

我国智慧图书馆建设方法体系识别研究*

吴丹 杨馨梅 梁少博 何丽英

摘要 在当前数字化、智能化快速发展的时代背景下,图书馆正积极面向用户需求进行智慧化转型升级。在智慧图书馆方法体系研究领域,国内学者围绕技术利用、平台研发、风险防治、模式探究等多方面进行了丰富探索,但目前对于方法体系的系统梳理与实践应用探析较少。本研究基于文献分析法,调研已有智慧图书馆方法体系研究内容框架,梳理方法体系关键要素及逻辑关系。综合智慧图书馆实践调研所发现的数据整合与挖掘、资源加工与知识发现、数据质量评估与处理规范问题,围绕数据知识语义融合与映射关联、AI赋能自动化处理与智慧交互、数据治理标准化控制与风险预警三方面,提出未来智慧图书馆发展趋势。图1。表2。参考文献38。

关键词 智慧图书馆 方法体系 实践应用

Research on the Identification of Methodological Systems for Construction of Smart Library in China

Wu Dan Yang Xinmei Liang Shaobo He Liying

Abstract: Under the background of the rapid development of digitalisation and intelligence, libraries are actively facing the needs of users for intelligent transformation and upgrading. In the field of smart library methodology system research, domestic scholars are rich in exploration around the use of technology, platform research and development, risk prevention and control, mode exploration, but at present there are fewer systematic combing and practical application analyses of the methodology system. Based on the literature analysis method, this study investigates the content framework of existing smart library methodology system research, and sorts out the key elements of the methodology system and its logical relationship. Based on the problems of data integration and mining, resource processing and knowledge discovery, data quality assessment and processing standards found in the practical research of smart libraries, this paper puts forward the development trend of smart libraries in the future from three aspects: semantic integration and mapping association of data knowledge, AI-enabled automatic processing and intelligent interaction, standardized control of data governance and risk early warning. 1 fig. 2 tabs. 38 refs.

Keywords: Smart Library; Methodology System; Practical Application

2020年,国家图书馆提出“全国智慧图书馆体系”建设思路^[1],次年,“积极发展智慧图书馆”“加强智慧图书馆体系建设”先后被写入《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》和《关于推动公共文化服务高质量发展的意见》。积极推进智慧图书馆体系建设,提升公共文化数字内容的供给能力和服务水平,也是图书馆响应国家文化数字化战略和“十四五”公共文化服务体系建设规划要求,发挥自身优势,进一步推动数字中国和文化强国

建设的重要战略举措^[2]。

智慧图书馆方法体系是推动图书馆智慧化转型的战略框架与规划指导。国内图书情报学领域及图书馆业界围绕其进行了深入讨论与应用探索,研究内容涵盖对数字孪生^[3]、AIGC^[4]、大语言模型^[5]、情境感知^[6]、用户画像^[7]、知识图谱^[8]、边缘计算^[9]等新兴技术融入图书馆业务流程与服务模式创新的探索,针对新一代图书馆智慧管理服务平台的研发^[10]与比较研究^[11],以及探讨智慧图书馆数据治理标准规范及风险安全

* 本文系国家重点研发计划资助项目“智慧图书馆体系构建关键技术研发与应用示范”(项目编号:2023YFF0906800)研究成果。

防治问题^[12]。但目前围绕智慧图书馆方法体系的研究缺乏系统梳理与总结,对实践运行中存在的问题分析较少。智慧图书馆建设需要在理论探索与实践应用中不断发现问题,根据服务效能与应用效率,及时调整方向细节,在理论研究与实践运行中形成正向循环反馈,推动智慧图书馆建设稳步进行、健康发展。

基于此,本研究对我国已有的智慧图书馆方法体系研究成果进行梳理,提炼智慧图书馆方法体系基本框架要素,并从实践层面,总结归纳图书馆智慧化转型管理过程中所存在的问题与困境,进一步提出未来智慧图书馆方法体系发展趋势,以期智慧图书馆相关理论研究与建设深化提供参考。

1 智慧图书馆方法体系识别

1.1 研究方法与数据收集

本研究利用扎根理论方法,对相关文献仔细阅读,进行文本摘录与编码分类,分析每个编码之间的相互关系,直至理论饱和,从而提炼发现智慧图书馆方法体系要素,归纳得出智慧图书馆方法体系框架。扎根理论是一种自下而上的质性研究方法,主张从系统化收集的经验资料中抽象出核心概念,通过开放编码、主轴编码和选择性编码对资料进行推演分析,阐释概念间的关系,总结概括上升至系统的理论^[13]。

本研究以中国知网收录的CSSCI期刊文献作为数据源,以“SU%=(‘数据’+‘知识’+‘方法’)*‘智慧图书馆’-‘馆员’-‘空间’”为检索式进行初步检索,文献出版年度选取2014—2024年,检索时间为2024年3月10日,共得到429篇期刊文献。剔除会议纪要、图书荐读等与研究主题相关度较低以及无法获取全文的文献后,得到151篇符合本研究要求的文献,并将其导入MAXQDA软件进行开放编码、主轴编码与选择性编码,得出智慧图书馆方法体系的基本框架要素。

