**端口地址 192.168.60.10:25870**

1. term\_segment

|  |  |
| --- | --- |
| api | 192.168.60.10:25870/term\_segment |
| 请求方式 | Post |
| 描述 | 对原词汇进行切分，返回核心词及搜索词 |
| 输入 | Json, {‘original\_word’: xxx} |
| 输出 | Json, {‘names’: names, ‘core’: core}  names为搜索词( ods\_hospital\_gather , word字段)  core为核心词（ ods\_hospital\_gather ， word\_key字段） |
| 使用条件（场景） | 对17万原词汇，提取核心词并构建搜索词。 |
| 备注 |  |

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

1. word\_cut

|  |  |
| --- | --- |
| api | 192.168.60.10:25870/word\_cut |
| 请求方式 | Post |
| 描述 | 主要针对已经爬取过的,且无返回结果的情况。  核心思路就是对词进行拆解。策略：1.基于互信息的词条拆解（类似于长词拆解）； 2.词语替换（2-1：基于核心词重新组合；2-2：关键词的替换） |
| 输入 | Json，{‘current\_word’:xxx, ‘original\_word’:xxx, ‘core’:xxx, ‘all\_word’:xxxx}，  Current\_word: str, 当前的搜索词( ods\_hospital\_gather , word字段)  Original\_word：str,原词汇(( ods\_hospital\_gather , health\_institution\_name字段)  Core：str, 核心词(ods\_hospital\_gather ， word\_key字段)  All\_word: **list,** 所有的搜索记录(ods\_hospital\_gather ， word\_all字段) |
| 输出 | Json, {‘current\_word’:xxxx}  Current\_word: str,下一个搜索词（ ods\_hospital\_gather , word字段）  **无新增词汇时，返回{‘current\_word’:None}** |
| 使用条件（场景） | (ods\_wjwtask\_hospital\_table ，state=1) & (ods\_wjwtask\_hospital\_table ,remark 非空) |
| 备注 |  |

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

1. word\_addition

|  |  |
| --- | --- |
| api | 192.168.60.10:25870/word\_addition |
| 请求方式 | Post |
| 描述 | 主要针对已经爬取过的(flag==1)且返回结果>20的情况。核心策略就是按互信息从两侧选词. |
| 输入 | Json, {‘original\_word’: xxx, ‘current\_word’: xxx}  Current\_word: str, 当前的搜索词( ods\_hospital\_gather , word字段)  Original\_word：str,原词汇(( ods\_hospital\_gather , health\_institution\_name字段) |
| 输出 | Json, {‘current\_word’: xxxx}  Current\_word: str,下一个搜索词（ ods\_hospital\_gather , word字段） |
| 使用条件（场景） | (ods\_wjwtask\_hospital\_table ，state=1) & (ods\_wjwtask\_hospital\_table ,num>20) |
| 备注 |  |

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成