

邓雅琪

yaqidengtravail@gmail.com | https://dengyaqi.github.io/

教育背景

法国索邦大学(Sorbonne Université)	2017.09 - 2019.10
语言与计算机, 全日制, 硕士, 主修 NLP。	
法国巴黎第七大学(Université de Paris 7 Denis Diderot)	2014.09 - 2017.09
计算机语言科学, 全日制, 本科, 主修 NLP。	
ESGCI(École de Commerce International et Marketing à Paris)	2013.09 - 2014.09
金融语言预科, 全日制, 语言预备科。	
深圳职业技术大学	2010.09 - 2013.06
应用法语, 全日制, 专科。	

工作经历

AI 研发主管	杭州格耐马网络科技有限公司 - 杭州, 中国	2026.04 - 至今
工作内容: 负责公司 AI 研发方向规划、技术路线选型与项目交付推进, 主导企业级 AI 业务平台、Agent/RAG/Workflow、卡系统数据治理、Chrome 侧边栏业务副驾驶、数据采集工具链等项目从 0 到 1 落地; 对接业务、产品、前后端与运维资源, 推动需求拆解、研发排期、上线验收、生产问题追踪和迭代治理。		
资深软件开发工程师	软通动力技术服务上海有限公司 - 上海, 中国	2025.05 - 2026.04
工作内容: 负责 RAG 系统、向量知识库、智能问答引擎与 AI 自动编码平台研发, 主导多 Agent 协同、LangGraph 工作流、流式响应、模型服务接入等能力建设, 推动 AI 工具从技术验证走向可交付应用。		
高级 Java 开发工程师	北京汉克时代科技有限公司 - 杭州, 中国	2023.06 - 2024.10
工作内容: 驻场蚂蚁集团网商风控项目组, 负责一站式策略研发、行业风险管理及风险策略等风控中台的后端开发。通过深入的需求分析、系统架构设计, 提升系统整体效率, 同时在团队内部提供技术指导, 与项目组保持高效沟通并优化工作流程。		
安卓软件开发工程师	浙江清华柔性电子技术研究院 - 杭州, 中国	2021.11 - 2022.12
工作内容: 负责多个医疗方向安卓软件的全栈开发, 涵盖软件设计、编码调试及发布后支持, 确保软件高效迭代和稳定运行。		
自由职业者	巴黎, 法国	2019.10 - 2021.08
工作内容: 负责 NLP 算法开发、Java Web 应用设计及安卓软件开发, 通过高效的技术解决方案推动了项目交付和提高客户满意度。		
自然语言处理工程师	Yseop - 巴黎, 法国	2019.04 - 2019.09
工作内容: 负责知识图谱项目研究, 包括知识本体融合算法的研发与验证, 并完成 Lexicon 网站的全栈开发, 显著提升检索准确度。		

NLP 项目经历

禾彩 AI Platform 企业级 AI 业务平台
项目简介: 基于 FastAPI、LangChain、LangGraph、OpenAI SDK、PostgreSQL、ClickHouse、Neo4j 和 Redis 构建企业级 AI 后端平台, 支撑多模型接入、SSE 流式对话、RAG 知识库、MCP 工具调用、Skills 能力包、组织权限、日报周报、飞书推送、小满/吉客云集成和客户背调等业务。
项目内容: 主导平台后端架构与核心模块拆解, 设计 API 路由、服务层、数据库 DAO、LLM 底座、Agent 工作流、RAG 检索、文件解析、OSS、Redis 锁、线程池和配置体系, 保证 AI 能力可以被前端、侧边栏和业务系统复用。
项目内容: 围绕生产化建设全链路审计, 记录 HTTP、Agent、LLM、Tool/MCP span, 并落入 ClickHouse 供问题追踪、调用分析和业务复盘使用。
工作业绩: 推动 AI 平台从单点问答升级为可扩展的业务中台, 形成模型接入、工具调用、知识库、权限、任务、审计、报表和外部系统集成的统一底座。
禾彩 AI Client 平台前端
项目简介: 基于 Vue 3、TypeScript、Vite、Element Plus、Pinia、Vue Router、Axios 和 SSE 构建 AI 平台 Web 客户端, 对接后端 REST/SSE 接口, 承载 AI 对话、知识库、任务中心、客户卡、规则中心、通知和业务工作台。
项目内容: 负责前端工程架构与核心页面落地, 设计登录鉴权、路由守卫、会话管理、Markdown 渲染、流式输出、停止生成、明暗主题、知识库管理、规则发布和客户画像重算等交互。

