

生成式知识图谱构建综述

洪 钊¹, 黄鸿发^{1,2}

¹北京信息科技大学计算机学院, 北京

²拓尔思信息技术股份有限公司, 北京

收稿日期: 2025年11月26日; 录用日期: 2025年12月31日; 发布日期: 2026年1月9日

摘 要

知识图谱作为一种结构化知识表示方法, 对于组织人类知识并支持智能应用至关重要。传统的知识图谱构建(Knowledge Graph Construction, KGC)依赖判别式流水线模型, 存在误差累积与跨领域泛化能力弱等问题。近年来, 生成式方法凭借其序列到序列(sequence-to-sequence, Seq2Seq)的端到端建模优势, 逐渐成为有效的替代方案。特别是随着大语言模型(Large Language Models, LLMs)的兴起, 生成式知识图谱构建已从传统的序列到序列范式, 演进至大模型驱动的全流程构建新阶段。本文系统梳理了生成式知识图谱构建的技术进展: 首先回顾基于序列到序列的经典生成方法, 分析其核心机制与应用场景; 进而重点探讨大语言模型在本体构建、知识抽取与知识融合等关键环节中的方法与贡献; 通过对比不同范式的优势与局限, 本文进一步展望了生成式知识图谱在架构优化、多模态扩展与智能融合等方面的未来研究方向。

关键词

知识图谱构建, 生成式模型, 大语言模型

A Survey on Generative Knowledge Graph Construction

Zhao Hong¹, Hongfa Huang^{1,2}

¹Computer School of Beijing Information Science and Technology University, Beijing

²TRS Information Technology Co., Ltd., Beijing

Received: November 26, 2025; accepted: December 31, 2025; published: January 9, 2026

Abstract

Knowledge graphs, as a structured form of knowledge representation, play a crucial role in organizing human knowledge and enabling intelligent applications. Traditional knowledge graph

文章引用: 洪钊, 黄鸿发. 生成式知识图谱构建综述[J]. 人工智能与机器人研究, 2026, 15(1): 168-179.

DOI: 10.12677/airr.2026.151017

construction (KGC) pipelines rely on discriminative models, but they often suffer from error propagation and limited cross-domain generalization. In recent years, generative approaches have emerged as effective alternatives owing to their end-to-end sequence-to-sequence modeling capabilities. With the rapid advancement of large language models (LLMs), generative KGC has further evolved from conventional seq2seq paradigms toward a new stage of large-model-driven, full-process construction. This survey provides a comprehensive review of recent progress in generative knowledge graph construction. We first summarize classical seq2seq-based generative methods and analyze their core mechanisms and application scenarios. We then focus on the growing role of LLMs across key components—including ontology construction, knowledge extraction, and knowledge fusion—and highlight their methodological contributions. By comparing the advantages and limitations of different paradigms, this work outlines promising future directions for generative KGC, including architectural optimization, multimodal integration, and intelligent knowledge fusion.

Keywords

Knowledge Graph Construction, Generative Models, Large Language Models

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

知识图谱(Knowledge Graph, KG)作为一种结构化知识表示形式,能够以语义化的图结构组织和存储现实世界中的实体及其关系,已在搜索引擎[1]、推荐系统[2]、智能问答[3]等诸多应用中展现出巨大的价值。典型的大规模知识图谱系统,例如 Freebase [4]和 Wikidata [5],主要依赖人工标注或众包方式构建,但这种方式在规模扩展与实时更新方面成本极高。因此,如何从非结构化或半结构化数据中自动化构建高质量知识图谱,成为自然语言处理与知识工程领域的核心问题。

随着深度学习的快速发展,尤其是预训练语言模型的广泛应用,KGC 研究逐渐由判别式方法向生成式范式转变。受序列到序列框架启发,研究者开始探索通过文本生成的方式统一实现多种知识抽取任务。早期研究[6]首次提出利用生成式模型处理实体与关系抽取问题。随后,得益于 T5 [7]与 BART [8]等预训练生成模型的快速进步,生成式 KGC 方法[9]-[11]在多个基准数据集上取得了优异表现。这类方法通过将知识三元组线性化为特定格式的文本序列,实现了端到端的知识生成与统一建模。相比传统的流水线模型,生成式 KGC 在缓解误差累积、简化架构设计以及提升跨任务与跨领域迁移性方面展现出明显优势。

本综述聚焦于生成式知识图谱构建领域,旨在系统梳理该技术从“传统 Seq2Seq 驱动”到“LLMs 赋能”的完整演进脉络。文中不仅深入分析了各类生成式方法,还将传统判别式方法纳入讨论,以奠定技术基础并作为对比参照。最后,本文展望了该领域的未来研究方向,以期为研究者在方法选型与前沿探索上提供系统的参考。

2. 相关介绍

本节将介绍知识图谱的定义以及生成式知识图谱构建原理,为后续对生成式知识图谱构建方法提供知识基础。

2.1. 知识图谱定义

知识图谱是一种以语义图形式组织和表达知识的结构化表示方式。其核心思想是通过节点与边构建实体及其语义关系的网络, 以支持知识组织、推理与发现[12]-[14]。

早期研究通常将知识图谱定义为一个多关系图, 即节点表示实体, 边表示不同类型的语义关系。然而, 这一定义忽略了知识层次中的语义结构与背景知识。为更全面地刻画知识图谱的本质特征, Ehrlinger 和 Woß [13]进一步提出, 知识图谱不仅应包含事实性数据, 还应借助本体组织知识结构, 并结合推理机制以实现新知识的发现。

Wu 等人[14]进一步形式化了知识图谱的语义本质, 指出知识图谱是一种语义网络, 其节点代表概念, 边表示节点间的语义关系, 同时结合关于概念与关系的背景知识, 以保证图中蕴含的知识具有知识层级的完备性与解释性。

