

# Innovative Thinking on big Data Enabling University Library

Ruilin Feng

Xianyang Normal University, Xianyang, Shaanxi, 712000, China

## Abstract

With the rapid development of information technology, big data has become an important force to promote social progress and innovation. In the field of education, the application of big data of university library high quality innovation also has revolutionary significance, it not only changed the library information storage, retrieval and management, and through the analysis of massive user behavior, resource use and external environment data, more accurately grasp the user demand, optimize the allocation of resources, improve the quality of library service. This paper puts forward some thoughts on how big data can enable high-quality innovation in university libraries.

## Keywords

big data technology; university library; innovation; enabling; thinking

## 关于大数据赋能高校图书馆创新思考

冯睿琳

咸阳师范学院, 中国·陕西 咸阳 712000

## 摘要

随着信息技术的飞速发展,大数据已经成为推动社会进步和创新的重要力量。在教育领域,大数据的应用对高校图书馆高质量创新同样具有革命性的意义,它不仅改变了图书馆信息的存储、检索和管理方式,而且能够通过分析海量的用户行为、资源使用和外部环境等方面数据,更精准地把握用户需求,优化资源配置,提升图书馆服务质量。本文就大数据如何赋能高校图书馆高质量创新提出几点思考。

## 关键词

大数据技术; 高校图书馆; 创新; 赋能; 思考

## 1 引言

大数据技术的快速发展为高校图书馆的智能化、精细化和个性化服务提供了新的机遇。传统高校图书馆主要依赖人工管理与传统信息系统进行馆藏管理、读者服务及资源调配,存在数据利用率低、信息孤岛现象严重、资源匹配效率不足等问题。随着信息技术的进步,大数据技术凭借其强大的数据存储、处理、分析能力,在学术资源整合、用户行为分析、精准服务推荐等方面展现出显著优势,使高校图书馆能够突破传统模式限制,实现深度智能化升级。

## 2 大数据技术概述

大数据技术是指利用现代信息技术手段对大规模、多

**【课题项目】**咸阳师范学院、陕西省教育学会2021年教育教学改革研究项目:高校图书馆在信息化时代的创新路径研究(项目编号:2021Z012)。

**【作者简介】**冯睿琳(1968-),女,中国陕西咸阳人,本科,副研究馆员,从事图书馆学研究。

源异构、高速增长的数据进行采集、存储、管理、分析和应用的一整套技术体系<sup>[1]</sup>。其核心特征通常归纳为“4V”:即数据体量(Volume)巨大,数据类型(Variety)多样,数据生成速度(Velocity)极快,以及数据价值(Value)密度较低但潜力巨大。此外,大数据技术还涵盖机器学习、人工智能、云计算、数据挖掘等多个相关领域,并在数据驱动的智能决策过程中发挥关键作用。

## 3 大数据赋能高校图书馆创新

### 3.1 基于大数据的个性化资源推荐

基于大数据的个性化资源推荐系统在高校图书馆中的应用依赖于多维度数据采集、复杂算法构建及智能优化策略。首先,数据采集涵盖用户的馆藏借阅记录、电子资源访问行为、全文下载日志、检索关键词、学术数据库使用情况及外部文献管理工具的数据交互信息,采用ETL流程进行数据清洗、缺失值填补及标准化处理,以确保数据质量及一致性。其次,用户兴趣画像构建采用高维特征表示,包括基于TF-IDF计算文献主题权重、LDA主题建模解析研究兴趣分布,并结合深度学习中的序列建模技术学习用户阅读偏好

变化规律,形成精准用户画像。引入基于内容推荐的文本相似度计算(如余弦相似度、Word2Vec 语义嵌入)优化资源匹配精度,结合强化学习策略动态优化推荐结果,实现个性化书籍、期刊及数据库资源推送。在推荐策略优化方面,基于图卷积网络构建学科知识图谱,结合热点追踪算法(如时间序列预测、贝叶斯动态更新)分析不同学术领域的研究趋势,实现学科热点驱动的资源推荐。同时,针对用户冷启动问题,采用混合推荐策略,将基于规则的初始偏好设定与自适应调整算法相结合,使新用户亦能获得精准推荐<sup>[2]</sup>。此外,系统应支持交互式反馈机制,通过强化学习动态调整推荐权重,以提升长期推荐效果。最终,结合多维度评估指标,如精准率、召回率、归一化折损累积增益及用户满意度调查,优化推荐系统性能,确保其长期适用性。

