,将所述区块数据列表中前n个区块数据作为待组装区块,其中,n为预设的待组装区块的数量。通过组装所述待组装区块,获得文件模板,而非是预先固定的完整固定的模板,实现文件模板的灵活、动态生成。文件模板可以是合同模板,也可以是协议模板。其中,区块数据的筛选,可以是将区块数据展示在前端页面中,用户根据需要对区块数据进行选择,选择需要填写信息的区块数据,如需要填写职业信息,教育信息等。文件模板的组成过程可以是手动在前端页面中的界面拖动排序,也可以根据信息区块的展示优先级默认排序生成文件模板,其中,每个小区块都有一个优先级数值。

[0065] 在本实施例中,文件的动态生成方法运行于其上的电子设备(例如图1所示的服务器/终端设备)可以通过有线连接方式或者无线连接方式接收区块数据。需要指出的是,上述无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接、以及其他现在已知或将来开发的无线连接方式。

[0066] S2:基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系。

[0067] 在本实施例中,通过自然语言处理(NLP)技术对文件模板的上下文进行分析,分析出空缺之处可能需要填写的信息是什么类别,并且将分析结果中的空缺组成一个结果集。其中,如果文件是针对某个特定的人的,那么上下文分析之后,计算机最终能够将该人的个人信息的类别作为空缺类别加入类别结果集。如果文件是针对某个特定的公司的,那么上下文分析之后,计算机就能够将该公司的通用信息,例如公司名称、地址等作为空缺类别加入类别结果集。

[0068] 具体的,在步骤S2中,即所述基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值的步骤包括:

[0069] 分别对所述文件模板中的每个空缺处的上下文进行分词操作,获得词组;

[0070] 将所述词组与预设的词语库中携带有意图标签的词语进行比较;

[0071] 在所述词组与所述词语相同时,将所述词语携带的意图标签赋予给所述词组,作为所述词组的词组标签;

[0072] 确定与所述空缺处距离小于阈值距离的词组所对应的词组标签,作为所述空缺处的空缺类别,距离作为所述空缺类别的分值。

[0073] 在本实施例中,通过NLP技术中的意图识别技术对文件模板的上下文进行分析,本申请的分析和过程利用了分词,然后与预设中的携带意图标签的词语匹配进行分析,分析出空缺之处可能需要填写的信息是什么类别,并且将分析结果中的空缺类别组成一个类别结果集,结果集包含空缺处想要的信息类型。其中,对于获得的空缺类别,将空缺类别所对应的词组与空缺处的距离作为该空缺类别的分值。距离具体指该词组与该空缺处之间相隔字符的数量的倒数,在该词组与该空缺处之间相隔字符的数量为0时,距离为1。则距离越近,分值越高。

[0074] S3:分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据。

[0075] 在本实施例中,本申请的系统后台拥有一个存储了海量信息的大数据系统,系统

中已经给数据源中的数据进行了分类,计算机通过模板字段匹配系统能够从大数据系统中根据分析出来的类别结果集(即空缺需要的信息类别),进而在大数据系统中检索需要的数据信息,返回给模板系统,并自动在文件模板中进行填充,最终获得目标文件。

[0076] 具体的,在步骤S3中,即所述分别基于所述类别结果集从数据库中检索数据,获得中间数据的步骤包括:

[0077] 分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致;

[0078] 在所述空缺类别与所述分类标签一致时,将所述分类标签所关联的数据作为对应空缺处的中间数据。

[0079] 在本实施例中,将空缺类别与分类标签进行匹配,字段相同的空缺类别和分类标签,则判断为匹配一致,将该分类标签所对应的数据,作为当前空缺处的中间数据,通过初步对数据进行筛选,便于后续的进一步处理,最终实现获得更加精准的目标数据。

[0080] 进一步的,所述分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致的步骤包括:

[0081] 将所述空缺类别和所述分类标签分别转换为空缺类别向量和分类标签向量;

