

融合动态知识图谱构建的跨语言船舶设计问答方法

梁钟渝¹, 程良伦²

¹广东工业大学自动化学院, 广东 广州

²广东工业大学计算机学院, 广东 广州

收稿日期: 2026年4月23日; 录用日期: 2026年5月22日; 发布日期: 2026年5月29日

摘要

针对船舶情报系统中知识图谱构建成本高、多语言知识碎片化及动态更新困难等问题, 文章提出融合动态知识图谱构建的跨语言检索增强生成方法。该方法通过迭代式精细化多标签抽取算法, 从多语言非结构化文档中自动构建可演化的船舶设计知识图谱; 引入路径置信度加权的混合检索机制, 将结构化知识路径与语义检索协同融合; 并通过跨语言自然语言推理驱动的协同验证策略, 提升生成答案的事实可靠性。实验在英、中、法、德四语种文档上进行评估, 多语言平均F1值达86.8%, 跨语言实体对齐准确率达94.7%; 端到端问答事实正确率为92.3%, 较基线方法提升7.8个百分点; 针对近期新发布规范的查询, 动态更新机制使事实正确性较静态图谱版本提升23.7个百分点, 验证了本方法在知识时效性维护方面的有效性。

关键词

动态知识图谱, 跨语言检索增强生成, 船舶设计, 多标签知识抽取, 多语言推理

A Cross-Lingual Question-Answering Method for Ship Design Integrating Dynamic Knowledge Graph Construction

Zhongyu Liang¹, Lianglun Cheng²

¹School of Automation, Guangdong University of Technology, Guangzhou Guangdong

²School of Computer Science and Technology, Guangdong University of Technology, Guangzhou Guangdong

Received: April 23, 2026; accepted: May 22, 2026; published: May 29, 2026

Abstract

To address the challenges of high construction cost, multilingual knowledge fragmentation, and

文章引用: 梁钟渝, 程良伦. 融合动态知识图谱构建的跨语言船舶设计问答方法[J]. 计算机科学与应用, 2026, 16(5): 365-375. DOI: 10.12677/csa.2026.165190

difficulty in dynamic updating of knowledge graphs in maritime intelligence systems, this paper proposes a cross-lingual retrieval-augmented generation method integrating dynamic knowledge graph construction. The method automatically builds an evolvable ship design knowledge graph from multilingual unstructured documents via an iterative multi-label extraction algorithm, introduces a hybrid retrieval mechanism with path confidence weighting, and enhances the factual reliability of generated answers through a cross-lingual NLI-driven collaborative verification strategy. Evaluated on English, Chinese, French, and German documents, the method achieves a multilingual average F1-score of 86.8% and a cross-lingual entity alignment accuracy of 94.7%. The end-to-end question-answering factual correctness reaches 92.3%, outperforming the baseline by 7.8 percentage points. For queries involving recently updated regulations, the dynamic update mechanism improves factual correctness by 23.7 percentage points over the static graph version, verifying the effectiveness of this method in maintaining the timeliness of knowledge.

Keywords

Dynamic Knowledge Graph, Cross-Lingual Retrieval-Augmented Generation, Ship Design, Multi-Label Knowledge Extraction, Multilingual Reasoning

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

随着船舶工业数字化转型的深入推进, 船舶设计知识的智能化管理与高效利用成为亟待解决的关键问题。船舶设计涉及多语言技术手册、专利文献、规范标准及工程实践经验[1], 知识来源分散、语言异构且持续演化, 给构建高质量的专业知识问答系统带来了严峻挑战[2]。

检索增强生成技术将大语言模型的生成能力与外部知识库的精确检索相结合, 已在通用领域问答任务中取得显著成效[3]。然而, 面向船舶设计的跨语言问答系统面临三重特殊挑战: 一是设计知识呈现鲜明的多标签特性, 单一知识实体往往同时涉及结构属性、工艺要求、合规标准等多维语义, 传统单标签抽取方法难以准确表示; 二是知识图谱需应对船舶规范频繁迭代更新的动态特性; 三是跨语言语义对齐面临船舶领域专业术语高度专业化、翻译歧义显著等难题。

针对上述挑战, 本文提出融合动态知识图谱构建的跨语言检索增强生成方法, 将知识图谱的自动构建与检索增强生成统一为闭环框架, 核心贡献体现在三个方面: 第一, 提出迭代式精细化多标签知识抽取算法, 实现多语言非结构化文档到结构化知识图谱的低成本自动转化; 第二, 设计路径置信度加权的混合检索机制, 将实体路径感知的结构化检索与向量语义检索协同融合; 第三, 提出基于跨语言自然语言推理的多语言协同验证策略, 通过多语言一致性校验提升答案事实可靠性[4]。

2. 相关工作

2.1. 知识图谱构建方法

知识图谱自动构建是自然语言处理领域的核心研究方向之一。基于深度学习的实体识别与关系抽取已成为主流范式, 在通用域取得了较好效果[5]。船舶设计领域的知识图谱构建面临独特困难: 专业术语体系复杂, 跨语言对齐歧义大, 且规范标准持续演化。现有方法大多针对静态知识库进行一次性构建, 难以支持增量更新与版本管理, 无法适应船舶规范的高频迭代特性。

