

刘西民

18813001998
wenxi303@gmail.com



教育经历

清华大学

工学硕士，微电子所，电子科学与技术

- 因直博转硕，故需最低四年毕业
- 期间决定毕业转行，开始零基础自学编程，主要是 Web 和爬虫
- 学业上的研究方向为机器学习与量子信息交叉领域

2014 年 9 月 - 2018 年 7 月

北京

山东大学（威海）

理学学士，空间科学与物理学院，应用物理学

- 以物理系综合排名第一名保送清华大学直博生
- 在威海比较偏安，主要精力是学习拿国奖，读书丰富头脑

2010 年 9 月 - 2014 年 7 月

威海

求职意向

- Python 开发：Web 或爬虫

实习经历

简媒科技（主营一键发文与上传视频）

Python 爬虫开发，后端服务

2017 年 10 月 - 2018 年 1 月

北京

- 独立维护、升级和扩展公司对各个自媒体平台的 cookie 抓取、文章阅读数抓取与匹配以及简媒平台发文状态的抓取和匹配，以此基本掌握了基础爬虫的整个流程，包括网络请求的分析与模拟，cookie 的获取与注入，网页数据和接口数据的解析和清洗，数据入库与分析，以及 MySQL 和 Redis 以及 MongoDB 的日常使用
- 配合产品经理独立完成简媒平台账号估价系统的初版设计与开发实现
- 参与模拟视频上传的研发工作，以此掌握了通过表单上传视频文件的整个流程
- 运用的图像识别技术破解网易发文平台的滑动验证码，为简媒节省不少人力
- 参与 TensorFlow 卷积神经网络破解图片验证码的研发工作
- 搭建 nsq，并以此构建平台自身的日志搜集系统

技术栈

• 编程语言

- Python：掌握语言基础，可熟练使用语言本身以及基本的常用框架与爬虫框架，了解 Flask、TensorFlow
- JS：了解语言基础，用过 vue，模仿过一个2048小游戏，开发了一个小程序
- Go：了解语言概念，使用过 chromedp，其它的正在进一步学习中
- 其它：在对比学习时，偶有了解过 Java，C++ 以及 C 当中的一些优秀设计理念

• 数据库

- MySQL、MongoDB 与 Redis 的基础使用无障碍，其中 MySQL 用得最熟

• 计算机基础理论

- 基本理解计算机软硬件系统与网络通信相关模型与协议
- 基本理解基础的常见算法与数据结构，主要是基于链表的栈、队列、集合、哈希表和常见排序以及二分查找，以及用于实现并发的进程、线程和协程



- **练手 demo**

- 使用 Python Flask 搭建一个搜索、爬取、展示豆瓣中部分影视数据的小网站和一个模仿 Google Keep 的简化版 Web Note 记事本

自我定位与评价

- 穷人家庭出身，能吃苦，爱思考，学习能力中等，有一定韧性，整体性格偏沉闷，个性不够吸引人
- 目前实际开发水平有限，基础依然不够牢靠，欣慰的是，经过实习之后，认识到自身学习能力可以应对软件开发工作的要求，思考以及习惯也比较匹配这个行业，下定决心在这个行业谋发展
- 打算以爬虫或者 Web 开发为切入点

其它

- 个人微信公众号：sixximin
- 个人 Github：https://github.com/wenxiheshang
- 个人 Github 博客：https://wenxiheshang.github.io