[0082] 计算所述空缺类别向量和所述分类标签向量之间的余弦相似度,获得向量相似度;

[0083] 确定所述向量相似度是否小于相似度阈值,在所述向量相似度大于所述相似度阈值时,确定所述空缺类别与所述分类标签不一致;

[0084] 在所述向量相似度小于或等于所述相似度阈值时,确定所述空缺类别与所述分类标签一致。

[0085] 在本实施例中,获取预先存储的向量映射表,基于该向量映射表将所述空缺类别和所述分类标签分别转换为空缺类别向量和分类标签向量。通过转换后的空缺类别向量和分类标签向量能够更加快速方便的计算出空缺类别和分类标签之间的相似程度,即向量相似度。进而根据向量相似度快速确定空缺类别与分类标签是否一致。

[0086] 此外,在步骤S3中,即所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

[0087] 分别将每个空缺处的空缺类别的分值赋予对应的中间数据,分别将每个空缺处的分值最高的中间数据作为准确结果,剩余的中间数据作为候选结果,将所述准确结果和所述候选结果展示在前端页面中,以供用户选择;

[0088] 在检测到用户选择所述准确结果时,将所述准确结果作为当前空缺处的目标数据;

[0089] 在检测到用户选择任一所述候选结果时,将用户选择的候选结果作为当前空缺处的目标数据,并向预先配置的人员发送算法调整提醒。

[0090] 在本实施例中,由于分析及结果可能并不完全100%准确,模板系统收到的返回包括将会自动填充的最准确结果和候选结果,分析出来分值最高的为将要填充的信息,其他为候选,或者附带返回候选结果供填写信息的人进行候选,候选结果有多个,同样来自数据源,根据分值排序返回,如果填写信息的人选择了与预期结果不同的结果,那么系统会通过学习来改进算法,并且在日积月累的学习中进一步提高分析结果的准确度。

[0091] 作为本申请的另一实施例,在步骤S3中,即所述基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据的步骤包括:

[0092] 分别将每个空缺处的分值最高的空缺类别所对应的中间数据,作为当前空缺处的目标数据。

[0093] 在本实施例中,空缺处会排序出多个类别,直接将最高值作为最准确的空缺类别,最准确的空缺类别对应的中间数据作为目标数据,实现快速获得目标数据。

[0094] S4:将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

[0095] 在本实施例中,将所述目标数据填充至所述文件模板中的空缺处,即获得目标文件,生成文件后,计算机的自然语言处理系统会进一步对全文进行语义分析和语法检查,如果检查不通过会提示人工处理,人工处理后系统和开发人员会通过训练来改进NLP算法的模型以及调整参数来进一步提高分析结果的准确度。

[0096] 在本实施例的一些可选的实现方式中,在步骤S4,即在所述将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件的步骤之后,上述电子设备还可以执行以下步骤:

[0097] 当识别到用户输入对象名称时,获取所述对象名称,并根据所述对象名称查找数据库中存储的对象信息;

[0098] 识别所述对象信息中的个人健康信息;

[0099] 在所述个人健康信息中包括视力障碍时,根据所述视力障碍中的视力度数,根据所述视力度数调整所述目标文件中的文字大小,或者基于音频转换技术将所述目标文件中的文字转化为音频数据并播放;

[0100] 在所述个人健康信息中的民族为少数民族时,基于所述少数民族,将所述目标文件中的文字转化为对应的民族文字。

[0101] 在本实施例中,音频转换技术为TTS(Text-To-Speech,语音合成)技术,能够将文字转换为语音。本申请同时针对个人进行一些个性化的服务,如果数据库中存储了个人的健康信息,那么会从存储的个人健康信息中获取视力等辅助信息,例如个人如果是视力障碍人群,那么生成的最终文件会对应调整字体大小,或者生成文件的对应语音版本供人播放。还可以根据个人的民族信息生成对应的语言版本的文件。

