

个人信息

基本信息	张伦/男/1996	南昌大学(18年毕业)	电子信息 (本科211)
联系方式	13647956188	zhanglun.me@gmail.com	2529450174(QQ)
博客 Github	blog.zhanglun.me	github.com/mrzhangboss	

工作经历

浙江普树信息科技

数据工程师(18年7月-至今)

风控大数据系统(19年1月-至今)

公司风控大数据系统建立在阿里云数加平台，系统产生数据的周期为天，由于业务增加以及业务转型需要搭建实时处理系统并且进行平滑升级

- 分析原项目的业务逻辑制定初步系统搭建方案，并完成Demo交由Leader审核评估
- 负责编写Sqoop脚本定时更新原来风控系统的第三方数据，使用Oozie进行管理
- 负责子系统运营商风控数据处理，使用Spark Streaming进行重写，处理速度提高了100倍

数据仓库转存平台(18年11月-19年1月)

公司数据仓库使用MySQL进行搭建，由于要陆续接入第三方数据源，以及兼容内部数据源，所以需要搭建一个动态数据存储与获取平台

- 开发了一套数据存贮框架[json2db](#)，能够将任意Json字符串存贮到数据库并且能够检索并导出原始数据
- 负责使用自研架构基于中间件模型搭建了一套数据接入API平台，接入上百个外部数据源，并提供数据类型转换，数据检索导出服务，平台数据日导入导出为百万级

文书检索项目(18年7月-11月)

文书数据存贮在MySQL数据库中，由于数量在两千万左右（每天都有更新），普通的SQL检索性能太慢，所以使用Spark批量处理数据并导入的ElasticSearch中进行检索

- 采集行政规划数据并改造[Aho-Corasick](#)算法，实现文书地理行政规划判断，解析速

率单核最终为2ms每篇

- 负责编写Scala程序实现对每天新产生的文书解析并导入ES中，优化速率最终速率单节点100篇每秒
 - 设计动态拼接DSL，优化查询客户数据匹配，准确度提升2倍左右
-

技能清单

- 擅长Java Scala Python
- 熟悉Hadoop生态区，熟悉Hive HBase Spark等组件
- 熟悉MySQL MongoDB Redis等主流数据库
- 掌握主流机器学习模型算法
- 三年Linux使用经验，掌握Linux平台常用软件使用
- 熟悉英语，能够流畅阅读英文文档

开源项目

- [Json2DB](#)

可以存贮任意Json字符串到主流数据库（支持MySQL，PostgreSQL，MSSQL，SQLite等），并支持在数据库中创建删除数据表，提供定制生成数据表结构的接口，编写了一套前端页面支持动态修改数据库字段设计以及复杂映射关系，[体验地址](#)

- [行政规划预测](#)

可以根据任意地名推测出最可能的完整行政规划名称，[项目地址](#)