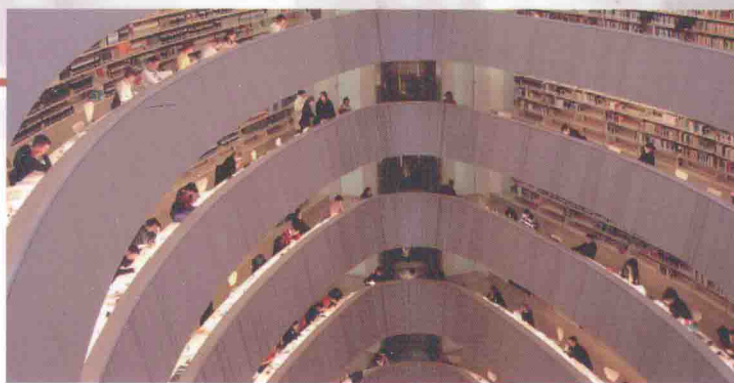




公共图书馆自建资源整合 研究与实例分析

梁蕙玮 © 著

GONGGONG TUSHUGUAN
ZIJIAN ZIYUAN ZHENGHE

YANJIU YU SHILI
FENXI



國家圖書館出版社
National Library of China Publishing House

公共图书馆自建资源整合 研究与实例分析

梁蕙玮 著



國家圖書館出版社
National Library of China Publishing House

图书在版编目(CIP)数据

公共图书馆自建资源整合研究与实例分析/梁蕙玮著. --北京:国家图书馆出版社,2014.7

ISBN 978-7-5013-5332-3

I. ①公… II. ①梁… III. ①公共图书馆—图书馆事业—研究
IV. ①G258.2

中国版本图书馆 CIP 数据核字(2014)第 097744 号

书 名 公共图书馆自建资源整合研究与实例分析
著 者 梁蕙玮 著
责任编辑 高爽 杨晓利

出 版 国家图书馆出版社(100034 北京市西城区文津街 7 号)
(原书目文献出版社—北京图书馆出版社)

发 行 010-66114536 66126153 66151313 66175620
66121706(传真),66126156(门市部)

E-mail btsfxb@nlc.gov.cn(邮购)

Website www.nlcpress.com →投稿中心

经 销 新华书店

印 装 北京佳顺印务有限公司

版 次 2014 年 7 月第 1 版 2014 年 7 月第 1 次印刷

开 本 710×1000(毫米) 1/16

印 张 12.5

字 数 150 千字

书 号 ISBN 978-7-5013-5332-3

定 价 48.00 元

前 言

数字图书馆的飞速发展,带来了图书馆自建资源建设的热潮。作为重要的文化信息服务机构,公共图书馆不仅要提供全方位、多方面的服务,更要提供所在地理区域内的特色服务,而自建资源正是提供这些特色服务的基础,因而各级公共图书馆都根据自身的情况建设了大量的自建资源。这些自建资源或是通过对馆藏资源数字化,或是通过对网络资源采集等方式获取,在内容上一般以反映本地区特色的自建资源居多,另外还有一些服务本地区经济的实用性资源。这些自建资源已经成为公共图书馆馆藏资源的重要组成部分,对弘扬地方文化传统,促进地区经济、文化发展意义重大。

尽管公共图书馆在自建资源建设方面取得了很大的成果,但还存在一些不足之处,如:整体上缺乏统筹协调,各馆独立开展资源建设,各自的加工系统、组织方式及著录标准都不一致,使得各数据库很难兼容;自建资源单独提供服务,没有很好地与图书馆的其他资源进行整合;宣传推广不够,基本上在各自图书馆网站上提供服务,用户知晓度不高;等等。因此,为进一步发挥图书馆自建资源的作用,对图书馆自建资源进行整合,已势在必行。

本书从公共图书馆自建资源现状出发,详细介绍国家图书馆自建资源情况以及各省公共图书馆自建资源情况,在此基础上开展对自建资源整合的探讨,对全面了解我国公共图书馆自建资源现状有一定的价值。

