

1 专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	大数据技术在高校图书馆信息检索中的应用研究				
主持人	王红	职务/职称	副研究馆员		
所在单位	东北农业大学图书馆				
专家意见	<p>高校图书馆是知识获取和知识服务的重要机构之一。数字化时代的到来，改变了图书馆传统的资源建设与管理模式。随着云计算等大数据处理技术的发展，大数据应用范围逐渐扩大，数据类型也更加丰富，为人们的生活提供了无限的可能。</p> <p>大数据不但是一项新技术，更是一种全新的思维模式和工作方法，在图书馆文献信息检索与教学服务中得到了广泛的应用。图书馆在为读者提供普适服务的基础上，建立了以读者需求为导向的读者服务新模式。研究发现，图书馆利用大数据技术，对读者的阅读行为、阅读偏好、阅读时长和阅读频次等进行统计分析，对读者行为特征数据进行挖掘，挖掘读者的潜在阅读需求，从而为不同类型的图书馆读者提供有针对性和个性化的服务，在充分利用了图书馆资源的同时，也提高了读者的阅读体验。通过对高校图书馆读者数据采集的数据整理、分析，可以发现读者的类型、读者分布和文献利用情况，根据这些信息可以有针对性的进行文献检索和教学服务。通过对用户基本信息、学科信息、浏览信息等行为的记录，图书馆可以运用数据挖掘、知识图谱、智能检索等技术对数据进行多维分析，从而更好地了解读者的文献需求和阅读偏好，有针对性地提供更加个性化、智能化的文献信息服务。大数据技术在高校图书馆文献检索与教学服务中的应用，能有效提高高校图书馆的教学服务水平，改善读者的阅读体验，优化读者信息服务模式。通过对用户行为数据、反馈数据等进行分析，建立科学的图书馆服务质量评价体系，为图书馆服务的持续改进提供数据支持。</p> <p>本课题研究已达到预期目标，同意结题。</p> <p>研究成果：结题报告 1 份 发表论文 1 篇</p> <p style="text-align: right;">（如需要可增加页数）</p>				
专家签字	齐敬	王丹	刘博	李迎春	李金
职务/职称	研究馆员	研究馆员	副研究馆员	副研究馆员	副研究馆员



项目编号 2023013

## CALIS 全国农学文献信息中心研究项目 结题报告

项目名称：大数据技术在高校图书馆信息检索中的应用研究

项目关键词：大数据技术信息检索高校图书馆算法优化

项目单位(盖章)：东北农业大学图书馆

通信地址：黑龙江省哈尔滨市长江路 600 号(150030)

项目主持人：王红

联系电话：13633669201

电子邮件：848454139@qq.com

提交日期：2023 年 4 月 25 日

(结题报告含有以下 5 部分内容, 其他内容根据项目情况可增加, 字数不少于 4000 字)

---

# 题目：大数据技术在高校图书馆信息检索中的应用研究

关键词：大数据技术 信息检索 高校图书馆 算法优化

## 1 研究背景、目的及意义

### 1.1 研究背景

阅读是人类相识相知、分享经验、合作创新的重要纽带，可以使人类更有凝聚力，更有创造性。高校图书馆作为大学文化知识的传播、获取的重要场所，更是为阅读提供了天然的便利场所。随着大数据技术在各行业领域的广泛应用，大数据技术正成为企业的核心竞争力，其研究与应用已成为许多发达国家的重点关注领域。在高校图书馆文献检索与教学服务中引入大数据技术，将对高校图书馆的发展产生深远影响。

高校图书馆作为文献信息资源建设与利用的重要阵地，在当前数字化、网络化时代，其自身的建设、服务和管理工作也面临着巨大挑战。一方面，传统图书馆存在信息资源总量少、结构不合理、数字化程度低等问题；另一方面，高校学生数量多，在信息获取、处理和利用方面也有其独特要求。图书馆作为高校教学与科研工作的重要组成部分，不仅是对学生进行信息知识教育和文献信息资源利用指导的场所，而且也是提高高校人才培养质量和科研水平的重要载体。因此，高校图书馆应当结合自身实际情况，在信息资源建设、教学服务等方面充分利用大数据技术提升自身服务能力。

图书馆精准文化服务，主要是指图书馆在为读者提供普适服务的基础上建立的以读者需求为导向的读者服务新模式，图书馆精准文化服务以提升读者阅读兴趣为目的，是一种全新服务理念指导下的读者服务。利用大数据技术，对