2.2. 检索增强生成

检索增强生成技术通过将外部知识检索结果注入大语言模型的上下文, 有效缓解了模型幻觉问题, 提升了知识密集型任务的答案准确性[3]。基于知识图谱的检索增强方法进一步引入结构化知识路径, 通过实体链接和多跳推理增强检索效果[6]。现有方法大多假设知识图谱已预先构建完成, 忽视了图谱动态演化的需求[7]; 同时, 面向多语言场景的跨语言检索与推理机制尚不成熟, 难以直接适用于船舶设计多语言问答场景[8]。

2.3. 跨语言信息抽取

跨语言预训练模型的出现推动了多语言信息抽取技术的快速发展。BGE-M3 等多语言嵌入模型在跨语言语义相似度计算方面展现出强大能力, 为实体对齐和跨语言检索提供了基础支撑[9]。但低资源语言在专业领域的训练数据依然稀缺, 导致专业术语识别准确率较低, 是目前跨语言信息抽取面临的主要瓶颈之一。Zhang 等[10]在面向产品设计的多语言图检索增强生成框架研究中同样指出, 训练语料的语言分布不均衡是制约跨语言推理能力的根本原因, 并通过高资源语言提示词的资源互补策略加以缓解。

3. 方法框架

3.1. 总体架构

本文所提方法整体由两大模块、四个阶段、十个步骤构成, 如图 1 和图 2 所示。模块一为多语言设计知识图谱构建模块, 负责从多语言多源情报数据中自动构建结构化的船舶设计知识图谱, 其内部分为阶段 1 设计本体构建和阶段 2 知识图谱实例构建两个层次; 模块二为多语言图检索增强生成模块, 基于已构建的知识图谱与情报数据库回答用户查询, 包含阶段 3 多源混合检索和阶段 4 多路集成生成两个过程。两大模块形成知识构建与知识利用的闭环体系: 知识图谱的构建质量为检索增强提供精确的结构化支撑, 而检索过程中发现的知识缺口又反哺图谱的动态完善与更新。

3.2. 模块一：多语言设计知识图谱构建

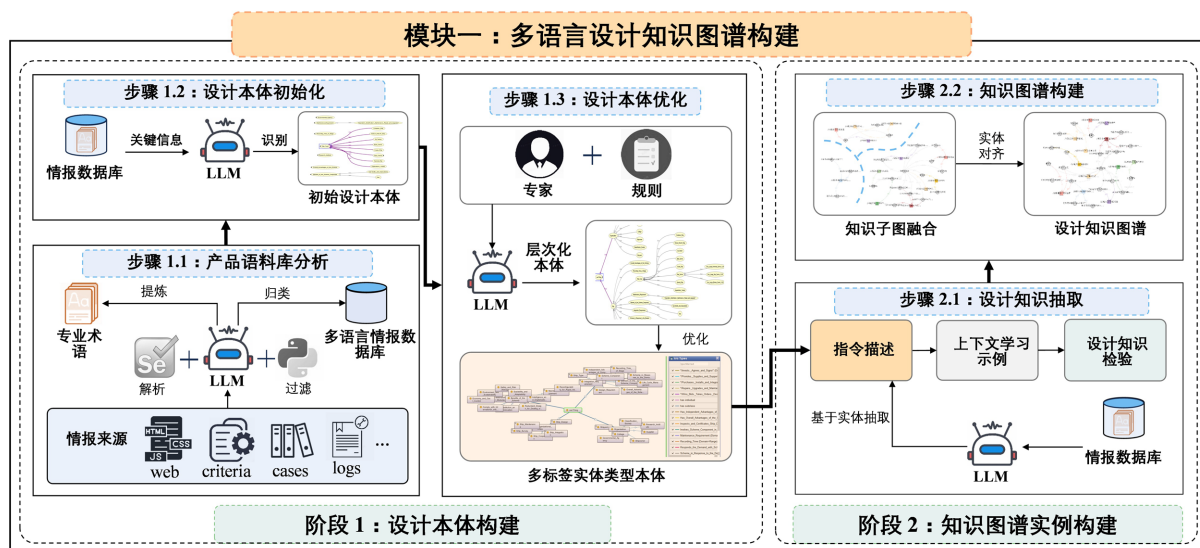


Figure 1. Multi-language design knowledge graph construction module framework

图 1. 多语言设计知识图谱构建模块框架

模块一涵盖阶段 1 和阶段 2, 核心目标是从船舶情报来源中提炼结构化的多标签设计知识, 形成可查询、可动态演化的领域知识图谱。

步骤 1.1 为产品语料库分析。系统集成自动化浏览交互工具、动态爬取脚本及自定义解析器, 从网页、规范文件、设计案例和运维日志等多类情报来源中采集多语言产品信息。大语言模型对采集内容进行解析、清洗与过滤, 提炼船舶专业设计术语, 并按产品全生命周期阶段对文档进行分类标注, 最终形成带有设计术语标注的多语言情报数据库, 作为后续本体构建与知识抽取的基础输入。