[0102] 本申请能够使办公更加的智能化并且高效,可以减少因为人为失误带来的未知的损失。能减轻撰写、审阅相关文件的人力成本,并且可以提升此类工作的体验,将繁琐而且重复的事务交由机器自动处理,将人力转移到其它具有更高价值的岗位和事务上去,提高公司或机构的整体运转效率,为企业机构减少相关的成本。使得文件的处理更加灵活化、规范化,可以与此同时训练针对此场景的人工智能模型逐步提升计算机的智能化程度。

[0103] 通过组装的区块的不同,从而可以实现获得灵活的动态的文件模板。通过自然语言处理技术对文件模板的上下文进行分析,能够有效的确定空缺处的类别,进而便于在数据库中检索到目标数据,填入空缺处,获得完整的目标文件,无需额外配置每个空缺处的数据源字段,更加无需手工填写,实现目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

[0104] 需要强调的是,为进一步保证上述目标文件的私密和安全性,上述目标文件还可以存储于一区块链的节点中。

[0105] 本申请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。

[0106] 本申请可应用于智慧政务领域中,从而推动智慧城市的建设。

[0107] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机可读指令来指令相关的硬件来完成,该计算机可读指令可存储于一计算机可读存储介质中,该计算机可读指令在执行时,可包括如上述各方法的实施例的流程。其中,前述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)等非易失性存储介质,或随机存储记忆体(Random Access Memory,RAM)等。

[0108] 应该理解的是,虽然附图的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,其可以以其他的顺序执行。而且,附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,其执行顺序也不必然是依次进行,而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0109] 进一步参考图3,作为对上述图2所示方法的实现,本申请提供了一种文件的动态生成装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0110] 如图3所示,本实施例所述的文件的动态生成装置300包括:选择模块301、分析模块302、检索模块303以及填充模块304。其中:选择模块301,用于接收区块数据,筛选所述区块数据,获得待组装区块,并组装所述待组装区块,获得文件模板;分析模块302,用于基于自然语言处理技术对所述文件模板的上下文进行分析,获得所述文件模板中的每个空缺处所对应的多个空缺类别和分值,并基于所述空缺类别生成分别对应每个空缺处的类别结果集,其中,所述空缺类别和所述分值为一一对应的关联关系;检索模块303,用于分别基于所述类别结果集从数据库中检索数据,获得中间数据,并基于所述分值筛选所述中间数据,分别获得每个空缺处的目标数据;填充模块304,用于将所述目标数据填充至所述文件模板中对应的空缺处,获得目标文件。

[0111] 在本实施例中,通过组装的区块的不同,从而可以实现获得灵活的动态的文件模板。通过自然语言处理技术对文件模板的上下文进行分析,能够有效的确定空缺处的类别,进而便于在数据库中检索到目标数据,填入空缺处,获得完整的目标文件,无需额外配置每个空缺处的数据源字段,更加无需手工填写,实现目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

[0112] 分析模块302包括分词子模块、比较子模块、赋予子模块和第一确定子模块。其中,分词子模块用于分别对所述文件模板中的每个空缺处的上下文进行分词操作,获得词组;比较子模块用于将所述词组与预设的词语库中携带有意图标签的词语进行比较;赋予子模块用于在所述词组与所述词语相同时,将所述词语携带的意图标签赋予给所述词组,作为所述词组的词组标签;第一确定子模块用于确定与所述空缺处距离小于阈值距离的词组所

对应的词组标签,作为所述空缺处的空缺类别,距离作为所述空缺类别的分值。

[0113] 检索模块303包括第二确定子模块和获得子模块。其中,第二确定子模块用于分别确定每个空缺处的类别结果集中的空缺类别,与所述数据库中的分类标签是否一致;获得子模块用于在所述空缺类别与所述分类标签一致时,将所述分类标签所关联的数据作为对应空缺处的中间数据。

