大语言模型与知识图谱协同增强知识组织理论架构研究*

张心源[1,2]　肖浩宇[3]　李白杨[4]
（1. 苏州大学社会学院，苏州 215163；2. 苏州大学智能社会与数据治理研究院，苏州 215163；3. 郑州大学信息管理学院，郑州 450001；4. 南京大学数据管理创新研究中心，苏州 215163）

摘要：探究了大语言模型（Large Language Model，LLM）与知识图谱协同增强知识组织的有效结合模式，旨在提高知识组织的准确性与效率。依托LLM的自然语言处理与知识生成能力，以及知识图谱对结构化信息的表达和推理能力，提出一种协同增强的知识组织理论架构。首先，系统回顾现有知识图谱应用于知识语义组织的理论与实践发展，梳理LLM在科学知识抽取、实体对齐和图谱融合中的应用，论证知识图谱存储和检索的优化原理，调研LLM与知识图谱协同应用的典型项目。然后，在调研和归纳的基础上，针对知识组织具体操作环节，探究嵌入LLM的合理步骤，重构知识图谱实现知识组织的智能化流程，归纳总结图模协同增强知识组织的有效理论架构。研究发现，图模协同增强可以显著提升知识组织的精确度和可解释性，尤其是在跨领域、多源异构科学知识组织中，能够有效减少LLM的幻觉问题，并提高科学知识检索和问答的准确性与交互性。

关键词：知识组织；大语言模型；知识图谱；协同增强；语义组织
中图分类号：G254；TP18　DOI：10.3772/j.issn.1673-2286.2025.05.001
引文格式：张心源，肖浩宇，李白杨. 大语言模型与知识图谱协同增强知识组织理论架构研究[J]. 数字图书馆论坛，2025，21（5）：1-10.

知识组织最早由英国图书馆学家布利斯在1929年提出，国内学者将知识组织定义为知识因子的有序化和知识关联的网络化，是以知识为对象的整理、加工、表示、控制等一系列组织化过程及方法[1]。从国际知识组织学会（International Society of Knowledge Organization，ISKO）历届会议主题的演变可以看出，知识组织领域正在与数智时代接轨，从早期的针对大型数据集的知识组织与知识发现研究，到近年来的人工智能与知识组织，体现了技术驱动下的学科发展趋势。大数据环境下，知识呈现泛在、规模庞大、碎片化、动态更新的特征，各领域更要求智能化、专业化、个性化的服务。传统的知识组织体系如《中国图书馆分类法》《汉语主题词表》面临新的挑战，语义关联的知识组织成为未来发展方向[2]。在此技术革新背景下，两种关键技术对知识组织产生了深远影响。①知识图谱作为语义知识库，通过描述客观世界中概念、实体及其关系，为知识提供了高度互联、可解释的网络化表达。它对于实现深度的语义关联、支持知识推理和理解复杂关系具有核心价值，是应对碎片化和语义关联需求的理想选择。②以Transformer模型为基础架构的预训练大语言模型（Large Language Model，LLM）的兴起为知识组织领域带来了新的机遇。LLM具备千亿乃至万亿级超大规模参数，实现了词元级极细粒度的语义向量化表达，深刻地影响着知识组织领域的知

收稿日期：2024-10-22
*本研究得到教育部人文社科青年基金项目“全球数据主权博弈背景下健全我国数据跨境流动规则体系研究”（编号：21YJC870019）资助。

识抽取、知识标引、知识关联、知识挖掘等工作内容。2022年底以来，以ChatGPT为代表的生成式人工智能（Generative Artificial Intelligence，GAI）引领了新一轮的AI革命，其在语义理解、知识抽取、知识生成与推荐方面的出色表现，对提高知识组织的效率有极大助力[1]。蔡祎然等[3]指出，LLM可高效、精准地从科技文献中识别、挖掘领域知识、科学数据与图表信息等细粒度、多模态数据，而且LLM还可以通过少样本学习、思维链提示等将碎片化的科技文献内容转化为结构化证据链数据。

然而，这两类技术各有局限。知识图谱依赖于精确的、预定义的结构（符号系统），在处理开放领域、非结构化文本和捕捉动态变化方面不够灵活；LLM虽然在文本处理上灵活强大，但其生成内容可能存在事实性错误，其内部的黑箱推理过程缺乏可解释性。学界研究发现，LLM与知识图谱的协同增强框架（本文称为“图模协同”）能有效突破传统知识组织方法以及单一技术应用所面临的局限：①知识表征层面，将LLM的连续向量空间与知识图谱的离散符号系统进行融合，构建可解释的联合嵌入表示；②方法创新层面，在利用LLM的生成式补全能力动态扩展知识图谱边界的同时，借助图谱的逻辑约束实施生成内容的实时语义校验；③计算模式层面，图模协同增强知识组织可以形成良性的迭代增强回路，先通过LLM发现新的知识关联，再利用知识图谱进行验证和融合，使知识组织过程兼具LLM的规模效应与知识图谱的推理严谨性。

