



18825163855
zhanqh@m.scnu.edu.cn
zhanwu5
https://github.com/zhanqh
http://zhanqh.com

詹前壕

爬虫工程师

教育经历 EDUCATION

2013 ~ 2017

华南师范大学

信息管理 with 信息系统 (本科)

2015 - 2016 学年度校三等奖学金

2016 - 2017 学年度校二等奖学金

项目经验 EXPERIENCE

钱鹿科技 研发中心

2016.10 ~ 2017.01

前端开发

- 负责公司内部后台管理系统前端模块的开发
- 基于 **Vue.js** 与 **Bootstrap** 实现前端组件化及前后端分离的数据交互
- 参与维护公司 iOS 端广告投放富交互 Web 应用
- 使用 **Webpack**、**Gulp** 构建自动化前端开发工作流

个人项目

Demo

实时公交

- 广州实时公交查询 Web 应用
- 前端部分使用 **Vue.js**、**Axios**、**WeUI** 库, 通过 **AJAX** 请求向服务器获取 JSON 数据, 实时解析车辆位置
- 后端部分使用 **Flask** 框架提供 **RESTful API**
- 当服务器接收到 AJAX 请求后, 通过 **Requests** 库从行讯通爬取实时公交数据

个人项目

源代码

Demo

www.zhanqh.com

- 个人网站, 包含爬虫数据展示、笔记博客与广州实时公交查询 Web 应用
- 使用 **nginx** 作为反向代理服务器, 根据具体请求匹配对应服务
- 使用 **Echarts** 库实现可交互的图表用以展示爬虫数据分析
- 项目托管于腾讯云虚拟主机, 实现热部署、热更新

技能 SKILLS

爬虫

爬虫基础

- 熟悉 **Headers** 请求头修改伪装, **Cookies** 登录
- 能够分析请求频率限制、控制爬虫请求频率
- 熟练掌握分析、拆解前后端请求, 并构造出合法的 **AJAX** 请求
- 熟悉 **CSS**、**XPath** 选择器与正则表达式使用, 能够快速准确地解析网页提取信息

爬虫实践

- 使用 **Redis** 构建过 **IP 代理池**、**Cookies 池**
- 掌握 **Selenium** 与 **PhantomJS** 在爬虫中的基本使用, 解决 JS 动态渲染问题
- 熟悉 **Requests** / **PyQuery** 库、**Scrapy** 框架的使用, 编写过大量爬虫应用
- 理解 **Scrapy-Redis** 分布式爬虫原理, 搭建过规模化分布式爬虫, 并部署到腾讯云服务器
- 了解验证码破解、图像识别, 在实践中应用过

Web 前端/后端

前端

- 掌握 **HTML**、**CSS**、**JavaScript** 基础知识, 能够进行基本的前端应用开发
- 熟悉 **jQuery** / **Vue.js** / **Bootstrap** 等前端库的使用, 编写过单页面应用
- 了解 **Webpack** / **Rollup** 等自动化前端工作流

后端

- 熟悉 **Python**, 能实现日常的数据及图像处理工具、自动化脚本等
- 能够使用 **Django** / **Flask** 搭建简单的后端程序与数据库交互、渲染模板及提供 **RESTful API**
- 有使用 **MongoDB** / **MySQL** 数据库、**Redis** 缓存、**nginx** 反向代理经验
- 了解 **HTTP**、**TCP/IP** 等协议

其他

- 两年 **Linux** 使用经验, 熟悉常用命令与配置
- 日常使用 **Mac OS** 作为开发平台, 使用 **Git** 与 **Sublime Text** / **Vim** 进行开发
- 对新鲜事物接受快, 自学能力强