[0114] 第二确定子模块包括转换单元、计算单元、第一确定单元和第二确定单元。其中,转换单元用于将所述空缺类别和所述分类标签分别转换为空缺类别向量和分类标签向量;计算单元用于计算所述空缺类别向量和所述分类标签向量之间的余弦相似度,获得向量相似度;第一确定单元用于确定所述向量相似度是否小于相似度阈值,在所述向量相似度大于所述相似度阈值时,确定所述空缺类别与所述分类标签不一致;第二确定单元用于在所述向量相似度小于或等于所述相似度阈值时,确定所述空缺类别与所述分类标签一致。

[0115] 检索模块303还包括展示子模块、检测子模块和发送子模块,其中,展示子模块用于分别将每个空缺处的空缺类别的分值赋予对应的中间数据,分别将每个空缺处的分值最高的中间数据作为准确结果,剩余的中间数据作为候选结果,将所述准确结果和所述候选结果展示在前端页面中,以供用户选择;检测子模块用于在检测到用户选择所述准确结果时,将所述准确结果作为当前空缺处的目标数据;发送子模块用于在检测到用户选择任一所述候选结果时,将用户选择的候选结果作为当前空缺处的目标数据,并向预先配置的人员发送算法调整提醒。

[0116] 在本实施例的一些可选的实现方式中,上述检索模块303进一步用于:分别将每个空缺处的分值最高的空缺类别所对应的中间数据,作为当前空缺处的目标数据。

[0117] 在本实施例的一些可选的实现方式中,上述装置300还包括:获取模块、识别模块、调整模块和转化模块,其中,获取模块用于当识别到用户输入对象名称时,获取所述对象名称,并根据所述对象名称查找数据库中存储的对象信息;识别模块用于识别所述对象信息中的个人健康信息;调整模块用于在所述个人健康信息中包括视力障碍时,根据所述视力障碍中的视力度数,根据所述视力度数调整所述目标文件中的文字大小,或者基于音频转换技术将所述目标文件中的文字转化为音频数据并播放;转化模块用于在所述个人健康信息中的民族为少数民族时,基于所述少数民族,将所述目标文件中的文字转化为对应的民族文字。

[0118] 通过组装的区块的不同,从而可以实现获得灵活的动态的文件模板。通过自然语言处理技术对文件模板的上下文进行分析,能够有效的确定空缺处的类别,进而便于在数据库中检索到目标数据,填入空缺处,获得完整的目标文件,无需额外配置每个空缺处的数据源字段,更加无需手工填写,实现目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

[0119] 为解决上述技术问题,本申请实施例还提供计算机设备。具体请参阅图4,图4为本实施例计算机设备基本结构框图。

[0120] 所述计算机设备200包括通过系统总线相互通信连接存储器201、处理器202、网络接口203。需要指出的是,图中仅示出了具有组件201-203的计算机设备200,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。其中,本技术领域技术人员可以理解,这里的计算机设备是一种能够按照事先设定或存储的指令,自动进

行数值计算和/或信息处理的设备,其硬件包括但不限于微处理器、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程门阵列(Field-Programmable Gate Array,FPGA)、数字处理器(Digital Signal Processor,DSP)、嵌入式设备等。

[0121] 所述计算机设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述计算机设备可以与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互。

[0122] 所述存储器201至少包括一种类型的可读存储介质,所述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,所述存储器201可以是所述计算机设备200的内部存储单元,例如该计算机设备200的硬盘或内存。在另一些实施例中,所述存储器201也可以是所述计算机设备200的外部存储设备,例如该计算机设备200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,所述存储器201还可以既包括所述计算机设备200的内部存储单元也包括其外部存储设备。本实施例中,所述存储器201通常用于存储安装于所述计算机设备200的操作系统和各类应用软件,例如文件的动态生成方法的计算机可读指令等。此外,所述存储器201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。