目前围绕图模协同增强知识组织的研究呈现“实践先导，理论滞后”的特点：已涌现出诸多具体领域的应用案例（如智能问答系统、文献知识库、个性化推荐引擎等），但支撑这些实践的系统性协同理论框架与优化路径论证仍显不足。①单向增强机制的局限。仅用知识图谱约束LLM生成，限制了跨模态知识交互的深度，未能充分释放符号逻辑与神经表示的双向互补潜力。结果是知识图谱的精细化与LLM的输出提升未能形成有效闭环，知识图谱作为静态的知识结构限制了LLM对未知知识疆域的探索能力。②知识表征割裂。知识组织构建时仍依赖传统文献计量指标与知识计量指标的独立计算，缺乏能够将LLM感知到的深度语义关联强度（如基于上下文预测的关联得分）与知识图谱固有的拓扑结构属性（如节点中心性、社区结构内的连接强度、路径相关性）有机统一的复合表征模型。③静态知识表征缺陷。协同系统难以捕捉知识概念的动态演化特征，例如当处理新兴领域或快速演进的学科时，新术语及其语义快速涌现或漂移，其语义关联对象发生巨大变化，现有的协同系统对此类变化反应迟钝。

鉴于此，本文提出图模协同增强知识组织的理论架构，旨在构建动态双向的图模协同框架，突破当前知识组织中动态性缺失、交互性不足、表征片面化的三重瓶颈。

1　相关研究与实践

1.1　知识图谱增强LLM

尽管LLM在许多应用中已经取得成功，但是也因幻觉问题和黑箱问题备受争议，这制约着LLM在金融决策、医疗问诊和法律咨询等场景的应用。LLM出现上述争议的原因在于其训练推理是一个非决断性过程，然而知识图谱是一种结构化和决断性的知识表示形式，基于知识图谱的推理过程具有客观的可解释性，因此可以用知识图谱审核LLM的推理输出，显著增强LLM生成准确率。已有研究和实践通过将知识图谱嵌入LLM预训练和推理环节实现知识图谱增强LLM。

（1）知识图谱融入LLM预训练。随着LLM的变革式发展和广泛应用，如何有效整合结构化知识以增强模型的推理能力和可靠性成为重要问题。通过知识图谱查询机制，可针对用户输入问题提取子图结构，将结构化知识转化为可被LLM理解的上下文表征，从而激发LLM在特定领域的知识推理能力，这种思路通常又称为LLM与图挖掘协同（LLM with Graph Mining），关键在于子图的边缘检测和图摘要生成。例如，KG-BART模型将知识图谱融入Facebook AI Research团队提出的BART模型，通过在编码器中融合知识图谱信息和文本信息，将构建好的知识图谱与BART进行融合，得到类似于BART的编码器-解码器结构[4]。Google提出的KELM模型通过将Wikidata知识图谱转换为自然语言文本，对LLM预训练语料进行扩充，将此合成语料库用于增强基于检索的REALM模型的预训练[5]。LMExplainer通过图注意力网络实现了大规模知识图谱中关键信息的精准定位与提取，优化了知识检索的效率与相关性，为LLM提供针对性的外部知识支持[6]。

由于将知识图谱融入LLM预训练的计算成本高、实时性较差，有学者提出将知识图谱整合到额外的模

块中，这类方法与低秩适应（Low-Rank Adaptation，LoRA）高效微调原理相通，即保持LLM的大部分参数不变，仅针对知识图谱编码模块进行训练优化，通过特征融合层实现LLM语义表征与图谱结构表征的有效融合。例如：CokeBERT可以根据文本上下文动态选择和嵌入知识，从而提高预训练模型的性能[7]；JAKET是一个针对知识图谱和文本联合训练的框架，包含知识模块和语言模块，知识模块利用图注意力网络为LLM提供实体嵌入输入，从而使知识模块的训练调整与语言模块互相独立[8]。

（2）知识图谱嵌入LLM推理。随着最新的LLM参数规模扩展至万亿量级，其内部知识表征已达到与传统知识图谱相当的复杂度，使得LLM具备了前所未有的知识生成与推理能力。在此背景下，先将结构化知识融入LLM推理，再借助思维链技术，LLM得以实现更为准确、可靠的查询推理。基本思路是将提取到的知识图谱三元组合并到LLM的提示中，一方面用以生成LLM可用的提示词内容，另一方面在知识图谱语义内容的约束下回答与知识图谱相关的问题。该设计模式的典型案例是Google的LAMDA[9]，原理是使LLM根据知识图谱中的语义信息实现推理过程，确保LLM生成知识的准确性。

引入先验知识和图谱信息，结合语义推理和注意力机制等技术，也有助于增强LLM对细粒度语义的理解和表达能力。从知识驱动LLM推理的角度来看，知识图谱作为包含逻辑知识和事实知识的知识库，提供了更丰富的信息来源。在训练阶段，未来的研究将越来越多地探索如何将结构化的高质量知识与LLM结合起来，例如使用图神经网络进行知识编码或采用新的编码策略。这些方法不仅能丰富LLM的知识库，还能优化其结构和训练目标。在推理阶段，整合知识库以检测和减少知识错觉，提高LLM输出的可靠性和准确性，将是一个关键的研究方向。

