

教育部人文社会科学研究一般项目(项目编号: 12YJA870021)

基于技术融合的 图书馆数字资源 利用服务机制研究

武三林 韩雅鸣 等◎著

海外借



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

教育部人文社会科学研究一般项目（项目编号：12YJA870021）

基于技术融合的图书馆数字资源 利用服务机制研究

武三林 韩雅鸣 等著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

· 北京 ·

图书在版编目 (CIP) 数据

基于技术融合的图书馆数字资源利用服务机制研究/武三林等著. —北京: 科学技术文献出版社, 2017. 3

ISBN 978-7-5189-2435-6

I. ①基… II. ①武… III. ①数字图书馆—信息资源—资源利用—研究
IV. ①G250.76

中国版本图书馆 CIP 数据核字 (2017) 第 047875 号

基于技术融合的图书馆数字资源利用服务机制研究

策划编辑: 周国臻

责任编辑: 赵 斌

责任校对: 张叫咪

责任出版: 张志平

出 版 者 科学技术文献出版社

地 址 北京市复兴路 15 号 邮编 00038

编 务 部 (010) 58882938, 58882087 (传真)

发 行 部 (010) 58882868, 58882874 (传真)

邮 购 部 (010) 58882873

官方网址 www.stdp.com.cn

发 行 者 科学技术文献出版社发行 全国各地新华书店经销

印 刷 者 虎彩印艺股份有限公司

版 次 2017 年 3 月第 1 版 2017 年 3 月第 1 次印刷

开 本 787 × 1092 1/16

字 数 332 千

印 张 14.5

书 号 ISBN 978-7-5189-2435-6

定 价 68.00 元

版权所有 违法必究

购买本社图书, 凡字迹不清、缺页、倒页、脱页者, 本社发行部负责调换

前 言

数字资源是图书馆不断创新的基础,现代化技术是图书馆发展的动力。开展基于技术融合的图书馆数字资源利用服务机制研究,有利于数字资源的合理配置,有利于现代化技术的广泛应用,有利于图书馆各类资源的有效利用。

随着网络技术和信息技术的迅速发展,图书馆数字资源已成为信息资源的主体,是国家的重要资产,已成为国家战略的一部分。针对数字资源数量庞大、类型各异等特点,如何利用多种现代化技术,有效快捷地收集、组织、开发与利用数字资源,成为图书馆需要解决的重大问题。

自20世纪90年代以来,各国政府、图书馆、信息机构和企业对数字资源的开发与利用高度重视,纷纷把数字资源的生产、传播和利用作为国家信息化建设的关键和重点。为抢占发展的制高点,各国纷纷投入巨资将数字资源建设作为提升本国创新发展能力的重要手段。

近些年来,我国政府对数字资源建设极为重视。1997年,国家启动了“数字图书馆工程”;1998年,教育部启动了“中国高等教育文献保障系统(CALIS)”工程;2000年,七部委联合启动了“国家科技数字图书馆工程(NSTL)”;2001年,中国科学院启动了“中国科学数字图书馆工程(CSDL)”,中国社会科学院启动了“中国社会科学院数字图书馆工程(CASSL)”;2002年,国家图书馆启动了“全国文化信息资源共享工程”;2004年,国务院转发了由科技部、国家发改委、教育部、财政部联合制定的《2004—2010年国家科技基础条件平台建设纲要》,国家与地方层面的科技基础条件共享平台项目纷纷启动建设,面向社会开放;2006年,中共中央办公厅印发《2006—2020年国家信息化发展战略》。这些都充分表明国家对数字资源建设与利用的高度重视。

本书属教育部人文社会科学研究一般项目“基于技术融合的图书馆数字资源利用服务机制研究(项目编号:12YJA870021)”研究成果,主要以图书馆数

字资源共享与利用服务为目的,系统论述了图书馆数字资源建设与发展的基本理论,研究了数字资源利用服务战略规划、数字资源选择与组织理论、数字资源挖掘技术、数字资源整合技术、数字资源云存储与云安全技术,建立了基于技术融合的图书馆数字资源管理机制与利用服务机制,构建了图书馆数字资源利用服务机制的绩效评价体系,以实现图书馆数字资源的科学化、规范化管理与利用服务。本书主要有以下特点:

第一,从共享的视角,将图书馆数字资源利用服务机制视为图书馆发展与创新系统工程的重要部分。本书不仅研究数字资源采集、组织、开发、管理与利用服务的全过程,尤其重视现代化技术与各种管理功能的协调,建立了基于技术融合的图书馆数字资源利用服务机制,这种机制可使管理科学化、服务便捷化、利用有效化,能更好地服务于社会 and 用户。

第二,从可持续的角度,对图书馆数字资源的利用服务进行了系统科学的规划。图书馆数字资源利用服务是一个知识密集型、技术应用型、协调合作型工作,图书馆数字资源利用规划就是要确保数字资源共享目标实现的策略和途径。本书对图书馆数字资源组织、业务流程、信息技术、用户及工作人员等各要素和关联度进行全面研究,统筹规划,使图书馆目标清晰,对图书馆数字资源利用服务的规划科学、可行。

