

大数据技术在图书流通数据分析中的应用研究

李 洁

云南大学滇池学院图书馆 661000

摘 要: 大数据的发展为图书流通研究带来了新的契机。图书流通是一种传统的服务,其数据信息往往呈现出非结构化和半结构化的特征。在大数据背景下,通过对图书流通数据的挖掘,可以揭示出图书馆服务体系中存在的问题,进而为图书馆改进服务模式、创新服务方式提供重要依据。文章基于图书流通数据特征分析,设计了大数据环境下图书馆服务体系中存在的问题,并提出了相应的解决对策,为图书馆开展服务提供借鉴。

关键词: 大数据技术; 图书; 流通数据分析; 应用研究

The application research of big data technology in book circulation data analysis

Li Jie

Yunnan University Dianchi College Library 661000

Abstract: The development of big data has brought new opportunities for the research of book circulation. Book circulation is a traditional service, and its data and information often presents the characteristics of unstructured and semi-structured. In the background of big data, the mining of the library circulation data can reveal the problems existing in the library service system, and then provide an important basis for the library to improve the service mode and innovate the service mode. Based on the analysis of the characteristics of library circulation data, this paper designs the problems existing in the library service system under the big data environment, and puts forward the corresponding solutions to provide reference for the library service.

Key words: big data technology, books, circulation data analysis, applied research

一、大数据背景下图书流通研究现状

在图书馆信息化建设的过程中,图书馆积累了海量的图书流通数据。然而,这些数据大多是结构化的数据,缺乏对图书流通行为的研究。在大数据背景下,需要对图书馆图书流通行为进行深度挖掘,从而更好地了解读者需求和行为模式,从而提升图书馆服务质量。近年来,在大数据时代的背景下,越来越多的学者开始关注到图书馆流通研究,并开展了一系列研究工作。例如,张少丽等人将大数据技术应用到图书馆流通数据分析中,揭示了当前图书流通的困境;张金梅等人利用数据挖掘技术对图书流通过程进行了分析。

1.1 大数据技术概述

在大数据时代,数据已经成为信息社会的重要资源。在信息技术的推动下,大数据技术的发展进入了快车道。大数据技术主要包括三个方面的内容,即数据获取、数据处理和数据分析。在大数据背景下,需要借助云计算、云计算平台等技术,对海量数据进行收集和存储,并将这些数据整合到统一的平台中。此外,在大数据背景下,需要借助数据挖掘等技术,对海量数据进行处理和分析,从而获取有价值的信息。需要注意的是,大数据技术不仅仅是一种新的技术形式,同时也是一种全新的思维方式和研究方法。在大数据时代的背景下,图书馆要想更好地满足读者需求,就需要积极开展大数据应用研究。因此,图书馆需要借助大数据技术来提升服务质量。

1.2 图书流通大数据的主要内容

大数据技术是指采用非结构化和半结构化数据来处理和分析海量数据的一种方法,能够在短时间内对大量数据进行有效地分析和挖掘。在图书馆管理中,图书流通是指图书的借阅和归还行为,其内容包括图书借阅量、图书借阅时间、图书归还时间等,这些数据都是以文本形式存在的。大数据技术主要应用在图书馆管理中,主要是通过图书流通数据的分析,挖掘出其中的规律。例如,通过对图书馆读者借阅时间和归还时间的分析,可以发现不同读者的借阅习惯。在大数据技术的支持下,图书馆管理人员能够获得更加全面、准确的读者信息,从而能够更好地为读者提供服务。因此,在大数据背景下,需要将数据分析技术与图书馆管理相结合,提升图书馆管理质量。

二、基于图书流通数据特征分析的图书馆服务体系存在的问题

2.1 资源利用率低

图书馆的资源利用率是指图书借阅率和流通率的比率。该比率

反映出图书在图书馆中的流通情况,也是图书馆在图书资源方面进行优化配置的重要依据。随着大数据时代的到来,通过对图书馆的资源利用率进行分析,可以得出以下结论:一是大多数图书的借阅率和流转率很低,而利用率较高的图书往往都是一些经典著作和重点图书,如一些医学类、经济管理类和文科类等方面的图书;二是图书馆内大量藏书并没有得到充分利用,很多读者在借阅过程中,对自己感兴趣或需要了解的书不够重视,而大量藏书被闲置。