1.2　LLM增强知识图谱

知识图谱作为知识工程的重要基础设施，其本质是对客观知识的一种映射。知识图谱一般利用“搜索+知识库”的方法汇总整理网络数据，将知识立体化、形象化、直观化地展现给用户，可以精准、详细描述不同层次及不同粒度的抽象概念。知识图谱的技术实现流程包含5个核心部分：知识获取、知识融合、知识存储、知识检索和可视化展现[10]。这一技术架构为后续的知识组织方法提供了重要的技术基础。LLM可以在知识图谱嵌入、知识图谱补全、知识图谱到文本生成、知识图谱问答等方面提升知识图谱应用的质量[11]。知识图谱的一般性构建涉及知识抽取、融合、表示、存储与检索等多个环节，这些环节高度依赖自然语言处理技术（如命名实体识别、实体对齐）和图算法（如图嵌入、图查询、图推理）。为了驱动这些技术并保证结果质量，传统开发方法不仅要求开发人员熟悉相关的机器学习技术（尤其是自然语言处理和图嵌入模型），还需要投入大量资源，邀请多位领域专家全程参与知识审核和校验。这种对专业技术人员和领域专家的双重依赖，使构建和维护知识图谱的成本与门槛显著提高。知识图谱的构建和维护开销不菲，无论是人员技术要求还是人力成本，都限制了知识图谱的更新时效和规模，而LLM的预训练和微调机制恰好能够解决知识图谱开发和维护的高成本问题[12]。相比利用传统机器学习技术开发知识图谱，LLM利用自回归或自注意力机制来捕捉输入序列中的长距离依赖关系，具有强大的跨领域和跨语言的语言理解和生成能力，能够更加有效地进行自然语言建模[13]。史忠艳等[14]研究了DeepSeek等LLM在领域知识图谱构建中的应用，认为传统知识抽取方法存在泛化能力受限、对数据标注要求高、模型训练耗时长、知识更新困难等问题，而基于LLM的知识抽取方法凭借其对自然语言的理解和推理能力，在知识抽取、文本分类、语义推理方面展现了优势，为领域知识图谱构建提供了新的解决方案。总体来说，使用预训练的LLM开发知识图谱，能够简化知识抽取、知识表示等环节算法模型的训练步骤[15]。

1.3　图模协同增强

知识组织场景中的深度智能化往往要求突破前述单向知识传递模式（知识图谱增强LLM或LLM增强知识图谱）的限制。这些模式虽然有效，但主要聚焦于一次性的知识注入或离线式地支持知识图谱构建更新。相比之下，图模协同增强代表了一种更高层次的集成范式：旨在构建一个使LLM与知识图谱能进行实时交互、双向知识流动与反馈的协作闭环系统。其核心理念是充分发挥两者互补优势——LLM的语言理解与生成灵

活性，以及知识图谱的结构化事实表示与推理可靠性，从而实现“1+1>2”的效果。吴信东等[16]提出的BEKO双向增强框架具有重要意义，该框架不仅利用知识图谱增强LLM的生成效果，还利用LLM的推理结果补充知识图谱，从而实现知识的双向流动，并最终形成知识图谱与LLM之间的循环正反馈，不断优化系统效果。具体而言，当应用场景要求可验证知识来源或严格逻辑推导时，知识图谱的符号化表示与规则化推理发挥关键作用；而在需要处理模糊查询、多轮对话或实现复杂语义理解时，LLM的强大语言建模能力则占据主导地位。为了实现上述复杂的图模协同，特别是知识图谱与LLM的高效、动态协同，检索增强生成（Retrieval-Augmented Generation，RAG）技术提供了一种重要的架构思路。RAG的核心在于，当LLM面对查询时，引入外部知识库进行检索，并将检索到的相关内容（上下文）与LLM本身的知识结合起来，生成最终响应。这使得LLM能够有效利用并整合超出其内部训练数据边界或最新/特定领域的信息（常由知识图谱承载）。将知识图谱作为RAG框架中的核心外部知识源，并结合图谱自身的推理能力，是实现图模协同的一种强大且可行的技术路径。这一理念具体体现在以下研究和实践中。①MindMap方法是一种典型的图模协同方法，又称为即插即用的提示方法[17]。MindMap使LLM理解图谱输入，从而构建含有图谱信息的提示指令，用于支持推理证据的生成。②DeepMind团队研发了拥有互联网规模检索能力的高效预训练模型RETRO。通过检索包括网页、书籍、新闻和代码在内的文本段落数据库来增强检索，利用图神经网络将非结构化文本与其关联图谱子图联合编码，构建出包含大规模实体的知识图谱，形成语义增强检索空间，使LLM能同步调用参数化知识[18]。③微软Bing Chat的技术架构普罗米修斯模型使用协调器将知识图谱和搜索引擎等外部知识引入LLM，实现知识增强的LLM对话应用[19]。

综上所述，图模协同在科学知识组织领域具有较高的适用性和可行性。首先，以图模协同的方式处理知识组织问题，提高了科学知识组织的准确性和效率，有利于表示更为细粒度的领域知识实体、属性和关系类型，有利于实现更为精细化的知识推理和问答等科技情报决策服务。其次，图模协同方式降低了传统知识图谱开发的时间和人力成本，通过少量领域标注数据微调LLM，使得微调LLM具备领域科学文本知识抽取能力，结合具体的抽取提示词可以实现零样本或少样本的知识图谱构建。最后，知识图谱增强了LLM的知识推理和问答功能，提供了自然语言方式的科学知识相关的查询与响应机制，降低了对专业技术人员的技术依赖性，提高了科学知识组织服务的交互性和易读性。

2　LLM与知识图谱协同增强知识组织理论架构

LLM与知识图谱协同增强知识组织的实现架构可以分为4个递进式的功能模块：首先是LLM增强知识抽取，其次是LLM驱动实体对齐，二者均属于知识图谱的构建工作；再次是知识图谱的词元化处理，使之满足后续LLM的调用条件；最后是利用构建好的知识图谱实现LLM RAG系统，优化知识推理和知识检索效率。图模协同增强知识组织理论架构工作流示意图如图1所示。

