



# 古籍数字化与文化共享

刘明华

(西南大学 文学院,重庆市 400715)

**摘要:**古籍数字化的理想境界是在这一领域实现文化共享,其服务对象为大众和学者两大群体。可从建立公益性的“中国古籍基本读物网”和注册式的研究性局域网入手实现资源共享。该工程可从版本的权威性和共享工程的多层性上逐步实现书目、文献、成果和检索软件的共享。在经费投入上可按政府专项补贴和校际互惠模式进行。

**关键词:**古籍;数字化;文化共享

**中图分类号:**G250.76 **文献标识码:**A **文章编号:**1673-9841(2008)06-0157-04

## 一、古籍数字化文化共享的基本目标

列入国家“十一五”文化发展纲要的数字出版工程共有八项,包括数字化多媒体研发工程、国家数字复合出版系统工程、中华字库建设工程、国家知识资源数据库出版工程、国家动漫振兴工程、中国古籍数字化工程、国家版权保护技术开发工程、数字化文化传播工程,“中国古籍数字化工程”是其中一项,并已完成个别重大项目如“中国基本古籍库”等。但我要说的是,这项工程,其最终目的应是一个“共享工程”。

这里的“共享工程”,受到了“文化信息资源共享工程”的启发。这是国家在文化广电领域服务基层群众的一项重大工程,基本内容是应用现代科学技术,将中华优秀传统文化信息进行数字化加工和整合,通过工程网络体系,以互联网、卫星网、有线电视/数字电视网、镜像、移动存储和光盘等方式,实现优秀文化信息资源在全国范围内的共建共享。由中央政府投资,地方政府配套,分别在省会、地市、区县、乡镇建立服务网点,方便群众使用。自2002年实施以来,该工程依托全国各级公共图书馆、乡镇、街道文化站和村、社区文化室,已建成1个国家中心、33个省级分中心、1970个市县支中心、13758个乡镇街道基层服务点、303371个村、

社区基层服务点,以及1109所学校、企业、部队等服务点。在该工程建设过程中,各地还根据实际情况,与其他相关部门的类似“工程”进行资源整合共建,产生了积极的作用,如通过共建共享方式与农村党员干部现代远程教育工程、农村中小学现代远程教育工程合作,共建农村点35万个,中小学点25万个。初步形成了覆盖全国的服务网络,基本实现“数字文化村村有”的战略目标。

文化信息资源共享的目标是服务最基层的群众。这是国家保障人民群众基本文化权益所进行的一项“德政工程”。这项工程中有两大内容需要经费:一是建立各级网点的设备配置,二是工程中需要添加大量优质文化信息并进行数字化处理。

古籍数字化的服务对象是谁,国家能够进行如此大规模的投资吗?

古籍数字化的目的,应有两大内容——普及与研究。从服务对象上看,正好表现为两大读者群。

面向普及性的服务,目的是让有一定文化程度且对古籍有兴趣的读者能够通过数字化工程方便地看到优质的古籍资料,增加对优秀传统文化的了解和热爱。这一群体应是文化传承中的基本队伍。对他们的服务应该最具公益性。其电子化的形式应是“全文型”的文本,即将古籍资源全文通过手工录入形成的电子文本。

\* 收稿日期:2008-05-27

作者简介:刘明华(1956-),男,四川平昌人,文学博士,西南大学文学院,教授,博士生导师,主要研究中国古代文学。

基金项目:教育部人文社会科学研究规划基金项目“文化视野下的古代作家作品阐释”(06JD75011-44002),项目负责人:刘明华。

面向研究性的服务,目的是为从事研究工作的群体提供方便、快捷、准确的古籍资料,为其研究提供数字化文献支撑。这一群体应是文化传承中的骨干队伍。面向他们的服务同样是公益性的,但因为工作性质和单位的特殊性,面向这个群体的服务往往可以通过局域网进行,可以根据产品成本实行酌情收费。其电子化的形式应是“图文型”的文本,是既保存古籍图像又可检索的文本。具体为在古籍书页图像存储基础上,将书中具有检索意义的内容数字化,并转为电脑可识别的文字,并辅以适当的软件工具,为读者提供快捷的检索、统计、整理和编辑功能的电子产品——作品文本兼工具书。

但无论是为了这两个目标的实现还是为这两大群体作好服务,都涉及到这一工作的基本性质问题。笔者认为,无论是从图书的功能、图书馆的功能,还是文化服务的功能看,古籍数字化的终极目标还是公益性的。

