

数据驱动视角下高校图书馆精准服务模型的构建与理论验证

庄育栋

广东艺术职业学院 广东 广州 510075

摘要: 在大数据与高校教育数字化深度融合背景下,传统高校图书馆以资源为中心的被动服务模式难以满足师生个性化、动态化的信息需求。本文基于数据驱动理论,融合用户画像与知识服务理论,构建了一个包含数据采集、处理、需求挖掘、服务输出及反馈优化的五层闭环精准服务模型。该模型通过打破数据孤岛、挖掘潜在需求、优化服务流程,实现用户需求精准识别与资源智能匹配。研究从科学性、可行性与适用性三方面对该模型进行理论验证,结果表明其能有效支撑高校图书馆由“资源导向”向“用户导向”转型,为数字化时代图书馆精准服务提供理论框架与实践路径。

关键词: 数据驱动; 高校图书馆; 精准服务; 模型构建; 理论验证; 用户画像

引言

高校图书馆作为高等教育的资源中心与文化育人阵地,正面临用户需求个性化、高阶化与服务数据孤岛化的双重挑战。师生信息行为日益碎片化、移动化,需求从文献借阅延伸至科研支持、知识发现等深度服务,而分散的业务系统阻碍了对用户需求的精准把握。与此同时,大数据、人工智能等技术为图书馆服务转型提供了新范式。数据驱动理念推动服务从经验导向转向实证决策,从被动响应迈向主动预测。尽管学界已在推荐算法、实施路径等方面有所探索,但普遍缺乏系统性理论模型及其科学性、可行性与普适性的严谨验证。为此,本文以数据驱动为核心,融合用户画像与知识服务理论,构建一个结构清晰、闭环运行的高校图书馆精准服务模型,并开展多维度理论验证,旨在为图书馆数字化转型提供兼具前瞻性与可操作性的理论支撑,助力实现“信息找人”的智慧服务新范式。

1 数据驱动下高校图书馆精准服务的理论基础

1.1 数据驱动理论

数据驱动理论是以数据为核心的方法论,强调通过采集、分析等多环节,揭示数据背后的规律等,完成从“数据”到“智慧”的价值跃迁。在高校图书馆精准服务中,它要求将各类数据视为宝贵资产,利用先进技术深度加工数据,提炼关键洞察,并嵌入服务设计等环节,形成闭环,是精准服务模型的“操作系统”和“动力源泉”。

1.2 用户画像理论

用户画像理论是实现精准服务的关键技术支撑。它通过收集分析用户多维度数据,构建能代表真实用户的

数字化虚拟模型,动态精准刻画用户特征与需求。高校用户群体分层明显,基于该理论,图书馆可针对不同类型用户构建差异化画像标签体系,持续更新数据修正模型,精准识别显性和隐性需求,是实现个性化服务的前提。

1.3 知识服务理论

知识服务理论标志着图书馆服务理念的跃升,超越传统模式,以用户知识需求为出发点,提供高附加值知识解决方案^[1]。数据驱动下的精准服务是其在数字时代的实践形态,传统知识服务依赖学科馆员经验判断,而数据驱动的知识服务基于客观全面的数据分析,能主动推送情境化知识服务,提升图书馆在学术价值链中的地位和作用。

2 数据驱动对高校图书馆精准服务的赋能机制

2.1 打破数据孤岛,整合多元数据资源

高校图书馆长期存在“烟囱式”信息系统,如ILAS、电子资源门户、座位预约、门禁、OPAC等系统各自独立,数据标准不一,形成数据孤岛。数据驱动通过构建统一数据中台或仓库,汇聚、清洗并关联结构化(借阅记录)、半结构化(日志)与非结构化数据(咨询文本),以学号为唯一标识串联用户全行为轨迹,并关联馆藏元数据与使用数据,形成动态“数据资产地图”,为精准服务奠定基础。

2.2 挖掘潜在需求,实现需求精准识别

借助机器学习等大数据技术,图书馆可挖掘用户隐性需求:通过关联规则发现资源借阅规律(如《高等数学》后常借《线性代数》),实现智能推荐;利用序列模式识别用户行为路径(如毕业论文前的检索序列),提前

