山西大学学报(自然科学版)49(4):535—546,2026
Journal of Shanxi University(Nat. Sci. Ed.)

DOI:10.13451/j.sxu.ns.2025120

融合大语言模型与动态反馈机制的鸟类知识图谱补全方法

蒋永佳[1,2,3],耿生玲[1,2,3*],贾泽宇[1,2,3],李升宏[1,2,3]

(1.青海师范大学 计算机学院,青海 西宁 810008;
2.国家青藏高原科学数据中心青海分中心,青海 西宁 810008;
3.藏语智能信息处理及应用国家重点实验室,青海 西宁 810008)

摘 要:知识图谱作为一种高效的知识表示方式,能够将碎片化的信息转化成可推理的关系网络,在鸟类迁徙、保护和生物多样性研究等方面发挥重要作用。然而,由于鸟类知识的专业性、复杂性和实体关系多样化,现有鸟类知识图谱往往存在知识覆盖不全面、关系挖掘不深入等问题。为此,提出了一种基于大模型的鸟类知识图谱补全框架,通过微调和动态反馈机制,聚焦鸟类垂直领域,增强大模型对细粒度鸟类知识的理解和推理能力,实现对鸟类知识图谱的高效补全。首先,通过构建补全指令集对Qwen模型进行微调,增强其对鸟类知识的理解和补全能力。接着,通过引入动态反馈机制强化大模型对鸟类知识的深度挖掘和推理能力,增强鸟类知识图谱的补全生成。本方法在放宽评估下Hits@1达63.0%,相较于ChatGPT提升了50.76%;引入动态反馈机制后,指标提升至95.3%,相较于DeepSeek提升了3.1%。在此基础上,模型生成的正确答案将被扩展至图谱中,形成“微调-补全-反馈-扩展”的闭环架构,持续增强其在垂直领域的自我学习与补全能力。

关键词:领域知识图谱;指令微调;知识推理
中图分类号:TP391;Q959.7 **文献标志码**:A **文章编号**:0253-2395(2026)04-0535-12

Research on Bird Knowledge Graph Completion Integrating Large Language Models and Dynamic Feedback Mechanisms

Jiang Yongjia[1,2,3], Geng Shengling[1,2,3*], Jia Zeyu[1,2,3], Li Shenghong[1,2,3]

(1. *School of Computer Science, Qinghai Normal University, Xining* 810008*, China;*
2. *Qinghai Branch, National Tibetan Plateau & Third Pole Environment Data Center, Xining* 810008*, China;*
3. *The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining* 810008*, China*)

Abstract: Knowledge graphs have been widely used to transform fragmented information into structured and inferable relations, thereby supporting research on bird migration, conservation, and biodiversity. However, due to the high specialization and complexity of avian knowledge, as well as the diversity of entity relations, existing bird-domain knowledge graphs have remained limited in coverage and exhibited shallow relation extraction. To address these limitations, a bird-domain knowledge graph completion framework was proposed by integrating large language models with fine-tuning and a dynamic feedback mechanism. The framework was designed to enhance the understanding and reasoning of fine-grained avian knowledge for more efficient graph completion. Specifically, the Qwen model was fine-tuned using a domain-oriented instruction set to improve its specialized inference capability, and a

收稿日期:2025-10-22;**修回日期**:2025-11-19
基金项目:中央引导地方科技发展专项(2024ZY050);国家自然科学基金(6246070542);青海师范大学中青年科研基金资助项目(2023QZR016)
作者简介:蒋永佳(2000—),女,青海西宁人,硕士研究生,研究方向为智能信息处理与知识图谱。E-mail:2252504189@qq.com
* **通信作者**:耿生玲(Geng Shengling),E-mail:geng_sl@126.com
引文格式:蒋永佳,耿生玲,贾泽宇,等.融合大语言模型与动态反馈机制的鸟类知识图谱补全方法[J].山西大学学报(自然科学版),2026,49(4):535-546. DOI:10.13451/j.sxu.ns.2025120.

dynamic feedback mechanism was incorporated to further strengthen reasoning and information extraction. Under relaxed evaluation settings, the proposed method achieved a Hits@1 of 63.0%, representing a 50.76% improvement over ChatGPT. With feedback enhancement, the performance increased to 95.3%, surpassing DeepSeek by 3.1%. The correct model outputs were iteratively integrated back into the graph, forming a closed-loop process of fine-tuning, completion, feedback, and expansion, which continuously strengthened self-learning and vertical-domain completion ability.

Key words: domain-specific knowledge graph; instruction tuning; knowledge reasoning

0 引言

知识图谱[1]作为一种高效的结构化信息管理工具,可以系统地整合垂直领域内不同来源的异构数据,在鸟类迁徙、保护和多样性研究发挥了重要作用。在鸟类研究中,通过构建鸟类知识图谱,将碎片化的信息转化为可推理的关系网络,实现从数据堆积向生态智慧挖掘转变,为科研、保护及公共服务提供智能化支持。然而,由于鸟类知识的专业性、复杂性和实体关系的多样化,现阶段的鸟类知识图谱存在知识覆盖不全面、关系挖掘不深入、部分关键属性未被有效利用等问题。因此,知识图谱补全成为解决这些问题、增强图谱完整性的关键技术。

知识图谱补全(Knowledge Graph Completion, KGC)[2]是知识图谱研究领域的重点任务之一,其核心是通过预测新增的关系和实体来提高知识图谱的质量。传统的知识图谱补全方法[3]通常依赖低维表示学习实体和关系,推断缺失的三元组,但在处理涉及特定生态领域,且具有复杂关系的知识图谱时,存在一定的局限性。随着生成式大语言模型,如ChatGPT、DeepSeek在自然语言处理领域的突破,越来越多的研究开始将这些模型应用于知识图谱补全任务[4]。大语言模型(Large Language Model, LLM)凭借其强大的推理能力和丰富的背景知识,通过语言理解辅助知识图谱补全。虽然在通用场景下现有方法已取得一定进展,但迁移至鸟类生态领域中,因在该领域的先验知识相对匮乏,生成结果常出现不一致或信息缺失等问题。大模型时代,模型微调增强对领域知识理解的深度,反馈强化学习增加模型推理生成的准确性,这些为知识图谱补全带来新的发展机遇。

在此基础上,我们提出一种基于大模型动态反馈增强的知识图谱补全框架。该框架在经过低秩适应(Low-Rank Adaptation, LoRA)和指令微调后,根据知识图谱的结构特性,引入动态反馈机制(Dynamic Feedback Mechanism, DFM),确保微调后的模型在补全过程中能够实时反馈并修正补全内容。同时,正确的补全内容将被扩展到知识图谱中,形成一个闭环,进一步提升图谱的完整性和准确性。图1展示了融合低秩适应与动态反馈机制的鸟类知识图谱补全框架(LoRA and Dynamic Feedback Mechanism for Bird Knowledge Graph Completion, L-DFM-BKGC),涵盖了知识图谱的构建过程及后续的补全推理过程。

