

数字图书馆个性化服务探讨

梁学明

(大连理工大学图书馆, 辽宁大连 116024)

摘要: 为了更好地满足读者各方面的需求, 数字图书馆根据自身的特点开展个性化服务。文章首先对读者兴趣和阅读习惯进行分析归纳总结, 剖析了读者兴趣的起源和形成过程, 然后对数字图书馆个性化服务系统及其关键问题进行探讨, 以希望数字图书馆能够根据不同类型读者开展各种有针对性的服务, 从而提高图书馆的服务水平和能力。

关键词: 读者兴趣模型; 个性化服务; 数字图书馆; 访问日志; 读者信息

中图分类号: G250

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2015.06.017

Study of Personal Service of Digital Library

Liang Xueming

(Library of Dalian Technology University, Dalian 116024)

Abstract: In order to meet the needs of readers, the digital library should carry out the individualized service according to its own characteristic, and study the reader's interest and reading habits, and analyze and summarize the interest and reading habits. According to different types of readers, we should carry out a variety of services, so as to improve the service level and ability of the library.

Keywords: reader's interest model, personal service, digital library, visit diary, reader information

面对图书馆大量的数字化资源, 如何方便快捷地为读者提供个性化服务是图书馆当前所面临的问题。早在上世纪90年代, 美国和一些欧洲国家相继开展了数字图书馆研究及其个性化信息服务的研究, 开发出了比较完善的数字图书馆个性化服务系统^[1-4]。而我国数字图书馆个性化服务研究始于1999年底, 在科技部支持的“中国数字图书馆示范系统”项目中就包括了数字图书馆个性化服务的内容。2000年初, 社会科学基金资助的“基于Web的数字图书馆定制服务系统”项目着手研究开发实用的数字图书馆个性化定制系统^[1]。比较有代表性的有深图朗思想ILAS、浙江大学的MyLibrary@ZJU系统和中国人民大学的“我的图书馆”等。笔者认为, 针对读者兴趣,

构建数字图书馆个性化服务系统将是一个有效的措施。对此, 本文将对数字图书馆个性化服务进行初步探讨。

1 读者兴趣起源与形成及个性化服务

从教育心理学的角度来说, 兴趣是一个人倾向于认识、研究获得某种知识的心理特征, 是可以推动人们求知的一种内在力量。读者兴趣是指图书馆的用户对图书馆的资源所关注或浏览的倾向。图书馆应分析研究读者兴趣, 确信读者对哪门学科感兴趣, 对哪种资源感兴趣, 从而为读者定制合适的个性化服务。

图书馆的个性化服务就是在提供信息资源和信息服务的过程中, 根据读者个人的爱好或特

作者简介: 梁学明(1977-), 男, 大连理工大学图书馆馆员, 研究方向: 读者服务工作、流通管理工作。

收稿时间: 2015年6月9日。

点, 提供有针对性和个性化的服务, 并能在读者需求发生变化时也能使读者在特定的时间和特定的地点得到服务, 从而确保所提供的服务是各取所需、各得其所^[5]。它是在正确的地点以恰当的方式, 及时、主动地满足不同用户特定的信息需求。在这个定义中有3个方面的个性化: 一是个性化的时空服务, 一切按照用户的希望, 在正确的地点及时满足读者需求; 二是个性化的服务方式, 根据用户的个人好恶, 以读者喜好的方式提供服务; 三是个性化的内容服务, 不再将所有的信息一股脑地推给用户, 而是将过滤后的合适的信息传递给用户^[6]。

面对日益丰富的图书馆资源, 图书馆需要建立一个信息服务系统来判断读者的主要兴趣指向, 推送适合读者兴趣的资源, 这是一种主动服务方式, 是以读者需求为中心, 利用数字图书馆信息资源开展不同层次、多种类型、满足读者个性需求的信息服务。这种个性化服务系统可根据读者要求, 提供个性化的界面设置、个性化的信息环境以及个性化的信息快报。特别是, 读者可定制个人馆藏, 提高读者的检索效率, 最大限度地满足读者的个性化需求^[7]。

分析、研究读者兴趣是为读者提供个性化信息服务的关键步骤。读者兴趣主要是通过读者打开浏览网页内容、访问日志记录、检索行为和使用方式等方面进行资源分类得到的。读者的兴趣可能是不断变化的, 在不同的时空中, 可能会有所改变, 对读者感兴趣的信息不断更新, 形成读者兴趣库, 然后对读者兴趣进行分析、评估、总结, 根据结果进行读者个性化服务。读者兴趣模型分析过程如图1所示。

当读者打开网页或访问网页内容时, 会在Web页面中留下许多隐含了作者兴趣偏好的访问数据。这些数据具有内容不确定性、信息多样化和分散化的特点。根据读者访问日志和浏览页面内容与资源之间的关联, 利用数字资源分类更新读者兴趣库。同时, 定期跟踪读者兴趣变化(增加或减少), 及时更新读者兴趣库, 及时反映读者的兴趣, 最终完成对读者的分类、评估总结。

