

# The AI technology path for the construction of digital library knowledge graph

Shifeng Quan

Shanghai Library, Shanghai, 200031, China

## Abstract

This paper aims to systematically explore the innovative path of AI-empowered knowledge graph construction in digital libraries. The research first analyzes the bottlenecks of traditional construction methods in terms of automation, semantic depth, and dynamicity, and then combs the core AI-enabling technologies in key links such as knowledge extraction fusion reasoning, and intelligent application. On this basis, this paper creatively proposes a three-layer integrated technical path covering automated data processing, dynamic knowledge evolution, and intelligent service generation and analyzes its implementation guarantee and future trends. This research provides feasible technical schemes and theoretical references for the transformation of digital libraries from information resource warehouses to intelligent knowledge service centers

## Keywords

digital library; knowledge graph; artificial intelligence; technical path; semantic service

## 数字图书馆知识图谱构建的人工智能技术路径

全石峰

上海图书馆, 中国·上海 200031

## 摘要

本文旨在系统探讨人工智能技术赋能数字图书馆知识图谱构建的创新路径。研究首先剖析了传统构建方法在自动化、语义深度与动态性方面的固有瓶颈,进而梳理了知识抽取、融合推理及智能应用等关键环节的核心AI赋能技术。在此基础上,本文创造性提出一个涵盖自动化数据处理、动态知识演化与智能服务生成的三层整合技术路径,并分析了其实施保障与未来趋势。该研究为数字图书馆实现从信息资源仓储向智慧知识服务中心的转型提供了切实可行的技术方案与理论参考。

## 关键词

数字图书馆; 知识图谱; 人工智能; 技术路径; 语义服务

## 1 引言

数字图书馆馆藏资源变得越发多源异构,用户的知识需求也愈发复杂而精准。传统依靠元数据的信息组织模式很难承担起深度知识识别和智慧服务的重任。知识图谱具备很强的语义关联能力和演绎能力,是解决这个难题的重要技术。不过,当前的形成过程很看重人工因素,存在效率低、语义简单、更新迟缓等问题。人工智能技术快速发展起来,在自然语言处理和机器学习方面有所超越,给知识图谱的全流程自动化、智能化形成创造了变革性的机会<sup>[1]</sup>。本研究将深入探索AI技术与知识图谱构建流程的深度融合,旨在设计一条高效、智能且可持续的技术实施路径,推动数字图书馆的范式革新。

## 2 数字图书馆知识图谱构建的当前技术格局

### 2.1 构建核心环节依赖的传统技术与方法

当下,数字图书馆知识图谱的形成仍然在较大程度上依靠一些传统或者早期智能技术。在知识获取阶段,大多运用依靠规则或者模板的方式,从结构化或者半结构化数据当中获取实体和它们之间的联系。对于非结构化文本,则凭借词典以及基本的自然语言处理工具来达成。知识表现通常借助专家亲手创建的领域本体来做到,从而提供基本的语义架构<sup>[2]</sup>。在知识整合与保存层面,主要还是靠依托字符串符合的实体对准技术和关系型数据库或者是早期的图数据库。这些方法构成了现有技术栈的基础,但智能化水平有限。

### 2.2 现有技术路径面临的突出瓶颈与挑战

上述技术路径在实践中暴露出诸多瓶颈。第一个难点在于自动化水平较低,过度依靠领域专家的人工介入,使得创建时间冗长,费用昂贵。第二个难点是语义认知仅限于表面层次,传统办法很难深刻领会资源所蕴含的语义及其复杂

【作者简介】全石峰(1978-),男,中国浙江鄞县人,本科,高级工程师,从事数字资源管理,信息技术应用研究。

的背景环境,知识获取的精准度和全面度受到限制。第三个难点在于应对多来源不同种类的数据时,传统融合手段无法很好地化解实体的歧义和矛盾之处,难以保证知识体系的一致连贯。第四个难点是静态生成模式不能顺应知识的极速变化情况,图谱的动态维护和升级速度迟缓,从而影响其及时性和有效性。

### 2.3 人工智能技术融入的必然性与切入点

为克服上述挑战,引入人工智能技术已成为必然选择。AI的融入并非全盘替换,而是对关键瓶颈环节的增强与重构。其核心切入点在以下几个方面:深度学习模型可达成非结构化文本的自动化、高精度知识抽取;表示学习技术把实体与关系映射到低维向量空间以做到深层次语义计算;图神经网络和推理模型被用来改善知识融合与逻辑推理的智能化水平;机器学习算法用于知识图谱的动态学习及自适应优化<sup>[5]</sup>。这些切入点为创建过程的质量和效果的提升形成了基础。