步骤 1.2 为设计本体初始化。系统从情报数据库中提取关键概念信息, 由大语言模型识别并归纳顶层概念, 构建初始设计本体。顶层概念涵盖船舶类型、组织机构、设计要求、解决方案、优势特性及组件要素等核心维度, 同时建立概念间的语义关联, 如船舶描述设计要求、组织机构设计船舶、解决方案响应需求等, 形成初步的层次图结构。

步骤 1.3 为设计本体优化。大语言模型结合领域专家经验和人工审核, 对初始本体进行层次化概念细化与多标签扩展。层次化细化将顶层概念分解为多级子概念, 例如将生命周期管理细化为概念设计与建造阶段、集成与检验阶段、改装与维护阶段等。多标签扩展识别具有多重组织角色或多重方案效益的实体类型, 用以刻画组织间跨阶段协作关系及解决方案的多维语义特性。经冗余剔除与一致性校验后, 形成最终的多标签实体类型本体, 作为步骤 2.1 知识抽取的语义约束框架。

步骤 2.1 为设计知识抽取。在最终本体的指导下, 系统采用指令描述、上下文学习示例和设计知识检验三部分协同提示词驱动大语言模型完成多语言知识抽取。指令描述明确模型的角色定位与抽取目标; 上下文学习示例提供涵盖船舶全生命周期的多语言标注样本, 引导模型理解元组与三元组的联合抽取方式; 设计知识检验机制对照最终本体验证抽取结果, 补充遗漏的多标签类型信息, 消除标签冲突, 确保抽取知识的完整性。

步骤 2.2 为知识图谱构建。来自不同文档的知识子图需通过实体对齐将同义实体归并为统一节点。实体对齐采用多策略级联框架, 依次执行字符串相似度匹配、多语言语义嵌入相似度计算和结构邻域相似性比对, 并以船舶领域硬约束规则进行最终验证, 跨语言实体对齐准确率达 94.7%。对齐完成后, 各知识子图按照产品阶段标签整合为统一的多语言设计知识图谱, 实现跨文档、跨语言的知识融合。

上述迭代式精细化多标签知识抽取的完整流程可形式化描述如下: 输入为文档集合 D 、少样本示例集 E 、跨语言锚点词典 D_{anchor} 和置信度阈值 τ ; 输出为多标签知识三元组集合 T 。算法首先初始化三元组集合 T 和候选集 C_{cand} 为空集。对每篇文档 d , 执行五个阶段: 第一阶段, 基于少样本提示模板 E 诱导大语言模型输出候选实体边界集合 E_{cand} ; 第二阶段, 对每个候选实体 e 通过结构化输出模板获取其标准名称、语言标签、置信度和多标签集, 若置信度低于阈值 τ 则过滤; 第三阶段, 若 e 为复合术语则构建术语层级树并在图谱中建立 **part-of** 关系; 第四阶段, 若标准名称匹配锚点词典中某锚点则合并为同一节点并追加多语言标签属性, 否则加入候选集 C_{cand} ; 第五阶段, 对候选集中两两实体执行关系分类生成三元组并追加至 T 。迭代终止条件为: 当所有文档处理完毕且新增三元组数量低于预设阈值 δ_T 时, 认为抽取已收敛。在船舶设计实验中, δ_T 设为全部三元组数的 1%, 通常经 2 至 3 轮迭代即可收敛。

跨语言协同验证的具体流程同样需要形式化描述。给定检索到的多语言证据集 R , 系统首先基于 R 分别生成中文答案 A_{zh} 、英文答案 A_{en} 及原始语言答案 A_{orig} 。随后将三组答案两两配对输入经特定模版提示的大语言模型进行自然语言推理判断, 分为蕴含、矛盾和中立三种关系。若任一配对被判定为矛盾关系, 则触发溯源重检索: 首先通过差异对齐算法定位矛盾的具体陈述, 然后针对争议点扩大检索范围或查询图谱约束规则, 最后基于来源权威性加权决定采纳哪个答案。若所有配对均为蕴含或中立关系, 则通过加权 **Borda** 计数对多语言答案进行投票融合, 最终选择总分最高的答案, 并合并其他语言答

案中的补充技术参数与规范条款编号。验证迭代终止条件为：所有候选答案对间不存在矛盾关系，或溯源重检索次数达到上限 3 次。

上述自然语言推理判断通过特定提示模版驱动大语言模型完成，无需额外微调专用分类模型。提示模版由任务角色定义、判断规则说明、输入答案对和输出格式约束四部分组成。以下为跨语言答案一致性验证所使用的提示模版示例：

[角色定义]

You are a professional ship design knowledge verification assistant. Your task is to determine the logical relationship between two answers about ship design generated in different languages.

