**功能**

1. **文章收藏（前后端配合）**
   1. 截屏获取文章部分内容，利用OCR接口识别文字后发送给服务器，服务器接收文字后进行搜索，利用爬虫爬文章，然后转成合适的JSON格式传回前端。
   2. 初步想到的是先做微信、知乎文章的收藏，利用搜狗的接口，搜索并且爬出文章内容，我看了一下html，因为都是微信、知乎的文章，所以非常的规则，id都标好了。其他的网站写一种类型的爬虫不一定都能爬得下来文章内容。

<http://weixin.sogou.com/weixin?type=2&query=%E5%A4%A7%E9%B9%8F&ie=utf8>

type类型2为文章，query为搜索标题

http://zhihu.sogou.com/zhihu?query=vr&ie=utf8

* 1. 查找html中文章中的图片元素，如果有的话就作为文章内容的引导图片，没有就不返回，这在前端卡片显示的时候效果会很好（和知乎首页一样的）
  2. 第一迭代我们先可以将NLP的分类贴标签功能作为一个供体验的功能。将手动的标签分类作为主要的功能

1. **文章搜索（前端完成）**
   1. 前端也建立一个数据库，保存本地数据，每次有网络时也能做到同步
2. **文章推荐（前后端配合）**
   1. 如果有用户模块的话就根据用户上传的搜索文字来匹配相关内容并且推送
   2. 如果不打算有用户模块的话，前端在本地存一个唯一的字符串进行用户匹配