



(21) 申请号 201810147204.6

(22) 申请日 2018.02.12

(65) 同一申请的已公布的文献号

申请公布号 CN 110209775 A

(43) 申请公布日 2019.09.06

(73) 专利权人 北京搜狗科技发展有限公司

地址 100084 北京市海淀区中关村东路1号
院9号楼搜狐网络大厦9层01房间

(72) 发明人 涂畅 周伟浩

(74) 专利代理机构 北京润泽恒知识产权代理有

限公司 11319

专利代理师 莎日娜

(51) Int. Cl.

G06F 16/332 (2019.01)

(56) 对比文件

CN 103024746 A, 2013.04.03

US 2016063879 A1, 2016.03.03

CN 105681609 A, 2016.06.15

CN 107612814 A, 2018.01.19

WO 2015184829 A1, 2015.12.10

汪卫明;陈世鸿;王世同;刘文印.基于语义
模板的医学问答自动生成.武汉大学学报(理学
版).2009,(第02期),110-115.

审查员 魏留强

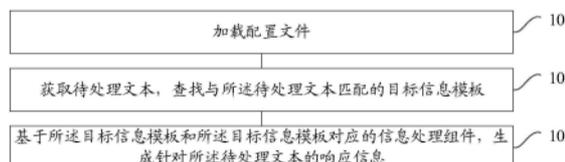
权利要求书3页 说明书15页 附图3页

(54) 发明名称

一种文本的处理方法和装置

(57) 摘要

本发明实施例提供了一种文本的处理方法和装置,该方法包括:加载配置文件,所述配置文件中配置有信息模板与信息处理组件;获取待处理文本,查找与所述待处理文本匹配的目标信息模板;基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。一方面,可以对用户需要监控的对象进行监控,自动识别文本并进行处理,无需用户手动处理,处理效率高,减少了处理的时延,保证文本及时处理,另一方面,处理文本的信息模板与信息处理组件可配置,使得可以不用关心处理文本的细节,大大提高了扩展性。



1. 一种文本的处理方法,其特征在于,包括:
加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
根据所述目标信息模板从所述待处理文本提取关键词;
将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
2. 根据权利要求1所述的方法,其特征在于,所述根据所述目标信息模板从所述待处理文本提取关键词,包括:
在所述目标信息模板中确定目标位置;
按照所述目标位置在所述待处理文本中提取关键词。
3. 根据权利要求1所述的方法,其特征在于,所述将所述关键词作为处理参数输入至所述目标信息类型对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:
将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成响应信息。
4. 根据权利要求3所述的方法,其特征在于,所述基于所述关键词确定文本处理信息,包括:
在预置的词库中查询是否存储有所述关键词;
当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;
当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否包括所述关键词;
若是,则确定文本处理信息为所述关键词为已挖掘的词组;
若否,则确定文本处理信息为未录入所述关键词。
5. 根据权利要求3所述的方法,其特征在于,所述基于所述文本处理信息生成响应信息,包括:
查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;
在所述响应模板中将所述关键词代替所述第一通配符,获得响应信息。
6. 根据权利要求5所述的方法,其特征在于,所述响应模板中具有第二通配符;所述基于所述文本处理信息生成响应信息,还包括:
查询所述待处理文本对应的用户信息;
在所述响应模板中将所述用户信息代替所述第二通配符。
7. 一种文本的处理装置,其特征在于,包括:
配置文件加载模块,用于加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
文本处理模块,用于获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
信息模板处理子模块,用于根据所述信息模板从所述待处理文本提取关键词;
信息处理组件调用子模块,用于将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
8. 根据权利要求7所述的装置,其特征在于,所述信息模板处理子模块包括:

目标位置确定单元,用于在所述目标信息模板中确定目标位置;

关键词提取单元,用于按照所述目标位置在所述待处理文本中提取关键词。

9.根据权利要求7所述的装置,其特征在于,所述信息处理组件调用子模块包括:

信息处理组件处理单元,将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。

10.根据权利要求9所述的装置,其特征在于,所述信息处理组件处理单元包括:

词库判断子单元,用于在预置的词库中查询是否存储有所述关键词;

第一文本处理信息确定子单元,用于当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

热词判断子单元,用于当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;若是,则调用第二文本处理信息确定子单元,若否,则调用第三文本处理信息确定子单元;

第二文本处理信息确定子单元,用于确定文本处理信息为所述关键词为已挖掘的词组;

第三文本处理信息确定子单元,用于确定文本处理信息为未录入所述关键词。

11.根据权利要求9所述的装置,其特征在于,所述信息处理组件处理单元包括:

响应模板查询子单元,用于查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;

关键词代替子单元,用于在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。

12.根据权利要求11所述的装置,其特征在于,所述响应模板中具有第二通配符;

所述信息处理组件处理单元还包括:

用户信息查询子单元,用于查询所述待处理文本对应的用户信息;

用户信息代替子单元,用于在所述响应模板中将所述用户信息代替所述第二通配符。

13.一种用于文本的处理的装置,其特征在于,包括有存储器,以及一个或者一个以上的程序,其中一个或者一个以上程序存储于存储器中,且经配置以由一个或者一个以上处理器执行所述一个或者一个以上程序包含用于进行以下操作的指令:

加载配置文件,所述配置文件中配置有信息模板与信息处理组件;

获取待处理文本,查找与所述待处理文本匹配的目标信息模板;

根据所述目标信息模板从所述待处理文本提取关键词;

将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

14.根据权利要求13所述的装置,其特征在于,所述根据所述目标信息模板从所述待处理文本提取关键词,包括:

在所述目标信息模板中确定目标位置;

按照所述目标位置在所述待处理文本中提取关键词。

15.根据权利要求13所述的装置,其特征在于,所述将所述关键词作为处理参数输入至所述目标信息类型对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:

将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成响应信息。

16.根据权利要求15所述的装置,其特征在于,所述基于所述关键词确定文本处理信息,包括:

在预置的词库中查询是否存储有所述关键词;

当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否包括所述关键词;

若是,则确定文本处理信息为所述关键词为已挖掘的词组;

若否,则确定文本处理信息为未录入所述关键词。

17.根据权利要求15所述的装置,其特征在于,所述基于所述文本处理信息生成响应信息,包括:

查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;