2.2 读者满意度不高

图书馆作为公共服务机构,其宗旨是向读者提供优质、高效、便捷、全面的服务,以满足广大读者对知识、信息等方面的需求。为了达到这一目标,图书馆需要与广大读者建立良好的互动关系,通过加强沟通了解广大读者对图书馆服务体系中存在的问题和不足。但是当前大多数图书馆并没有从用户角度出发思考问题,尤其是在服务过程中缺乏主动与用户进行沟通交流的意识,对用户需要了解不够充分。在此基础上进行服务模式、服务内容等方面进行优化调整时,很容易产生较大的阻力。因此图书馆要想得到读者认可和支持,就必须主动与广大读者进行沟通交流,了解他们对图书馆服务体系中存在的问题和不足。通常情况下,图书馆主要通过问卷调查等方式对读者满意度进行调查。在实际调查过程中,由于图书馆经费有限、工作人员专业水平不高等原因,导致问卷设计存在一定缺陷。在调查结果中也可以看出部分读者对图书馆服务体系中存在的问题和不足不够了解;此外图书流通数据作为重要的参考依据之一,其准确性直接影响着读者对图书资源利用率和图书借阅情况等方面的评价。

2.3 缺乏个性化服务

目前很多图书馆在开展个性化服务时仍存在较大阻力。由于信息技术和互联网技术快速发展,人们获取信息的渠道越来越多、越来越快捷;另一方面由于图书馆本身资源有限,而且对资源开发力度不足等原因,无法为广大读者提供个性化服务。

三、大数据环境下图书馆服务体系中存在问题的解决对策

充分发挥大数据技术的优势,提高图书馆服务体系的效率和质量。在图书馆服务体系中应用大数据技术,可以有效提升图书馆服务效率和质量,尤其是在电子信息技术和网络技术迅速发展的当下,图书馆可以借助互联网平台实现资源共享,同时通过对读者行为数据的收集和分析,为读者提供个性化的服务,从而提高读者满意度。转变图书馆服务模式,实现与读者的互动。图书馆可以通过

应用大数据技术,加强与读者之间的沟通和交流,从而在了解读者需求的基础上提供个性化的服务。同时还可以利用网络技术,加强与其他图书馆之间的联系和互动。引入现代信息技术手段。图书馆可以通过应用云计算、物联网、大数据等现代信息技术手段,改变传统图书馆管理模式和服务模式,进而实现对读者需求和资源利用情况的动态监测。此外还可以通过网络技术与其他图书馆进行交流与合作。在大数据背景下,图书馆可以充分利用大数据技术实现对读者行为数据和图书资源数据的挖掘和分析,从而更好地为读者提供服务。

3.1 有效利用大数据技术,促进图书资源的整合

在这个过程中,图书馆要积极建立与读者之间的互动,从而实现与读者之间的交流和沟通。通过这种方式可以有效提升图书馆服务水平。首先,利用大数据技术对图书资源进行整合和分析。例如可以对读者借阅书籍的类型、数量等进行分析,同时还可以对不同类型的图书进行比较分析,从而为读者提供更加全面、准确的图书资源信息。其次,在此基础上为读者提供个性化的服务。例如在大数据背景下,图书馆可以利用大数据技术建立数据中心,并将各个数据库连接在一起,从而实现数据资源共享,并且还可以利用互联网技术将图书馆和其他图书馆进行连接和交流。最后,促进图书馆之间的资源共享。在大数据背景下,图书馆之间可以实现资源共享。例如在图书流通过程中,图书馆员可以通过网络技术向其他图书馆发出订单信息,或者是将其他图书馆的图书借给图书馆员借阅,从而促进图书馆之间的资源共享。

3.2 利用大数据技术对读者行为数据进行挖掘和分析,提高图书馆服务水平

大数据时代下,图书馆的读者行为数据数量越来越多,且类型也越来越多样化,对此图书馆需要通过大数据技术对读者行为数据进行分析,从而更好地为读者提供个性化服务。例如,图书馆可以通过大数据技术对读者借阅图书的数量、借阅图书的类别以及读者的借阅时间进行统计,并利用相关统计结果对不同类别读者的阅读习惯进行分析,从而更好地为读者提供个性化服务。此外,图书馆还可以通过大数据技术对读者的阅读行为进行分析,了解读者的阅读兴趣和阅读习惯,从而根据不同类型的读者制定不同的阅读策略。例如对于喜欢小说类图书的读者可以推荐一些相关类型的书籍;对于喜欢文史类图书和学术类图书的读者可以推荐一些与其相关方面的书籍;对于喜欢历史和地理类图书的读者可以推荐一些相关领域的书籍等等。通过大数据技术对读者行为数据进行分析,能够为图书馆更好地提供个性化服务奠定基础。

四、数据采集、预处理和清洗技术

在大数据环境下,为了实现对图书流通数据的深度挖掘,就必须采用科学有效的技术手段来实现对图书流通数据的采集、预处理和清洗。在此过程中,首先需要采集相关图书流通数据信息,包括图书的新旧、摆放位置、借阅状态等信息;其次需要进行预处理,包括数据清洗、归并、标准化等操作;最后需要进行清洗处理,包括数据修复等操作。通过以上操作,才能实现对图书流通数据信息的全面分析。

