1.项目名称：百万级知乎用户数据分析与搜索系统

2.项目背景：

随着互联网的普及和发展，现在网民数量已经达到了 9.89 亿人，其中相当一部分人对 知识是十分渴求的，所以知乎这类问答平台变得火爆起来了。知乎作为一个问答软件，许多人将其当做阅读、学习、和业内人士沟通的工具。知乎的月活跃用户高达 9430 万。虽然知乎的用途广、用户多，但是缺点同样明显。例如缺少对一些信息的分类，这样使得在使用时会有一些不便利。例如对缺少对单个用户的回答进行整理分析。对知乎搜集到的整体没有直观展示呈现给用户。因此在这样的情况下，我们希望上述相关问题。我们根据用户需求，设计了“百万级知乎用户数据分析与搜索系统”这一项目，用于补足知乎相关方面的不足，其利用爬虫抓取知乎用户的数据并分析，处理，将其分析数据可视化。

3.解决问题：

可视化、下载相关资源、筛选机器人账号、去除其质量回答

4.功能：  
下载单一用户所有回答、图表展示知乎数据、查看用户数据排名（粉丝量/核心用户）、知识图谱构建

1. 对百万级数据进行爬取、清理、分析、统计及可视化，做到短周期内同步更新，确保数据的真实性以及有效性。
2. 对分析数据进行图标可视化展示
3. 查看相关数据排名
4. 对知乎中用户的详细数据（应包含其回答，视频，专栏等一切可见数据）提供下载功能，供使用者进行下载（数据会以文件包的形式提供给用户）
5. 简单知识图谱构建

5.测试功能：

搜索：知乎用户名、知乎id

下载：几个模块

查看：用户排行正常显示

接口：测试

6.技术

爬虫：python

后端：springboot+mybits+maven

数据库：mysql、elasticsearch

前端：vue+element+echarts，axios