自建资源整合包括图书馆内部各种自建资源自身的整合、自建资源和商购资源的整合,以及图书馆界自建资源整合等多方面的整合,整合的技术、方式也多种多样。本书从工作实践出发,结合近十年资源整合的工作经验和实际案例,深入浅出地介绍自建资源整合的常用技术,包括元数据技术、链接技术、跨库检索技术和元数据仓储技术;以及图书馆资源整合的常用方式,包括网络信息导航、专题资源库、OPAC 整合及统一检索平台等。

在对自建资源整合进行阐述的同时,本书对主要的整合技术和整合方式都提供了实际案例分析,包括国家图书馆采用 Metalib 对自建资源的整合、数字图书馆推广工程元数据仓储的建设、国家图书馆网络导航的建设、国家图书馆政府信息公开信息专题库建设、国家图书馆 OPAC 整合项目以及国家图书馆

文津搜索系统建设等六个项目。这些项目笔者都曾参与,通过实际工作,更能体会到各种整合方式的优点与不足之处,书中所列的案例正是笔者亲身体会的总结。

本书得以完成要感谢各省级、副省级公共图书馆同仁的支持与协助;感谢国家图书馆各位领导的支持与帮助;感谢国家图书馆数字资源部整合组萨蕾、李荣艳、冯红娟、曲云鹏、封硕、叶升阳等同事的辛勤付出;感谢国家图书馆出版社编辑的细致审核。正是在各方的共同努力下,本书才得以顺利出版。

本书撰写过程中,参考大量中外文资料,在此向所有参考文献的作者表示诚挚的感谢!

本书以笔者工作实践为基础进行撰写,行文通俗易懂,可以作为自建资源整合的普及性读物。六个实例分析各有侧重,代表不同的整合方式,具有较强的实用性,适合从事自建资源整合的图书馆同仁参考。

由于研究能力与研究时间有限,因此对有些问题的研究还不够全面与深入。对于错误与疏漏之处,敬请各位专家学者及图书馆同仁不吝赐教。

梁蕙玮

2013 年 12 月

国家图书馆出版社简介

国家图书馆出版社,原名书目文献出版社,1979年成立,1996年更名为北京图书馆出版社,2008年改为现名。

本社是由文化部主管、国家图书馆主办的中央级出版社。2009年8月,在新闻出版总署首次经营性图书出版单位等级评估中被定为一级出版社,并被授予“全国百佳图书出版单位”称号。

建社30余年来,本社通过与各图书馆密切合作,形成了两大专业出版特色:一是整理影印中文古籍等各种稀见历史文献;二是编辑出版图书馆学、信息管理学著译作,以及各种书目、索引等中文工具书。此外还编辑出版各种文史著作和传统文化普及读物。

目 录

前 言	(1)
1 引言	(1)
2 自建资源整合概述	(3)
2.1 自建资源概述	(3)
2.2 资源整合概述	(18)
2.3 资源整合层次	(23)
3 整合的主要技术	(37)
3.1 元数据技术	(37)
3.2 链接技术	(44)
3.3 跨库检索技术	(48)
3.4 元数据仓储技术	(58)
4 整合的常见形式	(71)
4.1 网络信息导航	(73)
4.2 专题资源库	(90)
4.3 OPAC 整合	(128)
4.4 统一检索平台	(136)
5 总结	(149)
附录 省级公共图书馆自建资源列表	(153)

图 目 录

图 1	旧京戏报数据库	(10)
图 2	广西戏剧专题资源	(10)
图 3	古籍插图图像数据库	(11)
图 4	老商标老广告数据库	(12)
图 5	网上湖南	(13)
图 6	中国政府公开信息整合服务平台湖北省分站	(13)
图 7	南京图书馆的“奥运世界 中华圆梦”图像数据库	(14)
图 8	上海与世博库	(15)
图 9	首都图书馆的首图讲坛	(16)
图 10	湖北图书馆的“长江讲坛”	(16)
图 11	FRBR 第一组实体关系图	(27)
图 12	按照姓氏的整合	(29)
图 13	多类型文献的整合	(29)
图 14	谱系图	(30)
图 15	OpenURL 总体框架结构	(46)
图 16	检索结果	(47)
图 17	SFX 菜单	(48)
图 18	Subscription 标签页面	(54)
图 19	Presentation:Primary 标签页面	(54)
图 20	Categories 标签页面	(55)
图 21	自建资源在前台的显示结果	(58)
图 22	电子报刊导航系统功能结构	(84)
图 23	国家图书馆电子报刊导航	(85)
图 24	国家图书馆图书馆学信息门户	(86)
图 25	可视化展示效果	(90)
图 26	网站群总体架构	(123)
图 27	网络采集业务流程	(124)

