

# SJTU\_soso需求文档

郭建铭/廖宁祎/刘浩文/江浩宇/郭子豪

## 1 需求综述

对交大同学来说，学校的网站、通知平台众多，且没有较好的分类消息通知平台，给同学们及时地、准确地获取并回顾相关新闻、通知等增加了难度。学校的网站包括[教学信息服务网站](#)、[教务处网站](#)、[门户网站](#)、[新闻学术网](#)、[大学生创新实践网](#)等等，学院的网站包括[电院官网](#)、[电院本科教务办](#)、[学生工作办](#)等等，还有众多的官方公众号，如上海交通大学、交大益友、交大源源、交大思思、电院E家人、上海交大电院学生会等。面对如此多的平台，以下情形变得十分常见，想寻找某事件相关的通知（如选课、考试通知等），却不知道发布在哪个平台，只能一个一个平台检索，还担心自己错过了某则重要通知；有印象曾经看过某则通知，但是想要去找的时候很难记住在哪儿看到过，因为没有搜索平台而词条又不够准确无法用百度等搜索引擎，同样只能一个一个网站检索.....

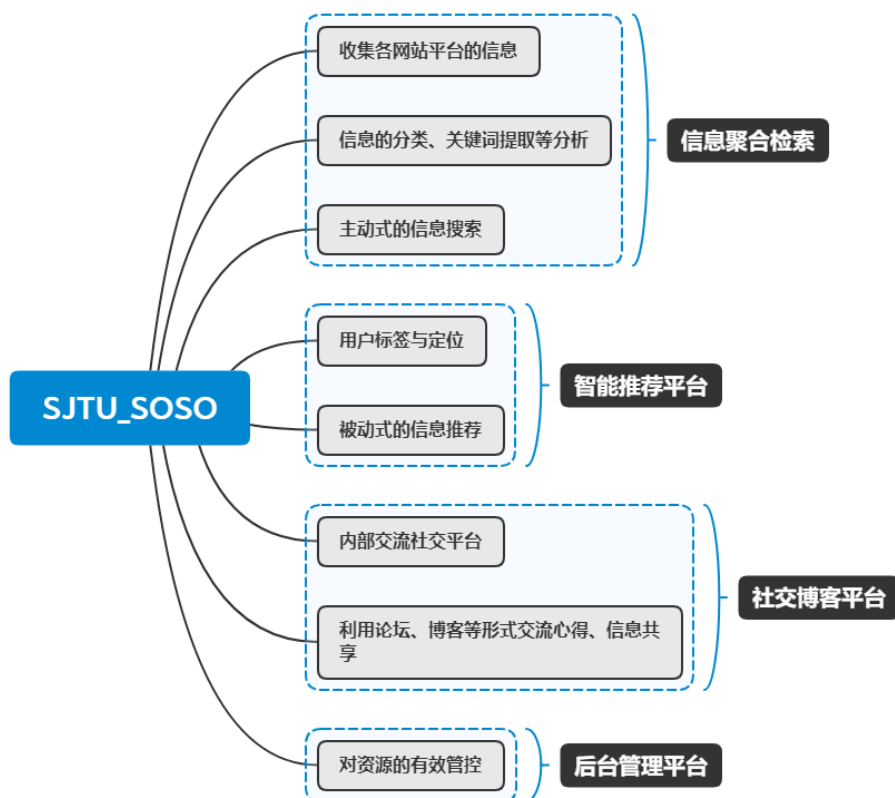
疫情期间，大家分散在各地，所有的学习、交流、活动均在线上完成，消息、通知的获取显得更为重要。但不同于线下，一旦错失了消息将无人提醒。一个基于交大所有网站信息建立的**搜索引擎**就能很好地解决上述痛点。此外，除了搜索引擎这种**主动信息获取**方式，根据**用户定位**进行相关的通知、新闻的推荐也是一种很好的解决方式。为了提高这种**被动信息获取**方式的信息传递效率，选取的信息传递媒介包括文章、图片、视频。同时，在短视频风靡的今天，视频、图片等往往能给人们带来更强烈的视觉冲击，留下更深的印象，能够传递的信息也更多。在获取用户定位后，这种智能信息推荐想用户所想，进一步提高用户获取相关信息的准确度、广度和效率。