关于古籍数字化的公益性问题,有识之士不乏先见。国家图书馆副馆长陈力指出:“中国古籍的数字化是一项中华优秀传统文化的保护和弘扬工作,具有强烈的公益性色彩,需要各方面加强协调,有一个整体的规划。”<sup>[1]</sup>他对公益性尚未展开讨论,且更多注意的是这项公益性工程中的“整体规划”问题,如“数字化对象的内容确定和合作分工,同时包括相关标准,规范的统一”等等。我所看重的“公益”,是“整体规划”后顺利完成的古籍数字化的产品如何发挥更大作用、如何进入寻常百姓家的问题。

古籍数字化的公益性问题,说到底还是一个文化保护、文化普及、文化提升和文化服务的问题。简而言之,一是为更多的人提供服务,二是为深入研究的人提供服务。从公益性角度考虑,国家应该在这一领域发挥更大的作用。

既然是一项公益性的有极大意义的事业,且已经纳入国家发展规划,政府就应当投入或加大投入。既然是政府投入,其成果就应该面向人民群众,政府以满足公众应享有的文化权益为前提而提供最基本的服务。相关研究者或部门、团队,将相关古籍进行数字化开发之后,应该向国民开放,即交给开放的或局域的网络图书馆让群众使用。不管国家投入多少经费,都不宜以谁开发、谁受益的思路来进行。不能像某些地区的高速公路建设那样,实行谁投资、谁受益甚至无限期收费,以至变成一个利益集团的摇钱树。

## 二、古籍数字化文化共享的可行性

这个愿望的可行性在何处呢?如何得以实现

呢?我认为可行性是有的。

### (一)建立公益性的“中国古籍基本读物网”

网站由政府投入,不会花太多的钱。一个文史哲本科生所应该看到的古籍读物,或一个中国人应该了解、一个外国人想了解的中国文化最基本的也最精华的读本,在这个网站上都可以浏览、查阅,并可以检索。一千种书,足矣。从目前我们知道的古籍电子化读物的数量看,这已不是问题。以台湾“汉籍全文资料库”为例可以说明这一点。该文库上传二十五史、十三经、大经藏、词话集成及台湾史料等资料一千余种,数亿文字。部分资料的查询需付费取得密码方能检索,大部分资料免费开放,不仅提供查询,还可全文浏览古籍。该文库专设人文资料库师生版,内有总量近亿字的重要文史典籍,向台湾大中小学师生免费开放,实际是向互联网所有读者开放。国内一些网站,在古籍读物的普及工作方面也有所努力。在一些重要的局域网或单位间联合制作的网络资料如《中国文学史参考资料》上,就有数百种典籍、上千万文字可以阅读。国学网开发的《国学备览》,通过光盘这一载体,亦提供上千万字的古籍电子读物。这类公益性网站的建立,技术和资金上都不存在太大问题。

这类网站的管理亦可是公益性的。点击率不低,也可通过广告的形式补贴,维持网站运营。

### (二)建立注册式研究性局域网

对于研究性的古籍数字化资料,可以通过局域网实现资源共享。一些成功的软件已经走在前列,产生了社会效益和经济效益的双赢。这个设想中的工程的基本要求应该有以下内容:

第一,版本的权威性。

目前进入市场的很多古籍电子化全文本型产品,由于成本原因,相关制作单位和个人主要是采用印刷业中形成的古籍整理电脑排版小样文件,即“方正排版小样文件”。笔者曾与一位在这方面颇有成就的专业人士交换过意见。我认为,古籍电子化产品的学术水准应以进入专业研究者的参考文献为最高目标。而目前最好的产品,除图文型产品如“书同文”研制的《四库全书》外,均难到达这一境界(而此版的本质仍是其“古籍图像”的“电子化”原貌)。专业人士的回答很简单,全文本如要做到这个水平,成本太高,高到一个企业或研究单位无法承受。这几乎成为古籍电子化产品向高端迈进的瓶颈。而版本的权威性的确又是专业水平的体现。所以,如果要进行这一重大的工程,这仍是一个不能回避的问题。在这个设想的工程中,将进行数字

