

Chen Xiang 陈想

✉ xiang.ch@gmail.com | 🌐 cxiang.site | 🐙 github.com/Xiang-CH

Education

The University of Hong Kong

Hong Kong

Master of Science in Computer Science

Sep. 2025 – Present

- Relevant Coursework: Quantum Computing, Data Mining, Algorithmic Trading, Information Security, Big Data Management

Bachelor of Engineering in Computer Science

Sep. 2021 – Jun 2025

- First Class Honours, Dean's Honours List (2021-2022)
- Relevant Coursework: Machine Learning, Computer Vision, Operating Systems, Object-Oriented Programming, Data Science, Software Engineering, Algorithm Design, Databases, Network Communications, Natural Language Processing, Deep Learning

Experience

The University of Hong Kong

Hong Kong

Research Assistant II

Jul 2025 – Present

- Developing Legal AI tools in collaboration with the Law and Technology Centre, HKU, to support teaching and learning initiatives at HKU Law School.
- Processed and structured raw Hong Kong legal documents for context engineering, building a relational semantic database to enable efficient information retrieval in Agentic Retrieval-Augmented Large Language Model legal applications.
- Updating the AI Sentencing Predictor for Drug Trafficking on HKLII based on the latest legal developments.

Student Research Assistant

Jun 2023 – May 2024, Sep 2024 – Jul 2025

- Designed and led multiple hands-on AI workshops at the Innovation Wing, HKU, covering topics such as building custom chatbots using Azure services and developing multi-modal Retrieval-Augmented Generation (RAG) applications with layout detection for document understanding.
- Collaborated on student-led AI projects, including CLIC-Search: a semantic search engine for the Community Legal Information Center, and AI Historian Assistant: a multi-modal history research assistant.

iFLYTEK Co., Ltd.

Shenzhen

Assistant Software Development Engineer

Jun 2024 – Aug 2024

- Contributed to the agile development of AI English-speaking tutor and AI picture book companion LLM solutions within the education product line.
- Developed and maintained application pipelines using *Node-RED* to support customized feature requirements.
- Automated data engineering workflows for model fine-tuning with *Python* scripts and optimized prompt tuning processes.

Projects

CLIC-Chat | *Next.js, Prisma, MS-SQL, Azure, LLM, RAG*

Jul 2025 – Present

- Developed a full-stack AI agent web application for Hong Kong legal document Q&A using *Next.js* and *Prisma*.
- Processed and extracted information from raw Hong Kong legal documents (legislation, judgments, and articles) by parsing and cleaning diverse data sources, and built a relational database in *MS-SQL*.
- Created a hybrid search index (vector + keyword) on *Azure AI Search* by semantically chunking and embedding legal documents.
- Implemented a Retrieval-Augmented Generation (RAG) pipeline featuring hybrid retrieval, relationship extraction, and reranking to improve document relevance and generate accurate responses using Large Language Models (LLMs).

CU-Campus | *React, React Router, Tailwind CSS, Vite, i18n*

May 2024 – Oct 2024

- CU Campus is a course review and calendar scheduling platform for CUHK students, seamlessly integrating multiple campus utilities to create a comprehensive one-stop information portal.
- Actively participated in UX design, feature planning, and iterative development cycles.
- Engineered the frontend using *React*, tri-lingual localization with *i18n*, and integrated with backend RESTful APIs.

See more at: cxiang.site/project

Technical Skills

Languages: TypeScript, JavaScript, Python, SQL, HTML/CSS

Frameworks: Node.js, Next.js, Flask, FastAPI, Tailwind CSS, Prisma, Vue.js, Chrome Extensions

Databases: MS-SQL, MongoDB, PostgreSQL, MySQL, ChromaDB

Developer Tools: Git, Docker, Linux, Azure, AWS, Postman, OpenAI API, Vercel, Clerk, Node-RED, Dify, n8n

Libraries: React, Pandas, NumPy, Matplotlib, PyTorch, Scikit-learn, LangChain, BeautifulSoup, Selenium, AI SDK

陈想

✉ xiiang.ch@gmail.com | 🌐 cxiang.site | 📄 github.com/Xiang-CH | 📍 上海/香港

个人简介

求职方向：AI 应用、LLM 工程与全栈开发。专注 Agent、RAG、MCP 等 LLM 应用工程，最熟悉 Python、TypeScript，具备数据处理、全栈开发、数据库、云平台部署与 CI/CD 的完整项目经验，熟练 Git 协作开发与 AI 辅助开发 workflow，并有开源项目贡献。

教育背景

香港大学

计算机理学硕士（研究生）

2025/09 – 2026/11 (预计)