第三,从技术融合的角度,就图书馆数字资源整合、数据挖掘、云存储与云安全等技术进行深入研究。本书综合多种因素,应用多种技术,对图书馆数字资源整合、数据挖掘、云存储与云安全进行了研究,建立图书馆数字资源共享利用服务平台,用户可方便快捷地获取所需资源。也就是说,无论是图书馆数字资源的理念,还是图书馆数字资源管理与利用服务机制,都体现了基于技术融合应用实践的特色。

第四,从科学管理的角度,构建了图书馆数字资源利用服务绩效评价体系。本书应用绩效评价理论,对图书馆数字资源技术、管理、利用服务等方面进行了分析,构建了图书馆数字资源利用服务绩效评价体系,为图书馆科学化管理提供依据,为图书馆决策提供参考。

本书共分9章:第1章为图书馆数字资源利用服务概论;第2章为图书馆数字资源利用服务战略规划;第3章为图书馆数字资源选择与组织;第4章为图

书馆数字资源挖掘；第5章为图书馆数字资源整合；第6章为图书馆数字资源云存储与云安全；第7章为基于技术融合的图书馆数字资源管理机制；第8章为基于技术融合的图书馆数字资源利用服务机制；第9章为基于技术融合的图书馆数字资源利用服务机制绩效评价。

本书第1章由吴汉华撰写，研究生姚小燕、王琛、刘静参与撰写；第2章由武三林撰写；第3章由吴汉华撰写，研究生刘静、张芳参与撰写；第4章由尚成国、张国瑜撰写；第5章由韩雅鸣撰写，研究生李雅韵参与了资料的整理；第6章由贾伟、郑芳撰写；第7章由徐彤阳撰写；第8、第9章由武翔宇撰写。武三林负责大纲编写与全书统稿，张玉珠、武翔宇、隗玲参与了本书统稿与调研工作，王舒波参加了本书的资料收集与整理工作。本书集中了撰写者的系列成果，借此对为研究成果付出艰辛的著者表示感谢！

本书在撰写过程中，吸收了国内外大量的研究成果，参考引用了许多专家的有关著述，在此表示衷心的感谢！对科学技术文献出版社的周国臻、赵斌先生等在策划编辑过程中付出的艰辛表示感谢！

著者在撰写本书时查阅了大量资料，由于研究涉及内容较多，加之撰写执笔人较多，书中难免有疏漏之处，希望读者给予指正。

武三林

2016年9月于太原

目 录

第1章 图书馆数字资源利用服务概论	1
1.1 图书馆数字资源	1
1.2 图书馆数字资源建设	6
1.3 图书馆数字资源服务	11
1.4 国内外基于多技术融合的数字资源利用服务研究进展	14
第2章 图书馆数字资源利用服务战略规划	26
2.1 图书馆数字资源利用服务战略规划基础理论	26
2.2 研究背景、意义、目的与作用	32
2.3 图书馆数字资源利用服务战略规划对象分析	39
2.4 图书馆数字资源利用服务战略规划的定位与原则	46
2.5 图书馆数字资源利用服务战略规划的内容	47
第3章 图书馆数字资源选择与组织	57
3.1 数字资源选择的原则	57
3.2 数字资源的采集	59
3.3 数字资源组织的标准	66
3.4 数字资源组织的方法	74
第4章 图书馆数字资源挖掘	78
4.1 图书馆数字资源挖掘的含义与意义	78
4.2 数字资源挖掘的准备与处理	81
4.3 数据挖掘的原理与方法	85
4.4 数据仓库与 OLAP 技术	88
4.5 数据挖掘在图书馆数字资源利用服务的策略	97
第5章 图书馆数字资源整合	108
5.1 图书馆数字资源整合的必要性	108

5.2	图书馆数字资源整合的原则与内容	110
5.3	图书馆数据整合的模式	115
5.4	图书馆数字资源与服务整合的机制	127
5.5	图书馆数据整合技术	131
第6章	图书馆数字资源云存储与云安全	137
6.1	图书馆数字资源云存储与云安全的内涵与意义	137
6.2	数字资源云存储与云安全的内容与体系结构	139
6.3	数字资源云存储与云安全技术	143
6.4	图书馆数字资源利用云存储与云安全研究	145
第7章	基于技术融合的图书馆数字资源管理机制	152
7.1	管理机制研究概述	152
7.2	基于技术融合的图书馆数字资源管理机制模式构建	157
7.3	图书馆数字资源多技术管理	165
第8章	基于技术融合的图书馆数字资源利用服务机制	175
8.1	数字资源利用服务概述	175
8.2	基于技术融合的图书馆数字资源利用服务机制	177
8.3	基于技术融合的图书馆数字资源利用服务政策法律保障体系	192
第9章	基于技术融合的图书馆数字资源利用服务机制绩效评价	199
9.1	评价的内涵	199
9.2	国内外图书馆数字资源利用服务评价述评	199
9.3	基于技术融合的图书馆数字资源利用服务机制评价的目的和意义	201
9.4	基于技术融合的图书馆数字资源利用服务机制评价因素分析与确定	202
9.5	基于技术融合的图书馆数字资源利用服务机制评价的方法	205
9.6	基于技术融合的图书馆数字资源利用服务机制评价体系构建的原则	207
9.7	基于技术融合的图书馆数字资源利用服务机制评价基本程序	210
9.8	基于技术融合的图书馆数字资源利用服务机制评价体系构建	210
参考文献		214

