

·业务研究·

地方非物质文化遗产知识图谱构建及其思政教育应用

张萌萌^{1,2}, 张矛矛¹

(1. 中国矿业大学, 江苏 徐州 221000; 2. 徐州工程学院, 江苏 徐州 221018)

摘要:【目的/意义】为推动非物质文化遗产领域的知识挖掘及其在高校思政教育创新应用,助力地方特色思政教育元素的资源整合及应用,提升非遗文化在当代思政教育中的应用效果和影响力。【方法/过程】本研究先搜集建构地方非物质文化遗产数据集,运用BERTopic主题模型提取实体类别与关系标签,再使用OneKE大模型进行实体抽取与关系识别,最后运用Neo4j图数据库进行知识图谱的可视化。并以国家级非遗徐州剪纸为例实证应用,探讨其在高校思政教育中应用路径。【结果/结论】研究发现:基于OneKE大模型的知识图谱构建方法能够清晰呈现地方非物质文化遗产的知识文化背景、技艺传承及内在关联,形成结构化的知识表达体系;基于构建的知识图谱能够协助挖掘思政元素,助力思政教学资源整合,提升地方性非遗文化在思政教育中的应用效果。【创新/局限】提出了一种基于大语言模型的地方非遗知识图谱构建方法,并挖掘思政元素用于思政教育。

关键词: 非物质文化遗产; OneKE大模型; BERTopic; 知识图谱; 思政教育

中图分类号: G122; G641

DOI: 10.13833/j.issn.1007-7634.2025.09.015

0 引言

非物质文化遗产(简称“非遗”)是中华民族独特的文化表达形式,是民族情感的纽带和文化认同的根基,承载着中华文明的智慧结晶和精神内涵,是历史长河中传承至今的生动见证,具有深厚的历史文化价值和教育意义。国务院在《关于进一步加强非物质文化遗产保护工作的意见》(以下简称《意见》)中明确提出,加大非物质文化遗产传播普及力度,将非物质文化遗产内容贯穿国民教育始终,构建非物质文化遗产课程体系和教材体系,鼓励非物质文化遗产进校园^[1]。《意见》为地方性非遗文化融入教育体系、提升文化传播效果提供战略方向。通过挖掘地方性非物质文化遗产知识,深入分析其与教育资源的创新整合,不仅为非物质文化遗产在教育中的价值转化提供全新思路,还为传统文化与思政教育的结合开辟新的发展路径。

知识图谱是利用图数据结构进行知识表示、组织与管理、挖掘利用的载体^[2],通过可视化呈现与系统管理,为特定领域的知识探索、智能应用与决策支持等提供强有力的工具。目前知识图谱在多个领域得到了广泛应用,非物质文化遗产领域便是其中之一,面向非遗文化传承的知识图谱,系统化呈现非遗项目的多维知识及其关联结构,从而为非遗的数字化传播与挖掘利用提供有力支持。鉴于此,本研究基于

知识图谱构建非物质文化遗产知识结构与关联,挖掘其内含的思政元素和教育价值,为非物质文化遗产数字化保护和教育应用提供技术方法。以期提升非遗在高校思政教育场景中应用深度,为思政教育注入更鲜活的文化资源与活力。

1 相关文献综述

1.1 非遗知识图谱构建及应用研究

近年来学者们运用本体、语义网、深度学习等不同方法对非遗进行知识图谱构建。在运用本体方法知识图谱构建方面,王常珏等以元曲四大家现存杂剧为例,通过自顶向下的方式对元曲知识进行抽取、融合、存储,运用本体构建知识图谱以实现元曲知识的多维度展示与应用^[3]。王左戎等按照本体的七步法知识建模流程提出中国传统戏曲知识图谱模式层构建方案,同时利用Protégé和Neo4j工具进行了图谱模式层的具体应用展示^[4]。在语义网应用方面,张卫等结合知识图谱在文本语义关联解析与数据语义链接层面的技术特质,引入自然语言处理技术对非遗文本内人文性知识进行语义关联解析,同时在语义融合的基础上推动数据的语义链接与知识服务^[5]。随着深度学习技术的发展,部分学者开始采用深度学习方法进行领域非遗知识图谱的构建。汪琳等提出了融合学习扩展和深度学习的非遗陶瓷工艺领域术

收稿日期:2025-01-01

基金项目:国家社会科学基金后期资助项目“中国古代辞书相关体育汉字整理与研究”(22FTYB002)。

作者简介:张萌萌(1987-),女,江苏徐州人,博士研究生,助理研究员,主要从事思想政治教育、高等教育管理等研究;张矛矛(1981-),女,江苏徐州人,教授,博士生导师,主要从事思想政治教育、高等教育管理等研究,通讯作者:4872@cumt.edu.cn。

语抽取及新词发现方法,并将其应用于非遗陶瓷工艺的知识图谱构建和检索服务^[6]。

此外,还有部分学者探讨了非遗知识图谱的应用问题,分析非遗知识图谱在领域知识组织、问答^[7]、文化保护与传承、媒介传播^[8]、旅游与教育领域的使用,展现出多场景的适用性。例如刘清堂等基于知识图谱技术设计了民族器乐文化的智能服务系统框架,从知识管理、智能问答、检索优化到推荐服务等多方面探讨数字化保护与传承的具体实践,充分体现非遗数据的个性化应用潜力^[9]。

1.2 非遗在思政教育中应用相关研究

非遗文化蕴含深厚的教育价值,是推动高校思政教育内容多样化的重要资源,融入高校思政教育,能够有效培养学生的传统文化理念与责任意识,同时增强文化自信,助力学生形成正确的价值观^[10]。从文化历史和现实意义的视角指出,非遗作为中华民族文化的积淀,不仅在精神文化传承中具有重要地位,也为思政教育提供了丰富的资源和深刻的教育内涵^[11];黄伟^[12]则通过少数民族地区的教育实践证明,将非遗文化引入思政课程,能够显著提升学生的文化认同感与民族自豪感,凸显出非遗文化在不同区域教育场景中的独特优势。

非遗与高校思政课程结合在实践路径上呈现出多样化特征。例如孙鉴^[13]通过系统思维视角分析非遗文化与高校思政教育的融合策略,提出从整体性、动态性与关联性多维度优化非遗内容的课程设计方案;张晴^[14]建议高校应结合现代技术手段创新非遗课程形式,拓展教学载体并增强师资力量,从而实现文化资源与教学内容的深度整合。非遗在课程思政中的功能不仅体现为内容扩展,还在于文化认同与产教融合方面。例如,郭瑾莉等^[15]指出非遗资源在高校课程思政中的应用契合文化强国建设的时代需求,有助于学生文化认同的提升和民族精神的塑造。