1 相关研究

1.1 鸟类知识图谱

当前互联网中关于鸟类的知识往往呈现出海量且碎片化的特点,传统数据库难以实现有效的语义关联和智能推理。知识图谱是2012年5月由Google公司提出的一种表示和组织知识的图形化数据结构,其本质上是一种语义网络,其中节点代表实体或概念,边表示实体或概念之间的各种语义关系[1]。知识图谱通过结构化数据的整合,能够揭示鸟类物种属性、生态关系及栖息地分布等复杂关联,在物种识别、生态监测、保护决策与科普教育等方面,存在巨大的应用潜力。

近年来,国内学者在鸟类知识图谱构建方面已取得初步成果。尹俊[5]通过本体建模与三层序列标注结构(Bidirectional Encoder Representations from Transformers-Bidirectional Long Short-Term Memory-Conditional Random Field, BERT-BiLSTM-CRF)模型联合抽取实体和关系构建了中国鸟类知识图谱联合抽取实体和关系,构建了面向领域知识的鸟类知识图谱(Bird Knowledge Graph for Domain Knowledge, BKG4DK),解决了数据整合难题;姚津[6]进一步拓展至亚洲鸟类,引入基于基于路径的平移嵌入模型(PTransE)的循环神

经网络模型(PTransE-based Recurrent Neural Network,PT-RNN)推理方法,增强图谱的深度与关联性;李忠伟等[7]聚焦迁徙要素构建相关图谱,并通过RoBERTa提升识别精度,量化要素关联。但这些研究均未涉及区域特异性数据,且缺乏对鸟类生理特征的深度刻画。国际研究方面,现有成果多集中于通用生态知识图谱,如生物多样性网络[8];或是基于计算机视觉的鸟类识别[9],如单阶段目标检测模型(You Only Look Once,YOLO)系列模型在物种检测中的应用[10];世界鸟类数据库Avibase[11]虽提供结构化分类数据,但缺乏语义关联与推理能力;Wu等[12]通用知识图谱虽涵盖部分鸟类实体,却未针对生态特征建模。当前针对鸟类生态知识的结构化建模仍处于起步阶段,特别是在形态特征的细粒度抽取与部分关键属性整合方面,仍存在显著空缺。本研究构建的鸟类知识图谱旨在填补该领域知识组织与语义推理的空白,为生物多样性保护提供精细化数据支撑,推动知识图谱由"通用架构"向"深度垂直"方向转型。

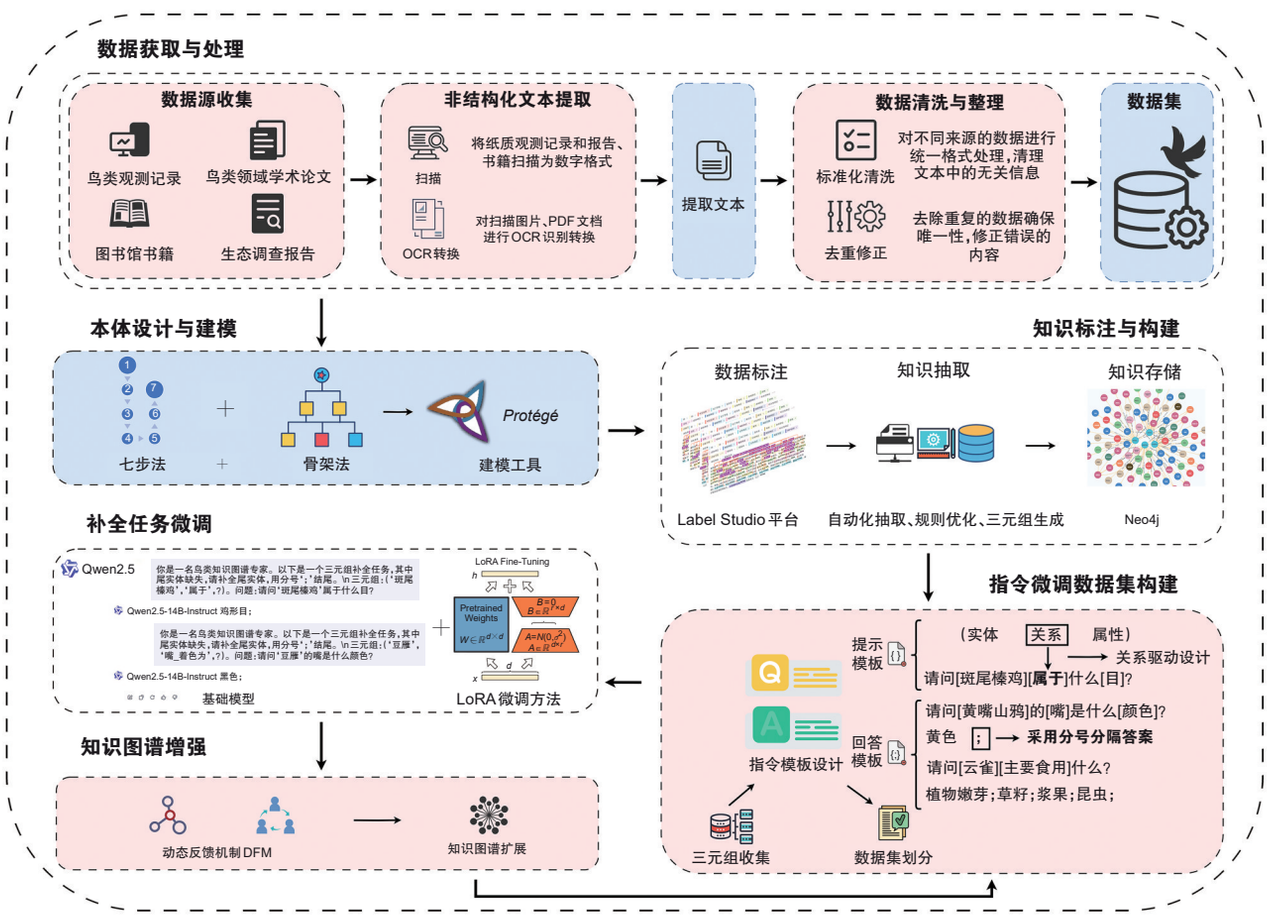


图1 L-DFM-BKGC整体框架图

Fig. 1 Overall framework of L-DFM-BKGC

1.2 知识图谱补全

知识图谱补全技术历经多阶段范式演进,其核心目标在于解决知识图谱固有的实体与关系缺失问题。早期研究依赖手工规则与统计推理方法[13],通过路径排序算法等人工定义规则进行关系推断,虽具备可解释性优势,但面临规则制定成本高、领域迁移性差的瓶颈。这一局限性推动了表示学习方法的兴起,以平移嵌入(Translation Embedding,TransE)等平移距离模型[14]通过向量空间几何约束建模关系语义,DistMult等语义匹配模型[15]则采用张量分解捕捉潜在关联。但这类嵌入模型在建模复杂语义交互与多跳推理任务时表现不足。随着深度学习技术的发展,图神经网络(Graph Neural Network,GNN)通过多层信息传播机制实现了结构化特征的有效聚合[16],