基于上述研究, 知识图谱可形式化定义如下:

定义 1: 一个知识图谱可定义为四元组 $\mathcal{G} = \{\mathcal{E}, \mathcal{R}, \mathcal{T}, \mathcal{F}_k\}$, 其中 $\mathcal{E}$ 表示概念的集合, 可具体包括实体与属性; $\mathcal{R}$ 表示关系的集合; $\mathcal{T}$ 表示事实三元组的集合, 标准的二元关系表示为 $(h, t, r) \in \mathcal{T}$, 其中 $h, t \in \mathcal{E}, r \in \mathcal{R}$, $\mathcal{F}_k$ 表示背景知识的集合, 是一个约束函数或知识库, 用以规范可能事实的语义空间, 即 $\mathcal{T} \subset \mathcal{F}_k(\{\mathcal{E}, \mathcal{R}\})$ 。

在实际应用中, $\mathcal{F}_k$ 可表现为规则集、模式或隐式知识约束, 例如逻辑推理规则、类型层级约束或领域本体等。这一形式化定义强调了知识图谱不仅是事实的集合, 更是一个语义化、可推理的知识载体。

2.2. 生成式知识图谱构建

知识图谱构建旨在从非结构化或半结构化数据源中自动抽取、链接并组织结构化知识, 以形成上述定义的知识图谱 $\mathcal{G}$ 。传统的 KGC 方法通常采用流水线式范式[15]-[17], 将整体任务拆解为若干子任务, KGC 的核心任务包括但不限于命名实体识别(Named Entity Recognition, NER) [18]、关系抽取(Relation Extraction, RE) [19]、事件抽取(Event Extraction, EE) [20]、实体链接(Entity Linking, EL) [21]以及知识图谱补全(Knowledge Graph Completion, KGCmp) [22]。图 1 为生成式知识图谱构建的整体流程。

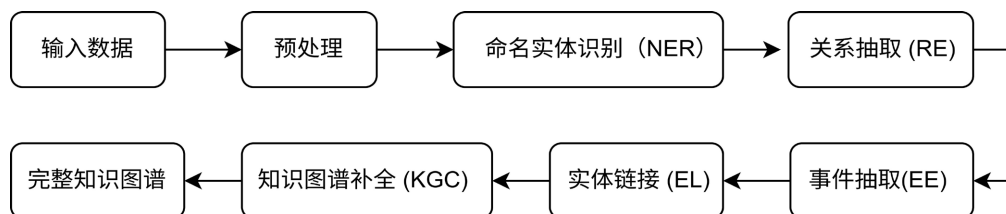


Figure 1. Overall pipeline of generative knowledge graph construction

图 1. 生成式知识图谱构建的整体流程

从机器学习视角来看, KGC 通常可视为一种结构预测任务。给定输入空间 x 与输出空间 y , KGC 目标是学习一个映射函数 F , 使得: $F(x) \rightarrow y$ 其中, 输入 $x \in \mathcal{X}$, 为原始数据, 输出 $y \in \mathcal{Y}$ 为相应的结构化知识单元。例如, 对于输入句子 “Bill Gates and Paul Allen co-founded Microsoft in 1975.”, KGC 系统应预测出结构化的关系事实: (Bill Gates, co-founded, Microsoft), (Paul Allen, co-founded, Microsoft)。

从系统论角度出发, 知识图谱构建过程可定义为一个从数据源到知识图谱的映射过程[23]:

$f: D \times f_k(D) \rightarrow G$ 。其中 D 表示数据源的集合, 可以是文本、网页、表格或多模态内容; $f_k(D)$ 表示与数据相关的背景知识或领域知识, 通常由预定义规则、本体或语言模型提供; G 表示最终生成的知识图

谱。该定义强调, KGC 过程并非孤立运行, 而是依赖于外部或隐含的背景知识。没有 $f_k(D)$ 的支持, 即缺乏语义约束与上下文先验, KGC 系统往往无法从数据中提炼出“知识层级”上的结构化信息。

3. 生成式知识图谱构建方法

知识图谱构建的主流技术涵盖判别式与生成式两大范式。判别式方法依赖分类模型以最大化后验概率, 而生成式方法则通过条件语言建模, 将知识抽取转化为序列生成任务, 实现了端到端的构建流程。生成式 KGC 技术本身肇始于传统 Seq2Seq 框架, 并逐步演进为大模型驱动的现代化范式, 与判别式方法共同构成了互补的技术体系。本节将系统梳理这一发展脉络, 并依据图 2 所示的分类框架展开具体介绍。



Figure 2. Taxonomy of generative knowledge graph construction

图 2. 生成式知识图谱构建方法分类

3.1. 判别式与生成式方法

3.1.1. 判别式方法

判别式模型旨在根据输入句子的特征, 预测其对应的实体关系标签。给定带注释的输入句子 x , 以及句中可能存在的重叠三元组集合 $t_j = \{(s, r, o)\}$, 其优化目标是最大化数据似然函数:

$$p_{cls}(t|x) = \prod_{(s,r,o) \in t_j} p((s,r,o) | x_j) \quad (1)$$

该类模型通过对输入语义特征的建模, 实现对实体关系的分类预测。另一种典型的判别式策略是序列标注方法[24]。对于包含 n 个词的输入句子 x , 模型为每个词位置 i 分配基于 BIESO (Begin, Inside, End, Single, Outside) 标注体系的标签序列。设预定义关系集合的规模为 $|R|$, 并以“1”“2”表示不同角色顺序, 则模型在训练时通过最大化目标序列的对数似然实现参数优化:

$$p_{tag}(y|x) = \frac{\exp(h_i, y_i)}{\sum_{y' \in R} \exp(h_i, y'_i)} \quad (2)$$