### 3.2 智能文献管理系统构建

利用大数据技术打造智能文献管理系统,能够对高校图书馆资源管理流程进行有效优化,并增强文献分类、调配、更新等方面的智能化程度。首先,建立基于大数据分析的文献资源挖掘模型,通过主题模型对馆藏文献进行自动化分类,即利用高维数据聚类算法对文本主题进行提取,提高分类的准确性和自动化程度。其次,应用深度学习技术,结合动态资源调配算法,对不同学科领域的文献借阅趋势进行预测,利用循环神经网络或时间长短记忆网络,对馆藏结构进行调整,以保证高需求文献的流通性。此外高校图书馆搭建智能索引系统,并采用分布式搜索引擎,优化关键词匹配和语义检索。同时,引入文本自动分析技术以及实体识别和语义相似度计算对图书馆文献进行自动归档,并利用 BERT 等预训练语言模型对未分类文献进行语义信息解析,以增强分类和存储管理的效率。最后,利用时间序列分析方法对文献使用周期进行预测,并结合强化学习算法对淘汰策略进行优化,进一步优化馆藏更新机制,从而确保图书馆馆藏文献资源科学合理。此外,以区块链技术建立去中心化图书馆文献管理机制,并结合智能合约方法开展文献资源访问控制以及版本管理。

### 3.3 用户行为分析与服务优化

通过大数据技术的深入应用,使高校图书馆用户行为分析的精细化程度得到显著提高,这有助于实现对服务策略的优化以及文献资源利用效率提高。首先,通过日志数据采集技术,结合 RFID 技术、视频监视系统以及 Wi-Fi 探针,实现对包括师生入馆时间、停留时长、座位占用以及借阅行为等用户图书馆内的行为数据获取、实时跟踪和动态记录。其次,高校图书馆利用清洗、归一化和特征工程等数据预处理方法提升用户数据质量,借助于机器学习方法对读者群体进行细分,识别不同类别用户的特征。例如,可通过 K-means 聚类算法对借阅偏好进行分类,区分高频借阅用户、特定学科依赖用户与偶然借阅用户,从而精准调整资源投放策略。此外,基于 Apriori 算法或 FP-Growth 算法挖掘用户借阅行

为的关联规则,识别常见联合借阅模式,以优化馆藏资源配置,提升文献复用率<sup>[3]</sup>。进一步,通过热力图分析技术,结合时空数据,解析高峰时段的座位使用情况,并采用 LOF (局部异常因子)算法识别资源占用异常行为,优化座位管理策略,提升空间利用率。其三,运用自然语言处理对读者反馈意见进行情感分析,结合 LDA (潜在狄利克雷分配)主题模型挖掘用户需求偏向,以优化智能推荐系统,提高个性化服务水平。此外,结合深度学习模型,预测用户未来借阅需求,提前调整文献采购计划,确保资源精准匹配。

### 3.4 学术资源精准对接

高校图书馆学术资源精准对接目的在于促进学术资源匹配效率地提高和资源供需关系优化,而大数据技术应用则有助于它的实现。首先,高校图书馆可以挖掘不同学科领域的文献引用关系、学科发展趋势及前沿研究热点,这可借助大数据技术对学术数据库、机构知识库和科研项目数据进行深度分析实现。通过对基于大数据分析的自动化推荐系统的构建,相关学术资源可根据师生学习与研究需求进行精准推荐,并将相关文献与其研究方向联系起来,从而提升文献资源获取与利用效率。其次,借助资料挖掘技术,可与教师及学生研究方向、学术履历、论文发表记录及引用次数等信息相结合,高效匹配潜在学术资源。通过对历史文献的引用行为的统计分析获取到师生与其相关的学术资源匹配规律,这有利于促进资源推荐体系的准确性及响应速度的提高<sup>[4]</sup>。另外,高校图书馆还可以利用自然语言处理技术,对文献中的关键词、主题、领域等进行自动标注,从而提高文献检索的精确性。简单而言,文献的自动化分析和标注是通过自然语言处理中的文本分类和实体识别技术,从而帮助师生对相关文献进行更精确的筛选。同时,通过机器学习算法不断优化学术资源推荐系统,提高信息检索的个性化水平,可以根据师生的行为偏好进行自适应推荐。最后,其三,高校图书馆结合大数据分析 with 人工智能技术,可实现学术资源智能推送和精准匹配,这将有助于促进师生实现跨学科学术协作与创新研究。在具体实施过程中,高校图书馆利用云计算平台整合馆内大量的学术资源,随后通过智能算法实现资源智能匹配,接着以大数据可视化技术对资源分布和需求予以动态展示,如此一来将有助于学术资源的精准对接与高效配置。