1.2 编码过程

本研究遵循扎根理论编码流程对文献进行编码分析,选择其中的145篇文献进行详细编码,

预留6篇文献用于理论饱和度检验。编码工作由两位编码人员独立完成,随后借助软件中的跨代码一致性功能,对不一致的编码内容进行反复讨论,咨询专家意见进行剔除、提炼或合并,达成共识,直至kappa系数值在0.8以上,表明编码一致性较好。具体编码步骤如下。

首先进行开放式编码,从文献中筛选出可编码片段并进行概念化和范畴化,对比整合有效概念形成36个范畴。随后,在此基础上进行主轴编码,发现、提炼范畴间的深层联系,经过重新组合和反复推敲,最终确立9个主范畴。主范畴与对应子范畴内涵如表1所示。选择性编码是基于主轴编码,进一步挖掘主范畴之间的关系,建立主范畴与核心范畴之间的意义关联,构建完整的理论体系。编码结果如表2所示。

理论饱和指编码分析过程中,新样本不再产生新的概念范畴。本研究对预留的6篇文献按照三级编码流程进行概念提炼,没有发现新的范畴与关联关系,表明理论框架已达到饱和。

1.3 智慧图书馆方法体系要素分析

“全国智慧图书馆体系”建设基于智慧图书馆云基础设施、全网知识内容集成仓储、智慧化知识服务运营平台、智慧图书馆管理系统、实体智慧服务空间的“1+3+N”总体架构,包含智慧图书馆评价、标准规范、研究与人才培养三个支撑保障体系^[1]。本研究基于扎根理论,归纳提炼已有智慧图书馆方法体系研究中所涉及的框架要素(见图1)。依据编码结果,进一步梳理智慧图书馆方法体系内部的逻辑关系与作用机制。算力设施与算法体系的升级迭代引领着人工智能等新技术赋能重构图书馆的业务处理流程。立体化数据处理、全方位知识集成、知识融合关联、资源整合共享等要素作为智慧图书馆提质增效的关键环节,需要及时识别其中可能存在的风险隐患,制定标准规范,严控数据质量。人机协同等多模式驱动着业务流程的改进与优化,同时也支撑着空间生态革新和信息服务转型,提升智慧化服务质量,丰富智慧化服务形态。各要素相互作用,紧密联通,共同构成智慧图书馆方法体系。下文将围绕上述九项要素展开具体讨论。

表1 主范畴与范畴

主范畴	范畴	内涵
立体化数据 数据处理	资源数据抽取	对纸本资源、网络资源、科学数据、开放数据、数据集新型资源数据等资源本体与描述数据进行采集加工。
	管理与服务数据分析	分析业务、服务、空间及综合数据,实现业务关联洞察。
	用户数据挖掘	基于用户属性、行为、情境数据等分析用户偏好、挖掘用户群。
全方位知 识集成	通用知识	知识覆盖面广、专业程度较低,包括通用知识库、知识图谱等。
	领域知识	图书馆面向个性化需求或特定问题,构建具有较大规模、系统化的领域知识体系,包括专家知识库、领域知识图谱等。
	用户知识	将使用者的人际关系、学习兴趣、知识领域等方面的信息进行关联,形成社会网络和用户兴趣图谱。
	情境知识	围绕用户个体情境、自然情境、社交情境、行为情境等,构建情境知识图谱。
资源整合 共享	多媒体资源融合	面向数字弱势群体,对多媒体资源进行集聚、重构、优化。
	本地与云端资源整合	将图书馆本地资源与云端资源进行整合,并将本馆数据上传至云端实现大规模数据统一调度。
	多模态异构资源整合	利用关联数据和语义网实现图书馆多模态异构数字资源聚合。
	馆内数据共享	馆内数据实现跨系统、跨流程共享,支持多元操作、运算与分析。
	馆际数据共享	馆际或与其他机构间通常以图书馆事实数据库形式共享数据。
	面向社会的数据共享	面向社会公众提供知识数据共享,如科学数据开放共享等。
知识融合 关联	知识抽取	通过流水线方法、联合抽取法等,对实体与关系进行抽取。
	知识交叉融合	将不同的知识库或知识图谱进行知识对齐、合并与重组等,实现如领域知识融合嵌入知识图谱、隐性与显性知识融合。
	知识关联发现	基于知识图谱,关联整序知识元。利用数据挖掘技术深层挖掘、语义分析资源组织网络。
风险问题 识别	多源异构	数据来源多样、数据结构存在差异,难以高效整合。
	数据一致性	分布式系统中,数据需在多个节点之间保持一致。
	数据冗余	存在重复、多余数据建设的现象,数字馆藏资源同质化严重。
	数据孤岛	数据间缺乏关联性,数据库彼此无法兼容。
	伦理风险	在大数据应用过程中可能出现的数据隐私安全、数字鸿沟、过度依赖数据决策等问题。
数据质量 控制	数据价值评估规范	对数据的准确性、完整性、一致性、时效性等进行客观评估,指导数据采选与再利用。
	数据接口规范	针对数据自动采集,制定通用的接口规范。
	数据资源描述规范	规范定义和描述数据结构,包括数据类型、格式、字段长度等。
	数据互操作规范	制定标准规范确保不同系统之间数据的互通性。
	数据存储规范	根据馆内数据容量、特点、系统架构、存储模式,制定存储管理与技术规范。
	数据开放共享规范	制定标准规范确保数据在共享过程中的安全、合规和有效使用。
	隐私保护规范	运用一系列法律和规定对隐私数据进行保护。

