

# 大数据环境下数字图书馆的基本职能

王 锰 郑建明 陈 雅 (南京大学信息管理学院 江苏 210093)

**摘要** 数字图书馆作为一种信息环境,其基本属性仍是收藏属性,基本职能仍是文献收藏。在大数据环境下,尤其需要重视对数字信息资源的整理与收藏,在这个过程中需要政府的强力支持,数字资源呈缴制度的建立与法律保障,数字图书馆联盟的组织,国家图书馆的核心领导,信息生产商、出版商、网络用户、技术商的合作,图书馆、档案馆、个人用户的协调,以发挥数字图书馆在大数据环境下保存和传承人类文化的基本职能。

**关键词** 收藏 基本职能 数字图书馆 大数据环境

Basic Functions of Digital Library in the Big Data Environment

Wang Meng Zheng Jianming Chen Ya

(School of Information Management, Nanjing University, Jiangsu, 210093)

**Abstract** Digital library is a kind of information environment. Its basic property is still the collection attribute, basic functions is still the literature collection. Under the environment of big data, more attention should be paid for the arrangement and collection of digital information resources. It needs government's strong support, presentation system of digital resources and regulations guarantee, the construction of digital library alliance system, the guidance of the national library, the corporation of information producers, publishers, technology corporations, the coordination among library, archives, individual users in order to play the basic functions of digital libraries in big data environment.

**Keywords** collection, basic functions, digital library, big data environment

随着移动技术、互联网技术、云计算等信息技术的快速发展,全球数据量狂飙式增长,如何对大规模数据进行采集、处理和应用成为企业和管理机构能否取得成功的关键。在国际上特别是欧美发达国家掀起了大数据研究和应用的热潮。图书馆一向是新信息技术的研究和实践重地,大数据不可避免地将会对图书馆产生影响。然而如何在大数据环境下发挥数字图书馆的基本职能,需要我们深入探讨。

据笔者对 CNKI 数据库的文献调研,发现研究数字图书馆与基本职能主题的相关文献(包括期刊论文、会议论文、博硕士学位论文)只有数十篇,而以关键词或题名进行检索,结果则更少。然而随着大数据环境的到来,数字图书馆基本职能是什么,如何发挥作用,更需要深入研究。笔者首先着眼于探索数字图书馆在大数据环境下的基本职能和基本属性,进而分

析大数据环境下馆员如何根据基本职能需要发挥作用,最后提出大数据环境下如何实施数字图书馆基本职能的对策,以为数字图书馆服务实践提供一定的借鉴。

## 1 数字图书馆是一个信息环境

准确理解数字图书馆的基本职能,首先要知道什么是数字图书馆。数字图书馆是什么?数字图书馆最早在 20 世纪 40 年代提出,而真正对数字图书馆进行定义则是在 20 世纪 80 年代末、90 年代初随着互联网发展起来的。很多机构、组织和个人对什么是数字图书馆进行了研究:美国数字图书馆联盟 1998 年定义数字图书馆是一个具有专业人员等资源的组织,其要对数字资源进行采集、组织、存取、翻译、传播、永久保存,从而使数字资源能快速、经济地被使用;大英图书

本文系 2013 年高等学校博士学科点专项科研基金(博导类)“数字图书馆协同机理研究及其概率群集建模应用”(编号:20130091110006)、2011 年江苏省高校哲学社会科学研究重点项目“江苏省‘十二五’时期数字文化产业之数字内容服务模式研究”(编号:2011ZDIXM011)的研究成果之一。

馆认为数字图书馆是利用数字技术获取、存储、存取、发布信息的图书馆;中国国家图书馆认为其可为国家信息基础设施提供关键性信息管理技术,同时提供其主要的信息库和资源库<sup>[1]</sup>。William认为数字图书馆是发挥服务功能的经过加工的信息收藏,其以数字格式存储并能借助网络进行存取<sup>[1]</sup>;孙坦认为数字图书馆的核心是借助先进的信息技术,把计算机网络当做基础平台,从而建立起一个环境,在这种环境下,数字图书馆不仅局限于网络信息资源之开放与利用,其更是一个可以进行信息获取、传递和交流的知识网络<sup>[3]</sup>。

