

ID: (ud(WZix6bfkN4SICf70UA | 更新时间: 2月 15日 (2 天前)

张先生

男 25 岁 (1994年11月) 4 年工作经验 本科

现居住地: 郑州 二七区 | 户口: 河南 安阳



求职意向

期望工作地点: 杭州, 郑州
期望月薪: 6001-8000元/月
目前状况: 我目前处于离职状态, 可立即上岗
期望工作性质: 全职
期望从事职业: 软件工程师, 其他, 软件研发工程师
期望从事行业: IT服务(系统/数据/维护)

教育经历

2011.09- 2015.07 河南大学 信息工程 本科 (4年)

工作经历

2017.12 - 2018.12 深圳自然一度科技有限公司 (1年)

python后端 | 10001-15000元/月

IT服务(系统/数据/维护)

工作描述: 1.根据公司kol数据平台的需要, 分析目标站点的数据结构与反爬虫手段。
2.微博爬取维护与改进, 准保网红的资料与博文正常获取。
3.使用requests, mitmdump, socket, 异常处理等技术编写自动化爬虫。
4.对获取的数据进行清洗过滤, 存储到mongo数据库。
5.根据反爬虫手段, 不断优化爬虫方案, 提升数据爬取效率。

2016.10 - 2017.12 深圳市瑞驰信息技术有限公司 (1年 2个月)

python开发工程师 | 8001-10000元/月

IT服务(系统/数据/维护)

工作描述: 公司是一家做私有云存储的企业。我参与的项目主要是给ceph分布式存储系统开发web管理平台, 项目主要要用到框架是Python的Django和Twisted, 前端是jQuery。

2015.09 - 2016.10 上海诺亚金融服务有限公司 (1年 1个月)

python后端 | 6001-8000元/月

IT服务(系统/数据/维护)

工作描述: 1.实现web后台代码开发工作
2.对测试产生的bug进行修复
3.做好后台代码的相关优化工作
4.维护好所开发的模块代码
5.负责参与设计, 开发公司网站后台系统

项目经历

2018.08 - 2018.11 本地舆情热点抓取 (3个月)

项目描述: 责任描述: 1.根据客户要求, 分析各大论坛网站结构, 根据字段制定合理的表结构;
2.利用scrapy创建多个爬虫, 实时抓取各大论坛, 如贴吧, 微博等舆情信息, 存入数据库;
3.使用numpy+matplotlib,对数据进行可视化分析;

4.使用Nginx作为静态文件服务器，降低Django后台服务器的压力。

项目描述：根据用户需求爬取相关热点，通过多进程编程中进程池思想对数据进行爬取用来加快爬取速度，分析要爬取的页面源代码，根据爬取的信息编写正则表达式，对爬取的数据进行简单处理存储到本地文件中，提取相关信息存到数据库中，可在本站查看本地舆情热点信息简讯及标签、热度，感兴趣的话可以点击跳转链接，查看完整的热点信息。本系统帮助客户实时发现网络上涉检舆情信息，实现了全省.市.县三级联动 监测，进行了网格化舆情管理。打破了之前省院肩负全省舆情监测的重任，网格化管理后进行账户分发，根据行政区域管好自身辖区涉检舆情，形成县.市.省三级联动上报。

2018.04 - 2018.08 聆听世界 （4个月）

项目描述： 责任描述：1.用scrapy-redis框架搭建分布式爬虫
2.设置中间件模块，添加请求头Pool、代理池、cookie设置等设置反爬虫机制，提交动态验证码form表单数据提交信息实现自动提交验证码登录
3.spider.py中建立广度优先分布式爬取程序，通过ITEMPIPELINES和redis去重url。
4.设计数据库表与表之间的关联关系，在Django框架中通过ORM关系映射，完成表的创建，并把爬取的数据存储在mysql数据库；
项目描述：项目为一个听书网站，本网站的书籍来源是通过爬虫技术爬取出来，存储到Mysql数据库，主要功能都有：书架，主页面，搜索，列表页，详情页，聆听页，个人中心。

