

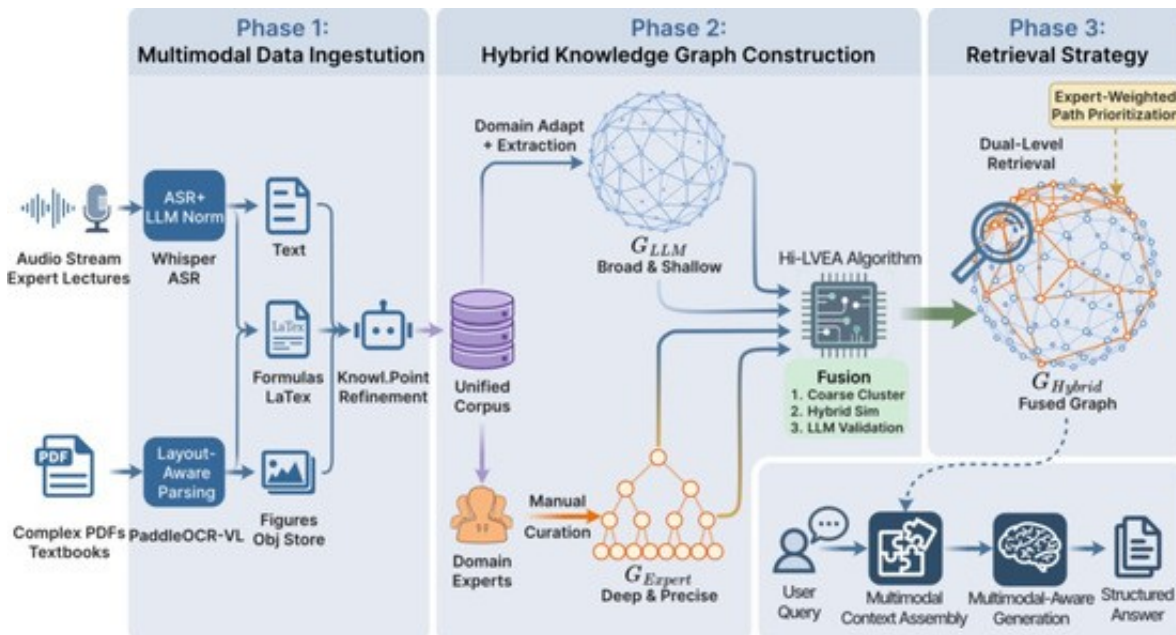
研究提出多模态图检索增强生成框架

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/40412.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

研究提出多模态图检索增强生成框架。近日，华南理工大学（以下简称华工）土木与交通学院2025届本科毕业生、2025级硕士研究生张鼎伦的研究成果发表于《专家系统与应用》（Expert Systems With Applications）。该成果提出面向领域知识增强的多模态图检索增强生成框架（DOK E-RAG），是其在导师、华工副教授陈太聪指导下完成的本科优秀毕业论文成果。



面向领域的知识增强多模态图检索增强生成框架。华轩供图

针对《结构力学》等土木工程核心课程中概念关系复杂、公式图表密集、知识层次性强等特点，传统大语言模型在回答专业问题时易出现依据不充分、推理不清晰、术语不准确等问题。张鼎伦等将专家知识、知识图谱与大语言模型相结合，提出DOKE-RAG框架，旨在服务于结构力学等深度垂直领域的智能知识问答。

该框架可从复杂PDF文档、课程讲授音频、数学公式、工程图表及表格等异构资料中提取多模态知识，并将其整合为统一的结构化知识图谱。同时，论文提出Hi-LVEA层级式大语言模型验证实体对齐算法，将专家构建的核心知识图谱与大语言模型生成的扩展知识图谱进行对齐、验证与融合。实验表明，DOKE-RAG在结构分析领域专业问答任务中的表现优于现有RAG和Graph RAG方法。

在论文研究和插图绘制过程中，张鼎伦尝试将AI绘图方法用于科研图示表达，相关经验整理成方法论文章发表于AI开源学习社区Datawhale公众号，并受人民邮电出版社邀请，独立撰写《AI科研绘图Nano Banana极速实战指南》，该书目前已正式出版。

据悉，张鼎伦本科期间曾参与学科交叉类大学生创新创业训练项目，参与智能语音引导式结构设计分析软件开发，在多模态信息处理、智能问答系统和大语言模型方向积累了交叉研究基础。其本科毕业论文成果体现了华工土木与交通学院在传统土木工程优势学科基础上，面向人工智能、智能建造和工程知识服务等新兴方向推进交叉融合培养的阶段性成效。（来源：中国科学报朱汉斌）

相关论文信息：<https://doi.org/10.1016/j.eswa.2026.133291>

作者：张鼎伦等 来源：《专家系统与应用》

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发