项目内容：配合后端交付客户卡、产品专家结果卡、三渠道回复建议卡、Agent 审计页和通知详情等业务模块，统一前后端字段、错误处理和发布验证。

工作业绩：将 AI 后端能力沉淀为业务人员可直接使用的 Web 工作台，并通过构建脚本、环境变量、接口封装和发布门禁提升前端交付稳定性。

卡类 Card System 与多源 ELT 规则治理平台

项目简介：面向跨境电商客户经营场景，建设客户卡、需求卡、产品卡和多源 ELT 规则治理体系，对接 WhatsApp、阿里国际站、小满 CRM、吉客云 ERP 等数据源，形成客户画像、需求识别和产品推荐的业务闭环。

项目内容：主导数据建模与规则链路设计，覆盖数据接入、字段清洗、实体对齐、规则配置、规则发布、客户画像重算、标签投影和结果卡展示。

项目内容：设计从业务数据到客户卡/需求卡/产品卡的投影逻辑，使销售、运营和 AI Agent 能够围绕同一客户上下文协同工作。

工作业绩：补齐规则中心发布闭环与客户画像重算链路，减少业务规则散落在代码与人工表格中的维护成本，为后续推荐、背调和自动跟进提供稳定数据底座。

业务 Agent 编排、审计与客户背调 Recon 系统

项目简介：基于 LangGraph 和业务 Prompt 构建客户背调、公司/产品搜索、报告生成、审核、分享和跟进任务流程，支撑销售在联系客户前快速理解客户背景、需求线索和潜在产品匹配。

项目内容：设计业务 Agent 状态流转、工具调用、联网搜索、CRM 匹配、LLM 分析、报告生成和任务沉淀链路，并接入审计 span 记录执行过程。

项目内容：将 Agent 执行过程与 HTTP 请求、LLM 调用、MCP/Tool 调用统一追踪，便于定位失败节点、复核输出依据和持续优化 Prompt/Workflow。

工作业绩：让客户背调从人工搜索、复制、总结转为可追踪的 Agent 工作流，增强 AI 输出的可复核性和业务可运营性。

AI 业务副驾驶前端与 Chrome 侧边栏

项目简介：基于 Vue 3 + Vite 构建 Chrome Extension MV3 侧边栏，业务人员使用禾彩账号登录后，可在 WhatsApp 等工作页面读取客户上下文、历史消息、回复草稿和话术建议。

项目内容：负责侧边栏登录态、生产/本地环境切换、CORS 配置、消息上下文识别、客户跟进卡片、会话分析、建议回复展示和扩展发布流程设计。

项目内容：规划 Chrome Web Store Unlisted 非公开发布、版本递增、构建门禁、测试验证、扩展 ID 保持、回滚策略和生产 API 地址校验。

工作业绩：把 AI 平台能力嵌入销售实际工作页面，减少业务人员在系统间切换和手工整理上下文的成本。

Activity Recorder 用户行为采集器

项目简介：构建 Windows/macOS 活动采集器，采集本机操作行为并写入本地缓存、SQLite/JSONL 和 PostgreSQL，用于业务活动看板、员工日报、AI conversation 日报和工作过程分析。

项目内容：负责核心采集代码、浏览器扩展、PyInstaller 打包、开机自启动、自动更新、诊断日志、权限确认、平台适配和数据库写入链路。

项目内容：在 macOS MVP 中设计 Accessibility/Input Monitoring 权限边界、受保护应用降级、低置信键盘兜底、窗口切换快照和 LaunchAgent 自启动机制。

工作业绩：将分散的人工工作过程转化为可记录、可汇总、可诊断的数据流，为后续报表、客户跟进和 AI 辅助分析提供行为侧数据基础。

AI 求职/招聘自动化助手

项目简介：构建本地优先的求职/招聘工作台，用于将简历和岗位 JD 转化为可执行求职材料，并跟踪投递 Pipeline。

项目内容：后端采用 FastAPI + SQLAlchemy，按 Clean Architecture 拆分 domain/application/infrastructure/api；前端采用 React + TypeScript + Vite + Tailwind，Docker Compose 编排 PostgreSQL、Backend 和 Frontend。

