1. **OA/ERP管理系统**

**项目名称：公司内部OA系统**

**项目描述：**公司内部研发的办公自动化管理系统，旨在提升办公效率，提供标准化工作流程，保证运营工作正常开展。

**职责描述：**

1. 参与公司OA系统研发前期准备工作，负责登陆模块，人力资源管理模块详细规划与实施。

2. 参与系统前端页面注册、登录、及修改密码页面的开发。

3. 登陆模块：仅设置登陆及修改密码功能，账号及密码由管理员统一配置，根据用户级别配置用户权限，用户

登陆通过表单发送AJAX请求，判断用户合法性。

4. 人力模块：薪资查询-社保缴纳-请休假管理-考勤模块。主要由高级权限用户通过表单提交数据，发送AJAX**至后端，后端根据数据内容，调用相应算法，实现存储工作。调用相应数据至私人用户页面展示。**

**项目名称：ERP系统**

**项目描述：**公司内部软件,实现元器件型号、采购源和产品库存等管理的功能。

**职责描述：**用户登录；数据库设计。

**技术要点：**Django + mysql + whoosh引擎/jieba词库。

**项目名称：CRM系统**

**项目描述**：CRM系统开发维护

**职责描述：**

1）views模块，线索，产品数据展示

2）用户行为记录

3）好车帮数据展示

简述：

1）项目基于Scrapy，Xadmin 实现，为业务团队的提供数据展示，根据业务流程变化升级系统，参与线索信息，订单信息数据库设计。

2）基于Celery 记录员工对数据的访问，监控异常。使用Celery 可以减轻高峰时的负载。

3）为好车帮APP提供访问线索，订单，客户信息的数据接口，验证请求是否合法，返回Json数据。

1. **Web后台项目**

**项目名称：“前店后厂”电商网站**

**项目描述：**服饰电商平台。实现了网上交易功能。

**职责描述：**用户注册\登录；首页\列表页\详情页\商品搜索页。

**主要技术：**1.参与了用户信息、商品信息数据库的设计。

2.对用户注册、登录及密码等信息进行增、删、改、查处理；注册时对用户名等信息进行唯一性校验。

3.运用邮箱验证完成对用户密码修改的验证服务。

4.实现用户登录页面自动跳转功能。

5.使用whoosh引擎和jieba词库实现了商品关键词检索功能。

**项目名称：一路商城**

**项目描述：**电商项目，采用传统B2C模式，利用电子商务无时间性，无区域性特点，实现网上商品自由交易过程。

**职责描述：**

1. 负责用户管理模块，个人中心的后端开发。

2. 使用re验证用户注册数据正确性。

3. 使用django.core.mail，AJax，pillows，Celery完成用户验证工作。

4. 使用AJAX实现用户中心物品更新及订单匹配功能。

5. 调用百度地图JavaScript API实现收货地址定位及关键字搜索功能

**项目名称：九曳天天生鲜物流配送**

**项目描述：**西安九曳供应链是物流行业致力于打造国内第一家专注于农业、生鲜电商服务的生鲜供应链运营商，创始人九曳首创“生鲜二段式配送”，是行业内首位将电商和冷链企业融合的推动者，并操作了天猫“喵鲜生”频道多个生鲜全国配送活动；

**责任描述：**

参与流程包括需求调研、设计、模块代码编写、文档编写；

参与用户子系统的开发，包括用户登录、用户信息的实时跟踪、用户的申诉处理等的开发。

利用可变参数对Web层进行了方法抽取重构,简化代码；

利用前端开发JS优化用户登录界面；

利用MySQL和Redis数据库存储用户数据；

项目名称：乐拍商城

**项目描述：**山东广电旗下购物网站，以电视渠道作为推广，网络营销作为方式，以居家保健为主的电商网站。

**职责描述**：用户注册\登录；用户中心模块；购物车模块。**技术描述：**

1.参与了用户信息、商品信息数据库的设计。

2.对用户注册、登录及密码等信息进行增、删、改、查处理；注册时对用户名等信息进行唯一性校验。

3.运用手机验证码完成对用户密码修改的验证服务。

4.利用中间件实现用户登录页面自动跳转功能。

5.使用ajax技术时间购物车界面的动态显示与服务器存储

**项目名称：易购购商城**

**项目描述：**

易购购商城是一个网上购物平台，采用的是店主租赁店铺的模式，店主无需负责采购、囤货等，商城上的所有商品店主都可以售卖，店主有设置商品补贴的权限，销售出的商品利润平台与店主按比例分配。