1.3 研究述评

现有针对非遗知识图谱构建与应用研究为非遗资源的数字化保护和多领域应用提供了理论基础,然而,现有知识图谱构建研究多依赖基于规则模板、本体、机器学习等技术方法,其在识别复杂关系类型时存在一定局限,导致对非遗数据中知识深层次语义的挖掘与表示能力不足。而大语言模型(Large Language Model, LLM)通过深度学习技术对语言数据中的复杂模式进行学习后,利用其语义识别和复杂关系提取能力,可灵活适应多平台异构数据,提升知识图谱构建的精度和广度,为非遗知识图谱构建提供更全面的技术支持^[16]。此外,当前对于非遗知识图谱应用,大部分学者是面向知识组织管理、智能问答、文化传播等领域,面向教育特别是思政教育领域应用较少。地方非物质文化遗产中存在丰富思政元素和具有重要价值的文化知识,能够为高校思政教育提供相关教育资源和案例知识。

基于此,本文以地方性非物质文化遗产的多源语料为研究对象,提出一种基于OneKE大模型和Neo4j图数据库的知识图谱构建方法,并进一步探索非遗知识图谱在高校思政教育中创新应用路径。

2 非物质文化遗产知识图谱的构建流程

首先,基于非物质文化遗产数字资源网、国家数字图书馆以及百度搜索、知乎等平台查阅搜集和编写爬虫程序得到非遗数据集。然后,运用OneKE大模型和Neo4j构建地方非物质文化遗产知识图谱。最后,基于构建的知识图谱探讨其在高校思政教育的创新应用路径,如图1所示。

2.1 非遗数据采集与预处理

通过非物质文化遗产相关文献的查阅搜集,并结合使用

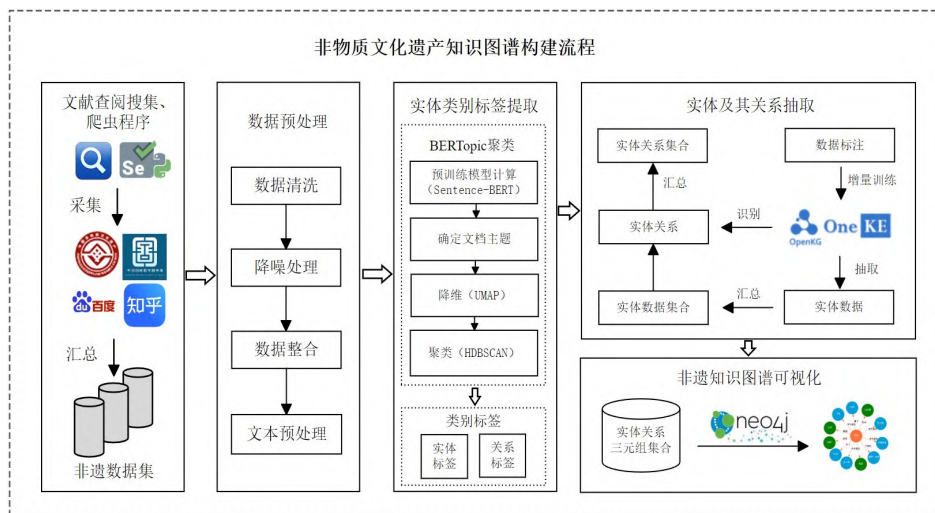


图1 非遗知识图谱的构建流程

Figure 1 The process of constructing the knowledge graph of NRM

Scrapy 和 Selenium 爬虫工具,从非遗数字资源网、国家数字图书馆、百度搜索、知乎等平台获取地方非遗相关数据,汇总得到非遗数据集,涉及学术文章、政策文件、书籍报刊资料、非遗介绍文本等多种类型,内容覆盖非遗的历史背景、工艺特点、代表性传承人和文化价值多个维度,以确保数据源的广泛性和多样性。经过数据清洗、降噪处理、数据整合、文本预处理等步骤进行预处理,清理并规范化文本数据。其中,数据清洗主要用于去除重复记录和缺失值,确保数据的完整性;降噪处理通过过滤网页广告、无效内容及短文本,剔除无关信息以减少干扰;数据整合则统一不同来源平台的数据字段格式,规范数据结构。

2.2 基于大语言模型的非遗知识图谱构建方法

首先利用 BERTopic 主题模型进行主题分析,在各主题下概括实体的类别标签与关系标签;然后,从非遗数据集中选取部分数据进行实体标注与关系标注,微调训练 OneKE 大模型;随后通过增量训练后的 OneKE 大模型提取汇总实体数据集合,并识别实体间的关系,汇总有效的实体关系三元组结构;最后,将提取的实体及关系三元组导入 Neo4j 数据库进行可视化,使地方性非物质文化遗产的知识结构清晰呈现。

2.2.1 BERTopic 提取实体类别标签

BERTopic 是一种主题抽取模型,结合 Transformer 和基于类别的 c-TF-IDF 技术,通过构建密集的主题集群,不仅能够保留各主题的核心词汇,还能提升主题的可解释性,BERTopic 的主题提取流程如图 2 所示。在降维后,BERTopic 使用 HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) 算法对文档嵌入进行聚类,HDBSCAN 是一种基于密度的聚类方法,能够在无需预定义聚类数量的情况下自动确定聚类数目,同时具有较强的抗噪声能力,从而保证了主题聚类的稳健性。在计算出每个簇的主题词后,BERTopic 会利用最大边界相关(MMR, The Maximal Marginal Relevance)算法对主题进行筛选与优化,删除不相关主题并保留最具代表性的主题。此外,通过调整相似度阈值合并相近的主题,并设置阈值控制主题数量,从而进一步优化模型的可解释性和精简性。

本研究利用 BERTopic 的主题聚类方法,对所有非遗文本数据进行主题分析,以全面挖掘非遗数据的语义特征和主题结构,从而提取出核心实体类别标签及其关系类型。BERTopic 通过将文本转化为语义嵌入向量,在向量空间中表示文本的语义特征,并基于无监督聚类方法对具有相似语义的文本进行聚类归并,形成多个主题。这些主题不仅能够揭示非遗数据的主要内容方向,还为实体及关系的进一步归纳与分类提供基础,结合每个主题的内部语义特征,提取具有代表性的实体类别标签,并概括这些类别之间的关系类型,这些类别将为后续的数据标注以及实体关系抽取提供明确的类别标签。

