



数智赋能视角下高校图书馆数据资源体系的研究与实践

——以华中科技大学图书馆为例

蔡婷婷 陈娇娇 肖芳 潘林强*

摘要 在教育数字化战略深入推进与高校“双一流”建设持续深化的背景下,高校图书馆作为学术资源枢纽,数据资源建设面临“重采购、轻建设”、资源利用率偏低等突出问题。文章以华中科技大学图书馆为研究案例,结合高校图书馆数据资源建设的共性问题提出优化建议,从数智赋能视角系统研究数据资源体系的构建原则与建设路径,构建“建设—管理—使用—评价”全链条闭环管理模式,并总结其在顶层设计、特色资源挖掘、智慧化管理及效果评价等方面的实践成效,为我国高校图书馆数据资源体系的智慧化、规范化转型提供实践参考与理论支撑。

关键词 数据资源体系 数据库管理 学术成果管理 资源使用统计 纸电融合

分类号 G250.7

DOI 10.16603/j.issn1002-1027.2026.02.003

引用本文格式 蔡婷婷,陈娇娇,肖芳,等.数智赋能视角下图书馆数据资源体系的研究与实践——以华中科技大学图书馆为例[J].大学图书馆学报,2026,44(2):25-36.

1 引言

在教育数字化战略深入推进与“双一流”建设攻坚的双重背景下,作为学术资源汇聚中心、知识服务枢纽和人才培养支撑平台,高校图书馆正面临前所未有的转型压力与发展机遇。《中国智慧教育白皮书(2025年)》明确提出,要推进教育数字化转型,实施人工智能赋能教育行动,建强用好国家智慧教育平台^[1]。这一要求与《教育强国建设规划纲要(2024-2035)》中“加快推进教育数字化转型,建设智能化校园,促进优质教育资源共享”的部署^[2]高度契合,不仅将高校图书馆的资源保障能力与知识服务水平提升至国家教育发展战略的高度,也对图书馆的数据资源体系建设提出了跨越式的治理要求。

当前,高校学习场景正发生深刻变革,已不再局限于传统的图书馆物理空间,而是逐步演变为集自主学习、科研创新、跨学科协作于一体的多元化、智能化学习生态,这种新型学习生态对数据资源的精准性、个性化、综合性提出了更高要求。传统图书馆以“文献收藏与借阅”为核心的服务模式,已难以适配新时代高校的改革方向,更无法满足新型学习生

态下师生的多元化需求。图书馆必须从被动的资源提供者,向主动的智慧化学习支撑中心、学术创新服务中心转型,而数据资源体系的重构与优化,正是实现这一转型的核心抓手与关键支撑。因此,各高校图书馆亟需引入新一代信息技术,将人工智能、大数据、云计算等技术与图书馆业务深度融合,推动图书馆进入“数智赋能”的新阶段。

与此同时,“AI+”战略正在高校广泛推进,为图书馆数据资源体系建设带来了新的业务增量与更高的质量要求。大模型、专用模型的研发与应用,需要大量高质量、专业化、特色化的语料资源作为支撑,而互联网上的公开数据普遍存在杂质多、真实性不足、学科深度不够等问题,无法满足高校学术研究与人才培养的精准需求。结合当前高校新型学习生态的发展趋势,师生的学习与科研行为已呈现出个性化、碎片化、跨学科的鲜明特征,不再满足于单一学科、单一类型的资源获取,而是需要多学科融合、多形式呈现的数据资源,AI驱动下的学习范式变革已成为必然,这种变革要求图书馆提供的资源不仅能满足师生的直接检索需求,更要适配个性化学习、跨

* 通讯作者:潘林强,ORCID: 0000-0002-4554-455X,邮箱:lqpan@mail.hust.edu.cn.



学科协作的新型需求。

当前我国多数高校图书馆的数据资源建设仍存在“重采购、轻建设”等突出问题,成为制约其转型发展的重要瓶颈。文献建设经费是图书馆经费的重要组成部分,但在数据资源“建设—管理—使用—评价—再建设”的闭环体系中,数据来源、使用绩效指标等核心信息均被数据供应商垄断^[3],导致图书馆在资源采购中处于被动地位,难以形成具有本校特色的自有馆藏,也无法实现数据资源的自主可控。

华中科技大学作为“双一流”建设高校,其图书馆始终立足学校学科优势与发展需求,积极响应国家教育数字化战略,主动适配高校新型学习生态的转型趋势,探索数智赋能下数据资源体系的建设路径。在“十四五”期间,华中科技大学图书馆以解决实际问题为导向,以数据资源全生命周期质量管理为核心,构建了“分级分类建设、自主可控管理、按需精准使用、多维量化评价”的业务闭环,在数据汇聚、流程重塑、服务创新等方面取得了阶段性成效。本文以华中科技大学图书馆为例,结合国内顶尖高校的相关研究和实践经验,系统研究数智赋能背景下高校图书馆数据资源体系的建设现状、面临的问题与挑战,总结实践经验与成效,展望未来发展方向,为我国高校图书馆数据资源体系的智慧化转型提供参考与借鉴。

2 相关研究综述

当前,我国高校正在推进“有组织人才培养、有组织科学研究、有组织社会服务”的战略转型,这一改革旨在发挥我国体制机制优势,实现高等教育的内涵式发展与弯道超车。图书馆作为高校教学科研的核心支撑,其数据资源服务需从“文献供给”向嵌入核心工作全流程、全场景的数智化服务转型。这一转型已成为国内高校图书馆的研究共识与实践方向。如北京大学图书馆构建的“数字教参系统+课程数据平台”,实现教参资源与课程进度、学生学习行为精准匹配,支撑有组织人才培养^[4]。清华大学图书馆围绕国家战略需求,以丰富馆藏资源为基础,以需求为导向建设双碳资源专题导航^[5]。华中科技大学图书馆通过融入学校治理体系提升服务能力,通过建立跨学科、跨部门的师生共商共议机制,有效带动了图书馆资源能力、服务能力和治理能力的全面提升^[6]。