而预训练语言模型(Pre-trained Language Model, PLM)的引入开创了文本与结构联合建模的新范式[17],基于双向编码器表征转换器的知识图谱补全模型(BERT-based Model for Knowledge Graph Completion, KG-BERT)通过将三元组编码为自然语言序列的创新实践[18]取得了一定进展,但受限于模型容量,在处理复杂推理任务时仍显不足。大语言模型LLM的快速发展,推动了知识图谱补全方法的范式转变[4]。凭借其强大的语言理解与生成能力,LLM可以更好地结合知识图谱中的文本与结构信息,从而进行更精准的补全。知识图谱大语言模型(Knowledge Graph Large Language Model, KG-LLM)通过指令微调实现三元组建模[19],但现有方法普遍依赖单向推理流程,缺乏对预测结果的动态验证机制,导致错误传播风险累积。尽管Zhu等[20]验证了LLM在零样本补全中的潜力,但直接应用仍面临幻觉问题与领域适配性挑战。

2 方法

2.1 鸟类知识图谱构建

目前,网络上关于鸟类的数据虽体量庞大,但往往零散且冗余,难以支撑精细化的研究与应用需求。与此同时,鸟类生态知识具有高维度语义依赖与实体多样性特征,使得知识图谱构建面临知识碎片化、语义歧义与推理不确定性等多重挑战。因此,本研究从鸟类观测记录、学术论文、图书馆馆藏书籍、鸟类保护机构及生态调查报告等多种资料中,收集并整理了三江源、青海湖与祁连山等典型生态区域的鸟类相关数据,包括栖息地、羽毛特征、眼部特征等信息。数据通过光学字符识别(Optical Character Recognition, OCR)技术提取出有效信息,并结合数据清洗、标准化与去重等步骤,形成一个结构化的鸟类文本数据集,作为构建知识图谱的原始数据[21]。通用的知识图谱本体构建方法,有骨架法、本体方法体系(Toronto Virtual Enterprise, TOVE)法、本体工程方法论(Methontology)和基于叙词表的领域本体构建法等[22]。鉴于骨架法自顶向下的清晰层次、易于扩展的知识结构,和七步法规范的系统流程、成熟的建模方案,本文结合两者优势进行本体构建,以此高效支撑鸟类垂直领域的高质量图谱构建需求。根据鸟类学和生态学的特点,本文将本体中的实体分为核心类和细节类,核心类代表鸟类物种的基本分类和生物学特征,细节类描述鸟类的外观特征与具体的生态行为。Protégé支持本体的图形化编辑,便于构建本体结构并建模实体间复杂关系[23]。通过Protégé,可定义类、属性、关系和约束,并以Web本体语言(Web Ontology Language, OWL)格式输出本体文件。

图谱标注工作采用Label Studio平台作为标注工具,通过在本体构建阶段预先定义的16个实体、7个属性和11种关系,共计标注了约36 000个节点和18 000条关系。Label Studio导出的标注语料数据(JSON格式)通过预定义规则对实体、属性及关系进行知识抽取,具体依据实体标识符(Identifier, ID)在已标注数据中进行关系匹配,结合自动化脚本完成。所有抽取的信息将以三元组的形式组织,并导入Neo4j知识图谱数据库,如图2所示。以上内容对应于图1中的知识图谱构建部分,为后续补全任务奠定基础。

2.2 指令微调数据集构建

知识图谱补全的核心目标是基于已知(h,r,t)三元组推断缺失实体。针对传统三元组直接输入引发的语义歧义问题,本文采用上文构建的鸟类知识图谱作为数据来源,构建一个基于自然语言提示,用于知识补全的指令微调数据集,具体构建流程如图3所示,指令微调数据集中Prompt与Answer的生成包括以下步骤:

1)语义融合重构。原始三元组在结构上较为离散化,缺乏可用于构造问题的上下文表达,语义片段之间无法直接支持有效的指令生成。为构建更适配提示设计的三元组集合,提出语义融合重构。设原始三元组集合为$T=\{(h_i,r_i^*,t_i)\}_{i=1}^{n}$,其中$h_i\in H$、$r_i\in R$、$t_i\in T$。定义实体桥接集为:

$$E=\{e\in\varepsilon\mid\exists(h,r_1,e),(e,r_2,t)\in T\}, \quad (1)$$

对于满足桥接条件的路径$(h,r_1,e),(e,r_2,t)$,构造新的三元组:$(h,r_1,e),(e,r_2,t)\Rightarrow(h,r^*,t)$,其中合成关系$r^*=f_{\text{merge}}(r_1,e,r_2)$,通过融合$r_1$和$r_2$为复合语义关

系,同时引入中介实体e,生成具备明确语境指向的新关系。融合后的三元组集合记为$\mathrm{T}'=\{(h_j,r_j^*,t_j)\}_{j=1}^{m}$,该集合作为后续模板生成与指令构造的基础数据。