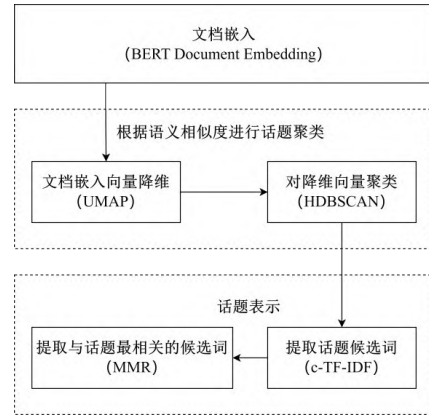


图2 BERTopic 主题提取流程

Figure 2 BERTopic topic extraction process

2.2.2 OneKE 大模型

OneKE 大语言大模型能够基于 Schema 实现泛化信息抽取,提高结构化信息提取的泛化能力。其设计特点在于对多样化数据的适应性,在中英文信息抽取任务中表现出色。此外,通过难负样本采样和 Schema 轮询式指令构造技术,OneKE 能够有效解决数据格式不统一、噪音干扰以及数据多样性不足等问题,这些技术提升了模型在零样本任务上的性能,使得 OneKE 即使在未见过的数据上也能实现较高的精度和召回率,特别是在实体抽取任务中,OneKE 的性能超越许多其他的模型,使其在复杂关系分析中展现出显著优势。成为实体及实体关系抽取任务中的理想选择。

为进一步提高 OneKE 大模型在特定知识领域进行实体关系抽取的通用性,本文将在各平台采集的非遗数据集中选择部分平台的非遗数据,进行标注后微调 OneKE 模型。非遗数字资源网、国家数字图书馆中的文献资料通常具备较强的专业性,能反映出非物质文化遗产中历史背景、传承者和文化价值等详细信息。因此,借助该部分平台中的数据对模型进行微调,增强 OneKE 在处理类似内容时的表现能力,从而提升其在非遗知识领域的抽取效果。此外,这种多层次的数据也有助于模型适应地方性非遗相关的复杂知识领域场景,提高对实体细节和知识关系导向的捕捉精度。

2.2.3 实体及实体关系的抽取

OneKE 大模型实现实体及实体关系抽取效果与 Prompt 紧密相关。因此,参考 OneKE 源码的 github 网站,分别对 OneKE API 的系统角色及提示词进行基础设置,如下所示:

```
system_role = '''<<SYS>>You are an expert in extracting the elements of the knowledge graph (entities, entity relationships) in the documentation of intangible cultural heritage. 假如你是一名非物质文化遗产研究专家,具备知识图谱构建技术,拥有丰富的抽取非物质文化遗产文献资料中知识图谱要素(实体、实体间关系)经验。<</SYS>>''' #
```

系统角色设置:

```
sinstruct = {'instruction': '', 'schema': [], 'example': [], 'input': ''}
```

```
prompt = '[INST]' + system_role + sinstruct + '[INST]' #
```

OneKE基本提示词

system_role变量定义OneKE模型的“系统角色”,即在实体及实体关系抽取任务中扮演领域专家角色,通过该角色设定,模型能够在处理文本时集中于指定领域的知识点,增强对任务的上下文理解。sinstruct变量是OneKE模型的基本提示词结构,包含四个关键部分:instruction、schema、example和input,其中,instruction提供了具体的任务指令,指导模型关注特定的信息抽取需求;schema用于定义实体和实体关系的结构,即实体类别和实体关系类型,有助于OneKE模型在抽取信息时保持结构化一致性;example提供具体的示例,以示范模型如何处理类似任务,提高信息提取的准确性;input则是待处理的文本输入,保证输入数据标准化,便于模型理解和抽取信息。将system_prompt和sinstruct内容整合为一个完整的提示词Prompt,使模型能够准确执行NER及RE的任务,并在非物质文化遗产领域分析相关数据并提取实体及关系。

3 应用研究—以徐州剪纸为例

3.1 数据获取与清洗

本研究通过人工搜集与自动化爬虫结合的方式采集数据。在非物质文化遗产网和国家数字图书馆,检索相关的文献资料共收集87篇学术论文,15份政策文件与规划文献,以及19篇报刊文章,涉及徐州剪纸的历史背景、技艺传承与文化价值,保护与传承政策等。在百度搜索和知乎平台,使用Selenium爬虫工具自动抓取与“徐州剪纸”相关文本内容。从百度搜索中以“徐州剪纸”为关键词采集网页数据,共计691页,涉及徐州剪纸的相关新闻报道、活动、展览和个人博客等内容,以及从知乎平台的问答栏目获取171个相关问题及其对应的4280条用户答案,12篇专栏文章。

数据采集完成后,对原始数据集进行数据清洗、降噪等数据预处理。在数据内容整合的过程中,对来自不同平台的数据字段进行统一规范:将学术论文、政策文件、报刊文章、专栏文章以及网页中的百科专业性强的内容汇总为134篇正式文献,网页博客及知乎中的用户问答专业性弱的内容汇总为4970条用户生成文本内容;进一步去除冗余字符、分词处理和停用词过滤,简化文本结构,并通过标准化方式对日期、地点等信息进行修正。经过系统化的数据预处理,获取

到有关于徐州剪纸的非遗数据集包括134篇正式文献与4656条用户生成内容。

3.2 实体抽取与特征概括

3.2.1 实体及关系类别标签提取

运用BERTopic模型进行文本聚类,从不同主题中以获取实体及关系的类型标签。经实验测试发现主题数量为4时,聚类效果最优。确定nr_topics=4,再训练BERTopic主题模型后使用主题条形图直观展示每个主题维度下的一组关键词及其c-TF-IDF权重值,如图3所示。

依据主题提取的结果及可视化呈现,根据各个主题维度下的主题词可概括出四个维度的主题名称,分别为文化要素、相关人物、作品成果、事件活动,各主题分别对应于四种实体的类别。在每类主题下分别根据c-TF-IDF权重值进行主题词筛选,获取到最常出现的五种关系类型:传承、属于、涉及、创作以及参与,具体四种实体类别与关系类型的逻辑连接表格如表1所示。

表1 实体类别与关系类型的逻辑连接表格

Table 1 Table of logical connections between entity categories and relationship types

主体类别	关系类型	客体类别
相关人物	传承	文化要素
作品成果	属于	文化要素
文化要素	涉及	事件活动
相关人物	创作	作品成果
相关人物	参与	事件活动
事件活动	涉及	作品成果

3.2.2 OneKE模型的微调训练

(1)数据集的标注

微调OneKE模型所使用的数据集来源于34篇正式文献与1000条用户生成内容,对其中的实体以及实体之间的关系进行标注和分类,经过统计,标注后的实体及实体关系的数量分别为2102、2673条,部分标注结果如表2所示。