2.1　LLM增强知识抽取

LLM增强知识图谱构建阶段的方案包括面向特定领域数据的提示微调和针对具体知识抽取任务的提示词工程构建。

提示微调指在LLM输入中添加特定的标注数据示例来引导LLM生成特定类型的响应。具体做法是：①准备训练数据（标注数据）作为示例；②使用训练数据构建提示微调算法的提示词，调用预训练通用LLM并设置损失函数、优化算法、学习率、训练轮数等，以优化LLM性能；③根据知识图谱构建的具体任务分工，单独生成提示工程的提示模板，如实体抽取模板、关系抽取模板、属性抽取模板等，且要满足引导LLM理解任务并生成期望输出格式内容的要求；④在测试集上评估微调后LLM的性能，确保其能够准确地完成具体的知识图谱构建任务，同时根据评估结果，选择是否继续优化提示模板或调整微调策略，以进一步提高知识抽取的准确性和效率；⑤保留满足任务需求的LLM设置和提示工程的提示模板。

提示微调算法实现原理如图2所示，冻结主模型全部参数，只在训练数据前面加入一小段提示微调的提示词，只训练提示的嵌入表示层。按照提示内容分为离散提示和连续提示两类：离散提示给予具体和明确的提示，要求模型按照给定的信息生成精确的结果，通常用于需要模型提供准确答案的任务；连续提示给予

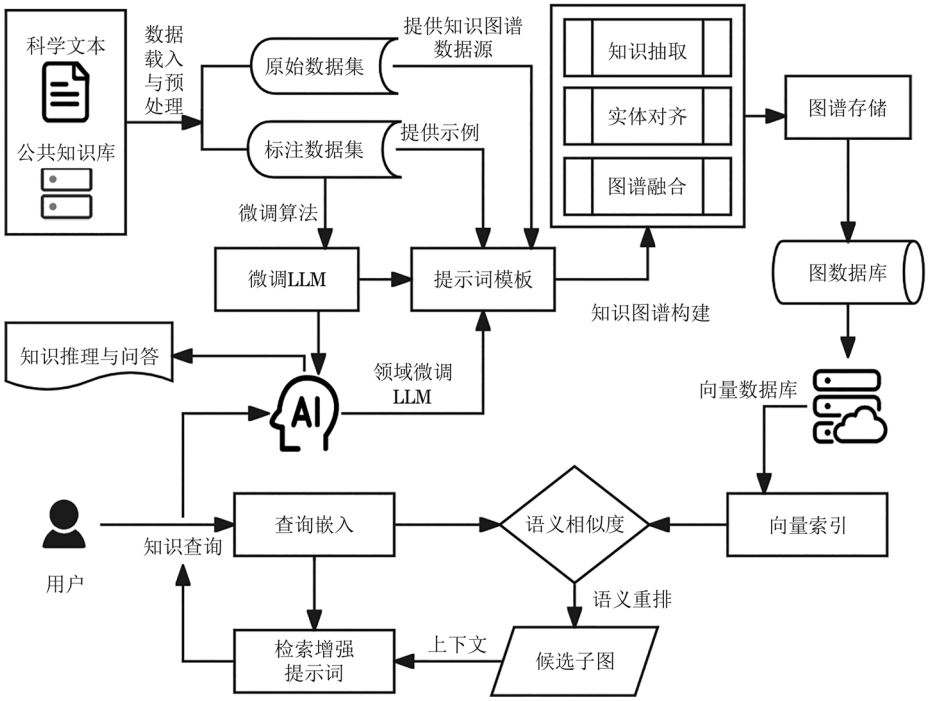
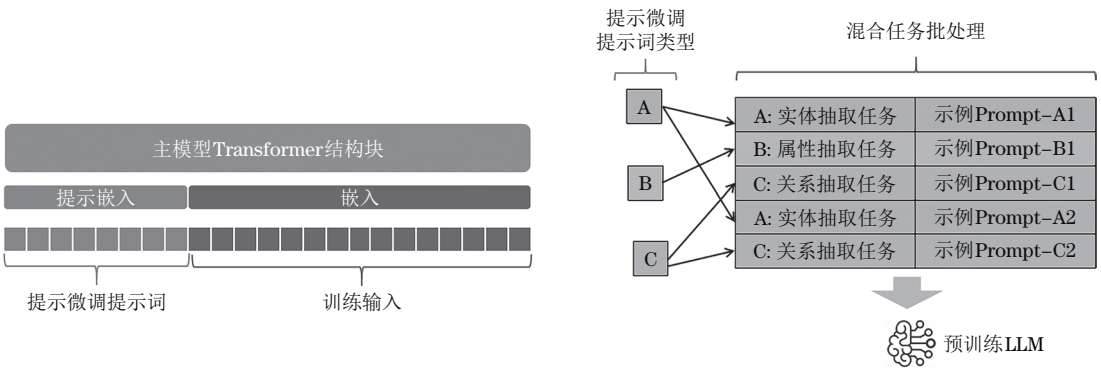


图1　图模协同增强知识组织理论架构工作流



(a)离散提示　　(b)连续提示

图2　提示微调实现原理

较为宽泛或模糊的提示，允许模型在生成结果时有更大的自由度，通常用于启发模型进行创造性的生成。此外，如图2所示，提示是一次性输入预训练LLM进行参数微调的，这种机制既减少了调参数量，又保证了微调LLM在下游抽取任务中的泛化能力。