其中, h_i 表示位置 i 处的隐向量表示。该类模型具有结构清晰、解释性强的优点, 但在处理重叠关系或跨句依赖时往往受限于固定标签空间。

3.1.2. 生成式方法

生成式模型的核心思想是将三元组抽取任务转化为条件文本生成问题。设输入句子为 x , 线性化后的目标三元组序列为 y , 模型的目标是自回归地生成输出序列:

$$p_{gen}(y|x) = \prod_{i=1}^{len(y)} p_{gen}(y_i | y_{<i}, x) \quad (3)$$

通过在此框架下微调预训练的序列到序列模型, 如 MASS [25]、T5 [7] 和 BART [8], 可利用交叉熵损失函数最大化生成结果的对数似然。生成式方法具备统一建模、语义灵活等优势, 能够同时完成实体识别与关系抽取。然而, 该类方法在长文本或多重关系场景中可能出现解码不稳定与语义幻觉等问题。

3.2. 传统序列到序列生成式方法

3.2.1. 基于拷贝的序列生成方法

该方法通过显式拷贝机制降低实体幻觉风险, 其核心思想是让解码器直接从输入序列中复制头、尾实体, 而关系标签仍取自预定义词表。Zeng 等[6]首次提出 CopyRE, 以端到端方式同时抽取重叠关系的三元组, 并引入拷贝指针缓解实体重复生成问题。随后, Zeng 等[26]将三元组生成顺序建模为马尔可夫决策过程, 利用强化学习搜索最优生成序列, 显著提升重叠场景下的召回率。为进一步增强实体边界一致性, Zeng 等[27]设计融合特征空间的非线性映射层, 使头、尾实体在拷贝前获得任务相关的统一表征。面向文档级抽取, Huang 等[28]提出 Top-k 拷贝策略, 通过动态剪枝实体候选对, 降低长文本中实体组合的计算复杂度。

3.2.2. 基于结构线性化的序列生成方法

该方法通过引入结构化知识与标签语义, 使生成模型具备统一的输出格式与较高的语义一致性。Lu 等[29]基于 T5 框架提出结构线性化事件抽取模型, 将事件结构转化为序列化文本输出, 并借助事件模式约束解码空间以减少噪声。Lou 等[30]提出多层双向网络(MLBiNet), 实现文档级事件间关联与语义信息的联合建模。Zhang 等[31]与 Ye 等[32]进一步引入对比学习与动态注意力掩码机制, 以缓解生成式架构中的语义冲突问题。Cabot 与 Navigli [33]则通过简化的三元组分解方法实现关系抽取在跨领域与长文档中的迁移。针对嵌套命名实体识别任务, Straková 等[34]提出基于 BILOU 标注方案的扁平化编码算法; Zhang 等[35]利用因果调整理论纠正生成偏差。Cao 等[36]提出的 GENRE 模型通过自回归生成方式实现实体链接任务, 可捕获上下文与实体名称之间的细粒度交互。Wang 等[37]与 Lu 等[11]进一步扩展该框架至异构结构化信息抽取, 实现任务无关的统一生成。

3.2.3. 基于标签增强的序列生成方法

该方法在输入或输出端引入显式标签标记, 如 “[实体|类型]”, 以自然语言形式表达类别语义, 从而充分激活预训练模型的先验知识。Athiwaratkun 等[38]在多项结构预测任务中验证, 复制输入词并插入标签括号可显著降低边界歧义。Paolini 等[10]提出的 TANL 框架进一步统一了不同任务的标签语义, 使解码器在生成过程中同时完成实体识别与类别分配。Cao 等[36]将标签增强与自回归实体检索结合, 通过跨编码机制对齐上下文与实体名称, 提升低资源场景下的链接准确率。然而, 由于输出序列长度通常远超输入, 文档级任务易出现标签遗漏或位置偏移等问题。

3.3. 大模型驱动生成式方法

3.3.1. 大模型增强的本体构建

大语言模型的引入为本体工程带来了范式转变, 其研究主要分为自上而下与自下而上两类路径。

1) 自上而下指的是 LLMs 作为本体建模助手。该方向延续语义网与知识工程传统, 强调在预定义语义需求下的本体建模。LLMs 能将自然语言需求转化为网络本体语言等形式化本体, 实现从语义需求到结构化模型的半自动映射。代表性工作包括 Ontogenia 框架[39], 通过“元认知提示”实现自反性建模与结构校正。近期研究还提出“无记忆建模”和“反思迭代”两种策略, 使 LLMs 在识别类、属性及生成逻辑公理方面已接近初级本体工程师水平。总体而言, 自上而下范式聚焦于语义一致性与人机协同, 使 LLMs 从辅助分析工具演化为主动建模合作者。

2) 自下而上主要是面向 LLMs 的本体模式构建。该路径强调知识图谱对大模型的“结构化记忆”作用, 推动从“为人类解释”向“为模型推理”转变。代表性工作如 GraphRAG [40]与 OntoRAG [41], 通过开放信息抽取与聚类归纳实现“数据到模式”的自动生成。随后, EDC 框架[42]提出抽取、定义与规范化的三阶段流程, 可实现模式归一与动态扩展。此外, AutoSchemaKG [43]融合有模式与无模式范式, 实现企业级知识图谱的实时演化。

3.3.2. 大模型驱动的知识抽取

大语言模型驱动的知识抽取方法主要沿着两条路径演化: 基于模式的抽取与无模式抽取。前者依赖明确的结构约束与语义模板, 强调一致性与规范性; 后者摆脱预定义本体的限制, 更注重开放性与自适应能力。两者共同构成了当代知识抽取研究的主要范式。