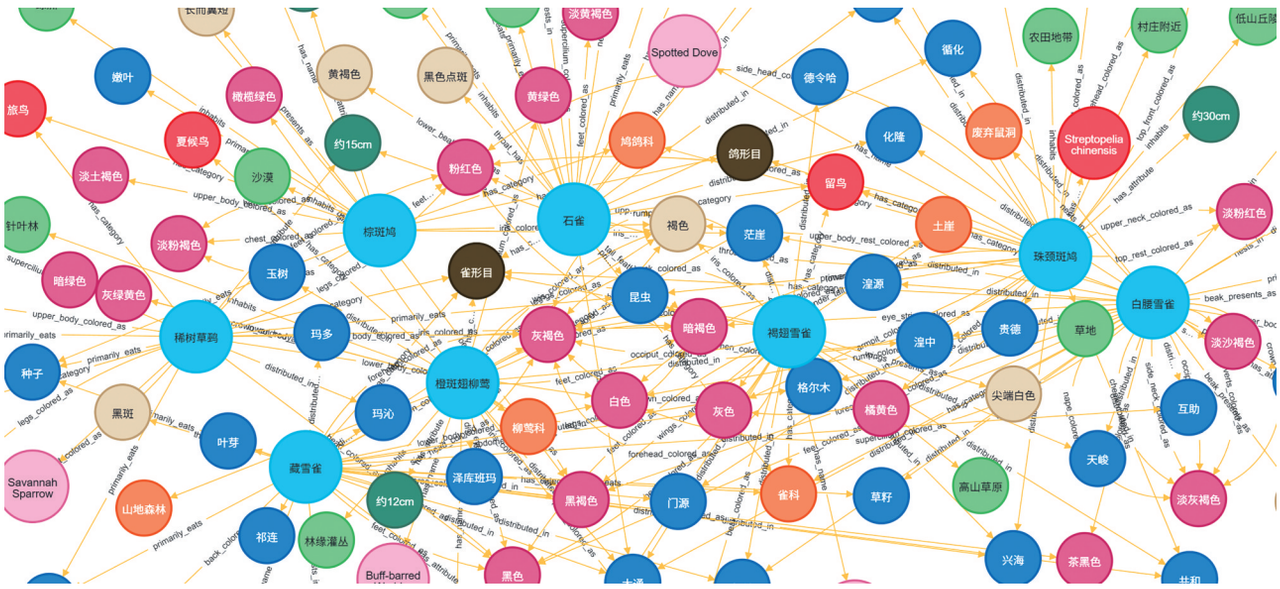


图 2 Neo4j 数据库可视化

Fig. 2 Neo4j database visualization

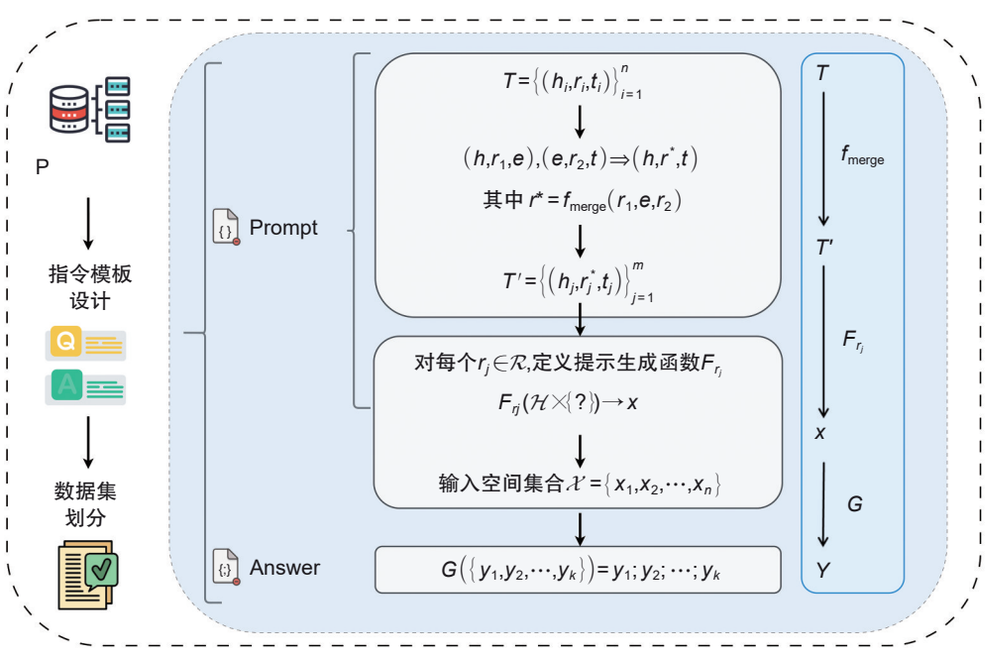


图 3 指令微调数据集构建流程图

Fig. 3 Workflow of instruction-tuning dataset construction

2)关系驱动生成。指令样本构造围绕语义关系展开,每种关系对应设计特定的语义生成模板。融合后的三元组集合包含若干关系类型,记为R。针对每一关系$r_j\in R$,定义其对应的语言模板映射函数F_{r_j},组成关系感知型函数族:

$$F=\left\{F_{r_j}\middle|F_{r_j}:H\times\{?\}\to X,r_j\in R\right\},\qquad(2)$$

其中H表示实体空间,X表示语言输入空间。函数F_{r_j}接收结构化片段并生成语言输入$x=F_{r_j}(\cdot)$,用于构造训练样本中的指令部分,实现结构语义与语言表达之间的对齐映射。

3)输出结构控制。在实际补全任务中,部分关系对应多个实体,即$(\{h,r,\{t_1,t_2,\cdots,t_k\}\})\in T'$或

者($\{h_1,h_2,\cdots,h_k\},r,t)\in T'$的结构形式。为统一建模与评估,定义输出线性化编码函G,对实体集合进行线性化处理:

$$G(\{y_1,y_2,\cdots,y_k\})=y_1;\ y_2;\ \cdots;\ y_k。\quad(3)$$

该编码方式以分号作为边界标识符,明确划分答案单元,增强结构分离性,支持对多目标样本的统一处理。

最终构建的数据集覆盖三江源、青海湖等典型生态地区432种鸟类实体,包含11 333个自然语言问题,其中38.7%的问题具有多答案特性,平均每个问题关联3.6个候选答案,如图4所示。该数据集在保持结构化知识准确性的基础上,实现向自然语言表达的有效转化,并通过分号分隔统一表示多重答案结构。构建的鸟类生态细粒度补全数据集为大模型提供更高质量的训练数据,有助于促进垂直领域知识的理解与推理。

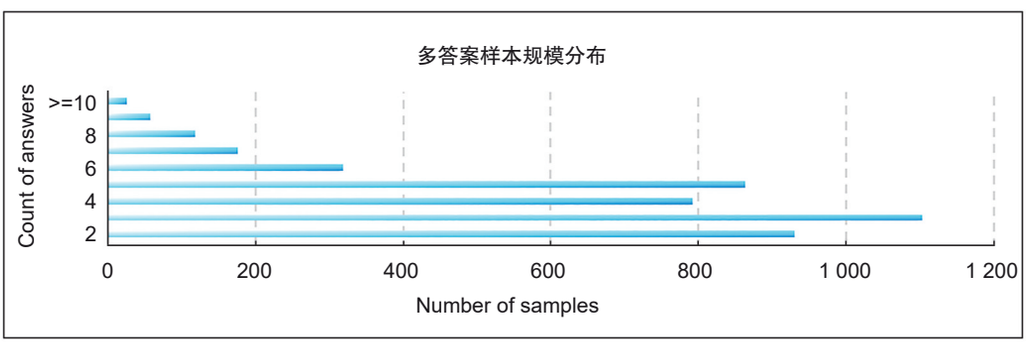


图4 多答案样本规模分布图

Fig. 4 Distribution chart of multi-answer sample sizes

2.3 动态反馈机制

当前知识图谱补全任务中,常通过引入外部知识实现大语言模型增强。检索增强生成(Retrieval-Augmented Generation, RAG)[24]通过从外部知识库中检索相关信息,将其作为输入提供给LLM增强生成,但多采用静态知识传递模式,仅单向输出预定义知识结构,且通常需遍历整篇文档,存在检索效率低与知识利用率不足等问题。因此,本文在大模型微调基础上,提出一种结构更简、响应更快的轻量级补全过程增强机制,可以在多层次语义交互中实现知识的递进式优化。图5呈现了推理路径的可视化实例,以形象揭示模型在多层级反馈阶段的知识流转机制与语义演化过程。

动态反馈机制DFM通过递进式反馈架构$\{F_1,F_2,F_3\}$增强生成模型的外部知识支持能力。其中每一层反馈F_k表示特定粒度和置信度的知识输入层,用于动态增强模型的生成过程。在每一层中,系统基于主语-谓语对从知识图谱中抽取相关对象,构成候选对象集合O_k,并引入一定比例的噪声对象O_{noise},其采样自预设的噪声分布P_{noise},以模拟真实语境中的信息不全与认知干扰,检验模型的鲁棒性与容错能力。模型当前输出与目标答案之间的语义相似度由度量函数$S(y_{\text{p}},y_{\text{t}})$计算,其中$y_{\text{p}}$表示模型当前输出,$y_{\text{t}}$表示目标答案。基于该相似度,采用启发式控制策略(Heuristic Control Strategy),动态地选择反馈层级k,以调整知识密度。当相似度超过设定阈值τ时,触发早停策略(Early Stopping),节省推理资源。DFM包含三个递进层级,各层级对应不同的知识传递策略:

1)低反馈层F_1。随机返回与给定主语和谓语相关的一个对象O_1,并可能包含噪声对象O_{noise},构成最小化反馈集$O_1'=O_1\cup O_{\text{noise}}$,模拟信息不足或上下文模糊的真实情境,测试模型在初始信息极其有限情况下的推理能力。