推送相关服务；通过聚类分析划分用户群体（如“科研型”“休闲型”），开展群体化精准服务，推动从“用户找资源”向“资源找用户”转变。

2.3 优化服务流程，提升服务响应效率

数据驱动支持精细化运营：基于门禁与预约数据预测客流高峰，动态调整开放时间与人员配置；依据电子资源使用统计（如COUNTER报告）优化采购决策；结合用户满意度反馈快速改进线上服务流程。通过数据支撑的资源配置与流程优化，显著提升服务效率与用户体验。

3 数据驱动下高校图书馆精准服务模型的构建

3.1 模型构建原则

一是用户中心原则：将满足高校师生的个性化、差异化需求作为模型设计的唯一出发点和最终落脚点。所有数据的采集、处理、挖掘与服务输出，都必须紧密围绕用户需求展开，确保服务供给与用户期望的高度契合。二是数据驱动原则：将数据视为模型运行的血液和燃料。模型的每一个层级都深度嵌入数据处理与分析环节，确保决策和服务的每一步都有坚实的数据支撑，而非主观臆断。三是动态优化原则：模型并非静态的蓝图，而是一个具备自我学习和进化能力的生命体。通过持续收集用户反馈，模型能够不断修正自身的算法、策略和参数，以适应用户需求和外部环境的持续变化。

3.2 模型的层级结构

本模型由五大核心层级构成，它们遵循“数据流转—价值挖掘—服务落地—迭代优化”的逻辑链条，形成一个紧密衔接、逐层递进的闭环体系。

3.2.1 数据采集层

数据采集层作为模型的基石，全面、实时、准确地汇集图书馆服务相关数据，涵盖用户、资源与环境三大维度。用户数据包括静态属性（如学号/工号、姓名、院系、专业、职称、年级）和动态行为（如借阅历史、检索关键词、数据库访问、电子书阅读进度、座位预约、活动参与及在线咨询记录）；资源数据涵盖各类馆藏（纸质书刊、电子资源、学位论文、音视频、特藏等）的元数据（题名、作者、ISBN/ISSN、主题词、摘要）、物理状态（在馆、借出、预约）及使用统计（借阅次数、下载量、浏览时长）；环境数据则反映图书馆物理与数字空间的运行状况，如区域实时人流、Wi-Fi连接数、设备使用情况、开放时间表及服务公告等^[2]。采集方式多元化，包括API接口对接、日志抓取、传感器（如RFID、摄像头）采集，以及在合法合规前提下用户主动授权填报。

3.2.2 数据处理层

数据处理层的核心任务是将原始杂乱、含噪声、缺

失或不一致的数据“原材料”加工为高质量的“数据产品”。其主要工作包括：首先进行数据清洗，剔除重复、错误或无效记录，并对关键字段的缺失值进行合理填补；其次通过数据整合，运用实体识别与链接技术，将来自不同来源但指向同一实体（如同一用户或同一本图书）的数据关联合并，构建统一的用户视图与资源视图；随后开展数据转换，将结构化、半结构化及非结构化等异构数据转化为适合分析的统一格式，例如将文本日志解析为结构化事件流，或将非结构化文本进行向量化处理；最后依托分布式架构（如Hadoop、Spark）或云平台，构建数据仓库或数据湖，实现对海量数据的安全、高效、可扩展存储与管理，为上层分析提供坚实支撑。

3.2.3 需求挖掘层

需求挖掘层是整个模型的“大脑”与核心，致力于从高质量数据中深度洞察用户需求，包含两大关键模块：一是用户画像构建模块，利用分类、聚类机器学习算法对用户进行多维度标签化，形成涵盖基础属性（如院系、年级）、行为偏好（如偏好数字资源、夜间使用图书馆）、兴趣领域（如机器学习、明清史）及价值等级（如高活跃度、高影响力）的动态标签体系，画像可随用户行为持续演化；二是需求关联分析模块，综合运用关联规则、协同过滤及深度学习技术（如RNN、Transformer），深入挖掘用户行为与资源间的复杂关系，例如通过分析检索序列与点击流预测用户下一步兴趣，或基于相似用户的资源使用模式实现跨域精准推荐，从而实现从“人找资源”到“资源找人”的智能服务跃迁^[3]。