化处理的古籍,必须由权威机构和专家确定版本的权威性。在全文本中,也包括今人可靠的校点、校注本。在这一点上,台湾《汉籍全文资料库》值得借鉴,其所用古籍底本主要采用大陆学者的整理研究成果,如《二十五史》的底本就是中华书局点校本,台湾学者在电子文库的建设中再参与校订工作。所收录的先秦诸子也多选用中华书局“新编诸子集成”本。而大陆古籍电子产品目前所用的“方正排版小样文件”,均不注明原排印本的出版社,这既未尊重原整理者和出版社的劳动,又不能保证其学术水平,甚至还造成知识产权方面的纠纷。所以,在以后的古籍电子全文本中,对整理者的知识产权保护和对版本权威性的明确上,需要尽快完善相关规定。如果由古委会来推进这项工作,可由古委会授权专家组或专家委员会,先对电子化处理的古籍版本进行确认,保证其权威性。也可另行成立相应机构专门进行这项工作,以保证其质量。

无论是普及性还是研究性的文本,在做到是目前海内外最好的版本同时,还应强调电子版本确定的动态性原则。即是说,当有了更好的版本时,数字化处理的文本可以调整。而确定某一版本纳入数字化处理的过程,即前文所谓“整体的规划”范畴。进入该工程的版本,无论是影印本还是校本,其质量均可作为学术论文的引用书目。这是这一工程质量评价的主要标准,也是研究型古籍电子化文本中的核心问题。

## 第二,共享内容的层次性。

一是书目的共享。目前传世古籍究竟有多少?这个工程应该尽可能完整反映。在古籍文献书目方面,海内外均有较好成绩。据目前可从网上检索到的信息,可知多家图书馆在进行这方面的工作。如:中国国家数字图书馆(<http://www.nlc.gov.cn/service/guancang/gujishanben.htm>),在导航区点击“善本古籍”,可检索其馆藏的27万种古籍中的善本;琳琅秘籍——北京大学数字图书馆古文文献资源库(<http://rbd1.calis.edu.cn/pages/Instruction.htm>),收录北京大学图书馆古籍约150万册,其中善本近20万册;其他如中国科学院文献情报中心古籍管理系统(网址略,下同)、中国人民大学图书馆普通线装古籍书目数据库、上海图书馆古籍书目查询、天津图书馆馆藏古籍善本图录、山东省图书馆古籍书目检索、北京师范大学、复旦大学、吉林大学、香港大学图书馆、美国哈佛大学燕京图书馆等,均有古籍书目可查询。台湾在古籍电子检索书目方面可供查询的网站更为丰富,仅网络古

籍全文检索系统就有:台湾中研院汉籍电子文献(<http://www.sinica.edu.tw/ftms-bin/ftmsw3>)、史语所中文资料库(<http://www.ihp.sinica.edu.tw/database/index.htm>)、史语所简帛金石资料库(<http://saturn.ihp.sinica.edu.tw/~wenwu/search.htm>)、中华电子佛典线上藏经阁大正藏全文检索(<http://ccbs.ntu.edu.tw/cbeta/result/search.htm>)、故宫寒泉资料库(<http://libnt.npm.gov.tw/s25/index.htm>)以及《诗经》、《全唐诗》、宋诗、唐宋词、台湾古典汉诗、《红楼梦》、《三国演义》、《水浒传》等多种全文检索系统(均见<http://cls.admin.yzu.edu.tw>)等。对这些资源加以整合,读者在书目方面的查询是非常方便的。

二是文献的共享。如前所述,可以是大众的公开性的网站开放,也可以是“小众”的局域开放。共享产品应包括目前古籍电子化的几种基本形式的所有内容,即古籍书目型、全文型、全图像型和图文型数字化产品。局域的开放是整体的有序开放,如高校联盟或研究单位的联盟网站。如果实现单位间的资源有限共享,其网络的受众最终仍是大众。

三是研究成果的共享。在充分考虑作者权益的基础上,可以将古籍研究成果放在相关网站,供读者或研究者共享。这类共享成果,还可从作者使用古籍共享资料的深浅度进行前期约定。当然,其核心仍是版权问题。目前超星图书馆和文化信息共享工程的一些作法可以借鉴,即在征得作者同意的基础上进入共享。如果研究者有相应报酬,这项工作的进展会更快些。而投入仍是关键。

四是检索软件的共享。检索软件本身是科技研究成果,前期投入无论是经费和还是工作量往往要比一部古籍的整理或研究多得多。所以,其成果共享的关键仍是如何对软件作者的知识产权进行保护。研制者的权益得到保护,共享才有可能。第三、四两项内容的根本解决,下面将有细致的讨论。