第1章 图书馆数字资源利用服务概论

在云计算、智慧城市、大数据等信息技术迅猛发展的背景下,数字资源的生成、存储、发布、传播和接收方式都发生了革命性的变化。数字资源正在从信息资源中的辅助地位走向主导位置,它具有传统文献资源难以比拟的巨大优势。数字资源是一个国家的数字资产,是学术研究信息的数字存档,一个国家的科技创新能力及与此相关的国际竞争力都依赖于其快速、有效地开发与利用数字资源的能力^①。由此可见,数字资源已经成为国家战略的一部分,已经真正地 与物质、能源并称为社会的三大要素。

1.1 图书馆数字资源

1.1.1 图书馆数字资源的概念

自从计算机技术应用到图书馆之后,图书馆便迎来了数字化时代,图书馆的文献资源开始变为数字资源,图书馆数字资源的概念也开始出现在各种文献中,被人们广泛使用。从语义语境范围上看,数字资源属于图书馆信息资源的范畴。对于图书馆信息资源观,目前并没有比较权威的定义,学者们所提出的定义归纳起来主要有两种:从广义角度看,图书馆信息资源是图书馆信息活动中所涉及的一切资源,不仅包括信息,而且包括管理方法、技术、人员等;从狭义角度看,图书馆信息资源就是指图书馆的信息本身,这种信息是以各种标记和形式组成的信息集合,比如文字、图像、音频、视频等信息。笔者认为,图书馆数字资源应当采用广义的定义,即与图书馆数字资源活动相关的馆员、技术设备、规章制度、用户等都应当作为图书馆的数字资源。

学者们在研究图书馆时,一般都是从要素观的视角选择一个方向进行研究,这些研究成果能被学术界所认可。同理,从要素观的角度对图书馆数字资源进行研究也符合学者们的研究思维模式,这种要素观符合图书馆数字资源的广义定义。在互联网快速发展的时代,图书馆将面临数字资源规模庞大、难以收集的困境,因此,图书馆如何有效、快速地收集数字资源是图书馆面临的重要问题之一。要解决这些问题,就需要图书馆集合各种人力、物力,从不同角度来分析解决,这些不同视角的分析与解决方式都与图书馆数字资源有关。据此,图书馆数字资源可以定义为:图书馆已收集或可收集的所有以数字化形式存在的各种类型的信息资源,在收集过程中所涉及的技术、馆员、规章制度、用户服务等都属于图书馆数字资源。

^① 肖希明. 数字信息资源建设与服务研究 [M]. 武汉: 武汉大学出版社, 2008: 1.

1.1.2 图书馆数字资源的产生和发展

图书馆数字资源的产生与计算机技术发展密不可分,自1946年电子计算机在美国问世以来,图书馆的各类资源便开始分阶段、逐级地数字化,这种采用二进制“0”和“1”来表示的数字资源,正成为图书馆资源中的核心内容。由于电子计算机实现了信息数字化,由此产生了馆藏资源存储介质与传输方式的变化,即由传统纸质馆藏存储介质转向磁质电子化馆藏存储介质,传输方式由传统的用户到馆领取资源转向通过网络以电流信号远程化的方式进行传播。在现代信息技术迅速发展的背景下,图书馆数字资源已经开始与各项通信技术、网络技术、流媒体技术相结合,形成了更加开放、智能的数字馆藏资源。从数量上看,图书馆数字资源由少到多,馆藏数字资源也进入了以T为存储单位的时代;从资源存储格式类型上看,既有文本格式的文档,也有音频、视频文档,还有各类自建数据库所收集的图片类型文档。

从图书馆数字资源的产生上看,最早的数字资源形式主要是便于查考的书目数据库。图书馆数字资源的起源可以从情报检索系统算起,最早的情报检索系统由美国海军兵器中心(NOTS)图书馆在1954年研制成功。1966年,美国化学文摘社开始利用计算机自动编排《化学题录》。1963年,美国国立医学图书馆开始使用计算机编制《医学索引》(Index Medicus)。1967年秋天,医学文献分析与检索系统MEDLARS(Medical Literature Analysis and Retrieval System)开始实验性联机检索,以神经病学数据库作为实验性联机检索数据库^①。据《Computer-readable databases: A directory and data sourcebook》统计,到1965年,全世界大约有20个数据库可供使用,但这时的数据库存储介质仅限于机读磁带,内容以科技文献书目、索引、文摘为主,检索也是以脱机批处理的方式进行,因此应用并不广泛^②。

