

胡宇哲

电话：18351614218 | 邮箱：925762849@qq.com

教育背景

南京邮电大学，计算机科学与技术，本科	2020.09 - 2024.06
南京邮电大学，计算机科学与技术，硕士	2024.09 - 2027.06

项目经历

近肾腹主动脉瘤临床真实世界数据挖掘 2026.01 - 2026.04

- 项目背景**：针对临床影印病历多源异构、高噪声、模态割裂与结构破损问题，搭建从数据清洗、多维度对齐到标准化入库的全流程 AI 处理系统，为临床预后挖掘与长期随访研究提供高质量数据底座。
- 多模态数据清洗与结构化**
 - 基于 OCR 与医疗文本规则实现影印病历噪声过滤、乱码剔除、表格结构化还原，处理公章遮挡、版式错乱、手写干扰等问题；对随访语音进行语义提纯与关键信息提取。
- 多维度数据对齐与归一化**
 - 完成患者唯一标识对齐、诊疗时序重构、医学术语标准化，消除跨模态冲突与表述歧义，实现文本、表格、图像、影像四类数据精准关联。
- AI Agent 与自动化工程实践**
 - 设计医疗数据处理 Agent，封装清洗、结构化、入库等独立 Skill；基于 MCP 实现多模型调度，支持批量处理与断点续跑，完成病历全流程自动化治理。
- 统计挖掘与知识库构建**
 - 采用卡方检验与 Logistic 回归完成临床预后风险因素挖掘；构建文字、表格、图像、影像四大标准化知识库，结构适配 RAG 检索增强生成，可直接支撑临床统计分析 with AI 科研。

面向超关系知识图谱的自适应流形表示学习框架 2025.01 - 2026.01

- 项目背景**：针对超关系知识图谱中异构拓扑导致的表示失真、实体角色转换带来的语义歧义问题，主导研发自适应流形表示学习框架 AdaGeM。
- 混合曲率流形构建**
 - 构建涵盖欧氏、双曲、球面的混合曲率乘积流形，设计拓扑感知门控机制；通过切空间消息传递，实现树状、环状等异构拓扑特征的动态自适应路由与无损融合。
- 角色感知与噪声过滤**
 - 定义核心图谱角色并设计专属投影矩阵，对几何嵌入进行动态上下文调制；在自注意力机制中注入微观结构偏置，引导模型聚焦合法高阶 n 元交互，有效过滤多跳噪声。
- 项目成果**：融合全局非欧几何流形学习与局部角色感知 Transformer，显著提升模型对复杂长尾图谱数据的结构保真度与关系推理能力。

静脉疾病知识图谱构建与可视化 2025.03 - 2025.08

- 项目背景**：针对医疗领域静脉疾病知识碎片化、检索效率低的问题，构建规范化知识图谱并实现高效检索与可视化。
- 领域知识抽取与预处理**
 - 基于 RAG 检索增强生成思想，完成医疗文本格式转换、分块预处理；结合大模型设计 Prompt 工程，强化实体关联与上下文记忆，精准抽取静脉疾病领域知识，有效缓解大模型幻觉问题。
- 知识图谱构建与优化**
 - 实现知识图谱实体消歧、冗余剔除与数据规范化，通过属性剥离优化图谱结构，显著提升知识检索效率。
- 图数据库工程实践**
 - 负责 Neo4j 图数据库全流程管理，完成 Windows/Ubuntu 跨平台部署与 JDK 环境配置，优化后台服务稳定性；熟练运用 Cypher 语法实现图谱精准查询、属性更新及节点/关系统计；搭建 dump/CSV 格式的导入、导出与备份体系。

学术成果

- AdaGeM: Adaptive Geometric Learning on Manifolds for Hyper-Relational Knowledge Graphs**
 - 论文已被 IJCAI 2026 Main Track 接收录用。

技能清单

- 编程语言**：掌握 Python（数据预处理/大模型脚本开发/知识图谱算法实现），适配科研与工程场景；了解 Java、C/C++，具备计算机专业基础编程能力
- 技术方法**：掌握 RAG 检索增强生成、大模型 Prompt 工程，可基于大模型实现领域知识精准抽取、上下文关联强化；熟悉知识图谱构建与优化相关算法
- 工具框架**：熟练使用 Neo4j 图数据库（Cypher 语法/跨平台部署/数据备份）
- 英语能力**：大学英语四级 651 分、六级 602 分，具备优秀的英文专业文献阅读与理解能力