

● 图书情报学

档案馆网站信息资源的组织模式

梁 广 寒

(武汉大学 信息管理学院, 湖北 武汉 430072)

[作者简介] 梁广寒(1974-), 女, 广东佛山人, 武汉大学信息管理学院档案学系博士生, 主要从事数字信息管理、电子文件管理研究。

[摘 要] 目前, 我国档案信息网站在资源建设方面存在着信息内容匮乏、信息类型与级别单一、特色性不足等问题。解决这些问题需要建立以档案馆网站为中心、多方联合协作、档案目录、原文与编研成果并重的档案信息资源组织模式, 该模式以网站为窗口, 以档案信息为根本, 以多种信息类型为表现方式, 以外包为实施手段, 以联合协作为方向, 在此过程中宜采取合作开发、外包商选择及数据库用户选择等开发档案信息数据库的实施手段。

[关键词] 档案馆网站; 数字资源; 组织模式; 内容建设

[中图分类号] G271 [文献标识码] A [文章编号] 1671-881X(2004)02-0216-05

2002 年 11—12 月间, 通过对我国档案信息网站的内容建设所做的调查, 我们发现, 目前我国档案信息网站存在信息内容匮乏、信息类型与级别单一、特色性不足等问题。针对此现状, 我们认为档案信息网站的资源建设, 应建立以档案馆网站为中心、多方联合协作、档案目录、原文与编研成果并重的组织模式; 这种模式, 要以网站为窗口, 以档案信息为根本, 以多种信息类型为表现方式, 以外包为实施手段, 以联合协作为方向。

(一) 以档案馆网站为中心, 多方联合协作

1. 行业纵向联合, 加强网站基础设施建设。

据统计, 截至 1998 年底, 我国各级各类档案馆已达 3700 多个。而通过各种渠道得到的档案馆网址(包括无法链接的网站)只有 200 多个, 这说明我国档案馆网站内容建设的首要问题是网站的基础设施建设问题。没有网站, 档案馆就失去了在数字世界存在的机会, 传播档案信息更无从谈起。另一方面, 从档案馆网站相互关系来看, 我国档案馆网站, 既缺乏网站之间的推介宣传, 也缺乏对行业网络的主动利用。实际上, 我国目前已完全具备网络基础, 网络构架已经形成, 综合性的通信主干网络已经构筑, 档案馆建设网站已经有了充实的现实条件。因此, 档案馆网站的建设可以在行业网络和九大互联网络的基础上, 分以下步骤进行:

首先是建立各地区的档案信息管理内部局域网, 以全面管理内外部数据库; 其次是在外网建立档案信息传递和共享的开放的数据库, 通过与其他文献信息系统、政府机构网站的联合实现信息共享; 三是以国家档案馆为龙头, 在档案馆行业网络内建立起各种综合性、专门性、包括目录数据库在内的档案信息数据库, 进而组成跨系统的国家档案信息网络。系统之间数据库相对独立, 纵向上则通过网络相联, 形成一个联合体, 组成分布式的信息总和; 最后也是最重要的, 就是与公众网相联, 作为一个或多个独立的网站, 使公布的档案信息进入公众网, 向国内外用户提供信息服务。

2. 与其他文化信息部门横向协作, 加强网站内容建设。

档案馆也是社会文化信息机构, 与其他文化信息机构共同承担着保护和传播文化遗产的职责。但实际上, 档案馆网站仿佛网络上的穷乡僻壤, 既没有多少“本土”资源, 也没有积极架桥铺路, 引入其他信息机构网站资源。另一方面, 在情报界如火如荼地讨论网络信息资源共建共享, 图书馆界、情报界、文化界、企业界等各界人士对数字图书馆的研究和实践方兴未艾的今天, 又有多少人将目光投向了同样珍藏着重要、珍贵、数目浩繁的文化资源的档案馆呢?

