

Shale-Agent: 一种用于加速科研知识发现的 Text-to-Query 智能体架构

Monday, 25 August 2025 16:20 (20 minutes)

现代科学研究中，从海量非结构化文献中高效提取与整合关键数据面临巨大挑战。以中国散裂中子源的页岩孔隙结构研究为例，跨多篇论文的参数比较分析过程依赖于人工查询，其过程繁琐且效率低下。为解决此问题，本文设计并实现了一个两阶段智能体系统：首先，利用大语言模型通过结构化指令模板，将文献中的实体与关系抽取为 JSON 格式并存入数据库；其次，构建多层解耦的查询流程，通过意图分类、查询优化与自适应结果生成，实现对自然语言查询的精准响应。系统成功将非结构化文献转化为结构化知识库，实现了对复杂自然语言查询的准确响应。实验表明，关键参数提取准确率达 95% 以上，将跨文献数据检索与对比的平均耗时从数小时级缩短至秒级。本研究证实了所提出的 Text-to-Query 智能体架构的有效性，为特定科研领域的知识发现提供了一款高效、精准的自动化数据整合与分析工具。

Summary

Primary authors: 易, 杨; 胡, 鹏 (高能所); 李, 斌斌 (中国科学院高能物理研究所); 李, 亚康 (高能所)

Presenter: 易, 杨

Session Classification: 人工智能与应用

Track Classification: 人工智能与应用