1) 基于模式的抽取

早期方法依托固定本体或领域知识库进行结构化抽取, 例如 KARMA [44]和 ODKE+ [45]通过静态模式指导实体识别与关系分类, 确保逻辑一致性但缺乏跨域灵活性。为增强可扩展性, Feng 等[46]提出“本体引导式抽取”, 先由 LLMs 自动生成领域本体, 再用于 RDF 三元组抽取, 实现结构约束与数据驱动的平衡。近期研究转向动态与自适应模式学习, 如 AdaKGC [47]通过前缀指令与动态解码机制实现模式演化, 而 AutoSchemaKG [43]则结合聚类与关系发现, 实现从大规模语料中自动诱导本体模式。这一方向体现了从“静态模板约束”到“模式共演化”的转变, 使抽取系统具备持续学习与自更新能力。

2) 无模式抽取

该范式不依赖外部本体, 直接从文本中识别实体与关系, 体现出 LLMs 的自组织推理潜力。早期工作如 Nie 等[48]将链式思维引入抽取过程, 通过逐步推理替代显式模式; AutoRE [49]通过指令微调学习隐含关系模式, 提升跨文档一致性。后续研究如 ChatIE [50]将抽取任务建模为多轮对话, 借助交互式问答提升准确率; KGGGen [51]采用两阶段生成流程, 首先检测实体然后生成关系以减少误差传播。代表性的 EDC 框架[42]通过开放信息抽取生成自然语言三元组, 并经定义与规范化形成可扩展知识图谱。

3.3.3. 大模型赋能的知识融合

大模型赋能的知识融合旨在实现多源知识图谱在模式层与实例层的统一与协同, 以构建语义一致、结构完备的知识体系。其研究脉络大致经历了从本体驱动到数据驱动, 再到大模型语义融合的演进。

1) 模式层融合

模式层融合聚焦于概念、实体类型及关系结构的统一。早期方法 Kommineni 等[52]依赖显式本体约束以保证一致性, 但难以适应跨域场景。随后, LKD-KGC [53]提出基于向量聚类与 LLMs 去重的自适应融合策略, 使模式统一可从数据中自发形成。近期, EDC 框架[42]通过 LLMs 生成自然语言定义并计算

语义相似度, 实现高精度的模式对齐与语义规范化, 标志着从规则约束向语义驱动的转变。

2) 实例层融合

实例层融合关注实体对齐、消歧与冲突消解。早期工作如 KGGEN [51] 利用迭代聚类完成语义层级的融合, 而 LLM-Align [54] 与 EntGPT [55] 将对齐任务转化为上下文推理过程, 通过多轮提示与候选筛选显著提升精度。此外, Pons 等 [56] 在 RAG 框架下结合结构与检索信号进行零样本消歧, COMEM [57] 则采用层次化管线以兼顾效率与语义准确性, 展现了多层次 LLMs 协作的潜力。

3) 综合融合框架

综合框架实现模式层与实例层的协同统一, 推动知识融合由分阶段流程向端到端架构演进。KARMA [44] 通过多智能体协作实现结构对齐与冲突解决; ODKE+ [45] 结合本体监督与实例验证以提升全局一致性; 而 Graphusion [58] 进一步提出基于生成式提示的统一框架, 将融合、推理与验证整合为单一循环流程。

3.4. 生成式知识图谱构建方法比较

近年来, 生成式知识图谱构建方法在统一抽取范式与增强语义一致性方面取得显著进展。根据生成机制与模型特征, 可将现有研究大致分为三类: 基于拷贝机制的序列生成、基于结构线性化的序列生成以及基于标签增强的序列生成方法; 同时, 随着大语言模型的兴起, 又衍生出以大模型为核心驱动力的本体构建、知识抽取与知识融合框架。

基于拷贝机制的方法强调实体的精确复制, 通过指针网络或强化学习策略显著降低实体幻觉与重复生成风险, 但在长文本与复杂关系场景下仍存在泛化不足的问题。结构线性化方法则通过引入模式约束与结构化输出, 实现统一的语义表达与任务扩展, 但解码空间较大, 易引发结构偏移。标签增强方法在输入或输出中显式标注类别语义, 有助于激活预训练模型的知识潜能, 实现任务统一与跨域迁移, 然而在文档级任务中常面临序列过长与标签遗漏问题。

与上述传统生成式架构相比, 大模型驱动的知识图谱构建展现出更强的自适应与推理能力。LLMs 不仅可辅助本体模式的自动生成与演化, 还能在知识抽取与融合阶段承担语义对齐、冲突消解与一致性验证等复杂任务。总体而言, 生成式与大模型赋能方法正逐步融合, 推动知识图谱构建从分阶段流水线向端到端、认知式系统演进。表 1 总结了不同生成式方法的优缺点。

Table 1. Summary of advantages and disadvantages of generative knowledge graph construction methods
表 1. 生成式知识图谱构建方法优缺点总结

方法	优点	缺点
基于拷贝的序列生成方法	有效缓解实体幻觉与重复生成问题; 对重叠关系抽取表现优异	对长文本与复杂句式敏感; 关系生成仍依赖封闭标签集
基于结构线性化的序列生成方法	输出格式统一、语义一致性高; 可扩展至事件抽取与实体链接等任务	解码空间大、推理成本高; 标签错误或结构偏移在长文本中更明显
基于标签增强的序列生成方法	显式融合语义标签与生成语义, 提升模型对实体类别与关系类型的区分能力; 在低资源与跨任务场景下表现稳定	输出序列较长, 易出现标签遗漏或位置偏移; 对标签设计敏感, 泛化至开放域任务时需额外模板优化
大模型增强的本体构建	提升本体构建的语义一致性与自动化水平; 支持模式共演化	对领域语义理解依赖强; 缺乏系统评估与标准化输出
大模型驱动的知识抽取	可在无人工本体条件下实现高质量三元组生成; 适应跨领域场景	抽取结果一致性与可控性较弱; 易出现语义漂移
大模型赋能的知识融合	兼顾跨源融合与语义一致性; 支持端到端整合与验证	依赖高算力与人工验证; 在大规模场景下仍面临效率瓶颈