在大数据环境下,为了能够实现对图书流通数据信息的高效利用,就必须对采集到的相关数据信息进行及时处理和有效整合。具体来说,可以采用以下几种技术手段来实现对图书流通数据信息的采集。一是将图书流通数据信息与相关实体信息进行关联处理,进而构建图书流通数据信息库。在此过程中,要充分利用信息技术手段对相关实体信息进行全面、高效的采集和管理,确保所采集到的图书流通数据信息可以与相应实体信息进行有效关联。二是使用先进技术手段实现对图书流通数据的全面整合和处理。在此过程中,要充分利用先进技术手段对图书流通数据信息进行全方位、多角度的处理和分析,进而实现对图书流通数据的高效利用。三是构建高质量的图书馆服务体系。在此过程中,要充分利用先进技术手段来构建高质量的图书馆服务体系,从而为读者提供更为优质、高效和便捷的图书馆服务。

五、图书流通数据挖掘流程设计

利用数据挖掘技术,对图书流通数据进行分析,可以揭示出在图书馆服务中存在的问题,并提出相应的解决对策,进而改进和完善图书馆的服务。图书馆可以通过对图书流通数据进行预处理,为数据挖掘提供必要的技术支持。首先要对图书流通数据进行预处理,包括对数据进行清洗、转换、集成、整合和精简等操作,然后采用关联规则算法进行数据挖掘,将图书流通数据信息挖掘转化为关联规则,以便于分析出图书流通的规律和特点。由于数据挖掘过程中涉及到的操作较多,为了提高挖掘效率和质量,可将图书流通数据挖掘分为三个阶段:数据准备阶段、模型构建阶段和模型分析阶段。下面将从这三个阶段对图书馆图书流通数据挖掘流程进行详细阐述。

5.1 数据准备阶段

在图书流通数据挖掘过程中,要对图书流通数据进行预处理。图书流通数据挖掘是建立在大量数据的基础上,因此必须对这些数据进行必要的筛选和过滤,剔除不相关或者有重复记录的数据。同时,为了避免由于数据不准确或不完整而影响后续的数据挖掘工作,对原始数据进行必要的处理也是非常必要的。首先要对数据进行清洗,把数据中出现错误或不相关的部分去掉,保证后续数据挖掘工作顺利地开展;其次要将原始数据按照一定的标准进行转换,使得数据可以被计算机系统直接接受和处理;最后要将原始数据集成在一起,方便后续工作开展。在这一阶段中,需要根据图书馆实际情况合理选择和使用相关工具,以达到最优的效果。

5.2 模型构建阶段

在图书馆数据挖掘中,模型构建阶段的任务是将已有的数据转化为关联规则,以便建立关联规则模型。在该阶段中,要根据关联规则分析的需求,对原始数据进行预处理,形成有效的关联规则集合。然后使用 Apriori 算法,对关联规则集合进行挖掘,得到最优关联规则集。在图书流通数据挖掘中,为了使挖掘结果更有价值,需要对数据集进行清洗、集成、整合等处理。此外,还需要进行数据压缩和数据转换等操作。在数据挖掘过程中,可以利用 K-近邻算法对数据进行聚类分析,以便于建立合理的模型。在模型构建阶段需要处理的数据集较多,需要结合实际情况进行选择。

六、数据挖掘结果分析和挖掘结果应用

在数据挖掘过程中,对图书流通数据进行分析,可以得到读者行为、借阅行为、读者兴趣等数据信息,根据数据挖掘结果可以预测图书的需求,进而为图书馆的决策提供数据支持。通过对图书流通数据进行分析,可以发现一些有价值的规律,从而为图书馆开展服务提供重要参考。例如通过对读者借阅行为进行分析,可以了解读者的阅读倾向和兴趣点,进而为图书馆调整馆藏结构提供重要依据。

通过对图书流通数据的挖掘,可以得到以下结果:读者的阅读倾向和兴趣点与所借图书有较强的关联性;图书馆的馆藏结构需要调整和优化;图书馆馆藏结构和布局需要改进和优化;图书馆需要根据读者需求进行服务模式和服务方式的创新。在大数据背景下,利用数据挖掘技术进行图书流通数据分析,可以从图书流通数据中获取大量有价值的信息,能够有效帮助图书馆了解读者需求,进而根据读者需求调整馆藏结构和布局,创新服务模式和方式。

参考文献:

- [1]李东红.大数据时代高校馆藏图书流通策略研究[J].电脑知识与技术, 2020, 16(10): 107-108+111.
- [2]李萍,汪滢.定期借阅图书流通数据关联规则检测仿真研究[J].计算机仿真, 2020, 37(05): 354-357+452.
- [3]高艳,刘海峰.大数据技术在图书流通数据分析中的应用研究[J].信息记录材料, 2021, 22(05): 104-106.
- [4]齐航,杨婷婷,马致明.数据分析在高校图书馆的应用[J].兰台世界, 2021(2): 116-118.
- [5]方超.大数据技术在图书馆智能图书采编业务中的应用研究[J].科技经济市场, 2023(1): 36-38.