在所述响应模板中将所述关键词代替所述第一通配符,获得响应信息。

18.根据权利要求17所述的装置,其特征在于,所述响应模板中具有第二通配符;所述基于所述文本处理信息生成响应信息,还包括:

查询所述待处理文本对应的用户信息;

在所述响应模板中将所述用户信息代替所述第二通配符。

19.一个或多个机器可读介质,其上存储有指令,当由一个或多个处理器执行时,使得处理器执行如权利要求1-6一个或多个的方法。

一种文本的处理方法和装置

技术领域

[0001] 本发明涉及计算机处理的技术领域,特别是涉及一种文本的处理方法和装置。

背景技术

[0002] 在产品开发、项目管理、客户咨询等场景中,多个负责不同事项的用户通常组建群组,在该群组中进行不同事项的交流,或者,对客户咨询的问题进行回复。

[0003] 某些用户在处理事项发现问题之后,通常会将该问题发送到群组中,询问相关的用户。该相关的用户在看到问题之后,会手动进行处理,将相应的答案告知该用户。

[0004] 发明人在实现本发明的过程中发现,现有技术中用户一般手动处理其负责解答的问题,处理效率较低,并且,用户通常间隔一段时间查阅一次群组的聊天记录,查看到该问题存在时延现象,导致处理不及时。

发明内容

[0005] 鉴于上述问题,为了解决上述问题处理效率较低、处理不及时的问题,本发明实施例提出了一种文本的处理方法和装置。

[0006] 为了解决上述问题,本发明实施例公开了一种文本的处理方法,包括:

[0007] 加载配置文件,所述配置文件中配置有信息模板与信息处理组件;

[0008] 获取待处理文本,查找与所述待处理文本匹配的目标信息模板;

[0009] 基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0010] 可选地,所述基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:

[0011] 根据所述目标信息模板从所述待处理文本提取关键词;

[0012] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0013] 可选地,所述根据所述目标信息模板从所述待处理文本提取关键词,包括:

[0014] 在所述目标信息模板中确定目标位置;

[0015] 按照所述目标位置在所述待处理文本中提取关键词。

[0016] 可选地,所述将所述关键词作为处理参数输入至所述目标信息类型对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:

[0017] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成响应信息。

[0018] 可选地,所述基于所述关键词确定文本处理信息,包括:

[0019] 在预置的词库中查询是否存储有所述关键词;

[0020] 当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

- [0021] 当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否包括所述关键词;
- [0022] 若是,则确定文本处理信息为所述关键词为已挖掘的词组;
- [0023] 若否,则确定文本处理信息为未录入所述关键词。
- [0024] 可选地,所述基于所述文本处理信息生成响应信息,包括:
- [0025] 查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;
- [0026] 在所述响应模板中将所述关键词代替所述第一通配符,获得响应信息。
- [0027] 可选地,所述响应模板中具有第二通配符;所述基于所述文本处理信息生成响应信息,还包括:
- [0028] 查询所述待处理文本对应的用户信息;
- [0029] 在所述响应模板中将所述用户信息代替所述第二通配符。
- [0030] 本发明实施例还公开了一种文本的处理装置,包括:
- [0031] 配置文件加载模块,用于加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
- [0032] 文本处理模块,用于获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
- [0033] 响应信息生成模块,用于基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0034] 可选地,所述响应信息生成模块包括:
- [0035] 信息模板处理子模块,用于根据所述信息模板从所述待处理文本提取关键词;
- [0036] 信息处理组件调用子模块,用于将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0037] 可选地,所述信息模板处理子模块包括:
- [0038] 目标位置确定单元,用于在所述目标信息模板中确定目标位置;
- [0039] 关键词提取单元,用于按照所述目标位置在所述待处理文本中提取关键词。
- [0040] 可选地,所述信息处理组件调用子模块包括:
- [0041] 信息处理组件处理单元,将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。
- [0042] 可选地,所述信息处理组件处理单元包括:
- [0043] 词库判断子单元,用于在预置的词库中查询是否存储有所述关键词;
- [0044] 第一文本处理信息确定子单元,用于当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;
- [0045] 热词判断子单元,用于当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;若是,则调用第二文本处理信息确定子单元,若否,则调用第三文本处理信息确定子单元;
- [0046] 第二文本处理信息确定子单元,用于确定文本处理信息为所述关键词为已挖掘的词组;
- [0047] 第三文本处理信息确定子单元,用于确定文本处理信息为未录入所述关键词。

- [0048] 可选地,所述信息处理组件处理单元包括:
- [0049] 响应模板查询子单元,用于查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;
- [0050] 关键词代替子单元,用于在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。
- [0051] 可选地,所述响应模板中具有第二通配符;
- [0052] 所述信息处理组件处理单元还包括:
- [0053] 用户信息查询子单元,用于查询所述待处理文本对应的用户信息;
- [0054] 用户信息代替子单元,用于在所述响应模板中将所述用户信息代替所述第二通配符。
- [0055] 本发明实施例还公开了一种用于文本的处理的装置,包括有存储器,以及一个或者一个以上的程序,其中一个或者一个以上程序存储于存储器中,且经配置以由一个或者一个以上处理器执行所述一个或者一个以上程序包含用于进行以下操作的指令:
- [0056] 加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
- [0057] 获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
- [0058] 基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0059] 可选地,所述一个或者一个以上程序还包含用于进行以下操作的指令:
- [0060] 根据所述目标信息模板从所述待处理文本提取关键词;
- [0061] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0062] 可选地,所述一个或者一个以上程序还包含用于进行以下操作的指令:
- [0063] 在所述目标信息模板中确定目标位置;
- [0064] 按照所述目标位置在所述待处理文本中提取关键词信息模板信息模板。
- [0065] 可选地,所述一个或者一个以上程序还包含用于进行以下操作的指令:
- [0066] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。
- [0067] 可选地,所述一个或者一个以上程序还包含用于进行以下操作的指令:
- [0068] 在预置的词库中查询是否存储有所述关键词;
- [0069] 当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;
- [0070] 当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;
- [0071] 若是,则确定文本处理信息为所述关键词为已挖掘的词组;
- [0072] 若否,则确定文本处理信息为未录入所述关键词。
- [0073] 可选地,所述一个或者一个以上程序还包含用于进行以下操作的指令:
- [0074] 查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;
- [0075] 在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。
- [0076] 可选地,所述响应模板中具有第二通配符;

