

20+ 年  
一线工程

3 款  
上线运营的 AI 产品

243 万段  
古籍对勘语料

17 个月  
生产环境持续运营

## 个人简介

PHP 官方手册中文翻译组成员, 20 年一线工程 + 15 年 0→1 经营; PHP 3.x 那代人里最早转做生产级 LLM 应用的一批。近 18 个月全职独立设计、开发并运营 3 款生产级 AI 产品, 从检索、评测、安全到 CI/CD 全链路负责。方向: AI 应用工程 · 检索 · 评测与安全 · 技术管理。寻求可远程的长期全职技术负责人 / AI 工程 Lead 角色。

## 工作经历

2003 至今 · 门户架构 → 消费品牌 → 生产级 AI

### 2003 门户时代 0→1 架构

中金在线首位技术员工, 从首行代码搭到日 PV 3,000 万、带 ~20 人技术团队。

### 2010 从工程走向经营

创立男人袜, 15 年把一个想法做成 42 万会员、570 万双履约的持续经营闭环。

### 2024 全职投入生产级 AI

把经营交给团队, 独立交付 3 款上线运营的 AI 产品, 用可迭代的系统而非 demo 证明判断。

## 如是我闻 · 大规模古籍长文本 RAG 系统 [rushiwowen.co](http://rushiwowen.co)

独立设计 / 开发 / 运营

- 角色** 面向 CBETA 佛典语料(4,515 部 / 243 万段)的大规模长文本 RAG 系统, 覆盖检索、生成、评测、安全与交付的完整生产链路, 让海量古文可精确检索、可溯源问答。
- 动作** 查询扩展 + 混合召回(pgvector / PGroonga)+ 二阶段语义重排 + 流式生成 + 引用逐字核验; 建 rubric + LLM judge 离线评测与确定性「引用核验」CI 闸(纯 DB、无模型); 24 数据模型 / 115 API, 700+ 自动化测试全自动 CI/CD。
- 结果** 上线 17 个月持续运营; 自研古汉语英译在盲评中 88/100, 通用 MT(DeepL-CBETA)基线 19; 中英日三语、70 万条 AI 译文(50 万带向量); 已被多伦多大学、布法罗大学的佛学研究指南收录。

## 拍猫猫 · 多模型 AI 视觉流水线

微信小程序

- 角色** 现实版「抓猫」小程序, 验证边缘部署下的多模型视觉流水线与确定性生成。
- 动作** Workers AI DETR 检测 + 视觉大模型复核(结构化 JSON Schema 约束)+ gpt-image 异步生成; 全栈跑在 Cloudflare 边缘(Workers / D1 / R2 / Queues)。
- 结果** 单文件无框架路由, 全边缘响应; GitHub Actions 自动部署, Headless E2E 覆盖核心流程。

## 净心禅堂 · 沉浸式静心产品 [rushiwowen.co/zen](http://rushiwowen.co/zen)

独立技术栈 · 出海

- 角色** 如是我闻旗下独立产品, 验证同一内容资产在不同场景下的产品化与多语言出海。
- 动作** 独立设计系统与技术栈: Canvas 粒子 / Web Audio / Wake Lock, 8 语言国际化, 路由级进出场动效, 全套 PWA。
- 结果** VibeCafe 日活榜 Top 10, 8 语言覆盖海外市场。

## 男人袜 · 男性袜品订阅电商 [nanrenwa.com](http://nanrenwa.com)

创始人 / 技术负责人

- 角色** 创立「像订杂志一样订袜子」的垂直订阅品牌, 独立设计商业模式并长期带技术团队。
- 动作** 全链路自研电商中台与订阅履约系统; 2014 年国内首批接入微信支付; 2023 年上线基于 GPT 的 RAG 客服系统(检索 5 年历史客服对话)。
- 结果** 经营 15 年(在营), 42 万会员、570 万双履约, 完成两轮融资; 团队峰值 13 人, 现带 3 人精简团队维持稳定运营。

早期经历 中金在线(cnfol.com) · 2003-2009 · 首位技术员工 → 技术负责人 | 网龙(港股 0779) · 2009-2010 · PHP 项目经理, 用户中心 UAP 从零搭建、引入 TDD 工程规范

核心能力

|                    |                                                                                                             |
|--------------------|-------------------------------------------------------------------------------------------------------------|
| AI 应用工程<br>RAG 全链路 | 查询扩展、混合召回、二阶段语义重排、流式生成、引用逐字核验；多模型编排(DeepSeek / OpenAI / Qwen)。                                              |
| LLM 评测<br>与安全      | rubric + LLM judge 离线评测、提示注入防护、内容相关性分类、确定性「引用核验」CI 闸。                                                       |
| 全栈<br>与边缘          | Next.js / TypeScript、PostgreSQL / Prisma / pgvector / PGroonga、Cloudflare Workers / D1 / R2 / Queues、微信小程序。 |
| 工程<br>交付           | 全自动 CI/CD(类型检查 / 迁移 / 部署 / 回滚 / 生产冒烟)、生产监控与排障、算力与 API 成本控制、700+ 自动化测试、遥测与离线评测体系。                            |
| 技术<br>管理           | 带过 ~20 人技术团队；RFC 驱动的技术决策、异步协作与代码审查规范；TDD 工程规范化；0→1 到 1→N。                                                   |

独立开发者 & 开源

2000 至今 · 从架构到上线的完整交付

独立完成 3 款 AI 产品的架构、开发与上线运营，跨 Web / 小程序 / 边缘计算全栈实战。PHP 官方手册中文翻译组成员、GitHub Arctic Code Vault 贡献者 · Pull Shark ×4；知乎技术与行业分享累计 3,700+ 赞同。

|                     |                                            |      |                           |                                   |     |
|---------------------|--------------------------------------------|------|---------------------------|-----------------------------------|-----|
| nova-activitylog    | PHP · Laravel Nova 活动日志                    | ★ 99 | laravel-vue-test-use-jest | PHP · Laravel Mix 4 / Jest 集成与覆盖率 | ★ 9 |
| qwen-ai-provider-v5 | TypeScript · Vercel AI SDK 的 Qwen provider | ★ 6  | deepseek-php-client       | PHP · DeepSeek API 客户端 · 贡献者      | 开源  |
| doodles faces       | Canvas / TS · 确定性头像生成 · 公开源码               | 开源   | phppe.net                 | 早期 PHP 教学网站 · 创始人(2000s)          | 先行  |

先行者 2000 年国内首批 PHP 程序员，参与 PHP 官方手册中文翻译(全球仅数十人)。如是我闻、净心禅堂、拍猫猫三款 AI 产品均由本人独立设计、开发并持续运营。

AI 判断与行动

|                                                          |                                                                          |                                                                |
|----------------------------------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------|
| 2023 男人袜上线生产级 AI 客服                                      | 2024 早期押注 DeepSeek 做多模型 RAG                                              | 2024 全职 All-in 生产级 AI                                          |
| GPT 刚可用即把 5 年客服对话建成 RAG 检索库、做成生产级客服系统，国内消费品牌里较早的 LLM 落地。 | 在 DeepSeek 还未成为主流选择时将其定为 RAG 主力(推理成本仅 GPT 的几分之一)，并保持 OpenAI / Qwen 可热切换。 | 把 15 年的经营重心交给团队，独立 18 个月交付 3 款上线 AI 产品，用「上过生产的系统」而非 demo 证明判断。 |

教育背景

华中科技大学 · 行政管理