面对知识经济和数字技术的种种挑战和机遇, 处于落后态势的档案馆必须丢弃“深谷幽兰”的心态, 而要充分利用后

发优势,主动与其他文化信息部门进行横向协作,加强网站的内容建设,这样才能达到自己的目标:一是使档案馆真正进入信息界和文化产业的核心地带,作为信息机构的主力之一,与其他文化信息机构共同进行网络信息资源的共建共享;二是使在现实世界分散保存于各种文化信息机构的信息资源,通过网络得以集中和整合;三是使档案馆网站与一切媒体真正融合,共同传扬中华文化。

文化精神往往通过具体的载体表现出来。由于各种原因,我国历史文化遗产保存在多种文化信息机构,如档案馆、图书馆、博物馆、文化馆、纪念馆、艺术馆、文史馆、考古研究所、历史研究所、文学艺术馆、中医研究所、科学院、藏书楼、电视台、广播电台、电影制片厂等。尽管收藏方针各不相同,它们的藏品仍有交叉。档案馆主要收藏在社会活动中自然形成的、以文字、图像、声频、视频为信息记录符号的社会记录和文化产物。这类记录和产物中的部分有时也会成为其他机构的馆藏。例如,古籍、方志、名人手迹手稿、照片、实物档案等。而档案馆也会收藏一些本属于其他信息机构藏品范围的文化遗产。档案馆与各类文化信息机构协作,整理出符合网站建设原则的信息目录和内容,上载网站提供利用。这一方面有利于分散在各种文化信息机构中的档案信息资源得以集中,另一方面有利于各信息机构保存的各类历史文化遗产通过网络得以整合,形成规模效应,共同丰富网络中文信息资源,而传媒汇流在打破传统媒介之间的坚冰的同时,也打破了前网络时代信息单向传播的文化生态,这是一个具有洗礼意义的断裂点,也是人类生存经验的转折点。麦克卢汉说:媒介即讯息,过去的一切媒介都是网络的内容^[1](第53页)。人们将希望通过网络这一统一载体得到集成信息,而无需考虑其原始形式和保存地点,并在获得信息的同时创造信息。这引发了文化产业大规模的“媒体转移”和资源整合。档案馆通过与上述各信息部门协作,能够充分利用网络多媒体平台优势,与一切媒体真正融合,向世界传扬中华文化。

(二)档案目录、原文与编研成果三级信息并重

档案目录、原文、编研成果三个级别的信息可以满足不同用户不同层次的利用需求。目录不仅对用户起着指示和导航的作用,而且它本身内含的信息就是对馆藏内容和效用的揭示。没有经过加工的档案原文,通过数字化后以全文形式直接上网,适合研究型、凭证型等要求严谨的用户需求。信息编研则是经过加工和整理的档案信息以各种专题发布,具有原文发布不可替代的优点。它能根据用户的不同需求提供经过整合的信息,更全面、更方便、更有针对性,是一种智能型的信息发布方式,从而满足网络用户集成性、综合性、休闲性等利用特点,在很大程度上实现了信息共享;更能较好地解决公开与保密的矛盾;当原件尚未解密,全文利用尚需限制,但其中一部分信息可以公开时,可以经过加工,提取能够公布的一部分信息,或以摘要形式揭示其内容,吸引用户到馆查阅;同时,它还有利于以各种专题宣传我国博大精深的传统文化,使用户在欣赏中得到传统文化教育。

网站建设是长期而浩大的工程。正如武汉市档案馆负责人所估计:将该馆馆藏全部数字化至少需要80年。可见,目录、全文、编研成果虽然在网站建设中举足轻重,但同时进行却是不现实的。为此,这三个级别的内容建设可分三个步骤进行:

第一步,档案目录上网。包括全宗级目录、案卷级目录、文件级目录、本馆编研成果目录、缩微目录、联合目录,以初步揭示馆藏,并吸引用户到馆查阅。

第二步,各专题编研成果上网。由于编研成果在详略程度、形式、内容等方面具有非常大的灵活性,档案馆可以根据网站建设的原则,结合本馆实际情况选择编研的主题、类型和程度。相对来说,这是一种能在较短时间内揭示馆藏概况,从而吸引普通冲浪者成为问答用户的途径。