## 3 驱动知识图谱构建质效跃升的人工智能核心使能技术

### 3.1 面向多源异构资源的知识抽取与表示学习技术

人工智能技术给知识的获取与表现带来了本质上的改变。依靠像BERT、GPT系列这样的预训练语言模型来执行序列标注和阅读理解时,可以以前所未有的精准度从文本、图片OCR结果甚至音视频转录内容当中抽取实体、属性以及复杂的关联关系<sup>[6]</sup>。而且,诸如TransE、RotatE等知识表示学习技术所包含的嵌入模型会把符号化了的知识映射到连续的向量空间里面,不但做到了数据的稠密表示,还能够凭借向量运算去捕捉隐藏的语义关联,为后续的语义相似度计算、链接预测等任务提供了高效且富含语义的数学基础。

### 3.2 支持动态演化与深度融合的知识融合与推理技术

在知识融合层面,AI技术带来了更强有力的解决途径。依靠嵌入表现的实体对齐方法,通过考量向量空间的相似程度,可以更为精准地找出不同出处中的同个实体,有效地化解异名同义以及同名异义的情况。就知识推理而言,凭借规则发掘的机器学习方法能够自行找到隐藏的关联规则,而且图神经网络可以直接在图结构之上执行端到端的学习及推理,推断可能存在的关系,以此来充实并拓展知识图谱。这些技术一起支撑起一个可不断吸纳新知识、自主提升与发展的动态知识体系。

### 3.3 赋能精准服务与交互的知识存储与智能应用技术

创建知识图谱的终极价值在于智能服务当中。就图存储和查询而言,新一代原生图数据库给海量关联数据的高效存储以及复杂图谱查询提供支持。在应用层,依靠知识图谱和插入技术的语义检索模型,可以领会用户的查询意图,达成凭借概念和关系的精确契合;智能问答系统凭借图谱的语义网络直接产生答案或者推导路径;个性化推荐系统通过剖

析用户在知识图谱之中的浏览和交互轨迹,找出其深层次的兴趣,从而做到知识的精确传递<sup>[5]</sup>。这些应用技术把静态的知识库转变成动态的智慧服务动力源。

## 4 面向数字图书馆场景的AI赋能的整合技术路径设计

### 4.1 以深度语义理解为基础的数据层自动化处理流程

对于数字图书馆的多模态资源而言,本路径规划重点在于创建依靠深度语义领悟的数据层处理流水线。这个流程首先要做的是对文本、图片、音视频等多源数据执行统一预处理及特征获取,接着凭借预先训练好的领域自适应模型来实施批量的实体识别、关系获取以及事件检测,从而生成最初的知悉单元。还要用表现学习模型把这些单元转为成向量形式<sup>[6]</sup>。这个过程借助积极学习策略削减人工标注量,而且通过质量评定模块给予反馈以改善模型,进而塑造起一条自动化的、循环的从“数据到知识”的转换路径,为上层图谱创建提供优质且富含语义的素材。

### 4.2 以动态图谱演化为核心的知识层组织与更新机制

在知识层,路径强调图谱的动态演化能力。机制首先将数据层产出的知识单元与现有领域本体或图谱Schema进行对齐与映射,利用基于嵌入的融合算法解决冲突、实现消歧。系统设立周期性与触发式两种更新模式:周期性模式依靠增量学习技术来批量处理新数据;触发式模式会及时回应用户反馈或者外部知识源的变动情况。凭借插入向量的在线学习以及图结构的增量更新算法,这个机制可以在无需重新创建整个图的情况下,达成知识的平稳纳入并且使语义关系自动适应并作出调整,从而保证图谱的鲜活性与准确性。

### 4.3 以用户需求为导向的应用层智能服务生成模式

在应用层面,路径规划出一个由用户需求所激发的智能服务产生框架。一旦收到用户查询或者服务请求,系统就会先凭借自然语言认识技术来分析并把它们对应到知识图谱的概念节点上<sup>[7]</sup>。接着,服务引擎会在图谱当中展开多跳推理、关联识别或者社群识别,从而创建起候选的知识路径或者答案集。再融合用户画像以及上下文环境,利用排序学习或者加强学习模型对这些候选结果执行优化和个性化的重新排列,最后生成符合用户深层次需求的各种个性化知识服务产品,比如知识导航、可视化的关联图谱、智能问答或者是研学推荐之类的综合知识服务项目。