从以上对数字图书馆的定义可以看出其许多共性,比如强调资源的收藏保存、技术的应用、用户的需要。西方关于数字图书馆的研究有两类代表性观点:一是5A系统,5A系统实际上就是阮冈纳赞的五定理的数字化,就是任何读者在任何时间、任何地点利用任何网络链接去获取网络上的任何资源,这就是5A系统;另一类是5S系统,即利用任何技术、利用图书馆数据库的任何服务、利用各种服务的手段在任何服务的地点去获取所有的服务。数字化背景下,图书馆是一种环境,在这种环境中,图书馆一方面是坐落于城市的某个地点之实体空间,而另一方面则是利用先进技术与设备为用户提供信息与知识服务<sup>[4]</sup>。2013年川、吉、冀、桂、苏图书馆学会学术研讨会上,国家图书馆的陈力副馆长在《对我国公共图书馆未来发展的思考》主旨报告中,认为“数字图书馆就是数字时代的图书馆”。在这个定义中他对图书馆的基本属性、性质和对图书馆的基本职能做了高度的概括,也肯定了数字图书馆的5S系统是一个服务的系统,一种服务的环境。因此,我们可以将数字图书馆看成是一个信息环境。

## 2 数字图书馆的基本属性和社会属性

### 2.1 图书馆的基本属性和社会属性

图书馆这个词是从日本引进的。Library翻译成日文时平假名、片假名处理不了这个信息,所以就用了三个字,图书后面是收藏的机构、集中的机构。日本人喜欢用藏、馆、所,所以用了一个馆字。早在殷商时期中国就有图书馆,那时称为藏书楼,藏书楼实际上更形象,就是一个藏书的地方。图书馆就是藏书的地方,这界定了其基本属性:第一,图书馆是一种社会机构;第二,这个社会机构的属性是收藏机构;第三,它的职能是收藏图书。在古希腊、古罗马时期,西方图书馆的前辈们就已经对图书馆的性质和职能进行了界定。古埃及早期图书馆建设者亚历山大一世在登基后的重要举措之一就是建图书馆,收藏天下之文献。亚历山大当时建图书馆收藏天下之所有是要把天下所有之文献归属天下所有人,即他建图书馆通过世代代的努力收藏天下所有文献之目的不是为国王、为宫廷贵族,而是为天下人。可以说,图书馆从产生之日起就担负起了保存文化遗产的社会责任。

中国近代是东西方文化交织、碰撞的时期,东西

方文化背景的各种学说在中国大地上冲撞,产生了很多学说,形成了很多学派。围绕图书馆学也形成了一些学派,如古典学派、西方学派,也有中西合璧学派,在这么多学派中有一个学派后来被淹没,这个学派如果流传下来可能影响中国图书馆学的发展方向。这个学派是以梁启超为代表的一批搞文献和图书馆研究的学者。1925年,梁启超作为学者在中华图书馆协会成立会议上做演讲,他提到了中华图书馆协会要通过图书馆同仁的努力建设中国式的图书馆学。他认为:图书馆是文献收藏的机构,图书馆的文献收得越多越好,图书馆的分类、编目不是为了利用而是为了更好的收藏。梁启超认为中国式的图书馆学首先是侧重收藏,其次是侧重分类编目、组织。他是要让读者自己到图书馆获取所需要文献,这就是梁启超的中国式的图书馆学。在馆藏资源建设的策略上,不管古代之重藏不重用,还是近现代之藏用并举、藏以致用,信息资源建设都是依据一定的收藏原则,确定相应的收藏范围,对相对分散而零乱的文献进行搜集和整理,从而形成一个系统且完整的文献收藏体系,世代相传、留存至今。收藏属性是图书馆作为社会机构的基本属性,对文献的收藏是图书馆作为社会机构的基本职能,服务则是图书馆的社会职能。

### 2.2 数字图书馆的基本属性和社会属性

现在数字图书馆建设中出现了一个问题,那就是我们越是想建设数字图书馆、越是希望利用数字图书馆开展各种服务,我们就越无所适从,我们的心就越虚,以至于图书馆消亡论等观点纷纷出炉。范并思教授提出需要颁布《公共图书馆法》,以法律的形式把服务固定下来,保证公众能够得到图书馆的服务。但是,我们认为公共图书馆法的目的只是以法律的形式把社会公众应该享有的图书馆的权利给固定下来,而不是为了推进服务。在数字时代,如果图书馆以服务机构自居,一味强调图书馆的信息服务功能,会对数字图书馆的发展产生不利影响,图书馆甚至会因此而灭亡。但是图书馆作为一个资源收藏的机构,它收藏的文献承载着人类的文明、人类的历史,人类需要去继承和发展的物质载体、精神载体、知识载体,这使得图书馆又不会灭亡。文献是承载着人类历史和文化的载体,但是这些承载文化遗产、历史遗产精华的物质载体有很多已经不能使用,不能使用则意味着人类将不能再继承,这将不利于人类的发展。通过文献资源数字化可以使这些资源继续被使用,从而解决人类历史、人类文化遗产的继承问题。因此,数字时代的图书馆担当的基本职能依然是收藏职能。数字文化资源承载着人类的社会文明和科技成果,是促进数字时代科学技术进步与发展的重要力量。对于数字图书馆而言,收藏数字文化资源并使其永久持续保存,使人类的社会文明成果能够继承和发展下去,也是一项基本的社会责任。