续表

主范畴	范畴	内涵
新技术赋能增效	算力设施升级	云计算、物联网、区块链等技术引入,促进图书馆云平台、数据中心等系统设备更新迭代。
	算法体系支撑	利用分类与回归算法、聚类、推荐算法、关联规则、深度学习算法、大模型等,助力用户兴趣挖掘、知识图谱深度学习等。
多模式驱动发展	多元协同	基于服务内容与用户需求,以智慧图书馆平台为依托,建立多元主体协同体系。
	平台数据治理	数据管理中心化,包括数据合作社、数据信托、数据共享池等模式。
	数据与知识双驱动	整合数据知识资源,驱动智慧图书馆高质量发展,满足用户个性化需求。
	人机协同	AI与馆员、用户协同交互,提升图书馆运营效率与服务效能。
智慧化服务应用	空间生态革新	图书馆空间服务升级,提供沉浸式体验、多感官交互、虚拟展厅等服务。
	信息服务转型	根据用户需求,提供个性化、专业化、学科化的推荐、咨询、检索等知识信息服务。

表2 主范畴与核心范畴的结构关系

结构关系	结构关系定义	文献典型语句
立体化数据处理→智慧图书馆方法体系	对各种数据进行采集和分析等处理是智慧图书馆方法体系的核心要素。	图书馆可以通过各种系统、设备和应用开展智慧服务活动,持续获取馆员和用户的多模态服务数据。通过处理和分析后,这些数据将转化为知识被图书馆吸收,帮助图书馆优化服务活动 ^[14] 。
全方位知识集成→智慧图书馆方法体系	通用知识、领域知识、用户知识、情境知识等全方位知识集成是智慧图书馆方法体系的核心要素。	智慧图书馆作为知识集成化服务的创新高地,应以嵌入式、深度的学科服务、情报服务、数据服务、出版服务、智库服务等为依托,持续强化服务功能与服务效果 ^[15] 。
资源整合共享→智慧图书馆方法体系	资源整合与共享是智慧图书馆方法体系的核心要素。	加强馆际、馆舍、馆网之间的知识资源联合建设,通过建立知识链接、知识结构、知识框架的方式,实现对互联网信息和数字图书馆资源库内容融合 ^[16] 。
知识融合关联→智慧图书馆方法体系	多维知识融合关联是智慧图书馆方法体系的核心要素。	通过数据接口调用图书馆知识图谱库中各类知识图谱,用知识计算、知识发现组件汇聚、融合、计算智慧图书馆平台知识,促进知识群落协同优化、服务方式创新重构及隐性知识动态挖掘 ^[17] 。
风险问题识别→智慧图书馆方法体系	数据资源风险问题识别是智慧图书馆方法体系的保障要素。	数据安全、数据价值等成为智慧图书馆数据风险治理关注的维度,但存在着隐私泄露与信任、数据共享与孤岛、数据价值与干扰等问题,需要图书馆构建数据风险治理体系,有效规避风险 ^[18] 。
数据质量控制→智慧图书馆方法体系	数据接口规范、存储规范等数据质量控制手段是智慧图书馆方法体系的保障要素。	在智慧图书馆建设中,科学的数据标准体系能够保障图书馆发挥更大的生命力,而基于智慧图书馆数据体系和数据生命周期所构建的数据标准体系框架则是实现体系整体建设的逻辑基础 ^[19] 。

续表

结构关系	结构关系定义	文献典型语句
新技术赋能增效→ 智慧图书馆方法体系	云计算、物联网、区块链等新技术是智慧图书馆方法体系的支撑要素。	大语言模型产品展现出智慧涌现能力、信息聚合适应能力、多任务高效处理能力与思维链交互能力,助力智慧图书馆的知识采集、组织、推理等全生命周期中精准生成多维知识关联与融合架构 ^[20] 。
多模式驱动发展→ 智慧图书馆方法体系	人机协同等治理模式是智慧图书馆方法体系的驱动要素。	AIGC技术的快速崛起进一步创新了智慧图书馆人机协同工作模式,在服务、资源和空间中形成正向反馈与良性循环 ^[21] 。
智慧化服务应用→ 智慧图书馆方法体系	个性化的智慧服务应用是智慧图书馆方法体系的载体要素。	提供以数据为基础、以用户需求为根本、围绕“知识”“空间”“管理”核心要素的智慧化服务是智慧图书馆的发展方向 ^[22] 。