2.2　LLM增强实体对齐

知识图谱嵌入又称为知识表示，是一种将知识图谱中的实体和关系映射到低维向量空间的技术，方便通过向量计算进行复杂的推理操作。现有的主流知识图谱嵌入方法有基于规则的规范化描述语言（如资源描述框架、网络本体语言）、基于语义网的图结构方法（如概念图、主题图）、面向对象的结构化表示方法（如本体）、基于机器学习的分布式表示方法（如TransE模型）、基于深度学习的网络表示方法（如网络嵌入）等。知识图谱嵌入是知识图谱自动对齐、融合以及图谱向量索引等任务的重要基础。然而，传统嵌入方法在捕捉领域知识的深层次语义和动态演化特征方面存在局限。为了提升嵌入表示的语义理解力和适应性，利用微调后的LLM来完成特定前沿领域知识图谱的嵌入表示。LLM强大的上下文理解能力使其能够更精准地编码实体和关系的丰富语义信息，同时其学习机制也为捕捉知识概念的动态演化提供了潜在可能。LLM

在自然语言处理任务中的应用给知识图谱嵌入提供了更为高效的技术方法，可以通过提示词引导LLM隐藏嵌入向量的输出，直接输出实体对齐或者实体融合的结果。

实体对齐是指在多个异源异构的知识图谱中发现完全等价的实体，即发现指代相同真实世界物体的两个实体。实体对齐涉及中英文实体对齐和同义词实体对齐两个方面。中英文实体对齐可以利用微调后LLM在双模语言上的翻译性能实现；同义词实体对齐需要在LLM嵌入表示的基础上，利用合适的提示词完成。同义词对齐是实体对齐的主要工作，因为知识抽取任务是基于多源、跨学科的科学文本完成的。多源数据知识抽取会造成大量同义词实体的出现和关系的冗余，导致知识图谱实体质量良莠不齐、关系重复等问题。

知识融合是指新构建知识图谱的后续更新融合。科学文本的发表是不停歇的，对应科学知识创新永不止步的过程。因此，需要对知识图谱进行及时、合理的更新，将新出现的科学文本中的实体、关系与原有知识图谱融合，并更新关联强度和时间信息。LLM在这一过程中发挥了重要作用，利用其强大的自然语言处理能力，可以高效地从大量文本中抽取实体与实体关系，从而加速知识图谱的更新。LLM不仅能够自动识别新出现的实体与关系，还能提升已有知识图谱中的数据质量与关联度，提高知识融合的准确性。此外，LLM还能够支持科学文本的实时解析和自动化处理，确保知识图谱的持续更新与优化，从而提升科学知识组织的时效性和准确性。这种技术创新不仅使得知识图谱构建过程更加高效，也能更好地适应科学研究的快速发展。

2.3 知识图谱的适LLM存储与可控信息生成

科学知识组织中知识图谱采用了五元组的表示方式，即{头实体，关系，尾实体，时间，关联强度}或者{实体，属性，属性值，时间，关联强度}。要将知识图谱存入图数据库，需要将五元组转换为图数据库可理解的数据结构。图数据库提供节点和边两种数据结构：节点用于存储知识实体和属性，边用于存储关系，边的属性用于存储时间和关联强度。图数据库提供为节点、边或者边属性建立索引的方法，能够实现对存储知识图谱的快速查询与修改，有利于知识图谱的及时更新，为可控信息生成提供了数据基础。

LLM输入单位是词元，并且不同LLM有不同的词元数限制，一旦超过该限制，LLM就会出现遗忘上下文问题，所以LLM本身不具备知识图谱的存储功能，但可以通过合适的功能开发，建立LLM与图数据库的数据接口，实现LLM增强下的知识图谱的存储和长期信息生成。向量数据库作为一种高效存储和搜索向量的数据库产品，提供了LLM与图数据库之间的数据接口，支持可控信息生成的实现。传统数据库支持点对点查询或者范围查询，而向量数据库支持相似度查询：将知识图谱或者相关子图作为LLM上下文，经过向量化后存储在向量数据库中，与LLM查询问题的嵌入向量进行相似度计算，筛选出相关的知识图谱片段，最终将这些结构化知识作为约束条件融入LLM的提示词信息或者微调信息，从而实现由知识图谱引导的可控信息生成。此外，最终相关子图向量一同融入LLM的提示词信息或者微调信息，这些信息被隐式地存储在LLM的参数中，这些参数信息反映了LLM在训练过程中学到的知识结构和模式。通过深度融合知识图谱与LLM，不仅能够确保生成信息的准确性和一致性，还能够根据特定的科学领域知识约束，实现更加精确和可控的信息生成，提升领域科学知识推理和回答生成的质量。

2.4 图模协同检索增强

基于LlamaIndex架构根据用户输入检索外部知识库，将相关内容作为上下文，同用户输入一起构建提示，引导LLM完成知识检索问答。具体实现过程包括4个模块功能的实现：嵌入模块、索引模块、检索增强模块、生成模块。

（1）嵌入模块。检索增强系统中，嵌入用于对输入的查询和检索到的外部知识图谱或文档进行编码，并使用编码的向量生成LLM响应。嵌入是LLM理解查询和检索外部知识库的数据格式，嵌入向量允许LLM理解不同文本片段之间的语义相似性。LlamaIndex可以同时使用OpenAI嵌入模型（如text-embedding-ada-002）和其他开源嵌入模型（如m3e或nomicembed-text-v1）。由于嵌入模型可接受的数据也受限于模型输入词元的长度，对于用户查询输入而言，文本长度通常在单个词元长度范围内。对于知识图谱而言，需要预先