在数字时代发展的过程中,尤其是大数据时代,我们一直关注数字信息资源的利用,很少有人注意数字资源的收藏。数字资源的收藏不仅仅是图书馆的工



作,而应该是整个社会都关注的问题。美国国会图书委员会在《21世纪国会图书馆数字战略》报告中就将建立数字馆藏和数字资料保存作为数字图书馆建设中的首要议题;大英图书馆制订的《2008—2011年发展战略》提出数字化建设的首要目标就是进一步促进大英图书馆馆藏内容之存取,保存大量数字内容<sup>1</sup>。国际上也早已认识到数字图书馆对数字文化资源收藏的责任,开展了大量的文化数字资源保存项目,例如,美国国会图书馆推出NDIIPP项目之目的是建立数字资源收藏、保存的全国性战略;斯坦福LOCKSS项目也重点研究了电子期刊的长期保存和利用<sup>2</sup>。在新中国成立后,文化部也颁布过图书样本、期刊样本、报纸样本的呈缴办法,为了执行呈缴样本的办法和条例,国家还专门建立了版本图书馆。然而,数字时代我国数字资源的保存还处在初级甚至可能是空白阶段,中国不仅没有数字资源呈缴本,数字资源保存在哪里也没有一个明确的机构,各图书馆各自为政。在基本职能的社会属性当中也有一些推波助澜的因素,就是数字内容产业化的社会发展趋势。数字内容产业化所指的就是数字文献、数字文化的生产和传播的产业化。人类文化历史传播的基本载体被作为一个产业去发展,使得社会发展面临着一种新的危机,那就是由于产业化发展,使知识成为一种资本,它在社会的发展当中所发挥的作用和价值只体现在资源的配置和效益的获取上。这些混淆的认识使得我们对数字图书馆在社会当中所担当的基本职能、基本属性产生了混淆,造成的后果不是图书馆的灭亡而是馆员太忙了,业务外包被认为是一种趋势,图书馆不要再分类,图书馆不要再编目,图书馆员们主要的精力是服务,这样一种发展趋势使得图书馆员开始远离自己事业的核心。

### 3 馆员的责任

中国的第一部图书分类法产生在两千多年前的西汉时期。刘向、刘歆父子集二十多年之力,对西汉王朝所收藏的文献进行整理,这是史料记载的第一次国家级的文献整理,国家收藏文献从这个时候开始历经千年。在这次国家文献整理中有几个创新:(1)创立一套国家文献工作整理的程序;(2)形成文献整理和揭示的方法体系,分类、编目、写提要;(3)形成一个七分法的分类体系国家标准,六大类三十八子目的分类法分类体系,对国家文献进行规范整理后又广罗异本进行文献校勘,形成正本文献,写序录,进行分类。他们做这些工作是为了收藏,是为国家而整理。西方也强调对文献的整理,当然西方对文献的整理和我们不一样,西方早期文献分类整理是把文献当中所揭示出来的学科发展、人们的认识由专家做一些概述,形成条目,通过编写百科全书的形式进行组织和整理,早期的百科全书就是分类式的。上文提到的梁启超当时讲建设中国式的图书馆学第一就是要研究分类编目系统,第二要研究中国的类书编写之法,类书是中国文献编纂形式。《永乐大典》就是类书,类书的编纂就是对文献的分门别类的整理,这与百科全书是不同的。

类书在中国的历史上担当的职能就是收藏的职能,把收集的资料分门别类进行汇编。

当前图书馆员每天都在忙于为用户提供各种服务,殊不知,我们的高尚是表现在对人类文化、对人类遗产的收藏使命上,而不是在别人面前炫耀图书馆的工作。在数字图书馆建设中,我们为了收藏开始和一些出版商签订授权协议,数字图书馆联盟强调资源的集团建设、采购,用尽可能少的钱买尽可能多的资源,但是应该调整现在的集团采购模式,因为集团采购使每一个图书馆都没有个性,没有自己的藏书体系、没有自己的资源体系。此外,存储功能的实现需要一代技术更新、一代代的资源覆盖,这些都是造成我们社会历史空白的重要节点,图书馆员怎么能够抓住这些节点,让我们收藏的资源真正的一代代传承下去?怎样实现数字保存真正永久的实现?学术界已经意识到这个问题但还没有展开详细讨论,我们以为研究者可以在数字图书馆基本职能方面寻找一些选题进行进一步研究。