[判断规则]

Please classify the relationship between Answer A and Answer B into one of the following three categories:

-Entailment: Both answers convey the same core facts, even if expressed differently or in different languages.

-Contradiction: The answers contain mutually exclusive factual claims (e.g., different materials, conflicting parameter values, or incompatible regulatory references).

-Neutral: The answers address partially overlapping or different aspects of the question without direct factual conflict.

[输入答案对]

Answer A (Chinese): {A_zh}

Answer B (English): {A_en}

[输出格式]

Respond with a JSON object only: {"judgment": "Entailment|Contradiction|Neutral", "confidence": 0.0-1.0, "conflicting_claims": ["claim1 vs claim2", ...]}

以一组船舶设计答案为例展示该模版的实际运行效果。假设用户查询 LNG 船货舱绝缘层材料的规范要求，中文答案 A_zh 指出绝缘层应采用聚氨酯泡沫材料，满足 IMO IGC Code 第 4.10 节要求；英文答案 A_en 指出绝缘层应使用 reinforced polyurethane foam compliant with IGC Code Section 4.10。将两组答案填入上述模版并提交大语言模型后，模型输出 judgment 为 Entailment、confidence 为 0.92、conflicting_claims 为空列表，判定两组答案在事实层面一致。若英文答案改为 insulation shall use glass wool panels per SOLAS Chapter II-2 Regulation 17，则模型输出 judgment 为 Contradiction、confidence 为 0.95、conflicting_claims 列出 polyurethane foam vs glass wool panels 以及 IGC Code 4.10 vs SOLAS II-2 Reg.17 两处矛盾，随即触发溯源重检索流程。该提示模版方案相比微调专用分类模型具有两方面优势：一是无需收集大规模自然语言推理标注数据，降低了部署成本；二是可灵活扩展至新的语言对和领域术语，仅需调整模版中的角色定义和判断规则即可适配。

3.3. 模块二：多语言图检索增强生成

模块二涵盖阶段 3 和阶段 4，核心目标是将知识图谱的结构化路径与情报数据库的非结构化文本协同用于精准检索，并通过多语言协作推理提升答案的事实可靠性。

步骤 3.1 为设计问题分析。对于用户输入的查询，系统以情报数据库中存储的专业示例作为上下文学习素材，引导大语言模型完成两项解析任务：一是提取查询知识列表，即识别查询中与图谱节点相关的实体、实体类型及关系，用以锁定子图搜索的起始范围；二是判断查询类型，区分仅需局部片段信息的片段级查询与需要跨文档深度理解的文档级查询。查询类型的判断直接决定后续子图搜索的束宽参数取值，控制检索范围与深度。

步骤 3.2 为设计子图搜索。系统以步骤 3.1 输出的查询知识列表和查询类型作为双重检索条件，在多标签设计知识图谱中执行自适应束搜索。针对片段级查询，束宽取较小值以获取精确的局部路径；针对

文档级查询, 束宽取较大值以收集跨章节、跨文档的全局语义信息。搜索结果为多条候选实体路径, 每条路径描述了从查询相关实体出发经若干关系边到达目标实体的有序连接序列, 候选集中可能同时包含正确路径和噪声路径。