对知识图谱进行文档切块操作，文档块是LlamaIndex架构的基本操作单位。文档切块由分割器完成，切块操作需要考虑两个因素：①嵌入模型的词元限制；②语义完整性对整体检索效果的影响。可以采用完整五元组为分割粒度或者固定词元长度进行分割。分割后的文档块经过嵌入模型生成向量矩阵后，将存入向量数据库供后续索引查询。嵌入模块的工作流程如图3所示。

（2）索引模块。索引模块承担两部分工作。①使用LlamaIndex为知识图谱建立向量存储索引。在查询时，LlamaIndex会查找节点相似度排名前k的节点，并将其返回给响应生成器。以中医领域的科学知识为例，查询“柴胡治疗抑郁症的有效化学成分有哪些”，需要返回与“柴胡治疗抑郁症”相似度高的知识图谱子图，并进行知识推理，找到相似度最高的化合物实体。②使用LlamaIndex为知识图谱建立关键词索引，从查询中提取关键词，只查询与关键词有映射关系的节点。关键词索引适用于知识检索和问答响应，如查询“柴胡用于抑郁症研究对精神疾病领域研究的影响”，查询系统只需要返回学科属性为精神疾病的相关实体所在的子图即可。此外，也可用于需要中英文双模语言转换的用户查询，如“查找柴胡药物在国外的最新适应证有哪些”，查询系统返回*Bupleurum*（柴胡）对应的适应证实体子图即可。

在索引模块中，与查询输入相关的索引子图会被LlamaIndex自动分割成处理单元块，并存储在向量数据库中。经过合适的重排算法，将较为相关的块输送给检索增强模块。上述索引模块的工作流程即为混合检索与重排，索引模块工作流程示意图如图4所示。

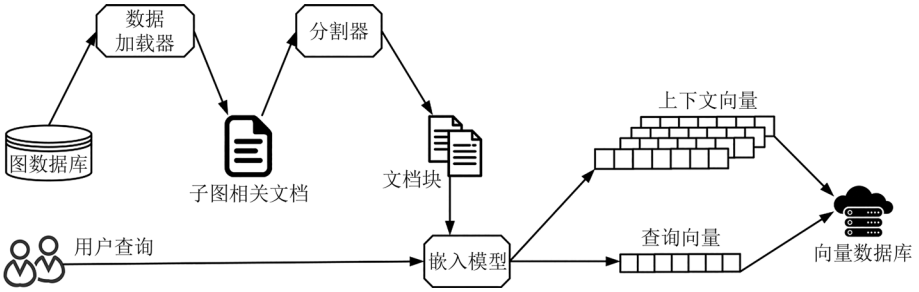


图3　嵌入模块工作流程

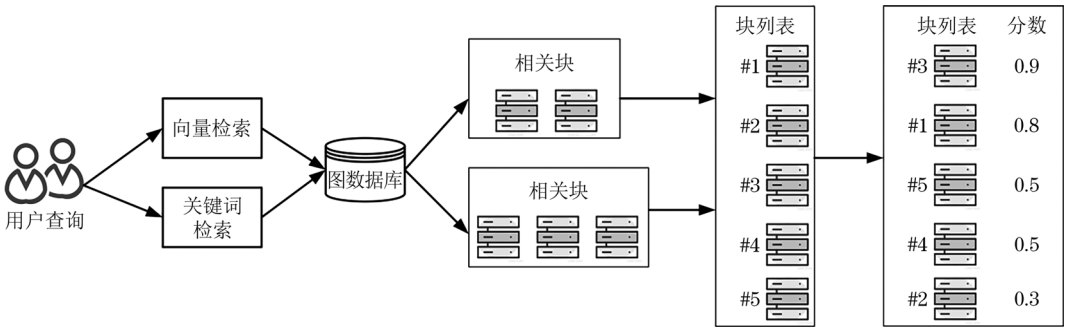


图4　索引模块工作流程

（3）检索增强模块。上述索引模块检索到的块列表被传递给LlamaIndex的响应生成器以生成最终输出。查询与索引得到的相关块将一起被封装为检索增强提示词，并输送给微调后的LLM。

检索增强提示词的结构一般包括任务描述、背景知识（检索得到上下文）、任务指令（一般是用户查询输入）等，检索增强模块提示词模板如图5所示。

（4）生成模块。微调LLM在检索增强提示词的引导下，完成科学知识检索的生成响应。同时生成模块提供对生成效果的评估，以衡量查询响应的准确性、忠实度和速度等核心性能。

{ 任务描述 }
假如你是一名中医药网络药理学领域专家，请参考{ 背景知识 }，按照{ 回答要求 }回答{ 问题 }
{ 背景知识 }　// 检索返回的上下文
< 五元组 1>、< 五元组 2>、< 五元组 3>、……
{ 回答要求 }
你需要严格根据背景知识的内容回答；
对于不知道的信息，直接回答“未找到相关答案”；
{ 任务指令 }
现有研究表明黄芪的适应证有哪些？

图5　检索增强模块提示词模板

生成模块的生成部分叫作生成器，有两种使用上下文文档的方式。①序列生成方式：使用同一个上下文文档生成每一个词，先确定一个文档，然后计算序列生成概率，如式（1）所示。②词元生成方式：使用不同的上下文文档生成每个词，对于第i个位置，候选词的概率等于所有文档的条件概率之和，即计算候选词在文档中的边际概率，如式（2）所示。

$$p_{\text{RAG-Sequence}}(y|x) \approx \sum_{z\in\text{top-}k(p(\cdot|x))} p_\eta(z|x)\,p_\theta(y|x,z) = \sum_{z\in\text{top-}k(p(\cdot|x))} p_\eta(z|x)\prod_i^N p_\theta(y_i|x,z,y_{1:i-1}) \tag{1}$$

$$p_{\text{RAG-Token}}(y|x) \approx \prod_i^N \sum_{z\in\text{top-}k(p(\cdot|x))} p_\eta(z|x)\,p_\theta(y_i|x,z,y_{1:i-1}) \tag{2}$$