[0123] 所述处理器202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202通常用于控制所述计算机设备200的总体操作。本实施例中,所述处理器202用于运行所述存储器201中存储的计算机可读指令或者处理数据,例如运行所述文件的动态生成方法的计算机可读指令。

[0124] 所述网络接口203可包括无线网络接口或有线网络接口,该网络接口203通常用于在所述计算机设备200与其他电子设备之间建立通信连接。

[0125] 在本实施例中,本申请实现了目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

[0126] 本申请还提供了另一种实施方式,即提供一种计算机可读存储介质,所述计算机可读存储介质存储有计算机可读指令,所述计算机可读指令可被至少一个处理器执行,以使所述至少一个处理器执行如上述的文件的动态生成方法的步骤。

[0127] 在本实施例中,本申请实现了目标文件的动态生成,减少了人工操作,目标文件的生成更加的自动化。

[0128] 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。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本申请各个实施例所述的方法。

[0129] 显然,以上所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例,附图中给出了本申请的较佳实施例,但并不限制本申请的专利范围。本申请可以以许多不同

的形式来实现,相反地,提供这些实施例的目的是使对本申请的公开内容的理解更加透彻全面。尽管参照前述实施例对本申请进行了详细的说明,对于本领域的技术人员来而言,其依然可以对前述各具体实施方式所记载的技术方案进行修改,或者对其中部分技术特征进行等效替换。凡是利用本申请说明书及附图内容所做的等效结构,直接或间接运用在其他相关的技术领域,均同理在本申请专利保护范围之内。

