

图书馆特色数据库建设实践

——以首都图书馆历史图片数据库为例

丁蕊

(首都图书馆,北京 100021)

[摘要] 建设特色数据库是我国图书馆界的一项重要课题。本文分析了当前图书馆特色数据库建设的现状,并以“首都图书馆历史图片数据库”为例,探讨图书馆特色数据库建设的方法。

[关键词] 特色数据库;首都图书馆;老照片

[中图分类号] G250.74 [文献标识码] A [文章编号] 2095-5197(2016)06-0037-05

The Practice of the Library's Characteristic Database Construction: A Case Study of History Image Database of Capital Library of China

DING Rui

(Capital Library of China, Beijing 100021, China)

Abstract: The construction of characteristic database is an important subject in the libraries of our country. This paper analyzes the current situation of the construction of characteristic database in libraries of China, and probes into the methods for the construction of library's characteristic database taking the construction of "Capital library of China's history picture database" as an example.

Keywords: characteristic database; capital library of China; old photos

CLC number: G250.74

特色数据库是指图书馆或相关信息机构依托馆藏特色文献信息资源,针对特定用户的信息需求,对某一学科或某一专题的资源,进行收集、分析、筛选,再按照一定的标准和规范将其进行加工处理、标引、组织、存储,通过局域网或广域网进行传输和发布,提供检索和全文传递服务,以满足用户个性化需求的文献信息资源库^[1]。在“互联网+”大数据的背景下,自建特色数据库是展现图书馆数字化能力、科研能力的重要标志,也是不断提高图书馆的核心竞争力与服务水平的重要途径。如何建设特色数据库,是我国图书馆界一个十分重要的课题。

1 我国图书馆特色数据库建设所取得的成绩

20世纪90年代中后期,随着计算机技术、网络技术的不断发展成熟,我国图书馆的数字化建设取得了突飞猛进的发展,建设了一大批不同类型的特色数据库。据统计,我国95所211工程高校图书馆和部分省市级公共图书馆共有特色数据库795种,大体可分为地方特色数据库、学科特色数据库、专题数据库、商情数据库四大类;高校图书馆以学科特色数据库建设为主,公共图书馆以地方特色数据库建设为主^[2]。我国图书馆的特色数据库数量可观、内容丰富、形式多样,为科研工作者和广大读者提供了丰富的数字资源。同时,也为特色数据库的建设提供了经验和借鉴。

2 我国图书馆特色数据库建设存在的问题

2.1 缺乏总体规划、重复建设现象较为普遍

黑龙江、吉林、辽宁三个省的图书馆建立的以东北抗战为主题的数据库多达十几个,相同或相似选题的数据库内容上有不少重复之处,这不但造成了资源的浪费,还给读者利用增加了难度。

2.2 缺乏标准化和规范化

各个图书馆在建设特色数据库时都是自主选择工作平台,文献的著录、标引更是缺乏统一的标准,彼此之间很难实现资源的共享。

2.3 重建设,轻利用,读者体验较差

笔者以一个普通读者的身份,对全国33家省级公共图书馆的274个特色数据库进行了体验。除9家无自建资源外,将特色数据库放在图书馆主页明显位置的仅7家,占总数的21%;有的图书馆还存在特色数据库组织不当的现象,例如山东省图书馆的“齐鲁旧影”放在“山东省图书馆古籍珍本数据库”的类目之下,读者不悉心查找很难发现该项内容。在体验过的274个数据库中,可以互联网直接浏览的154个,需要馆内浏览的72个,无法浏览的40个,分别占总数的56%、26%、14%。从比例上看,供读者直接浏览的数据库占多数,但是大型精品数据库还是以馆内浏览为主。

3 图书馆特色数据库建设的实践

3.1 科学选题

公共图书馆特色数据库建设应遵循特色原则、保护原则、实用原则、发展原则^[3]。首都图书馆(以下简称我馆)结合读者需求和自身特色,经过反复调研,最终决定建设“首都图书馆历史图片数据库”。随着读图时代的到来,历史图片的史料价值越来越凸显。从实用性的角度来看,各行各业的读者对历史图片的需求量也越来越大。如:园林古建工作者可以在历史图片中了解到古建筑的原貌;影视工作者可以通过老照片老海报还原已经逝去的生活场景。因此,对图片进行整理和分编,建设图片数据库可以更好地满足读者的需求。我馆珍藏老照片万余幅,这些照片时间跨度大,从清末一直延续到20世纪50年代,其内容更是相当丰富:有辅仁大学和燕京大学校园风景以及师生学习生活和宗教活动的照片;有美国著名汉学家福开森先生在华的生活照片;有法文图书馆的资料照片;有北京、曲阜、青岛、杭州等地的风景民俗照片等等,其中不乏珍品,具有很高的利用价值。由于照片自身的材料局限和岁月的侵蚀,这批珍贵的史料出现了霉变、泛黄、粘连等现象。为了保护这些照片,将其数字化并建设数据库是一项非常必要的工程。