交大人一直苦于没有属于自己的论坛性质的交流平台。校内的热点事件、具有争议性的话题、可供学习交流的博客等无处分享，只能去选择更为开放的交流平台，这与部分话题的私密性相违背。有些话题也因为无处分享被搁置，信息没有得到很好的流动。而且因为外界的论坛等面向全社会，流量更大，受众更广，很多话题往往还没得到及时的互动与回复就被淹没，平台众多也给想要搜索类似的博客的其他用户提供了难度。为解决此难题，首先是建立**用户系统**，基于此用户系统对用户进行定位并实现**社交**的功能，基于用户标签推荐相应的志同道合、有着共同兴趣的达人。基于此系统建立**论坛和博客平台**，众多话题可以在保证相对私密的情况下进行交流和分享，又因为大家经历的相似性也能引起更强烈的共鸣。通过博客平台可以发布博客，以交流某课程的心得、共享资料、分享就某一科研题目的心得等。而通过论坛平台可以就某一热点事件合理发表自己的想法、针对某些通知和新闻进行推广、为最近的烦恼发发牢骚寻求安慰等。其他用户则可以利用对博客和论坛的帖子点赞、评论、分享的功能进行社交、共享信息、合理表达诉求等。

既然涉及到信息的发布等，为了保证环境的健康以及更好地传播积极的正能量，就需要一个完善的**后台管理平台**进行上传内容的管控、资源的管理等。此外借助此平台还可以实时更新数据库的资源，管控对用户的定位、聚类分析、审查用户的评论等等。

## 2 功能需求

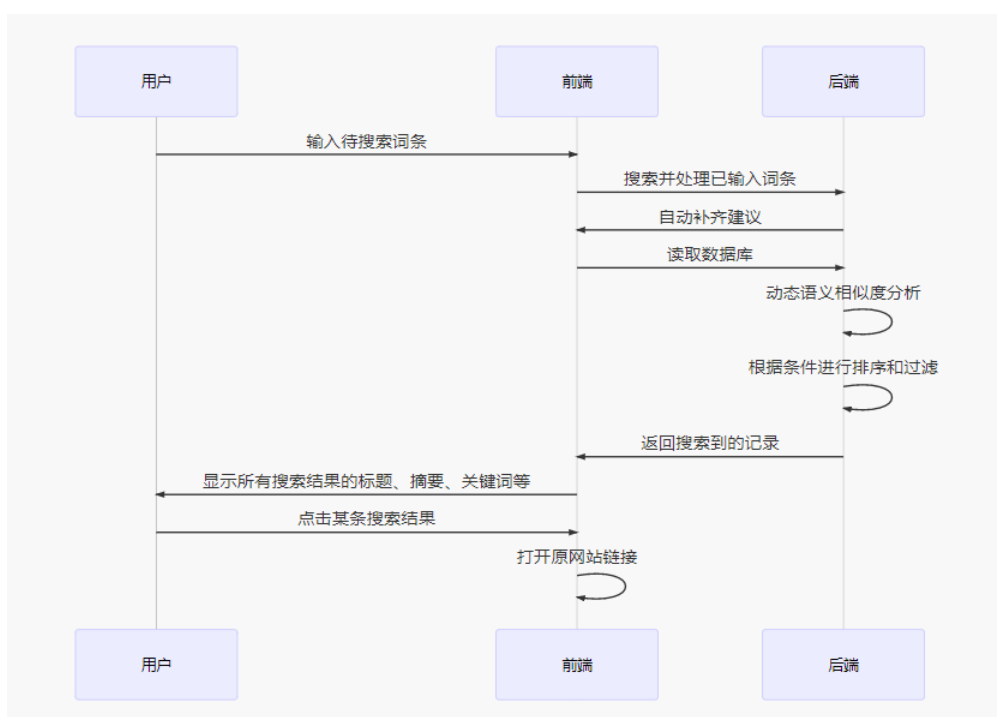
针对上述需求，本项目共分为四个模块，分别为**信息聚合检索**、**智能推荐平台**、**社交博客平台**、**后台管理平台**。功能需求示意图如图所示，按照模块对相关功能需求做详细说明。