20世纪60年代中后期,随着计算机技术和网络通信技术的快速发展,人们利用计算机进行远距离传输,使信息交换变得更加容易,国际联机检索系统也开始走上发展轨道。国际联机检索就是用户使用终端设备,远距离从国际联机检索中心迅速而准确地获取数字化信息,其实质就是数据库和通信的结合^③。1965年以后,国际联机检索系统开始进入规模化和繁荣阶段,比如DIALOG、MEDLINE、ORBIT、JOIS、QUESTEL、JOIS等国际联机检索系统相继出现。伴随着联机检索系统进一步走向成熟,数据库也开始出现,并且走上了商业化的轨道,科学研究及各种资料也逐渐成为数据库的内容。

20世纪70年代至80年代末,卫星通信技术、光纤通信技术、个人计算机的产生和发展给数据库联机检索创造了空前的发展机会,联机检索已不受地域限制,向国际化发展,个人用户也开始加入数据库检索行列;数据库的生产由美国向西欧扩展,在短短10年内即增长了10倍,到80年代末,数量已达到3600多个;数据库的容量增加,存储介质增加了光盘,随之产生了光盘检索数据库系统;数据库类型也发生了变化,除书目、索引、文摘数据库外,全文数据库也开始迅速增加,而数值数据库、指南数据库等也相继出现^④。吴慰慈教

① 于海生. 美国国立医学图书馆医学文献分析与检索系统: MEDLARS [J]. 软件, 1995 (12): 30-34, 8.

② 肖珑. 数字信息资源的检索与利用 [M]. 2版. 北京: 北京大学出版社, 2013: 6.

③ 肖希明. 数字信息资源建设与服务研究 [M]. 武汉: 武汉大学出版社, 2008: 3.

授将数据库分为文字型、数字型、图像型和声音型等类型,其中,文字型数据库占各类型数据库 70% 的比例;20 世纪 60 年代至 70 年代,文字型数据库中的书目型数据库占主体地位,到 80 年代中期,全文数据库扶摇直上,从 1985 年的 28% 增加到 1997 年的 50%,而书目型数据库则从 57% 降到 24%;截至 2004 年,全文数据库的数量与书目型数据库的比例已经达到 2:1;从数据库容量上看,1975 年,每个数据库记录数的平均值为 17.3 万条,1985 年达到 55.8 万条,1997 年达到 116.6 万条,2001 年达到 153.4 万条^①。除了联机数据库以外,这一时期还有光盘检索数据库,主要是因为此时光盘技术发展成熟,光盘存储密度高,读取较联机数据库快,容量也大,所以光盘检索占据了一定的市场地位,比如美国 UMI 公司和银盘公司等。

进入 20 世纪 90 年代中期,随着信息高速公路计划的实施,互联网络逐渐走入寻常百姓家,基于网络的数字资源及其检索系统也开始成为普通图书馆所建数字图书馆馆藏的一部分,并且其重要性逐年提升。互联网络使用一个统一的 TCP/IP 协议连接和通信,所有的网络计算机都应遵循这一协议进行数据交换,这便使得所有数据库及检索系统互联互通成为可能,因为它们需要遵从同一套特定的协议和规范。随着网络资源的丰富,数字图书馆的资源也越来越丰富,比如,电子图书、电子期刊、会议资料、电影欣赏、各类讲座视频、各种音频文件等。近年来,越来越多的出版机构也开始将资源搬到网上,比如,各大专业期刊编辑部将期刊电子版放在编辑部网站,并提供免费下载,一些作者将自己的论文、书籍提交到特定机构知识库,开放获取(Open Access)逐渐受到学者们的认可。随着搜索引擎技术的成熟发展,一站式、智能化的检索成为用户的主要需求,这方面的技术也趋于成熟,比如,北京大学图书馆的“未名学术搜索”,可以对馆藏所有资源进行一站式的搜索,还能将各机构数据库结合在一起,只需用一个检索界面进行搜索,如图 1-1 所示;而谷歌公司的搜索引擎,现在已经能部分地给出用户检索所需答案,这些智能化的趋势也反映出图书馆数字资源利用的未来趋势。

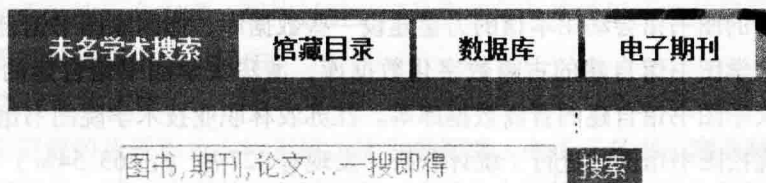


图 1-1 北京大学图书馆“未名学术搜索”

1.1.3 图书馆数字资源的类型

随着信息技术的日益发展,图书馆数字资源的类型也更加丰富,信息资源的范围更加广泛,形式多种多样。学者们从不同角度,将图书馆数字资源分为不同类型,比如:肖希明教授从数字资源内容的表现形式将数字资源分为数据库、电子期刊、电子图书、电子报纸等;

^① 吴慰慈. 网络环境下图书馆信息资源建设: 兼谈数据库产业的现状和发展趋势 [J]. 图书馆论坛, 2004 (6): 11-14, 24.