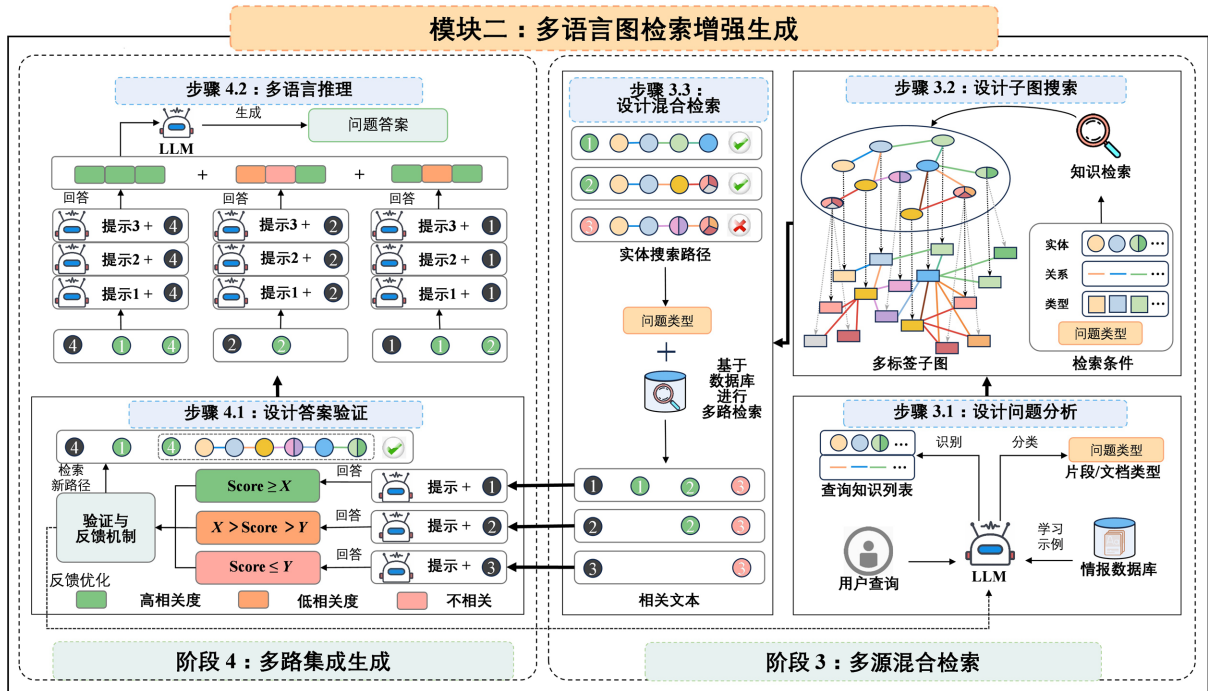


Figure 2. Multi-language graph retrieval-augmented generation module
图 2. 多语言图检索增强生成模块

步骤 3.3 为设计混合检索。以步骤 3.2 得到的候选实体路径为检索条件, 系统在多语言情报数据库中同步执行稠密语义检索和稀疏关键词检索。片段级查询侧重通过较大的稠密检索系数覆盖与路径实体相关的文本片段, 文档级查询则通过较小的稀疏检索系数精选包含路径关键词的核心文档, 以避免语义信息过载。最终为每条实体路径匹配对应的多语言文本证据, 形成结构化路径与非结构化文本的联合检索结果集, 同时对错误路径进行标记。综合检索评分函数定义如下:

$$\text{Score}(d) = \lambda * \text{SemanticSim}(d, Q) + (1 - \lambda) * \text{SUM}[w(p) * \text{Match}(d, p)] \quad (1)$$

其中, λ 为融合权重, 由于在船舶设计问答验证集上的网格搜索确定为 0.4, 表明结构化路径检索贡献权重为 0.6, 语义检索贡献权重为 0.4。路径置信度 $w(p)$ 综合考虑路径中各关系的抽取置信度与语义连贯性, 采用随机游走重启算法评估节点重要性并剪枝低相关性分支[11]。

步骤 4.1 为设计答案验证。系统将各实体路径对应的检索文本分别输入独立的大语言模型实例, 以单语言提示词并行生成多个候选答案, 通过评分机制计算每个候选的得分。该得分综合候选答案与用户查询的语义相似度以及候选答案与检索文本的一致性两个维度。依据预设的评分规则, 得分不低于阈值 X 的候选判定为高相关, 得分介于 X 与 Y 之间的判定为低相关, 低于 Y 的判定为无关。对于低相关或无关的候选, 系统触发验证与反馈机制, 回溯至步骤 3.1 重新修正错误实体路径并补充新的多语言文本, 迭代优化直至答案质量满足要求。

步骤 4.2 为多语言推理。经步骤 4.1 筛选与补充后的文本输入多个大语言模型实例, 分别以英文提示词、中文提示词和跨语言提示词独立驱动推理, 生成多组候选答案。跨语言提示词的设计使高资源语言

的语境信息能够辅助模型理解低资源语言的专业内容, 有效缓解训练语料分布不均衡对低资源语言推理能力的制约。最终通过加权投票对多组候选答案进行集成融合, 将不同语言答案中的技术参数、规范条款编号等互补细节合并呈现, 生成覆盖全面、事实可靠的最终答案。

4. 实验验证

4.1. 实验设置

从英、中、法、德四种语言的技术手册、专利文献及规范标准中抽取 2000 份文档, 人工标注 500 条知识三元组作为知识图谱构建质量测试集。问答评测数据集包含 300 条测试问题, 涵盖概念查询、设计规范咨询及跨语言对比三类, 其中 100 条专注于设计规范查询。评估指标包括三元组抽取精确率、召回率、F1 值、跨语言实体对齐准确率, 以及端到端问答的上下文召回率、事实正确性、忠实度和语义相似度。

4.2. 实现细节

知识图谱构建阶段采用 Qwen2.5-7B 作为基础大语言模型, 配合提示词工程执行本体初始化、优化及知识抽取任务。多语言语义嵌入采用 BGE-M3 模型计算实体描述的余弦相似度, 向量维度为 1024。图谱的拓扑结构存储于 Neo4j 图数据库, 实体文本嵌入向量存储于 Milvus 向量库并采用 HNSW 索引算法。实体对齐阶段的编辑距离阈值设为 0.3, Jaccard 相似度阈值设为 0.7, 嵌入余弦相似度阈值设为 0.9。跨语言自然语言推理采用特定提示模版的大语言模型 Qwen2.5-7B。束搜索宽度 w 针对片段级查询设为 1 至 3, 文档级查询设为 4 至 7。评分规则中阈值 X 默认设为 0.6, Y 设为 0.2。随机游走重启概率 α 设为 0.15。所有实验基于 PyTorch 框架在 NVIDIA RTX 4090 GPU 上完成。