4. 未来方向

尽管生成式知识图谱构建在方法和应用上均取得了显著进展,但仍存在模型同质化、跨模态迁移能力有限及应用拓展不充分等问题。未来的发展趋势可从以下几个方面展开。

1) 生成架构优化与可解释性增强。

当前大多数生成式 KGC 框架基于 Transformer 架构,导致模型推理机制高度同质化。未来研究可探索神经-符号融合模型(Neuro-Symbolic Models) [17] [59] [60],以融合神经网络的表示能力与符号逻辑的可解释性,从而实现端到端的可控生成。此外,脉冲神经网络[61]、动态神经网络[62]、常微分方程网络[63]与扩散模型[64]等新兴架构,也为生成式 KGC 的因果建模与知识表达提供了潜在突破口。

2) 统一化与多模态扩展。

借鉴 T5 [7]的“Text-to-Text”范式,未来生成模型可进一步统一不同任务与领域,将命名实体识别、关系抽取、事件抽取与知识补全映射到统一的文本生成框架。代表性工作如 UIE [11]已展示出多任务融合潜力。同时,多模态知识图谱将成为重要趋势,通过视觉-语言预训练模型实现跨模态实体对齐与语义一致性[65]。例如, VaLiK 模型通过视觉语义级联与跨模态验证实现无监督图像实体链接,为多模态知识融合奠定了基础。

3) 智能化知识融合与认知应用。

未来的研究焦点将从“知识构建”转向“知识驱动推理”。高质量、结构化的知识图谱将成为大模型的语义记忆与认知中层[66] [67],支撑可解释推理与因果分析。动态知识记忆系统[68] [69]通过时间知识图谱实现持续演化与自我反思,为自治智能体提供长期记忆支持。此外,知识图谱还将在认知决策、医疗诊断与科学发现等场景中承担核心作用,如 CogER [70]将推荐建模为认知感知推理过程,PKG-LLM [71]利用领域知识图谱辅助心理健康预测。这一趋势表明,生成式 KGC 正迈向可持续演化、跨模态融合与认知增强的智能阶段。

5. 结语

生成式方法为知识图谱构建领域注入了新的活力,并推动了其技术范式的根本性转变。本文系统回顾并梳理了生成式知识图谱构建从传统序列到序列模型到大语言模型赋能的技术演进历程。传统生成式方法实现了端到端的知识抽取,有效缓解了判别式流水线模型的误差累积问题。而大语言模型的兴起,则进一步将生成式构建的范围从单一的知识抽取,扩展至涵盖本体构建、知识抽取与知识融合的全流程,显著提升了构建过程的自动化与智能化水平。尽管生成式知识图谱构建已取得显著进展,其在模型架构、跨模态能力与应用深度方面仍面临挑战。未来的研究将聚焦于架构创新与可解释性增强、多模态知识融合,以及动态演化与认知推理等关键方向。本文期望能够为后续研究提供系统化参考,并推动生成式知识图谱构建迈向新阶段。

参考文献

- [1] Liu, J., Ren, J., Zheng, W., Chi, L., Lee, I. and Xia, F. (2020) Web of Scholars: A Scholar Knowledge Graph. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, online, 25-30 July 2020, 2153-2156. <https://doi.org/10.1145/3397271.3401405>
- [2] Wang, X., He, X., Cao, Y., Liu, M. and Chua, T. (2019) KGAT: Knowledge Graph Attention Network for Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, 4-8 August 2019, 950-958. <https://doi.org/10.1145/3292500.3330989>
- [3] Bao, J., Duan, N., Yan, Z., Zhou, M. and Zhao, T. (2016) Constraint-Based Question Answering with Knowledge Graph. *Proceedings of the 26th International Conference on Computational Linguistics*, Osaka, 11-17 December 2016, 2503-2514.

