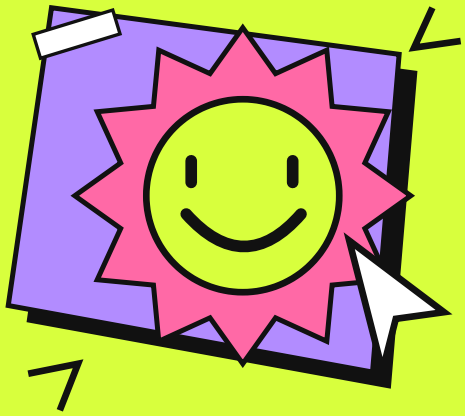
李梓靖

AI应用开发工程师

电话：173-8047-0757
邮箱：674846858@qq.com ｜ 成都
8年软件开发经验
目前承担项目合作，可随时到岗



01 教育经历

2014.09 - 2018.07 ｜ 四川大学锦江学院
本科 · 计算机科学与技术 ｜ CET-4

02 职业概述

8年软件研发经验，具备Java后端与Python AI应用开发能力。参与医疗AI产品建设，承担业务侧与算法侧技术对接，完成Dify工作流与Java服务集成；独立交付投标智能体和筑优CRM，具备数据聚合、异步调度、业务流程集成与上线经验。

03 核心技能

AI 应用与编排：Python / FastAPI / Pydantic / SQLAlchemy、Dify与Java集成；LangChain / LangGraph、RunnableLambda、Function Calling、Tools / Agent Skills分层封装、Structured Output、串并行编排、业务条件分流与HITL；Prompt Chaining、ReAct、Tree of Thoughts，任务拆解、输出约束与样本迭代。

RAG 检索增强生成：MinerU、表格上下文；固定长度分块、滑动窗口与重叠、结构感知切分、语义切分；Qwen3-Embedding / BGE-M3文本向量化；Chroma / Milvus / Elasticsearch；关键词 + 向量混合检索、来源关联与引用；PICO要素与医学实体提取集成；SAX流式解析、DTD映射、并行流水线、文件状态机、MD5校验与导入对账。

Harness 工程：Context / State、并发限流、超时重试、JSON修复、MinIO产物管理；Langfuse Task→Node→LLM分层追踪、节点级离线评测、LLM-as-a-Judge、回归、引文校验与反馈闭环设计。

全栈工程：Spring Boot / Spring Cloud Alibaba、MyBatis / MyBatis-Plus、MySQL / MariaDB、Redis / Redisson、Elasticsearch、Kafka / RocketMQ、Flink、XXL-JOB / Quartz；Feign / HTTP、Netty / WebSocket / SSE / MQTT；多租户、动态审批、候选检索与评分、迁移一致性；Vue 3 / TypeScript / Element Plus。

安全与信创：Spring Security、USB Key身份认证、访问控制；SM2/SM3/SM4、数字信封、密钥协商、隔离网络通信；麒麟、统信UOS及飞腾/海光/龙芯适配，商用密码认证支持。

部署与运维：Linux、Docker、Kubernetes；Jenkins + Docker + SonarQube CI；Prometheus + Grafana + Alertmanager监控告警，高可用与离线部署。

AI 协作与工具链：Claude Code / Codex、OpenSpec辅助需求澄清、编码、Code Review与回归；MCP配置集成、Skills、Obsidian + Git。生成式创作：ComfyUI / MiniMax H3、分镜、多模态Prompt与结果校验。

04 / EXPERIENCE

工作经历

从软件研发，到企业 AI 应用交付。

筑优CRM管理系统 | 项目合作负责人

2026.03-2026.05　2026.08-至今

- 以项目合作方式独立负责两期建设，覆盖业务调研、需求梳理、系统设计、前后端开发及上线。
- 完成资源管理与CRM经营流程、历史数据全量迁移，系统已在千余人规模企业日常使用。

四川省明信能源集团有限公司 | AI应用开发工程师

2026.06.01-2026.08.21

- 独立完成投标智能体调研、设计、前后端与上线，实现招标解析、审查、查重三项能力。
- 推动Pipeline迁移至LangGraph，统一执行控制、状态追踪与离线评测，系统已供商务部使用。

上海润达医疗科技股份有限公司 | AI应用开发工程师

2024.08-2026.03.17

- 承担病情分析、人文AI业务研发及算法侧技术对接，完成Dify工作流与Java集成，支撑62项分析场景交付。
- 建设临床数据模型、缓存异步与模块解耦；完成PubMed文献解析、批量导入及混合检索链路。

北京国泰网信科技有限公司 | 高级Java开发工程师

2022.05-2024.07

- 承担云密码安全服务管理平台核心架构与后端研发，处理隔离网络调度、代理转发和国密集成。
- 完成信创适配，推进6项商用密码产品认证；参与研发与交付规范建设，协助通过CMMI5认证。

四川省集创科技有限责任公司 | Java开发工程师 / 中台技术组组长

2017.09-2022.03（含实习）

- 由Java开发成长为中台技术组组长，负责数据字典、任务管理、USB Key登录等公共组件。
- 交付智慧公安、培训监管核心模块，参与需求与方案设计，负责外部系统及硬件厂商联调。