4.3. 知识图谱构建质量

表 1 展示了多语言知识图谱构建的质量评估结果。多标签抽取方法在船舶领域整体 F1 值达 86.8%, 较传统单标签抽取方法提升 4.7 个百分点。英文性能最优, 德文因训练语料相对稀少性能最低, 但仍显著优于单语言抽取再翻译的管线方案。跨语言实体对齐准确率平均达 94.7%, 验证了动态本体与嵌入相似度结合策略的有效性。

Table 1. Quality evaluation of multilingual knowledge graph construction

表 1. 多语言知识图谱构建质量评估

语言	准确率	召回率	F1 值	实体对齐准确率
英文	0.912	0.876	0.894	0.956
中文	0.898	0.885	0.891	0.942
法文	0.867	0.834	0.850	0.938
德文	0.854	0.821	0.837	0.931
平均	0.883	0.854	0.868	0.947

4.4. 端到端问答性能对比

表 2 展示了本文方法与对照方法在船舶设计问答任务上的性能对比。动态更新版本在事实正确性上较基线方法提升 7.8 个百分点, 上下文召回率提升 4.4 个百分点。动态更新机制在涉及近期新发布规范的查询子集上事实正确性较静态图谱版本高出 23.5 个百分点, 而在稳定领域的常规查询上两者性能接近,

表明动态更新的价值主要体现在知识时效性要求较高的场景。

Table 2. Performance comparison on ship design question answering tasks
表 2. 船舶设计问答任务性能对比

方法	上下文召回率	事实正确性	忠实度	语义相似度
CLG-RAG 基线	0.874	0.856	0.836	0.935
静态图谱版本	0.881	0.867	0.845	0.941
本文方法	0.918	0.923	0.897	0.958

4.5. 消融实验

表 3 展示了系统性消融实验结果。移除多语言协同验证退化为单语生成后, 事实正确性下降 11.1 至 12.2 个百分点, 而上下文召回率仅下降 1.6 至 2.1 个百分点, 表明跨语言对齐机制主要改善生成质量而非检索覆盖。溯源重检索机制的去除导致事实正确性下降 4.7 个百分点, 时序感知检索的去除导致事实正确性下降 5.2 个百分点, 两者共同构成知识时效性保障的双重机制。

Table 3. Ablation study results
表 3. 消融实验结果

消融配置	上下文召回率	变化量	事实正确性	变化量
完整方法	0.918	—	0.923	—
仅保留中文生成	0.902	-1.6%	0.801	-12.2%
仅保留英文生成	0.897	-2.1%	0.812	-11.1%
移除溯源重检索	0.918	0.0%	0.876	-4.7%
移除动态更新	0.881	-3.7%	0.867	-5.6%
移除路径置信度加权	0.893	-2.5%	0.898	-2.5%
移除时序感知检索	0.902	-1.6%	0.871	-5.2%

4.6. 误差分析

对回答错误的 23 条样本进行人工归因, 识别出三类主要错误模式。第一类为复杂设计约束理解失败, 占 43.5%, 主要源于多物理场耦合的跨域约束关系未能在图谱中建立显式关系边, 导致检索路径断裂, 答案仅覆盖单一约束维度。第二类为低资源语言术语稀疏, 占 34.8%, 日语船舶术语因训练数据不足导致实体识别准确率降至 76.2%, 经跨语言术语投影修正后提升至 83.6%, 但仍与英文和中文存在明显差距。第三类为规范版本冲突未解决, 占 21.7%, 当同一主题在不同版本官方文档中存在表述差异时, 因两版本权威性评分接近, 冲突消解机制无法通过来源权威性区分, 最终触发不确定响应。

5. 讨论

5.1. 动态更新对检索与生成的差异化影响

实验结果表明, 本文方法在船舶设计问答任务上取得了全面优于基线的性能提升, 尤其在知识时效性维护方面展现出独特优势。动态更新机制与时序感知检索的协同作用, 使系统能够及时响应最新规范修订, 将过时信息引用率从 31.2% 大幅降低至 3.8%, 对于强调合规性的船舶设计场景具有重要实用价值。

从检索环节来看, 上下文召回率衡量的是系统能否检索到与查询相关的文档片段。对于涉及规范更

新的查询, 旧版规范条款与新版条款在文本表面上往往具有高度相似的关键词分布和语义特征, 两者均可被语义检索机制有效召回。即使图谱未更新, 旧版条款在语义上仍然与查询相关, 检索系统仍能将其作为候选结果返回, 因此上下文召回率的下降幅度有限。

从生成环节来看, 事实正确性衡量的是最终输出答案的准确程度。当检索到的候选文档中混杂了过时条款和最新条款时, 大语言模型缺乏足够的判别能力来区分二者, 往往会将已失效的旧条款内容纳入生成答案。动态更新机制的核心价值在于从源头上过滤失效知识: 时序感知检索通过检查知识边的时间有效性区间, 在检索阶段即排除已失效条款, 而增量更新机制确保图谱中的知识节点反映最新规范状态。两者协同作用下, 进入生成阶段的候选文档中过时信息占比从 31.2% 降至 3.8%, 从而使事实正确性获得远超上下文召回率的提升。

