**信息搜索接口设计**

# 管理种子网站

1. **获取种子网站列表**

请求：/GetSeedUrls

参数：无

方式：GET

返回：一个JSONObject的序列化串。格式示例如下：

{"urls":[

"http://www.sina.com.cn",

"http://www.google.com",

"http://www.apache.org"]

}

1. **提交种子网站列表**

请求：/SaveSeedUrls?data=[种子网站列表数据]

方式：POST

参数：data，一个JSONObject的序列化串。种子网站列表数据格式示例同上

返回：一个JSONObject的序列化串。格式示例如下：

{"passed":[true/false], "message":[文本消息串]}

# 信息采集管理

1. **控制爬虫操作**

请求：/CrawlerAction?action=[ACTION]

方式：GET

参数：action，为“START”，“PAUSE”或“STOP”三个串中的任意一个

返回：一个JSONObject的序列化串。格式示例如下：

{"passed":[true/false], "message":[文本消息串]}

# 采集日志管理

1. **获取最近一次的采集日志信息**

请求：/GetCrawlLog

参数：无

方式：GET

返回：一个JSONObject的序列化串。格式示例如下：

[

{

"url":"http://www.sina.com.cn",

"passed":"false",

"message":"Can not downloads"

},

……

]

# 信息搜索操作

1. **获取搜索结果总数**

请求：/SearchCount?keywords=[关键字]

参数：keywords，多个关键字时中间以空格分隔

方式：GET

返回：一个数字，表示符合搜索条件的文档数

1. **搜索操作**

请求：/Search?page=[页码]&keywords=[关键字]

参数：

page，下标从0开始的页码

keywords，多个关键字时中间以空格分隔

方式：GET

返回：一个JSONObject的序列化串。格式示例如下：

{

"total":101,

"results":[

{

"title":"[文章标题]",

"author":"[文章作者]",

"abstract":"[摘要]",

"url":"[原文的URL]"

}

……

]}

*摘要内容为（关键字周围的一段文字，不超过256个字节）*