(2)模型结果评价

选取精确率(Precision)、召回率(Recall)以及它们的综合指标F1值(F1-Score)应用于评估OneKE大模型的效果。OneKE大模型微调训练过程的损失率变化曲线如图4所示,在模型的微调过程中,训练损失Train Loss随着训练批次Batch的增加呈现出快速下降后趋于平稳的趋势。从图4中



图3 BERTopic聚类结果主题条形图

Figure 3 Thematic bar chart of BERTopic clustering results

表2 实体及实体关系标记结果样例

Table 2 Sample results of entity and entity relationship labeling

待标注内容	实体标注结果	关系标注结果
徐州剪纸广泛流传于江苏省徐州各县(市)区。地处苏、鲁、豫、皖四省交界处的徐州是南北文化交融地带,素有“五省通衢”之称。处于这种特殊地理环境中的徐州剪纸集南北之长,作品中既融入了北方剪纸的豪放和浑厚,又糅合了南方剪纸的细腻和清丽。	[{"entity": "徐州剪纸", "entity_type": "文化要素"}, {"entity": "江苏省徐州", "entity_type": "文化要素"}, {"entity": "北方剪纸", "entity_type": "文化要素"}, {"entity": "南方剪纸", "entity_type": "文化要素"}]	[{"head": "徐州剪纸", "relation": "属于", "tail": "江苏省徐州"}, {"head": "徐州剪纸", "relation": "涉及", "tail": "北方剪纸"}, {"head": "徐州剪纸", "relation": "涉及", "tail": "南方剪纸"}]
徐州剪纸有两种主要风格,一种以邳州剪纸为代表,作品构图紧凑,疏密得当,结构粗犷浑厚,风格朴实无华,喜用变形夸张等手法,显得大胆泼辣。另一种以徐州市区吴国本、张丽君及沛县敬安镇的剪纸为代表,作品风格灵秀俊美、细腻流畅、精巧玲珑。邳州粗犷风格的剪纸是徐州剪纸的主要代表,它精于展现地方现实生活和民风民俗,题材内容丰富多样,保持了原生艺术纯正品格,洋溢着浓厚人文气息,具有很高的艺术价值。	[{"entity": "徐州剪纸", "entity_type": "文化要素"}, {"entity": "邳州剪纸", "entity_type": "文化要素"}, {"entity": "吴国本", "entity_type": "相关人物"}, {"entity": "张丽君", "entity_type": "相关人物"}, {"entity": "沛县敬安镇", "entity_type": "文化要素"}, {"entity": "灵秀风格剪纸", "entity_type": "作品成果"}, {"entity": "粗犷风格剪纸", "entity_type": "作品成果"}, {"entity": "地方现实生活", "entity_type": "事件活动"}]	[{"head": "邳州剪纸", "relation": "属于", "tail": "徐州剪纸"}, {"head": "粗犷风格剪纸", "relation": "属于", "tail": "邳州剪纸"}, {"head": "吴国本", "relation": "创作", "tail": "灵秀风格剪纸"}, {"head": "张丽君", "relation": "创作", "tail": "灵秀风格剪纸"}, {"head": "粗犷风格剪纸", "relation": "涉及", "tail": "地方现实生活"}, {"head": "粗犷风格剪纸", "relation": "涉及", "tail": "民风民俗"}]
2023年4月15日上午,金秋晨曦颐养院举办了非遗剪纸文化体验活动,特邀徐州剪纸非遗传承人董志强老师展示中国传统剪纸艺术之精华,让大家近距离感受非物质文化遗产的独特魅力。	[{"entity": "金秋晨曦颐养院", "entity_type": "事件活动"}, {"entity": "体验活动", "entity_type": "事件活动"}, {"entity": "董志强", "entity_type": "相关人物"}, {"entity": "徐州剪纸", "entity_type": "文化要素"}]	[{"head": "董志强", "relation": "传承", "tail": "徐州剪纸"}, {"head": "董志强", "relation": "参与", "tail": "体验活动"}, {"head": "金秋晨曦颐养院", "relation": "涉及", "tail": "体验活动"}]

可以观察到,模型训练初期前20个批次中,损失值迅速下降,从8.3下降至接近2.5,表明模型在早期阶段快速学习到了输入数据中的主要特征。最终收敛至接近0.4左右,从曲线的平稳趋势可以确定,模型在当前设置下未出现明显的过拟合现象,优化器AdamW在训练过程中的表现较为稳健,说明学习率和批量大小的设置合理。模型在训练后期逐步优化并达到较高的拟合程度,具有较好的拟合效果,为进一步的实体抽取及关系识别提供可靠基础。

为进一步检验OneKE大模型的抽取效果,选取OneKE官方提供的其他开源大模型进行对比,包括LLaMA2-13B-Chat、BaiChuan2-13B-Chat、LLaMA3-8B-Instruct以及Qwen1.5-14B-Chat,以上模型均已在多个通用数据集上接受广泛的指令微调,同时与OneKE模型在相同训练配置下微调评估,因此均获得一定的在非遗领域中特定的信息抽取能力。评估结果如表3所示。实验结果表明,在实体抽取和关系识别两个任务上,OneKE模型的P值、R值和F1值均优于其他对比模型。具体而言,OneKE在实体抽取任务中的F1值达到0.8931,在关系识别任务中的F1值为0.7520,均显著高于

LLaMA、Qwen等模型的最高表现值。相比之下,其他模型在P值与R值的表现上存在一定的波动,而OneKE在这两项指标上均表现出较高的稳定性,例如LLaMA2-13B-Chat在关系识别任务中的F1值仅有0.3854,这相较于实体抽取显著偏低,在关系识别中的稳定性差,因此OneKE在提取领域特定知识时能够保持全面且均衡的表现,更适合用于地方性非物质文化遗产知识图谱构建任务。