- [0077] 所述一个或者一个以上程序还包含用于进行以下操作的指令：
- [0078] 查询所述待处理文本对应的用户信息；
- [0079] 在所述响应模板中将所述用户信息代替所述第二通配符响应信息。
- [0080] 本发明实施例还公开了一个或多个机器可读介质，其上存储有指令，当由一个或多个处理器执行时，使得处理器执行上述一个或多个的方法。
- [0081] 本发明实施例包括以下优点：
- [0082] 本发明实施例预先加载配置文件，该配置文件中配置有用于处理某类型的文本的信息模板与信息处理组件，若获取待处理文本，则可以查找与该待处理文本匹配的目标信息模板，基于目标信息模板和目标信息模板对应的信息处理组件，生成针对待处理文本的响应信息，一方面，可以对用户需要监控的对象进行监控，自动识别文本并进行处理，无需用户手动处理，处理效率高，减少了处理的时延，保证文本及时处理，另一方面，处理文本的信息模板与信息处理组件可配置，使得可以不用关心处理文本的细节，大大提高了扩展性。

附图说明

- [0083] 图1是本发明的一种文本的处理方法实施例的步骤流程图；
- [0084] 图2是本发明的一种文本的处理装置实施例的结构框图；
- [0085] 图3是根据一示例性实施例示出的一种用于文本的处理的装置的框图；
- [0086] 图4是本发明实施例中服务器的结构示意图。

具体实施方式

[0087] 为使本发明的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0088] 参照图1，示出了本发明的一种文本的处理方法实施例的步骤流程图，具体可以包括如下步骤：

[0089] 步骤101，加载配置文件。

[0090] 在具体实现中，在本发明实施例中可以应用在独立的应用中，也可以为某个应用中的模块，该应用可以部署在服务器中，实现聊天机器人的至少部分功能、对指定的服务对象（如搜索引擎等）进行监控的监控程序，等等，也可以部署在移动终端中，对指定的服务对象（如即时通讯工具等）进行监控的监控程序，等等，本发明实施例对此不加以限制。

[0091] 针对该应用可以设置配置文件，如XML（eXtensible Markup Language，可扩展标记语言）、JSON（JavaScript Object Notation，JS对象标记）等格式的文件。

[0092] 在配置文件中配置有用于处理某类型的文本（如问题、咨询、建议等）的信息模板与信息处理组件，其中，信息模板和信息处理组件具有对应关系，每个信息模板与其对应的信息处理组件用于处理不同类型的问题。其中，文本的类型可以由本领域技术人员根据不同业务领域的需求进行设定，例如，在输入法领域中，该类型可以包括缺词类型、日活统计类型；在产品开发领域中，该类型可以包括项目进度类型，等等。

[0093] 信息模板可以为TXT（文本文件）等格式的文件，记录了识别某一类型的文本及提取该类型文本的关键词的规则。信息处理组件可以为shell脚本等格式的文件，可以独立进行运行，用于对某一类型的问题进行处理。

[0094] 该应用可以定义运行的框架,通过加载、调用配置文件中的信息模板、信息处理组件,对某一类型的文本进行处理。在该框架下,各个用户可以根据业务需求,按照该应用提供的规范,开发信息模板与信息处理组件,以便在该框架读取、调用。

[0095] 当接收配置指令时,则可以对配置文件中的信息模板与信息处理组件进行配置。若该配置指令为增加指令,则可以在配置文件中新增指定问题类型对应的信息模板与信息处理组件;若该配置指令为修改指令,则可以在配置文件中修改信息模板和/或信息处理组件;若该配置指令为删除指令,则可以在配置文件中删除指定问题类型对应的信息模板与信息处理组件。

[0096] 在启动应用时,可以将配置文件加载至内存中,以便后续使用。某一问题类型的问题可以配置一个唯一的ID,其对应一套信息模板、信息处理组件,在内存中可以建立映射关系。

[0097] 例如,某一类型的文本的ID为001,其对应的信息模板为module001.txt、信息处理组件为001.sh,在内存中即可建立001、module001.txt与001.sh之间的映射关系。

[0098] 步骤102,获取待处理的待处理文本,查找与所述待处理文本匹配的目标信息模板。

[0099] 该应用启动之后,可以针对指定的对象进行监控,获取相应的信息作为待处理文本。需要说明的是,在不同的应用场景中,具有不同的信息可以作为待处理文本。

[0100] 在一个示例中,针对即时通讯的应用场景,可以在即时通讯工具的会话窗口中提取通讯消息,作为待处理文本。

[0101] 在此示例中,监控的对象为即时通讯工具中的用户、群组等,实时提取接收到的通讯消息,作为待处理文本。

[0102] 当然,上述待处理文本的获取方式只是作为示例,在实施本发明实施例时,可以根据实际情况设置其他待处理文本的获取方式,例如,在输入法系统中,获取用户选定的候选词,作为待处理文本,或者,在搜索引擎、浏览器中,获取用户输入的搜索关键词作为待处理文本,等等,本发明实施例对此不加以限制。另外,除了上述待处理文本的获取方式外,本领域技术人员还可以根据实际需要采用其它待处理文本的获取方式,本发明实施例对此也不加以限制。

[0103] 在本发明实施例中,可以从内存中提取信息模板,将其与待处理文本进行匹配。若两者匹配成功,则可以确定该待处理文本归属于该信息模板对应的问题类型,该信息模板即为目标信息模板。若两者匹配失败,则提取下一个信息模板,继续进行匹配。

[0104] 例如,假设某一类型的文本(问题)为缺词类型,其信息模板中定义了匹配的规则,“###这个词没有”、“###这个词打不出来”等等。若当前在即时通讯工具的群组的会话窗口中提取了“代善这个词没有呢”作为待处理文本,与缺词类型的信息模板进行匹配,符合规则“###这个词没有”,因此,可以确定该待处理文本为属于缺词类型的问题。

[0105] 步骤103,基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0106] 本步骤中,确定了与待处理文本匹配的信息模板(即目标信息模板),相对应地,也确定了处理该待处理文本的信息处理组件,则可以调用该目标信息模板与该信息处理组件,对该待处理文本进行处理,获得问题的响应信息。

[0107] 具体地,响应信息可以是当前应用场景内全员可见的响应信息,也可以是提出问题的用户自身可见的响应信息,还可以是仅对与问题相关的用户可见的响应信息,本发明在此不做限定。