图 目 录

图 28	政府公报显示页面	(126)
图 29	中国政府公开信息整合服务平台总站	(127)
图 30	中国政府公开信息整合服务平台触屏版	(128)
图 31	美国国会图书馆的详细显示页	(129)
图 32	印本图书和电子图书在 OPAC 中的展示	(136)
图 33	文津搜索首页	(146)
图 34	检索结果页	(147)

目 录

表 1	国家图书馆自建特色资源	(7)
表 2	DC 元素描述内容的类别和范围表	(41)
表 3	元数据仓储中国国家图书馆自建资源情况表	(63)
表 4	公共图书馆自建资源元数据登记情况	(65)
表 5	公共图书馆自建数据库登记情况	(66)
表 6	图书馆学信息门户的内容分类	(87)
表 7	图书馆学信息门户的元数据方案	(87)
表 8	元数据术语定义属性表	(95)
表 9	GILS 核心元素集	(108)
表 10	DC-Government 核心元素集	(110)
表 11	政府信息公开核心元数据	(112)
表 12	国家图书馆政府信息元数据标准	(114)
表 13	主题分类表	(115)
表 14	题材分类表	(117)

1 引言

随着信息技术的发展,公共图书馆越来越重视对馆藏资源、网络资源的开发与利用,建设了许多数字资源及资源库为用户提供服务。如:国家图书馆对其所收藏的部分珍贵古籍资源,包括甲骨、拓片、敦煌文献、西夏文献、地方志、老照片、手稿、宋人文集等,进行数字化为读者提供服务;首都图书馆从所藏古籍文献中遴选各类插图 1.4 万余幅制作了古籍插图图像数据库;上海图书馆的“上海年华”特色资源,以“图片上海”的形式发掘和整理了该馆馆藏地方文献中的图片资源,并进行数字化,全方位地向读者介绍了上海悠久的历史和文化以及上海改革开放以来取得的成就;天津图书馆建设了津门曲艺、天津民俗、天津地方志、天津抗战记事等一系列的地方特色资源等。总体来说,大型的省市级公共图书馆都已开展了该项工作。但由于这些资源或资源库都是各个图书馆基于自身需要而建设的,在标准规范、实现技术、使用方式等方面都存在着很大的差异,并且多数都仅在各自图书馆的网站上进行展示,也很难产生规模效应,知晓的用户也不多,因此如何对图书馆自建资源进行整合,以充分发挥其价值也成为图书馆研究的一个方向。

图书馆自建资源整合的目标是通过各种方式对各个图书馆自建资源进行科学有效的组织,使其能够形成一个新的有机整体,并为用户提供统一的检索服务,便于用户对资源的发现与获取。

对公共图书馆自建资源进行整合有着如下的意义:

有利于读者全面了解公共图书馆馆藏资源的整体面貌,提高资源的利用率。将公共图书馆的各种资源整合在一起,读者在一个界面下通过各种分类浏览的方式就可以了解各个图书馆的资源,有针对性地获取信息资源;同时又可根据关键词、题名以及学科分类等检索点对多个数据库进行检索来查询多个图书馆的资源。这样既节省了读者的检索时间,又提高了文献的查全率和查准率,形成一个知识一体化的知识服务网络,发挥文献资源的最佳效益。

有利于满足读者多元化的信息需求。图书馆现有各类资源由于其自身的局限性或者使用界面不统一等问题,不能充分满足读者的不同需要,提供的资源单一容易导致读者对信息的理解困难,使读者对信息的利用难度加

大。对公共图书馆界资源进行整合可以弥补这些缺陷,所提供的服务可以满足读者多元化、多层次的需求。

有利于促进公共图书馆信息服务平台的建设。公共图书馆在信息社会中所起的核心作用是信息的加工及服务,有效地对各种资源进行开发与整合是图书馆进行信息服务的基础。图书馆通过科学的手段和方法对信息资源进行整合与管理,可以准确地揭示信息资源并充分挖掘资源的内涵,从而全面地满足社会的信息需求。