## 2.1 信息聚合检索

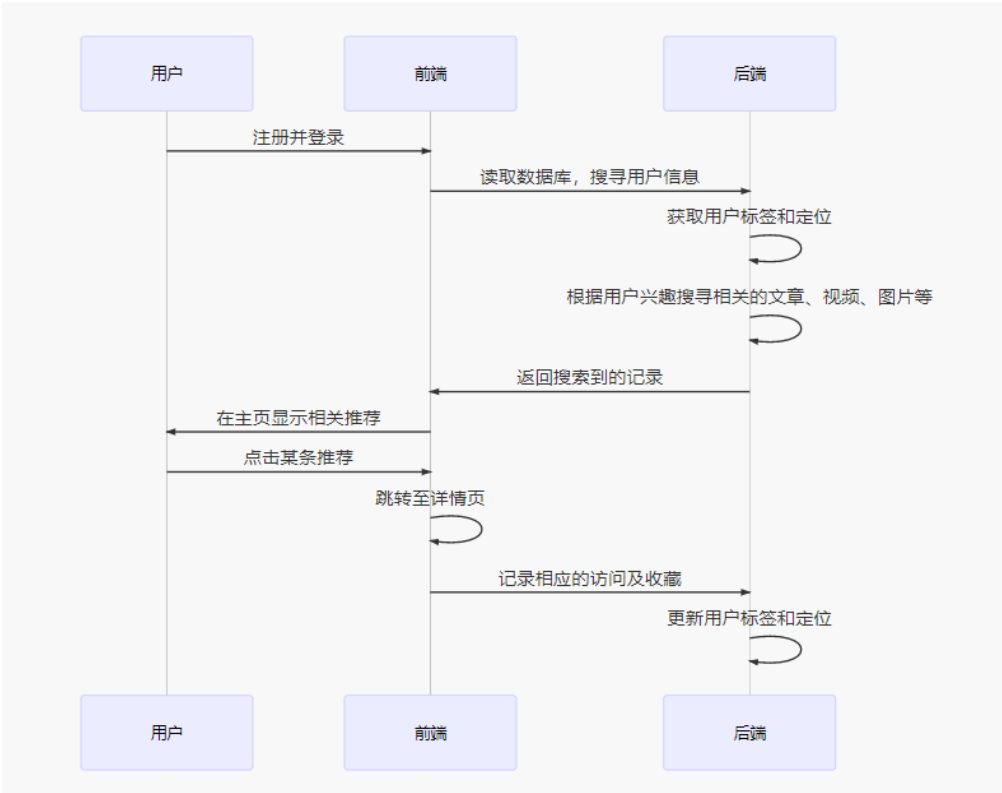
此模块首先需要针对学校和学院的相关网站以及公众号等做信息的聚合，聚合内容包括文字、图文、视频等，针对每一条信息需要做一定的分析，包括发布时间、浏览量、关键词提取等等。

信息检索部分需要搭建智能搜索引擎，针对用户输入的关键词在文本框中进行智能填补建议，对搜索到的结果进行条件排序、过滤等并按照响应要求打印在屏幕上，每一则信息包括标题、发布日期、关键词组、摘要以及浏览量等，点击标题即可链接至源网页。输出结果时对搜索的目标词条高亮显示。