2.1 新技术新应用带来图书馆的业务增量

人工智能、大数据等新一代信息技术的深度应用,正在重构图书馆的业务边界。当前各高校正在积极推进的各类“AI+”战略,对图书馆数据资源的精度和准确度提出了更高要求。与大模型所需的语料建设相对应,图书馆需要为高校的大模型提供“专业知识丰富”“内容准确且全面”“具有本校自身特色”的数据语料。这使得图书馆在原有面向人的数据组织和使用基础上,新增了面向大模型的数据组织和使用,即“机器阅读的数据资源”业务。这一业务增量并非简单的技术延伸,而是图书馆核心功能的质变。传统图书馆的数据组织以“人读”为最终目标,注重文献的揭示与获取;而面向机器阅读的数据组织,需要建立机器可理解、可计算的知识表示框架,实现从文献单元向知识单元的跃迁。南京大学2025年2月完成DeepSeek满血版本地化部署,7月率先实现“盘古 Pro MoE”大模型本地化部署。先后上线“科学导航”“南雍闻道”“研学助手”等智能平台,提供一站式学术知识库检索、智能问答、AI辅助选题、调研综述智能生成等功能^[7]。西安交通大学图书馆于2025年7月上线人工智能智慧学习平台,该平台基于自然语言处理、知识图谱、多模态分析等技术构建,依托图书馆资源与服务场景,提供智能文献管理、深度文献分析、动态知识图谱构建、智能问答等功能^[8]。

2.2 数据资源建设多为“重采购、轻建设”模式

根据2023年中国高校图书馆基本统计数据分析,全国高校图书馆电子资源购置费约为118.10亿元,均值为384.2万元,占馆均文献资源购置费的65.8%,达到历史最高点。在文献资源购置费中,电子资源购置费所占比例已超六成,且均值自2006年以来呈上升趋势^[9]。但由于数据资源建设中数据来源由供应商垄断,数据库之间的全文资源大量重复,理论上是买方市场,但图书馆缺少真实数据与评价抓手、缺少高水平数据管理与数据分析人才队伍,造成目前各高校图书馆文献建设经费均以采购“标准产品”为主,普遍面临“重采购、轻建设”的结构性问题。

夏立新等人的研究指出,近年来我国高校图书馆总经费在全校经费占比中呈现普遍下降趋势。以52所教育部直属高校为例,图书馆经费占比下降数量高于增加的数量。调查结果显示,反馈问卷的12



家图书馆中有 10 家表示因为经费预算短缺而减少购买使用频率较低的数据库,进而导致无法为教学和科研提供全面保障。该研究同时指出,外文资源高度垄断,导致外文资源数据库的价格涨幅居高不下。重要科技文献的获取存在诸多不确定性,由于外科技文献市场基本被国外资本控制,一旦发生战争或国际纷争等不可抗力事件,在我国访问和使用外科技文献将受到严重限制^[10]。这一困局的深层原因在于,图书馆长期处于资源流通链条的末端,缺乏对上游内容生产的介入能力,也缺乏对下游用户行为的洞察能力。图书馆采购的是“标准产品”,而非“定制服务”;积累的是“使用权”,而非“所有权”;提供的是“访问入口”,而非“知识资产”。这种业务模式使图书馆在资源建设中始终处于被动地位,难以形成具有自身特色的馆藏体系和数据资产。

3 高校图书馆数据资源体系建设面临的主要问题与挑战

在数智赋能的时代背景下,高校图书馆数据资源体系建设迎来了前所未有的发展机遇,但同时也面临着诸多深层次的问题与挑战。

3.1 学界与业界的“中试空白”

所谓“中试”,是指产品在正式投产前,从实验室研究向行业大规模工程化量产过渡的关键阶段。其核心目的是通过小规模、接近实际生产条件的试验,验证技术可行性、工艺稳定性、产品质量与成本经济性,从而打通科技成果从“书架”到“货架”的转化路径。国内高校图书馆学界的理论研究成果、国外引进的先进理念,在高校图书馆业务场景中也需要进行小规模试验、优化与完善,使其适配我国高校的组织架构、管理模式和师生需求,再进行大规模推广应用。但在实际工作中,尤其是从欧美引进的相关理念与产品,往往跳过中试环节,直接由企业按某个图书馆需求定制后便向全行业推广,导致这些理论与产品难以适配我国高校的实际情况,最终无法发挥预期效果,甚至被市场淘汰。

以机构知识库为例,其核心目标是汇聚高校的学术成果,实现学术资源的开放共享与高效利用,在国外高校和科研机构中得到了广泛应用。但该理念引入我国后,由于缺少高校图书馆行业的中试环节,直接由企业研发产品并在全国各高校图书馆推广建设,未经中试验证,其与我国高校的组织架构、学术

评价体系、师生使用习惯存在诸多不适配之处,更无法适配当前新型学习生态下师生的跨学科、个性化资源获取需求。在产品的工程应用中,由于我国高校的学术成果管理涉及教务处、科研院、各院系等多个部门,权责划分与国外高校存在显著差异,而机构知识库的建设往往由图书馆单独推进,难以实现与其他部门的数据互通与协同联动,导致学术成果的采集、审核、更新效率低下,数据准确性难以保障。同时,由于缺乏对师生需求的充分调研与适配优化,尤其是未充分考虑当前新型学习生态下师生的使用习惯和需求特点,与师生的实际需求脱节,使得机构知识库的使用率低,部分高校机构知识库的年访问量不足千次,最终许多高校的机构知识库逐渐陷入停滞状态。

3.2 数据指数增长的管理需求与图书馆数据处理能力不匹配

随着数字技术的快速发展,高校图书馆的数据资源呈现出“海量增长、多源异构、类型多样”的特点,数据规模已从千万级跃升至亿级,甚至部分顶尖高校图书馆的数据资源规模突破十亿级,且数据增长速度仍在持续加快。这些数据既包括传统的纸质图书、期刊、学位论文、专利等文献资源数据,也包括电子书、电子期刊、数据库等数字资源数据,还包括师生的借阅记录、资源使用日志、学术成果产出数据、课程教学相关数据等多类型数据,更新增了跨学科融合数据、个性化学习轨迹数据、科研协作数据等适配新型学习生态的特色数据。同时,这些数据来自不同的系统与渠道,格式各异、标准不一,既有结构化数据,也有非结构化数据,形成了复杂的多源异构数据格局,给数据治理工作带来了巨大的挑战,也对数据治理的精细化、智能化水平提出了更高要求。