[0108] 在本发明的一个实施例中,步骤103可以包括如下子步骤:

[0109] 子步骤S11,根据所述目标信息模板从所述待处理文本提取关键词。

[0110] 在具体实现中,可以根据目标信息模板对获取的待处理文本进行识别,确定该待处理文本归属的问题类型,并从该待处理文本中提取关键词。

[0111] 在本发明的一个实施例中,子步骤S11可以包括如下子步骤:

[0112] 子步骤S111,在所述目标信息模板中确定目标位置。

[0113] 子步骤S112,按照所述目标位置在所述待处理文本中提取关键词。

[0114] 在本发明实施例中,可以通过指定的标识识别目标信息模板中的目标位置。按照目标信息模板定义的目标位置,在该待处理文本中提取相应的词组,作为关键词。

[0115] 例如,假设某一类型的文本(问题)为缺词类型,其目标信息模板中定义了匹配的规则,“###这个词没有”、“###这个词打不出来”等等,其中,“###”所处的位置为目标位置,即关键词为位于“这个词打不出来”之前的词组。

[0116] 若当前在即时通讯工具的群组的会话窗口中提取了“代善这个词没有呢”作为待处理文本,与缺词类型的目标信息模板进行匹配,符合规则“###这个词没有”,因此,可以确定该待处理文本为属于缺词类型的问题,关键词为“代善”,表示缺少“代善”这个词。

[0117] 子步骤S12,将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0118] 若确定了待处理文本的关键词,则可以以该关键词作为处理参数,调用该目标信息模板对应的信息处理组件。该信息处理组件接收到处理参数,则可以按照自身设置的逻辑对该处理参数进行处理,生成用于回复该待处理文本的响应信息。

[0119] 在具体实现中,可以将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成响应信息。

[0120] 进一步而言,信息处理组件可以基于该关键词进行处理,确定待处理文本对应的文本处理信息,再以该文本处理信息组装响应信息。

[0121] 在本发明实施例的一个示例中,该类型包括缺词类型,则在此示例中,目标信息模板对应的信息处理组件可以通过如下方式基于关键词确定文本处理信息:

[0122] 子步骤S31,在预置的词库中查询是否存储有所述关键词。

[0123] 子步骤S32,当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库。

[0124] 子步骤S33,当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;若是,则执行子步骤S34,若否,则执行子步骤S35。

[0125] 子步骤S34,确定文本处理信息为所述关键词为已挖掘的词组。

[0126] 子步骤S35,确定文本处理信息为未录入所述关键词。

[0127] 在输入法应用的应用场景中,在云端的服务器可以设置各种类型的词库,例如,记载基础字词的的系统词库,记载扩展字词的细胞词库(细胞词库中的字词至少具有一个共同

属性),记载某个游戏专用字词的细胞词库、记载生物专用字词的细胞词库等,记载用户个性化化字词(如自造词)的用户词库,等等。

[0128] 这些词库可以推送至各个输入法应用,用户输入该词库中的词条对应的编码信息(如拼音、笔画等),则可以显示该词条,用户点击后上屏。

[0129] 此外,为了扩充词库中的词条,服务器可以启动热词流程,在一定的时间段内通过时事新闻、论坛帖子等方式挖掘出一些热门的词组,例如,在商品销售旺季中,“剁手”、“快递”、“折扣”、“双十一”等词组较为热门,在某个电影上映的时间,电影名称、主角名字、“票房”等词组较为热门。这些热门的词组通过去重、人工鉴定、词频统计等方式进行筛选,筛选出的词组可以作为词条存储至词库中,其余词组则可以删除、丢弃。

[0130] 在本示例中,面对缺词类型的问题(即待处理文本),则可以访问云端的服务器,查询词库是否存储有该关键词。如果词库已存储该关键词,则可以确定文本处理信息为关键词已录入词库。如果词库未存储该关键词,则可以进一步查询该关键词是否为通过热词流程挖掘出的词组。如果该关键词为通过热词流程挖掘出的词组,则可以确定文本处理信息为关键词为已挖掘的词组(即热词),否则,确定文本处理信息为词库、热词均未录入该关键词。

[0131] 当然,上述确定文本处理信息的方式只是作为示例,在实施本发明实施例时,可以根据实际情况设置其他确定文本处理信息的方式,例如,面对日活统计类型的问题,其对应的问题处理组价可以统计日活跃用户的数量作为文本处理信息,等等,本发明实施例对此不加以限制。另外,除了上述确定文本处理信息的方式外,本领域技术人员还可以根据实际需要采用其它确定文本处理信息的方式,本发明实施例对此也不加以限制。

[0132] 在本发明实施例的一个示例中,信息处理组件可以通过如下方式基于文本处理信息生成响应信息:

[0133] 子步骤S41,查询所述文本处理信息信息处理组件对应的响应模板。

[0134] 子步骤S42,在所述响应模板中将所述关键词代替所述第一通配符,获得响应信息。

[0135] 在本发明实施例中,信息处理组件可以针对一种文本处理信息配置一种响应模板。该响应模板中具有第一通配符,将关键词代替响应模板中的第一通配符,则可以获得问题响应信息。

[0136] 例如,对于缺词类型的文本(问题)“代善这个词没有呢”,提取了关键词“代善”,若信息处理组件处理之后获得的文本处理信息为关键词为已挖掘的词组,则可以配置对应的响应模板为“(\d+)这个词在云端词库没有,但是,热词流程已经发现了这个词。”其中,“(\d+)”为第一通配符。将“代善”代替“(\d+)”,则可以获得问题响应信息“代善这个词在云端词库没有,但是,热词流程已经发现了这个词。”

[0137] 在本发明实施例的另一个示例中,信息处理组件可以通过如下方式基于文本处理信息生成响应信息:

[0138] 子步骤S51,查询所述信息处理组件对应的响应模板。

[0139] 子步骤S52,查询所述待处理文本对应的用户信息。

[0140] 子步骤S53,在所述响应模板中将所述用户信息代替所述第二通配符。

[0141] 子步骤S54,在所述响应模板中将所述关键词代替所述第一通配符,获得响应信

息。

[0142] 在本示例中,针对即时通讯等应用场景,在响应模板中,除了对关键词配置第一通配符之外,还可以对回复的对象配置第二通配符。此时,响应模板中具有第一通配符与第二通配符,一方面,将关键词代替响应模板中的第一通配符,另一方面,查询发送问题的通讯用户,将该通讯用户的用户信息代替响应模板中的第二通配符,则可以获得响应信息。

