

杨勇

两年 Python 爬虫系统开发经验，熟悉 **Scrapy** 框架，作为主力工程师参与设计与开发过多个项目，负责系统核心模块的开发，自动化测试与部署。熟悉 **HTTP** 协议、**TCP/IP** 协议，**正则表达式**，**XPAT**H 的用法，了解 **Redis**，**MySQL** 数据库与 Linux 系统的常见机制与原理。

基本信息

contact@jobder.net

133-5555-6666

github.com/Windsooon

求职意向：爬虫工程师

政治面貌：群众

教育经历

中山大学

计算机科学 | 2010年- 2014年

操作系统（85分 / 专业排名 18/100），数据结构（90分 / 专业排名10/100）

2011-2012学年获得美国大学生数学建模竞赛一等奖

志愿工作

协助组织翻译

Flask, Requests 第一版本文档，翻译多篇技术文章。

Github 上的开源项目 Cherry 获得 200个 Star。

Django, Flask, Tornado

Vue, React

Docker

Linux, macOS

工作经历

独角科技有限公司（旗下产品 EngineGo）

2014年6月-2016年6月 | 高级 Python 工程师

- 作为组长负责设计和开发**分布式网络爬虫系统**，优化爬虫策略和防屏蔽规则，提升网页抓取的效率和质量。
- 根据行业需求分析设计方案可行性，对项目代码进行测试优化，**协助持续集成与自动化部署**，提高系统可用性。
- 负责公司技术文档的编写以及维护，定期review团队的代码，与团队共同学习以及成长。

EngineGo 爬虫系统

- 与产品经理保持沟通，使用 **Scrapy**框架 对爬虫模块进行重构，提高 **200% 爬虫速度** 并减少服务器 **20%CPU** 负载。改进爬虫策略，降低 **40%** 被屏蔽的请求数。
- 作为主要工程师设计以及开发**物业模块**，**活动模块**，**实现报名**，**即时通知**等 10 个功能。
- 使用 **Docker** 对项目进行拆分重新架构，减少业务模块之间的资源耦合，实现**持续集成与自动化部署**。

未来科技有限公司

2012年6月-2014年6月 | Python 工程师

- 使用 Python 爬虫帮助公司进行电商数据的**数据采集**，**文本分析与文本标注**。并将就分析结果与开发人员讨论产品方向。
- 带领 3 人团队完成后端系统的难点分析与架构设计工作，承担核心功能代码编写，开发与维护系统核心模块。使用 Selenium 等工具对产品进行多方面测试，保证代码的可靠，安全。

内部后台管理系统

- 基于 **Flask** 实现RESTful风格的后台管理系统，使用 **MySQL** 作为存储数据，配置主从热备，使用 **Celery** 实现异步邮件推送。
- 使用 **Redis** 存储定时任务和缓存结果，选择 **RabbitMQ** 做消息队列服务，开发，测试**登录注册**，**请假审批**等7个模块。
- 采用 **Python**，**Shell** 等脚步语言公司多个办公项目的自动化，大幅提高公司内部项目效率。