

胡昊

+86 133 8752 1087 · felixhuhao@gmail.com · haohu.tech

个人概述

资深后端工程师，现任 AI 方向 CTO —— 10 余年构建大规模、高可用分布式系统与数据基础设施，曾就职于 Google、Uber 与 Microsoft。

构建面向生产的 Agentic AI 系统——RAG、对话式数据分析、MCP 驱动的多智能体编排，以 harness engineering (黄金集评测、确定性护栏、上下文管理、人在回路安全) 保障系统质量。

实战型工程领导者——带领小团队贯穿从架构到交付，采用规范驱动、审查把关的 AI 辅助 workflow，配合质量门禁与自动化验证。

技术能力

大模型与智能体: 智能体编排、智能体治理、工具 / 函数调用、MCP 集成、多智能体路由、上下文工程、LangChain / LangGraph / DeepAgents、OpenAI / Claude / DeepSeek

RAG 与检索: RAG (多模态)、混合检索、嵌入、重排序、查询优化、Milvus、Qdrant

评测与安全: LangSmith、RAGAS、OpenEvals、LLM 裁判评测、护栏、人在回路、可观测性

编程语言: Java & Go (Google readability)、Python、SQL、C#、C++

后端与分布式系统: 微服务、API 设计、gRPC / Protobuf、REST、Thrift、Spring AI、FastAPI、SSE

数据系统: 数据建模、PostgreSQL、MySQL、Redis、BigQuery、Kafka、Flink、ClickHouse、MongoDB

云与基础设施: Google Cloud、AWS、Docker、Kubernetes、Linux、CI/CD

工作经历

元宝创意

2025.02 – 至今

首席技术官

中国武汉

- 架构并交付 ERP SaaS 平台的统一 AI 层，覆盖运营、数据分析与知识，带领 4 人团队；建立共享的评测与安全实践：可观测性、黄金集与回归评测套件。
- 运营 Copilot：基于 DeepAgents 构建智能体运行框架，搭载委托式子智能体（销售 / 库存 / 订单 / 采购），通过角色权限约束的 MCP 工具、沙箱化执行、主动监控、可审计工具追踪实现安全运营，并以 Spring AI / Java 21 服务提供人在回路审批（加密绑定、一次性使用）以保障高风险写操作。
- 数据分析智能体：交付面向非技术用户的多轮自然语言转 SQL 分析系统，基于 ERP ClickHouse 数仓，由 Qdrant 语义层支撑，将提示上下文减少 ~73%；增加了 OLAP 意图路由、确定性 SQLGlue 护栏、只读执行、有界 SQL 修复、结果等价性回归评测及自动图表生成。
- 知识助手：交付基于 LangGraph 编排的企业文档 RAG，支持多源数据接入与 Milvus 混合检索；根据查询意图选择精准 / 均衡 / 广域检索策略，结合查询扩展、HyDE 与 RRF 重排序，强制引用溯源、无证据不答及基于角色的访问控制；引入 LLM 裁判与关键切片综合评分，配合查询级可观测性。

自由职业

2023.03 – 2025.01

独立软件工程师与顾问

加拿大多伦多 / 中国武汉

- 构建并部署云原生旅行指南 SaaS，支持 QR 扫码 POI 音频导览、匿名评分与播放分析，以 Next.js + Go 驱动，Postgres 存储 POI 与评分，音频走 S3/CDN 分发，Docker 容器化部署。
- 为客户交付 RAG / 多智能体原型与安全护栏方案，覆盖多智能体编排、对话式检索等场景。

Uber

2022.05 – 2023.03

高级软件工程师

加拿大多伦多

- 构建 Uber 首个 CPG (快消品) 广告功能——店铺横幅广告：作为 3 人团队的后端负责人，构建基于 Go/gRPC + DocStore 的服务，连接 UberEats Storefront 与 Ads 平台，使广告主可在各大零售连锁中按预算投放商品/品牌广告；首季度即达到 10–30K QPS, p99 ~140 ms。
- 重新设计 CPG 广告计划的数据模型：将餐厅时代的单点商品 ID 替换为 GTIN 索引的多门店模型——使全国性广告计划成为可能，消除数百万条重复记录，形成 CPG 广告数据基础。
- 推出店铺菜单赞助商品：将第三方广告供应商接入 CPG 广告主干，赞助商品在所在子菜单中优先展示，通过 Kafka + Flink 实时推送曝光/点击事件用于计费，即便接入外部供应商，p95 仍维持在 ~180 ms，配合 Prometheus / Grafana 监控与依赖告警。
- 主导 CPG 广告分阶段生产上线，并担任广告团队开发代表——协调产品、数据科学、销售、DevOps、合作团队及外部广告供应商。