式中：y是待生成的完整输出序列，x是输入查询文本，z是检索到的上下文文档，η是检索器参数，θ是生成器参数，top-$k\left(p(\cdot|x)\right)$是检索概率排名前k的文档，$p_\eta(z|x)$是检索器生成文档z的概率，$p_\theta(y_i|x,z,y_{1:i-1})$是基于文档$z$生成第$i$个词的概率，$N$是输出序列总词数，$y_i$是输出序列中第$i$个位置的目标词。

生成模块中需要完成整个知识检索问答模块的性能评估，包括检索评估和生成评估。检索评估指标同一般自然语言处理任务评估指标，包括平均倒数排名、命中率、精确率等。检索评估需要先给定一个非结构化的“查询问题+上下文”组合形式的语料集，然后测验问答模块所用子图检索算法在上述指标方面的性能表现，根据指标结果优化检索策略。LlamaIndex提供了开源的检索评估数据集LabelledRagDataset。生成评估用于验证生成文本是否满足科学知识检索问答的需求，是对知识检索和问答生成模块端到端的评估方式。LlamaIndex提供了基于LLM的评估模块llama_index.core.evaluation，用来衡量生成结果的质量，使用GPT-4作为基准来决定预测的答案是否正确。可调用的评估模块包括多种类型：①准确性模块利用带标签的训练语料来检验问答生成的准确性；②语义相似性模块同样基于带标签的训练语料，用于评估问答生成答案的语义相似程度；③忠实度模块专注于验证生成的答案是否忠实于检索到的上下文，即是否存在LLM幻觉现象；④内容相关度模块评估检索到的上下文与用户查询之间的相关性；⑤问答相关度模块用于判断生成的答案是否与查询问题相关。这些模块共同构成了全面的生成效果评估体系。

综上，构建了LLM与知识图谱协同增强知识组织的四层递进式理论架构。该架构通过LLM增强知识抽取、LLM增强实体对齐、知识图谱的适LLM存储处理以及图模协同检索增强4个核心模块，实现了从知识获取到知识应用的全流程协同优化。

3 理论框架的适用性与前景

3.1 理论框架的适用性边界

提出的图模协同增强知识组织理论架构具有明确的适用性边界。在知识密集型领域，如科学研究、医学诊断、法律咨询等场景中，该架构能够充分发挥结构化知识与生成式AI的协同优势。然而，在创意性、主观性较强的应用场景中，过度依赖知识图谱的约束可能限制LLM的创新能力。此外，该架构对于知识图谱的质量和完整性要求较高，在知识稀缺或者快速变化的新兴领域中，其效果可能会受到限制。

3.2 应用转化的关键挑战

随着AI技术的不断发展，图模协同增强知识组织理论架构具有广阔的发展前景。未来，其在多模态知识融合、个性化知识服务、跨语言知识组织等领域均有可观的应用潜能。然而，理论架构向实际应用转化面临着若干技术挑战：①知识图谱与LLM的语义对齐问题，需要跨越符号表示与向量表示之间的语义鸿沟；②知识动态更新的时效性问题，需要平衡系统稳定性与知识更新的时效性；③计算资源优化问题，大规模知识图谱与LLM的协同计算对硬件资源提出了更高的要求；④知识质量控制问题，需要进一步探索知识自动化抽取过程中的知识准确性和一致性保障技术。

4 结语

随着多模态融合、智能体等LLM新技术的发展，知识组织也迎来技术变革与场景拓展的新机遇。本文以知识组织为核心对图模协同增强方法进行了探究，提出了LLM与知识图谱协同增强知识组织的方法架构，通过图模协同技术实现了知识组织的双向增强，有望为LLM与知识图谱协同研究提供参考。

基于上述理论架构的设计原理和技术特点分析，LLM与知识图谱的协同应用在知识组织中具有以下理论优势：首先，图模协同工作扩展了信息来源，既可以处理文本数据又可以处理结构化知识图谱；其次，图模协同交互形式实现低代码化，LLM支持自然语言问答，降低了知识抽取技术门槛；再次，图模协同能有效提高知识组织的准确性和可解释性，知识图谱提供事实验证，减少了LLM的幻觉问题；最后，图模协同可维护性优异，知识图谱数据更新比LLM重新训练更加快捷。在未来研究中，将进一步拓展理论框架，并通过实证研究验证其在特定领域的实际应用效果，重点探索知识图谱与LLM在动态数据更新、跨领域融合等方面的协同作用。

参考文献

[1] 曹树金，曹茹烨，李睿婧. 数智时代的知识组织研究进展[J]. 情报学进展，2024，15：318-347.

[2] 杨欣谊，苏新宁. 领域知识结构认知：基于大数据环境的适用性分析[J]. 图书情报工作，2024，68（23）：4-16.

[3] 蔡祎然，胡正银，刘春江. 大语言模型赋能科技文献数据挖掘进展分析[J]. 农业图书情报学报，2025，37（2）：4-22.

[4] LIU Y，WAN Y，HE L，et al. KG-BART：knowledge graph-augmented BART for generative commonsense reasoning[J]. Proceedings of the AAAI Conference on Artificial Intelligence，2021，35（7）：6418-6425.