3.2.4 服务输出层

服务输出层是模型价值的最终体现，致力于将需求挖掘结果转化为多样化、场景化的具体服务行动。在OPAC、电子资源门户及移动APP等触点，系统可基于用户画像与实时行为，推送高度相关的图书、文章、数据库或多媒体资源，实现个性化推荐；面向科研团队，提供定题情报服务（SDI）、文献计量分析与学科态势报告，为学生则嵌入论文开题辅导、查重指导和学术规范讲座等定制化知识服务；同时，依托实时人流数据，智能化引导用户前往空闲区域或座位，并根据借阅与预约习惯主动提醒图书归还或预约到期；此外，通过搭建跨学科的资源共享与学术协作平台，促进知识流动与科研协同，全面支撑教学、科研与学习的多元需求。

3.2.5 反馈优化层

服务输出层是模型价值的最终体现，致力于将需求挖掘结果转化为多样化、场景化的具体服务行动。系统在OPAC、电子资源门户及移动APP等触点，基于用户画

像与实时行为,精准推送图书、文章、数据库或多媒体资源,实现个性化推荐;面向科研团队,提供定题情报服务(SDI)、文献计量分析和学科态势报告,为学生则嵌入论文开题辅导、查重指导及学术规范讲座等定制化知识服务;同时,依托实时人流数据,智能引导用户前往空闲区域或座位,并根据借阅与预约习惯主动提醒归还或预约到期^[4]。此外,通过搭建跨学科资源共享与学术协作平台,促进知识流动与科研协同,全面支撑教学、科研与学习的多元需求。

4 数据驱动下高校图书馆精准服务模型的理论验证

4.1 模型的科学性验证

本模型具有坚实的理论基础,融合数据驱动、用户画像与知识服务三大成熟理论,其五层结构遵循“数据输入—处理—分析—输出—反馈”的信息处理逻辑,契合图书馆业务流程。同时,模型内部形成完整的PDCA闭环:数据采集为前提,处理为保障,需求挖掘为核心,服务输出为目的,反馈优化为升华,各层级环环相扣,逻辑自洽,具备自我校准与持续进化能力。

4.2 模型的可行性验证

模型具备良好落地条件:技术上,大数据分析、机器学习、云计算等核心技术已成熟,开源框架(如TensorFlow、SparkMLlib)和商业方案丰富,高校普遍具备相应基础设施与人才储备;数据上,多年信息化积累的TB/PB级历史数据为模型训练提供充足“养料”,辅以有效数据治理即可保障可用性;组织与需求上,图书馆正处于智慧化转型关键期,管理层与馆员对数据价值高度认同,内生动力强劲,为实施营造了良好环境。

4.3 模型的适用性验证

模型具有广泛普适性与动态适应性:在用户层面,通过分层画像可灵活满足本科生、研究生、教师等不同群

体需求;在机构层面,适用于研究型大学、职业院校或专业院校,可根据办学定位调整重点——如强化科研服务或职业技能资源匹配;在演进层面,依托反馈优化机制,能随学校战略、学科发展或技术变革动态调适,持续保持与外部环境同步,展现出强大生命力与推广价值。

5 结语

本文构建并理论验证了数据驱动视角下高校图书馆精准服务五层闭环模型,该模型以数据为纽带、用户为中心、知识服务为价值导向,为破解传统服务模式困境提供理论蓝图,且在理论、技术、数据和组织层面可行,适用于不同类型和层次的高校图书馆。不过,理论模型构建只是第一步,未来研究与实践方向如下:一是深化实证研究,选代表性高校图书馆试点,通过多种方法收集真实数据,检验、修正与优化模型,形成最佳实践;二是推动技术融合创新,将前沿人工智能技术融入模型,如在需求挖掘层引入大语言模型,服务输出层利用生成式AI提供智能咨询;三是保障伦理与隐私,建立健全数据安全管理制度,遵循最小必要原则,确保用户数据权益,让精准服务合法合规发展。

参考文献

- [1]孙玲.大数据驱动图书馆业务应用与服务创新[J].办公室业务,2019(15):156.DOI:10.3969/j.issn.1004-647X.2019.15.125.
- [2]廖素明.基于大数据的高校智慧图书馆建设研究[J].人文之友,2019,(13).DOI:10.3969/j.issn.2096-4684.2019.13.084.
- [3]曹贤龙.基于关键链技术的S高校智慧图书馆建设项目进度管理优化研究[D].2024.
- [4]朱静薇,李红艳.大数据时代下图书馆的挑战及其应对策略[J].现代情报,2013,(5).DOI:10.3969/j.issn.1008-0821.2013.05.003.