按存储信息的介质,可分为光介质和磁介质;按数据传播范围,可分为网络信息资源和单机信息资源;按数字资源的生产途径和发布范围,可分为商用数字资源、网络公开学术资源、特色资源等;按数字资源的生产方式,可分为原生数字资源和复合数字资源;按数字资源的性质和功能,可分为一次数字资源、二次数字资源、三次数字资源^①。肖珑依据不同标准将图书馆数字资源进行不同划分,依据数字资源的性质和功能,可分为一次文献、二次文献、三次文献;依照数字资源的生产途径和发布范围,可分为商用数字资源、网络公开学术资源、特色资源等;依照数字资源的载体,可分为光盘数据库、磁带数据库、硬盘数据库等;依照数字资源的格式,可分为文本信息、音频信息、视频信息等;还可以依照学科类型来划分数字资源^②。尽管图书馆数字资源有不同的划分方式,但有很多划分方式没有太大的指导实践工作的价值。笔者依据图书馆的实践工作需求,将图书馆数字资源划分为馆藏书目数字资源、自建数据库资源、采购中文数据库资源、采购外文数据库资源、开放存取学术导航资源等类型。

(1) 馆藏书目数字资源

这类资源主要存储图书馆馆藏书目信息,方便用户查找馆藏书目。书目数字资源也是图书馆馆藏图书与期刊、报纸信息的一个总目次,方便用户按照自己的阅读需求查找资源,并找到具体馆藏借阅资源。一般情况下,书目数字资源不仅要标明书目名称,还需要对图书内容进行简单提要,提供藏书的索书号,并指明馆藏物理位置。有些图书馆的书目数字资源还能提供用户的评论信息。在新技术发展逐步成熟条件下,一些书目资源也开始通过虚拟化的途径表现出来,比如通过虚拟3D图像设计的形式来显示模拟现实场景,以显示图书馆馆藏书目的位置,甚至能够精确到第几架第几格,还能显示旁边书目的信息。这种具有虚拟场景和人工智能的书目数字资源符合用户的阅读需求,也是馆藏书目数字资源检索的未来发展趋势之一。

(2) 自建数据库资源

一些有能力的图书馆会动用本馆的力量建设一些数据库,如国家图书馆自建的中国记忆数据库,北京大学图书馆自建的古籍数字化数据库,清华大学图书馆自建的学位论文数据库,山西财经大学图书馆自建的晋商数据库等。江苏农林职业技术学院图书馆的高原等对全国672所本科院校图书馆网站进行了统计分析,发现有427所(占63.54%)高校图书馆开展了特色数据库建设,自建数据库数量总数为2082个。自建数据库数量排在前10位的省(市)、自治区高等教育发达、科技经济发展迅速,具体包括:北京市206个、广东省201个、江苏省137个、四川省127个、陕西省101个、上海市100个、山东省96个^③。但是对于甘肃、青海、内蒙古、宁夏、新疆等西部边远省(区),自建数据库的总量还非常少。这些地方有很多珍贵的民族文化资源,亟待保护,因此,这些地方自建数据库的潜力非常巨大。

① 肖希明. 数字信息资源建设与服务研究[M]. 武汉: 武汉大学出版社, 2008: 5-9.

② 肖珑. 数字信息资源的检索与利用[M]. 2版. 北京: 北京大学出版社, 2013: 13-15.

③ 高原, 邹金花, 解阳, 等. 全国本科院校图书馆自建数据库建设现状分析[J]. 情报探索, 2015(7): 88-93.

(3) 采购数据库资源

数据库资源查找方便、检索快速、收录全面,受到越来越多用户的喜爱,也是图书馆必须收藏的数字资源。中文网络数据库包括中国知网数据库、维普科技期刊数据库、万方数字化期刊全文数据库、人大报刊复印资料全文数据库、超星数字图书馆、方正电子图书、民国时期期刊全文数据库等;外文网络数据库包括 Academic Search Premier (学术期刊集成全文数据库)(EBSCO)、American Chemical Society (美国化学学会数据库)、ERIC (教育资源信息数据库)、JSTOR (西文过刊全文数据库)、LISA (图书馆与信息科学文摘库)、ProQuest Dissertations & Theses (PQDT) Full Text (博硕论文全文数据库)等。用户希望图书馆所收录的数据库资源越全面越好,但是图书馆采购数据库需要花费非常高昂的费用。普通的知网数据库,年度采购费至少在20万元以上;对于外文数据库,诸如北大、清华、复旦等“985”类高校,年度购置费都超过千万元。因此,各馆所采购的数据库资源一般需要根据自己的需求来确定。

(4) 开放存取学术导航资源

随着互联网络技术的普及,开放存取的学术导航资源已经成为图书馆需要关注的重要资源。从理论与技术层面看,创建开放存取资源平台不是目前有效开发利用开放存取资源的障碍,但是如何解决遍布于网络的开放存取资源“孤岛”现状,实现开放存取资源基于知识层面的在线集成服务,却是学者与实践者面临的重要课题^①。DOAJ 和 DOAR 已经是开放存取资源的重要集散地,比如开放存取资源目录、开放存取期刊全文数据库等。开放存取的学术导航资源包括 BioMed Central (生物学与医学免费资源)、ChemSpider (免费检索化学结构式的网站)、China Open Access Journals (中国科技期刊开放获取平台)、Directory of Open Access Journals (开放获取期刊)、中国学术会议论文在线、中国预印本服务等,这些都是比较重要,也是图书馆应当收集的开放存取学术资源。