**责任描述：**

1. 登录系统维护及二次开发。使用到Cookie及redis缓存技术
2. 与第三方对接。商品信息的获取及同步。
3. 供应商自助管理系统维护及新功能开发，使用到celery服务中间件。
4. 用户中心系统维护及新功能开发。图片上传功能等。
5. 订单系统维护及新功能开发。

**项目名称：在线课程教育平台**

**项目描述**：苹果产品相关软件学习以及IT编程语言视频网站

**责任描述：**

1. 前期参与需求分析、技术方案选型，项目的架构设计、开发环境搭建。
2. 用户模块：收藏，登录，注册，密码找回。
3. 课程模块：热门课程展现，相关课程推荐。
4. 后台管理模块：数据增添，修改，查看数据的访问次数。

**项目名称：我的私教-iLight**

**项目描述**：MyCoach 是 Apple 授权经销商 哈尔滨光线科技开发有限公司 向其顾客提供的，关于 Apple产品及软件的一对一店内私教服务

**责任描述**：

1. 用户中心：用户登陆注册，登陆状态验证；
2. 订单管理：订单管理、生成；
3. 支付模块：订单支付；

**技术描述**：

基于Django框架，MySQLl数据库，uWSGI + Nginx部署；

使用装饰器完成状态验证；

使用数码视讯支付子公司第三方接口和其他三方接口，完成户订单支付；

**3.爬虫**

**项目名称：房价信息爬取**

**项目描述：**爬取乐居房产，腾讯房产，安居乐网房价信息 ，土拍网信息。

**职责描述：**使用Scrapy-reids +MongoDB爬取信息,numpy+matplotlib+pandas数据分析

1. 分析目标网站真实request地址和response内容，使用代理试水网站反爬虫级别，检测是否有投毒数据。

2. 针对不同反爬虫级别网站，做出具体Scrapy部署，使用工具包括不局限于：伪造User-Agent，批量IP代理池，Selenium + PhantomJs，设置并发数量及下载速度。

3. 使用BeautifulSoup，XPath获取网页有效数据，re提取有效数据，load方法读取JSON数据。

4. 提取后的数据存入Redis，使用numpy+matplotlib+pandas进行数据统计分析，生产最终报告。

**项目名称：python爬虫爬取中国知网文献（自学）**

**项目描述：**为解决毕业论文文献查阅的实际困难，借助Python主流的框架Scrapy爬取近10年的主流期刊文献材料（期刊名称、文章名、文章作者、文章关键词、文章链接）。

**责任描述：**

参与流程包括需求调研、设计、模块代码编写；

全程参与开发爬取、数据存取本地磁盘、分析数据。

**技术描述：**

基于Scrapy架构设计项目实现路径，在“中国知网”利用Requestes实现页面爬虫请求，借助Xpath、RE解析页面，获得数据；

使用模块实现爬虫系统的封装操作，优化代码；

操作利用本地磁盘存储数据，利用Excel和PASS显著性差异分析处理数据

**项目名称：租房信息采集爬虫**

**项目描述：**这个项目主要是58同城，租房网，搜房网等网站的租房信息的爬取并提供给数据分析部门

**职责描述：**

分析网站、网页、链接的特征挖掘，采集任务的分析及采集方案设计，使用scrapy框架爬取数据，使用mango存储数据，准备一个代理ip池和user-agent列表以应对网站的反爬虫机制

Spider Engine负责Spider、ItemPipeline、Downloader、Scheduler中间的通讯，信号、数据传递，获取Spider的URL之后，发送至调度器，请求排列入队，发送处理好的request给引擎，下载器根据中间件的设置下载request请求，spider自行处理下载好的请求，处理完毕之后，有两个结果，既需要跟进的URL和获取到的item数据，item数交给Pipeline处理，URL交给调度器，重复之前的步骤，直到获取到所有需要的数据