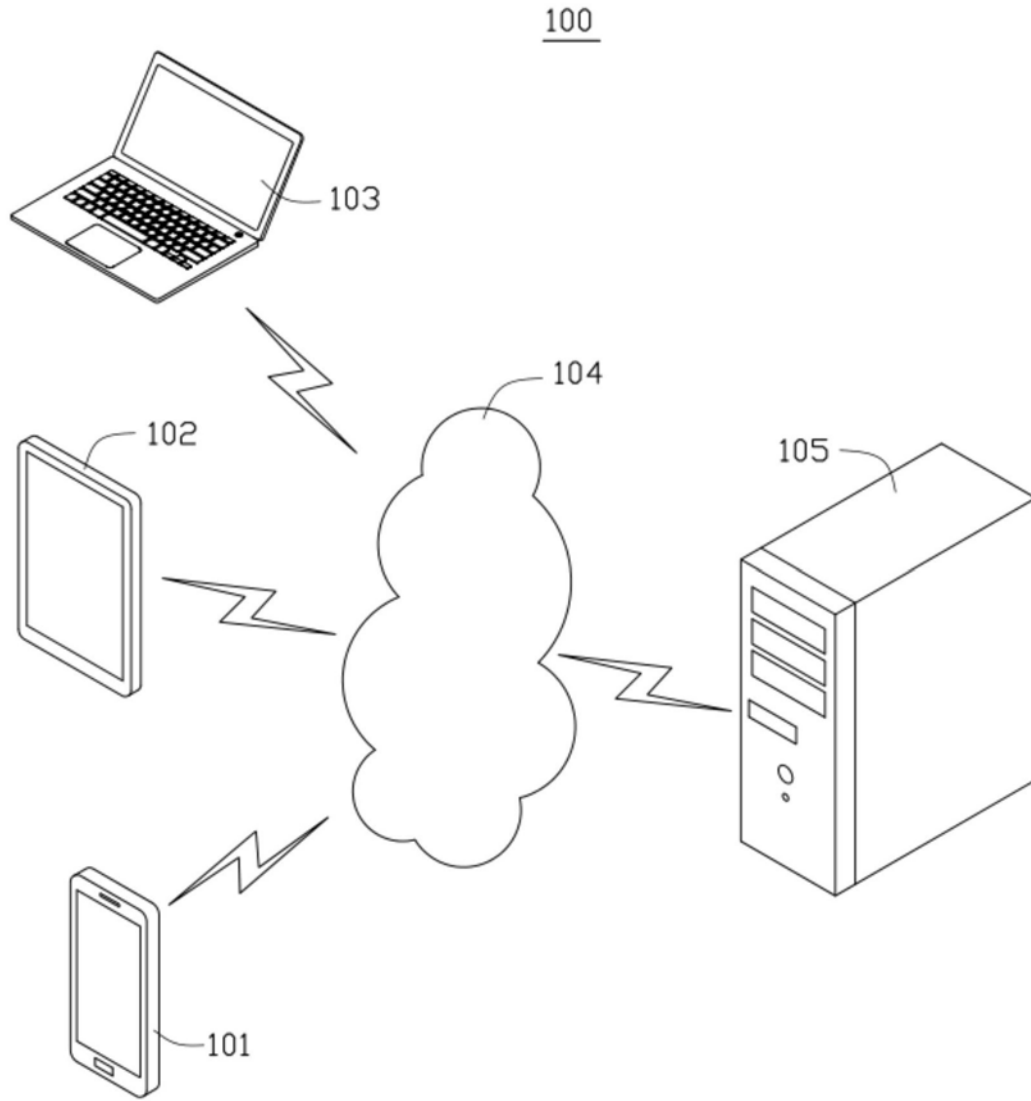


图1