[0143] 需要说明的是,该通讯用户的用户信息,可以为即时通讯工具记录的好友昵称,也可以为信息处理组件针对该通讯用户设置的特定称呼,本发明实施例对此不加以限制。

[0144] 例如,对于缺词类型的文本(问题)“代善这个词没有呢”,提取了关键词“代善”,若信息处理组件处理之后获得的文本处理信息为关键词为已挖掘的词组,则可以配置对应的响应模板为“回(\e+),经查,(\d+)这个词在云端词库没有,但是,热词流程已经发现了这个词。”其中,“(\d+)”为第一通配符,“(\e+)”为第二通配符。

[0145] 一方面,将“代善”代替“(\d+)”,另一方面,查询发送该文本(问题)的通讯用户的ID为1234,好友昵称为“小琳”,信息处理组件对该通讯用户设置的特定称呼为“琳老板”,将“小琳”或“琳老板”代替“(\e+)”,则可以获得问题响应信息“回@小琳,经查,代善这个词在云端词库没有,但是,热词流程已经发现了这个词。”或者“回琳老板娘,经查,代善这个词在云端词库没有,但是,热词流程已经发现了这个词。”

[0146] 当然,上述生成响应信息的方式只是作为示例,在实施本发明实施例时,可以根据实际情况设置其他生成响应信息的方式,例如,针对日活统计类型的文本(问题),在响应模板中设置统计日期的通配符与统计数量的通配符,在响应模板中将统计日期、统计数量代替相应的通配符,获得响应信息,等等,本发明实施例对此不加以限制。另外,除了上述生成响应信息的方式外,本领域技术人员还可以根据实际需要采用其它生成问题响应信息的方式,本发明实施例对此也不加以限制。

[0147] 需要说明的是,除了基于关键词确定文本处理信息、基于文本处理信息生成问题响应信息之外,本领域技术人员还可以根据实际的业务需求,在信息处理组件中设置其他操作,本发明实施例对此不加以限制。

[0148] 例如,针对缺词类型的问题,若信息处理组件处理之后获得的文本处理信息为关键词为已挖掘的词组,除了生成问题响应信息之外,还可以设置负责运营的用户,通过邮件、短信等方式将问题与问题响应信息发送至该用户,提醒该用户进行处理。

[0149] 本发明实施例预先加载配置文件,该配置文件中配置有用于处理某类型的文本的信息模板与信息处理组件,若获取待处理文本,则可以查找与该待处理文本匹配的目标信息模板,基于目标信息模板和目标信息模板对应的信息处理组件,生成针对待处理文本的响应信息,一方面,可以对用户需要监控的对象进行监控,自动识别文本并进行处理,无需用户手动处理,处理效率高,减少了处理的时延,保证文本及时处理,另一方面,处理文本的信息模板与信息处理组件可配置,使得可以不用关心处理文本的细节,大大提高了扩展性。

[0150] 需要说明的是,对于方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明实施例并不受所描述的动作顺序的限制,因为依据本发明实施例,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作并不一定是本发明实施例所必须的。

[0151] 参照图2,示出了本发明的一种文本的处理装置实施例的结构框图,具体可以包括如下模块:

[0152] 配置文件加载模块201,用于加载配置文件,所述配置文件中配置有信息模板与信息处理组件;

[0153] 文本处理模块202,用于获取待处理文本,查找与所述待处理文本匹配的目标信息模板;

[0154] 响应信息生成模块203,用于基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0155] 在本发明的一个实施例中,所述响应信息生成模块203包括:

[0156] 信息模板处理子模块,用于根据所述目标信息模板从所述待处理文本提取关键词;

[0157] 信息处理组件调用子模块,用于将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0158] 在本发明的一个实施例中,所述待处理文本获取子模块包括:

[0159] 目标位置确定单元,用于在所述目标信息模板中确定目标位置;

[0160] 关键词提取单元,用于按照所述目标位置在所述待处理文本中提取关键词。

[0161] 在本发明的一个实施例中,所述信息处理组件调用子模块包括:

[0162] 信息处理组件处理单元,将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。

[0163] 在本发明的一个实施例中,所述信息处理组件处理单元包括:

[0164] 词库判断子单元,用于在预置的词库中查询是否存储有所述关键词;

[0165] 第一文本处理信息确定子单元,用于当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

[0166] 热词判断子单元,用于当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;若是,则调用第二文本处理信息确定子单元,若否,则调用第三文本处理信息确定子单元;

[0167] 第二文本处理信息确定子单元,用于确定文本处理信息为所述关键词为已挖掘的词组;

[0168] 第三文本处理信息确定子单元,用于确定文本处理信息为未录入所述关键词。

[0169] 在本发明的一个实施例中,所述信息处理组件处理单元包括:

[0170] 响应模板查询子单元,用于查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;

[0171] 关键词代替子单元,用于在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。

[0172] 在本发明的一个实施例中,所述响应模板中具有第二通配符;

[0173] 所述信息处理组件处理单元还包括:

[0174] 用户信息查询子单元,用于查询所述待处理文本对应的用户信息;

[0175] 用户信息代替子单元,用于在所述响应模板中将所述用户信息代替所述第二通配

符。

[0176] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0177] 图3是根据一示例性实施例示出的一种用于问题的处理的装置300的框图。例如,装置300可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等。

[0178] 参照图3,装置300可以包括以下一个或多个组件:处理组件302,存储器304,电源组件306,多媒体组件308,音频组件310,输入/输出(I/O)的接口312,传感器组件314,以及通信组件316。

[0179] 处理组件302通常控制装置300的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理元件302可以包括一个或多个处理器320来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件302可以包括一个或多个模块,便于处理组件302和其他组件之间的交互。例如,处理部件302可以包括多媒体模块,以方便多媒体组件308和处理组件302之间的交互。

[0180] 存储器304被配置为存储各种类型的数据以支持在设备300的操作。这些数据的示例包括用于在装置300上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器304可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0181] 电源组件306为装置300的各种组件提供电力。电源组件306可以包括电源管理系统,一个或多个电源,及其他与为装置300生成、管理和分配电力相关联的组件。

