

基于知识服务的智慧图书馆知识资源建设研究与实践^{*}

——以民国时期地方文献知识库为例

冀晓萌

(国家图书馆, 北京 100081)

[摘要]智慧图书馆的核心内容是提供智慧服务,而知识资源建设是提供知识服务的前提和基础。结合民国时期地方文献知识库的建设实例,从知识服务的视角,对智慧图书馆知识资源建设的趋势、结构以及发展途径进行探讨。

[关键词]知识资源建设 知识服务 智慧图书馆 知识库

[分类号]G253

1 引言

近年来,国家高度重视文化事业与公共文化服务体系的建设。中共中央办公厅、国务院办公厅于2022年发布的《关于推进实施国家文化数字化战略的意见》明确提出,要推动公共文化服务的数字化转型,构建覆盖全社会的智慧化知识服务体系^[1]。在此背景下,智慧图书馆作为文化数字化战略的重要载体,其建设目标已从传统的文献存储与借阅转向以知识服务为核心的综合信息平台。

优质的知识资源是智慧图书馆提供高效服务的基础。只有拥有丰富、准确、权威的知识资源,智慧图书馆才能通过智能检索、推荐系统等为用户提供精准的知识服务。可见,知识服务与知识资源的构建是智慧图书馆发展的基石,对于提高图书馆服务能力、激发知识传播活力、满足用户信息需求有决定性意义。然而,智慧图书馆知识资源建设在实践中仍面临诸多问题,如资源整合难度大、数据标准化不足、技术应用不成熟等。因此,深入研究智慧图书馆知识资源建设的理论与实践,对于推动图书馆的智慧化转型具有重要意义。

笔者基于对知识资源建设发展特点及组织框架的分析,通过民国时期地方文献知识库平台的构建实例,探讨如何将知识资源有效地应用于知识服务之中,并针对未来知识资源建设工作的开展提出相关意见,从而为图书馆从传统的“信息存储中心”向“智能化枢纽”的转变提供理论支撑和实践参考。

2 智慧图书馆知识资源建设特点

2021年9月,国家图书馆发布了《智慧图书馆知识资源数据建设指南》,详尽地阐述了知识资源建设的详细要求和建设路径。指南资源类型广泛,有图书、视频资料以及VR/AR等前沿的数字化资源。在内容方面,主要聚焦于传承中华优秀传统文化、地方红色文献的整理、政府信息公开的整合,以及战“疫”专题资料的收集^[2]。其中知识资源建设的加工过程包括3部分:一是数字资源的细颗粒度索引;二是知识组织与专题揭示;三是针对新兴数字资源,提供专门化和知识化的服务。与“十三五”规划时期的图书馆数字资源建设相比,智慧图书馆阶段的知识资源建设呈现出以下特征。

2.1 建设目标以“知识服务”为主导

以“知识服务”为主导的建设目标,明确了智慧图书馆知识资源建设的核心任务,即通过提供高质量的知识服务,满足用户日益增长的知识需求。为实现这一目标,其建设内容包括以下3种路径:一是进一步细化馆藏资源的细颗粒度数据建设,精准展现知识内容的细节;二是重新整合各类资源,构建主题和专题类的知识资源架构和分类展示,增强知识内容的展示力度;三是推广VR/AR等创新数字资源,与智慧化服务场景相结合,以拓宽图书馆馆藏的范围,提供互动性强、沉浸式的阅读享受,满足读者的多样化需求^[3]。

^{*}本文系国家图书馆科研项目“知识图谱构建中知识标注与表示方法的实践研究”(项目编号:NLC-KY-2022-15)的研究成果。

2.2 资源加工方式更加深化

相较于数字图书馆在建设基础数字资源时采用的传统数字化方法,智慧图书馆在知识资源数据建设方面采取了更为深入和细致的处理手段,如通过细颗粒度标引开展知识组织及专题揭示的加工。更重要的是,智慧图书馆在知识资源的建设上,特别注重揭示知识之间的多维关系和深层次的语义联系。

通过利用关联数据技术,智慧图书馆建立起数据间的互联链接,形成了一个综合的数据关系网络。这不仅是技术上的进步,更意味着知识的呈现方式由此转化为更易于读者理解和探索的知识图谱形式。这种转化为知识检索提供了更为丰富和直观的入口,使得读者能够在复杂的知识体系中,更加轻松地发现和链接相关知识点^[4]。

2.3 技术应用突出智能手段

在智慧图书馆的构建与成长过程中,现代技术的运用是促进服务品质提升和提高服务效率的核心动力。自然语言处理(NLP)、机器学习等高新技术,在提升知识内容挖掘的精确度和效率上扮演着越来越重要的角色,利用该技术图书馆能够更有效地进行知识的探索 and 整合,从而提升资源处理的能力和服务水平。