## 4 数字图书馆的基本职能实施

### 4.1 数字资源呈缴制度的建立与法律保障

国外合法呈缴制度中有的国家如丹麦、新西兰等明确提出了可以采集网络资源,有的国家将呈缴范围限定在数字出版物,而我国目前还没有一部系统的数字资源呈缴法律。所以,当前建立我国的数字资源、网络学术信息资源呈缴制度很有必要,尤其是文化部正在草拟《中国图书馆法(征求意见稿)》、《公共图书馆法》,应对数字信息资源,尤其是网络学术信息资源等的缴送问题要有较明确的规定,并与档案管理部门积极协调<sup>1</sup>,在网络信息的归档方面积极推动,并根据环境的变化,适当扩展呈缴客体<sup>2</sup>,将符合条件的数字资源引入国家信息资源保存体系,收集更多高质量的数字资源,最终解决现阶段数字图书馆开展信息资源保存活动时难以在著作权法规中找到法律支持的难题,促进国家文化遗产的有效保存、传承和利用。

从以往经验看,国家图书馆在呈缴制度参与和实施方面发挥着不可替代的作用。在国外,许多国家的国家图书馆就数字资源呈缴制定了相应计划。诸如《美国国会图书馆2011—2016年战略规划》、英国国家图书馆《应对知识增长:英国国家图书馆2011—2015年战略》<sup>1</sup>都在规划中认为要通过法定呈缴制度促进数字资源收藏,以保证后代自由获取信息资源。中华民族在传统上就注重收藏典籍,注重通过文献收藏实现中华文明的传承和延续。因此,我们需要根据环境的变化,着重解决数字出版物等数字信息资源纳入法定呈缴面临的各种关键问题,并通过制定科学的数字资源呈缴制度促进数字信息资源的收藏。

### 4.2 数字图书馆联盟的组织

数字时代的到来,数字信息资源迅速增长,随之而来的是用户在信息资源需求数量和质量上的要求比之以往更高,各个图书馆如果按“大而全”的方式去建设各自的馆藏,就会使资源重复建设严重,图书馆

馆藏体系混乱,而且没有一个图书馆能够收藏所有的文献资源,古代图书馆如此,当代数字图书馆也不例外。所以,数字图书馆的馆藏体系不要局限于本馆的文献信息资源,各馆要借助现有技术手段,采取合作策略,促进数字馆藏的协同发展。数字图书馆联盟体制是实际而有效的资源合作收藏与共享模式<sup>[9]</sup>,在相互合作中,可以争取政府的政策支持,引导媒体的舆论方向,扩大社会影响<sup>[1]</sup>。数字图书馆联盟建设使数字图书馆在数字资源体系建设中突破了“大而全”、“小而全”的藏书传统观念,建立了联合协作、合理布局、资源共建、资源共享的收藏新观念,在构建自身特色馆藏的基础上,建立起一定范围的协同发展的文献保障体系。数字图书馆联盟合作也不仅限于国内,可以推进国际合作,如2012年,中科院国家科学图书馆与英国自然出版集团达成数字资源长期保存合作共识,使国家科学图书馆可以本地保存48种Nature品牌期刊的电子版本,就是值得推进的国际合作范例。

#### 4.3 国家图书馆的核心领导

国家图书馆作为全国各类图书馆的领导核心,是专业性的非营利保存部门,其可以在数字资源收藏体系中发挥核心作用<sup>[1]</sup>。首先国家图书馆应制定数字资源收藏计划,明确其作为国家数字资源保存主体的地位,收藏全国各地、各种数字信息资源。其次,订立数字信息资源收藏的具体实施方案和数字资源采集之形式标准、内容标准,坚持实施和贯彻这些标准和制度,并进行相关教育、培训。最后,协调各方、各机构定期召开例会,讨论工作进展,听取各方意见,修正具体工作细节等。如美国国会图书馆在NDIIPP项目中就发挥着核心作用,通过国会授权,积极组织各机构开展数字信息资源收藏、保存研究与实践。

地方各级信息收藏机构,如各级公共图书馆、高校图书馆、档案馆,应根据国家图书馆制定的各种原则、标准,对数字资源进行采集、处理,按要求提交数字资源,交流收藏情况,反馈具体工作中存在的问题,在部门之间合作中实现数字资源整合和共享。通过与各级政府部门交流,使其对于数字资源收藏、保存有一个全面而具体的认识,这样在制定有关收藏保存政策时就会更加符合实际,更加具有针对性。