2)中反馈层F_2。返回与给定主语和谓语相关的所有对象集合$O_2=\{o_1,o_2,\cdots,o_n\}$,随机引入一定数量的噪声对象$O_{\text{noise}}$,构成更丰富但含干扰的上下文反馈$O_2'=O_2\cup O_{\text{noise}}$,考察模型在上下文信息充足的基础上,区分目标信息与噪声的能力。

3)高反馈层F_3。返回与给定主语和谓语相关的所有对象集合O_3,且不引入任何噪声$O_3'=O_3$,模拟信息充分、上下文完整的理想语境,测试模型在高可信背景下的语义整合与精细推理能力。

DFM通过多层次设计使大模型在生成过程中灵活应对不同复杂度问题,融合传统知识图谱的结构化优势与动态反馈的灵活性,通过优化反馈粒度,避免反馈信息过简或过繁的负面影响,从而提升在复杂任务中知识的利用效率与生成质量。

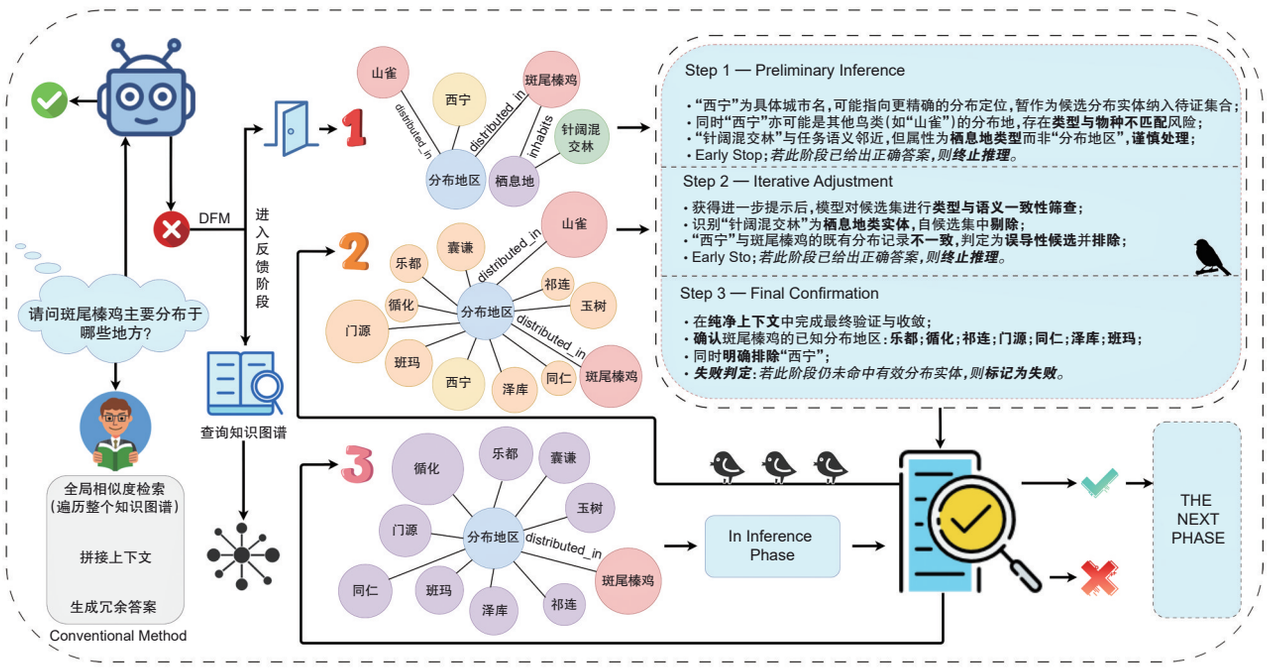


图5 推理路径可视化

Fig. 5 Visualization of reasoning paths

2.4 知识图谱扩展

以往的知识图谱补全方法[14]通常在完成推理任务后,并未进一步利用补全后的知识更新或扩展图谱,导致图谱更新滞后。而图谱补全的本质在于结构性知识的增量扩展,将生成结果转化为可写回的三元组结构,图6展示了动态反馈与图谱扩展两个关键环节,分别用于在生成阶段引导大模型优化补全结果与持续完善图谱结构信息。

基于统一格式的指令构造,输出内容具备可映射的结构特性,能够在实体层级对应输入中的结构缺口。补全任务可视为不完全结构的语义闭合操作,输入通常对应三元组模式中的缺失形式$(h,r,?)$或$(?,r,t)$。基于该结构残缺,模型生成候选实体集合$\{\hat{e}_i,\cdots,\hat{e}_k\}$,经位置绑定后重构为形式一致的三元组集$\{(h,r,\hat{t}_i)\}$或$\{(\hat{h}_i,r,t)\}$。虽然生成结果在结构维度已具有闭合性,但语义层面仍存在不一致映射的可能,需引入规约范式以满足图谱扩展约束。设生成三元组集合为T_{gen},原始图谱为T_{ori}。本体建模阶段已预设实体集E、关系集R与类型类别集合C。类型约束函数为$\tau: R \to C$,规定关系的语义域所接纳的实体类别$\phi: R \to \{\text{functional}, \text{set-valued}\}$为结构范式映射,表示关系在实体绑定上的值域约束形式。

冗余项由$T_{\text{gen}} \cap T_{\text{ori}}$剔除;类型不匹配项依据$\tau(r)$规定的实体类别过滤;若某关系$r$被映射为functional,则在该绑定主体已存在对应记录时,仅允许置信度最高者写入或执行策略性替换,减少最后大量冲突检测带来的处理成本。对于在放宽评估中被判定为语义正确、但未与图谱命名空间实体严格一致的候选项$\hat{e}$,引入归一化函数$f_{\text{norm}}: \hat{\varepsilon} \to \varepsilon \cup \perp$,用于在写回阶段实现命名对齐与语义对映,确保语义表述符合标准的表示形式,其定义如下:

$$f_{\text{norm}}(\hat{e}) = \begin{cases} e^*, \text{if } \text{sim}(\hat{e}, e^*) \geqslant \delta, e^* \in \varepsilon \\ \perp, \text{otherwise} \end{cases}, \quad (4)$$

其中归一失败的候选项不可写入,对于已通过归一与类型判定的项,若其与图谱中已有三元组存在逻辑或语义层面的潜在冲突,写入策略由融合函数决定,其定义如下:

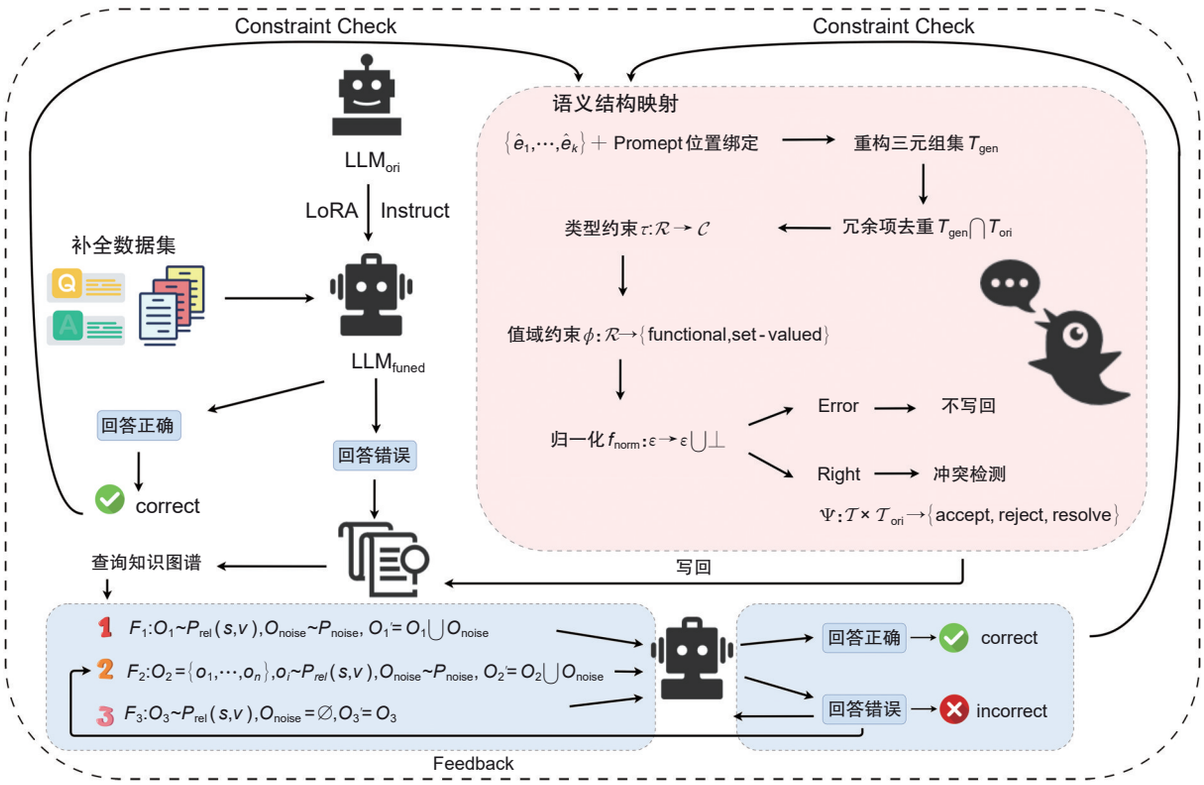


图 6 动态反馈及知识图谱扩展整体流程图

Fig. 6 Overall workflow of dynamic feedback and knowledge graph expansion

$$\Psi((h,r,t),T_{ori})=\begin{cases}\text{accept}, & \text{if}(h,r,\cdot)\notin T_{ori}\\ \text{reject}, & \text{if}(h,r,t')\in T_{ori}\wedge t\neq t'\wedge\neg\,\text{conflict}(t,t')\\ \text{resolve}, & \text{if conflict}(t,t')=\text{true}\end{cases}, \tag{5}$$

最终接入图谱的补全集 $T_{final}\subseteq T_{gen}$ 在经过这些规则约束和筛选后，能够满足本体定义、结构一致性及命名空间的统一性，进而作为图谱更新的扩展集。

2.5 评价指标

为全面评估模型在知识图谱补全推理任务中的表现，采用严格评估和放宽评估两种判定标准。Hits@1(Top-1 Hit Rate)用于衡量模型生成的首个预测答案是否与参考答案完全一致，是最常用的严格匹配指标[14]。

严格评估 Hits@1(Strict)要求模型输出在实体内容、数量和顺序上与参考答案完全一致，其计算公式见式(6)：

$$\text{Hits@1}_{\text{Strict}}=\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left(\hat{A}_i=A_i\right), \tag{6}$$

其中 N 表示样本总数，$\hat{A}_i$ 表示模型生成的第 i 个预测答案，A_i 为参考答案，$\mathbb{I}(\cdot)$ 为指示函数，当括号内条件成立时取值为1，否则为0。

放宽评估 Hits@1(Relaxed)适用于存在模糊表达、语言多样性或参考答案非唯一的复杂场景。放宽评估不要求模型输出与参考答案完全一致，而是基于以下任一条件判断命中：

1) 预测答案中的实体集合完整覆盖真实答案中的所有尾实体；

2) 预测答案中至少一项与真实实体在语义上具有较高相似度。实体覆盖保证答案内容的完整性，语义相似容纳语言表达的多样性，两者互为补充。满足上述任一条件即视为命中，其计算公式见式(7)：

$$\text{Hits@1}_{\text{Relaxed}}=\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left(\text{Coverage}(\hat{A}_i,A_i)\geqslant\theta\vee\text{Sim}(\hat{A}_i,A_i)\geqslant\tau\right), \tag{7}$$

其中 $\text{Coverage}\left(\hat{A}_i,A_i\right)$ 表示预测答案中实体集合对参考答案实体集合的覆盖比例，$\text{Sim}\left(\hat{A}_i,A_i\right)$ 表示两者之间的最大语义相似度；θ 和 τ 分别为覆盖率和相似度的判定阈值，本研究中设置为 $\theta=1.0$，$\tau=0.7$。

通过结合严格评估和放宽评估,能够更全面地衡量模型的推理能力,严格评估保证公平性,放宽评估考虑语义相似或部分匹配等情况,更贴近真实应用中的容错需求,更充分地反映模型的实际表现。

3 实验及结果分析

3.1 实验设置

本实验使用文章2.2节构建的鸟类知识图谱补全数据集,将其划分为70%的训练集、15%的验证集、15%的测试集。实验评估指标包括Hits@1(Strict)和Hits@1(Relaxed),用于衡量不同模型在知识图谱补全任务中的表现。Qwen2.5-14B-Instruct是本研究选用微调的大语言模型,主要基于其在指令优化、推理能力、参数规模适中以及多语言支持方面的显著优势[25]。针对垂直领域鸟类文本数据的独特性,本文采用LoRA方法[26]对Qwen2.5-14B-Instruct模型进行高效微调,其中LoRA秩r为8,缩放系数α为16。使用AdamW优化器,采用余弦退火(Cosine Schedule)策略,并将初始学习率设为5×10^{-5}。训练过程中,每个设备上的batch size批大小为2,梯度累计步数设为8。

3.2 对比实验

鉴于当前知识图谱补全任务中,大语言模型已成为常用方案,本文选取了多个维度上的代表性模型进行对比实验,以全面评估所提方法的性能。实验涵盖了经典知识图谱补全方法TransE[14]、国外主流大模型ChatGPT4[4]、国内优秀模型ChatGLM-4-9B[27],以及当前高关注度的新兴模型DeepSeek-R1-Distill-Qwen2.5-14B[28]。同时,针对经过LoRA微调后的大模型Qwen2.5-14B-Instruct[25]结合DFM进行对比实验,以验证所提方法的有效性。对比实验结果如表1所示。