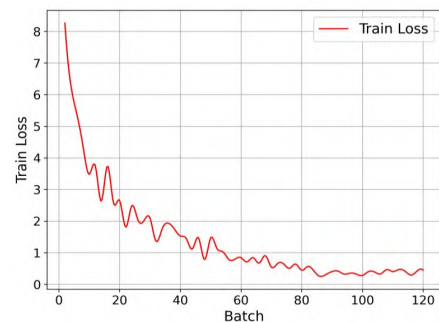


图4 微调训练损失率曲线

Figure 4 Fine-tuned training loss rate curve

表3 OneKE及对比模型评估结果汇总

Table 3 Summary of OneKE and comparative model evaluation results

模型名称	实体抽取			关系识别		
	P值	R值	F1值	P值	R值	F1值
LLaMA2-13B-Chat	0.8972	0.7059	0.7901	0.4217	0.3549	0.3854
BaiChuan2-13B-Chat	0.8571	0.7396	0.7940	0.5226	0.3534	0.4217
LLaMA3-8B-Instruct	0.8889	0.7059	0.7869	0.7159	0.5431	0.6176
Qwen1.5-14B-Chat	0.8962	0.6985	0.7851	0.6049	0.4224	0.4974
OneKE	0.9286	0.8603	0.8931	0.8342	0.6845	0.7520

3.2.3 实体及关系抽取结果

为实现更全面的非物质文化遗产徐州剪纸的实体及实体关系抽取,针对实体及关系抽取任务设计 sintruct 提示词结构模板,以优化 OneKE 模型在其中的信息提取表现。针对实体的抽取, sintruct 中的 instruction 参数部分为模型提供明确的任务指引,要求其关注并抽取指定的实体类别; schema 参数定义地方性非遗徐州剪纸中所涉及的四类实体,并通过详细说明以帮助模型识别不同类型的关键信息; example 参数部分通过提供具体的实体抽取的示例,使模型能够参考相似情境下的抽取方式,以提高抽取的准确性和一致性; input 参数则作为实际待处理文本,将实体数据标准化输入, sintruct 变量详细结构见表4。

实体关系的抽取的 sintruct 提示词结构模板如表5所示,通过 schema 参数,模型能够识别出实体中的五种关系类型。同实体抽取任务中 sintruct 提示词的 example 参数类似,借助 example 中的示例关系三元组,模型可以参考实体抽取任务,提高识别的可靠性与逻辑性。

整合 system_prompt 和 sintruct 内容,形成完整的任务指引,使模型能够更精确地提取复杂文献以及用户生成内容中的实体与关系,构造实体及关系抽取任务的 Prompt。通过利用实体以及实体关系抽取的 Prompt,将正式文献逐段以及知乎用户答案逐条输入到微调训练后的 OneKE 大模型进行实体抽取及关系识别,并导出抽取结果,最终共获得4126个实体和6185条实体关系对,部分实体的抽取结果如表6所示,实体之间关系的抽取结果如表7所示。

3.3 知识图谱的可视化呈现与分析

利用提取的徐州剪纸相关实体及关系集合,进行实体融合与消歧并去除含义相同的关系对后。选取数量排名前20的实体的利用 Neo4j 软件绘制面向徐州剪纸地方性非物质文化遗产的知识图谱,如图5所示。以剪纸相关实体为节点,紫色、橙色、蓝色、红色分别代表文化要素、相关人物、作品成果、事件活动类别的实体,实体之间的关系名称也均在节点间的连线上。

表4 实体抽取的 sintruct 变量
Table 4 Sintruct variables for entity extraction

参数	内容示例	描述
instruction	"假如你是一名非物质文化遗产研究专家,具备知识图谱构建技术,拥有丰富的抽取非物质文化遗产文献资料中知识图谱要素经验。请从 input 中抽取符合 schema 定义的实体,不存在的实体类型返回空列表。"	为模型定义具体任务,引导其抽取指定类型实体
schema	{"文化要素": "文化要素指的是在徐州剪纸这一非物质文化遗产中所蕴含的特有艺术风格、技法、图案及表现形式", "相关人物": "相关人物指的是在徐州剪纸的创作、传承、研究及推广过程中发挥重要作用的个体", "作品成果": "作品成果是指由徐州剪纸艺术家创作的具体剪纸作品,包括各类图案、形状和风格的剪纸艺术品等", "事件活动": "事件活动是指围绕徐州剪纸展开的各类文化活动,包括但不限于展览、比赛、研讨会等"}	定义抽取的实体类别及其含义,确保模型识别实体类别的准确性
example	[{"input": "2023年4月15日上午,金秋晨曦颐养院举办了非遗剪纸文化体验活动,特邀徐州剪纸非遗传承人董志强老师展示中国传统剪纸艺术之精华,让大家近距离感受非物质文化遗产的独特魅力。", "output": {"文化要素": ["徐州剪纸"], "相关人物": ["董志强"], "作品成果": [], "事件活动": ["金秋晨曦颐养院", "体验活动"]}}]	提供抽取示例,使模型参考同类任务并提高抽取一致性
input	"徐州剪纸代表性传承人技艺日益精湛,参加各级比赛多次获奖,有的已经成立自己的剪纸工作室,把剪纸作为了自己毕生追求的艺术。"	实际输入的文本内容,确保模型在预设情境下进行抽取

表5 关系抽取的 sintruct 变量
Table 5 Relationally extracted sintruct variables

参数	内容示例	描述
instruction	"假如你是一名非物质文化遗产研究专家,具备知识图谱构建技术,拥有丰富的抽取非物质文化遗产文献资料中知识图谱要素经验。请从 input 中抽取符合 schema 定义的关系三元组,不存在的关系返回空列表。"	为模型定义任务指令,集中于关系三元组的抽取
schema	{"传承": "传承关系指的是剪纸技艺、文化元素及其价值观的代际延续", "属于": "属于关系是某一具体的剪纸作品或艺术创作归属于徐州剪纸这一特定文化体系,包括但不限于时间、地点及特征等", "涉及": "涉及关系表现为徐州剪纸的不同要素之间的关联与互动,可以是人与人、人与物、物与物,强调不同元素间的交织与互联", "创作": "通过剪纸技法和创新思维,创作出具有独特风格、艺术价值和文化内涵的作品", "参与": "徐州剪纸的各类活动中的关系,可以是创作、展示、比赛以及推广等"}	规定关系类型,便于模型识别并抽取符合结构的关系类型
example	[{"input": "2023年4月15日上午,金秋晨曦颐养院举办了非遗剪纸文化体验活动,特邀徐州剪纸非遗传承人董志强老师展示中国传统剪纸艺术之精华,让大家近距离感受非物质文化遗产的独特魅力。", "output": {"传承": [{"subject": "董志强", "object": "徐州剪纸"}], "参与": [{"subject": "董志强", "object": "体验活动"}], "涉及": [{"subject": "金秋晨曦颐养院", "object": "体验活动"}]}}]	提供关系抽取示例,作为模型的参考案例
input	"徐州剪纸展示由“勤政廉政”和“美好现代生活”两个主题组成,共有作品30件。这些剪纸作品形式多样,有鞋面花、窗花、顶棚花、盆花、枕花、帐花、灯花等;主题鲜明,涉及历史、戏剧人物、神话传说、动植物、花鸟鱼虫、生活场景、田园风光等,反映了在中国共产党的领导下,祖国的巨大变化和人民群众的美好生活。"	待提取的实际文本输入,用于关系的实际抽取