3.2 合理建设

特色数据库建设需要复合型的专业人才,既要精通计算机数字化知识,又要对数据资源有深刻的理解。建设历史图片数据库需要工作人员了解计算机网络、历史、摄影、信息管理等多方面的知识。在复合型人才缺乏的情况下,我馆采用多部门联合的方式,并借助专业公司的技术,进行数据库建设。

3.2.1 借助专业公司搭建数据库管理系统

数据库管理系统(Database Management System)用于建立、使用和维护数据库,其应具备功能完善、操作简单、应用便捷等特性。我馆采用外购管理系统的方式,选择与该领域技术先进、有丰富的成功案例的厂家合作,经过和供应商反复沟通、调试,最终完成了“首都图书馆历史图片管理系统”的搭建。该系统集数据的录入、管理、发布、维护于一体,具有录入、编辑、发布、标引、批量处理、简单检索、高级检索等功能,基本可以满足数据库建设和使用的需要。

3.2.2 图片数字化扫描

图片的数字化是建设图片数据库的前提,丰富的数字化资源是建设图片数据库的基础。我馆配置了先进的扫描设备,对图片进行扫描加工。馆藏中有一部分银盐胶片,甚至还有一定数量玻璃底片,还配置了专业的底片扫描仪,可以较好地再现底片中的内容。为了保证图片质量,均采用600dpi扫描,tiff格式保

存。某些价值较高,尺寸过小的照片适当调整 dpi 值。此外,还采用优质、大容量的存储设备,做好扫描数据的存储和备份工作。

3.2.3 依据国家标准对数据进行数据著录与标引

数字化照片资源建设是数据库的基础,文字注释则是数据库的灵魂。没有详细而准确的文字注释,数据库很难被合理利用。我馆参照《数字资源加工标准规范》和《中华人民共和国国家标准照片档案管理规范(GB/T11821-2002)》,并结合我馆照片的实际情况,设定每条数据的著录字段依次为:照片题名、照片编号、拍摄者、拍摄时间、拍摄地点、主题词、文字说明、图片尺寸、照片载体、照片数量、所属相册、馆藏地点、备注。参照《中国分类主题词表(第2版)》,根据照片反映的内容,为数据进行主题标引,先在主题词表中选择相对应的规范词,不能满足需求时,根据实际情况选择非控主题词。主题标引以全面、准确为原则,尽量展现照片的各种细节,完成标引之后的图片便具备了多项信息属性和特质,为读者检索提供多种途径。

3.2.4 以方便读者利用为中心的数据分类

目前,我国各大图书馆的很多特色数据库不具备分类浏览功能,不利于读者利用。我馆以方便读者为出发点,以突出主题、揭示图片特色为原则,依据本馆藏品的特色,并参照中图法,对图片进行了系统分类,大致如图1所示。