项目内容：实现 JD 导入、简历/JD 匹配分、关键词缺口、简历要点改写、定制求职信、面试问题生成、投递状态流转和国内岗位结果池。

工作业绩：将求职材料改写、岗位筛选和投递跟踪产品化，沉淀出可复用的 JD 匹配、文本改写和流程管理能力。

AI Text Detector MVP

项目简介：构建本地优先的 AI 生成文本检测 Web 工具 MVP，前端使用 Next.js，后端使用 FastAPI，默认不长期保存原文，只保存摘要、检测结果、模型版本和分段预览。

项目内容：实现中英文文本检测、账号登录、报告按用户隔离、txt/docx/pdf 上传抽取、AI 生成概率、风险等级、置信度、分段评分、信号和警告输出。

项目内容：采用统计基线 + AI 味聚类的可解释模型，识别助手式回复、Markdown 模板、清单框架、抽象空泛词和缺少真实语境等特征。

工作业绩：完成从文本输入、文件解析、检测推理、报告解释到历史记录 MVP 闭环，为后续接入更重复核模型预留接口。

Activity Watch Dashboard 活动数据看板

项目简介：基于 FastAPI、SQLAlchemy、psycpg、Streamlit、Pandas 和 Altair 构建活动数据看板，读取 PostgreSQL 中的 ar_actor、ar_device、ar_session、ar_event 表，展示键鼠、窗口、浏览器、文件和采集器事件。

项目内容：设计后端健康检查、summary、actors、devices、sessions、events、event-types、timeseries 和 top 聚合接口，支持时间、用户、设备、事件类型、进程和关键词筛选。

项目内容：实现设备与员工映射、会话数/事件数/首次最近事件聚合、推荐索引 SQL 和一键启动脚本，便于业务人员查看采集数据。

工作业绩：把 Activity Recorder 产生的原始事件转化为可筛选、可分析、可复盘的运营看板，补齐“采集-入库-分析”的展示层。

飞利浦 PDF 处理和 AI 分析工具

项目简介：面向 ICF 安装确认表 PDF 文档处理，构建 PDF 转图片、AI 信息抽取、OSS 上传、Excel 对比和 API 查询工具，用于将安装文档关键信息与 IBS 系统数据进行比对验证。

项目内容：设计配置管理、字段映射、OSS 客户端、PDF 处理器、Excel 文件处理、Excel 对比、路径管理和 FastAPI/命令行入口。

项目内容：支持扫描 PDF、转换 JPG、调用百炼平台分析、验证图片 URL、生成结果 Excel、查询任务状态和下载处理结果。

工作业绩：将人工核对 PDF 与系统数据的流程自动化，形成从文件上传、文档解析、结构化抽取到对比结果下载的完整处理链路。

沉浸式开发(AI 自动编码系统)

项目简介：基于 FastAPI 构建轻量级 AI 自动编码系统，面向 VSCode/开发工作台场景，通过 Multi-Agent 协同、工作流编排、意图识别和流式响应辅助开发者完成代码生成、重构、调试与项目搭建。

项目内容：多 Agent 框架：基于 LangChain 设计代码、重构、调试等 Agent 分工，围绕开发任务拆解上下文、工具调用和结果汇总链路。

项目内容：智能工作流引擎：基于 LangGraph 编排意图识别、任务规划、代码生成、结果校验与流式输出，优化交互链路和上下文管理。

工作业绩：意图路由：自研意图识别与路由模块，将自然语言开发需求分发到匹配 Agent 或工作流，避免单轮 Prompt 堆叠导致的结果不可控。

工作业绩：模型服务：接入 vLLM/兼容 OpenAI 接口的模型服务，支持底层模型切换、部署验证和接口稳定性保障。

AI 智能文档解析与知识问答系统(RAG)

项目简介：基于 Django 构建企业级 AI 知识管理平台，集成文档解析、向量化知识库、智能问答、摘要生成和标签抽取等能力，支撑企业内部资料检索与知识复用。

项目内容：分布式向量知识库：基于 PostgreSQL + PgVector 构建知识存储中台，实现文档分块、Embedding 入库、语义相似度检索和上下文召回。

项目内容：智能问答对话引擎：集成 DeepSeek 等多模型 LLM，支持上下文记忆、流式响应、知识库问答和多轮对话。

项目内容：高性能服务架构：采用 Django REST Framework 构建 RESTful API，集成 Redis 缓存和异步任务处理，提升接口响应与批处理稳定性。