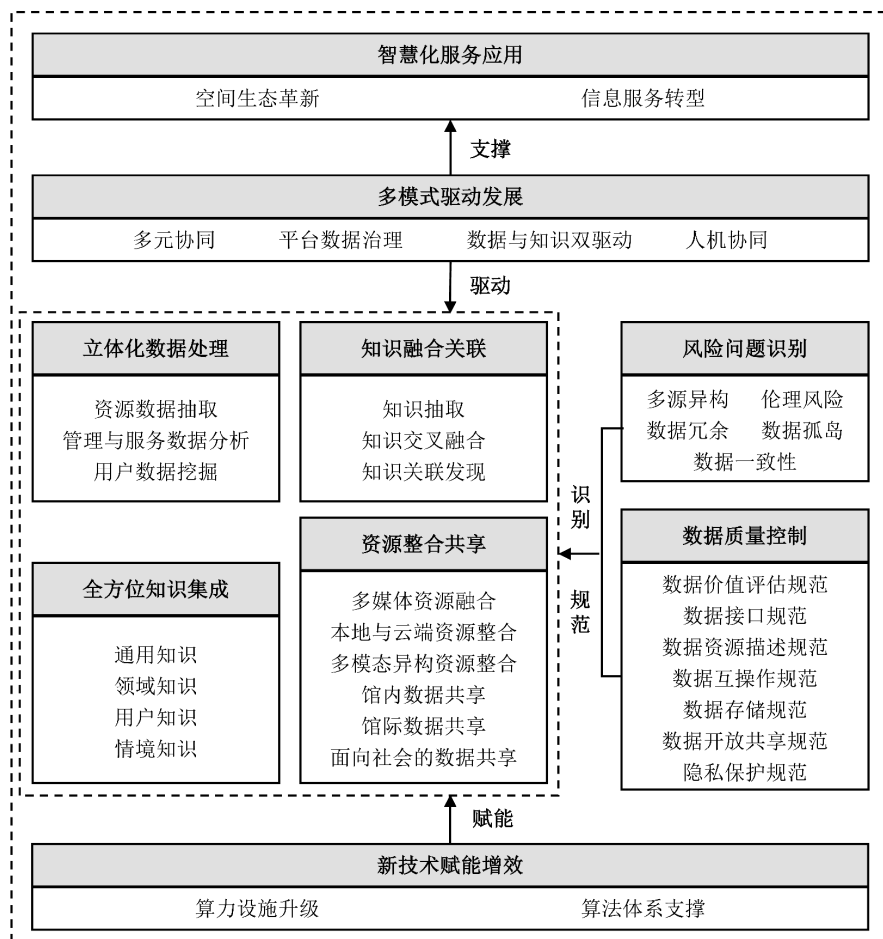


图1 智慧图书馆方法体系框架

(1) 新技术赋能增效。图书馆智慧化转型初期,学者们探讨大数据、物联网技术对图书馆服务模式创新;2015年后,如何将新兴技术应用转化为图书馆服务能力成为重要研究方向^[23]。新兴技术的迭代与图书馆实践的融合发展,推动着智慧图书馆建设不断探索未知与可能。随着众多生成式人工智能产品的推出,大语言模型应用领域逐渐广泛。已有研究主要探讨大模型技术赋能图书馆的参考咨询、资源加工、内容生成^[24]等多场景领域。随着智慧化管理系统的引入、开放联通的数据中心的搭建以及云端存储管理中心的建设,智慧图书馆的算力设施得到了全面升级。

(2) 立体化数据处理。已有研究强调图书馆应集成利用日常运营中产生的多维度、异构化、海量级的数据。如何处理好其统一管理、集成共享、分析利用的问题,是图书馆后续科学决策、提供高质量服务的基石。图书馆利用机器学习等技术进行业务数据的分析、用户画像的构建,不仅可以极大提高数据处理效率,还可对借还情况、设备使用情况、目标用户进行深层次的粒度刻画^[25],了解用户的真实需求,改善资源配置与服务内容。

(3) 全方位知识集成。集成通用知识、领域知识、用户知识、情境知识等全方位知识是智慧图书馆提供精细化、专业化知识服务的基础。随着万物互联,图书馆知识库建设将突破行业内知识集成共享格局,构建立体化知识体系^[1]。将图书馆领域知识与通用知识集成,可用于构建智能问答系统语料库,以尽可能满足用户多样化咨询需求。利用情境知识图谱精准感知、语义描述读者情境要素及相关行为,为场景化用户需求,进而实时推荐图书馆知识资源提供可能^[17]。

(4) 知识融合关联。对知识进行细粒度抽取标引、交叉融合、关联发现,进而搭建内涵丰富、关联紧密、规范有序的专有知识库是智慧图书馆知识组织与加工管理流程革新的必然趋势。知识融合将现有的知识元素根据约定规则进行推理、融合、转化,进而产生新知识,根据数据源变化更

新和用户反馈进行动态调整,最后提供知识服务^[26]。图书馆可以利用神经网络、自然语言处理等,实现语义分析、预测与可视化,进而从海量数据中发现隐性的知识关联^[27],构建体系宏大、关系多样的知识图谱、知识库,提供全文检索、图片检索等细粒度检索服务与知识溯源,提升检索的查全率与查准率。

(5) 资源整合共享。将多模态异构资源、本地与云端资源等进行整合,依据不同需求进行多粒度封装管理,可为知识内容集成仓储提供有序规范的初始素材,进一步为数据的一体化、标准化管理和知识融合服务奠定基础^[28]。智慧图书馆体系以统筹全国图书馆公共文化数字资源,打破“数字烟囱”“数字壁垒”为目标^[29]。数据资源间互联互通可以有效缓解馆际、区域间不平衡的情况。图书馆可以运用大模型技术进行文本、图片和视频资源的元数据标引以及自动分类^[24],利用关联数据与语义网技术实现数字资源的有序聚合,加以统一的标准规范,实现数据间的互操作。