1.1.4 图书馆数字资源的特点

图书馆数字资源的生产方式、载体材料、记录类型等与传统纸质文献资源完全不同,其具有与传统纸质文献资源相异的特性。

(1) 资源表现形式多样

图书馆数字资源的表现形式多种多样,从手机阅读、电子工具书、商业数据库、书目数据库、索引数据库等,到统计报表、电子地图、电子留言板等。这些数字资源的表现形式既可以是文字、图表、网页等静态信息,又可以是图、文、声、像并茂的动态信息资源。除此之外,各类数字资源还可以借助计算机实现任意组合编辑,随时进行资源格式的更新与变换。这种形式使得信息组织方式发生巨大变化,不仅以知识和信息为基本单元,而且充分展示了这些单元之间的逻辑关系^②。

(2) 资源内容主题丰富

我国数字资源经过几十年的发展,其内容非常丰富,几乎覆盖了科技、工程、经济、商

① 邓君,毕强,宋文凤.国内开放存取资源在线集成服务理论进展[J].情报科学,2012(7):1103-1108.

② 章云兰,万跃华,舒炎祥.数字资源检索教程[M].北京:科学出版社,2006:3.

业、金融、财政、交通、税务、文教、卫生、新闻出版、能源和国家事务等各方面的信息,其中经济与社会方面的数据库约占 55%,科技方面的数据约占 45%^①。另外,还有很多原生态的数据库资源,比如国家图书馆的民俗记忆数据库、山西财经大学图书馆所建设的晋商文化数据库等,这些都是非常重要的数据资源。据互联网信息中心的统计数据,我国产品数据库所占比例已达到 61%,图片数据库占 37.7%,企业名录数据库占 20.4%,报刊新闻数据库占 17.0%,科技信息数据库占 16.2%,政策法规数据库占 15.1%,期刊论文数据库占 11.9%,人物数据库占 10.3%,金融股票信息数据库占 4.0% (各类型数据库中有重复)^②。我国的数字资源建设,从主题上看已经非常完善,这也表明数字资源正引起社会公众重视,开始成为社会发展的重要引擎。

(3) 资源共享程度高

与传统纸质资源相比,数字资源是利用现代信息技术而生产的数字化资源,其共享性主要体现在:第一,数字资源可以进行大量复制,其复制品与原件无任何差异,且复制品对原件无任何损害,因此,数字资源本质上是一种易于共享的资源;第二,数字资源主要由计算机来读取,其是以计算机编码形式(“0”和“1”)存储,这种传输方式不受时间和空间限制,方便远距离传输,更容易共享。

(4) 动态更新、时效性强

网络信息资源从根本上改变了人们发布信息、获取信息、交流信息的习惯,对传统的出版行业产生了颠覆作用。在互联网思维之下,传统的服务付费模式已经行不通,同时传统出版方式也发生了巨大变革。以电子期刊为例,从投稿、评审、稿件编辑等都在网上进行,大大缩短了文献编辑出版时间,使得信息的时效性大大增强。在 Web 2.0 环境下,传统出版物的修订更加容易,可以随时更新、发布最新动态。每一位用户都可以随时修改自己所发布的博客、微博、微信朋友圈等信息,使得信息动态性高、时效性强。

(5) 资源获取更加便利

由于数字资源是以二进制电子信号作为存储符号,其存储和传输依靠电流通信或光波传递的方式,因此数字资源具有非常强的存储和传递能力。在计算机软件高度发达的时代,人们的信息检索速度更加快捷,可以进行多途径检索,且检索信息不受时间、地点的限制,特别是网络信息资源检索过程中所利用的超文本链接、图像检索、流媒体检索等。在数字资源时代,人们利用超文本检索能够将不同国家、地区、文本、语种的信息一次性跨库检索出来,方便快捷。

1.2 图书馆数字资源建设

1.2.1 图书馆数字资源建设的含义

数字资源是一个国家和社会的重要战略资源,图书馆数字资源在其中占有重要地位,和

① 肖珑. 数字信息资源的检索与利用 [M]. 2 版. 北京: 北京大学出版社, 2013: 10.

② 肖珑. 数字信息资源的检索与利用 [M]. 2 版. 北京: 北京大学出版社, 2013: 10-11.