传统图书馆的管理模式主要针对纸质图书的采编、典藏、流通等业务,采用的是精细化的人工管理与小规模数据处理方式,已无法应对海量多源异构数据的治理需求,更无法适配新型学习生态下对数据资源精细化、个性化、高效化的管理要求。在数据资源管理尤其是元数据管理方面,问题尤为突出。元数据作为数据资源的“身份证”,因数据多源异构的特征,数据管理存在混乱、重复、质量低下等问题,且缺乏统一的元数据标准,无法实现多源异构数据的有效整合,难以实现元数据的精准管理,导致篇章



级数据管理混乱,师生在检索资源时经常出现重复结果、无效结果,无法满足新型学习生态下师生快速获取精准资源的需求,严重影响了资源使用效率。

3.3 数据资源生命周期无法自主可控

数据资源的生命周期涵盖建设、管理、使用、评价、再建设等多个环节,是一个闭环式的管理过程。当前多数高校图书馆的数据资源难以形成特色、无法实现自主可控,进而影响了服务的精准度与图书馆的核心竞争力。

在数据资源建设环节,数据资源交易市场卖方话语权过重,高校图书馆在资源采购中处于被动地位,核心数字资源市场常被“一家独大”的供应商垄断,供应商不仅逐年涨价,外文核心数据库年费动辄几十万甚至上百万,且每年以5%—10%的幅度上涨。部分供应商还采取“捆绑销售”模式,将冷门资源与核心资源打包,使得图书馆不得不“花冤枉钱”,进一步加重经费负担。数据供应商垄断了数据来源、使用绩效等核心信息,图书馆无法准确掌握资源的实际使用情况与真实价值,只能被动接受供应商提供的标准产品,难以根据学校的学科优势进行数据资源建设。

在数据资源管理与使用环节,当前多数高校图书馆的数据资源主要依赖商业供应商,数据的存储、管理、维护等环节均由供应商负责,图书馆缺乏对数据的自主管理权,无法为各个数据库的数字资源建立档案集中管理、集中使用。例如,虽然各种类型的云端系统提供统一搜索等发现与导航功能,但由于资源供应商之间的竞争与垄断,并不是所有资源的搜索、发现都能通过一个发现系统完成。同时,作为“资源身份证”的元数据存储于云端数据库商处,对于不同数据库间资源重复性的判断仅能依靠供应商数据,对于数据库之间资源重复率的真实性难以判断,存在较多资源(如重要学术期刊)重复建设的问题。

在数据资源评价环节,由于图书馆缺乏真实使用轨迹数据,评价工作难以发挥导向作用,进一步制约了数据资源体系的优化与完善。当前多数高校图书馆的资源建设模式为采购数据库模式,即租用云端数据使用权模式。对于租用某类型的资产,其主要建设依据为该资产的使用绩效,但图书馆对于资源绩效的评价依赖供应商提供的数据,这些数据往往存在片面性,无法真实反映资源的实际使用效果

与价值。

3.4 图书馆信息化市场容量有限带来企业研发的困境

与移动互联网、能源、金融等行业相比,高校图书馆行业的市场规模较小,对信息化产品的需求总量有限,且高校图书馆的经费预算相对固定,对信息化产品的价格敏感度较高,导致企业在高校图书馆领域的盈利空间有限。从行业投入数据来看,尽管2025年高校图书馆数字化建设持续升温,大额项目集中释放,但从整体来看,全国高校图书馆的信息化投入仍处于较低水平,且多数高校图书馆信息化投入仅能保障设备设施、信息系统正常运维。行业市场总额远低于能源、金融等一线行业的投入水平,难以支撑业内企业高额的研发投入,尤其难以支撑人工智能等新兴技术原始创新所需的投入。因此,多数企业采取“借鉴成熟模式、引进溢出人才”的策略,通过借鉴移动互联网、能源、金融等行业已经成熟的产品与技术、招聘一线大厂分流的工程师等方法研发新产品,严重制约了高校图书馆业务场景中新技术的快速落地,只能等待一线行业技术与人才的“溢出”。

综上所述,当前高校图书馆数据资源体系建设面临的“中试空白”、海量多源异构数据治理困难、特色化与自主可控能力不足、新技术落地滞后等问题,相互交织、相互影响,形成了制约图书馆智慧化转型的多重瓶颈,也严重影响了数据资源体系对新型学习生态的支撑能力。解决这些问题,需要高校图书馆立足自身需求,加强顶层设计,完善数据资源全生命周期治理体系,强化人才队伍建设,推动数据资源的特色化、自主可控建设;同时,需要推动学界与业界的深度协同,搭建中试验证平台,促进理论成果与实践应用的深度融合,引导企业加大研发投入,开发适配高校图书馆业务场景和新型学习生态需求的技术产品,优化行业生态。

4 华中科技大学图书馆的数据资源工作理念

华中科技大学图书馆在“十四五”深化实践与“十五五”战略推进的关键时期,紧扣学校“双一流”建设核心任务,以数据资源全生命周期管理为主线,全面推动质量管理体系落地生根,构建“评采管用”全链条业务闭环。在体系规划与设计前开展“中试验证”,在顶层设计中强调数据资源的统筹与共享,



在建设环节重点关注各级各类数据资源的特征,在管理环节以标准为抓手促进数据资源的自主可控,在使用环节坚持用户需求导向提升服务精准度,在评价环节以量化的方式提高评价质量,实现数据资源从建设到评价的精细化、智能化治理,为学校教学科研与学科发展提供高质量文献信息保障。