[0182] 多媒体组件308包括在所述装置300和用户之间提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件308包括一个前置摄像头和/或后置摄像头。当设备300处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0183] 音频组件310被配置为输出和/或输入音频信号。例如,音频组件310包括一个麦克风(MIC),当装置300处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器304或经由通信组件316发送。在一些实施例中,音频组件310还包括一个扬声器,用于输出音频信号。

[0184] I/O接口312为处理组件302和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0185] 传感器组件314包括一个或多个传感器,用于为装置300提供各个方面的状态评估。例如,传感器组件314可以检测到设备300的打开/关闭状态,组件的相对定位,例如所述

组件为装置300的显示器和小键盘,传感器组件314还可以检测装置300或装置300一个组件的位置改变,用户与装置300接触的存在或不存在,装置300方位或加速/减速和装置300的温度变化。传感器组件314可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件314还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件314还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0186] 通信组件316被配置为便于装置300和其他设备之间有线或无线方式的通信。装置300可以接入基于通信标准的无线网络,如WiFi, 2G或3G,或它们的组合。在一个示例性实施例中,通信部件316经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信部件316还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0187] 在示例性实施例中,装置300可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0188] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器304,上述指令可由装置300的处理器320执行以完成上述方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0189] 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移动终端的处理器执行时,使得移动终端能够执行一种文本的处理方法,所述方法包括:

[0190] 加载配置文件,所述配置文件中配置有信息模板与信息处理组件;

[0191] 获取待处理文本,查找与所述待处理文本匹配的目标信息模板;

[0192] 基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0193] 可选地,所述基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:

[0194] 根据所述目标信息模板从所述待处理文本提取关键词;

[0195] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0196] 可选地,所述根据所述信息模板从所述待处理文本提取关键词信息模板关键词,包括:

[0197] 在所述目标信息模板中确定目标位置;

[0198] 按照所述目标位置在所述待处理文本中提取关键词。

[0199] 可选地,所述将所述关键词作为处理参数输入至所述目标信息类型对应的信息处理组件,关键词生成针对所述待处理文本生成问题响应信息,包括:

[0200] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。

[0201] 可选地,所述基于所述关键词确定文本处理信息,包括:

- [0202] 在预置的词库中查询是否存储有所述关键词;
- [0203] 当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;
- [0204] 当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;
- [0205] 若是,则确定文本处理信息为所述关键词为已挖掘的词组;
- [0206] 若否,则确定文本处理信息为未录入所述关键词。
- [0207] 可选地,所述基于所述文本处理信息生成问题响应信息,包括:
- [0208] 查询所述信息处理组件文本处理信息对应的响应模板,所述响应模板中具有第一通配符;
- [0209] 在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。
- [0210] 可选地,所述响应模板中具有第二通配符;
- [0211] 信息处理组件文本处理信息响应信息查询发送所述待处理文本对应的用户信息;
- [0212] 在所述响应模板中将所述用户信息代替所述第二通配符响应信息。
- [0213] 图4是本发明实施例中服务器的结构示意图。该服务器400可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上中央处理器(central processing units, CPU)422(例如,一个或一个以上处理器)和存储器432,一个或一个以上存储应用程序442或数据444的存储介质430(例如一个或一个以上海量存储设备)。其中,存储器432和存储介质430可以是短暂存储或持久存储。存储在存储介质430的程序可以包括一个或一个以上模块(图示没标出),每个模块可以包括对服务器中的一系列指令操作。更进一步地,中央处理器422可以设置为与存储介质430通信,在服务器400上执行存储介质430中的一系列指令操作。
- [0214] 服务器400还可以包括一个或一个以上电源426,一个或一个以上有线或无线网络接口450,一个或一个以上输入输出接口458,一个或一个以上键盘456,和/或,一个或一个以上操作系统441,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™等等。
- [0215] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本发明的其它实施方案。本发明旨在涵盖本发明的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本发明的真正范围和精神由下面的权利要求指出。
- [0216] 应当理解的是,本发明并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制
- [0217] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。
- [0218] 本发明实施例公开了A1、一种文本的处理方法,包括:
- [0219] 加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
- [0220] 获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
- [0221] 基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

- [0222] A2、根据A1所述的方法,所述基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:
- [0223] 根据所述目标信息模板从所述待处理文本提取关键词;
- [0224] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0225] A3、根据A2所述的方法,所述根据所述目标信息模板从所述待处理文本提取关键词,包括:
- [0226] 在所述目标信息模板中确定目标位置;
- [0227] 按照所述目标位置在所述待处理文本中提取关键词。
- [0228] A4、根据A2所述的方法,所述将所述关键词作为处理参数输入至所述目标信息类型对应的信息处理组件,生成针对所述待处理文本的响应信息,包括:
- [0229] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成响应信息。
- [0230] A5、根据A4所述的方法,所述基于所述关键词确定文本处理信息,包括:
- [0231] 在预置的词库中查询是否存储有所述关键词;
- [0232] 当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;
- [0233] 当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否包括所述关键词;
- [0234] 若是,则确定文本处理信息为所述关键词为已挖掘的词组;
- [0235] 若否,则确定文本处理信息为未录入所述关键词。
- [0236] A6、根据A4所述的方法,所述基于所述文本处理信息生成响应信息,包括:
- [0237] 查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;
- [0238] 在所述响应模板中将所述关键词代替所述第一通配符,获得响应信息。
- [0239] A7、根据A6所述的方法,所述响应模板中具有第二通配符;所述基于所述文本处理信息生成响应信息,还包括:
- [0240] 查询所述待处理文本对应的用户信息;
- [0241] 在所述响应模板中将所述用户信息代替所述第二通配符。
- [0242] 本发明实施例还公开了B8、一种文本的处理装置,包括:
- [0243] 配置文件加载模块,用于加载配置文件,所述配置文件中配置有信息模板与信息处理组件;
- [0244] 文本处理模块,用于获取待处理文本,查找与所述待处理文本匹配的目标信息模板;
- [0245] 响应信息生成模块,用于基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。
- [0246] B9、根据B8所述的装置,所述响应信息生成模块包括:
- [0247] 信息模板处理子模块,用于根据所述目标信息模板从所述待处理文本提取关键词;
- [0248] 信息处理组件调用子模块,用于将所述关键词作为处理参数输入至所述目标信息