其他数字资源相比,图书馆数字资源在数据质量和规范上具有优势。

首先,数字资源是指以数字形式存储在光、磁等非纸质介质的载体中,通过网络通信及计算机或终端再现出来的资源,其表现形式可以是文字、图像、声音、动画等多种形式^①。图书馆的数字资源可以自行建设,也可从数据服务商手里购买,再提供给一定的用户群阅读和使用。在购买和建设的过程中,需要图书馆员通过选择、组织、开发、布局等方式,对图书馆数字资源进行优化,从而为用户提供高质量的数字资源服务。

其次,数字资源与自然资源不同,它不是一次性资源,数字资源在被利用后依然保有它的价值,还可以再次被人们使用。而且,其价值并不因其被使用的次数多而降低或者失去其价值。在一段时间内,数字资源的使用、被引用的次数多,往往能够说明这类数字资源更有价值,因而才会被人们所使用。图书馆数字资源在被引进的过程中需要资源管理者对其删繁就简、归纳整理,在被储藏的过程中需要对其进行进一步的优化。这都是图书馆数字资源建设中不可缺少的环节,只有这样图书馆数字资源才能不断丰富,不断推陈出新。

最后,数字资源建设商在学术、教育、文化资源建设上的发展十分迅速,无论是数字资源的规模还是数字资源的种类都是图书馆的数字资源建设所无法比拟的^②。这对图书馆数字资源建设来说既是机遇又是挑战,图书馆数字资源建设过程中需要与数字资源建设商合作,结合自己的优势,跳出传统资源建设的固有思维,扩展数字资源利用的方式与方法,本着为用户更好服务的理念进行图书馆数字资源建设。

图书馆数字资源的这些特性正是我们在进行图书馆数字资源建设时需要注意的。图书馆数字资源建设的定义就是运用数字化的技术手段,将传统文献进行数字化加工和处理,同时对已形成的数字资源进行归纳整理,使之形成有价值的可以更好服务用户的数字资源的过程,这个过程包括数字资源的加工、存储、检索、传输和利用。

这个定义包含以下几个方面的内容:

第一,图书馆的数字资源建设需要运用数字化手段对传统文献进行数字化处理,产生数字化的文字、图像、音频、视频等信息,使用户更方便快捷地检索、学习及利用。

第二,图书馆数字资源建设还要对已经生产或者购买的数字资源进行科学的选择、组织、开发、布局。选择就是对数字资源删繁就简,对图书馆试用数字资源进行科学选择、完善管理和比较论证,能够帮助图书馆做出正确的订购决策,使图书馆的数字资源建设和引进工作更加具有针对性,能够为用户提供更多、更精准的数字资源,提升数字资源的使用效率^③。组织就是对采集的数字资源进行加工、整序,揭示其内容和形式特征,组成各种信息检索工具和检索系统,并且以感知、思维、创造等方式去深入信息内容,去发现、发掘、提取有现实或者潜在价值的信息,以供用户有效利用^④。开发就是以用户需求为导向,进行数据库建设、资源整合、信息集成服务,以及与大众搜索引擎合作^⑤。布局就是对图书馆数字

① 马费成. 数字信息资源规划、管理与利用 [M]. 北京: 经济科学出版社, 2012: 2.

② 苏新宁. 大数据时代数字图书馆面临的机遇和挑战 [J]. 中国图书馆学报, 2015 (6): 4-12.

③ 杜莹琦. 高校图书馆试用数字资源的选择与评价策略 [J]. 图书馆建设, 2011 (7): 13-15.

④ 马费成. 数字信息资源建设与服务研究 [M]. 武汉: 武汉大学出版社, 2008: 13.

⑤ 肖希明, 黄连庆. 以需求为导向的数字信息资源开发 [J]. 中国图书馆学报, 2007 (6): 65-68, 87.

资源的建设进行长远规划,并且建立图书馆数字资源共享系统。对图书馆数字资源的选择、组织、开发、布局,构成了图书馆数字资源建设的全过程,每个环节可相互转移、相互补充。

第三,数字资源虽然可以多次利用,但数字资源的生产、集成、传播和使用涉及多个权益主体,需要从法律、市场、技术等多个角度对各权益主体进行权利保护和利益平衡^①。图书馆数字资源面向的用户应该明白这一点,对于数字资源更应该珍惜使用。

第四,图书馆数字资源建设过程中,需要与数字资源开发商进行交涉,要灵活地谈判采购,不能随其漫天要价。在数字资源的服务过程中,要保持自己为用户服务的初衷,在适度商业化的同时又为用户提供优质且廉价的数字资源,这是图书馆数字资源建设过程中应该坚持的,因为它是图书馆在激烈的信息时代竞争中处于不败之地的立足点。同时还要发挥图书馆拥有高素质人才的优势,对数字资源建设进行规划,对数字资源进行深加工。一线的馆员与用户是图书馆数字资源的使用者,这些人的意见是数字资源建设中需要考虑的重要因素,必须搜集他们的意见进行改善。这样才能使图书馆数字资源建设更加人性化、科学化。