2018.02 - 2018.04 获取快手app视频 （2个月）

项目描述： 项目描述：根据公司KOL数据平台及网红小程序的需求，负责快手App短视频爬虫开发，最终将爬取的视频与数据通过API接口上传到数据平台。
责任描述：1.爬取快手APP短视频及信息，包括视频标题，粉丝数，点赞数，播放数，评论数，播放地址，作品数，快手Id，用户昵称，用户头像等；
2.使用requests,urlretrieve,socket等技术实现数据自动抓取；
3.将返回的Json数据进行结构化清洗处理；
4.通过设定视频唯一ID，将重复爬取的数据过滤；
5.将提取的数据通过pymongo存入到MongoDB数据库；
6.爬取完成后将所有视频与数据通过API接口上传到数据平台；

2017.12 - 2018.02 百度新闻爬虫 （2个月）

项目描述： 项目描述：开发环境：Ubuntu + Scrapy + Redis + MongoDB + python3
项目简介：因公司新闻网站项目需要，抓取百度新闻数据做内容抓发
项目职责：开发该爬虫并测试
应用技术：1.使用基于Scrapy框架的Scrapy-redis实现分布式爬虫
2.在中间件中实现代理池，User-Agent池，突破封锁限制
3.使用XPath和正则表达式对网页信息进行抽取
4.使用MongoDB存储新闻类数据
图片下载使用Celery异步任务，加快下载速度

2016.10 - 2017.12 云存储web管理平台 （1年 2个月）

项目描述： 负责web管理平台对象存储版块的开发，对象存储web平台，从基础功能到高级功能存储类型，多站点异地的前端设计、在这些功能的界面、交互设计阶段，有参考AWS，阿里云，腾讯云等一些公有现有的方案。经过测试小组的测试，这些特性的稳定性，bug数量都令人满意，已提供给产品部一个稳定可商用的版本。
这些特性的大部分开发步骤为：
1.先熟悉了解这些特性怎么手动配置，考虑各种极端情况，如何处理。（如用户名最多只能输长？）
2.考虑所有，用户有需要提供的值，用户可能关心哪些值的变化（如网关状态，）
综合考虑这些因素后，出一个简单原型图，也就是应该让用户输入哪些东西，应该展示哪些东西给用户，以及这些东西的排版。
3.后端代码，用python解析配置文件，可用根据前端参数生成不同的配置文件。以达到配置服务的目的。由于要同时控制好些台服务器（节点），会用到多线程。
项目描述：是给一个基于开源ceph的分布式存储系统开发一个人性化的web管理平台，以图表的方式更直观的展示集群各方面的状态，方便使用者（客户）对云存储系统进行管理、监控等运维操作，降低客户对云存储系统的运维难度。

2015.09 - 2016.09 诺亚金融网站后台开发 （1年）

项目描述： 开发环境：window 10 + Django + Redis + MySQL + Python 3

项目简介：本项目基于Django框架的一个电商平台实现了相关金融产品在线交易的功能。主要模块包括用户管理模块、商品模块、搜索模块、购物车模块、订单模块、支付模块等。

项目职责：1.负责用户注册登录模块，包括用户正常注册登录

2.负责购物车模块，包括购物车的增删改查、合并

3.负责用户下单模块，主要有订单结算和保存

应用技术：1.使用Celery进行短信验证码的异步验证

2.使用JWT进行用户登录的为用户创建身份认证的凭证

3.应用captcha生成图片验证码，Redis来实现图片验证码文字的缓存

4.发送注册邮件以Celery异步操作实现

5.使用django-cors-headers扩展处理CORS来解决跨域访问的支持

6.使用redis和cookie保存用户的购物车数据

引入了页面静态化技术，使用Celery任务队列，定时生成静态页面

自我评价

阳光开朗，幽默风趣。

遇事比较稳，越刺激越认真。

擅于管理，勇于面对挑战。

随机应变，演技十足。

有脑子的莽夫。