(6) 风险问题识别。将数智技术泛化应用于数据挖掘、知识加工等过程中,带来图书馆优化提升的同时,也暴露出伦理风险、隐私安全、版权纠纷、运维安全^[30]等问题。准确及时地识别评估智慧图书馆建设所面临的风险问题,是图书馆提升智慧管理水平与可持续发展的重要保障。

(7) 数据质量控制。在智慧图书馆的建设运营中,数据质量管理不仅是数据管理的核心环节,更是确保图书馆信息系统稳定运行、提供优质服务的核心。制定数据治理标准规范能够有效避免智慧图书馆内数据不准确、一致性低、时效性差等问题^[19],对智慧图书馆规避隐私安全等风险具有显著意义。图书馆需研制并贯彻统一成熟的数据价值评估、数据互操作、隐私保护等规范,以确保数据管理的科学、高效。

(8) 多模式驱动发展。学者们从实践中总结出多种协同治理模式,包括多元协同模式^[31]、平台数据治理模式^[30]、数据与知识联合驱动模式^[32]、人机协同模式^[21]、大数据+微服务模式^[33]

等,成为智慧图书馆方法体系建设的重要驱动要素,共同推动着智慧图书馆向更高层次、更宽领域迈进。AIGC技术加快了图书馆人智协同融合与多主体协同工作进展^[21],并从算法支撑、数据资源等方面实现智慧图书馆服务功能拓展与模式创新。但已有研究往往将数据与知识的关联集成分而论之,并未实现两者真正的融合。

(9)智慧化服务应用。业务流程的升级再造最终作用于图书馆空间生态革新与信息服务转型。为用户提供高效便捷、个性化、多感官交互的服务体验,是智慧图书馆方法体系建设的重点。通过建设沉浸式阅读空间等新型空间形态,提供多元化学学习交流场所。构建用户画像模型预测用户信息行为,或以此为依据提供精准化服务^[34],不仅可提升服务效率,也搭建起图书馆与用户之间更加紧密的桥梁,进而营造全民阅读、终身学习的社会氛围。

通过梳理已有关于智慧图书馆建设的方法体系、流程模式的研究,发现许多研究围绕用户画像、资源推荐、智能问答等图书馆服务内容升级而展开,或是探讨大数据、区块链、AI等技术引入图书馆应用的策略与可能,亦或是提出隐私保护、数据合规等数据治理的监管规范体系,较为缺乏对智慧图书馆建设整体性、系统化的方法体系的梳理总结。明晰当前智慧图书馆建设运行所需的各关键要素以及相互作用的关系机制,有助于推动我国智慧图书馆“1+3+N”体系的落实与发展,指引各级图书馆制定适应本馆特色的智慧化转型方案,并开展有方向、有目标的具体实践应用。

此外,已有研究多是进行理论探索,形成具体模型、应用路径与模式,较少结合实践调研来调整、改进,进而有针对性地回应亟待解决的问题与需求。因此,本研究基于过往智慧图书馆方法体系研究,总结提炼关键要素及其作用关系,并前往实地调研图书馆智慧化转型实践,发现上述框架要素中所存在的现实问题,进而提出未来智慧图书馆建设发展的新方向。

2 智慧图书馆建设实践探析

2.1 调研设计

本研究针对国家级、省级、市级等多级公共图书馆及高校图书馆开展调研,从多视角获取我国智慧图书馆建设现状及面临挑战,并从中提炼共性问题,为后续研究提供参考。调研方法采用实地走访、线上问卷、馆员访谈等形式,最终选取国家图书馆、省级图书馆7家、市级图书馆21家、区/县级图书馆17家、镇级图书馆3家、村级图书馆2家,以及高校图书馆16家为调研对象,地区分布广泛,具有一定的代表性,以期从整体上看待目前国内智慧图书馆建设过程中存在的问题与挑战。

调研提纲基于上文总结的智慧图书馆方法体系框架,围绕关键环节展开调研分析:(1)资源整合与共享,调研图书馆资源类型、整合方式、共享策略。(2)数据分析与挖掘,了解图书馆数据采集利用情况、策略以及存在的数据质量等问题。(3)知识加工处理,了解细粒度知识加工的方法技术与标准,调研知识库等建设方法与现状。(4)新技术嵌入服务,调研智能问答等新兴服务的技术路线、使用效果及存在问题。

2.2 智慧图书馆建设实践存在的问题

2.2.1 多源数据缺乏整合与深入挖掘

随着图书馆数据规模的扩大和采集维度的增多,数据价值作用凸显^[35],驱动图书馆智慧化转型。目前,图书馆主要对馆藏资源数据、读者行为数据、场馆使用数据、用户反馈数据等,进行深入分析和挖掘。

但调研发现,图书馆由于数据处理技术有限、存储分散等因素,对业务数据的分析维度较少、挖掘深度不足,仅通过图书馆管理系统、网站分析工具等第三方平台产品,进行简单地分类汇总。例如,图书馆目前无法通过对比分析不同区域停留人数、读者群体特征、高峰使用时段、读者停留与借阅关系,来识别读者阅读兴趣,改进排架空间,提供情境化服务与个性化推荐。在用户数据的分析挖掘上,也存在无法深入获取用户行为偏好的