4.1 协同多方力量,开展“中试验证”工作

立足学校基础预研优势、依托合作企业工程化研发能力,深度结合图书馆海量数据管理、学术资源服务等核心业务场景,协同开展“中试验证”相关工作,推动前沿技术从理论研究向实际业务落地转化,打通技术研发与场景应用的壁垒,提升数据资源治理的技术支撑能力。如依托学术论文、专利等多类型数据资源的“时间局部性”原理,联合学校计算机学院研究团队,重点开展海量(十亿级)元数据的去重、合并及关联算法研究,将研发的算法模型在图书馆学术资源管理场景中进行全流程中试验证,持续优化算法的精准度与处理效率,最终实现智能管理学校所有学术成果的目标,从每年 20 多万条原始异构数据中及时完成去重、合并,形成约 3 万条标准化、高质量的权威学术成果数据,为学校“一张表”工程、科研评价、学科分析等提供可靠的数据支撑。

4.2 强化顶层设计,规范数据资源体系建设

坚持“顶层统筹、分类施策、逐步推进”的原则,通过完善顶层规范、搭建统一平台,引导数据资源体系标准化、集中化建设,打通各系统数据孤岛,实现数据资源的有序流转与高效共享,夯实数据质量管理的体系基础。如在“十四五”期间以开放平台建设为抓手,统筹规范各业务系统间的数据交互流程,采用“新系统新办法、老系统老办法”的差异化策略,通过开放平台明确约束 2023 年及以后上线的信息系统,要求其严格遵循开放平台数据标准进行数据交互,同时逐步引导存量老系统下线或进行数据优化,稳步推动全馆数据资源的标准化整合与集中化管理。

4.3 立足业务特征,厘定数据资源分类建设

结合图书馆各业务板块的核心特征与需求差异,厘定数据资源分类建设标准。针对不同业务场景,精准规划数据建设方向、明确建设重点,推动数据资源供给与业务运行的深度适配、同频推进,切实提升资源建设的针对性与实效性,让数据要素高效赋能各业务板块的运转。如利用 Meta 系统,聚焦纸

电图书一体化管理核心业务,重点推进纸电图书资源的整合建设,通过完善纸电图书采编元数据及纸质图书典藏、流通等核心数据,实现纸电图书资源的一体化管控与高效利用。利用 Alma 系统,围绕数据库采购与管理业务,聚焦流程优化与数据规范,系统梳理数据库采购、权限分配、使用统计、续费评估等全流程数据节点,提升数据库资源管理的精细化、标准化水平,确保数据库资源精准匹配学校教学科研需求。

4.4 锚定自主可控,夯实数据管理标准

以数据资源自主可控为核心落脚点,完善数据管理标准规范,强化本地数据标准的制定与核心资源的平台建设,筑牢数据安全与质量防线,确保数据资源的规范性、安全性与可复用性,为数据资源全生命周期管理提供坚实保障。如重点推进本地元数据标准体系建设,制定适配图书馆业务场景的元数据规范,实现元数据的自主管控与标准化管理;推进课程资源研学平台建设,自主整合各类课程相关数据资源,规范数据采集、存储与使用流程;开展 AI 语料资源建设,筛选、整理适配图书馆服务场景的优质语料数据,为 AI 技术在教学与科研的应用提供自主可控的资源支撑。

4.5 紧扣使用需求,明确数据资源使用的广度和深度

以用户使用需求为导向,针对不同服务场景的用户需求,差异化界定数据资源的开放范围与应用粒度,实现数据资源的精准供给,兼顾服务的覆盖面与精细化水平,提升用户使用体验与资源利用效率。如针对广大师生资源获取需求,推行二维码纸电融合服务,满足大众化、高频次的数据资源服务需求,覆盖各类纸电文献、资源检索等基础场景,满足师生日常学习科研需求;针对科研团队、学科带头人等群体的精细化需求,重点建设学术成果库,提供小范围、高精度的学术成果数据服务,涵盖成果详情、引用关系、学科关联等精细化内容,支撑科研创新与学术研究。

4.6 深化量化机制,推动数据资源绩效评估提质增效

以量化指标为核心抓手,打破传统依赖数据库商提供数据的粗放式评估模式,构建相对全面、客观、可信、持续、可对比的电子资源绩效评估体系,反映电子资源建设质量与使用效益,推动电子资源优



化配置与服务升级。如重点推动建设基于网络日志数据的电子资源使用评价体系,通过量化电子资源访问次数、访问人数(日活、月活等)、总流量、平均使用时长等核心指标,实现用户级和院系级的电子资源统计分析,精准掌握电子资源使用情况。同时,联合教科院团队开展关联分析研究,深入挖掘学生个人学业成绩与其图书馆资源利用行为之间的内在关联,以数据驱动的方式验证和提升资源建设的绩效。

5 华中科技大学图书馆的数据资源业务实践

5.1 顶层设计

在图书馆信息化建设进程中,为满足多样化的业务需求,往往需要多家服务商为图书馆定制开发业务系统,提供完整闭环的信息化服务。然而,不同系统相互隔绝独立,数据不互通,业务不互联,极易

形成“孤岛效应”。传统的系统与系统之间点对点对接模式面临多重困境:定制化开发导致重复建设问题突出;多方对接协调费时费力;黑盒模式下问题暴露滞后;安全措施不足,数据安全缺乏保障。为解决上述问题,华中科技大学图书馆于“十四五”时期提出开放平台的建设构想。该平台的设计目的是为参与图书馆建设的厂商提供统一的应用标准、数据管理、安全认证机制和基础能力支撑,让图书馆的信息化建设更加标准、规范、可扩展。以开放平台为枢纽,实现数据共享、标准统一、多维监管、安全保障的建设目标。