此项目是对租房信息的抓取，抓取租房的价格、位置、图片、配置等信息，并做保存，对于爬取的数据使用jupyter notebook进行简单分析。将数据提交给数据分析部门，通过爬取的数据分析当下北京地区的租房现状，快速捕捉市场机遇，实现利益最大化。

**项目名称：新闻分类资讯分布式爬虫**

**项目描述：** 这个项目是对人民网，新浪，腾讯等网站分类新闻资讯爬取的分布式实现。

**职责描述：**

采用scrapy-redis分布式框架实现爬虫集群，分布式使用Redis实现。

存储Request请求和指纹集合，并且对各个Slave端爬虫实现集中管理和控制。

利用Redis的高并发和I/O读写来实现高速下载。

采用MongoDB做为本地数据库，将资讯新闻按所属大类、子类以及标题和内容，保存在MongoDB中。

同时每次下载前会检查请求指纹，防止重复下载，避免资源浪费

**项目名称：淘宝网（个人项目）**

**项目描述：分类店铺、商品信息(使用Selenium+PhantomJS实现动态网页数据的抓取)**

**项目名称：各大农业资讯类网站的爬取**

**职责描述：**

1. 使用scrapy框架；

2. 使用XPath（lxml）,正则（re）进行页面分析并提取数据；

3. 使用MangoDB进行信息存储。

**项目名称：资讯网站 新浪（爬虫）**

**项目描述：**定时获取最新信息，目标网站：新浪，人民网，中国财经信息网等

**职责描述：**

1）请求去重，页面爬取

2）数据清洗

简述：

1）项目基于Scrapy实现，去重模块使用Scrapy-Redis 中的 dupefilter 模块实现，分析

页面规律，如新浪，人民网的详情页网址都有日期字段，为提高效率，只存储当日

request到指纹Redis。在下载中间件中配置动态代理。

2）爬取当日最新新闻或资讯链接，下载中间件加入pre\_request，根据url使用Re去除

娱乐不相关链接，减少无效请求，process\_content模块对资讯正文使用正则表达式对

标题和全文进行关键字匹配加标签，保存数据到MongoDB 数据库，使用MongoDB

因为存储灵活。

**项目名称：客户各平台同类商品比价服务**

**项目描述：**

根据客户需求，抓取如京东、淘宝等商城的商品价格等信息。

**责任描述：**

项目为公司为部分客户抓取同类商品价格比价，采用Scrapy-Redis分布式框架实，分布式使用Redis进行实现， 利用Redis作为临时存储库，最后将数据存储在Mysql数据库中。

**项目名称**：**苹果手机天猫店的数据爬取（个人项目）**

**项目描述：**收集各个天猫店苹果手机的销售情况，型号折扣以及价格的趋势。

**责任描述**：

1. 采用 requests实现爬取。

2. 用 xpath 提取想要的节点，分别为产品的名字，价格，折扣，销量。以列表形式存储。

3. 根据分析出下一页的链接规则，运用re去获取数字，添加到url中继续爬取，下一页面直到获取不到。

4. 把爬取的内容存入本地磁盘，也可存入数据库。

**项目名称：抓取新浪微博数据（利用移动端抓取数据）**

**职责描述:**

分析url，找到正确的首页start\_url，

引擎从调度器中抽取start\_url用于抓取数据；

引擎把url封装成request传给下载器；

下载器把资源下载下来并封装成应答包Response；

解析Response，把item交给实体管道Pipeline进行下一步处理；

若解析了url则交给调度器等待抓取。

**项目名称：斗鱼直播房间信息**

**职责描述:**

1. 找到start\_url；导入selenium的webdriver包；

2. 发送get请求，获得响应；

3. 利用find\_elements\_by\_xpath来获取数据，并获取下一页这个元素；

4. 提取数据及下一页的元素；

5. 保存数据，点几下一页进入循环。