第三步,开放的档案原文上网。根据档案的价值、用户需求、以往利用率、载体状况和技术可能,优先选择一部分馆藏档案,进行数字化处理后将原文上网。

从调查结果来看,目前我国已有部分档案馆网站提供了目录信息和编研信息,但提供全文信息的极少。网站应加快全文信息的提供利用。

这三个级别的档案信息在网站上的提供利用上,都应采取数据库的方式,与目录信息、编研成果信息、全文信息相对应,分别建立目录数据库、事实数据库和全文数据库。数据库是存放档案信息资源的仓库,是档案信息网的核心,只有建立数据库,档案信息才能真正和网络接口,以实现更大范围内的档案信息资源的共享。目前,我国许多档案馆都已经具备了建设数据库,尤其是目录数据库的现实基础:根据对28家馆藏30万卷以上的综合性档案馆的调查结果统计(河北省档案馆、江苏省档案馆、上海市档案馆、南京市档案馆、浙江省档案馆、四川省档案馆、重庆市档案馆、湖北省档案馆等),正在使用的检索工具共有25种,其中最常用的有全宗目录、案卷目录、全引目录、专题目录、人名索引、照片档案目录和声像档案目录。据1999年统计,全国各级综合档案馆已有开放目录615万条^[2](第12-13页),同时也出现了大型联合目录,如全国历史档案资料目录中心下属的明清档案资料目录中心和民国档案资料目录中心编制而成的《全国明清档案全宗通览》《全国民国档案通览》^[3](第37-38页)。这些都为建设目录数据库打下了良好的基础。目前部分网站已

经建成目录数据库并投入使用,如北京市档案馆网站、杭州萧山区档案馆网站等,而事实数据库和全文数据库则亟待发展。

(三)永久保存级、服务级、镜像级、链接级四级资源共存

为科学地提供数字化档案信息,合理配置网站信息资源,使网站效益达到最佳状态,有必要将档案馆网站提供的各种信息资源划分为四个级别:永久保存级、服务级、镜像级和链接级。

永久保存级:这是档案馆馆藏的特色内容,具有惟一性,仅仅定位于本馆网站,其他档案馆网站不可得到,本馆对其负有永久保存责任。

服务级:它是用户必须的,能丰富网站的资源,但不属于本馆馆藏范围,不具有所有权,也不负有永久保存责任。

镜像级:对网站资源有支持和补充作用,但不是本馆馆藏,而是对保存在其他文化信息机构的档案信息、或与档案馆收藏范围有交叉的数字化信息拷贝在本站的拷贝。这是目前应主要考虑的。

链接级:则是网站资源的补充,可在网站显示分散于多种信息机构的信息资源地址。

根据资源级别及来源的不同,可将资源分为限制使用级、部分限制使用级和公开使用级:见表 1。

表 1 档案馆网站内容资源级别结构表

级别	内 容	使用范围
永久保存级	档案全文、照片档案、声像档案、多媒体档案	部分限制
	本馆馆藏档案目录、编研成果目录	公开
	本馆编研成果	部分限制
服务级	联合目录	公开
	公布档案史料为主的期刊	部分限制
	本地档案协会主办的档案期刊	部分限制
	馆藏资料	部分限制
镜像级	其他档案馆或信息机构开发的、对本馆有重要意义的专题数据库;古籍数据库、地方文献数据库等	限制
链接级	各类信息文献机构网站,包括图书馆、博物馆、纪念馆、资料馆、文献中心、方志办等机构的网站	公开
	各级地方政府信息网	公开
	各种为公众提供生活、学习、娱乐资料的专门网站,以及其他各种专门技术网、行业信息网等	公开
	国外优秀档案馆、档案协会网站	公开

(四)多档案种类、多信息类型并举

我国档案馆网站目前提供利用的档案信息种类和类型均过于单一,没有充分利用网络多媒体技术平台,也不利于全面真实地再现档案原貌。由于现代信息交流的不断发 展,用户在使用网络资源时,趋向于利用丰富多样的多媒体资源。而现代信息载体的更新,配套网络信息载体技术的发展、带宽的不断扩大,使网络今天已经真正成为“一切媒体的媒体”,为提供音频、视频、多媒体等多种类型信息、满足用户多元需求提供了可能。档案馆网站利用文本、静态图像、动态图像、音频、视频、多媒体等多种信息类型,提供文书档案、照片档案、声像档案、图纸档案、多媒体档案及其他专门档案信息势在必行。

原生性数字档案信息应当以其原有类型提供利用,再生性数字档案信息则应根据原件表达信息的方式(如文字、图像、声音、画面等)以及传承信息的特殊需求选择合适的类型。对于一些较复杂的对象,需要使用多媒体技术才能够全面、保真地传播原信息。如濒临失传的特殊语言如女书等,一方面需要将其原文文献数字化,同时以数字摄影摄像技术记录、保存文献的外观,另一方面更需要以音频文件记录保存以女书诵读、传唱的诗歌和歌曲,以视频文件记录和保存女书传人的最后一位在世者颂唱女书的音容笑貌。利用文本、图像、多媒体等多种技术才能全面再现这种珍贵文字的字体、语音、语调、使用这种语言的主人公及其文化现象,才能全面弘扬其文化内核,通过网络扩大其维系存在的可能性,从而达到传承中国传统文化、持续进行爱国主义教育的目的。