- [4] Bollacker, K.D., Cook, R.P. and Tufts, P. (2007) Freebase: A Shared Database of Structured General Human Knowledge. *Proceedings of the 22nd National Conference on Artificial Intelligence*, Vancouver, 22-26 July 2007 1962-1963.
- [5] Vrandečić, D. (2012) Wikidata: A New Platform for Collaborative Data Collection. *Proceedings of the 21st International Conference on World Wide Web*, Lyon, 16-20 April 2012, 1063-1064. <https://doi.org/10.1145/2187980.2188242>
- [6] Zeng, X., Zeng, D., He, S., Liu, K. and Zhao, J. (2018) Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), Melbourne, 15-20 July 2018, 506-514. <https://doi.org/10.18653/v1/p18-1047>
- [7] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J. (2020) Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research*, **21**, 1-67.
- [8] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., et al. (2020) BART: Denoising Sequence-to-Sequence Pre-Training for Natural Language Generation, Translation, and Comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, 5-10 July 2020, 7871-7880. <https://doi.org/10.18653/v1/2020.acl-main.703>
- [9] Yan, H., Dai, J., Ji, T., Qiu, X. and Zhang, Z. (2021) A Unified Generative Framework for Aspect-Based Sentiment Analysis. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Volume 1: Long Papers), Online, 1-6 August 2021, 2416-2429. <https://doi.org/10.18653/v1/2021.acl-long.188>
- [10] Paolini, G., Athiwaratkun, B., Krone, J., Ma, J., Achille, A., Anubhai, R., dos Santos, C.N., Xiang, B. and Soatto, S. (2021) Structured Prediction as Translation between Augmented Natural Languages. *Proceedings of the 9th International Conference on Learning Representations*, 3-7 May 2021, 1-7.
- [11] Lü, Y., Liu, Q., Dai, D., Xiao, X., Lin, H., Han, X., et al. (2022) Unified Structure Generation for Universal Information Extraction. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), Dublin, 22-27 May 2022, 5755-5772. <https://doi.org/10.18653/v1/2022.acl-long.395>
- [12] Wang, Q., Mao, Z., Wang, B. and Guo, L. (2017) Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering*, **29**, 2724-2743. <https://doi.org/10.1109/tkde.2017.2754499>
- [13] Ehrlinger, L. and Wöß, W. (2016) Towards a Definition of Knowledge Graphs. *Proceedings of the SUCCESS Workshop at SEMANTiCS 2016*, Leipzig, 12-15 September 2016, 1-4.
- [14] Wu, X., Wu, J., Fu, X., Li, J., Zhou, P. and Jiang, X. (2019) Automatic Knowledge Graph Construction: A Report on the 2019 ICDM/ICBK Contest. 2019 *IEEE International Conference on Data Mining (ICDM)*, Beijing, 8-11 November 2019, 1540-1545. <https://doi.org/10.1109/icdm.2019.00204>
- [15] Luan, Y., He, L., Ostendorf, M. and Hajishirzi, H. (2018) Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, 31 October-4 November 2018, 3219-3232. <https://doi.org/10.18653/v1/d18-1360>
- [16] Mesquita, F., Cannaviccio, M., Schmidek, J., Mirza, P. and Barbosa, D. (2019) Knowledgenet: A Benchmark Dataset for Knowledge Base Population. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong SAR (China), 3-7 November 2019, 749-758. <https://doi.org/10.18653/v1/d19-1069>
- [17] Zhang, N., Xu, X., Tao, L., Yu, H., Ye, H., Qiao, S., et al. (2022) Deepke: A Deep Learning Based Knowledge Extraction Toolkit for Knowledge Base Population. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Abu Dhabi, 7-11 December 2022, 98-108. <https://doi.org/10.18653/v1/2022.emnlp-demos.10>
- [18] Chiu, J.P.C. and Nichols, E. (2016) Named Entity Recognition with Bidirectional LSTM-CNNs. *Transactions of the Association for Computational Linguistics*, **4**, 357-370. https://doi.org/10.1162/tacl_a_00104
- [19] Zeng, D., Liu, K., Chen, Y. and Zhao, J. (2015) Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, 17-21 September 2015, 1753-1762. <https://doi.org/10.18653/v1/d15-1203>
- [20] Shen, W., Wang, J. and Han, J. (2015) Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions. *IEEE Transactions on Knowledge and Data Engineering*, **27**, 443-460. <https://doi.org/10.1109/tkde.2014.2327028>
- [21] Lin, Y., Liu, Z., Sun, M., Liu, Y. and Zhu, X. (2015) Learning Entity and Relation Embeddings for Knowledge Graph Completion. *Proceedings of the AAAI Conference on Artificial Intelligence*, **29**, 2181-2187. <https://doi.org/10.1609/aaai.v29i1.9491>
- [22] Du, X. and Cardie, C. (2020) Event Extraction by Answering (Almost) Natural Questions. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, 16-20 November 2020, 671-683.