开放平台采用“开放接口—数据总线—数据仓库”三位一体的核心架构,构建起覆盖全馆业务系统的数据枢纽,具体如图1所示。

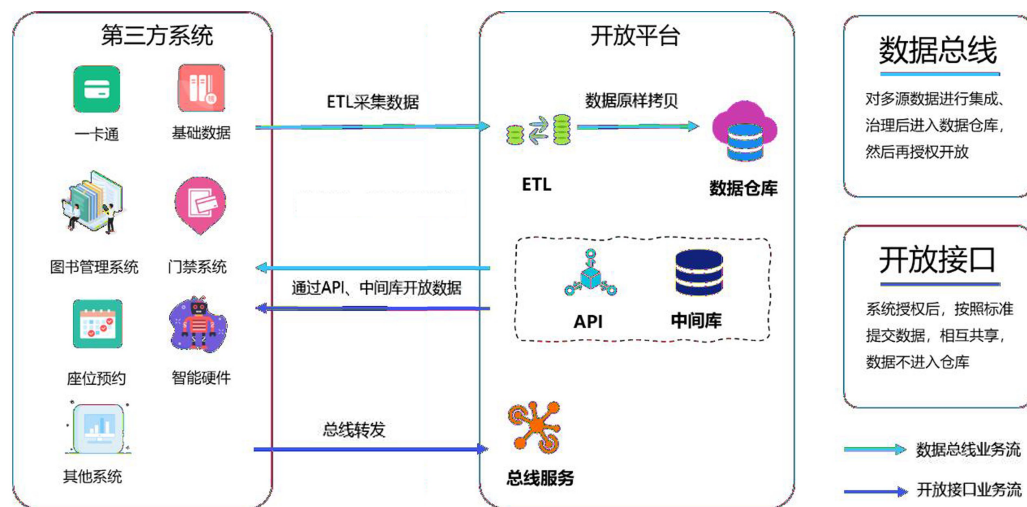


图1 开放平台与第三方系统对接结构图

开放接口实现业务系统间小批量数据的实时调用。接口管理涵盖接口/数据库参数定义、文档生成、上下架管理、流量监控、权限管理等功能,为第三方服务商和学校其他部门提供标准化的数据接入途径,实现接口粒度的权限控制与流量监控。

数据总线由数据同步服务和调度中心组成,采用ETL(Extract-Transform-Load)工具实现数据同步,中心化的收割和管理图书馆不同业务系统的数据,并提交给数据仓库做统一的清洗加工和计算,实现大批量数据的非实时同步。在数据一致性保障方面,数据总线设计了主动规避、异常处理和监控告警三套机制。

数据仓库负责提供数据治理、存储、查询服务。数据存储采用云端与本地结合的方式:所有用户数

据、资产数据、业务数据均在图书馆本地保存以保障数据安全;文献资产数据和运行数据定期由云端集群收割,生产出成品数据后同步回本地供图书馆使用。

目前开放平台已完成一期功能建设,图书馆近三年上线系统的数据同步和接口对接均转移到开放平台,图书馆对学校其他部门提供的数据和服务也均通过开放平台进行对接,实现数据统一分发与业务互联。

5.2 数据资源建设

数据资源建设是信息化建设的核心根基,是实现资源高效利用、精准服务的前提,重点围绕纸电图书资源整合与数据库采购线上化管理两大方向推进。



通过借鉴传统数据库采购业务中“总库”与“子库”的概念,华中科技大学图书馆对数据库总库和子库进行了可量化的定义,总库定义为同一域名下的数据资源集合,子库定义为最小采购单位。针对电子书单本采购带来的子库定义难以应用于工程实践的问题,华中科技大学图书馆将电子书纳入图书建设流程,即纸电图书统一建设。纸电图书的建设依托 Meta 系统实现采购、编目、典藏、流通全流程智能化,数据库建设借助 Alma 系统完成资源整合、采购管理全链条优化,两大系统协同发力,构建起标准化、高效化、智能化的数据资源管理体系,为后续服务升级与业务拓展提供坚实的数据支撑。

2025 年底,华中科技大学图书馆正式上线 Meta 系统,替换自 1999 年使用的原图书管理系统(Sierra)。Meta 打破传统模式下纸质图书与电子图书分开采选、分别下单的业务壁垒,系统基于开放平台统一的元数据标准,让采访人员在一个界面内即可完成纸本图书与电子图书的比对、选择与订购,有效避免了重复采购。同时系统支持多维度数据分析,能够根据学科建设需求、借阅流通数据及馆藏结构,自动生成科学的采购建议,实现从“经验驱动”向“数据驱动”的采购模式转变。系统还集成了读者荐购功能,读者提交的荐购请求可实时推送至采访模块,图书馆员能够即时处理并反馈订单状态(如已订购、已到馆、已加工等)。这种“读者发起—馆员响应—系统反馈”的闭环管理机制,极大地缩短了读者需求与资源建设之间的响应时间,提升了读者参与感与满意度,使馆藏结构更加贴合用户实际需求。相对于老旧 Sierra 系统功能单一、操作繁琐的弊端,Meta 通过自动化 workflow 显著提升了内部处理效率,它整合了供应商管理、订单发送、财务结算、验收登到等全链条环节,实现了从预算到入藏的全生命周期管理。此外,系统内置的智能验收功能,通过与物流数据的对接,能够实时追踪图书到货状态,确保资源建设的时效性与透明度。

2026 年 4 月,图书馆将启动 Alma 系统的建设,依托其强大的电子资源管理模块,实现各类数据库(电子期刊数据库、学位论文数据库、会议论文数据库、特色资源数据库等)的采购、续订、停用等全流程线上办理,全面提升数据库资源建设的管理水平和服务质量。

5.3 数据资源管理

“十四五”期间,华中科技大学图书馆以质量管理为支点,促进馆藏资源发展与学校学科发展同步,建设文献资源质量信息管理体系,包括纸电图书一体化管理系统、数据库管理系统,逐步构建起高效、安全、可追溯的数据资源管理体系,为各项业务开展提供坚实的数据支撑。

在数据库业务方面,2023 年华东理工大学图书馆建设了文献资源质量信息管理平台,以标准化建档、全生命周期管控、跨平台信息共享为核心,为各类数据库资源建立统一、规范的电子档案。档案内容涵盖数据库厂商及联系人信息、最小采购单位子库详情(含数据库介绍、访问方式、收录期刊列表、资源类型等关键信息)等,支持一次录入、多次复用,大幅降低重复填报与信息核对成本。系统通过定义统一的数据结构与约束规则,实现对每条数据库档案从建档、更新、停用至归档的全流程规范化管控,并依托开放平台,将权威、准确的数据库揭示信息直接同步至数据库导航系统,实现资源信息高效共享。同时,数据库档案里收录的期刊列表为馆藏资源的纸电一站式检索提供了精准的资源索引和高效的检索支撑。此外,系统对数据库历史订购记录进行永久留存与可追溯管理,为资源采购决策、经费统计、资产盘点提供完整、可靠的数据支撑。

