求 职 意 向 ： Python（爬虫方向）

businessman-clients-portfolio(1).pngemployment-certificate(1).pngbig-telephone.png

**雕 虫 小 技**

* HTML、CSS、JS
* Django
* HTTP/HTTPS、TCP/IP
* MangoDB
* **工作描述**

**分析网站、网页、链接的特征挖掘，对指定的多个网站的网页进行网页抓取、数据提取, 负责特殊网页采集任务的分析及采集方案设计**

* **项目经验**

**※ 项目名称：房价网分布式信息采集**

**项目描述：这个项目是对房价网房产信息进行数据提取。**

**关键词：scrapy, scrapy-redis, XPath, re, json, redis, csv, mysql, User-Agent, IP, request**

**技术简介：**

**1、使用Python的爬虫框架scrapy框架**

**2、使用XPath（lxml）,正则（re）进行页面分析并提取提供给数据部门**

**3、使用scrapy-redis进行分布式爬取**

**4、使用json，CSV和redis进行信息存储**

**项目模块介绍：**

**1、spider模块：处理需要的数据和请求地址**

**2、中间件模块： 设置代理和User-Agent**

**3、管道模块：数据库交互并保存数据**

**职责描述：**

**负责spider模块和中间件模块**

**2015.10-2017.7**

**北京**

**首建投资有限公司**

**Python爬虫工程师**

**武 林 大 会**

**盖 世 神 功**

* Python
* request、urllib、urllib2
* Scrapy框架、scrapy-redis
* XPath、re、BeautifulSoup4
* Selenium、PhantomJS
* Redis、MySQL

**武 功 展 示**

**2014 – 2017**

**2009 – 2012**

**中国农业大学**

**衡水职业技术学院**

**门 派 背 景**

* 年 龄：26岁
* 籍 贯：河北
* 性 别：男
* 体 重：60kg
* 经 验：2年
* 姓 名：\*\*\*
* 身 高：170cm
* 学 历：本科

wnc2010@126.com

17600321100

businessman-clients-portfolio(1).png

* **项目名称：安居客M站信息采集**

**项目描述：此项目是对安居客房价数据进行提取并提供给数据部门**

**关键词：scrapy, XPath,re, redis, MangoDB, User-Agent, IP, request**

**技术简介：**

**1、使用Python的爬虫框架scrapy框架**

**2、使用request进行代理ip信息采集**

**3、使用XPath（lxml）,正则（re）进行页面分析并提取数据**

**4、使用redis和MangoDB进行信息存储**

**职责描述：**

**负责该项目的各省市房价指数信息提取**

* **项目名称：开氪36kr网站信息提取**

**项目描述：此项目是对开氪网合作客户信息抓取**

**关键词：selenium, PhantomJS, request, XPath, json, reids, ajax**

**技术简介：**

**1、使用PhantomJS获取动态页面并触发click事件，获取下一页信息**

**2、使用request进行ajax请求**

**3、使用XPath（lxml）,正则（re）进行页面分析并提取数据**

**4、使用redis和json进行信息存储**

**职责描述：**

**负责该项目的各公司合作方数据提取**

* **项目名称：内网系统开发**

**项目描述：本系统是提供给内部员工信息化和内部物资审批系统**

**关键词：Django、ajax、MySQL、cookie、Session**

**技术简介：**

1. **使用Django进行web框架搭建.**
2. **发送ajax请求将用户添加的信息保存到MySQL数据库**
3. **采用cookie和Session实现用户状态保存。**

**职责描述：**

**负责本公司员工用户信息系统模块开发并合并到总系统。**

**个 人足 迹**

**12306余票信息，苏宁易购，斗鱼，百度贴吧，糗事百科，拉钩，boss，Tencent等**