[5] YANG L Y，CHEN H Y，LI Z，et al. ChatGPT is not enough：enhancing large language models with knowledge graphs for fact-aware language modeling[EB/OL]. [2024-12-12]. https://arxiv.org/abs/2306.11489v1.

[6] CHEN Z C，CHEN J D，CHEN Y Y，et al. LMExplainer：a knowledge-enhanced explainer for language models[EB/OL]. [2025-02-04]. https://openreview.net/forum? id=VRJzlm2ecv.

[7] SU Y S，HAN X，ZHANG Z Y，et al. CokeBERT：Contextual knowledge selection and embedding towards enhanced pre-trained language models[J]. AI Open，2021，2：127-134.

[8] YU D H，ZHU C G，YANG Y M，et al. JAKET：joint pre-training of knowledge graph and language understanding[J]. Proceedings of the AAAI Conference on Artificial Intelligence，2022，36（10）：11630-11638.

[9] LE T，NGUYEN T，HO N，et al. LAMDA：label matching deep domain adaptation[C]//International Conference on Machine Learning. Cambridge：PMLR，2021：6043-6054.

[10] 孙利平，刘亮，彭胡萍. 知识图谱技术实现流程及相关应用[J]. 科技创新与应用，2024，14（18）：24-27.

[11] ZHANG Y，SHEN Y P，XIAO G，et al. Leveraging non-parametric reasoning with large language models for enhanced knowledge graph completion[J]. IEEE Access，2024，12：177012-177027.

[12] HUANG J C，ZHAO Y，HU W，et al. Trustworthy knowledge graph completion based on multi-sourced noisy data[C]//Proceedings of the ACM Web Conference 2022. New York：ACM Press，2022：956-965.

[13] PAN S R，LUO L H，WANG Y F，et al. Unifying large language models and knowledge graphs：a roadmap[J]. IEEE Transactions on Knowledge and Data Engineering，2024，36（7）：3580-3599.

[14] 史忠艳，雷洁，孙坦，等. DeepSeek赋能领域知识图谱低成本构建研究[J]. 农业图书情报学报，2025，37（3）：4-17.

[15] PENG H，ZHANG P F，TANG J Y，et al. Detect-then-resolve：enhancing knowledge graph conflict resolution with large language model[J]. Mathematics，2024，12（15）：2318.

[16] 吴信东，黄满宗，卜晨阳. BEKO：大语言模型与知识图谱的双向增强[J/OL]. 计算机学报：1-18[2025-03-25]. https://link.cnki.net/urlid/11.1826.TP.20250319.1054.010.

[17] WEN Y L，WANG Z F，SUN J M. Mindmap：knowledge graph prompting sparks graph of thoughts in large language models[EB/OL]. [2025-02-05]. https://arxiv.org/abs/2308.09729.

[18] NORLUND T，DOOSTMOHAMMADI E，JOHANSSON R，et al. On the generalization ability of retrieval-enhanced transformers[EB/OL]. [2025-02-05]. https://arxiv.org/abs/2302.12128.

[19] 陈子睿，王鑫，王晨旭，等. 面向时间感知的知识超图链接预测[J]. 软件学报，2023，34（10）：4533-4547.

作者简介

张心源，女，博士，副教授，研究方向：知识组织与数字治理。
肖浩宇，男，硕士研究生，研究方向：知识组织与数字治理。
李白杨，男，博士，助理教授，通信作者，研究方向：数据智能与数字素养，E-mail：libaiyang@nju.edu.cn。

Collaborative Enhancement of Knowledge Organization Theoretical Framework by Large Language Model and Knowledge Graph

ZHANG XinYuan[1,2] XIAO HaoYu[3] LI BaiYang[4]
(1. School of Sociology, Soochow University, Suzhou 215163, P. R. China; 2. Institute of Intelligent Society and Data Governance, Soochow University, Suzhou 215163, P. R. China; 3. School of Information Management, Zhengzhou University, Zhengzhou 450001, P. R. China; 4. Data Management Innovation Research Center, Nanjing University, Suzhou 215163, P. R. China)

Abstract: This study explores the collaborative enhancement model of large language model (LLM) and knowledge graph (KG) in knowledge organization, aiming to improve the accuracy and efficiency of knowledge organization. The research combines the natural language processing and knowledge generation capabilities of LLM with the structured expression and reasoning capabilities of KG, proposing a collaborative enhancement method for knowledge organization. First, we systematically review the theoretical and practical development of the existing KG applied to knowledge semantic organization, and sort out the applications of LLM in scientific knowledge extraction, entity alignment, and graph fusion. We demonstrate the optimization principle of KG storage and retrieval, and investigate typical projects of the collaborative application of LLM and KG. Then, based on research and induction, aiming at the specific operation links of knowledge organization, we explore the reasonable steps of embedding LLM, reconstruct the KG to realize the intelligent process of knowledge organization, and summarize the effective theoretical framework for enhancing knowledge organization through the collaboration of graph and model. The study concludes that the collaboration between graph and model can significantly improve the precision and explainability of knowledge organization, especially in cross-domain, multi-source, and heterogeneous knowledge processing. It effectively reduces the hallucination issues of LLM while enhancing the accuracy and interactivity of scientific knowledge retrieval and knowledge Q&A.

Keywords: Knowledge Organization; Large Language Model; Knowledge Graph; Collaborative Enhancement; Semantic Organization

（责任编辑：王玮）