在纸电图书方面,Meta 系统重点聚焦纸质图书与电子书的一体化管理。该系统依托强大的数据处理能力,遵循开放平台的数据标准,将纸质馆藏、电子图书(包括单本订购的电子书以及自建电子书等)等异构资源进行归一化整合,形成统一的元数据,打破了传统图书馆纸质图书与电子书孤立管理的局面,实现了“纸电一体”的深度融合与智慧化运营,为纸电资源的一站式检索与发现提供了坚实的基础。

5.4 数据资源使用

5.4.1 利用二维码进行纸电图书馆的融合阅读

“多维融合的二维码书标系统”通过在传统图书书标的基础上集成图书资源二维码、盘点二维码、图书入藏年份,实现馆藏资源的纸电融合与轻量化的图书盘点。该系统通过为每本纸质图书绑定一个永久的资源二维码,建立资源二维码与电子资源的映射关系,并深度对接学校统一身份认证与图书馆门户平台,实现“一书一码,扫码直读”,在读者使用场景完成纸电资源的融合。同时,利用印刷成本较低



的盘点二维码作为视觉锚点,通过批量识别书脊上的盘点二维码,不涉及图像分割与文字识别,对算法要求较低且识别速度更快。书脊上的入藏年份便于读者服务部快速分拣图书。与基于RFID的盘点方案相比,实施成本低,识别准确度高,且在每层书架能做到精准排序。馆员可通过移动端、盘点设备等拍照进行图书盘点,操作复杂度低。对比基于机器人的盘点方案,实施和维护成本低,可作为馆员日常工作常态化开展。盘点的数据可用于构建智慧书库,通过VR导览查看分馆各阅览室各层架上的图书,实现图书馆实体空间资源的数字化映射。读者在图书馆门户网站检索图书后,通过智慧书库定位到图书所在层架,便于读者快速找书,还可浏览同架位或相邻架位图书信息,扩展图书馆资源揭示渠道。若纸本馆藏全部借出,而图书馆已购买对应电子资源,读者可直接线上阅读电子资源,解决热门图书复本不足的问题,充分发挥纸电资源各自的优势,提高图书馆已购资源的揭示效率。

2024年9月起,华中科技大学图书馆所有入藏新书全部启用二维码新书标,至2025年底已完成三个馆区在架图书的二维码书标替换与图书盘点。

5.4.2 管理和使用全校学术成果数据

在数据资源体系建设中,学术成果库是连接图书馆数据服务与学校核心业务的关键枢纽。华中科

技大学图书馆自2017年启动机构知识库建设,历经四期持续迭代,逐步构建起全校学术成果“取领管用”全周期治理体系,具体如图2所示。在“取”的环节实现多源数据整合与清洗去重。学术成果库整合了来自Web of Science、中文期刊库、学校科研管理系统、人事系统、研究生系统以及国家知识产权局专利数据等多源数据。图书馆每季度进行一次全校科研成果数据获取,将成果记录进行去重、合并、补齐,根据署名单位与院系对应关系,将成果记录推送至院系。以2021年为例,收集全年学术成果7734万条,清洗汇总华中科技大学25万条成果记录,计算机去重与数据核查后形成2.9万条高质量数据。在“领”的环节,由图书馆保障成果数据来源、成果记录基础字段完整性、成果准确性与成果推送及时性,学院管理员只需根据工作安排,认领并确认本学院成果,大大减轻了学院和教师对学术成果数据管理的负担。在“管”的环节,学院根据实际情况进行本学院教师学术成果的管理。图书馆协助学院开展系列平台使用培训,为学院培养1—2名学术成果数据管理员。通过该模式,分散年底集中核对的压力,让学术成果库成为学院科研成果管理的常用平台。在“用”的环节,图书馆将全校学术成果数据缴存到学校基础数据库,供全校各单位使用。

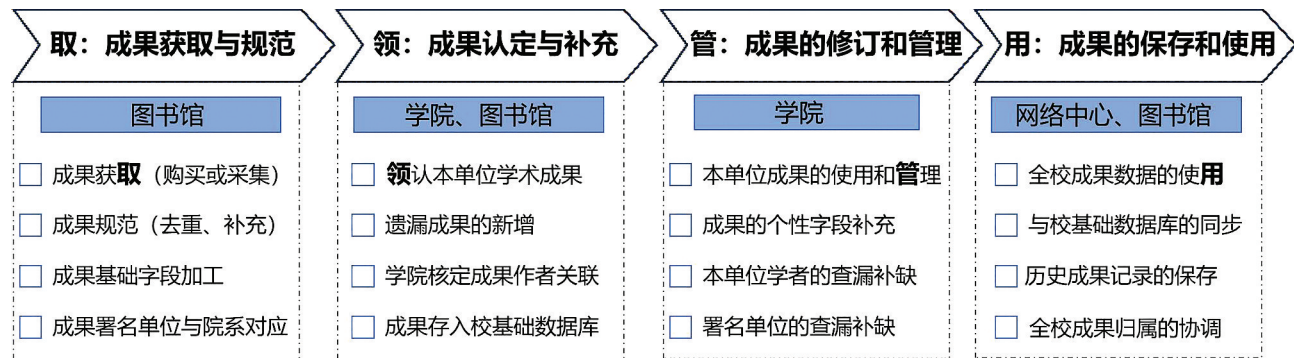


图2 学术成果库“取领管用”的业务流程图

通过建立高效的“取领管用”工作机制,数据准确率从2018年的不到80%提高至98%以上。以2025年为例,图书馆协助全校43个单位完成2024年度学术成果集中核对工作,累计核查并修订论文27425篇、专利1907项、图书163部,汇总形成1824份教师个人成果档案册,学术成果数据在学校职称