(五)档案信息数据库开发的组织

档案信息数据库的开发,可由三方合作完成:档案馆、外包商及数据库主要用户。在合作关系中,档案馆负责提供启动资金、提供并筛选信息、质量把关,拥有监督权、管理权和数据库的所有权;外包商负责建库业务和设备提供;用户负责资金的提供。通过这样的合作关系,将档案信息数据库的所有权、管理权和使用权分开。

1. 合作开发的原因。建议档案数据库开发采用这样的合作开发方式,主要是出于如下考虑:

一是在知识经济时代,一向处于经济边缘地带的档案馆也进入了产业化领域,这意味着档案信息可以以文化产品的方式发布。采用这样的合作关系,有利于将部分档案信息数据库作为商业数据库,有利于促进档案信息社会化和部分产品化,以弥补档案馆网站建设和数据库开发中的资金不足。

二是档案馆网站的建设,从数字化过程,到制作数据库,再到最后上网提供利用,都需要大量的、持续的资金支持。日本失败的例子证明,完全政府拨款模式是行不通,也是不可能的,必须走合作开发的道路。

三是国内外相关调查表明,在网络无数的站点上,人们访问的大多数大型综合网站,信息主要来源于各国权利组织的大型数据库。完全凭档案馆自身力量显然难以达到上述要求。走合作开发的道路,有利于档案馆选择资金、技术强大的组织合作进行数据库开发与建设,最终有利于网络档案信息资源的建设。

四是由用户投资,特别有利于避免使档案信息数据库成为“死库”或“空库”。据有关资料统计,长期以来我国政府部门掌握社会信息资源80%的有价值的信息和多个数据库,但大部分都是“死库”;全国现有数据库近1200种,占世界总数的10%,而容量仅为1%,其中可用的300多个大型数据库中,对外服务的仅占5%,能够提供商品化服务的数据库更少。由用户投资开发数据库,首先可以避免由于后继资金不足而导致数据库建设搁浅成为空库,其次可以避免由于不了解用户需求而导致数据库内容专指度不高、缺乏用户而成为死库。

2. 外包商的选择。外包是指把本应由内部操作的大量事务性和部分技术性工作委托给外部企业或团体来完成。它的意义在于基本上不涉及管理权问题,能够调动一切社会积极因素,使各种社会资源合理配置和有效利用,实现科学化、专业化管理的目标;同时,它能降低成本、提高效率和专业化水平。

由于档案信息的特殊性,档案馆对建库外包商的选择必须十分严格。笔者认为,首先应该以国内外外包商为前提,其次还应满足下列条件:政治上安全可靠,诚信度高;有稳定的、高素质的专业队伍;市场渗透力强,与档案信息数据库的目标用户有较密切的关系;在用户中有较高声誉和美誉。具备这些条件的外包商有利于保证数据库的质量、促进数据库的市场化,更重要的是有利于保证我国民族文化的纯粹性。

在网络时代,我国面临的不仅仅是市场竞争,更严峻的是文化生存问题。许多国家已经开始与中国相关部门和单位合作开发数字资源。例如2000年5月,敦煌地区文物管理部门与美国梅隆基金会签订了协议,建立“数字化虚拟洞窟”。2000年末,故宫博物院和一家日本企业签订了共同进行“故宫文化遗产数字化应用研究”的合作协议书。但这些尝试中存在着的严重问题是:第一,依赖于别国加工的文化产品还是我国自己的文化产品吗?第二,由别国进行数字转移的文化遗产,我国就失去了解释权,遗产的级别含义发生变异,那么,它还是我国自己的文化资源吗?它还是我国原有意义的文化遗产吗?第三,我国是具有巨大潜力的文化产品消费市场,但若若干年后,我国消费的是原汁原味的本国文化产品还是以中国文化为廉价资源、经过外国加工又以高价返销中国的四不像的“舶来货”?第四,作为网站资源主要支撑的数据库,如果成为我国无法完全掌握所有权和管理权的“合资产品”,那么网站代表的究竟是中国文化机构还是外资企业?

当前,中外合作的项目已经产生的一些现象足以为上述问题加注:在“国际敦煌项目网”网站中,共有9个板块,其中“出版物”、“专题”、“链接”和“地图检索”四个板块和“收藏地”中相当重要的内容“查找原件方法”完全是英文,甚至没有中文选择。当然,由于敦煌项目在国际学术界的地位和敦煌文化在世界文明史中的地位,网站以英文为主要语言有其国际化的积极一面,但是接近一半的重要板块却没有敦煌文化的故乡——中国的文字,这又意味着什么?!