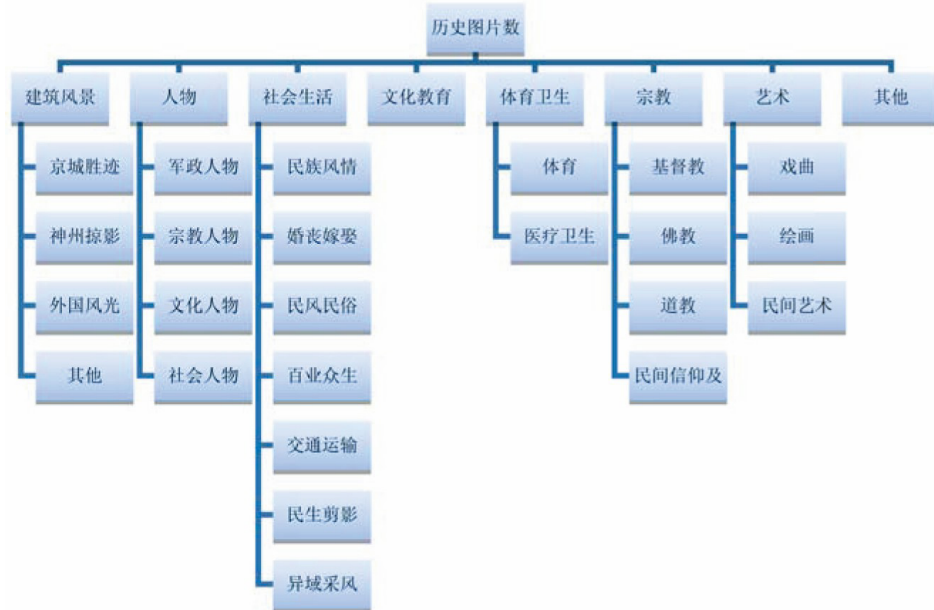


图1 历史图片数据库内容分类图

由于建筑风景类照片数量多,读者的关注度也高,所以又将建筑风景内细分为故宫、太庙、中南海、北海景山、颐和园、山东、杭州、福建、广州、香港、西北、长城等专题。

3.2.5 数据库的审校

审校工作是数据库建设必不可少的环节。工作人员通过自校、互校,确保每条数据准确无误。与文本数据库比较,图片数据库的审校应注意的内容更多,包括图片是否可以正常显示,题名与图片是否相符,题名、拍摄者等字段著录和分类是否准确,照片鉴定有无重大失误等。

3.2.6 数据库的发布

数据库最终的服务对象是广大读者,如何将数据库展示在读者面前是数据库建设的一道重要工序。我馆始终坚持“读者至上”的原则,以首都图书馆历史图片管理系统的发布平台为基础,通过精心设计,最

终使数据库以美观大方、操作便捷的 Web 发布方式为用户提供服务。读者可以进行分类专题浏览,也可以通过照片的题名、拍摄者、主题词等途径进行简单检索和高级检索。

4 建设特色数据库的同时加强实体文献保护

实体文献是数据库的基础,建设特色数据库的同时也要加强对文献的保护。我馆加强了老照片的保护与整理,主要工作内容包括以下几个方面。

4.1 将照片放置于恒温恒湿库房

光线、温度、湿度是直接作用于照片最普遍、最明显的因素。光线中含有紫外线,会造成照片变黄、褪色。高温和潮湿会使照片滋生霉菌、微生物,导致照片发生物理化学变化。因此必须将照片放置在阴凉、干燥、通风的环境里。我馆书库温度保持在 16°-20℃之间,湿度保持在 50%-60%之间,并配有空气净化器装置,24 小时恒温,较适合照片的存放。

4.2 摸清家底

面对纷繁杂乱、大小不一、风格迥异、破损程度不同、未经过系统整理的老照片,很难制定出合理的整理方案。为此,先要摸清家底,在对老照片的数量、基本内容、珍稀程度、破损状况等有一个总体的认识的基础上,才能制定出比较合理的方案。

4.3 制作半透明无酸纸装具

我馆的老照片长期以来未经过系统整理,大部分没有独立包装而直接暴露在空气中。空气中的硫化物、氮氧化物、过氧化物、氯化氢等有害气体对照片档案有强烈破坏作用。空气中某些具有活性的灰尘颗粒,是霉菌孢子的传播物及滋生体,可以催化照片老化降解。少量照片有原始照片袋包装,但由于年代久远,老照片袋含有硫化物、酸性物质、色素或者填充剂会加速照片的老化。制作自身不会释放有害物质的新式装具,对照片进行独立包装,使照片和空气隔离,是保护工作的一项重要任务。在查阅国内外相关文献和借鉴兄弟馆的研究成果的基础上,我馆选用半透明无酸纸制作照片袋,对照片进行包装保护。无酸纸是指原料纸浆为 100%木浆,不含木质素等杂质,在造纸过程中不使用酸性施胶剂(硫酸铝等),不含游离酸,PH 值为 7(中性)或大于 7(偏碱性)的纸。用其制作的装具自身材料不会对照片产生不良影响或损害,同时还能有效地防止照片发生老化、酸化等现象。选用半透明材质是为了查找方便,不用开袋即可看清照片的概况,减少不必要的抽取。

4.4 除尘与包装

由于长期暴露在空气中,照片表面灰尘覆盖严重,灰尘颗粒不仅影响照片质量,还会加速照片老化。除尘是照片保护工作必不可少的环节。工作人员选用软硬适中的文物专用毛刷,耐心细致地进行除尘工作。除尘完毕的照片放入半透明无酸纸的照片袋中;完全相同的照片放入同一袋中,中间用无酸纸隔开。