模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0249] B10、根据B9所述的装置,所述信息模板处理子模块包括:

[0250] 目标位置确定单元,用于在所述目标信息模板中确定目标位置;

[0251] 关键词提取单元,用于按照所述目标位置在所述待处理文本中提取关键词。

[0252] B11、根据B9所述的装置,所述信息处理组件调用子模块包括:

[0253] 信息处理组件处理单元,将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。

[0254] B12、根据B11所述的装置,所述信息处理组件处理单元包括:

[0255] 词库判断子单元,用于在预置的词库中查询是否存储有所述关键词;

[0256] 第一文本处理信息确定子单元,用于当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

[0257] 热词判断子单元,用于当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;若是,则调用第二文本处理信息确定子单元,若否,则调用第三文本处理信息确定子单元;

[0258] 第二文本处理信息确定子单元,用于确定文本处理信息为所述关键词为已挖掘的词组;

[0259] 第三文本处理信息确定子单元,用于确定文本处理信息为未录入所述关键词。

[0260] B13、根据B11所述的装置,所述信息处理组件处理单元包括:

[0261] 响应模板查询子单元,用于查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;

[0262] 关键词代替子单元,用于在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。

[0263] B14、根据B13所述的装置,所述响应模板中具有第二通配符;

[0264] 所述信息处理组件处理单元还包括:

[0265] 用户信息查询子单元,用于查询所述待处理文本对应的用户信息;

[0266] 用户信息代替子单元,用于在所述响应模板中将所述用户信息代替所述第二通配符。

[0267] 本发明实施例还公开了C15、一种用于文本的处理的装置,包括有存储器,以及一个或者一个以上的程序,其中一个或者一个以上程序存储于存储器中,且经配置以由一个或者一个以上处理器执行所述一个或者一个以上程序包含用于进行以下操作的指令:

[0268] 加载配置文件,所述配置文件中配置有信息模板与信息处理组件;

[0269] 获取待处理文本,查找与所述待处理文本匹配的目标信息模板;

[0270] 基于所述目标信息模板和所述目标信息模板对应的信息处理组件,生成针对所述待处理文本的响应信息。

[0271] C16、根据C15所述的装置,所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0272] 根据所述目标信息模板从所述待处理文本提取关键词;

[0273] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,生成

针对所述待处理文本的响应信息。

[0274] C17、根据C16所述的装置,所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0275] 在所述目标信息模板中确定目标位置;

[0276] 按照所述目标位置在所述待处理文本中提取关键词信息模板信息模板。

[0277] C18、根据C16所述的装置,所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0278] 将所述关键词作为处理参数输入至所述目标信息模板对应的信息处理组件,以基于所述关键词确定文本处理信息、以及、基于所述文本处理信息生成问题响应信息。

[0279] C19、根据C18所述的装置,所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0280] 在预置的词库中查询是否存储有所述关键词;

[0281] 当所述词库已存储所述关键词时,则确定文本处理信息为所述关键词已录入所述词库;

[0282] 当所述词库未存储所述关键词时,在预设时间段内挖掘的词组中判断是否具有所述关键词;

[0283] 若是,则确定文本处理信息为所述关键词为已挖掘的词组;

[0284] 若否,则确定文本处理信息为未录入所述关键词。

[0285] C20、根据C18所述的装置,所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0286] 查询所述信息处理组件对应的响应模板,所述响应模板中具有第一通配符;

[0287] 在所述响应模板中将所述关键词代替所述第一通配符,获得问题响应信息。

[0288] C21、根据C20所述的装置,所述响应模板中具有第二通配符;

[0289] 所述一个或者一个以上程序还包含用于进行以下操作的指令:

[0290] 查询所述待处理文本对应的用户信息;