这样的做法有两方面特点: 一是不需要人工参与和调整, 这个系统从对读者兴趣信息的获取、搜集到对读者兴趣的更新、分析、分类, 都是由系统自动化技术完成的。二是整个分析结果是动态的、变化的, 并且越来越完善和准确。

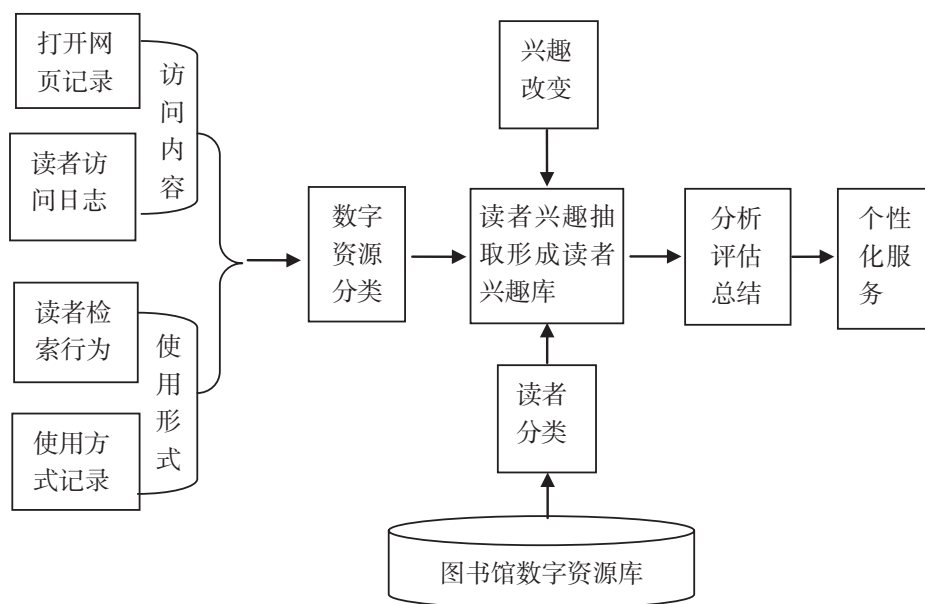


图1 读者兴趣模型

在系统初始状态时,信息量较少,分析所得的读者兴趣不十分完善,随着读者访问资源信息的积累,系统对读者兴趣的分析的结果将越来越准确。最后系统的分析结果会随着读者兴趣的变化而变化,也就是说系统所反映的读者兴趣始终是读者当前的兴趣,有效地发现读者即时兴趣程度和兴趣方向,为数字图书馆的个性化服务提供依据。

2 个性化服务系统的工作原理

数字图书馆是以网络技术为依据来进行信息资源传递,为读者提供服务。在数字图书馆信息服务中,不只是简单地将所有的数字化信息资源放在服务器上,让读者来浏览,而是根据读者特征及信息资源的访问历史等来识别读者兴趣以及信息资源的关联等来为读者提供个性化的服务。数字图书馆的个性化服务系统模型如图2所示。

当读者首次通过读者认证系统登录到数字图书馆时,个性化服务系统将自动生成读者的初始兴趣,也就是读者最基本的信息兴趣,并自动地

将对应的兴趣类型赋予读者。然后系统依据兴趣类型和该兴趣类型相匹配的数字资源去判定该读者对图书馆数字资源信息的需求,形成具有读者本身个性化的数字资源要求信息,并把这种需求发送到个性化信息调度中心进行处理。最后由个性化信息调度中心从资源库中提取读者需要的信息,反馈给读者。同时,信息采集模块对读者行为进行处理,送往个性化分析处理中心,个性化分析处理中心根据读者行为来分析读者兴趣和资源的关联。

(1) 资源关联分析及联机分析处理

资源关联分析及联机分析处理是个性化服务系统中的关键,是个性化信息处理中心提供个性化服务的重复依据。它主要是:利用数据仓库、联机分析和关联挖掘技术,通过对读者的历史访问数据进行分析 and 挖掘,发现不同类型、不同层次和资源之间的访问关联,并对读者、资源和资源的访问情况进行统计分析等。