4.5 老照片的编号与登录

老照片的分类比较复杂,在登录工作阶段,我们暂时不把照片分类。每张照片按照登录的顺序给一个流水号,作为唯一标示。登录内容包括照片题名、照片编号、拍摄者、拍摄时间、拍摄地点、尺寸大小、照片载体、照片数量、所属相册、馆藏地点等。

5 老照片数据化需要注意的特殊问题

老照片是某一事件、某一人物、某一地点景物或某种时尚鲜活而生动的记录,可以将活生生的人物面貌和丰富多彩的历史场景直接展示在读者面前。一张照片所昭示的信息量是文字史料难以全部描绘的。然而由于年代久远,时过境迁,广大读者不一定能很好地解读照片所包含的信息。如果对其所展现的内容考证不清,照片的价值会大打折扣。因此照片的考证与鉴定是数据库建设中十分重要的环节。经过长时间的工作积累,我们采用的考证方法大致有以下几点。

5.1 深度发掘照片细节

照片是当时情景的生动再现,一张照片包含着丰富的信息量。但是,照片中的某些细节需要悉心观察,合理解读,才能发现照片背后的信息。因此,工作人员需要深度发掘照片中的细节,考证照片的内容、背景等要素。例如,有一张照片展现的是民国时期的葬礼场景,初看没有特别之处。细心观察会看到铭旌亭上有“□威上将军吴公”字样。结合相关史料可以认定照片记录的是“孚威上将军吴公”,即吴佩孚将军的葬礼。这张照片拍摄的时间、地点等要素便可以准确著录。一张看似普通的民俗风情照片,经过工作人员的考证变身为具有史料价值的历史照片。

5.2 利用照片及附件中的文字说明

我馆老照片中,部分照片背面有说明文字,部分照片还有信封包装,信封上也可能有文字,我们可以根据这些信息对照片进行著录。这些藏品大部分为外国人拍摄,文字多为英语、德语,且都是很随意的手写体,给辨认工作增添了很大难度。我馆专门挑选了外语基础好,并具有多年西文文献整理经验的工作人员负责这项工作,尽可能准确无误地翻译这些文字,为照片著录提供可靠的信息。

5.3 借助相关历史资料

照片鉴定对工作人员的要求很高,不仅需要鉴定者熟知相关的历史背景,还要求具备摄影、地理、民俗、宗教等多方面的知识。但是工作人员的知识水平有限,需要借助相关资料。例如,我馆老照片中有一部分辅仁大学师生的人物照。我们借助《辅仁大学年刊》《北京辅仁大学校史》等资料,辨认出陈垣、胡适、启功、英千里、余嘉锡、张星烺等名人照片。

5.4 合理利用网络资源

当今社会是信息化、网络化的社会,有大量网络资源可以利用。某些搜索引擎具有图片搜索功能,可以利用该功能搜索和本馆照片类似的资源。例如,我馆藏有一张风景照,反映的内容是一座古塔和周边的民居。起初工作人员无从下手,用某搜索引擎进行图片搜索,可以发现相似度很高的照片,其中英军随军摄影师费利斯·比托、美国摄影家西德尼·甘博等人所拍摄的北京通州燃灯塔和周边民居与我馆藏品有90%以上的相似度,由此可以推断,我馆这张风景照也是通州古迹——燃灯塔的旧照。

6 历史图片数据库建设的思考与展望

受工作条件和工作人员水平的局限,我馆历史图片数据库建设还有很多需要改进的地方。下一步需要开展的工作主要有以下几点:(1)加强老照片的保护,设立照片专柜。我馆老照片虽有无酸纸袋包装,但仍存放在普通纸箱当中,下一步需要购置适合老照片的专门保存柜,进行单独存放。(2)修复破损老照片。老照片中有一部分破损比较严重,由于我馆缺乏照片修复的专业人员,本着“不遇良工、宁存故物”的原则而未开展照片修复,待条件成熟之后,照片修复是一项重要工作。(3)充实数据库资源。

目前我馆数据库只收录了集中存放的老照片,还有为数不少的老照片分散在民国文献当中,下一步将组织人力对民国文献中的老照片进行开发整理。

[参考文献]

- [1] 刘葵波.特色数据库及其相关概念辨析[J].图书馆建设,2015(4):14-17.
- [2] 李晨晖.图书馆特色数据库研究与开发[J].图书情报导刊,2016(5):85-87.
- [3] 赫青.公共图书馆自建数字资源选题研究[J].图书馆学研究,2012(8):51-54.

[作者简介] 丁蕊(1980-),女,馆员,硕士,研究方向:民国时期文献的保护与利用。

[收稿日期] 2016-09-28

(编发:王丽)