### 3.5 科研辅助服务系统优化

高校图书馆科研辅助服务系统优化中的大数据技术应用作用显著,其能够促使科研资源整合和服务效率得到提高。首先,高校图书馆利用大数据技术构建科研知识图谱,从而将多维度的学术论文、科研项目和专利等数据融为一体。教师和学生通过可视化的知识图谱可以很快地定位相关领域的研究内容,随后再与自身学习与研究方向、学科专业等相互联系展开分析查看是否存在潜在的研究空白和热点。借助于科研知识图谱,在提高文献资源检索精确度的同时,也使学科之间的交互性、协同空间得到有效提升。其次,大

数据在高校图书馆科研辅导系统优化的另一个重点手段在于文本挖掘技术。该技术通过对图书馆内馆藏以及网络数据库等大量研究文献、报告、专利的分析,在研究领域识别出关键名词,发现热点话题,以及学科交叉的潮流。利用这一技术,高校图书馆对馆藏文献资源进行相应调整,为师生提供实时学术动态更新以及帮助他们跟踪领域发展的最新成果。该技术还可帮助高校图书馆对文献中的引用进行分析,从而给师生研究推荐相关高质量、前沿的文献资源。此外,高校图书馆借助学术社交网络数据的分析可以进一步拓宽自身科研服务边界。通过对学术社交平台上的科研人员关系和合作网络进行梳理,挖掘出潜在的合作伙伴关系,给本校教师或者学生提供精准的学术合作建议。同时高校图书馆通过分析社交网络并结合师生学习与研究需要或者兴趣领域,对适宜的合作对象进行自动推荐,从而促进跨学科学术合作与交流<sup>[9]</sup>。另外,高校图书馆应用大数据技术也可以提供个性化的研究资源推荐服务。该技术应用后,高校图书馆能够基于分析用户的研究历史和偏好情况下自动推送相关领域的文献、科研工具及课题进展。从而帮助教师与学生节省查找时间以及获取相关资源准确性大大增加。最后,基于大数据的科研评价体系可以为高校图书馆提供精准的科研辅助服务评价支持。高校图书馆通过量化分析本校教师与学生所做科研项目、课题及发表论文的影响力、引用频次、传播范围等多维度数据,从而提供给他们全面且客观的科研评价结果。该评价体系不但可以更进一步调动出师生学习与研究的积极性,同时高校图书馆还能结合评价结果给予文献资源、科研方向、合作建议等方面的支持,从而促使他们科研活动质量与效率地提升。

### 3.6 大数据技术运用团队建设和人才培养

大数据技术的运用,对高校图书馆团队建设和人才培

养提出了更高的要求。高校图书馆需要加强培养和引进具有大数据分析能力的专业人才,一方面通过加强对内部工作人员的大数据技术运用知识、方法以及技巧等方面培训,全面提高图书馆人员的数据意识和数据分析运用能力;另一方面,通过引进专业人才、建立与相关单位及第三方专业机构合作机制等形式,组建跨学科跨专业的大数据服务团队,确保大数据赋能图书馆作用的充分发挥。

## 4 结语

综上所述,基于大数据赋能下高校图书馆朝着智能化、精准化和高效化方向创新,并在精准把握师生需求情况下有效地增强图书馆服务质量。因而,高校图书馆须充分利用数据挖掘、数据预处理等大数据技术手段对自身资源推荐、文献管理、用户行为分析、学术资源匹配等方面工作予以创新升级。另外,高校图书馆为给师生学习与学术研究提供更为专业且优质的文献资源服务,未来须进一步加强大数据基础设施建设以及深化数据处理,从而促使大数据技术在图书管理应用上发挥出更大价值,推动高校图书馆在数字化时代的转型升级。

## 参考文献

- [1] 黄强.大数据背景下高校图书馆知识服务研究[J].武汉船舶职业技术学院学报, 2023, 22(4):21-24.
- [2] 虞明.大数据背景下高校图书馆精细化管理策略研究[J].服务科学和管理, 2025, 14(1):5.
- [3] 李冰.大数据时代高校图书馆嵌入式服务创新模式研究[J].中文科技期刊数据库(全文版)图书情报, 2023(2):4.
- [4] 黄梅珍.大数据背景下高校图书馆服务创新策略[J].百色学院学报, 2024, 37(1):122-127.
- [5] 张萍 蒋鑫 朱光.大数据情境下高校图书馆服务创新优化对策[J].新世纪图书馆, 2024(10).