智能技术的应用不仅限于资源的加工处理,还深入信息检索、用户个性化推荐等服务环节,直接提升了智慧图书馆的服务水平和用户体验,这一变化象征着图书馆服务正从传统的信息处理模式转变为智能化、定制化的知识服务模式,同时也是推动知识共享和利用过程中面临的新机遇与挑战。

2.4 建设内容注重打造专题知识资源库

在“十三五”规划纲要的实施阶段,多地公共图书馆利用自身优势,积极构建了包括古籍、报纸、地方志、红色革命文献和音频等多种类型的自建数字资源库,成功打造了一系列凸显地域特色的专题资源库,成为各馆馆藏的亮点和特色,是支撑智慧图书馆资源体系的关键支柱。专题资源库建立以后,为地方文化的研究、保存和传播提供了强有力的资源保障,对于时代发展而言,促进了图书馆在数字时代的发展和转型,增强了图书馆在信息服务领域的竞争力。

智慧图书馆在知识资源建设上不局限于传统的数字化转换,更重要的是通过创新性的转变和发展来有效利用这些特色资源。专题资源库提供的特色知识服务,不仅仅是单纯的知识集合和系统化,而是结合地方图书馆特色馆藏资源,通过智慧馆员的深入挖掘、精炼以及数据分析、知识发现等智能技术手段,对相关特色资源进行创新性的整合和开发^[5]。

3 知识资源建设组织框架

智慧图书馆知识资源建设是依托人工智能、大数据、物联网等新一代信息技术,对知识资源进行数字化、智能化、协同化的全生命周期管理。其核心目标是通过资源整合、知识挖掘和服务创新,构建开放共享、精准服务、动态优化的知识生态体系。笔者主要依照“知识提取—知识标注—知识关联—知识应用”的工作流程搭建知识资源建设的组织框架(见图1),打造一个结构化的语义知识库及服务系统,提供一整套综合语义搜索、知识重组、分类、可视化分析以及智能推送等功能的智慧知识服务平台。

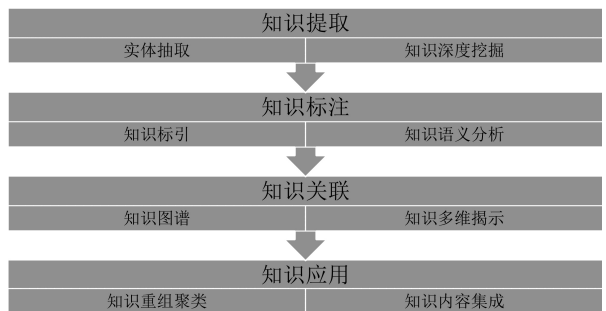


图1 知识资源建设组织框架

3.1 知识提取

知识提取是从大量非结构化或半结构化的知识资源中提取有价值的知识,并转化为结构化形式的过程。其核心目的是从文本或数据集中提取出有用的知识元素,如实体抽取、关系抽取、属性抽取、事件抽取等,通过运用深度学习、文本挖掘、构建主题模型等先进技术,将这些元素组织成为一种结构化的格式。

3.2 知识标注

知识资源的建设离不开知识内容的构建、集成和服务,而知识标注在其中贯穿全局、影响始终。知识标注环节对文献的格式和内容特性进行概括性阐述和细致标注,需要对处理的基本文献资源的格式、内容架构以及服务需求进行全方位审视,明确加工知识资源的详细程度和需要记录的要素。同时,基于文献的具体特点科学地确定记录和索引的具体内容。这一步骤遵循文献记录规则,分为两个主要部分:第一,依据文献的结构进行信息记录;第二,根据文献的实质内容进行知识内容的标注。

在初始阶段的任务是构建一个元数据描述的结构体系,使用标准化的形式化语言进行表达;在第二个阶段引入本体的概念,从知识单元的内容层面出发建立知识单元的语义描述模型。因此,知识标注的整个流程是一个逐渐深入、全面覆盖、从结构到内容的标注过程,体现了从基础到深层、从形式到内容的细致入微的逐步深化。

3.3 知识关联

知识关联是在知识资源中通过发现和建立实体、概念、事件等知识元素之间的各种语义关系,如因果关系、隶属关系、时空关系等,将分散的知识有机地连接起来,形成一个相互联系、相互依存的知识网络,以揭示知识的内在结构和逻辑,帮助用户更好地理解 and 利用知识^[6]。知识关联可以用关系列表、语义网络以及知识图谱等形式去展现,而知识图谱能够更加直观地提供可视化的关系展示。