表1 对比实验结果

Table 1 Results of comparative study

模型	Hits@1 (Strict)	Hits@1 (Relaxed)
TranE	0.003 3	0.088 4
ChatGPT4	0.106 4	0.122 4
ChatGLM-4-9B	—	0.140 0
ChatGLM-4-9B-LoRA-Finetuned	0.294 0	0.677 0
ChatGLM-4-9B-LoRA-Finetuned-DFM	0.680 0	0.824 0
DeepSeek-R1/Qwen2.5-Dis(14B)	—	0.080 0
DeepSeek-R1/Qwen2.5-Dis(14B)-LoRA-Finetuned	0.270 0	0.666 0
DeepSeek-R1/Qwen2.5-Dis(14B)-LoRA-Finetuned-DFM	0.825 0	0.922 0
Qwen2.5-14B-Instruct-LoRA-Finetuned-DFM	0.895 0	0.953 0

注:—表示该设置下模型输出未能形成可用于相应指标计算的有效答案,故未纳入统计。

传统的TransE模型经过3 000轮训练后,依然表现不佳,尤其在Hits@1排名指标上,正确答案的排名通常位于1 000名以后,因此其Hits@1(Strict)和Hits@1(Relaxed)的结果均不理想。相比之下,ChatGPT由于其强大的理解能力,在快速明确任务后,能够生成格式接近正确答案的回复,因此其Hits@1(Strict)和Hits@1(Relaxed)的表现相差不大。但在面对鸟类知识图谱这一细粒度生态领域数据时,ChatGPT的Hits@1(Relaxed)仅为12.24%,表明其在处理该特定任务时存在一定的局限性。微调后的ChatGLM和DeepSeek的Hits@1(Strict)和Hits@1(Relaxed)分别提升了29.4%、27.0%、53.7%、58.6%,可见LoRA微调对大模型知识图谱补全任务有显著的提升。引入DFM后,本研究选用的Qwen-14B表现较ChatGLM和DeepSeek-R1的Hits@1(Strict)和Hits@1(Relaxed)分别提升了21.5%、12.9%、7.0%、3.1%。对比实验结果表明,LoRA微调和DFM对鸟类知识图谱补全任务的性能有显著的提升。

3.3 消融实验

在模型的设计和优化过程中,为进一步评估所提模型中各组件的有效性,设计了消融实验。实验将Qwen2.5-14B-Instruct作为基线模型,并以此为对照建立四个实验版本:A模型为原始模型;B模型为LoRA微调后的版本;C模型为在B模型

的基础上加入DFM;D模型则是在C模型基础上加入扩展集。消融实验结果如表2所示。

表2 消融实验结果

Table 2 Results of ablation study

模型	Hits@1 (Strict)	Hits@1 (Relaxed)
A	—	0.040 0
B	0.283 0	0.630 0
C	0.895 0	0.953 0
D	0.900 0	0.970 0

注:—表示该设置下模型输出未能形成可用于相应指标计算的有效答案,故未纳入统计。

由于大模型输出具有较强的语言多样性,未经微调的模型在生成答案时常出现冗余内容,与正确答案格式存在明显差异,难以解析为可用于严格匹配评测的有效答案。因此,Qwen、DeepSeek-R1和ChatGLM在未经微调设置下的相关Hits@1(Strict)指标未纳入统计,并在对比实验和消融实验的表格中以"—"表示。但即便在放宽评估下,原始模型的表现仍较差,仅为4.0%。经微调后,模型在格式和内容上得到了显著改进,能够更好地生成符合任务要求的答案,Hits@1(Strict)达到了28.3%,Hits@1(Relaxed)达到了63.0%。引入DFM后,Hits@1(Strict)提升了61.2%,Hits@1(Relaxed)提升了32.3%。这一结果表明,LoRA微调和DFM对于提高模型的回答准确性和任务适应性具有重要作用。并且,由于DFM层级之间并非独立,采用结果分布统计分析生成结果停留比例。如图7所示,按级别由低到高,比例依次为49.8%、11.0%、39.2%。结果表明,近一半的任务在初级反馈阶段即可得到有效输出,显著提升了模型的准确性与响应效率。最后,为了验证扩展的有效性,本研究新增了10%的知识图谱内容,该部分内容在对比实验和消融实验中均隐藏。实验结果表明,利用扩展后的知识图谱作为反馈信息源,在相同测试集上的实验显示,Hits@1(Strict)提升了0.5%,Hits@1(Relaxed)提升了1.7%。总之,在扩展后的充足知识背景下,利用LoRA微调和DFM能够有效增强大模型知识图谱补全任务的性能。

4 结语

本研究通过收集鸟类领域的相关数据,采用七步法与骨架法相结合的建模方法,使用Protégé工具设计本体建模。通过Label Studio平台完成数据标注,结合自动化抽取与规则优化生成知识三元组,最终将数据存储在Neo4j图数据库中。针对补全任务构建指令微调数据集,提出L-DFM-BKGC框架,对Qwen2.5-14B-Instruct模型通过LoRA微调方法进行优化,引入DFM以增强补全能力。实验结果表明,本研究所提的方法使得大模型在知识图谱补全任务中取得良好的效果。未来可进一步引入显式反馈权重建模、面向链式推理的过程设计,以及视觉问答技术[29],以增强模型的可解释性、推理追溯性与跨模态交互能力;同时将在更多领域的知识图谱中验证方法的可迁移性,并在更复杂的噪声环境下开展鲁棒性评估,从而推动本研究方法在更广泛场景下的应用。

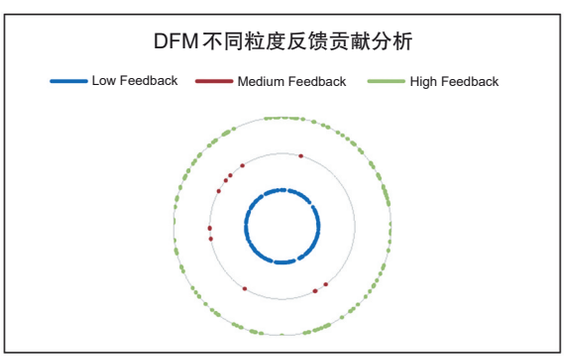


图7 DFM不同粒度反馈贡献分析图

Fig. 7 DFM feedback contribution analysis at different granularities

参考文献:

[1] Hogan A, Blomqvist E, Cochez M, *et al*. Knowledge Graphs[J]. *ACM Comput Surv*, 2022, **54**(4): 1-37. DOI: 10.1145/3447772.

[2] Chen Z, Wang Y H, Zhao B, *et al*. Knowledge Graph Completion: A Review[J]. *IEEE Access*, 2020, **8**: 192435-192456. DOI: 10.1109/access.2020.30300076.

[3] 张文豪, 徐贞顺, 刘纳, 等. 知识图谱补全方法研究综述[J]. 计算机工程与应用, 2024, **60**(12): 61-73. DOI: 10.3778/j.issn.1002-8331.2401-0380.
Zhang W H, Xu Z S, Liu N, *et al*. Overview of Knowledge Graph Completion Methods[J]. *Comput Eng Appl*, 2024, **60**(12): 61-73. DOI: 10.3778/j.issn.1002-8331.2401-0380.

[4] Achiam J, Adler S, Agarwal S, *et al*. GPT-4 Technical Report[EB/OL]. (2023-03-15) [2025-3-10]. https://arxiv.org/abs/2303.08774.

[5] 尹俊. 中国鸟类领域知识图谱构建与应用研究

[D]. 南昌: 江西财经大学, 2022. DOI: 10.27175/d.cnki.gjxcu.2022.001130.
Yin J. Construction and Application of Knowledge Graph in Chinese Avian Domain[D]. Nanchang: Jiangxi University of Finance and Economics, 2022. DOI: 10.27175/d.cnki.gjxcu.2022.001130.