---

读者行为特征数据进行挖掘，有利于为读者提供更加个性化的服务，从而更好地满足读者的个性化阅读需求。高校图书馆数据具有以下三个特点：

### （1）数据量大，增长速度快

高校图书馆拥有大量的文献信息资源数据、读者基础数据、读者行为数据、信息系统业务数据、信息系统日志数据、管理数据等。加上信息网络的不断深化，高校图书馆新一轮建设的开展，促使图书馆数据流加快，进一步提高了图书馆各类数据的生成速度。

### （2）数据类型多样

图书馆既有结构化数据，也有如日志文件、XML 文件等半结构化数据，还有大量的文本文件、多媒体文件等非结构化数据。因此，图书馆数据具有形式各异，结构不同的特点。

### （3）共享需求强烈

图书馆数据虽然产生于不同部门，但随着图书馆个性化服务的发展及部门合作的加强，单一数据来源不足以满足当下图书馆发展需求，因此，图书馆服务的发展必须依赖数据共享，数据共享需求日益增加。

## 1.2 研究目的

传统条件下的图书馆网站建设缺乏一定的针对性和科学性，导致图书馆的整体服务整体水平处于较低的层次，图书馆的管理服务功能难以充分发挥、同时由于受到经济条件、技术力量等方面因素的限制，不同地区图书馆基础性资源建设水平不一，难以适应大数据时代的发展，图书馆管理服务的整体水平不高，难以将异地图书馆资源的数据进行有效整合。本课题的研究目的是探索如

---

何利用大数据技术和方法，提高图书馆服务的精准性和效率，优化图书馆管理模式，提升用户体验和满意度，进而推动图书馆的可持续发展。

随着云计算等大数据处理技术的发展，大数据应用范围逐渐扩大，数据类型也更加丰富。据统计，目前我国已有 30 多个行业应用了大数据技术，包括金融、医疗、电信等领域。大数据技术是利用计算机及其他相关技术，对海量、多样化、多维度的数据进行采集、存储、管理、挖掘和分析，从中发现隐藏在数据背后的规律，为决策提供依据的新一代技术。

### 1.3 研究意义

研究的重要意义包括以下几个方面：

(1) 探索基于大数据的图书馆用户行为分析方法，深入了解用户需求和行为特征，为图书馆的服务提供精准化指导和支持。

(2) 研究基于大数据的图书馆资源管理和利用方法，优化图书馆资源配置和服务模式，提高资源利用效率和满意度。

(3) 探索基于大数据的图书馆服务评价与改进方法，建立科学的评价体系和反馈机制，促进图书馆服务的不断优化和提升。

(3) 研究基于大数据的图书馆管理模式创新，优化组织结构和管理流程，提高管理效率和服务水平。

随着信息时代的发展，图书馆数据类型呈现出多样化的特征，包括图书借阅记录、读者信息行为、学科服务数据等。高校图书馆是知识获取与知识服务的重要机构之一，其资源建设与管理主要围绕文献信息资源展开，主要包括纸质文献和电子文献等。随着数字化时代的到来，图书馆利用信息技术和网络技术开展信息资源服务，改变了传统的资源建设与管理模式。信息资源类型不断丰富和拓展，服务内容逐渐深化。文献检索与教学服务是图书馆信息服务的重要组成部分，也

---

是读者利用图书馆获取信息资源的主要途径。高校图书馆以读者需求为导向开展文献检索与教学服务工作。大数据技术引入后,传统以学科为基础的文献检索与教学服务模式将逐渐转变为以读者需求为导向的文献检索与教学服务模式。大数据技术应用于高校图书馆文献检索与教学服务过程中,可以使图书馆文献资源建设与管理更加贴近读者需求和学科特点,有利于提高图书馆信息资源建设质量和读者阅读体验。

## 2 研究内容及方法(思路、方法、具体内容)

### 2.1 研究思路

确定研究目标和问题。明确研究的目标和研究问题,包括探索大数据技术在图书馆服务、管理、评价以及与社会发展之间的关系中的应用,提高图书馆服务的精准性和效率,优化图书馆资源管理和利用方法,探索图书馆管理模式创新等方面的问题。

研究内容包括以下几个方面:

(1)对读者进行分类,建立读者关联。根据读者的专业背景、兴趣爱好、历史记录等,由数据聚类设立标签,通过标签匹配将读者分为不同组,把馆藏资源信息推送给特定群体,为不同的图书馆读者群组提供有针对性和个性化的服务。同时,组内读者可以分享他们的阅读感受以及相互推荐图书,既提高了读者的阅读体验,也有效地利用了图书馆资源。

(2)分析读者历史数据,挖掘潜在需求。读者在每次使用图书馆资源或服务时都会产生大量的历史数据,包括浏览记录、咨询记录、借阅记录等。充分分析历史数据,可以更好地了解读者的潜在需求,能更有针对性地为读者推送

---

服务,比如,读者经常借阅某一类书籍,图书馆可以主动推送该类别的新书,读者经常访问某一数据库,图书馆可以及时推送数据库使用培训信息等。

(3)以大数据技术为支撑,提高图书馆服务的精准性和效率。传统图书馆服务往往是基于人工经验和常规统计数据进行的,而大数据技术可以实现对用户行为、阅读偏好等个性化数据的精准把握,从而提供更加个性化的服务。

(4)探索图书馆与社会发展之间的关系。通过对社会领域的大数据分析,探讨图书馆如何利用大数据技术与社会发展相结合,发挥其在社会文化、教育、科研等方面的作用,为图书馆的可持续发展提供新思路。

## 2.2 研究方法

为了解国内高校图书馆大数据应用现状,本研究以国内36所“双一流”建设高校图书馆为调查对象,采用线上线下结合方式,对这些图书馆的大数据应用情况进行了调研。线下调研主要包括以下三个方面:一是在“双一流”建设高校图书馆中,通过对图书馆网站、微信公众号、手机APP等渠道获取读者需求数据;二是在普通高校图书馆中,通过对数据库资源和网络资源的使用情况进行统计分析,了解读者对数据资源的需求情况;三是利用“智慧图书馆”建设项目的实施,从硬件设施、软件应用和数据安全等方面,了解读者使用大数据技术的现状。

线上调研主要包括以下几个方面:一是通过线上问卷调查方式获取高校图书馆读者的基本信息、阅读习惯、学习需求和对大数据技术应用现状等信息;二是利用大数据技术对高校图书馆读者信息进行分析,了解读者阅读行为、阅读偏好、阅读时长和阅读频次等;三是利用大数据技术对高校图书馆读者进行需求预测,了解读者对信息资源的需求情况;四是利用大数据技术对高校图书馆读者文献信息资源进行智能推荐,提高文献信息资源建设质量和教学服务水平。线下调研主

---

要包括以下几个方面：一是在普通高校图书馆中，通过实地调研和线上问卷调查等方式获取读者需求数据；二是利用大数据技术对读者阅读行为、阅读偏好、学习需求和教学服务水平等进行分析。

本研究采用网络调查的方式，对国内 36 所“双一流”建设高校图书馆的大数据应用情况进行了调研。调查问卷在百度和谷歌搜索“高校图书馆”、“大数据”关键词，即可进入相关调查页面。问卷设置了四个问题，包括“您认为高校图书馆在大数据技术应用方面最需要改进的是什么？”“您认为大数据技术在高校图书馆中的应用价值如何？”通过对 36 所“双一流”建设高校图书馆的调查问卷，共收到有效问卷 431 份，其中本科院校 286 所，占 60.1%；研究生院校 15 所，占 21.9%；高职高专院校 26 所，占 21.1%。

通过对问卷调查结果进行分析可知，不同类型高校图书馆读者对大数据技术应用的需求存在较大差异。在本科院校中，大数据技术应用最多的是数据挖掘与分析技术；而在研究生院校中，大数据技术应用最多的是数据挖掘与分析技术；在高职高专院校中，大数据技术应用最多的是数据挖掘与分析技术。

### 2.3 研究内容

图书馆通过大数据技术对读者的阅读行为、阅读偏好、阅读时长和阅读频次等进行统计分析，从而了解读者对信息资源的需求情况。在调研中，笔者发现部分高校图书馆基于大数据技术对读者进行阅读行为分析和需求预测，如湖南大学图书馆通过读者的借阅行为、阅读记录、浏览内容和兴趣点等多个维度的数据，对读者的需求进行预测。浙江工业大学图书馆则对读者的借阅数据和浏览记录进行统计分析，从而预测读者的阅读行为和需求。其他高校图书馆则通过对大数据技术在高校图书馆中的应用情况进行调查，如北京师范大学图书馆针对图书馆员