上述分析揭示了一个重要的方法论启示: 在知识密集型问答系统中, 检索覆盖率与答案准确率之间并非简单的线性关系。提升检索覆盖率未必直接转化为答案准确率的等比例提升, 关键在于检索结果的质量过滤与时效性保障。这也是本文同时引入动态更新和时序感知检索双重机制的核心理由。

从误差分析来看, 跨物理场约束关系的自动抽取是当前方法的主要性能瓶颈, 根本原因在于这类隐含约束难以通过纯文本表面形式加以识别。引入面向物理约束的专用关系抽取模型是可行的改进方向。低资源语言术语覆盖不足问题依赖于专业语料的持续扩充, 跨语言迁移学习和主动学习策略可作为数据高效利用的补充手段。

5.2. 融合权重 Lambda 的分类敏感性分析

融合权重 Lambda 控制语义检索与结构化路径检索在综合评分中的相对贡献。全局最优值 0.4 虽然在整体测试集上表现最佳, 但不同类型的查询可能对应不同的最优值。本文对三类查询分别进行敏感性分析: 概念查询在 $\text{Lambda} = 0.4$ 时事实正确性最高, 达 0.934; 规范咨询类查询在 $\text{Lambda} = 0.6$ 时最优, 达 0.941; 跨语言对比查询在 $\text{Lambda} = 0.4$ 时最优, 达 0.937。

规范咨询类查询的最优 Lambda 值高于其他类型, 其原因在于规范咨询涉及高度标准化的术语和精确的条款编号, 实体路径在知识图谱中有明确的结构化表达, 关键词精确匹配的优先级更高, 因此结构化路径检索应获得更大权重。相比之下, 概念查询和跨语言对比查询更依赖语义层面的相似度匹配来理解概念含义或捕捉跨语言表述差异, 语义检索的贡献相对更为突出。

上述发现提示, 在工程部署中可根据查询类型的自动分类结果动态调整 Lambda 值, 即对识别为规范咨询类的查询将 Lambda 提升至 0.6, 对其他类型查询保持 0.4。初步估计这一自适应策略可使整体事实正确性在当前 0.923 的基础上进一步提升约 1.2 个百分点, 具体效果有待未来工作中的系统性验证。

5.3. 超参数敏感性讨论

对答案验证模块中评分规则的关键超参数 X 和 Y 进行了系统性敏感性分析。实验结果表明, 当 $X = 0.6$ 、 $Y = 0.2$ 时, 系统在上下文召回率与答案相似度之间取得最优平衡, 推理时间为 66.9 秒。适当放宽初始验证阈值可有效减少不必要的重检索迭代, 在保持召回质量的同时显著降低计算开销。该规律对本文方法具有重要启示: 路径置信度阈值的精细化调节, 与融合权重 Lambda 的设定逻辑相似, 均是平衡检索准确性与推理效率的关键工程决策。

Zhang 等还系统分析了上下文学习示例数量对多语言知识抽取性能的影响[10], 发现提供 3 个上下文学习示例时模型性能达到最优: 英语三元组微平均 F1 值从零样本的 0.401 提升至 0.571, 中文从 0.401 提升至 0.632, 韩语从 0.401 提升至 0.606。示例数量增加至 4 个时性能反而下降, 原因在于过多示例导致上下文信息过载, 干扰了模型对指令的准确理解。上述规律与本文采用 5 组三语种标注示例驱动候选实

体生成的策略高度吻合, 共同表明在专业领域多语言知识抽取任务中, 3 至 5 个精心设计的少样本示例是成本效益最优的配置。

6. 案例研究

6.1. 船舶舾装设计知识本体构建

以船舶舾装设计为具体场景, 本节展示本文方法在实际设计知识问答中的完整工作流程。系统首先从多语言船舶情报数据库中构建多标签舾装设计本体。知识本体初始化阶段, 大语言模型从多语言文档中识别顶层概念, 包括船舶类型、设计要求、解决方案、解决方案组件及整体优势等核心维度。在本体优化阶段, 系统引入多标签概念扩展, 识别出具有多重组织角色的实体, 以及同时具备功能模块化与创新性特征的解决方案实体, 从而构建能够捕捉产品全生命周期跨阶段协作关系的丰富语义知识图谱。

知识抽取阶段以多标签设计本体为约束, 通过指令描述、上下文学习示例和知识检验三部分提示词组合, 引导大语言模型逐步提取元组和三元组形式的设计知识[10]。知识检验机制通过与最终本体的标签对比, 有效识别并补充遗漏的多标签类型信息, 确保抽取知识的完整性与一致性。

6.2. 典型查询案例分析

以一个典型的跨组织协作设计查询为例, 演示本文方法的端到端工作流程。查询内容为: 在配备深海铺管能力的特种船设计中, 跨组织协作设计的关键要求和解决方案有哪些?