工作业绩：智能内容生成：基于知识库上下文生成问答对、文档摘要和标签，为后续 RAG 检索、内容运营和知识沉淀提供结构化素材。

多语言流行病检测系统项目

项目简介：传染病疫情的信息必须及时跟世界各国分享，以便及时围堵与消除，例如非典、Covid-19 等。

责任模块：通过信息抽取、字符细分化分析和依赖于文本体裁来实现《不受语言制约》的一个流行病检测系统。

工作业绩：数据采集：通过第三方合作与爬虫技术，采集各国第一线报道数据。

数据预处理：文章分割、模式提取、模式过滤、事件定位等获取文本体裁所需的元数据。

模型训练与优化：加入语言识别模块后，系统的准确率从 93%提升至 96%；召回率从 94%提升到 95%；错误率减少 1%。

日语语音识别与合成项目

项目简介：建立一个连续语音中的单词识别系统，并在 HTK(马尔科夫隐藏模型工具包)环境下对其进行验证。

责任模块：识别和定位给定的日语语音连续体中的关键词列表中的所有出现的词，并生成新的语音段。

工作业绩：语音分析：分析报道中的发言(发言顺序，识别不同人，发言现象的重叠，犹豫，连音，外来词，缩写等)。

数据预处理：识别音素在音段中为位置，使用 hmm 模型通过音素进行分割出 wav 文件，文件单位为单词。

模型训练：预处理完成的数据提供给 HTK 模型进行训练。

数据生成：生成词典(词和罗马音对照表)、生成语法(测试内容中不足的单词使用相似率最高的单词进行替换)。

语音生成：通过使用 MBROLA、Praat 软件和上述识别的模型，加上 Emofilt 添加情感，最终合成新句子。

Stream 法语评论自动摘要生成项目

项目简介：生成简约连贯的评论(好评和差评)摘要。

工作业绩：数据采集：通过 Scrapy 与 Sream API 采集百万条数据，并在 kaggle 下载相关数据，最终达到千万级数据量。

数据预处理：使用词向量、keyword、NER、KNN 等基本方法，对文本处理(不处理文本内的表情包，网络用语和缩写)。

机器学习：使用非监督学习进行情感、好评差评等自动分类。标签进行手动标注。

应用部署：使用 Flask、Bottle、Bootstrap 变成一个网络应用并且部署在服务器上。

NLP 技术栈

编程语言：Python、Java、C++、JavaScript、TypeScript、Kotlin、Matlab、R、PHP。

相关技术：LangChain、LangGraph、Agent Workflow、MCP、RAG、Embedding、Prompt Engineering、知识图谱、本体论、PostgreSQL、PgVector、ClickHouse、Neo4j、Redis、SSE、FastAPI、Django、LLM 多模型接入、文档解析、OCR、数据采集与预处理、数据挖掘与分析、数据标注审核、智能问答系统开发与运维。

开发工具：Pycharm、VSCode、Cursor、Codex、Claude Code、Brat、Praat、Labelme、Power BI、Weka、Excel。