有鉴于此,为保证中华传统文化的纯粹性,建设档案信息数据库应该选择技术强大、队伍稳定、资金雄厚、诚信度高、口碑佳的国内机构为合作的外包商。依此条件可选择对象有:国内高等院校或专业研究机构的相关部门或所办实体,也可选择专业性强的国内IT企业。从长远来看,中国企业家将肩负起自身抢占市场和为民族捍卫数字中国的双重责任。

3. 数据库用户的选择。档案信息数据库的用户可以分为四种类型:个人、企事业单位、政府机构和商业网站。一般来说,除个人用户外,其他三种都可考虑作为合作对象,其提供的资金可作为使用数据库的预付资费。档案馆则可以依照两种思路选择数据库用户作为合作伙伴:

一条是确定数据库专题,然后根据专题选择相应的目标用户。确定数据库专题后,考虑该专题所属的对口行业,以及与所属行业有业务交叉关系的相关行业和机构。例如,地理档案信息数据库,可寻找建委、房产局、房地产教义中心、房地产开发商、规划局、国土局、旅游局等机构作为合作伙伴;人口档案信息数据库可寻找保险公司、商业银行、出入境管理局、社保局、劳动人事机构等作为合作对象;药品档案信息数据库可寻找制药厂、医药公司、医院、专门网站作为合作对象;车辆档案信息可以寻找各级车辆交易中心、商业网站等作为合作对象,等等。

另一条是先寻找合作对象,然后根据其工作需要或经营范围确定数据库专题。对于有着源源不绝的信息流的档案馆来说,这并不困难,况且同一专题的数据库往往能满足来自该专题所属行业内外的多用户的不同需求,而不同用户的

信息需求又往往有交叉之处。例如我国的巨灵证券信息网(www.genius.com.cn)提供上市公司、基金、债券、金融信息、国民经济统计、央行统计、券商交易数据、行情系统、财经新闻和法律法规等 10 大类数据库,而这些数据库的内容多来源于档案信息。类似的专业网站和商业网站都可作为档案馆开发数据库的合作对象,针对它们的经营范围开发专题数据库。专业网站和商业网站参与开发的数据库,出于利益原因,必然拥有大量稳定的客户源。又如政府机构在决策、管理中需要参考大量各类文件、社会、经济、教育、医疗卫生等专题的档案信息,这些专题数据库中的内容也为其他许多机构所需求。因此,不必顾虑根据某一用户需求而开发的专题数据库的利用率问题。

在数字时代,数字信息将成为信息存在的主流。数字技术和数字信息带给档案馆的严酷现实是:如果档案馆网站资源匮乏,不能在“内容之战”中赢得众多用户,将使我国一部分珍贵的历史文化遗产继续蒙受冷遇,阻碍中华优秀传统文化的数字勘探工程。采用合理有效的模式进行网站内容建设、吸引公众,弘扬优秀历史文化遗产,是档案馆在数字时代的必然选择。

[参 考 文 献]

- [1] 保罗·莱文森. 数字麦克卢汉——信息化新纪元指南[M]. 何道宽译. 北京: 社会科学文献出版社, 2001.
- [2] 王绍忠. 蓬勃发展的档案馆事业[J]. 档案学通讯, 1999 (5).
- [3] 张建平, 姚红. 全国历史档案资料目录中心的建立与发展[J]. 中国档案, 2000, (6).
- [4] 谢泽明. 网络社会学[M]. 北京: 中国时代经济出版社, 2002.
- [5] 刘家真. 数字图书馆的馆藏级别与用户交互类型[J]. 图书情报知识, 1999 (2).
- [6] 刘家真. 馆藏文献优先数字化的选题范围思考[J]. 图书馆杂志, 2002 (10).

(责任编辑 涂文迁)

Model for Organizing Information Resource of Archival Websites

LIANG Guang-han

(School of Information Management, Wuhan University, Wuhan 430072, Hubei, China)

Biography: LIANG Guang-han (1974-), female, Doctoral candidate, School of Information Management, Wuhan University, majoring in digital information management.

Abstract: The author argues that there are some weak points in the content construction of archival website in China according to a previous investigation she did. She then provides a model for organizing the information resource of archival websites in China and for developing archival information database also. During the constructing, the archival websites should associate with other cultural websites and recognize both catalog and the original text, and the contractors are suggested.

Key words: archival website; digital resource; organizing model; content instruction