## 5 技术路径实施的保障机制与未来演进方向

### 5.1 实施过程中的关键影响因素与应对策略

该技术路径若想达成,离不开诸多保障措施。其一,数据质量颇为重要,要形成起数据清洗、标注规范以及质量评定的常规运作流程。其二,算力资源起到支撑作用,可考量采用混合云架构并利用高性能计算单元,从而满足模型训练和推理方面的需求。其三,组建复合型人才团队十分关键,这个团队应融合图书情报学、计算机科学以及相关领域的专

业知识。其四，标准与互操作性就像一座桥梁，要积极采纳并参与制订有关语义网、本体以及数据交换方面的标准。其五，还要重视用户隐私保护和数据安全问题，创建起对应的伦理审查和技术防范体系，促使该路径在合法合规的框架内稳健推进。

## 5.2 技术路径的预期成效与多维度价值评估

执行此路径会从诸多方面创造重大价值。在效率层面，将大幅缩减人工成本并缩紧形成时间，做到规模化的知识整合。在质量层面，可以优化知识体系的语义层次、统一性以及动态适配能力<sup>[8]</sup>。在服务层面，则能够催生以前难以达成的深入知识识别、跨学科联系分析以及个性化的知识伴侣等高端智能服务。其价值归结起来就是优化数字图书馆的核心竞争力，令其由消极的内容供应者变成积极的知识革新倡导者和智能科研平台，突出加强其在学术交流和知识生产格局中的核心地位。

## 5.3 面向下一代智慧图书馆的长期技术演进趋势

展望未来，数字图书馆知识图谱的技术路线会深入融合前沿科技。大语言模型和知识图谱协同成热点，大语言模型具备很强的泛化语义认识与生成能力，知识图谱提供精确的结构化知识限定，两者结合能够达成更自然的人机交互及知识创作。神经-符号系统若取得发展，便有可能做到逻辑推理和神经网络学习的统一，从而加强知识图谱的可解释性及其复杂推理能力<sup>[9]</sup>。跨模态知识图谱技术将会加深对图片、音频、视频等非文本资源的语义认识并形成联系，最终向构建一个覆盖全媒介、可用于复杂认知任务的“馆藏大脑”这一长远愿景演进。

## 6 结语

本文围绕人工智能技术推动之下的数字图书馆知识图谱形成路径展开系统探究。首先剖析传统技术遭遇的瓶颈，

然后梳理关键的 AI 助力技术，据此规划出包含自动化数据处理、动态知识发展以及智能服务生成的综合技术路径。而且探讨该路径的保障机制及其未来趋势<sup>[10]</sup>。研究显示，AI 技术全面参与可以切实加强知识图谱形成的速度、品质及其实用价值，是数字图书馆迈向智能化的关键动力源。未来的研究可在路径的具体工程化落地、与大语言模型等新技术的融合实践、以及长期可持续运营模式等方面进行更深入的探索与验证。

## 参考文献

- [1] 许萌.基于人工智能的高校图书馆数字阅读推广策略研究[J].数字通信世界,2025,(12):184-186.
- [2] 王瑞珍,刘蓝予.人工智能赋能高校图书馆数字资源服务的现状与思考[J].郑州师范教育,2025,14(06):93-96.
- [3] 黄笑夏.生成式人工智能赋能公共图书馆数字记忆建构研究[J].图书馆学刊,2025,47(10):41-44.
- [4] 肖调云.数字化时代图书馆服务创新研究与实践[J].参花,2025,(04):144-146.
- [5] 王曦.人工智能赋能智慧图书馆发展的作用机制[J].图书情报知识,2024,41(06):94-101+165.
- [6] 陈天宇.基于人工智能的数字图书馆信息检索系统[J].信息技术,2024,(07):173-179.
- [7] 张文亮,陈重阳,李璇.基于知识图谱的中国图书馆学人数字记忆平台构建研究[J].高校图书馆工作,2025,45(02):58-67.
- [8] 赵范心.基于知识图谱的档案知识服务模式思考[J].办公室业务,2022,(13):64-66.
- [9] 吴荣,段宏涛.知识图谱在图书馆推荐系统中的应用研究[J].自动化应用,2020,(08):71-73+76.
- [10] 王光梅.智慧服务视域下高校数字图书馆信息服务模式创新研究[J].石家庄学院学报,2025,27(05):155-159.