现象。检索系统进行结果推荐与排序时,大多停留在用户检索历史关注的学科层面。对于平台流量分布、提问分类等数据,图书馆仅作时间分布统计与内容标签化处理,尚未进行深层次关系的挖掘。

2.2.2 资源加工与知识发现效率较低

“全国智慧图书馆体系”建设规划中提出要搭建全网知识内容集成仓储,实现对多源知识内容的统一加工揭示、自动语义关联和集成管理服务^[1]。随着参考咨询等服务逐渐集成至统一云端操作,图书馆在提供知识服务时,所涉及到的如国外技术、市场覆盖等信息的查询需求,使得全文检索效果较差、资源缺乏横向整合与纵向深挖的问题凸显。

在资源整合方面,纸本资源与数字资源较为分散,数字资源间也未完全打通关联。关于副省级公共图书馆数字资源建设情况的调研结果显示,公共图书馆各数据库相对孤立,无法实现统一检索、统一维度统计^[36]。在资源揭示方面,目前图书馆还存在书目数据更新不及时、验收机制不完备等问题^[37]。

目前图书馆在细粒度知识的加工处理方面,面临着自动化程度较低、标引效果不够理想等挑战。在实践中,对于如古籍存在的异体字、字迹模糊等,其自动化识别效果与处理效率仍有可进步空间。形成的知识图谱、词云,存在某些实体、关系或属性缺失、分类错误、分词不完整的问题。

图书馆应用智能问答系统时,建立大规模、高质量的FAQ语料库与文档知识库也是一大挑战。目前,图书馆主要存在重复信息整合困难、相似问泛化不足、语义缠绕难消除、数据格式不一、句子主语不明等问题。在信息采集、加工标引、危险测试等流程上,仍需投入大量人力、时间去监督介入。

2.2.3 数据质量评估与处理规范缺失

面向用户提供精准化的服务是图书馆智慧化转型的核心特征之一,如何有效地对数据质量进行全面评估,提高数据的可靠性和可用性,是图书馆借助数智技术提升服务效能所面临的重

要挑战。

在业务数据处理与资源共享方面,随着平台系统开发及智能设备接入,数据呈现出多源、异构、多模态的特征。但在实践过程中,对于保存在不同云端后台的业务数据,以及多模态资源数据,图书馆缺乏统一的数据接口标准以及统一语义框架下的知识表达,难以统一提取特征,无法支撑智慧图书馆基于多模态数据实现智慧分析与跨媒体知识检索。

在用户数据的可用性方面,目前图书馆对用户数据的开发程度较低,较少针对个人用户进行画像,存在线下读者咨询记录等数据未数字化的情况,这可能导致数据遗失、前后统计维度不一等问题。

3 智慧图书馆方法发展趋势

本研究基于梳理得出的智慧图书馆方法体系关键要素与逻辑关系,结合实践调研所发现的数据整合与挖掘、知识资源处理、数据质量控制等问题,提出数据知识语义融合与映射关联、AI赋能自动化处理与智慧交互、数据治理标准化控制与风险预警三大发展趋势,以期为我国智慧图书馆建设提供参考指引。

3.1 数据知识语义融合与映射关联

目前信息分析相关研究仍聚焦在数据融合、知识融合层面^[38]。数据与知识融合驱动图书馆智慧化高速发展是未来发展的必然趋势。

多模态、异构数据与知识融合方法可分为三步,数据关联、知识融合、数据与知识融合。针对多模态数据映射关联,图书馆可以将非结构化、半结构化数据转化为结构化数据,利用语义关联与可视化技术,生成语义映射关系,实现对复杂数据的重构,采用注意力动态分配的知识嵌入方法,为多类型数据的模态架构注入知识,实现基于语义匹配与神经网络向量转化的知识融合与迭代推理。针对异构知识融合,图书馆可以基于深度学习、语义注释与映射、本体推理技术,将情境知识与用户知识融合到已有的知识图谱中,形成开放、

动态、自适应的知识体系。数据与知识融合则可分为特征层面的特征拼接与使用、模型层面的模型嵌套与综合、决策层面的融合汇总与判断^[38]，未来可继续深入研究结合图书馆具体应用场景的融合策略。

图书馆可利用数据知识融合方法来改进数据挖掘效果，以辅助业务决策、提供个性化推荐和情境化服务等。针对业务数据分析维度较少的问题，可基于业务运营的采访知识库、决策知识库，识别关键特征并主动调整其特征权重或弱化干扰，引导数据驱动模型在面临问题时快速找到突破口；或是在决策层面，结合业务数据分析与馆员知识经验，在活动策划、设备引进等情境下作出更科学全面的决策。例如采选场景下，可以基于 Hadoop、Spark 开展大数据分析，综合多馆历史采购数据、供应商的到书情况、读者使用数据、热门图书排行等进行快速计算、实时返回查询结果、评估分析热点与读者需求，结合馆员采选经验，进行判断决策，有效改善采选评估的盲目性、主观性和馆藏零借阅率。

针对现有用户画像的不足，在检索结果排序与页面呈现上，可基于 LDA 等文本挖掘技术，发现读者的高频借阅/检索类型，聚合移动端用户互动记录，精准感知读者检索需求。提高算法中读者偏好标签的权重占比，依据使用场景、所处空间与资源类别、界面风格的映射关系，进行页面调整与资源推荐，提供定制化的感官使用体验。