知识图谱的构建过程主要包括语义关联和系统建设两个阶段。语义关联旨在构建知识实体间、实体属性间以及实体与属性间的关系,并通过元数据框架实现书目关系和字段的映射扩展;系统建设则通过知识获取、存储、建模以及发布与服务等流程,完成知识图谱的维护和操作^[7]。知识图谱技术通过应用先进的数据挖掘算法,分析图谱中的关联路径,以挖掘资源中人物、事件、地点、时间等基础要素之间的隐含关系,并采用可视化手段进行展示,例如构建社会网络关系图、时空轨迹图以及串联历史事件的线索等,这一过程不仅提升了知识的可视性和互动性,还极大地提高了知识发现和应用的效率。

3.4 知识应用

知识应用是为读者构建一个高效的智能服务平台,用于展示知识资源并推动知识的探索与创新。该应用平台应当具备专业、深入和开放的特点,可以提供基于知识资源的语义搜索、可视化分析、知识重新组合、分类汇总以及智能推送等关键功能,通过利用知识图谱技术,提供更细致和高质量的数据服务,进而使读者接触到更为深入和广泛的信息资源。这一过程旨在拓宽读者的阅读视野,激发创新思维,同时提供丰富且多元化的知识体系化专题服务,以满足读者对信息深度探索 and 知识获取的需求。

4 民国时期地方文献知识库建设实践

2023年6月30日,民国时期地方文献知识库在国家图书馆的官方网站上正式上线。该知识库展示了从文献中抽取的篇章、图表细粒度数据和人物、机构、事件、地理名称、物产5类知识数据,提供图文对照阅览、可视化图谱、自定义知识分析、知识关联与溯源等服务,深度挖掘民国时期地方文献的知识价值和文化价值,向全国智慧图书馆体系建设项目提供了宝贵的民国时期地方文献资源,促进数字人文研究领域的深度挖掘 and 应用,为读者搭建了一个便捷、高效的共享信息平台,满足读者对知识服务的精准化和语义化需求,帮助其全面了解民国时期各地区在社会、文化、生活等方面的详细信息。

4.1 精细化标引提升检索效率

民国时期地方文献知识库采用BETT深度学习模型,系统实现了文献知识要素的自动化抽取与结构化处理。通过细粒度标引包含人物、机构、历史事件、地理名称及物产资源等多元主题的关键词,累计形成42,600条知识数据。在知识组织层面,该库创新采用多维度标引体系,针对人物、机构、地理名称、历史事件等关键词进行细粒度标引,实现摘要、关键词、全文、分类、人物、机构、事件、地名等细粒度检索,为读者提供更加丰富的检索选项,也为读者提供大量可能感兴趣的知识检索点。同时,在阅读界面还对不同主题的标引关键词进行颜色区分,该功能使读者在浏览文献时可同步识别页面内标引主题的分类分布,便于随时调阅或开展精细化检索,并为用户推荐潜在关联知识主题。此外,民国时期地方文献知识库还支持不同类型资源的综合检索 and 基于语义关联的进一步检索,大大改善了检索功能,提高了检索效率。

4.2 图谱化应用实现知识的多维揭示

民国地方文献知识库利用语义关联技术,基于民国时期地方文献知识数据之间的关联关系,打造支持知识挖掘、关系发现和知识溯源的多维知识图谱,为读者提供可视化检索、展示和多边关系分析工具。在知识库首页可以看到多种类型的知识图谱展示,人物社会关系网络知识图谱可根据人物的时代、首字拼音等选项选择想要看到的人物社会关系图谱,以关系网图的形式展示人物间的关联交互;历史事件时空维度图谱则以时间轴的形式展示事件的类型、发生地及发生时间,按图谱形式分析聚合历史事件的结构脉络,为读者提供多场景事件分析;知识图谱的展示还包括地理名称历史沿革分析、志书机构时空维度分析、志书物产专题分析以及地理—人物关联分析,图谱展示不仅实现了跨文献知识数据的关联和融合,还体现了知识的交互性,拖拽布局展示以及知识节点详情查看等功能为读者提供多维度、跨类型资源的知识化阅读体验^[8]。

4.3 阅读模式满足读者多元需求

民国时期地方文献知识库为用户提供全息化阅读体验。集原始书影对照、流式文本阅读、图文互证三位一体,通过智能标引系统对历史地名、人物、事件等知识单元进行色标分类,读者点击标签即可激活知识卡片,获取结构化语义信息。此外,平台构建的动态知识图谱,支持关键词触发跨文献检索、关联知识跳转及语义扩展查询,独创“知识探针”功能,实现知识点到文献、图表的双向溯源。平台还提供文献聚合分析工具,支持卷册目录树、知识分类导航双通道浏览,读者可进行在线批注、版本校勘、知识