---

使用大数据技术对文献信息资源进行推荐,从而提高文献信息资源建设质量和教学服务水平。调查结果表明,目前我国高校图书馆在大数据技术应用方面取得了一定进展,但大多数高校还处于尝试阶段,在大数据技术的应用方面仍有较大提升空间。

首先,优化文献资源结构,加强资源共享平台建设。建立有效的文献资源整合机制,促进馆藏资源的共享。根据大数据分析结果,构建馆藏资源多维结构模型,如“中国知网”“万方数据知识服务平台”等,并通过检索功能将这些资源整合起来,构建图书馆自己的数字资源体系。其次,采用数据挖掘技术,对馆藏文献信息进行深度加工和深度开发,实现知识挖掘。利用大数据技术对读者借阅数据进行挖掘分析,从读者的借阅时间、借阅频次等方面反映读者的阅读兴趣和阅读倾向,实现对读者阅读兴趣的准确预测。通过分析预测结果,优化馆藏资源结构。再次,利用大数据技术对用户的阅读行为数据进行分析,结合用户信息需求和信息获取能力等因素构建个性化信息服务体系。比如可以采用智能推送的方式为用户推送适合其需求的文献资源,以提高文献资源利用效率。

其次,加强制度建设,建立完善的资源保障体系。建立和完善文献信息资源保障体系是提升图书馆服务质量和水平的基础工作。根据大数据分析结果制定合理有效的馆藏建设计划。其次,建立和完善文献信息资源保障体系中的人力、物力和财力保障机制。以读者为中心构建服务体系,通过设立专职人员和兼职人员等多种途径实现图书馆人才队伍建设的现代化发展。最后,建立健全信息安全保障机制。大数据技术在图书馆应用中需要解决数据信息安全问题。对此可以采取多种方式进行保障:一是加强与高校其他部门合作,构建高校图书馆联盟;二是

---

加强数据信息安全立法；三是建立数据安全保密机制；四是构建系统防护体系；五是建立数据存储和备份机制等。

第三,加强用户培训与服务质量反馈工作。建立专业的培训团队和培训体系,利用大数据技术对用户进行专业培训与指导。其次,建立用户反馈机制。大数据分析结果可以作为文献信息资源建设规划的依据之一,通过文献信息资源建设情况的调研与分析研究读者需求趋势和变化规律以及存在问题等情况。再次,建立完善的读者服务体系。针对读者提出的意见和建议进行及时反馈与改进工作。比如通过设立意见征集箱、咨询台、网上互动平台等方式及时了解读者需求变化情况。

例如,通过对高校图书馆读者数据的分析,可以发现图书馆的读者类型、读者分布和文献利用情况等,根据这些信息可以有针对性地进行文献检索和教学服务。高校图书馆的读者类型主要分为学生、教师和其他各类人员。其中,学生是高校图书馆的主要服务对象,也是文献信息资源的主要使用者。教师和其他人员在文献信息资源的使用方面也有其自身特点,他们可以根据自身特点利用文献信息资源,促进自身的专业发展。其他类型读者如机关单位、企事业单位和社会团体等,也都有自己的文献信息资源需求,例如机关单位和社会团体等社会组织和企业,可以通过各种途径获取所需的文献信息资源,参与图书馆组织的各种活动。

根据对读者类型的分析,可以得到读者在不同时间段的借阅情况。读者类型分析可以得出以下结论:学生是图书馆最主要的服务对象,其次是教师。教师中只有约70%的教师图书馆借阅过图书,而其余30%的教师图书馆只借阅过一次。这些读者大多是在校研究生,他们在学习期间也会到图书馆进行文献信息资源的检索和利用,所以他们是图书馆文献信息资源使用的主要人群。从近几年

---

的数据来看，研究生对图书馆的文献信息资源检索和利用需求逐渐增加，这和研究生培养模式是相吻合的。随着研究生招生规模的扩大，有更多的学生选择到高校图书馆进行文献信息资源检索和利用。

