

佐国清

求职意向：大数据开发工程师

年龄：27岁

工作年限：5年

手机：155 1277 1508

邮箱：[15512771508@163.com](mailto:15512771508@163.com)

籍贯：江西萍乡

## 教育背景

2015.09 - 2019.06

河北工程大学

本科

## 工作经历

2020年10月-2024年10月

中兴智慧(北京)技术有限公司

2019年07月-2020年09月

长江空间信息技术工程有限公司（武汉）

## 项目经历

项目：廊坊银行数据集市项目

2023/09-2024/11

项目背景：该系统主要集数据采集、数据加工、数据展现于一身，实现业务统一管理，为业务部门（财务部，风险部，投资部等）分析决策提供数据指标方面的有力支持。

### 贷款资金监管报送需求

#### 工作描述：

- 需求分析：参与需求评审，与行方确定各层业务逻辑，取数口径，确定数据表的加载类型，并且与业务部门确认细节。
- 设计阶段：研读mapping映射文档，将数仓分为四层，ODS层（贴源层）-FDM层（基础层）-ADM层（汇总层）-MDM层（集市层），编写设计说明，参与行方对设计文档的审核，通过行方提供的模型及样表完成集市层表结构的设计。
- 数据开发：根据业务需求文档，设计文件，确定好数据接口，将各业务系统数据导入贴源层，将贴源层数据经过设计存储过程，清洗、转换、整合到主题层、汇总层；研读需求文档及设计文档，与业务确认指标口径，取数逻辑，从主题层、汇总层进行集市层指标运算构建集市报表层（业务部门主题等）；将集市报表每日/每月与总账对比，检查是否缺失数据、数据发散、是否有脏数据；代码性能调优，查看是否执行缓慢，资源配置是否合理。确认集市层报表数据是否为下游所需数据，供前端报表系统访问。
- 配合下游报表组，将集市层数据通过可视化的工具展示出来，供业务部门分析决策。
- 测试验收：测试源端和目标端的数据量对比是否正常；源端关联逻辑验证，数据量是否会发散，发散后的数据是否是正常关注的数据；增量是否合理，全量是否合理；源端是否有主键，目标端是否有主键；源端逻辑处理过程中，字段的映射、指标计算和函数处理等是否正确；目标端表名、数据库名称、前置执行SQL和后置执行SQL，存储过程是否正常运行，产出测试报告。
- 变更上线：上线评审，评审内容是评审具体代码，并在演练环境验证通过，进行变更上线。

### 对数字银行部贷款担保公司余额分析需求

#### 工作描述：

参与需求评审：需求评审阶段，业务人员传达需求后进入初步的可行性分析，并确认该业务方案是否可以落地实施。

分析指标:贷款余额, 贷款笔数, 客户数等等

分析维度:记账日期, 星期, 担保公司名称等等

需求和设计分析:与业务、总工办讨论、确认需求。

分析报表是否能实现: 根据需求编写需规文档并与业务确认, 与源系统确认取数规则, 是否涉及新增接口, 评估数据量。

开发:如果涉及接口变更, 按需求进行申请、变更, 下发接口操作, 设计、创建表结构, 开发ETL脚本;开发表样。整理文档, 脚本、报表测试。

## 数据集市项目验证结果数据并补充历史数据

工作描述:项目上线后按照业务及下游反馈做出数的调整及补充, 根据下游反馈数据质量, 调整取数逻辑;按照业务需求补充近两年到三年的历史数据

## 零售 CRM 系统 2023 年优化需求及 2023CRM 营销平台需求

项目描述:按照业务需求给 CRM 系统提供相关数据支持

工作内容:开发内容包括客户负债, 产品, 交易余额, 客户购买资产金额, 客户的签约信息, 客户及客户经理创利明细, 客户购买贵金属, 收单客户信息客户代发信息等等

---

## 项目: 廊坊银行人行大集中报送

2022/09-2023/08

项目背景:人行大集中系统是中国人民银行的数据集中系统, 主要用于金融数据的集中管理和分析, 支持货币政策制定和金融稳定评估。该系统通过集中处理大量金融数据, 提供全面的金融统计和分析服务, 支持宏观调控和金融监管。

工作描述:

- 1、需求调研:深入研究人行大集中报表规范, 大集中系统的业务目标和功能要求, 样表, 开发设计文档, 与上游数仓沟通确定数据来源、数据接口文件。
- 2、开发阶段:根据设计文档, 编写 sql 代码, 存储过程, 运用正则表达式将数据进行清洗, 转换, 整合。开发金融机构资产负债项目月报表、金融机构信贷收支月报表、金融机构存款类月报表金融机构贷款类月报表、金融机构中间业务月报表、金融机构其他业务月报表, 存储在 Oracle 数据库中, 将计算得到的报表传入行方的报送平台。
- 3、测试阶段:配合测试组进行功能测试、数据质量测试、ETL 加工逻辑测试、业务核对测试、报表/文件数据准确性测试、查询测试。
- 4、运维阶段:输出存储过程文档, 脚本清单, 定时查看日志文件, 查看执行流程, 性能调优。

---

## 项目: 河北银行信贷风险管理项目

2021/10-2022/08

项目背景:提取客户贷款业务相关的数据, 统计包括客户资产负债率、现金流量情况、资产质押情况、客户还款率、逾期率等指标。对风控指标进行预警提示, 依次来得出失信用户的整体特征, 使企业在以后的业务中尽可能规避此类问题, 以及分析各维度的不良贷情况, 以及放贷比例, 给企业提出一些有效的参考意见, 完善业务形态

工作描述:

1. 参与产品经理举办的项目需求会议讨论及评审, 分析和理解业务需求, 分析源数据以及取数逻辑;
2. 参与项目经理举办的开发设计会议讨论及评审, 分析和规划开发设计方案;
3. 根据项目需求各主题要求, 对 DM 层进行开发, 通过编写存储过程完成数据的同步和指标的计算, 检查并优化 SQL;
4. 对上线脚本进行检查, 并进行上线支持。

## 项目：河北银行金融租赁大数据风险管理项目 2020/10 - 2021/09

项目背景：融资租赁是一种金融工具，允许企业通过租赁而非直接购买的方式获取资产使用权，并在一定期限后可以选择购买这些资产，运用大数据技术可以显著提高融资租赁业务的效率和风险管理能力。

### 工作描述：

- 1、需求调研：与甲方驻场人员沟通，明确项目需求和目标表结构。
- 2、数据开发：根据数据增长特点，将目标表分为增量表和全量表。增量表通过部署数据通道 Maxwell-Kafka (zookeeper) -flume 自动传到 hdfs 对应目录，全量表通过 datax 大数据工具导入到对应的 hdfs 目录，在 hdfs 上建立数仓，分为原始层 (ods)，公共维度层 (dim)，明细数据层 (dwd)，应用层 (ads) 四层，编写 sql 代码语句将 o 层转化为 dim 层，dwd 层，再根据业务需求，从 dim 维度层，dwd 明细层进行指标运算转换到应用层 (adm 层)，将以上 sql 代码封装为 shell 脚本，上传到 Dolphinscheduler 资源中心，通过设置定时工作流，每日夜间自动同步数据导出到 Oracle。
- 3、性能优化：查看执行计划，对 SQL 语句进行性能优化调整，备份代码和完善代码注释增加代码可阅读性。
- 4、上线测试：项目上线配合测试组对项目上线集成测试、定时任务调度维护，定时开发或更新存储过程，查看日志文件。

## 长江空间信息技术工程有限公司（武汉）

## 大数据工程师

项目：全国 2010 年-2020 年土地利用变化大数据开发项目 2019 年 07 月-2020 年 09 月

### 工作描述：

- 1、遥感技术：利用高分辨率卫星遥感影像进行土地利用现状的判读和分类。
- 2、使用 GIS 技术进行数据的采集、处理、分析和管理。
- 3、结合遥感数据，进行实地调查和核实，确保数据的准确性。
- 4、建立统一的土地调查数据库，实现数据的集中管理和共享。
- 5、对调查数据进行汇总和审核，形成最终的调查成果。

### 个人技能

1. 熟练掌握 SQL 语句、Hivesql 语句，常用函数及存储过程，熟悉 MySQL、Oracle 等关系型数据库
2. 熟练掌握 Hadoop, Hive, Kafka, Sqoop, Spark, Flume, Maxwell, Datax, Kettle, Datagrip, 海豚 dolphinscheduler 工具。
3. 熟悉数仓层建模流程
4. 熟悉 Hivesql，SQL 优化
5. 熟悉 Linux 操作系统的常用命令和 Shell 脚本开发
6. 英语四级，C1 驾照，普通话二级

### 个人评价

- 性格稳重，抗压能力强
- 良好的沟通能力，能够清晰地表达自己的想法和意见。
- 能够快速学习和掌握新的技术和工具
- 我始终保持对新知识、新技术的学习热情与敏感度，信奉终身学习的理念。