## 2.2 智能推荐平台

首先需要对用户进行定位，通过用户的个人信息，平时的浏览习惯等对用户打标签，并在此基础上分别对文本、视频、图片等做推荐。并通过比较不同用户的行为来推荐相似度较高的用户。用户在登陆后即可在主页显示推荐的文章、视频、用户等等。



### 2.3 社交博客平台

此模块首先需要建立用户系统，用户可以实现注册、登录、个人信息完善等。用户的个人空间包括昵称、头像、自我介绍、用户互访等等。在此基础上需要制作博客平台和论坛平台。博客平台需要实现发布博客、收藏博客、评论博客等功能，在用户界面还有个人的博客列表、收藏列表，在主页界面按照博客的类型将所有博客分类展示，博客的详情页包含内容、作者、发布时间、点赞、收藏和评论等。与博客平台类似，论坛平台在用户登陆后，主页显示热门的主题帖和发帖的窗口，用户在详情页可以看到帖子内容、作者、发帖回帖时间、点赞、收藏和评论等。

### 2.4 后台管理平台

此平台需要实现相应的管理员功能，包括资源管理、用户聚类、评论审查。资源管理包括查看所有的内容、统计相关数据、删除部分资源、上传更新资源等。用户聚类需要实现评分数据读取、聚类分析和输出等。评论审查包括敏感条目、垃圾问本检测等。管理员可以看到所有资源的标题、浏览量、收藏量等，可以进行资源审查、删除等操作，对资源数目、用户数目、收藏数等进行统计分析。在用户聚类界面，管理员可以获取所有的用户信息并根据聚类结果推荐相关用户，在评论审查界面可以智能进行敏感识别敏感条目和垃圾问本，确认后递交给管理员审核并删除。

## 3 非功能性需求

- 高性能

为了满足用户的日常使用，相关功能需完备且运行稳定，网站的响应跳转时间等应满足用户的日常使用需求。

- 可维护性

相关功能易于管理，内容的发布需要经过一定的审阅和管控。所有的数据信息应保存于远端数据库以实现定时更新和维护的需求。

- 实时性

相关信息应满足实时性要求，信息聚合应定时更新以保证搜索引擎结果的及时性和准确性。用户的信息应及时更新以满足使用需求。论坛、博客的发布以及评论、收藏等也应实时响应和显示，以提高用户使用的流畅性。

- 兼容性

为满足日常使用需求，项目应满足跨平台、跨终端使用的要求，对运行环境无过高要求。

- 安全性

最后的项目应满足一定的安全需求，针对SQL注入攻击，XSS攻击，暴力破解等常规攻击方法有一定的防范机制。

- 并发性

考虑到多用户同时使用的情况，本项目需要满足一定的并发性，对部分功能做并发处理。

## 4 设计实现上的限制

- 设计开发语言

本项目旨在为用户提供信息搜索和社交娱乐的平台，对项目的兼容性、稳定性和可维护性都有较高要求，所以本项目选择基于**B/S架构**进行开发。在具体的开发过程中，采用**Django框架**进行前端Web的搭建，后端全过程对用户透明，向用户呈现的是完整的Web前端项目并提供友好的用户界面，提高了项目使用的便捷性。

- 数据存储

为了满足跨终端等需求，用户可以多地登录运行，此外，搜索引擎基于聚合的数据进行，所以对用户数据和信息检索数据的存储提出较高要求，需要存储于云端数据库以保持更新并便于维护。项目中使用Aliyun服务器自行搭建Mysql进行数据存储和维护。

- 运行环境

为实现多地访问、兼容性等需求，项目部署于Aliyun服务器端，链接云端数据库以实现所有功能。用户使用时对运行环境无要求，通过浏览器访问网站 [39.100.88.210:8820](http://39.100.88.210:8820) 即可实现全部功能，无任何依赖项。