

## 个人信息

---

- 甘家城 / 男 / 1996
  - 手机/微信: 15968182251
  - 邮箱: 935669873@qq.com
  - 本科 / 杭州电子科技大学  
电子科学与技术专业 / 19届毕业生
  - 技术博客: <http://ganjiacheng.cn>
  - Github: <https://github.com/webgjc>
  - 期望职位: 后端开发 / 数据开发
  - 期望薪资: 面谈
  - 期望城市: 杭州
- 

## 技术之谈

---

- 后端目前主要使用比较多的为Java和Python, Web框架的分别为SpringBoot和Django, 这个是主要对象;
  - 对前端Html/Css/Js有所涉猎, 工作期间学了React和公司内部用的前端框架, 并在实际公司项目中配合antd动手写了一部分, 自己日常也常用原生Js和Jq写一些插件和实现(最近更新一个github上的项目 [web\\_robot](#));
  - 关于大数据一些组件有在工作中用到, 包括hadoop, hive, ranger, 元数据与指标等, 之前做大数据相关的平台和数据可视化的后端开发。
  - github和技术博客一直有更新, 最近做的比较多的是大数据相关的事情, 和一些插件;
  - 之前做过段时间运维, 日常的运维也能自己完成, 打包部署装机, shell脚本等, 对linux也玩的还行;
  - 版本管理git用的比较多, svn很早前也用过一段时间;
  - 对于算法一直在跟进, 偶尔也算在leetcode有做题吧;
- 

## 工作经历

---

### • 大搜车软件技术有限公司 (数据研发工程师) (2019-04 ~ 至今)

#### ◦ 元数据管理与数据地图

该项目主要是对数据仓库元数据进行采集、管理, 同时也支持业务模型元数据, 指标元数据等, 本身提供一个数据查询平台数据地图, 并也为其他应用提供元数据支持。采集的元数据主要包括表字段信息, 血缘, 权限, 业务关联信息等, 从metastore、hivehook、离线调度平台、ranger采集到整合到后端数据库存储, 进行了统计与分类, 并在数据地图中的展示搜索。项目使用python/django/celery实现, 存储用的mysql和redis, 本人在项目中为主要后端开发, 项目为从零搭建。项目中的元数据采集使用celery的定时任务, 采用新表和全量两种同步策略, 也提供手动同步的方式。具体可以参考我的原创博客[讲讲元数据与元数据采集](#)

- 对元数据采集进行了多层次(增量全量)多方式(手动定时)多来源 (metastore/information\_schema)的优化;
- 元数据进行了多个来源的整合, 并标记清理, 关联, 统计等业务操作, 同时监控了每日的变更;
- 实现了建议方便的元数据查询平台, 包括大部分元数据的展示;
- 提供多种稳定的元数据外部对接接口;

#### 。敏捷分析与可视化报表

该项目是包括一个sql即时查询工具, 包括直连数据库和模型元数据的拖拉敏捷分析, 包括一个报表系统。项目分两个大模块, 查询模块和图表后端模块; 查询模块专注于sql的生成、sql的解析优化, sql的执行, 多种数据源的对接等, 图表模块专注于图表配置, 前端拖拽交互接口, 外部对接, 权限控制等, 两个服务使用dubbo进行交互。本人为项目主要后端开发之一, 主要参与图表模块, 进行权限设计, 查询交互; 也部分参与查询模块, 对hiveSql的解析进行了部分研究和优化(博客中主要分享了antlr实现解析的一些实现)。项目使用java/springboot/mybatis搭建, 存储使用mysql, 查询支持hive, presto, mysql等。

- 查询使用ranger控制到表级别的查询权限, 中间设计进行了权限缓存, 行级权限由后端自行实现;
- 元数据由数据地图提供, 这边由后端对接接口;
- 查询分同步异步查询, 统一走jdbc协议;
- 使用了sql解析实现了sql的格式化, sql血缘和sql元数据解析;
- 使用图表缓存的方案, 由离线任务关系控制缓存的刷新;

#### 。数据标准与数据建模

该项目一方面是对数据标准、数据码值进行定义、录入、管理和对外提供。另一块是数仓建模工具, 通过表单, hql, execl等方式导入建表信息, 提供ER图画图模块对导入表进行建模作图。建模的时候可以直接应用到数据标准和码值信息。项目使用java/springboot/mybatis搭建, 存储使用mysql, 本人为项目的后端开发, 项目为从零搭建。

- 项目中在码表中因为码表是自建的, 数值字段是自定义的, 因此使用了json存储, 且用到分表存储;
- 在hql导入中用到了hql的create语句的解析, 使用的是antlr4和hive的hplsql.g4然后用java实现visit模式的create\_table;
- 渲染导出hql使用mustache模板引擎;
- excel导入使用easyexcel实现;
- 同时参与实现了一部分前端的功能, 使用react和antd;

### • 默安科技有限公司(实习) (运维开发) (2018-07 ~ 2019-03)

#### 。产品打包部署更新等自动化脚本和平台

为各个产品编写打包的脚本, 部署脚本, 更新脚本等。通过搭建的平台使用可以直接将产品打包部署到机器上, 可以更新该环境的代码, 和查看环境中产品运行的状态。平台项目使用python/django构建, 日常的各个脚本使用shell编写完成, 同时在版本迭代有变更或依赖包有变动时实时更新。

#### 。其他一些服务如vpn, 文档服务器, 官网等的维护

搭建维护vpn平台使用openvpn实现，对接钉钉群完成在群内获取验证码登录，并做登录记录和统计。

文档服务器和官网做日常维护和更新。

- 艾锴信息技术有限公司(实习) (python开发) (2017-12 ~ 2018-04)

- 量化策略与相关平台工具

- 主要实现一些金融研报的策略，在开放平台完成回测。

- 搭建金融相关使用工具的平台，主要使用python/flask。

---

## 最后一席话

---

- 毕业一年多了，算上实习的也有个两年多了，技术栈也是在深挖阶段，希望可以在web开发，大数据等这些领域有更多的机会和挑战。最后感谢阅读，欢迎交流！