有利于指导公共图书馆馆藏建设。有效地整合并揭示公共图书馆界的各类自建资源,可以使公共图书馆资源建设人员全面了解本馆及其他公共图书馆自建资源状况,及时调整资源建设策略,避免不必要的重复建设,尤其是对古文献等珍贵资源的重复数字化。

2 自建资源整合概述

2.1 自建资源概述

2.1.1 基本概念

自建资源是指图书馆依靠自己的馆藏文献与信息组织技能,通过数字化、网络资源采集、资源导航和专题资源库建设等途径建立和丰富数字馆藏^①;在图书情报界通常将图书馆建设的馆藏书目数据库、专题特色数据库和有效组织的网络资源统称为自建资源^②。

2.1.1.1 类型

图书馆自建资源的种类和形式多样,可以从不同的角度对自建资源进行划分:

按照目前图书馆常见资源类型可分为全文类资源,如电子图书、电子期刊、电子报纸、电子论文等;二次文献类资源,如题录、文摘、索引等;工具类资源,如电子百科全书、电子年鉴等;专题数据库资源,即以某个专题为主线,收集各类型资源制作的数据库。

按照自建资源的使用方式可划分为互联网访问资源,指用户可以通过互联网直接访问到的资源,包括互联网公开资源、通过专用账号或代理服务器的远程访问资源;局域网访问资源,指用户需要到图书馆内,通过图书馆的局域网访问的资源;本地光盘访问资源,即资源采用光盘存储方式,在使用时需通过光盘才能进行检索。随着网络环境的改善及信息技术的发展,光盘访问的方式已经不多见了。

按照自建资源的分布区域可划分为本地馆藏资源,即从地理位置上是图书馆的资源,资源是保存在图书馆服务器中,所有权归图书馆;虚拟馆藏资

① 申晓娟,周晨,韩超.《数字图书馆资源建设指南》解读[J].中国图书馆学报,2011(1):38-46.

② 魏大威.数字图书馆理论与实务[M].北京:国家图书馆出版社,2012:26.

源,是指资源本身不在图书馆内,可以通过计算机系统和通信设备获取的网络资源,虚拟馆藏资源一般经过图书馆工作人员的组织,便于用户的使用。

按照自建资源的建设方式,可以分为数字化资源,包括图书馆通过扫描、拍摄等方式对馆藏实体资源进行数字化加工所获得的馆藏数字化资源和通过摄像、录音、拍照等方式所获取的各种类型的原生数字资源;网络采集资源,指图书馆通过技术手段,对网络资源进行本地化处理所获得的资源。

2.1.1.2 特点

由于自建资源属于数字资源的一部分,因此自建资源具有数字资源的特点,如具有流动性、可重复使用等,还具有如下特点:

地域特色强。当今网络环境下,公共图书馆自建资源既不可能也不需要“大而全”,特色资源的建设已成为公共图书馆自建资源的主要亮点。在这种思想的指引下,各公共图书馆都很重视特色数据库的建设工作,如首都图书馆的旧京戏报数据库、首图讲坛资源库,上海图书馆的上海年华数据库,广西壮族自治区图书馆的广西戏剧专题资源库、广西民国照片库,辽宁省图书馆的《盛京时报》数据库,湖南省图书馆的湖南近代人物资源库、湖南地方戏剧数据库,福建省图书馆的闽南文化数据库、福建文化记忆库、客家文化库等。

资源经过选择与组织。公共图书馆的自建资源建设都是紧密结合图书馆的自身情况和用户需求来进行的,因此图书馆自建资源有严格的资源建设与评估标准,对所选取的资源有一定的质量要求,同时图书馆的自建资源一般需要制作元数据,从而进行科学的组织和控制,是一种高质量稳定的信息资源和知识资源,具有很高的学术价值。