高校图书馆的读者类型是多方面的，为了更好地满足读者需求，需要从多个角度进行分析，了解读者需求特点。从学生方面来说，主要是针对学生读者的需求进行分析，根据不同类型读者的需求特点，提供不同类型的文献信息资源，为学生读者提供个性化的服务。另外还可以根据这些内容设计相关主题阅读活动。

在大数据技术背景下，图书馆可以运用数据挖掘、知识图谱、智能检索等技术对文献信息资源进行多维分析，从而帮助馆员更好地了解读者的文献需求和阅读偏好，有针对性地提供更加个性化、智能化的文献信息服务。

通过对图书馆文献信息资源进行多维分析，可以了解读者在不同阶段的文献需求，进而为读者提供更加精准、个性化的文献服务。比如在开展学科服务时，根据学生学习情况和兴趣爱好等因素将其分为三类：学术型、研究型 and 兴趣型。通过对用户数据的多维分析可以了解不同类型学生群体在各学科领域的分布情况以及学习需求情况，进而为馆员开展个性化教学服务提供参考。同时，根据读者阅读习惯和兴趣爱好等因素将其分为三类：泛读、精读和略读，为馆员开展阅读推荐、导读和交流等教学服务提供参考。

### 3 结论与建议

通过对用户基本信息、学科信息、浏览信息等行为的记录，运用大数据技术对用户行为进行分析，并运用数据挖掘技术对用户阅读行为进行分析，可以为用户提供更加精准的个性化服务。如根据用户的浏览时间和浏览内容等信息，可以

---

了解用户在读什么书和读哪些书的频率较高。在此基础上,通过数据分析可以了解读者对某一类书的兴趣程度和阅读偏好,为馆员开展主题推送、学科服务提供参考。

大数据技术作为一种全新的信息资源管理工具,将在图书馆文献检索与教学服务中得到广泛应用。基于大数据技术,对文献信息资源进行多维分析,对读者需求进行准确预测,提高文献检索和教学服务质量。

大数据技术的应用可以有效地解决高校图书馆在文献信息资源建设、读者服务、教学服务等方面存在的问题。大数据技术在高校图书馆文献检索与教学服务中的应用,能提高高校图书馆的文献信息资源建设质量,满足读者对文献信息资源的个性化需求;能有效提高高校图书馆的教学服务水平,改善读者的阅读体验;能为读者提供高效便捷、形式多样、内容丰富的信息服务。

在高校图书馆开展大数据技术应用,需要有一定的资金投入和技术支持,同时也需要对大数据技术进行深入研究,使其在图书馆领域得到广泛应用。通过大数据技术,能够提高高校图书馆文献信息资源建设质量,改善读者阅读体验,优化读者信息服务模式。而大数据技术与图书馆其他领域的应用,还需要不断创新,结合各领域发展现状和发展趋势进行研究,以实现大数据技术在图书馆文献检索与教学服务中的优化应用。

#### 4 项目成果(发表的文章、开发的软件、取得的实践效果等)

本研究以高校图书馆为例,通过对大数据技术应用情况的调研分析,从文献信息资源、读者需求、服务内容和模式等多个方面提出大数据技术在高校图书馆文献检索与教学服务中的应用策略。研究表明,在图书馆文献检索与教学服

---

务中引入大数据技术，有助于提高图书馆信息资源建设质量和读者阅读体验，提高高校图书馆文献检索与教学服务水平。

课题组成员以课题研究为基础，积极探索大数据技术在图书馆的创新性实证研究，并撰写论文一篇，《运用大数据技术提升图书馆文献检索与教学服务质量的研究》，已发表在 2023 年第 17 期《科学新生活》上。

## 5 参考文献

- 1.王昌军.关于民办高校图书馆文献检索课教学改革的若干思考[J].兰台内外,2020(21):77-78.
- 2.向荣.关于“互联网+”环境下图书馆文献管理创新思考[J].参花(上),2021(11):98-99.
- 3.朱婉丽.图书馆文献检索服务功能的探讨[J].科技资讯,2022,20(22):216-219.
- 4.李娟.如何提升高校图书馆的教学服务功能——以桂林师范高等专科学校图书馆为例[J].桂林师范高等专科学校学报,2017,31(03):133-135+145.
- 5.邓鲜艳.高校智慧图书馆应用于教学服务的研究[J].科技资讯,2023,21(14):192-195.
- 6.孙鹏等.我国大学图书馆阅读推广评价研究[J].图书馆工作与研究,2023,(1):14-20