标引等协同操作,形成个性化研究空间。在文献分类层级下,知识库平台能够智能聚合关联卷册、相关知识等多元信息,通过可视化图谱展示文献知识网络,辅助读者快速定位关键信息节点。

5 智慧图书馆知识资源建设发展路径

民国时期地方文献知识库服务平台资源建设仍然处在不断改进和完善阶段,对于国内多数公共图书馆来说,知识资源的构建与智慧化的知识服务是一项创新性的任务,必须在理论和实践的交互中进行探索与进步。作为现代信息技术与图书馆服务紧密结合的产物,对提高知识资源的使用效率和服务水平具有极其重要的作用,笔者结合项目实践经验从建设规划、资源内容、服务平台、技术赋能、人员参与等方面提出以下几点建议。

5.1 建设规划

智慧图书馆知识资源建设的发展前提是要有完善的建设规划。第一,制定智慧图书馆知识资源建设的长远发展蓝图。确立智慧图书馆知识资源的建设宗旨、成长路径和执行计划,在知识资源建设过程中始终按既定目标前进,按计划执行,按流程操作,才能保障建设项目的持续性和预见性。第二,全面规划图书馆的资源分配。图书馆在数字化发展阶段已经建立积累了大批的数字资源,在智慧化转型的过程中,需要科学配置人力资源、物质资源和财务资源,对现有的数字资源进行分配和整合,确保知识资源建设任务的顺利执行。第三,构建智慧图书馆知识资源建设评价体系。评价体系需要从多方面、多角度、多维度进行评价,制定完善详细的评价机制,定期对知识资源平台建设的阶段性建设成果进行评价,适时优化发展策略,强化既定目标的达成。

5.2 资源内容

在智慧图书馆智慧化转型的过程中,如何利用好现有的资源内容,做好知识资源建设,首先要优先开展特色资源类型的知识化加工,公共图书馆应结合当地的历史文化和生活特色,重点开发特色文化资源,重视特色专题文化资源内容的挖掘和开发,发掘、整合和利用好地方特色习俗、历史文化等资源,丰富知识资源库,满足用户对特色科研信息的需求。其次,要优化资源结构,根据学科发展动态和用户需求变化,及时调整资源收藏策略,确保资源的时效性和针对性。最后,是加强资源整合,通过元数据整合、跨库检索等技术手段,实现各类资源的有效融合和一站式检索,提高资源利用效率,为用户提供个性化和知识

化服务。

5.3 服务平台

知识资源建设应立足于打造一个综合性网络服务平台,融合多种资源、功能、服务于一体,为用户提供一个统一且易于使用的服务接口。在此基础上发展便携式服务,例如研发移动应用小程序、社交媒体公众号等便携式平台,以应用用户随时随地的信息获取需求。此外,还可以根据用户需求开展定制服务,借助大数据分析手段,深入探索用户行为及资源使用数据情况,依据用户的兴趣倾向和操作习惯推出个性化的资源推送、知识引导等定制服务。

5.4 技术赋能

智慧图书馆知识资源建设离不开高新技术的应用。采用物联网、云计算、人工智能等当代信息技术,对资源进行知识提取、知识标引、知识整合,构建知识图谱,提供智慧服务,提高智慧图书馆的智能化程度。在建设过程中,还应推动技术研发,激励原创研发,开发适应智慧图书馆特点的技术和应用解决方案,在建设实践中不断积累经验,加强技术分享与协作,与高等院校、科研院所等搭建合作桥梁,联合促进智慧图书馆相关技术的研发与应用。

5.5 人员参与

智慧图书馆知识资源建设的核心宗旨是为用户提供智慧化服务,而智慧馆员则是提供智慧化服务的创造者和协助者,因为知识资源建设过程的每一步都有智慧馆员的参与,因而对智慧馆员提出了更高的要求。图书馆应加强人才培养,加大对智慧馆员的培训力度,提升智慧馆员信息素养和专业技能水平,建立激励机制,设立奖励措施,表彰在智慧图书馆建设中作出突出贡献的个人和团队,激发全员参与热情和创新活力。此外,还应鼓励用户积极参与,通过用户调研、意见征集等方式,了解用户需求和建议,引导用户积极参与智慧图书馆知识资源建设和管理过程。

6 结语

笔者从民国时期地方文献知识库构建的实践经验出发,以知识服务的视角对智慧图书馆知识资源建设的趋势、结构以及发展途径进行了研究和讨论。随着技术的不断进步和理论的日益成熟,智慧图书馆的知识资源及智能化服务必将持续优化和提升,在数字中国建设与文化强国战略的背景下,通过图书馆界同仁的共同努力,为推动图书馆事业的数字化、智能化转型以及创新发展的目标提供坚实的支持。

(下转第60页)