[0291] 在所述信息模板中将所述用户信息代替所述第二通配符响应信息。

[0292] 本发明实施例还公开了D22、一个或多个机器可读介质,其上存储有指令,当由一个或多个处理器执行时,使得处理器执行如A1-A7一个或多个的方法。

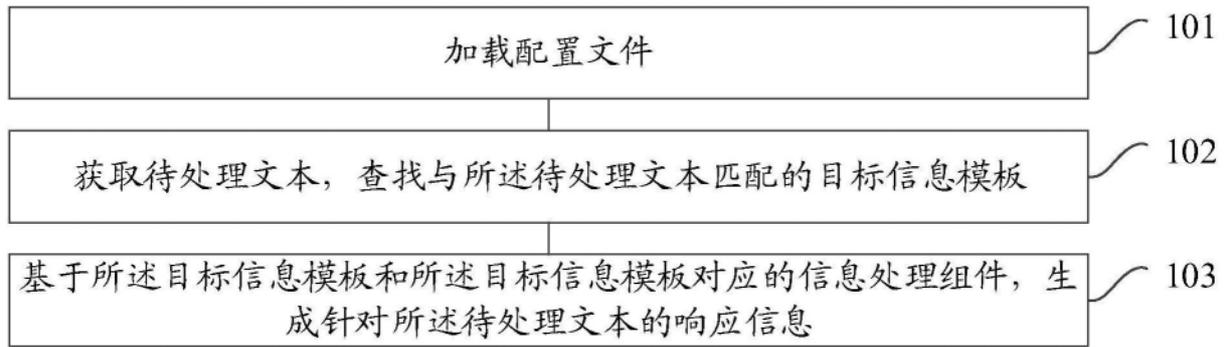


图1



图2

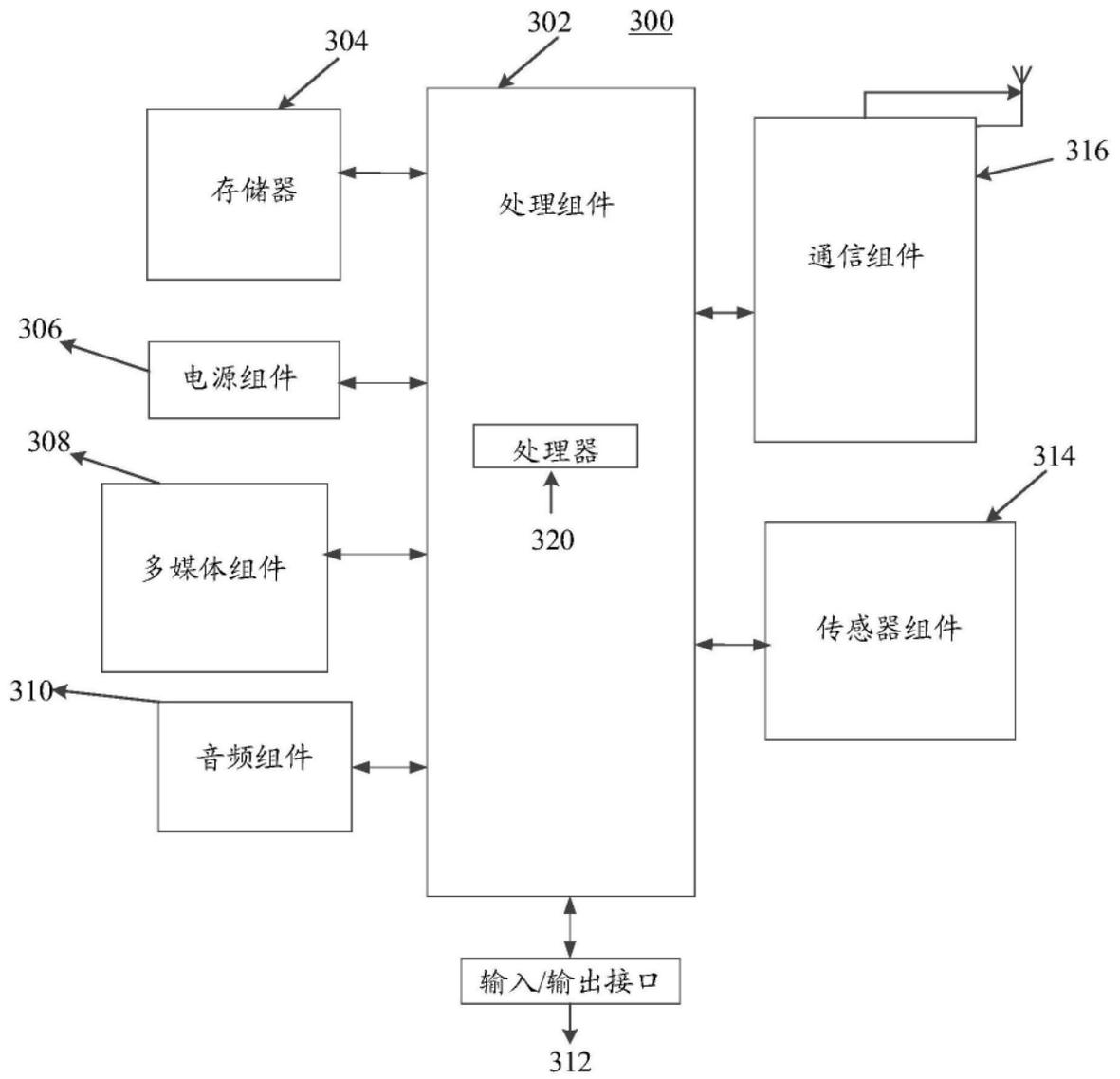


图3

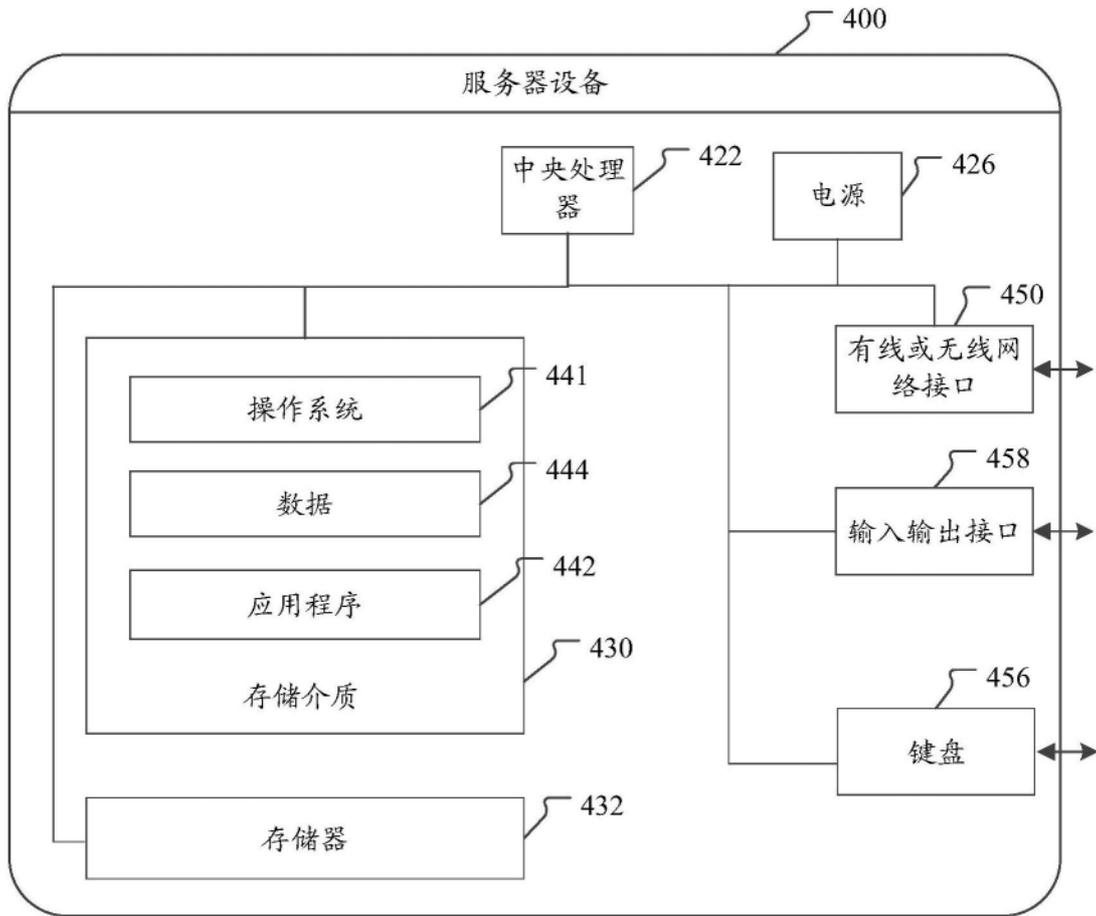


图4