05 / SELECTED PROJECTS

项目经历

筑优CRM管理系统 | 项目合作负责人

2026.03-2026.05；2026.08-至今 | https://www.zhuyoucrm.com

技术栈：服务与数据：Java / Spring Boot / MyBatis-Plus / MySQL / Redis / Redisson / Elasticsearch；前端与集成：Vue 3 / TypeScript / Element Plus / 钉钉组织接口

分两期统一客户资源与经营流程，历史数据全量迁移，已在千余人规模企业日常使用。

- **组织与业务：**独立交付两期业务系统及历史数据迁移；对接钉钉组织，统一负责人、参与人与递归主管权限，动态解析部门审批链。
- **匹配与一致性：**ES候选检索、硬门槛、六维评分及去重支持正反向Top20推荐；以预检、事务写入、排队和持久化清单区分数据库与ES/附件同步，失败仅重试未完成部分。

投标智能体 | 项目负责人

2026.06.01-2026.08.21 | 明信能源商务部

技术栈：服务与编排：Python / FastAPI / Pydantic / SQLAlchemy / LangChain / LangGraph；文档与运行：MinerU / MySQL / MinIO / Langfuse / SSE；前端：Vue 3

招标解析、投标审查及标书查重，输出带原文依据的Word报告，已上线使用。

- **流程集成：**为展示与复用将自研Pipeline迁移至LangGraph，统一节点输入输出；按Tool-Skill-Agent分层封装11个原子工具、13个复合技能。
- **文档证据：**MinerU解析长文档与表格，结合规则校验、原文引用及检索兜底；支持元数据、目录、正文、图片四维查重和归属冲突人工确认。
- **运行治理：**统一上下文、任务/节点状态及MinIO产物，接入Langfuse三层追踪与SSE；集中处理并发限流、超时重试及JSON修复。
- **质量验证：**围绕155项判定真值构建节点级离线评测、缓存、LLM-as-a-Judge、引文校验与轮次对比，跟踪漏判和证据真实性；设计用户纠错、人工复核及样本回流。

05 / SELECTED PROJECTS

项目经历 · 医疗 AI

病情分析引擎 | 后端研发与系统重构

所属：上海润达医疗科技股份有限公司

技术栈：业务与集成：Java / Spring Boot / MyBatis-Plus / MySQL / Dify / Redis / Redisson；算法与指南检索：Python / 千问 Embedding / Milvus（医学指南存储、向量检索）

支撑62项已交付分析场景，多院区合计日均接入万余份原始临床报告。

- **数据集成：**以统一患者模型聚合HIS、LIS、EMR等7类数据；多院区数据聚合P95≤200ms，不含模型推理耗时。
- **医学指南检索：**采用Java业务侧与Python算法侧协作；Python服务调用千问向量化模型，使用Milvus存储医学指南数据并进行向量检索。
- **缓存与触发：**按数据变更触发分析，数据未变更时复用结果；同等业务请求量下，大模型实际调用次数减少约85%。
- **异步调度：**用信号量及优先级队列控制模型请求；同模型、同硬件压测，峰值QPS达原方案5倍以上。
- **服务解耦：**推动病情分析模块独立迭代与部署，同类需求平均交付周期缩短约65%。

人文AI智能培训平台 | Java后端与AI流程集成

所属：上海润达医疗科技股份有限公司

技术栈：业务与集成：Java / Spring Boot / MyBatis-Plus / MySQL / Feign / HTTP / SSE / Elasticsearch / Redisson；算法侧接入：Python / Pipeline

串联病例分析、角色对话、评分与报告，服务医疗人文关怀培训。

- **流程接入：**以Pipeline串联病例分析、角色对话、评分与报告，负责Java业务接入及会话管理。
- **状态与隔离：**实现多租户隔离、SSE流式推送及会话/流程进度持久化，管理业务运行状态。
- **评分与运营：**组织总分、维度、细项、扣分原因、正确做法5级评分，关联对话引用；记录运营数据，支持按租户、用户、模块及时间分析。
- **系统集成与状态管理：**通过Feign / HTTP接入Python算法服务，以Redisson分布式锁与持久化状态表协调会话和流程进度，关键行为写入Elasticsearch供运营分析。

PubMed医学证据数据库 | 数据处理与检索开发

所属：上海润达医疗科技股份有限公司

技术栈：业务与数据：Java / Spring Boot / MyBatis-Plus / MySQL / XML / SAX / DTD / Redisson；算法与检索：Python / BGE-M3（1024维）/ Elasticsearch / LLM API

建设文献导入、恢复处理及关键词 / 向量混合检索链路。

- **导入性能：**并行流水线导入约3700万篇XML文献，全量导入实测约45天→约12天，耗时下降约73%；SAX流式解析与DTD字段映射，进程实测峰值内存≤512MB。
- **检索与恢复：**文件状态机、MD5及重试支持恢复，核对源数据与入库记录数一致；集成摘要PICO与医学实体提取，基于BGE-M3、Elasticsearch实现混合检索。