表6 部门实体抽取结果
Table 6 Department entity extraction results

序号	抽取实体结果
1	{"文化要素":["徐州剪纸","造纸工艺","刺绣技术","南方剪纸"],"相关人物":["王桂英","吴国本","孟宪云","张丽君","黄继婷"],"作品成果":[],"事件活动":["报春","祭奠风俗"]}
2	{"文化要素":["徐州剪纸","窗花","门笺"],,"相关人物":[],"作品成果":["枕头花","帽子花","鞋花样"],"事件活动":["节日庆典"]}
3	{"文化要素":["徐州剪纸"],"相关人物":["王桂英"],"作品成果":["窗花","鞋花","洗床单","农村生活剪影","春耕秋收"],"事件活动":["剪纸培训班"]}

表7 部分实体关系对抽取结果
Table 7 Partial entity relationship pair extraction results

序号	内容
1	{"传承":[{"subject": "王桂英", "object": "徐州剪纸"}, {"subject": "吴国本", "object": "徐州剪纸"}, {"subject": "孟宪云", "object": "徐州剪纸"}, {"subject": "张丽君", "object": "徐州剪纸"}, {"subject": "黄继婷", "object": "徐州剪纸"}], "属于":[{"subject": "徐州剪纸", "object": "南方剪纸"}], "涉及":[{"subject": "徐州剪纸", "object": "造纸工艺"}, {"subject": "徐州剪纸", "object": "刺绣技术"}], "创作":[""], "参与":[""]}
2	{"传承":[""], "属于":[{"subject": "枕头花", "object": "徐州剪纸"}, {"subject": "帽子花", "object": "徐州剪纸"}, {"subject": "鞋花样", "object": "徐州剪纸"}], "涉及":[{"subject": "徐州剪纸", "object": "节日庆典"}, {"subject": "徐州剪纸", "object": "门笺"}], "创作":[{"subject": "徐州剪纸", "object": "窗花"}], "参与":[{"subject": "人们", "object": "节日庆典"}, {"subject": "人们", "object": "门笺"}]}
3	{"传承":[{"subject": "王桂英", "object": "徐州剪纸"}], "属于":[""], "涉及":[""], "创作":[{"subject": "王桂英", "object": "窗花"}, {"subject": "王桂英", "object": "鞋花"}, {"subject": "王桂英", "object": "洗床单"}, {"subject": "王桂英", "object": "农村生活剪影"}, {"subject": "王桂英", "object": "春耕秋收"}], "参与":[{"subject": "王桂英", "object": "剪纸培训班"]}

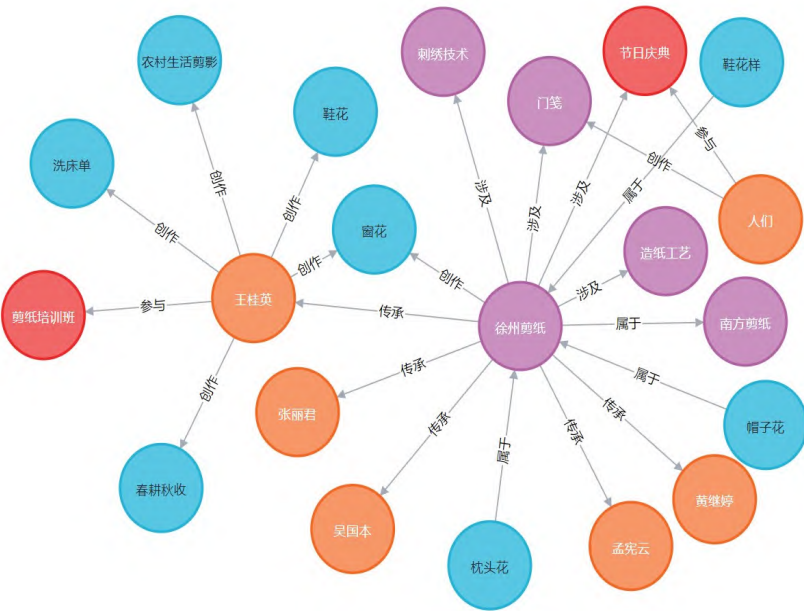


图5 地方性非遗“徐州剪纸”知识图谱
Figure 5 Knowledge map of local NRH "Xuzhou Paper-Cutting"

图中,“徐州剪纸”作为核心节点,连接相关文化要素、传承人及相关活动的多种实体,构成“徐州剪纸”文化传承与传播的知识网络。从图谱可以看出,与“徐州剪纸”直接相关的主要节点包括“王桂英”“造纸工艺”“南方剪纸”“窗花”等,这些节点围绕剪纸艺术展开,涵盖技艺传承和艺术表现等多个维度。在这些节点中,“王桂英”是技艺传承的核心,连接“春耕秋收”“鞋花”“农村生活剪影”等作品,展现出徐州剪纸技艺的多样性与丰富的文化内涵。

此外,图谱清晰地呈现出文化传播与技艺传承的多条路径,涵盖节庆活动、技艺教学和农村生活等多个场景,表现出非物质文化遗产在现代社会中的多层次价值与实践模式。通过分析,“徐州剪纸→节日庆典”“徐州剪纸→门笺”以及“人们→节日庆典”链条揭示剪纸艺术在节庆场景中的功能和传播效果,门笺这一传统形式为人们日常对剪纸文化认同的重要载体;而“徐州剪纸→王桂英→剪纸培训班”以及“王桂英→农村生活剪影”则体现传承人通过教学与培训将剪

纸技艺深入到农村生活场景的具体实践,进一步扩展其社会覆盖面与实际影响力,而“徐州剪纸→张丽君”“徐州剪纸→吴国本”等相关人物凸显传承人在推动技艺保护与传播方面的核心作用,徐州剪纸的文化发展与传播路径不仅依赖于节庆活动和社会化教育的推动,也需传承群体与艺术形式之间的深刻联系。