评审、聘期考核、项目申报等场景广泛应用。

5.5 数据资源评价

数据资源评价是图书馆资源生命周期管理的关键一环,是优化资源配置、提升服务质量、发挥数据价值的核心环节,不仅有助于优化采购经费配置、调整馆藏结构,更能为学校的人才培养和学科建设提



供数据驱动的决策支持。

为全面掌握全校电子资源的使用状况,华中科技大学图书馆于2025年建设了数据库使用统计系统(见图3)。该系统从校园网络出口采集非结构化的Web日志,通过构建大数据网络分析平台进行网络日志的清洗、解析与统计分析,同时整合用户上网登录认证信息,包括校园网上网认证、免认证登录、VPN登录及其它校外访问系统认证登录等,实现按总库/平台(即相同域名)的用户级、院系级等更精细维度的电子资源使用统计。本系统面向总库,可从访问次数、访问人数、访问人次、总访问时长、平均访问时长等统计维度开展用户行为与资源利用分析。相较于

传统Counter统计报告,更能反映用户真实使用强度与粘性,访问人数、访问人次、平均访问时长,可有效刻画用户规模、访问频次、使用深度与持续活跃度,能够降低异常流量干扰,客观反映数据库实际使用效益与读者真实利用情况,为资源评价与采购决策提供更稳健且自主可控的数据支撑。表1、表2为2025年秋季学期(2025年9月1日至2025年12月31日)师生访问数据库情况的统计示例。其中,访问次数代表用户访问某平台域名的总次数;访问人数代表在一个统计周期内的去重活跃用户数;访问人次代表在一个统计周期内每天的活跃用户数之和;访问时长代表用户单次访问的停留时间。



图3 数据库使用统计系统

表1 2025年9—12月秋季学期各数据库使用情况(TOP 10)

序号	平台名称	访问人次	访问人数	总访问时长(h)
1	IEEE Xplore Digital Library	769134	50080	81578
2	ScienceDirect	573377	31821	192073
3	CNKI 知网	543134	41703	104550
4	超星期刊	430404	36780	174876
5	Nature	287071	22323	41705
6	Springer Link	274831	29797	36016
7	ACS Publications	170149	12491	35200
8	科睿唯安	100177	17590	21782
9	万方数据知识服务平台	99002	21759	25394
10	arXiv 平台	70758	11961	11335



表2 2025年9—12月秋季学期校内各学院数据库使用情况(TOP 10)

序号	学院名称	访问人次	访问人数	总访问时长(h)
1	机械科学与工程学院	284840	4085	60415
2	光学与电子信息学院	246648	3194	54250
3	材料科学与工程学院	204024	2358	40212
4	能源与动力工程学院	203101	2440	40637
5	化学与化工学院	186521	1323	34563
6	电气与电子工程学院	186150	3527	41427
7	物理学院	170211	1894	27026
8	生命科学与技术学院	168997	2653	31818
9	公共卫生学院	160785	1267	32986
10	基础医学院	147795	2387	35879

该系统可直观呈现各类电子资源的实际使用效能,能助力图书馆对高需求、高利用资源的优先保障,对低利用率资源开展调研与优化配置,避免资源闲置浪费。同时,系统可精准定位高频使用学科与重点用户,支撑学科馆员开展针对性的信息素养教育与学科服务;长期积累的使用数据,也为数据库服务质量评估与续订议价提供了可靠依据。

为深入挖掘图书馆数据在辅助教学、促进学风建设方面的潜在价值,2026年图书馆与学校教育科学学院(以下简称教科院)联合成立“学习分析与数据支持”跨学科研究小组,推动资源利用绩效的分析。研究小组以学生访问图书馆各类资源的行为数据、进馆数据、借阅数据等为基础,结合其学业成绩等进行关联比对,分析学生的学习习惯,探究数据资源使用效率与学业质量之间的内在联系,聚焦学生学业发展全流程,探索将图书馆资源使用数据与学生学业表现相关联的研究范式。研究成果可反哺数据资源优化,推动数据库资源的精准补充与适配,形成“数据采集—分析研究—资源优化—服务提升”的良性循环,助力教育教学质量的持续提升。

6 总结与展望

华中科技大学图书馆立足学校“双一流”建设需求与教育数字化战略要求,以解决实际问题为导向,以数据资源全生命周期质量管理为核心,构建“分级分类建设、自主可控管理、按需精准使用、多维量化

评价”的业务闭环。在顶层设计方面,通过搭建开放平台,统筹规范各系统间的数据调用与同步,推动数据资源的标准化、集中化,为数据资源体系建设奠定了坚实基础;在数据资源建设环节,通过Meta系统、Alam系统等专业化平台,实现了纸电图书、数据库等资源的精细化建设,兼顾了资源的通用性与特色化;在数据资源管理环节,强化本地元数据及标准建设,推进课程资源与AI语料资源建设,逐步提升数据资源的自主可控能力;在数据资源使用环节,通过本地化的统一搜索、统一资源访问、二维码纸电融合、学术成果库等特色服务,实现了资源的精准推送与高效利用;在数据资源评价环节,突破传统评价模式,构建了基于网络日志数据的资源使用评价体系,联合相关院系开展学业成绩与资源利用关系研究,为数据资源的优化升级提供了科学依据。

上述理论验证与工程实践的相关工作,是华中科技大学图书馆、校内有关单位(如计算机学院等)和企业的长期合作、联合攻关、多次“中试”验证后取得的工作实绩。图书馆为工程应用提供验证的业务场景与空间,校内科研单位与图书馆进行预研与原型验证,企业依据原型进行中试,较为有效地解决了数据资源体系建设中面临的“中试空白”、数据治理低效、缺乏自主可控性等问题,推动了数据资源体系的高质量发展。

