# 林雨辰 | 高级数据工程师 / 实时数仓与数据平台建设

中国，上海，上海，浦东新区世纪大道 100 号，200120
+86 138 0013 8000 • yuchen.lin@example.com • https://www.linyuchen-data.dev
@yuchen-lin-data • @yuchen-lin-data • @yuchen-lin

## 简介

具备 7 年数据工程与数据平台建设经验，专注于实时数仓、数据湖治理和批流一体架构。

- 熟练使用 Spark、Flink、Kafka、Hive、ClickHouse 与 Airflow 构建稳定高效的数据链路
- 擅长从业务目标出发设计数据模型、指标体系与数据质量监控方案
- 具备云原生数据平台落地经验，熟悉 AWS、阿里云与 Kubernetes 环境
- 重视工程规范、成本优化与跨团队协作，能够推动数据产品持续迭代

## 教育背景

**华东理工大学** **https://www.ecust.edu.cn**

*硕士，计算机科学与技术，成绩：3.8/4.0* 2015 年 9 月–2018 年 6 月

- 研究方向为大规模数据管理与实时计算，完成基于流式处理的数据分析课题
- 系统学习数据库、分布式系统与数据挖掘课程，为数据工程实践打下扎实基础
- 参与实验室数据平台项目，负责数据采集、清洗与指标计算模块

**课程**：数据库系统原理、分布式计算、数据挖掘与知识发现、大数据存储与管理、机器学习基础、云计算与虚拟化技术

## 工作经历

**云启数据科技有限公司** **https://www.yunqi-data.cn**

*高级数据工程师* 2022 年 3 月至今

- 负责企业级实时数仓与数据中台建设，支撑营销、风控与运营分析等核心场景
- 设计基于 Flink 与 Kafka 的实时数据链路，将核心报表时效从小时级提升至分钟级
- 主导数据模型分层与指标口径治理，沉淀公共维度与主题数据集，降低重复开发成本
- 建设数据质量监控与告警体系，覆盖完整性、及时性、唯一性与一致性等关键规则
- 推动数据任务成本优化，通过资源队列、分区策略与任务合并降低计算与存储开销

**关键字**：实时数仓、Flink、Kafka、数据治理、成本优化

**星河金融信息服务有限公司** **https://www.xinghe-fin.com**

*数据工程师* 2019 年 7 月–2022 年 2 月

- 参与金融业务数据仓库建设，负责客户、交易与产品主题域的数据开发
- 使用 Spark SQL 与 Hive 完成离线批处理任务，保障日终报表与监管报送稳定产出
- 搭建 Airflow 调度依赖与任务巡检流程，缩短故障定位与恢复时间
- 与分析师和产品团队协作，梳理指标定义并输出数据字典与开发规范

**关键字**：Spark、Hive、Airflow、数据仓库、监管报送

**蓝鲸网络科技有限公司** **https://www.lanjing-tech.com**

*数据分析工程师* 2018 年 7 月–2019 年 6 月

- 负责电商用户行为数据的采集、清洗与建模，支持增长与运营分析
- 开发用户分群、漏斗转化与留存分析数据集，提升分析响应效率
- 编写 SQL 与 Python 脚本完成日常数据提取和自动化报表生成
- 参与埋点规范设计，推动数据源质量与口径一致性提升

**关键字**：SQL、Python、用户行为、埋点治理、报表自动化

## 语言

**中文**：母语或双语水平 **关键字**：普通话

**英语**：完全专业水平　　**关键字**：大学英语六级、技术文档写作、跨团队沟通

## 专业技能

**数据仓库与建模**：专家　　**关键字**：维度建模、数据分层、指标体系、数据字典

**大数据处理**：高级　　**关键字**：Spark、Flink、Kafka、Hive、ClickHouse

**编程与开发**：高级　　**关键字**：Python、SQL、Scala、Shell、Java

**调度与云平台**：中级　　**关键字**：Airflow、Kubernetes、Docker、AWS、阿里云

**数据治理与质量**：高级　　**关键字**：元数据管理、数据质量、权限管理、成本优化

## 荣誉

**年度技术创新奖**

云启数据科技有限公司　　2024 年 1 月

表彰在实时数仓性能优化与数据质量体系建设中的突出贡献，核心报表时效提升 90% 以上。

**优秀项目奖**

星河金融信息服务有限公司　　2020 年 12 月

因监管报送数据链路稳定性提升与交付效率优化获得团队表彰。

## 证书

**AWS 数据分析专项认证**　　**https://aws.amazon.com/certification/**

亚马逊云科技　　2023 年 8 月

**阿里云大数据分析师认证**　　**https://www.aliyun.com/certification**

阿里云　　2021 年 5 月

## 出版刊物

**中国数据工程实践**　　**https://www.example.com/articles/realtime-data-quality**

面向实时数仓的数据质量监控实践　　2023 年 11 月

- 总结实时数仓中数据质量规则设计、告警分级与故障恢复的工程经验
- 介绍批流一体场景下的数据校验方法与成本控制策略

## 引荐

**周立诚**　　**+86 139 0013 9000**

前技术总监　　licheng.zhou@example.com

林雨辰在实时数仓与数据治理项目中展现出扎实的技术判断力和跨团队推动能力，能够将复杂业务问题转化为可落地的数据方案。

## 项目

**实时用户行为分析平台**　　**https://www.yunqi-data.cn/projects/realtime-analytics**

面向营销与运营场景的实时用户行为分析平台　　2022 年 3 月–2023 年 12 月

- 基于 Flink、Kafka 与 ClickHouse 构建从埋点采集到实时看板的数据链路
- 支持分钟级用户分群、漏斗转化与活动效果分析，服务多个业务团队
- 建立任务监控、数据质量校验与资源成本看板，提升平台可观测性

**关键字**：实时计算、用户行为、ClickHouse、数据质量

**电商数仓分层重构**　　**https://www.example.com/projects/ecommerce-warehouse**

对原有离线数仓进行分层重构与指标口径统一　　2020 年 1 月–2021 年 10 月

- 梳理客户、商品、订单与营销主题域，设计 ODS、DWD、DWS 与 ADS 分层模型
- 统一核心指标口径并建设公共维度，减少重复开发与数据口径争议
- 通过分区治理与任务合并降低批处理资源消耗，提升数据产出稳定性

**关键字**：数据仓库、维度建模、Spark、指标治理

## 兴趣爱好

**开源社区**：数据工具文档、社区分享

**运动**：长跑、羽毛球

**阅读**：技术书籍、历史

## 志愿服务

**开源数据工程社区** **https://www.example.org/open-data-community**

文档贡献者 2021 年 5 月至今

- 参与开源数据工具中文文档翻译与示例整理
- 协助组织线上技术分享，推动数据工程实践经验交流