

技能清单

- 熟悉Java、Python。
- 熟练使用SQL，熟悉常用窗口分析函数，有SQL的优化经验。
- 熟悉Hive、Spark，了解基本运行原理。
- 熟悉Airflow、DataX的使用。
- 熟悉数据建模理论，熟悉数据分层和维度建模，能够根据业务场景设计数据模型。
- 了解元数据管理、数据质量管理，有一定的实践经验。
- 熟练使用Linux、Git、Maven。
- 了解Spring Boot、Mybatis的使用，熟悉RESTful API设计。

工作经历

2018.04-2020.01 上海致维网络科技有限公司

- 负责数据模型的设计和数据库的ETL
- 负责数据任务的日常维护
- 负责数据服务接口开发

项目经验

商户数据采集系统

项目描述：该项目以HTTP接口的形式提供数据抽取服务，同时有一定的任务管理功能。

技术栈：SpringBoot、MyBatis、DataX

个人工作：

- 采用将DataX包装为HTTP服务的方式，可以让外部程序拥有非常灵活的数据抽取能力。
- 数据抽取任务不再由Airflow管理，改为由采集系统内部管理，大大减小了Airflow的压力。
- 记录任务执行状态、时长、报错信息等信息并进行分析，便于及时发现问题和对任务进行优化。

超级导购重构

项目描述：该项目作为数仓重构的一部分，主要解决导购数据错乱和不一致的问题。

个人工作：

- 重新梳理数据指标和口径。
- 重新设计数据模型，划分数据层级。
- 重构后数据结构更加清晰，方便使用，同时减少了30%计算任务。

商户自定义指标计算

项目描述：该项目提供会员升级、会员PLUS以及品类、品牌、SKU自由组合等指标，在基础经营指标外，满足商户的一些自定义指标需求。

个人工作：

- ADS层表模型设计以及Spark计算任务开发。
- 数据服务接口开发以及接口文档的编写。
- 使用EasyExcel和阿里云OSS编写了通用的Excel导出接口。
- 计算采用策略模式而非硬编码，让算法和计算逻辑解耦，解决了计算规则不同的问题，同时后续指标的新增也能非常方便的扩展。

自我评价

- 喜欢思考，喜欢总结，能从多个角度看待事物。
- 热爱学习且学习能力强，工作中遇到问题大部分可以通过自己搜索查阅资料快速解决，业余时间也会主动学习。
- 对数据领域有着浓厚的兴趣，相信数据能产生价值。