#### 4.4 数字信息资源生产商、出版商、网络用户、技术商的协调

构建数字资源收藏保存机制<sup>[1]</sup>,资源商与数字图书馆要相互从对方角度出发,相互信任、积极沟通、共同努力达到合作共赢。资源商包括信息生产商、出版商和网络用户,其数字信息资源尤其是网络信息资源之生产者,具体可从以下方面协调:规范信息资源处理流程,收藏有价值的数字信息资源。这不仅能拓展数字资源的收藏渠道,而且在信息加工、处理方面可以有效简化和优化工作流程。然而,如果没有适当的人力、物力、财力相匹配以及相应的精神激励,数字资源收藏就会难以为继,因此需要制定相关的政策来调动他们的积极性。虽然在资源利用上,信息生产商、出版商与数字图书馆之间有利益冲突,比如图书馆是公

益性机构,其在数据库、电子出版物等数字信息资源服务中不可避免地会对信息生产商、出版商等的利益造成损失。再者,为节省成本,各图书馆通过图书馆联盟采取的联合采购方式也对信息资源提供商的发展造成一定的威胁。需要明确的是,即使如此,数字图书馆的收藏是不会损害以上信息提供商的利益,甚至还可以帮助其扩大宣传,引导舆论,促进其利益的实现。因此,图书馆可以采取一些平衡的手段,与信息生产商、出版商主动交流,就数字资源使用的范围和使用条件展开协商,确保在合作中互利共赢<sup>[1]</sup>。个人网络用户是网络信息资源保存不可忽视的一个重要群体,他们也是数字信息资源尤其是网络信息资源的重要生产者,但是将个人用户纳入数字信息资源收藏的体系还存在着许多困难。对于如何在保存工作中利用好个人用户这一群体还需要进一步的探索。

技术商是数字信息资源收藏的一个辅助性角色,为整个数字资源收藏保存体系提供技术方面的支持,从信息的获取、筛选、存储、评估等环节提供应用程序和硬件设备。但是需要图书馆、档案馆的专业人员协同开发部门共同研发保存软件,协调各方,与各个数字收藏保存成员充分交流,了解技术需求,并在具体的实践中不断完善,提高技术应用水平。

#### 参考文献

- [1] 数字图书馆 EB/OJ. [2013-05-25]. <http://baike.baidu.com/view/8181.htm>.
- [2] William Y Arms. Digital Libraries[M]. MIT Press, 2001.
- [3] 孙坦. 基于网格的数字图书馆研究[J]. 图书馆建设, 2008(1): 86.
- [4] 谢欢, 沈妍. 移动环境下图书馆的机遇与挑战[J]. 图书与情报, 2011(9): 65-68.
- [5] Digitisation Strategy 2008-2011[EB/OJ]. [2014-04-05]. <http://www.bl.uk/aboutus/stratpolprog/digi/digitisation/digistrategy/index.html>.
- [6] 潘晓秋. 美国数字资源长期保存主要成果及其启示[J]. 现代情报, 2008(9): 98-101.
- [7] 吴鸿骊. 图书馆文献资料呈缴制度的几点建议[J]. 图书馆建设, 2012(9): 1-3.
- [8] 吴钢. 数字出版法定呈缴制度客体研究[J]. 中国图书馆学报, 2014(1): 93-102.
- [9] British Library International Survey on E-Legal Deposit 2011 Summary of Findings[EB/OJ]. [2014-03-29]. <http://www.cdnlib.info/2011/pdf/e2Dlegaldeposit20survey20CDNL20Slides20Aug%20Compatibility%20Model.Pdf>.
- [10] 赵莉. 数字图书馆联盟探析[J]. 情报杂志, 2006(3): 127-129.
- [11] 李富玲, 卢振波. 数字图书馆联盟研究[J]. 大学图书馆学报, 2005(2): 11-15.
- [12] 许群辉. 美国数字信息资源保存项目NDIIPP及其启示[J]. 现代情报, 2006(9): 67-69.
- [13] 王猛, 郑建明. 我国数字图书馆运行模式研究——基于公共文化事业和泛在知识环境视角[J]. 图书与情报, 2012(9): 21-25.
- [14] 申晓娟, 胡洁. 数字图书馆知识产权策略探讨[J]. 图书馆, 2012(2): 35-38.

[作者简介] 王猛,男,1989年生,南京大学信息管理学院博士研究生。  
郑建明,男,1960年生,南京大学信息管理学院教授,博士生导师。  
陈雅,女,1965年生,南京大学信息管理学院教授。

收稿日期:2014-04-24