- <https://doi.org/10.18653/v1/2020.emnlp-main.49>
- [23] Zhong, L., Wu, J., Li, Q., Peng, H. and Wu, X. (2023) A Comprehensive Survey on Automatic Knowledge Graph Construction. *ACM Computing Surveys*, **56**, 1-62. <https://doi.org/10.1145/3618295>
 - [24] Zheng, S., Wang, F., Bao, H., Hao, Y., Zhou, P. and Xu, B. (2017) Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), Vancouver, 30 July-4 August 2017, 1227-1236. <https://doi.org/10.18653/v1/p17-1113>
 - [25] Song, K.T., Tan, X., Qin, T., Lu, J.F. and Liu, T.Y. (2019) MASS: Masked Sequence to Sequence Pre-Training for Language Generation. *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 5926-5936.
 - [26] Zeng, X., He, S., Zeng, D., Liu, K., Liu, S. and Zhao, J. (2019) Learning the Extraction Order of Multiple Relational Facts in a Sentence with Reinforcement Learning. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, 3-7 November 2019, 367-377. <https://doi.org/10.18653/v1/d19-1035>
 - [27] Zeng, D., Zhang, H. and Liu, Q. (2020) Copymtl: Copy Mechanism for Joint Extraction of Entities and Relations with Multi-Task Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 9507-9514. <https://doi.org/10.1609/aaai.v34i05.6495>
 - [28] Huang, K., Tang, S. and Peng, N. (2021) Document-Level Entity-Based Extraction as Template Generation. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online, 7-11 November 2021, 5257-5269. <https://doi.org/10.18653/v1/2021.emnlp-main.426>
 - [29] Lu, Y., Lin, H., Xu, J., Han, X., Tang, J., Li, A., et al. (2021) Text2event: Controllable Sequence-To-Structure Generation for End-to-End Event Extraction. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Volume 1: Long Papers), Online, 1-6 August 2021, 2795-2806. <https://doi.org/10.18653/v1/2021.acl-long.217>
 - [30] Lou, D., Liao, Z., Deng, S., Zhang, N. and Chen, H. (2021) Mlbinet: A Cross-Sentence Collective Event Detection Network. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Volume 1: Long Papers), Online, 1-6 August 2021, 4829-4839. <https://doi.org/10.18653/v1/2021.acl-long.373>
 - [31] Zhang, N., Ye, H., Deng, S., Tan, C., Chen, M., Huang, S., et al. (2021) Contrastive Information Extraction with Generative Transformer. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **29**, 3077-3088. <https://doi.org/10.1109/taslp.2021.3110126>
 - [32] Ye, H., Zhang, N., Deng, S., Chen, M., Tan, C., Huang, F., et al. (2021) Contrastive Triple Extraction with Generative Transformer. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 14257-14265. <https://doi.org/10.1609/aaai.v35i16.17677>
 - [33] Cabot, P.L.H., and Navigli, R. (2021) REBEL: Relation Extraction by End-to-End Language Generation. *Findings of the Association for Computational Linguistics: EMNLP 2021*, 7-11 November 2021, 2370-2381.
 - [34] Straková, J., Straka, M. and Hajic, J. (2019) Neural Architectures for Nested NER through Linearization. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, 28 July-2 August 2019, 5326-5331. <https://doi.org/10.18653/v1/p19-1527>
 - [35] Zhang, S., Shen, Y., Tan, Z., Wu, Y. and Lu, W. (2022) De-Bias for Generative Extraction in Unified NER Task. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), Dublin, 22-27 May 2022, 808-818. <https://doi.org/10.18653/v1/2022.acl-long.59>
 - [36] De Cao, N., Izacard, G., Riedel, S. and Petroni, F. (2021) Autoregressive Entity Retrieval. *Proceedings of the 9th International Conference on Learning Representations*, 3-7 May 2021, 1-17.
 - [37] Wang, C., Liu, X., Chen, Z., Hong, H., Tang, J. and Song, D. (2022) Deepstruct: Pretraining of Language Models for Structure Prediction. *Findings of the Association for Computational Linguistics: ACL 2022*, Dublin, 22-27 May 2022 803-823. <https://doi.org/10.18653/v1/2022.findings-acl.67>
 - [38] Athiwaratkun, B., Nogueira dos Santos, C., Krone, J. and Xiang, B. (2020) Augmented Natural Language for Generative Sequence Labeling. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, 16-20 November 2020, 375-385. <https://doi.org/10.18653/v1/2020.emnlp-main.27>
 - [39] Lippolis, A.S., Ceriani, M., Zuppiroli, S. and Nuzzolese, A.G. (2025) Ontogenia: Ontology Generation with Metacognitive Prompting in Large Language Models. In: Meroño Peñuela, A., et al., Eds., *Lecture Notes in Computer Science*, Springer Nature, 259-265. https://doi.org/10.1007/978-3-031-78952-6_38
 - [40] Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S. and Larson, J. (2024) From Local to Global: A Graph RAG Approach to Query-Focused Summarization. arXiv:2404.16130.
 - [41] Tiwari, Y., Lone, O.A. and Pal, M. (2025) OntoRAG: Enhancing Question-Answering through Automated Ontology