后端项目经历

一站式策略研发平台 I 期、行业风险管理平台、风险策略平台等项目

项目简介：基于 Sofaboot 微服务框架构建分布式风控中台，集成多模态解析引擎、全链路仿真系统、智能搜索引擎三大核心模块。

项目内容：多模态文档解析引擎：自主研发支持 SQL/RQL/Word/PDF 等 15+格式的智能解析平台，文档类型覆盖率提升 300%；

全生命周期风控仿真系统：构建覆盖贷前准入、贷中监控、贷后预警的全链路仿真沙箱，策略验证周期缩短 75%。

低代码策略引擎平台：设计规则编排平台，支持业务方自助配置规则，策略上线周期从 2 天缩短至 8 小时；

分布式风控搜索引擎：基于 Elasticsearch 构建跨 MySQL/HBase/OSS 的统一检索中台，支持 50+字段组合检索；

高性能商城项目

项目简介：基于 Java SpringBoot 开发了包括商品与订单管理、用户权限、搜索等功能的商城系统，并实现了高并发秒杀模块。

项目内容：秒杀设计：基于 RocketMQ 控制订单提交，通过 Redis 做热点缓存与分库分表以实现降低数据库压力；

搜索优化：基于 ES 与 Redis 实现热门搜索词的二级缓存，并通过 Tag-Based Recommendation 智能推荐商品；

存储冷热分离：redis 缓存热门图片，实现智能置换淘汰。图片压缩及 lazy 加载，降低客户流量冲击；

适配普通机型：JVM 调优限制内存占用、避免进程 OOM，调优 GC 策略在性能和内存占用取 tradeoff ；

部署运维：通过 Kubernetes 实现自动化部署和运维，提升了系统的可伸缩性和管理效率。

RAFT 分布式系统项目

项目简介：完成一款具有高可用、强一致性的分布式 Key-Value 存储系统。

项目内容：论文复现：根据 Raft 论文，实现 Raft 一致性共识算法，具备 Leader 选举，日志复制，快照更新等功能；

数据迁移：基于 Multi-Raft 架构，对数据进行 shared 分区，并且可以在多组 Raft Group 中迁移；

优化存储接口：线性一致性的 key-value 存储接口，通过异步 Apply、FollowerRead 和 Prevote 等进行性能优化；

高性能分布式事务：实现了 MVCC 多版本并发控制，并基于 Precolator 模型，支持高效分布式事务。

后端技术栈

编程语言：Java、Python、C++、JavaScript、TypeScript。

相关技术：系统架构设计、SpringBoot、SpringCloud、FastAPI、Django REST Framework、MySQL、PostgreSQL、ClickHouse、Redis、Nginx、Kubernetes、Docker、Jenkins、Zookeeper、Kafka、RocketMQ、Elasticsearch、Git、分布式开发、JVM 调优、ES 优化、MySQL 性能调优、库表设计、并发编程、Alembic 数据库迁移。

相关工具：IntelliJ IDEA、Pycharm、Kibana、Visual VM、Kafka Eagle、Postman、Visual Studio、Xampp、StarUML、Docker Desktop、GitLab CI/Jenkins。

移动端项目经历

运动心电衣项目

项目简介：基于安卓原生实现一款实验室人员使用的连接心电衣设备与云后端实时通信的安卓移动端软件。

责任模块：安卓原生全栈开发。

工作业绩：UI 设计：严格按照 UCD 与产品定义的需求，基于 Figma 进行界面的交互设计，输出终端界面视觉设计图；
软件设计：整体框架使用 Jetpack 组件构建 MVVM 架构，包含实时监测、缓存、用户登入、权限管理、蓝牙操作等模块
前端开发：根据交互设计图实现自定义 View 页面架构和布局，基于 MPAndroidChart 绘制心电图。
设备交互：基于 nordic 库与自定义蓝牙协议交互，基于 Rxjava 异步回调监测实时数据、基于 greenrobot 实现本地数据库。
云端交互：基于 Retrofit+OkHttpClient 请求云计算资源和上发设备实时数据。

人工关节术中与术后项目

项目简介：完成两款软件帮助医生与实验人员对于人工关节设备的研究。1) 基于 Matlab 实现术中实时检测关节移动轨迹与数据上传云端。2) 基于安卓原生实现一款 APP，可实时链接嵌入式设备，术后监测关节压力并具有报警系统。

责任模块：安卓原生全栈开发，在运动心电衣 APP 的基础上进行。

工作业绩：UI 设计、软件设计、设备交互、云端交互等模块与运动心电衣项目类似。

UI 优化：持续优化产品的用户体验；

性能优化：不同移动设备不同版本的兼容问题优化、基于 Service 保持 CPU 持续运行、防止系统回收软件；

内存优化：使用 Leakcanary 检测内存泄漏，并修复例如 viewbinding 导致内存泄漏等问题；

移动端技术栈

编程语言：Java、Kotlin、HTML5、JavaScript、TypeScript、CSS。

相关技术：Android 原生开发、Jetpack、MVVM、蓝牙协议、IOT 设备联调、Chrome Extension MV3、Vue 3、React、Electron、Vite、Pinia、Element Plus、网络编程、搜索引擎、缓存设计、应用上架、UI 设计、多媒体处理、内存优化。

开发工具：Android Studio、Postman、VUE、Figma、Sketch、Photoshop、Illustrator、微信开发者工具、Chrome DevTools、PyInstaller。

其他

项目管理：Scrum。

证书：人工智能训练师四级、人工智能训练师三级、DELTA B2、C2 驾照。

语言：普通话(母语)、粤语(母语)、客家话(母语)、法语(商务)、英语(商务)、日语(流畅)。