4 地方性非遗知识图谱在高校思政教育中的创新应用

4.1 助力课程思政元素资源库建设

通过知识图谱的梳理与扩展,非遗文化中隐含的内在思想政治元素得以系统化呈现,为思政教育资源的整合和思政元素库素材来源提供全新路径,为高校思政教学内容资源建设提供思路支持。在徐州剪纸知识图谱中,实现了非遗文化中隐性思政要素的显性转化。其纹样体系中蕴含着丰富的礼仪规范、道德准则和价值观念,构成了中国传统文化的微观叙事单元。运用大语言模型构建的知识图谱,可将散见于技艺传承、民俗活动中的思想政治教育元素进行立体化、可视化呈现。例如:通过提取“门笺文化”中“福禄寿喜”的吉祥寓意与社会主义核心价值观的联系,解析“团花剪纸”对称均衡构图中蕴含的家国同构哲学思维,挖掘“抓髻娃娃”等图腾符号承载的民族团结意识,形成了覆盖价值观念、审美取向、思维方式等多维度的教育资源矩阵。非遗资源在教学设计与思政教育目标紧密结合,构建出涵盖社会责任、文化认同与价值观教育等多层次内容的资源网络与课程内容,充分发挥非遗文化在思想政治教育中的价值,为资源的语义整合与思政教育创新应用奠定基础。

在技术实现层面,通过构建“纹样-工艺-匠人-传说”四维知识图谱模型,形成高关联度图谱体系。这一新型架构不仅为思政元素库的素材整理提供标准化模板,更借助图数据库的智能推理能力揭示出新型教学元素。例如:通过“五毒避邪”纹样的防疫功能映射公共卫生教育主题,由“剪纸填色”互动工艺延伸网络安全意识培养模块,基于年俗剪纸的时令特征开发二十四节气文化认同教育项目。

基于知识图谱的思政元素和案例资源库,动态更新机制是非遗知识图谱在思政元素库建设中的技术优势。这种多层次、多维度的关联充分拓展思政元素库的素材范围,为思政教育资源的深度整合与广泛应用提供了可能性。通过对文化节点与关系的动态扩充,知识图谱能够持续引入新的非遗文化内容,使得思政元素库始终保持与文化实践发展的同步性,新加入的节庆活动、艺术作品与传承事件通过数据库的实时更新直接映射至知识图谱中,并与已有的文化要素形成新的关联,这种动态化特性使得知识图谱能够涵盖更加全面的文化资源,确保思政元素库的内容具备鲜活性与时效性。在动态更新的基础上,知识图谱还能够提供精准的资源定位能力,在“徐州剪纸”地方性非遗知识图谱中,通过对“徐

州剪纸→王杜英→农村生活剪影”等链条的分析,可以快速提取与思政教学目标相契合的文化素材,通过徐州剪纸的历史传承故事可以设计以传统文化为主题的课堂教学内容,通过农村生活剪影的节点关联则可以进一步引导学生感受非遗文化传承中的责任意识,以及在乡村振兴与社会发展中的实践意义。

此外,非遗资源的语义标注和链接能力能够实现思政资源的深度聚合,以徐州剪纸为例,其知识图谱将学术论文、政策文件与报刊资料等多种类型数据与剪纸艺术的具体技艺、历史脉络及社会活动相关联,形成多元多维度资源整合的基础框架。例如通过对节日庆典与窗花艺术的关联分析,可以精准呈现剪纸艺术在节庆文化中的传播形式;而通过传承人物与农村生活节点的关系,则能够直观揭示非遗文化与乡村振兴的互动模式,对非遗资源进行直观表述。

4.2 思政教育教学设计与模式智慧化实现

地方性非遗知识图谱为高校思政教育提供了全新的教学资源。通过线上线下相结合、理论与实践相结合的方式,有助于创新教学方式与教学设计。在思政课程中引入徐州剪纸作为一种艺术形式,将传统文化和社会主义核心价值观相结合。可以设计主题为“剪纸中的爱国情怀”和“民俗与道德”的课程,通过剪纸作品的分析,引导学生思考文化传承的重要性和自身责任。

地方性非遗知识图谱有助于构建集非遗知识图谱、教学资源、互动交流于一体的思政教育智慧教学平台。以徐州剪纸文化为例,通过构建在线智慧教学平台、团队协作与项目式学习以及智慧辅助与实时反馈,可以实现更加具体、高效和生动的教学。徐州剪纸智能体构建,让学生在智能体上根据兴趣选择主题进行讨论、提问或解答。设计互动游戏和挑战,如拼图、音频识别等,激发学习兴趣 and 参与度,巩固所学知识。地方性非物质文化遗产有助于教师创新思政课堂线下教学模式。教师可以结合徐州剪纸的特点,采用案例分析、小组讨论、实践操作等教学方法,让学生在课堂上充分参与、互动。例如,组织学生参观剪纸艺术展览,邀请剪纸艺人进行现场教学,使学生更好地了解和传承非遗文化。

此外,思政教育智慧教学平台设计过程中可以基于知识图谱构建地方性非遗的问答系统,实现课堂知识和内容的及时交互和问答。基于知识图谱的智能系统可以从多维度反映学生的学习状态,支持智慧辅助技术收集分析思政课堂学习行为数据,评估学习效果和识别需求。

4.3 实现思政教学内容个性化推荐

非遗知识图谱通过结构化语义关联为推荐系统提供深度上下文,解决了传统协同过滤的稀疏性问题,同时使推荐结果更透明、多样。基于非遗知识图谱开展思政教育过程中,可以根据学生的学习兴趣和需求推荐特定非遗文化资源内容,提高学生的积极性和参与度。通过构建非遗领域的结构化知识网络,能够在思政教育中实现内容的精确匹配与高

效推送。“徐州剪纸”知识图谱以其多层次的语义节点和关联关系为基础,将非遗文化资源与学生的学习需求联系起来,相比于传统的推荐方法,知识图谱依托深度语义分析技术,使非遗文化内容能够以情境化方式融入思政课堂教学。例如,在分析学生的学习兴趣时,知识图谱能够识别其对窗花艺术与节庆文化的关注程度,并基于关联关系推送与之匹配的学习资源,在教学创新中实现高效的知识传递。而对于对历史感兴趣的学生,可以推荐关于徐州剪纸文化起源、发展及演变的过程;对于艺术爱好者,可以介绍徐州剪纸的制作工艺、图案设计等方面的内容;而对于价值观教育方面,则利用知识图谱可视化手段,可以帮助学生更好地理解 and 吸收社会主义核心价值观。徐州剪纸中常见的“富贵吉祥”“和谐家庭”的主题,反映了人们对幸福生活的追求。这种文化象征的传达,可以与思政课程中的价值观教育相结合,引导学生树立正确的价值观念,认识到传统文化在当代社会的重要性。同时,非遗艺术的折射,也能帮助学生更好地理解社会责任感、集体意识等重要的思想品质。