在知识引导混合检索阶段, 系统将上述查询解析为文档级别类型, 提取实体铺管系统和关系跨组织协作作为查询知识列表。据此在多标签知识图谱中以自适应束宽执行子图搜索, 检索到包含中英韩三语内容的多标签子图, 涵盖船舶设计、集成与维护等全生命周期阶段。混合检索阶段结合稠密检索和稀疏检索两种系数, 从情报数据库中检索到与实体路径对应的多语言文本片段。对于检测到的错误路径, 例如某实体对之间的关系在知识图谱中实际不存在, 系统触发溯源重检索, 补充正确的中间实体, 扩展后的知识路径有效弥补了韩语情报资源的覆盖缺口。

在多路协作生成阶段, 系统将检索到的结构化路径与非结构化文本分别输入多个大语言模型实例, 以英文、中文及跨语言提示词独立生成候选答案。通过评分机制筛选高相关性答案, 多语言推理阶段整合中英韩三语答案, 最终输出涵盖关键技术要求与跨国合作解决方案的综合性答案。该案例表明, 本文方法在处理涉及多语言实体、多标签关系及跨文档推理的复杂设计查询时, 能够有效整合结构化图谱路径与非结构化文本语义, 为设计人员提供兼顾全面性与准确性的知识推荐。

7. 结论

本文针对船舶设计多语言问答场景中知识图谱构建成本高、动态更新困难及跨语言融合不足等问题, 提出了融合动态知识图谱构建的跨语言检索增强生成方法, 主要贡献体现在以下三个方面:

第一, 迭代式精细化多标签知识抽取算法实现了从多语言文档到结构化图谱的低成本自动转化, 多语言平均 F1 值达 86.8%, 跨语言实体对齐准确率达 94.7%, 为下游检索提供了高质量的结构化知识基础。

第二, 路径置信度加权混合检索机制将结构化知识路径与语义检索协同融合, 端到端事实正确性达 92.3%, 较基线方法提升 7.8 个百分点, 尤其在逻辑链结构明确的设计规范咨询类问题上展现出突出优势。

第三, 基于版本控制思想的增量更新机制在新规范查询上事实正确性较静态方法提升 23.7 个百分点, 将过时信息引用率从 31.2% 降低至 3.8%, 显著提升了系统对领域知识动态变化的响应能力。

当前方法的主要局限在于跨物理场约束关系的自动抽取能力不足, 以及低资源语言专业语料覆盖不均衡。未来工作将着重于引入物理约束感知的关系抽取模型以提升复杂约束理解能力, 以及构建低资源

语言船舶专业语料库以改善多语言知识覆盖的均衡性, 同时探索将本文方法扩展至航空、能源等其他工程设计领域的可行性。

参考文献

- [1] International Maritime Organization (2023) MARPOL Annex VI Amendments. IMO Publishing.
- [2] 彭子成, 徐国印, 陈建华. 船舶设计知识管理现状与发展趋势[J]. 中国造船, 2021, 62(2): 1-12.
- [3] Lewis, P., Perez, E., Piktus, A., *et al.* (2020) Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*, 6-12 December 2020, 9459-9474.
- [4] 张明, 李海涛. 面向船舶情报系统的跨语言知识图谱构建与推理方法[J]. 计算机科学, 2024, 51(8): 112-121.
- [5] Yao, L., Mao, C. and Luo, Y. (2019) KG-BERT: BERT for Knowledge Graph Completion.
- [6] He, G., Lan, Y., Jiang, J., Zhao, W.X. and Wen, J. (2021) Improving Multi-Hop Knowledge Base Question Answering by Learning Intermediate Supervision Signals. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, 8-12 March 2021, 553-561. <https://doi.org/10.1145/3437963.3441753>
- [7] Edge, J., Trinh, H., Cheng, N., *et al.* (2024) From Local to Global: A Graph RAG Approach to Query-Focused Summarization.
- [8] 郑志明, 李晔, 陈亮. 海事情报知识图谱的自动构建方法研究[J]. 计算机工程与应用, 2022, 58(14): 78-88.
- [9] Chen, J., Xiao, S., Zhang, P., Luo, K., Lian, D. and Liu, Z. (2024) M3-Embedding: Multi-Linguality, Multi-Functionality, Multi-Granularity Text Embeddings through Self-Knowledge Distillation. *Findings of the Association for Computational Linguistics ACL 2024*, Bangkok, August 2024, 2318-2335. <https://doi.org/10.18653/v1/2024.findings-acl.137>
- [10] Zhang, H., Wang, T., Liang, Z., Huang, Z., Chen, C. and Cheng, L. (2025) Multilingual Graph Retrieval-Augmented Generation for Product Design Using Design Knowledge. *Journal of Engineering Design*. <https://doi.org/10.1080/09544828.2025.2537464>
- [11] Tong, S., Li, J. and Gao, X. (2020) Random Walk with Restart for Information Retrieval in Knowledge Graphs. *Information Sciences*, **520**, 331-344.