3.2 AI 赋能自动化处理与智慧交互

图书馆的主要职责就是利用数智技术，挖掘判定用户真实需求，对知识进行重组、融合，提供高质量知识服务^[27]。学者们从多角度探讨了新兴技术与图书馆结合的可行性与可能性，指引着图书馆实践发展的方向。据调研发现，目前图书馆引入的人工智能技术主要应用于智能问答，且部分省、市、区县级图书馆尚未引入或上线应用该技术。未来可依托人工智能技术在数据处理、语义理解、知识生成、智能交互等方面的优势，进一步实现图书馆业务处理流程的自动化，提升资

源整合加工效率，创新服务体验。

将人工智能技术引入业务流程中，可简化审批填报环节；嵌入资源关联整合，以实现高效特征提取、资源深度整合；嵌入知识加工，实现自动分类标引、改善知识融合效果；嵌入资源开发利用，实现知识产品自动生成、机器辅助阅读；嵌入问答系统，形成智能助手，增强图书馆与用户之间的交互。

针对资源关联与挖掘浅层化问题，借助人工智能技术，高效抓取网络信息资源，如不同图书馆官网信息等，以扩充图书馆资源发现类型；通过文本、图像等特征提取技术，深度识别资源特征；通过相似度计算、聚类分析等，挖掘相关主题/作者/学科/事件/年代等资源，辅助相关性推荐；利用共现分析、知识图谱技术，结合人工智能技术的自然语言处理能力，从标题、摘要、内容中快速提取关键信息，生成标签和分类，精准发现同一流派、理论演进等彼此脉络关联的资源，实现知识溯源与跨库资源的深度整合利用。

在智能问答方面，针对通用大模型在智慧图书馆领域应用效果较差的问题，可以采用检索增强生成的方式，引入专业知识经验如馆藏资源知识图谱等；或针对图书馆专有领域形成领域知识训练集，来进行答案矫正检验，提供权威数据来源，加速问题的泛化，提升分类处理效率，赋予其有别于通用大模型的、更具图书馆领域知识专业性的特征，实现读者与智能体交互，支持多模态信息输入，提升模型的理解能力与智能程度。

3.3 数据治理标准化控制与风险预警

基于标准的数据开放与利用，是基于互知理念建设智慧图书馆的必备条件。已有研究就用户画像建模等技术进行了深入讨论，然而，实践调研发现数据格式不统一等基础数据质量问题，严重制约着后续服务的改进与业务的升级，这要求智慧图书馆运行建设中应不断强化数据治理与规范的基础引领作用。

数据治理标准化控制应包括对数据处理过程中的质量控制规范以及数据合规、隐私保护等

法规条例的制定。针对数据处理标准体系不统一、不完善等问题,未来图书馆及有关机构应制定一系列如多源异构数据采集加工、智能问答语料建设、数据开放共享等方面的标准规范,通过统一数据接口、数据采集加工方式、数据存储方式等,实现数据格式转换、深度融合、跨地区/机构流通与共享,形成高质量、开放互通的数据库,建立一体化集成管理的数据中心。针对资源共享利用问题,图书馆可以借助云计算等技术,在遵循馆际合作规范的前提下,开放共享权限、接口或建立中间库,统一接入网络,实现资源互联、跨库访问、一站式检索。

同时,对于可能存在的隐私侵权、版权纠纷、数据安全等风险问题,可以建立自动化、全天候扫描监测与识别预警系统,通过大数据看板,查看数据与舆情的实时动态,及时识别并处理如敏感数据泄露等风险隐患。针对数据遗漏等问题,建立数据风险应急管理机制,实现数字化备份与建库管理,将读者信息库、电子资源元数据库等分类集成于数据中台,做到可追溯、全记录、可检索、可恢复的全生命周期管理和定期评估。针对用户数据的可用性问题的,图书馆界还需借鉴《中华人民共和国公共图书馆法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法律法规,完善智慧图书馆制度体系,加强行业规范管理,实现数据风险治理的制度化、规范化^[18]。在法律规范的许可下,对用户操作日志、基本信息等进行脱敏处理,实现用户群体共性特征和个性特征的提取与定位,提供个性化推荐、检索优化等服务,促进智慧图书馆建设持续健康发展。