同时,也需清醒地认识到,高校图书馆数据资源体系建设是一项长期的、系统性的工程,不可能一蹴而就。当前,无论是华中科技大学图书馆,还是国内



外其他高校图书馆,在数据资源体系建设中仍面临诸多不足:数据治理的精细化水平仍需提升,多源异构数据的整合与共享能力有待加强;特色资源建设的规模与质量仍需扩大与提升,自主可控能力仍需进一步强化;人工智能、大数据等新技术的应用深度不够,智慧化服务水平仍有较大提升空间;学界与业界的协同联动机制仍需完善,“中试空白”问题尚未得到根本解决。这些问题,既是当前高校图书馆数据资源体系建设面临的挑战,也是未来发展的重点方向。例如,为缓解资源市场中“卖方话语权”过重的问题,图书馆还需深化资源与数据建设模式,一方面需将现有的资源采购模式转变为建设模式,区分OA资源、商用资源并进行差异化建设,另一方面也需要尝试对数字资源出版商开展直通采购模式。

展望未来,在“AI+”时代背景下,高校图书馆的数据资源体系建设将迎来新的发展机遇,也将面临更高的要求。随着人工智能、大数据技术在数据挖掘、知识图谱构建、智慧荐购等方面的深入应用,图书馆的数据资源体系将更加完善,数据治理的精细化水平、资源服务的精准度将得到显著提升。未来,高校图书馆的数据资源体系建设可重点聚焦三个方面:一是持续强化顶层设计,完善数据资源全生命周期治理体系,推动数据资源的标准化、集中化、智能化管理,破解多源异构数据治理难题,提升数据质量与治理效能;二是加大特色资源建设力度,聚焦学校的学科优势与发展需求,规模化建设具有本校特色的自有馆藏与AI语料资源,不断提升数据资源的自主可控能力,打造核心竞争力;三是深化新技术与业务的融合应用,推动人工智能、大数据等技术在资源采购、检索、推荐、评价等全流程的深度应用,构建智慧化的知识服务体系,实现从“数据服务”向“知识服务”的转型,更好地支撑学校的“有组织人才培

养、有组织科学研究、有组织社会服务”。

参考文献

- 1 教育部.《中国智慧教育白皮书》发布[EB/OL]. [2025-05-17]. http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2025/2025_zt06/dongtai/202505/t20250517_1190910.html.
- 2 中共中央,国务院.中共中央 国务院印发《教育强国建设规划纲要(2024—2035年)》[EB/OL]. [2025-01-19]. https://www.gov.cn/gongbao/2025/issue_11846/202502/content_7002799.html.
- 3 王丽娜.外刊数据库垄断趋势下的购销博弈[J].图书馆工作与研究,2016(10):33-36,81.
- 4 张元俊,王昊贤,李国俊,等.“信息协同”视角下嵌入数字校园的高校图书馆数字教参服务——以北京大学数字教参资料系统为例[J].大学图书馆学报,2022,40(2):27-35.
- 5 田兆雪,任奕,王媛,等.用户需求导向的交叉学科专题信息资源导航建设实践——以清华大学双碳信息导航为例[J].图书馆情报工作,2024,68(2):41-49.
- 6 王敏,张绚丽,施亮.华中科技大学图书馆资源服务一体化的创新路径与实践探索[J].大学图书馆学报,2025,43(4):42-48.
- 7 南京大学.【南大“奋进行动”1.0答卷】数字化赋能发展取得新成效[EB/OL]. [2026-02-04]. <https://hnc.nju.edu.cn/info/1884/35971.htm>.
- 8 西安交通大学.西安交大图书馆上线AI智慧学习平台 创新科研新范式[EB/OL]. [2025-07-04]. <https://news.xjtu.edu.cn/info/1219/223476.htm>.
- 9 吴汉华,王波.2023年中国高校图书馆基本统计分析[J].大学图书馆学报,2025,43(2):103-114.
- 10 夏立新,周鼎,叶光辉.我国文献信息资源保障体系的重构[J].中国图书馆学报,2023,49(6):27-40.

作者贡献说明:

蔡婷婷:论文撰写及修改

陈娇娇:部分内容撰写及修改

肖芳:论文框架设计,校对及修改

潘林强:论文校对及修改

作者单位:华中科技大学图书馆,湖北武汉,430074

收稿日期:2026年3月10日

修回日期:2026年3月18日

(责任编辑:支娟)

Research and Practice on Data Resource Systems Under the Digital Intelligence Empowerment in Academic Libraries: A Case Study of Huazhong University of Science and Technology Library

CAI Tingting CHEN Jiaojiao XIAO Fang PAN Linqiang

Abstract: Amid the accelerating digitalization of higher education and the continued deepening of Double First-Class initiatives, academic libraries are undergoing a profound transformation in data resource development. As



central hubs of scholarly information, they face prominent challenges, including overreliance on resource acquisition at the expense of system construction, low utilization rates, vendor dependency for core data, inconsistent metadata standards, and fragmented evaluation systems. These constraints impede the construction of distinctive library collections and limit libraries' ability to meet increasingly diverse user information needs within emerging learning ecosystem. Taking Huazhong University of Science and Technology Library as a research case, this paper explores the principles and pathways for developing data resource systems from the perspective of digital intelligence empowerment. Adopting an integrated research method combining theoretical analysis, pilot validation, and engineering practice, this study proposes a full-chain closed-loop management model encompassing “construction—management—utilization—evaluation” centered on lifecycle-oriented quality management of data resources. At the implementation level, the Library has established a standardized and centralized data resource aggregation framework through the top-level design of open platform. Refined management of print and electronic collections, as well as databases are achieved by specialized systems including Meta and Alma. Simultaneously, autonomous control of local metadata and AI corpus resources were strengthened. Resource utilization is optimized via characteristic services such as integrated search, print-electronic integration, and research output management. In addition, a data-driven evaluation system based on web log analysis has been established to provide a empirical basis for resource allocation and decision-making. The paper demonstrates that digital intelligence empowerment serves as the core driving force for transforming academic libraries from passive resource purchasers into proactive data managers and knowledge service providers in data resource construction. In the future, academic libraries should focus on three aspects: refining data governance, scaling up characteristic resource development, and deepening the application of artificial intelligence technology, so as to drive the transformation of library services from data services to knowledge services. The research provides theoretical support and practical reference for the intelligent and standardized transformation of data resource systems in academic libraries in China.

Keywords: Data Resource Systems; Database Management; Research Output Management; Resource Utilization Statistics; Print-Digital Integration