- Derivation from Unstructured Knowledge Bases. arXiv:2506.00664.
- [42] Zhang, B. and Soh, H. (2024) Extract, Define, Canonicalize: An LLM-Based Framework for Knowledge Graph Construction. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Miami, 12-16 November 2024, 9820-9836. <https://doi.org/10.18653/v1/2024.emnlp-main.548>
 - [43] Bai, J.X., Fan, W., Hu, Q., Zong, Q., Li, C.Y., Tsang, H.T., Luo, H.Y., Yim, Y.W., Huang, H.Y., Zhou, X., *et al.* (2025) AutoSchemaKG: Autonomous Knowledge Graph Construction through Dynamic Schema Induction from Web-Scale Corpora. arXiv:2505.23628.
 - [44] Lu, Y.X. and Wang, J.Z. (2025) KARMA: Leveraging Multi-Agent LLMs for Automated Knowledge Graph Enrichment. arXiv:2502.06472.
 - [45] Khorshidi, S., Nikfarjam, A., Shankar, S., Sang, Y., Govind, Y., Jang, H., Kasgari, A., McClimans, A., Soliman, M., Konda, V., *et al.* (2025) ODKE+: Ontology-Guided Open-Domain Knowledge Extraction with LLMs. arXiv:2509.04696.
 - [46] Feng, X.H., Wu, X.X. and Meng, H. (2024) Ontology-Grounded Automatic Knowledge Graph Construction by LLMs under Wikidata Schema. arXiv:2412.20942.
 - [47] Ye, H., Gui, H., Xu, X., Chen, X., Chen, H. and Zhang, N. (2023) Schema-Adaptable Knowledge Graph Construction. *Findings of the Association for Computational Linguistics: EMNLP 2023*, Singapore, 6-10 December 2023, 6408-6431. <https://doi.org/10.18653/v1/2023.findings-emnlp.425>
 - [48] Nie, J.X., Hou, X., Song, W.F., Wang, X., Zhang, X.Y., Jin, X.L., Zhang, S.Z. and Shi, J.Q. (2024) Knowledge Graph Efficient Construction: Embedding Chain-of-Thought into LLMs. *Proceedings of the VLDB Endowment*, Guangzhou, 25-29 August 2024, 3041-3054.
 - [49] Xue, L., Zhang, D., Dong, Y. and Tang, J. (2024) Autore: Document-Level Relation Extraction with Large Language Models. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 3: System Demonstrations), Bangkok, 11-16 August 2024, 211-220. <https://doi.org/10.18653/v1/2024.acl-demos.20>
 - [50] Wei, X., Cui, X.Y., Cheng, N., Wang, X.B., Zhang, X., Huang, S., Xie, P.J., Xu, J.N., Chen, Y.F., Zhang, M.S., *et al.* (2024) ChatIE: Zero-Shot Information Extraction via Chatting with ChatGPT. arXiv:2302.10205.
 - [51] Mo, B., Yu, K., Kazdan, J., Mpala, P., Yu, L., Cundy, C., Kanatsoulis, C. and Koyejo, S. (2025) KGen: Extracting Knowledge Graphs from Plain Text with Language Models. arXiv:2502.09956.
 - [52] Kommineni, V.K., König-Ries, B. and Samuel, S. (2024) Towards the Automation of Knowledge Graph Construction Using Large Language Models. *CEUR Workshop Proceedings*, Amsterdam, 17-19 September 2024, 1-15
 - [53] Sun, J.Q., Qian, S.Y., Han, Z.C., Li, W., Qian, Z.L., Yang, D.Y., Cao, J. and Xue, G.T. (2025) LKD-KGC: Do-Main-Specific KG Construction via LLM-Driven Knowledge Dependency Parsing. arXiv:2505.24163.
 - [54] Chen, X., Lu, T. and Wang, Z.C. (2024) LLM-Align: Utilizing Large Language Models for Entity Alignment in Knowledge Graphs. arXiv:2412.04690.
 - [55] Ding, Y.F., Poudel, A., Zeng, Q.K., Weninger, T., Veeramani, B. and Bhattacharya, S. (2025) EntGPT: Entity Linking with Generative Large Language Models. arXiv:2402.06738.
 - [56] Pons, G., Bilalli, B. and Queral, A. (2024) Knowledge Graphs for Enhancing Large Language Models in Entity Disambiguation. In: Demartini, G., *et al.*, Eds., *Lecture Notes in Computer Science*, Springer Nature, 162-179. https://doi.org/10.1007/978-3-031-77844-5_9
 - [57] Wang, T.S., Chen, X.Y., Lin, H.Y., Chen, X.A., Han, X.P., Wang, H., Zeng, Z.Y. and Sun, L. (2024) Match, Compare, or Select? An Investigation of Large Language Models for Entity Matching. arXiv:2405.16884.
 - [58] Yang, R., Yang, B.M., Ouyang, S.X., She, T.W., Feng, A.S., Jiang, Y., Lecue, F., Lu, J.H. and Li, I. (2024) Graphusion: Leveraging Large Language Models for Scientific Knowledge Graph Fusion and Construction in NLP Education. arXiv:2407.10794.
 - [59] Galassi, A. (2021) Deep Networks and Knowledge: from Rule Learning to Neural-Symbolic Argument Mining. Ph.D. Thesis, University of Bologna.
 - [60] Negro, P. and Pons, C. (2022) Artificial Intelligence Techniques Based on the Integration of Symbolic Logic and Deep Neural Networks: A Systematic Review of the Literature. *Inteligencia Artificial*, **25**, 13-41. <https://doi.org/10.4114/intartif.vol25iss69pp13-41>
 - [61] Tavanaei, A., Ghodrati, M., Kheradpisheh, S.R., Masquelier, T. and Maida, A. (2019) Deep Learning in Spiking Neural Networks. *Neural Networks*, **111**, 47-63. <https://doi.org/10.1016/j.neunet.2018.12.002>
 - [62] Xu, C.W. and McAuley, J.J. (2023) A Survey on Dynamic Neural Networks for Natural Language Processing. *Findings of the Association for Computational Linguistics: EACL 2023*, Dubrovnik, 2-6 May 2023, 2370-2381.
 - [63] Chen, X., Li, L., Deng, S.M., Tan, C.Q., Xu, C.L., Huang, F., Si, L., Chen, H.J. and Zhang, N.Y. (2022) LightNER: A Lightweight Tuning Paradigm for Low-Resource NER via Pluggable Prompting. *Proceedings of the 29th International Conference on Computational Linguistics*, Gyeongju, 12-17 October 2022, 2374-2387.

-
- [64] Dhariwal, P. and Nichol, A.Q. (2021) Diffusion Models Beat Gans on Image Synthesis. *Annual Conference on Neural Information Processing Systems 2021*, 6-14 December 2021, 8780-8794.
 - [65] Liu, J.M., Meng, S.Y., Gao, Y.T., Mao, S., Cai, P.L., Yan, G.H., Chen, Y.R., Bian, Z.L., Wang, D. and Shi, B.T. (2025) Aligning Vision to Language: Annotation-Free Multimodal Knowledge Graph Construction for Enhanced LLMs Reasoning. arXiv:2503.12972.
 - [66] Kim, Y.J., Kang, E.J., Kim, J. and Huang, H.H. (2024) Causal Reasoning in Large Language Models: A Knowledge Graph Approach. arXiv:2410.11588.
 - [67] Wu, W.J., Jing, Y.C., Wang, Y.J., Hu, W.B. and Tao, D.C. (2025) Graph-Augmented Reasoning: Evolving Step-by-Step Knowledge Graph Retrieval for LLMsReasoning. arXiv:2503.01642.
 - [68] Xu, W.J., Liang, Z.J., Mei, K., Gao, H., Tan, J.T. and Zhang, Y.F. (2025) A-MEM: Agentic Memory for LLMsAgents. arXiv:2502.12110.
 - [69] Rasmussen, P., Paliychuk, P., Beauvais, T., Ryan, J. and Chalef, D. (2025) Zep: A Temporal Knowledge Graph Architecture for Agent Memory. arXiv:2501.13956.
 - [70] Bing, Q., Zhu, Q. and Dou, Z. (2023) Cognition-Aware Knowledge Graph Reasoning for Explainable Recommendation. *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, Singapore, 27 February 2023-3 March 2023, 402-410. <https://doi.org/10.1145/3539597.3570391>
 - [71] Sarabadani, A., Taherinia, H., Ghadiri, N., Shahmarvandi, E.K. and Mousa, R. (2025) PKG-LLM: A Framework for Predicting GAD and MDD Using Knowledge Graphs and Large Language Models in Cognitive Neuroscience. <https://www.preprints.org/manuscript/202502.0982>