参考文献

- 1 饶权. 全国智慧图书馆体系: 开启图书馆智慧化转型新篇章[J]. 中国图书馆学报, 2021(1): 4-14.
- 2 熊远明. 围绕国家文化数字化战略积极推进全国智慧图书馆体系建设[J]. 中国图书馆学报, 2022(4): 5-9.
- 3 王静, 王鹏. 基于数字孪生的高校智慧图书馆数据治理模型研究[J]. 图书馆学研究, 2023(11): 41-50.
- 4 刘莉, 邵波. 生成式 AI 赋能智慧图书馆的融合路径探析——以扎耶德大学图书馆为例[J]. 图书馆学研究, 2023(12): 34-43.
- 5 袁虎声, 等. ChatLib: 重构智慧图书馆知识服务平台[J]. 大学图书馆学报, 2024(2): 72-80.
- 6 曲哲, 廖宏建. 基于 Beacon 的情境感知在智慧图书馆中的场景应用[J]. 图书馆学研究, 2021(2): 58-66, 57.
- 7 尹婷婷, 郭永建. 数据驱动背景下智慧图书馆用户画像模型构建与研究[J]. 图书馆理论与实践, 2023(4): 81-86.
- 8 程秀峰, 等. 智慧图书馆建设过程中技术元素分析及知识图谱构建研究[J]. 图书情报工作, 2024(9): 123-136.
- 9 夏思洋, 朱学芳. 5G 环境下基于边缘计算的图书馆智慧服务响应能力研究[J]. 情报理论与实践, 2023(12): 21-27, 51.
- 10 孟佳琪, 支凤稳. 智慧图书馆知识服务平台研究[J]. 图书馆学研究, 2024(2): 85-91.
- 11 卢凤玲, 周兰羽. 国内智慧图书馆服务平台比较研究——以维普、超星和云瀚为例[J]. 图书馆理论与实践, 2024(1): 107-114.
- 12 张春春, 孙瑞英. 智慧图书馆用户数据合规治理机制研究[J]. 图书情报工作, 2024(4): 15-26.
- 13 Glaser B G, Strauss A L. The Discovery of Grounded Theory: Strategy for Qualitative Research[J]. Nursing Research, 1968, 17(4): 377-380.
- 14 朱纯琳. 基于数据流动的图书馆智慧服务生态系统构建研究[J]. 图书馆, 2021(1): 49-55.
- 15 覃玮境, 等. 融合与重构: AI 技术驱动下智慧图书馆服务逻辑与路径[J]. 图书馆工作与研究, 2019(3): 29-33.
- 16 郭亚军, 等. ChatGPT 赋能图书馆智慧服务: 特征、场景与路径[J]. 图书馆建设, 2023(2): 30-39, 78.
- 17 孙雨生, 祝博. 知识图谱加持的智慧图书馆信

- 息推荐架构体系研究[J]. 情报杂志, 2023(2):158-165.
- 18 陆康,等.“社会5.0”时代我国智慧图书馆数据风险治理研究[J]. 图书馆,2023(4):22-27.
- 19 卢小宾,等. 智慧图书馆数据标准体系研究[J]. 图书情报知识,2021(4):50-61.
- 20 罗飞,等. 大语言模型嵌入图书馆知识服务的风险范式与管控策略[J]. 图书与情报,2023(3):99-106.
- 21 赵杨,等. AIGC 驱动的智慧图书馆转型: 框架、路径与挑战[J]. 情报理论与实践,2023(7):9-16.
- 22 卢凤玲. 融合数据治理体系的智慧图书馆框架研究[J]. 图书馆,2021(5):74-78.
- 23 吴丹,等. 我国智慧图书馆信息服务的智慧化特征识别研究[J]. 图书馆建设,2024(2):9-20.
- 24 郭利敏,付雅明. 以大语言模型构建智慧图书馆: 框架和未来[J]. 图书馆杂志,2023(11):22-30,133.
- 25 廖运平,等. 大数据视域下智慧图书馆用户画像研究[J]. 国家图书馆学刊,2020(3):73-82.
- 26 肖花. 新时代智慧图书馆知识服务融合模式研究[J]. 新世纪图书馆,2019(3):48-51.
- 27 洪亮,等. 大数据驱动的图书馆智慧信息服务体系构建研究[J]. 图书与情报,2018(2):8-15,23.
- 28 魏大威,等. 基于区块链技术的智慧图书馆数字资源管理研究[J]. 中国图书馆学报,2022(2):4-12.
- 29 刘阳阳. 智慧图书馆体系下数据治理的法律模式选择[J]. 情报理论与实践,2023(3):90-97.
- 30 陈臣. 基于大数据挖掘与知识发现的智慧图书馆构建[J]. 现代情报,2017(8):85-91,97.
- 31 孙守强,李青青. 基于情感感知的智慧图书馆网络舆情服务体系构建[J]. 情报科学,2023(7):146-152.
- 32 顾佐佐,等. 智慧图书馆动态知识服务体系构建与平台设计[J]. 情报科学,2020(10):119-124.
- 33 李艳,等. 基于大数据挖掘与决策分析体系的高校图书馆个性化服务研究[J]. 图书情报知识,2016(2):60-68.
- 34 潘雪,陈雅. 泛在网络环境下我国公共图书馆智慧服务模式探究[J]. 情报科学,2018(5):30-34.
- 35 张文佳. 智慧图书馆建设中数据价值实现的机理及路径研究[J]. 图书与情报,2023(6):108-116.
- 36 周懿琼,蔚雷. 副省级公共图书馆数字资源建设情况调研分析[J]. 数字图书馆论坛,2022(11):60-66.
- 37 武丽娜,等. 电子资源管理过程中数据质量控制研究——以清华大学图书馆基于 ALMA 系统的实践为例[J]. 图书情报工作,2023(10):63-71.
- 38 卢小宾,等. 数智时代的信息分析方法: 数据驱动、知识驱动及融合驱动[J]. 中国图书馆学报,2024(1):29-44.
- (吴丹 教授 武汉大学信息管理学院, 杨馨梅 武汉大学信息管理学院图书馆学专业 2023 级硕士研究生, 梁少博 副教授 武汉大学信息管理学院, 何丽英 武汉大学信息管理学院图书情报专业 2023 级硕士研究生)

收稿日期:2024-05-29