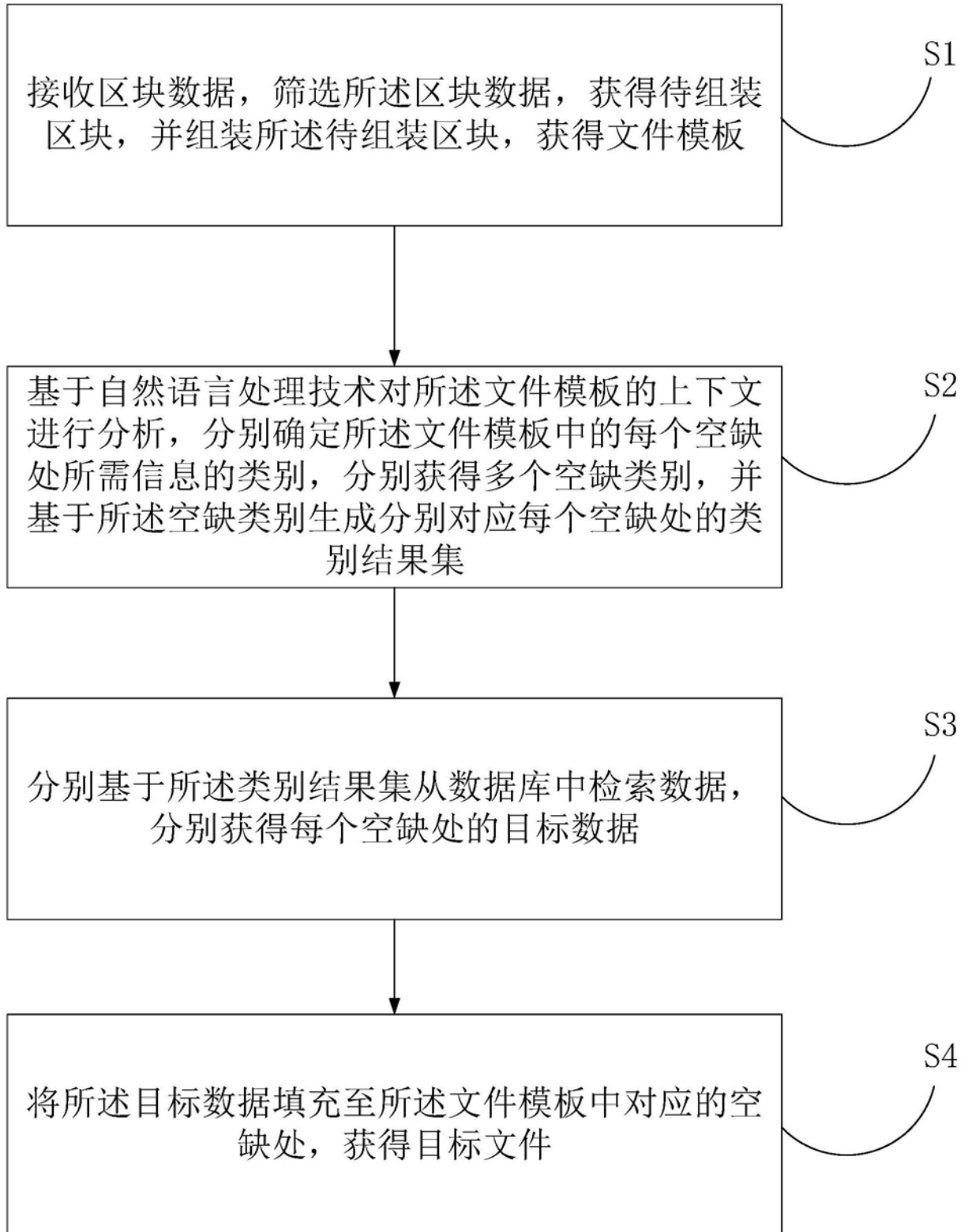


图2

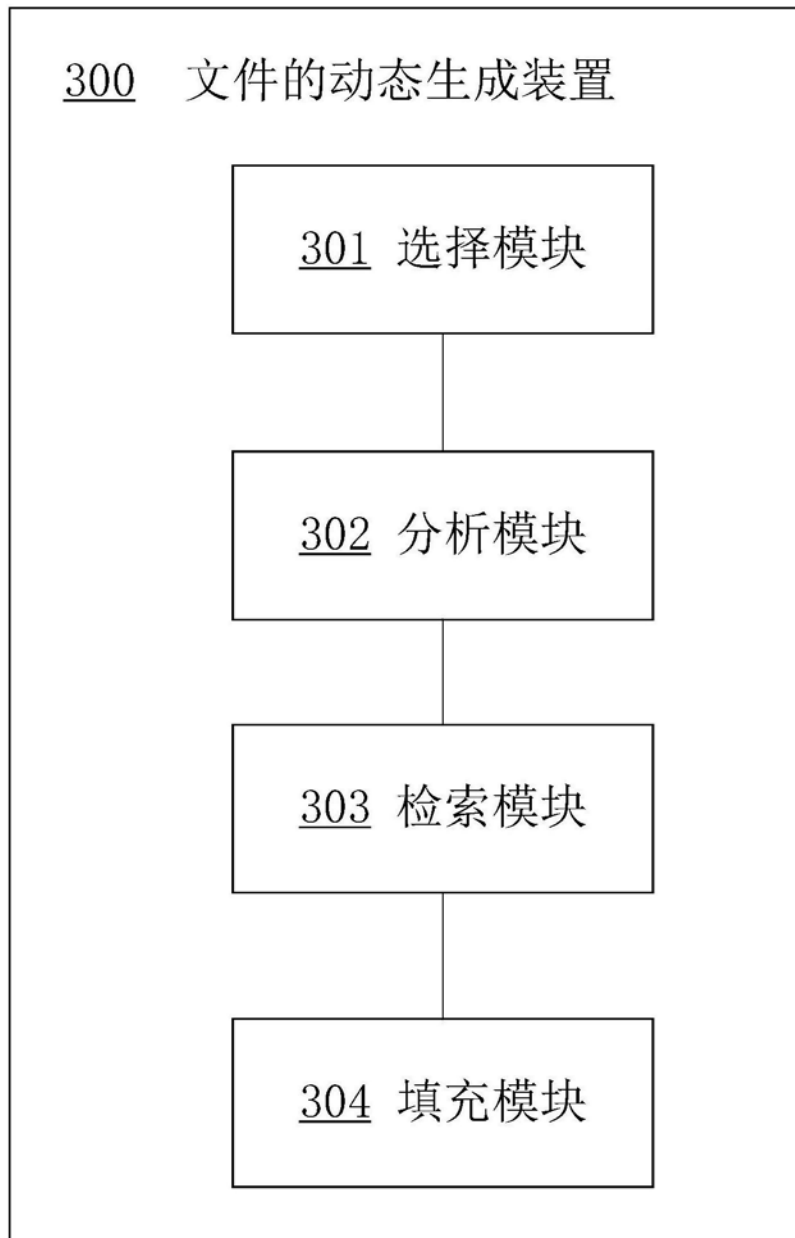