[6] 姚津. 面向亚洲鸟类知识图谱的本体构建和推理研究[D]. 南昌: 江西财经大学, 2023. DOI: 10.27175/d.cnki.gjxcu.2023.000640.
Yao J. Research on Ontology Construction and Reasoning for Asian Bird Knowledge Atlas[D]. Nanchang: Jiangxi University of Finance and Economics, 2023. DOI: 10.27175/d.cnki.gjxcu.2023.000640.

[7] 李忠伟, 李明轩, 李永, 等. 基于鸟类迁徙关联要素数据的知识图谱构建[J]. 计算机工程与设计, 2024, **45**(12): 3704−3711. DOI: 10.16208/j.issn1000-7024.2024.12.024.
Li Z W, Li M X, Li Y, *et al*. Construction of Knowledge Graph Based on Bird Migration Associated Factor Data [J]. *Comput Eng Des*, 2024, **45**(12): 3704−3711. DOI: 10.16208/j.issn1000-7024.2024.12.024.

[8] Page R D M. Ozymandias: A Biodiversity Knowledge Graph [J]. *PeerJ*, 2019, **7**: e6739. DOI: 10.7717/peerj.6739.

[9] Tan M, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[C]//Proceedings of the 36th International Conference on Machine Learning (ICML 2019). Long Beach: PMLR, **97**, 2019: 6105-6114.

[10] Yang H, Feng Q, Zouyi L Y, *et al*. Artificial Intelligence in Aquatic Biology: Identifying and Conserving Aquatic Species[J]. *Water Ecol*, 2025, **1**(2): 100008. DOI: 10.1016/j.wateco.2025.100008.

[11] Lepage D. Avibase: The World Bird Database[DB/OL]. (2003−06−24)[2025−3−20]. https://avibase.bsc−eoc.org/.

[12] Wu W T, Li H S, Wang H X, *et al*. Probase: A Probabilistic Taxonomy for Text Understanding[C]//Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. Scottsdale: ACM, 2012: 481−492.. DOI: 10.1145/2213836.2213891.

[13] Lao N, Cohen W W. Relational Retrieval Using a Combination Of Path-constrained Random Walks[J]. *Mach Learn*, 2010, **81**(1): 53−67. DOI: 10.1007/s10994-010-5205-8.

[14] Bordes A, Usunier N, Garcia-duran A, *et al*. Translating Embeddings for Modeling Multi-relational Data[C]// Advances in Neural Information Processing Systems 26 (NeurIPS 2013). Red Hook: Curran Associates Inc., 2013: 2787−2795.

[15] Zhong J, Zhao H C, Zhao Q C, *et al*. A Knowledge Graph-based Method for Drug-drug Interaction Prediction with Contrastive Learning[J]. *IEEE/ACM Trans Comput Biol Bioinform*, 2024, **21**(6): 2485−2495. DOI: 10.1109/TCBB.2024.3477410.

[16] Arora S. A Survey on Graph Neural Networks for Knowledge Graph Completion[PP/OL]. (2020−07−24)[2025−3−21]. https://arxiv.org/abs/2007.12374.

[17] Hu A, Kuang L N, Yang D H. LPBERT: A Protein-Protein Interaction Prediction Method Based on a Pre-trained Language Model[J]. *applied Sci*, 2025, **15**(6): 3283. DOI: 10.3390/app15063283.

[18] 黄勃, 吴申奥, 王文广, 等. 图模互补: 知识图谱与大模型融合综述[J]. 武汉大学学报(理学版), 2024, **70**(4): 397−412. DOI: 10.14188/j.1671-8836.2024.0040.
Huang B, Wu S A, Wang W G, *et al*. KG-LLM-MCom: a Survey on Integration of Knowledge Graph and Large Language Model[J]. *J Wuhan Univ Nat Sci Ed*, 2024, **70**(4): 397−412. DOI: 10.14188/j.1671-8836.2024.0040.

[19] Zhang L, Wang W, Jiang Z X, *et al*. Exploring Knowledge from Knowledge Graphs and Large Language Models for Link Prediction[C]//2024 IEEE International Conference on Knowledge Graph (ICKG). Washington: IEEE, 2024: 485−491. DOI: 10.1109/ICKG63256.2024.00068.

[20] Zhu Y Q, Wang X H, Chen J, *et al*. LLMS for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities[J]. *World Wide Web*, 2024, **27**(5): 58. DOI: 10.1007/s11280-024-01297-w.

[21] 杨胜埕, 李榜明, 田仁祥. 基于OCR与NLP技术的ETC发行服务流程优化探索[J]. 中国交通信息化, 2025(2): 16−23. DOI: 10.13439/j.cnki.itsc.2025.02.001.
Yang S C, Li B M, Tian R X. Exploration on Optimization of ETC Publishing Service Process Based on OCR and NLP Technology[J]. *China ITS J*, 2025(2): 16−23. DOI: 10.13439/j.cnki.itsc.2025.02.001.

[22] 赵婉亦. 多模态唐卡知识图谱构建方法研究[D]. 西宁: 青海师范大学, 2024. DOI: 10.27778/d. cnki.gqhzy.2024.000724.
Zhao W Y. Research on the Construction Method of Multi-modal Tangka Knowledge Graph[D]. Xining: Qinghai Normal University, 2024. DOI: 10.27778/d.cnki.gqhzy.2024.000724.

[23] 郭英英. 基于Protégé的青花瓷本体构建方法研究[J]. 陶瓷科学与艺术, 2023, **57**(11): 39−41. DOI: 10.13212/j.cnki.csa.2023.11.130.
Guo Y Y. Research on the Construction Method of Blue and White Porcelain Based on Protégé[J]. *Ceram Sci Art*, 2023, **57**(11): 39−41. DOI: 10.13212/j. cnki.csa.2023.11.130.

[24] Lewis P, Perez E, Piktus A, *et al*. Retrieval-augmented

generation for knowledge-intensive NLP tasks[C]//Advances in Neural Information Processing Systems 33 (NeurIPS 2020). Red Hook: Curran Associates Inc., 2020: 28613−28624.

[25] Yang A, Yang B, Hui B, *et al*. Qwen2.5 Technical Report[PP/OL]. (2024−12−19) [2025−3−25]. https://arxiv.org/abs/2412.15115.

[26] Hu J E, Shen Y, Wallis P, *et al*. LoRA: Low-Rank Adaptation of Large Language Models[C]//International Conference on Learning Representations (ICLR 2022). Appleton: ICLR 2022.

[27] Zeng A, Xu B, Wang B, *et al*. ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools[PP/OL]. (2024−06−18) [2025−4−2]. https://arxiv.org/abs/2406.12793.

[28] Guo D, Yang D, Zhang H, *et al*. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning[PP/OL]. (2025−01−22) [2025−4−15]. https://arxiv.org/abs/2501.12948.

[29] Demirhan H, Zadrozny W. Survey of Multimodal Medical Question Answering[J]. *BioMedInformatics*, 2024, **4** (1): 50−74. DOI: 10.3390/biomedinformatics4010004.