便于检索与获取。图书馆的自建资源,尤其是各式各样的数据库,通常采用规则的操作方式,直观浅显、灵活多样、操作方便。可以按照学科分类逐级进行查询,也可以通过资源系统配备的检索引擎检索。用户可以从书名、篇名、作者、主题词、自由词、全文等多途径来进行检索,可以选择一般检索、高级检索、二次检索等功能进行布尔逻辑检索、相邻检索、截词检索或位置检索等,快速地获取所需信息,检索效率很高,节省时间和精力。

2.1.1.3 意义

对公共图书馆来说,开展自建资源建设与服务具有重要意义。

公共图书馆作为重要的文化信息服务机构,不仅要提供全方位、多方面的服务,更要提供所在地理区域内的特色服务。目前公共图书馆的资源主要是由商购资源和自建资源构成。商购资源使图书馆趋同性比较强,尤其在网

络时代,图书提供的数字资源若大体相同,则势必降低对读者的吸引力,图书馆的价值将大打折扣;而自建资源正是公共图书馆特色的体现,是践行《图书馆服务宣言》之“实现传播知识,传承文明的社会功能”的需要,也是提高公共图书馆知名度、美誉度的重要举措。

公共图书馆的读者需求是多样化和多层次的,大多数普通读者可能只是需要一般的信息资源,而一些特殊的(如研究型)读者则可能要求图书馆提供某个学科或某个范围内系统全面的信息资源。为了满足这部分读者需求,公共图书馆根据自己的条件建设特色资源,并开展特色资源服务,对整体提高用户的满意度有重要作用。

公共图书馆特色资源本身蕴含了其所在地区的文化气息和地方特色,公共图书馆建设特色资源,开展特色资源服务,对于该地区形成有特色的地方文化中心,弘扬地方文化传统,促进地区经济文化发展意义重大。

随着信息技术的飞速发展和数字资源建设力度的不断加强,自建资源越来越成为公共图书馆文献资源建设的重要组成部分,自建资源在保障读者的需求方面发挥着不可替代的作用。

2.1.2 公共图书馆自建资源现状

随着信息技术的发展,网络服务已经成为公共图书馆服务的重要方式,为避免大量外购资源给公共图书馆带来的趋同化,公共图书馆越来越重视自建资源的开发,建设了很多有价值的自建资源。

2.1.2.1 国家图书馆自建资源

国家图书馆一向比较重视资源的开发利用,截止到2012年年底,已有自建资源742.5TB,其中馆藏特色资源数字化694.9TB、全文数据2亿页,网络导航和网络资源采集30.7TB,数字资源征集12.4TB,电子资源呈缴4.5TB,主要包括电子图书、电子期刊、电子报纸、学位论文、音视频和古籍等各种类型的资源。

- 电子图书:包括中文图书、少儿图书、民国图书、少儿连环画等,国家图书馆从2000年开始,从馆藏图书中挑选部分优秀的图书进行数据化制成电子图书,目前已经有40万册图书,这些图书可以在网络上免费阅读前24页。

- 电子期刊:主要是民国期刊,为了有效地对民国文献进行保护与利用,国家图书馆早已完成馆藏民国期刊的缩微胶片制作,近年来又开展缩微胶片的数字扫描工作并制作成资源库。民国中文期刊资源库以书目数据、篇名数

据、数字对象为内容,提供简单检索、高级检索、二次检索、关联检索和条件限定检索。现提供 4351 种期刊电子影像的全文浏览。

- 电子报纸:包括报纸数字化,主要是对历史报纸数字化进行尝试,完成了对《益世报》缩微胶片的数字化,共 400 拍在网上提供服务。另外还有网络采集的报纸约 200 种。

- 学位论文:是以国家图书馆收藏博士论文为基础建设的学位论文全文影像数据。目前博士论文全文影像资源库以书目数据、篇名数据、数字对象为内容,提供 20 万余种博士论文全文前 24 页的展示浏览。

- 音视频:音视频资源是国家图书馆馆藏中的重要组成部分。自 1987 年以来音视频资源收藏已达 15.8 万余种,其中音频资料 5.7 万余种,视频资料 10.8 万余种。载体形式包括录音带、CD、MP3 等音频资料及录像带、LD、VCD、DVD 等视频资料,目前音视频资料每年还以 7000 至 8000 种的人藏速度不断丰富着馆藏。这些音视频资源内容涵盖政治、经济、文化、教育以及工业、农业、医学等各个学科领域,内有很多优秀的国内外影视资料及音乐资料为收藏珍品,可为来自不同领域的读者提供不同的音视频服务。为保存这些珍贵资料,国家图书馆开展了多年音视频资源的数字化工作,已完成数字化音频资源 515 889 首,数字化视频资源 53 305 小时。