参考文献

- [1] 中国政府网. 中共中央办公厅 国务院办公厅印发《关于进一步加强非物质文化遗产保护工作的意见》[EB/OL]. (2021-8-12)[2024-12-13]. https://www.gov.cn/gongbao/content/2021/content_5633447.htm.
- [2] 夏立新,陈欢,李雪漫,等. 基于图文数据的非遗跨模态知识图谱构建方法研究[J]. 情报资料工作, 2024, 45(5): 73-81.
- [3] 王常珏,张强,王盟烁,等. 基于本体的剧曲类非遗知识图谱构建研究——以元曲为例[J]. 图书馆杂志, 2023, 42(11): 91-98.
- [4] 王左戎,邓三鸿,胡畔,等. 价值共创视域下中国传统戏曲知识图谱模式层构建及应用研究[J]. 情报科学, 2025, 43(3): 165-175.
- [5] 张卫,王昊,李跃艳,等. 面向非遗文本的知识组织模式及人文图谱构建研究[J]. 情报资料工作, 2021, 42(6): 91-101.
- [6] 汪琳,王昊,李晓敏,等. 融合学习扩展的非遗陶瓷工艺领域术语库构建及应用[J]. 图书馆论坛, 2024, 44(2): 66-78.
- [7] 范青,史中超,谈国新. 非物质文化遗产的知识图谱构建[J]. 图书馆论坛, 2021, 41(10): 100-109.
- [8] 李燕,胡智文. 中国民间说唱艺术鼓词知识图谱研究[J]. 中国音乐, 2018(2): 186-192.
- [9] 刘清堂,尹兴翰,吴林静,等. 数字化保护与传承视域下民族器乐文化智能应用服务系统构建[J]. 图书馆杂志, 2023, 42(6): 113-120.
- [10] 罗晓晴,李晓元. 文化育人视域下非物质文化遗产融入高校思政教育的价值与路径[J]. 通化师范学院学报, 2024, 45(11): 120-126.
- [11] 张敏琦. 高校思想政治教育中的非物质文化遗产传播和保护[J]. 广西教育学院学报, 2023(6): 185-188.
- [12] 黄伟. “非遗”在民族地区大学生思政教育中的应用[J]. 贵州民族研究, 2017, 38(7): 238-241.
- [13] 孙鉴. 基于系统思维的非遗与高校思政教育融合研究[J]. 浙江工商职业技术学院学报, 2024, 23(3): 53-57.
- [14] 张晴. 非物质文化遗产融入高校思政教育的思考[J]. 大众文艺, 2024(4): 158-160.
- [15] 郭瑾莉,朱伟利. 非物质文化遗产赋能高校课程思政的理论侧重与教学实现[J]. 中国大学教学, 2023(8): 59-64.
- [16] GROOTENDORST, MAARTEN. BERTopic: Neural topic modeling with a class-based TF-IDF procedure[DB/OL]. (2022-03-11)[2024-12-24]. <https://doi.org/10.48550/arXiv.2203.05794>.

(责任编辑:毛秀梅)

(下转第160页)

A Study on User Transfer Intention of Knowledge Q&A Platforms from the Perspective of Social Learning: Based on PLS-SEM and FsQCA

SHEN Wang, WANG Yuxin, LIANG Shihao, FENG Xin, LI He

(School of Business and Management, Jilin University, Changchun 130012, China)

Abstract: [Purpose/significance] This paper explores the willingness of users to transition from free knowledge Q&A platforms to paid knowledge Q&A platforms based on Social Learning Theory and the Push-Pull-Anchorage Model. It is of great significance for deeply understanding the motivation behind user transitions and promoting the sustainable development of paid knowledge platforms. [Method/process] Starting from the perspective of social learning and combining the Push-Pull-Anchorage Model to construct a theoretical hypothesis model, this study uses Structural Equation Modeling (SEM) and Fuzzy-set Qualitative Comparative Analysis (fsQCA) to investigate the factors influencing the transition of users from free to paid knowledge Q&A platforms and their configurations. [Result/conclusion] The study results indicate that users learn from others' feedback, behaviors, and outcomes during the process of knowledge payment transition. There are five configurations of pathways, in which the combination effect type pathway includes users referring to the opinions of other users in the decision-making process. In the innovation-pull type pathway, highly innovative individuals are more receptive to new things, and users learn from positive word-of-mouth to form positive feedback, thereby strengthening their willingness to transition. In the auxiliary effect type pathway, users are influenced to some extent by anchoring effects during the transition process. This study deepens the understanding of the questioners' willingness to transition and the underlying motivations, and provides many feasible suggestions for paid knowledge Q&A platforms. [Innovation/limitation] This study explores the willingness of questioners to transition and the underlying motivations, and proposes corresponding strategies. By using fsQCA to explore the complex relationships between variables, it has a certain degree of innovation.

Keywords: knowledge payment; knowledge Q&A platform; transfer intention; social learning theory; PPM model

(上接第 147 页)

Construction of Knowledge Mapping of Local Intangible Cultural Heritage and Its Application to Civic and Political Education

ZHANG Mengmeng^{1,2}, ZHANG Maomao¹

(1.China University of Mining and Technology, Xuzhou 221000, China; 2.Xuzhou Institute of Technology, Xuzhou 221018, China)

Abstract: [Purpose/significance] In order to promote knowledge mining in the field of intangible cultural heritage and its innovative application in ideological and political education in universities, and to assist in the integration and application of local characteristic ideological and political education elements, further enhance the application effect and influence of intangible cultural heritage in contemporary ideological and political education. [Method/process] This study first collected and constructed a dataset of local intangible cultural heritage, used the BERTopic topic model to extract entity categories and relationship labels, then used the OneKE big model for entity extraction and relationship recognition, and finally used the Neo4j graph database to visualize the knowledge graph. Taking the national intangible cultural heritage Paper Cuttings in Xuzhou as an example, this paper explores its application path in ideological and political education in colleges and universities. [Result/conclusion] Research has found that the knowledge graph construction method based on the OneKE big model can clearly present the knowledge and cultural background, skill inheritance, and internal relationships of local intangible cultural heritage, forming a structured knowledge expression system; the knowledge graph based on construction can assist in exploring ideological and political elements, facilitate the integration of ideological and political teaching resources, and enhance the application effect of local intangible cultural heritage in ideological and political education. [Innovation/limitation] A method for constructing a knowledge map of local non-heritage based on a large language model is proposed, and the knowledge elements are mined for use in ideological and political education.

Keywords: intangible cultural heritage; OneKE big model; BERTopic; knowledge graph; ideological and political education