(2) 读者信息搜集

读者信息搜集主要是收集读者的信息、读者

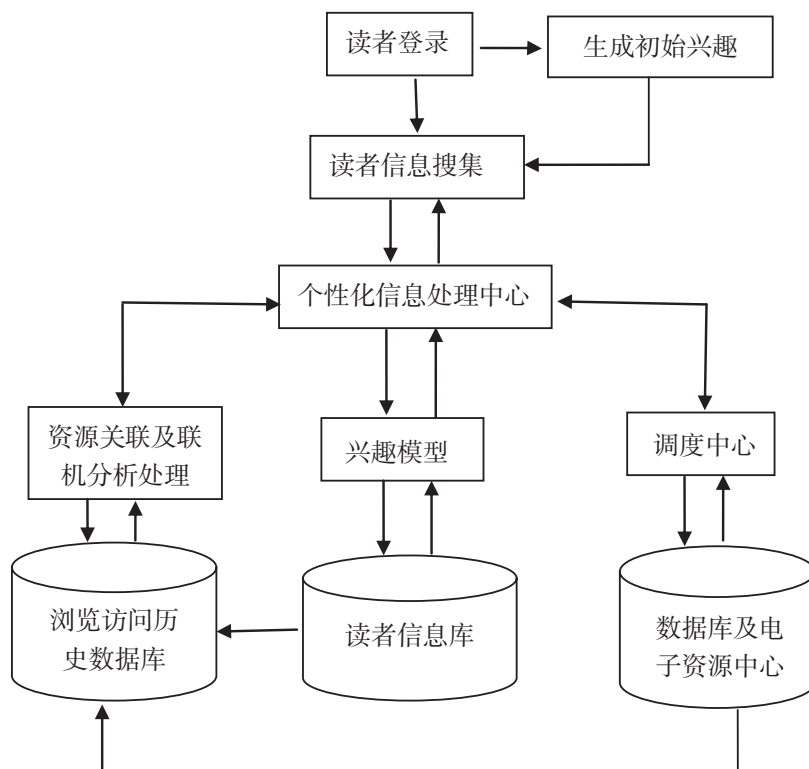


图2 个性化服务系统模型

行为(检索行为和使用方式记录等)、打开网页类型和访问日志、读者使用的时间段等信息,为个性化分析中心提供基本资料,其收集信息的质量直接关系到分析中心的处理精确度,信息在这个过程中只是被记录和进行简单的分析归纳。

(3) 个性化信息处理中心

在个性化信息处理中心中,根据读者访问历史数据和初始兴趣来分析读者的信息需求,并根据读者访问行为和检索行为及时更新读者的兴趣。当读者登录系统后,系统会自动找到相应的兴趣与读者兴趣匹配,系统将过滤和筛选兴趣模型,找到与之关联的资源来分析读者具体的需求信息,把不感兴趣的信息过滤掉,最终生成读者准确的信息需求结果。

(4) 调度中心

个性化信息处理中心所分析的读者信息需求传递给调度中心。调度中心根据相应的需求信息对数据库和电子资源中心发出调度命令,然后再反馈给读者,并通知读者兴趣模型修改更新读者兴趣。

(5) 数据库及电子资源中心、读者信息库、浏览访问历史数据库

数据库及电子资源中心是指读者所能访问到的数字图书馆的电子资源,包括数据库、电子期刊和电子图书等。中心拥有大量的数字化信息资源,这些信息资源一般由数据商提供,也有图书馆根据自身情况开发的。读者信息库保存着读者的基本信息及其兴趣情况,如行为日志、检索行为、访问行为及浏览行为等数据信息。读者的浏览访问历史数据库指的是读者对资源的访问历史数据,这些历史数据经过处理之后存放在历史仓库中,以供联机分析和数据挖掘,发现读者和资源之间的关联及访问统计结果。

3 相关问题研究

数字图书馆个性化服务融入数据挖掘方式可为读者提供更深层次的服务。通过数据挖掘可从历史数据中发现读者的访问模式、读者兴趣以及资源之间的关联等信息。在构建数字图书馆个性

化服务系统模型过程中,应注意以下几个关键性问题。

3.1 读者信息的获取和收集

读者信息的获取主要方式分为显式获取和隐式获取两种。显式获取要求在构建读者兴趣模型进行人工提供兴趣偏好的信息,给读者带来很多不便;隐式获取是建立兴趣模型时保存Bookmarks时的标签获得读者感兴趣特征词。隐式获取不需要读者参与,系统在获取读者信息时不会干扰到读者正常的活动。

(1) 通过网络调查和注册信息获取读者信息,主要是网页上设置调查页面,供读者填写或选择自己感兴趣的资源信息。这种方法可能不太准确,读者输入的词具有随意性,对读者兴趣的描述不能达到一定的专指性,无法准确表达读者的需求。

(2) 利用信息挖掘技术,通过一定的程序自动搜集读者感兴趣的资源。主要方式有分析网页关键字、分析电子邮件关键字、分析借阅记录、分析检索词等。通过系统自动搜集,可以提高个性化服务的准确性,节省大量的人力和时间成本,只是设计比较麻烦和复杂^[8]。