图3

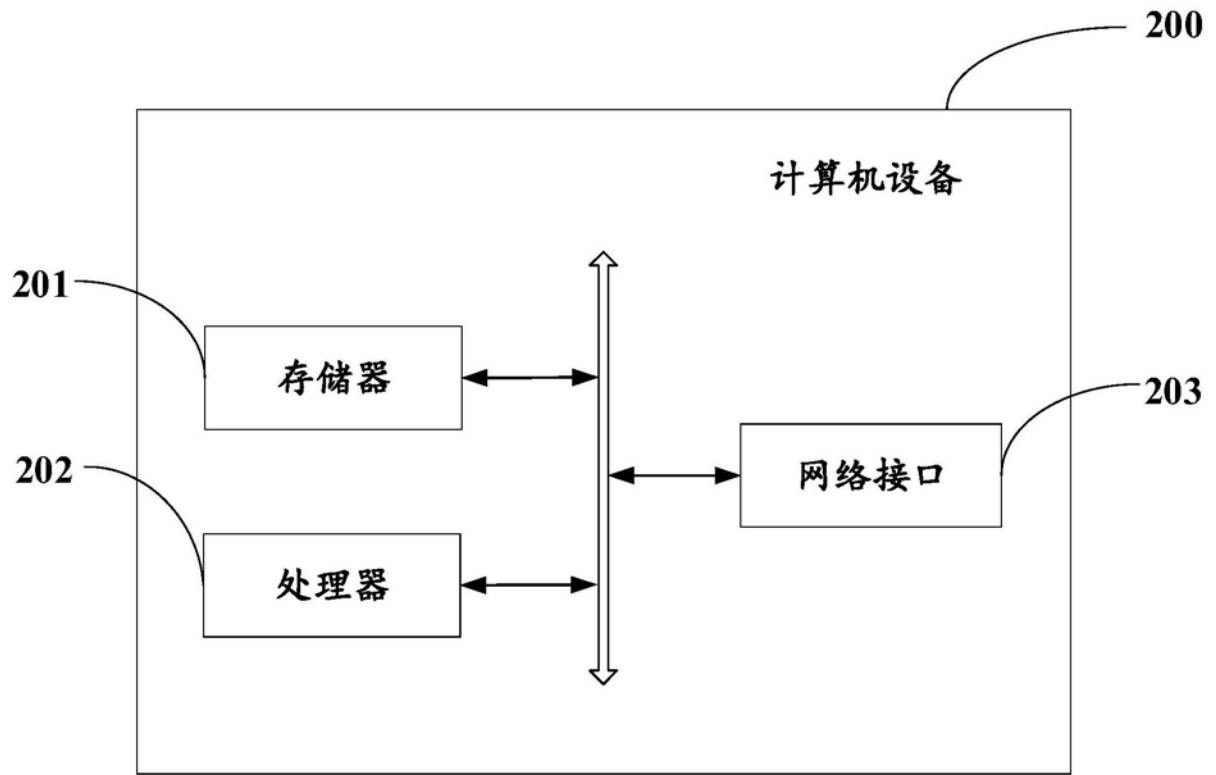


图4