

郝卓远

哈尔滨工业大学（深圳）

☎19537980096 | ✉hzy2210@gmail.com | 📱hhh2210

“热衷于探索 LLM 前沿技术及落地应用，具有 RAG 系统构建经验，熟悉 Info Retrieval”

教育背景

哈尔滨工业大学（深圳）

计算机与电子通信集群-计算机科学与技术

深圳

2023.09 - 至今

研究经验

融入知识图谱和大模型的多跳问答系统

哈尔滨工业大学（深圳）

团队负责人

2023.10 - 2024.12

- 组织团队成员进行科研探索，安排分工
- 独自重构最后一版代码，提出了新的算法，在 Long Context Benchmark-NarrativeQA 上超越了 ICLR24 RAPTOR
- 成果：在“大一立项”科研项目中获优秀奖
- 技术栈：语义检索、知识图谱和 LLM 紧耦合、知识图谱构建、知识图谱推理

RAG-Long Context、Knowledge Graph

哈尔滨工业大学（深圳）

研究方向

2024.12 - 至今

- 具体工作不方便透露，可参考 ICLR25 Oral Inference Scaling for Long-Context Retrieval Augmented Generation
- 预期成果：投稿 NIPS 2025/EMNLP 2025

项目经验

Cherry-studio 知识库系统

CONTRIBUTOR

2025.2 - 至今

- 项目描述：一款支持多个大语言模型（LLM）服务商的桌面客户端，20k star+
- 个人工作：尝试着手知识库重构开发，专注于 RAG 检索模块优化

memobase (1kstar+)

深圳

CORE CONTRIBUTOR

2025.3 - 至今

- memobase 项目参与：Agent-memory layer，参与日常 feature 开发工作
- HumanRAG：作为算法核心成员，为用户打造 SOTA 的 User-Docs 知识库，并计划在后期与 memobase 记忆系统完成耦合

云服务器与大模型 API 服务

独立运营

2025.1 - 2025.2

- 项目描述：使用腾讯云服务器，通过 lobechat/Librechat 集成 OPENAI、Claude、Gemini 等大模型 API
- 特点：支持多用户并发、文件消息各端云同步，向大学城内部（清华深圳、北大深圳、哈工大深圳）试点推广并实现盈利

Github 项目参与

CORE CONTRIBUTOR

- r1-overthinker (400star+)：将 unsloth 的调用模块替换成了 vllm，来支持 multi-gpu 的推理，并添加了本地模型管理模块
- paperlists (300star+)：添加了一个本地 paper 搜索模块，便于使用多关键词，面向多个顶会论文集进行快速定向检索
- Awesome-Tech-Share：为便于技术知识的交流与共享，和上交、华科等同学一起搭建了一个技术分享网站

文体并进

学校赛艇队

哈尔滨工业大学（深圳）

男八艇主力领桨七号位

2023.9 - 2024.6；2025—至今

- 获深圳市及地方市级冠军
- 培养了团队协作、领导力、时间管理等能力