1.2.2 图书馆数字资源建设的内容

图书馆数字资源建设是文献数字化、数字资源整合、数字资源组织等活动的综合,其主要内容可概括如下。

(1) 图书馆数字资源的建设规划

过去,图书馆主要存储文献资源,在利用数字资源进行服务时,就不能再用以前收藏纸质图书文献那样的固有思维去设计,数字资源的建设规划要注意以下问题:①数字资源的存量更加丰富。单单百度的数据信息存量就是国家图书馆全部文献数字化后数据量的5万倍^②。毫无疑问,图书馆数字资源建设需要在数字资源质量和规范上下功夫,只有这样图书馆的数字资源才会更有竞争力。②数字资源的载体与传播工具更加多样化。数字资源在互联网时代得以迅速发展,离不开计算机、智能手机、电子书终端等电子产品日新月异的技术进步。图书馆数字资源建设规划时,应该注意对接这些终端,这样才是科学的规划。③数字资源的建设规划应以相关法律为依据。数字资源产业发展迅速,但是数据资源的盗版、侵权现象层出不穷。图书馆数字资源建设要搭建好法律的地基和平台,这样图书馆数字资源的高楼大厦才能更加稳固。④图书馆数字资源的规划应紧跟时代潮流。这需要与包括数据库商、出版商等其他机构合作。图书馆虽然也可以进行商业化操作,但是需要坚持为用户服务的出发点,不能太功利化。

(2) 图书馆数据库的建设

21世纪是知识经济和信息化的时代,数字资源的建设已成为评价高校图书馆信息资源建设的标准,也是数字图书馆建设的基础核心^③。其中,图书馆数据库的模块设计、数据搜集、数据加工是数据库建设过程中的重要环节。

①模块设计。在模块设计环节中,图书馆高素质人才的作用不可或缺,这个过程不仅需

① 吴志强,刘梅.数字资源权益分享的法律机制研究[J].图书情报知识,2011(6):93-100.

② 苏新宁.大数据时代数字图书馆面临的机遇和挑战[J].中国图书馆学报,2015(6):4-12.

③ 刘莹.我国高校图书馆特色数据库建设现状及发展策略研究[J].图书馆学研究,2008(7):36-38,47.

要对用户群进行调查研究,还需要具有专业高度的图书馆人才对数据库的建设进行设计规划,防止数据库的重复建设。数据库的模块设计是数据库建设中最重要的一环。设计规划过程中需要对数据存储的框架结构了然于胸,数据库资源检索的用户界面一定要操作简便,使得用户更容易检索、利用图书馆的数字资源。

②数据搜集。数据搜集对于图书馆来说较为容易,因为图书馆的资源集中而且质量高。但收集过程中仍需要对资源分类、筛选、集中和整理,及时更新资源,随时对资源进行查缺补漏,保证数据的完整性、权威性、即时性。

③数据加工。数据加工是对所收集的数据进行处理和加工,主要包括:第一,筛选,认真审核并筛选所收集的数据,删除那些不准确及价值不大的信息,确定可以收录进数据库的数据源;第二,数字化处理,文献的图像扫描、图像处理、识别转换、编改、校对等;第三,标引,选择合适的标引方式,制定标引细则,详细规定标引的深度、分类的集中与分散、主题词和关键词的选用规则等;第四,录入,制定严格的质量管理制度,确保输入的数据准确无误;第五,审核,全面、认真地进行审校,确保录入的每一条数据的准确性;第六,数据发布,使用数据库管理软件,做成便于用户浏览、检索的网页进行发布;第七,更新维护,收集图书馆数字资源用户使用过程中的反馈信息,及时对数据进行替换、删除、修改和整理^①。

(3) 图书馆网络数字资源的开发与组织

以前文献资源几乎是图书馆的全部(包括光盘、软盘、微缩胶卷等),但是信息化时代人们的信息诉求是方便快捷地得到想要的信息资源,这就需要图书馆将文献资源数字化,以最快最好的方式方法传递给用户。网络数字化时代,图书馆突破了传统的图书馆思维,无论是数字资源的开发与组织,还是图书馆的服务理念都应该有数字化思维。这要求每一个图书馆工作人员都应与时俱进,勇敢迎接数字时代的来临。

①跳出传统思维框架。图书馆现在面临的不仅有数字信息还有网上海量繁杂的信息。用户对数字资源的不同需求,促使图书馆对网络数字资源的采集、处理、组织过程发生了很大变化。为此必须重新构建图书馆的数字资源框架,即不限于数字资源,又要把整个网络资源纳入图书馆的资源框架。

②建立全方位的服务概念。传统图书馆的服务基本是文献性服务,提供的服务多为被动型的服务。图书馆的数字资源服务已经有了拓展,各类数据库大大丰富了图书馆的馆藏资源;图书馆的服务由被动式服务转向自助式服务,但是从数字图书馆的角度出发,这是不够的,还应建立全方位的服务理念。其一,面向全社会服务,图书馆建立的自有版权的数据库也应面向全社会服务;其二,提供各种类型的信息服务,例如,文献信息、网络信息、音频信息、视频信息等;其三,服务方式与传播方式也应进一步拓展,建立图书馆信息服务平台,拓展到移动设备的服务终端。

③拓展工作重点。在做好常规服务的同时,还需定位一些高端服务,努力把图书馆数字资源中心建设成信息分析策源地,密切关注学科前沿领域。同时对国家重大科研领域或本单

^① 马费成. 数字信息资源建设与服务研究 [M]. 武汉: 武汉大学出版社, 2008: 16.