3.2 注意隐私保护问题

在读者信息行为收集过程中应充分考虑读者的隐私权问题,并不是所有的读者都愿意接受“主动”服务。当读者每次打开电子邮箱时,就会看到一大堆信息,尽管可能与读者的兴趣相关,但有时读者也会感到厌烦。网站首先需要考虑到读者是否愿意接受这种主动服务,解决的办法就是提供读者选择的机会,只有那些接受主动服务的读者才是个性化服务的对象,否则会影响到服务的效果。在提供服务时,一定要注意保护读者的隐私权,特别是要避免读者个人信息的外泄,保护读者的合法权益。

3.3 个性化信息的描述

对于个性化系统来说,为了跟踪读者的兴趣与行为,有必要为每个读者建立一个读者描述文件来刻画读者的特征与读者之间的关系。读者描述文件表达不同的个性化服务系统中的读者描述

文件各自的特点,有基于兴趣的读者描述文件和基于行为的读者描述文件。读者描述文件可以用文件来组织,也可以用关系数据库或其他数据库来组织。目前,有一些个性化系统采用基于XML的RDF(Resource Definition Framework)来表达读者描述文件,并利用支持XML的数据库系统来存储读者描述文件,这样不仅利用了XML的优点,也保持了系统的性能。

3.4 读者兴趣模型构建

读者兴趣模型构建是从有关读者兴趣和行为的信息中归纳出可计算的读者兴趣模型的过程。可计算性是构建读者兴趣模型的基本要求。在个性化服务系统中的读者兴趣模型不是针对读者个体的一般性描述,而是一种面向算法的、具有特定数据结构的形式化的读者描述。读者建兴趣模型是个性化服务的基础和核心,无论何种形式的个性化服务都需要建立对读者的描述,然后才能据此对不同读者提供个性化服务。读者兴趣模型有读者手工定制模型和自动读者建立模型。手工定制模型由读者自己手工输入或选择,如读者自己输入感兴趣的关键词列表或感兴趣的栏目。自动读者建立模型是根据读者的浏览内容和浏览行为自动构建读者模型,建模过程无须读者主动提供信息,不需要读者参与,由系统自动完成。

4 结语

数字图书馆个性化服务在国内开展得比较

晚,也比较少。对数字图书馆个性化服务缺乏系统的研究。数字图书馆个性化服务节省了读者的时间、提高了读者科研效率,必将受到读者的欢迎。数字图书馆引进个性化信息服务模式,是适应图书馆读者需求多样化的需求,是让图书馆的信息服务向更深、更广的方向发展,是提升图书馆读者服务质量的重要手段。信息社会中,个性化服务将成为图书馆读者服务的重要方向。数字图书馆应开展不同类型的个性化服务、为读者提供个性化信息,吸引更多的读者进入到图书馆,从而提升图书馆的社会地位和竞争力。

参考文献

- [1] 文露霜.基于Web2.0的数字图书馆个性化信息服务模式研究[D].武汉:华中师范大学,2009.
- [2] 陈玉梅.高校图书馆个性化服务问题与策略[J].企业导报,2011(6):41-43.
- [3] 刘阳.基于个性化学习的远程职业培训研究[D].沈阳:沈阳师范大学,2009.
- [4] 周灵威.试析图书馆个性化信息服务研究的现状与问题[J].图书馆论坛,2005,25(2):160-162,132.
- [5] 严根荣,朱毅洁.高校图书馆个性化服务探析[J].安徽文学,2008,(10):392-393.
- [6] 陈亮.基于多Agent的网络信息服务的技术研究[D].武汉:湖北大学,2009.
- [7] 张伟娜.馆藏中医古籍数字化共享平台的组建设想[J].医学信息,2014,27(1):26-27.
- [8] 张泯泯.基于自适应随机元胞自动机的数据挖掘技术[D].杭州:浙江大学,2004.
- [9] 王晴.云计算大数据时代图书馆的挑战和机遇[J].公共图书馆,2013(1):48-49.
- [10] 徐子沛.大数据及其成因[J].北京:科学与社会,2014(1):14.
- [11] 彭凤,黄力军.高校图书馆文献资源建设组织体系构建[J].四川图书馆学报,2014(3):25.
- [12] 杨海燕.大数据时代的图书馆服务浅析[J].图书与情报,2012(4):121-122.
- [13] 张建勋,古志良,郑超.云计算研究进展综述[J].计算机应用研究,2012,27(2):429-433.
- [14] 郭自宽,张兴旺,麦范金.大数据生态系统在图书馆中的应用[J].情报资料工作,2013(2):24.
- [15] Peter Spitzforn, Pongracz Sennyey.A Vision for the Future of Academic Library Collections[J]. International Journal of the Book,2007(4):4.
- [16] 罗彬,阳静,袁赟.数字图书馆中大数据存储的应用研究[J].科技与企业,2013(18):122.

(上接第86页)