

# 艾映锋

30 岁 | 男 | 大专 | 6 年经验  
18262031725 | [aiyingfeng110@qq.com](mailto:aiyingfeng110@qq.com) | 党员

## 个人优势

- 1、熟练掌握 Python
- 2、熟练 Scrapy 爬虫框架开发
- 3、熟练抓包工具 mitmproxy, fiddler, packet capture
- 4、熟练自动化测试工具: appium、Selenium
- 5、熟练基于正则表达式、XPath、CSS 等网页信息抽取技术
- 6、熟练 Git 管理项目
- 7、熟练基于 Cookie 或 JWT 的网站登录原理
- 8、熟练掌握基本前端技术: HTML、CSS、JavaScript、jQuery、Ajax;
- 9、熟练掌握 MySQL, redis, mongo
- 10、熟练掌握 Django REST framework 开发 RESTful API
- 11、熟练掌握 Django 框架开发
- 12、熟练 ip 池代理
- 13、熟练 Selenium, PhantomJS 框架
- 14、熟练分布式、多线程/协程/进程的编程
- 15、熟悉 jenkins 镜像管理
- 16、熟悉 docker 容器部署
- 17、熟悉 Elasticsearch 搜索引擎

## 期望职位

Python 苏州 计算机软件

## 工作经历

### 苏州有推信息科技有限公司

爬虫工程师, python 工程师

2018.04—至今

#### 内容:

- 1、负责分布式爬虫框架的搭建完善上线
- 2、负责分布式爬虫框架源代码维护, 开发
- 3、负责数据源的收集, 下载, 解析代码编写, 单元测试
- 4、负责对爬虫 500 个数据源的文章, 图片, 视频下载定期维护, 抓取
- 5、负责工具类产品的快速开发
- 6、负责爬虫监控系统检查, 维护

**业绩:** 完成了视频爬虫技术突破

## 苏州市中盾投资实业有限公司

PHP/Python 程序员, 项目经理

2012.12—2017.12

### 内容:

- 1、负责 web 部分网页模块的详细设计、
- 2、负责 api 接口编写、爬虫开发
- 3、负责单元测试工作;
- 4、负责网站 seo 优化设计
- 5、负责公司项目对接, 洽谈;

**业绩:** 17 年到 18 年, 承担公司部分业务指标, 并取得了一个半月净利润 14 万小成绩。

## 中国人民解放军

军人 | 义务兵

2008.12—2010.12

**内容:** 日常训练

**业绩:** 连续 2 年优秀士兵, 年底入党

## 项目经历

### snippetmedia 分布式爬虫

python 爬虫, python 开发

2018.04—至今

#### 描述

- 1、每日抓取国内外 500 多网站数据
- 2、日常开发维护成员 3 人
- 3、负责爬虫维护, 破解反爬, 下载入库
- 4、抓取的网站有 youtube, tiktok, facebook, jd, bbc, buzzfeed, cnbc, topbuzz, twitch, vimeo 等

#### 技术架构:

下载模块: 文章链接去重、代理池、实现 cookie 提取, 部署 scrapyd

解析模块: python 代码编译成 c 语言提高性能, 清洗数据, 测试用例

评论模块: 基于 scrapy 的评论下载模块,

图片视频模块: youtube-dl 视频下载, 多线程图片视频下载

web 监控模块: fiask 监控报警, 定时任务, 每日邮件报告

通信: mns, redis, thrift

部署: docker

数据库: mongo, mysql

**项目链接:** <https://play.google.com/store/apps/details?id=ph.com.snippetnews&hl=zh>

### 抖音、快手 app 多任务爬虫

**描述:**

1. 视频类数据抓取, 清洗, 入库
2. 日常开发维护成员 2 人
3. 负责开发, 维护爬虫, 破解反爬

**技术架构:**

基于 request 多线程爬虫  
基于 appium 实现 app 模拟滑动  
基于 python 实现抖音加密字段的破解  
部署: docker  
抓包工具: mitmproxy, fiddler, packet capture  
数据库: mongo

## 京东商品数据分布式爬虫

**描述:**

1. 京东商城单品类数据抓取, 清洗, 入库
2. 单人开发爬虫系统
3. 负责开发, 维护爬虫, 破解反爬

**技术架构:**

基于 scrapy 商品数据下载爬虫  
基于 scrapy-redis 实现分布式爬虫  
基于 echarts 数据可视化  
基于 Nginx 作为服务器前端负载均衡、动静资源分离  
基于 Elasticsearch 商品搜索引擎  
数据库: mongo, mysql, redis

## 中盾现货 (金融类)、寰珠阁商城 (金融类)

**描述:**

- 1、公司官方 web 站, 产品展示, 客户信息收集, 内容发布, 客户互动交流
- 2、网站维护开发人员 3 人
- 3、负责开发, 维护 web 站点 0

**技术架构:**

基于 Django 实现 RESTFul API  
基于 Sentry 完成线上系统的错误日志的监控和告警  
基于第三方的短线登录验证码  
实现 json web token 方式完成用户认证  
实现支付宝对接, 第三方登录  
基于 Django 配置 redis 缓存  
数据库: mysql, redis

## 教育经历

江苏苏州广播电视大学

计算机网络技术 | 大专

2008—2012

## 社交主页

<https://github.com/luzhisheng>

<http://www.yujzw.com/>

## 其他证书

新浪 sae 中级开发者，证券从业资格