- GPA: 3.39/4.3; 相关课程：量子计算、数据挖掘、信息安全、大数据管理

计算机科学工学学士（本科）

2021/09 – 2025/06

- GPA: 3.56/4.3; 一等荣誉学位、院长荣誉名单
- 相关课程：数据结构、算法、软件工程、数据库、机器学习、深度学习、自然语言处理

实习经历

苹果公司 AI 开发实习生

2026/05 – 2026/09 (预计)

- 主导测试工程师故障监控分析 Agent 插件从需求访谈、架构设计到交付，封装为 Agent Skill；配套开发 MCP Server、CLI 工具和后台 PostgreSQL 系统，并整合 Human-in-the-Loop 工单流程，支持人工创建与关联工单。
- 针对 iPhone 产线验证阶段 MB 级测试数据，设计定制 MCP Server：将原本多步骤、易出错的数据拉取流程封装为单一 Tool 调用，提升流程成功率，同时数据本地化处理避免原始数据进入 Agent 上下文；结合多子 Agent 按站点并行完成错误归因分析，日报生成周期由数天压缩至一小时左右。
- 搭建 Skill 仓库的 CI/CD 发布流程，包括 AI PR 检查、自动打包制品发布、一行命令安装及智能体自更新机制。

香港大学 研究助理（兼职）

2025/07 – 至今

- 面向香港大学法学院教学场景设计法律 AI 工具，并为港大法律与科技中心开发、维护 AI 法律研究平台。
- 维护并优化 HKLII 平台的 AI 量刑预测器：利用大模型对超过 2,000 份判决书进行结构化信息提取并完成全量人工审核，搭建供超过 40 名法律学生使用的人工检查平台，并基于结构化数据开展统计分析与预测模型构建；模型中点误差不超过 3 个月，并持续适配香港最新法律条文。(项目仓库)

学生研究助理

2023/06 – 2024/05, 2024/09 – 2025/07

- 在香港大学创客空间设计并主导多个实践型 AI 工作坊，涵盖基于 Azure AI 服务构建聊天机器人，以及结合版面检测 (Layout Detection) 的多模态 RAG 文档理解应用。
- 参与跨部门学生 AI 项目，包括 CLIC-Search (社区法治信息中心语义搜索引擎) 与多模态历史研究助手。

科大讯飞 软件开发助理工程师

2024/06 – 2024/08

- 参与教育产品线 AI 英语口语私教与 AI 绘本陪读等 LLM 应用的敏捷开发，使用 Node-RED 设计并维护应用工作流；使用 Python 自动化模型微调的数据工程流程，并通过 LoRA 微调适配产品功能需求。
- 搭建并改造基于 LLM-as-a-Judge 的评测流程，辅助筛选和分析 bad case (失败样例)，定位问题类型并迭代 Prompt，完善从 bad case 发现、归因到 Prompt 优化的闭环。

项目经历

CLIC-Chat

- 开发面向香港法律内容问答的全栈 AI Web 应用，使用 Next.js 与 Prisma 构建系统架构。
- 解析并清洗香港法律数据源 (包括法规、判决书及法律文章)，提取结构化信息并构建关系型数据库。
- 对非结构化法律文本进行语义分块与向量化，在 Azure AI Search 上构建向量 + 关键词 (BM25) 混合检索索引。
- 实现基于 RAG 的问答系统，包括查询扩展 (Query Expansion)、混合检索、关系信息提取与结果重排序 (Reranking)，形成从数据处理、检索到生成的完整链路。

zhipu-ai-provider

- 开源并维护为智谱 (Z.ai) GLM 系列模型开发的 Vercel AI SDK 适配器，支持文本生成、流式输出、工具调用、多模态、推理及 Embedding/图像生成能力。
- 项目发布至 npm，下载量超 23 万次，被广泛应用于 Typescript 生态的 AI 项目中。

更多项目：cxiang.site/project

技术栈

AI / LLM: RAG, Agent, MCP, LLM-as-a-Judge, Prompt Engineering, LoRA, LangChain, Azure AI Search, Vercel AI SDK, OpenAI API, Dify, n8n

编程语言: TypeScript, JavaScript, Python, SQL, HTML/CSS

框架与平台: Node.js, Next.js, React, Vue.js, TanStack, Flask, FastAPI, Tailwind CSS, Prisma, Chrome 扩展

数据库: MS-SQL, MongoDB, PostgreSQL, MySQL, ChromaDB

工具与部署: Git, GitHub, Docker, Linux, Azure, AWS, 阿里云, Cloudflare, Vercel, Postman, Node-RED, Clerk

常用库: Pandas, NumPy, Matplotlib, PyTorch, Scikit-learn, BeautifulSoup, Selenium