Google

2018.02 – 2021.02

软件工程师——Cloud SQL (数据集成 / 灾备)

美国森尼韦尔

- **将大规模实例集群的备份健康度从 99.92% 提升至 99.97%** (支撑 99.95% SLA): 独立负责并加固备份验证流水线 (快照挂载、数据库启动、校验和比对, 在共享 GCE 实例池上运行), 解决包括资源耗尽、启动失败、误报验证失败在内的主要故障类别。
- **实现 PITR 安全的事务日志保留**: 构建备份 PD 快照清理机制, 精简非必要的通用日志数据, 保留恢复关键日志, 实现通过最近备份结合日志回放的任意时间点恢复。
- **重新设计备份保留策略**: 将按固定数量淘汰 (优先删除最旧备份、不区分健康状态) 替换为始终保留经流水线验证为健康的最新备份, 确保每个实例都拥有可恢复的还原点。
- **主导 MySQL 高可用从流式主备复制向 Persistent Disk 同步复制的迁移**: 构建贯穿控制面与数据面的 Java 工作流, 自动转换旧版流式主备配对。
- **为 MySQL 与 PostgreSQL 的 CSV 导入/导出增加可定制格式**: 扩展后端 REST API, 用 Go 重构单次导入/导出操作, 打通 CSV 数据流转闭环。
- **将 Cloud SQL PostgreSQL CSV 导入/导出功能交付至 GA**: 加固导入/导出路径, 防范注入攻击, 采用按需最小权限数据库访问, 并通过 Cloud SQL IAM 管控对 Google Cloud Storage 的访问。
- **作为数据面 on-call 工程师维护 99.95% SLA** ——处理 Java 后端与 Go 实例代理上的发版阻塞及客户投诉事故; 指导初级工程师。

Microsoft

2014.01 – 2017.10

软件开发工程师

美国山景城

- **从零开始联合构建 People I Communicate With (PICW)**, 3 人团队: 以 C# 开发 Exchange 基础信号采集组件, 采集常用联系人邮件交互信号, 服务于 Contact Safety 与 Exchange 反垃圾邮件系统, 支持安全等级分类、垃圾规则处理与垃圾 / 钓鱼邮件检测等功能。
- **独立负责 PICW**, 扩展其信号覆盖面并持续丰富信号类型, 主导跨团队集成, 对接 Exchange 反垃圾邮件、Contact Safety 与 People Relevance 下游系统; 其信号成为 People Relevance 的核心输入, 驱动 Graph People API 的推荐、排序与轻量级 CRM 等功能。
- **为 People Relevance 构建监控基础设施**: 日志看板、实时告警与性能计数器, 为团队提供 Graph People API 服务的运维可观测性, 提升问题定位与 on-call 响应效率。
- **主导 Hotmail / Outlook.com 向 Exchange 迁移的账户筛选工具**: 基于 C# 构建分布式功能画像服务, 通过 SQL 大规模扫描 22 亿账户基数的使用特征, 发现 40+ 项 Exchange 功能缺口, 识别 4 亿 + 可迁移账户, 并在缺口各轮修复后重新画像以支持逐批迁移。
- **参与 Exchange on-call 轮值**, 覆盖 Hotmail / Outlook 服务, 处理发版阻塞与客户投诉事故, 快速排查并高质量修复, 保障 4 亿 + 用户邮箱服务的在线稳定性。

Epic Systems

2012.06 – 2013.11

软件开发工程师

美国麦迪逊

- **构建并扩展 EMC2**: Epic 内部软件配置管理平台 (VB.NET on InterSystems Caché), 涵盖版本控制、依赖管理与自动化构建 / 部署——一套内部 CI/CD 系统。
- **开发客户版本发布与升级工具** (控制台与 Web), 通过安全通信框架自动化版本检查与更新推送, 确保医疗系统运行于当前受支持的版本。
- **扩展 Trial Dev Environment**, 支持安全的高风险变更集成, 并解决构建 / 部署相关升级问题。

HCR ManorCare

2011.01 – 2012.04

数据仓库开发工程师

美国托莱多

- **构建 SSIS ETL 管道**, 为拥有 500+ 医疗机构的全国性医疗集团驱动企业数据仓库: 从多数据源抽取、验证、转换并加载至一致的维度 / 事实表模型。
- **开发 SSRS 报表与看板** (交叉表、钻取、参数化、级联), 支持财务与运营商业智能。
- **创建 SQL Server 对象** (存储过程、索引视图、UDF), 支撑报表逻辑与 ETL 验证。

教育背景

Bowling Green State University

计算机科学硕士

2010.08 – 2012.05

GPA 4.0/4.0 · 美国鲍灵格林

武汉大学

软件工程学士

2006.08 – 2010.06

GPA 3.5/4.0 · 中国武汉