- 国图讲座:国图讲座是国家图书馆面向社会,面向大众推出的双休日学术文化系列讲座。保存我国优秀古代文化典籍,培养中华文化传人,使文明薪火代代相传,发挥继续教育和社会教育的功能,是国家图书馆神圣而重要的使命。早在 20 世纪 50 年代,在文津街 7 号,一流学者的公益性学术讲座,启迪众多年轻学子,使他们走上学术研究道路。今天,身处全球信息化时代,古老的中华文明作为现代文明的源头越来越受到重视,也吸引了更多的人从中寻找现代文明发展的动力。以国图宏富的馆藏为基础,加之学术界的广泛支持,主讲人或为德高望重、岳峙渊渟的学界前辈,或为风华正茂、见解精到的学术中坚,国内外著名专家学者莅馆开讲,深入浅出地讲授他们毕生研究的精华。作为中华民族优秀文化典籍的收藏单位,国家图书馆面向公众推出精选的国图讲座视频资源,期望让更多的优秀文化为大众所共享。目前国图的在线讲座包括文津讲坛、文津沙龙、中国典籍与文化系列讲座以及一些其他的讲座,这些都是面向社会公众的。此外,国家图书馆还不定期的举办一些业务讲座,并将讲座的视频放在网上,供员工内部使用。

- 特色资源:国家图书馆根据馆藏情况,开发了诸如数字方志、碑帖菁

华、甲骨世界、西夏文献、年画撷英、前尘旧影、宋人文集等一系列特色资源,详见表1。

表1 国家图书馆自建特色资源

名称	内容
数字方志	地方志文献为我国所特有,也是国家图书馆独具特色的馆藏之一,所存文献数量与品质极高;数字方志采用数字图书馆方式,整理、加工编纂清代(含清代)以前的方志资源,将有利于保存、传播、研究、开发中华文化,推动数字资源建设
碑帖菁华	以国家图书馆藏有的历代甲骨、青铜器、石刻等类拓片23万余件为基础建设了馆藏石刻拓片数字化资源库,内容涉及历史、地理、政治、经济、军事、民族、民俗、文学、艺术、科技、建筑等方面,现有数据2万余条
甲骨世界	甲骨被誉为20世纪四大文献发现之一,集文献性、文物性、收藏性于一身,是研究我国商朝晚期不可多得的珍贵史料,甲骨世界包括甲骨元数据2964条,影像5932幅;甲骨拓片元数据2975条,影像3177幅
西夏文献	国家图书馆馆藏西夏文献大多为西夏、元代孤本,距今有七八百年的历史,与宋、元时期的善本属于同一时代;馆藏西夏文献的纸张、封面和题签的绢绸,分别是研究古代西夏、元代纸张、丝绸质料及制作的珍贵实物资料;馆藏西夏文献对研究西夏佛教史具有重要价值,馆藏西夏文献数字资源库内容包括馆藏西夏古籍书目数据124条,馆藏西夏古籍原件影像近5000拍;同时,还收集整理研究了西夏的论文千余篇,制作成西夏研究论文篇名数据1202条
年画撷英	国家图书馆收藏了4000余幅年画作品,从中精选出一部分制作了馆藏年画数字化资源库,并为每种年画编写了内容说明,重点介绍了该年画的故事梗概、历史背景、制作特色和相关知识
前尘旧影	收录了国家图书馆收藏的新旧照片7000余种、10万余张,真实生动地记录了过去的社会事件、历史人物、城乡面貌、名胜古迹和建筑服饰等,人们可从中解读出不同历史时期特定事物的形象特征和真实信息,具有重要的历史价值
宋人文集	据记载,宋人文集约有1000余种,流传至今尚有700余种,国家图书馆精选所藏宋人文集善本275部,通过缩微胶卷还原数字影像,并辅以详细书目建成善本全文影像数据库,免费呈献给公众利用