## 三、古籍数字化文化共享的基本思路

共享有无可能——这是问题的关键。

前面设想的两种方案,第一种较为容易解决。甚至只要古委会下决心,很快就可以实现。如果国家有关部门关注,就更容易解决。以“中文研究网”([www.zwyjw.net](http://www.zwyjw.net) 或 [www.zwyjw.com](http://www.zwyjw.com))为例,西南大学文学院以一个学科点尚能建立一个囊括20世纪以来本学科的论著数据库,向公众开放。以一个大国政府之力,如何不能建立一个弘扬民族文化、方便全民读书的公益性的文献门户网站呢?第

二种方案有难度,实质在于已经投入并形成产品的古籍数字化成果在应用中遇到的不能回避的费用问题,研究、开发、管理、升级等等均需费用。何况还有大量的产品并非政府投入,而是民间资本和个人热爱促成的。如何使这些产品发挥作用并获得相应回报,也是题中应有之义。但这是可以解决的。其思路有:

### (一)专项补贴式

一些重要产品是个人无力购买的,其发挥作用必须建立在集体购买的基础上。由此,也就产生了解决的路径。笔者建议:凡国家投资的古籍数字化项目,其成果在推广时,可由国家进行专项补贴。以《中国基本古籍光盘库》为例,该产品在初期预售价为人民币50万左右,但最终成果出版时,已涨到130万左右。一些高校图书馆因经费紧张,或受众有限,难以购置。而且,目前在教育部的教学评估中,凡数字化读物,无论其内容多少,均不能按其中收录的书籍计算册数。花费巨资购买一个1亿字的光盘,只能算1本书。购买一套《中国基本古籍库》,也只能算500个光盘吧。购买电子产品会增大学校的开支,却不能完成生均藏书册数。馆长们亦有难处。但如果配套投入的话,就会调动学校的积极性。假如国家能够像文化信息资源共享工程一样,按配套鼓励高校或其他单位购买,这样,一个优质产品的普及也就实现了。从目前大陆研究文史哲的学者大都集中在高校这一现状考虑,通过高校的网络进行共享是完全可能的。目前“中国期刊全文数据库”(CNKI)在各高校都很普及,即能说明这一问题。一个原因是CNKI覆盖的学科面较大,各学科均能受益,这是该数据库能普及并自身产生良好的经济效益的重要原因。而人文科学尤其是古籍,不可能立即产生或大量产生经济效益,这就必须要政府高度重视和加大投入。

### (二)校际或馆际互惠式

正在“施工”中的“中国数字图书馆工程”是一个很好的平台。2001年10月由国务院批准立项这项工程,已成立了“中国数字图书馆联盟”,内有41家公共图书馆、30家高校图书馆、8家科研机构图书馆、7家其他类型的图书馆和14家技术公司。古籍数字化虽然不是该工程的核心内容,但也是其庞大资源群开发建设的重要内容之一。该工程的总体目标是在宽带IP网上形成超大规模的、高质量的中文资源群,支持国家整体创新体系的形成与发展,通过国家骨干通信网络系统向全国乃至全球提供服务。其工程的重要内容之一,就是发掘历史遗产。在这个平台中,加盟单位间的互惠是没有问题的。而值得有志于古籍数字化事业的同仁们关注的是,这是一个很好的机会。相关产品进入这个平台,也就意味着进入了中国大陆的主要研究机构,进入了大多数研究者的视野。校际或馆际的合作还不仅止于此,应该有更多的合作空间供人们选择。在互联网的时代,合作意味着共赢。

至于在古籍数字化过程中检索软件的开发和利用,也是重要的副产品。当“共享工程”完成时,软件的普及也才会真正开始。那时,灵活多样的检索软件必有庞大的市场,研制者和研究者都会从中受益。目前,一个可行的办法是通过“政府采购”使其发挥效益。如果设立公共服务体系的古籍网络平台,对那些政府没有投资的优质产品,完全可以用“政府采购”的形式使全民受惠。这样,既考虑到了研发者的前期投入,保护了他们对古籍数字化研究的积极性,肯定了他们做出的贡献,也让广大群众受益,更重要的是弘扬了祖国的优秀传统文化。

#### 参考文献:

- [1] 陈力. 中文古籍数字化方法之检讨[J]. 国家图书馆学刊, 2005(3):11-16.

责任编辑 韩云波

## The Digitization of Ancient Books and Culture Sharing

LIU Ming-hua

(School of Literature, Southwest University, Chongqing 400715, China)

**Abstract:** The ideal state of ancient books digitization is the realization of culture sharing in this field. Digitization of ancient books serves the public and scholars. Resource sharing can be realized through building a public website 'Basic Ancient Chinese books for Readers' and a registered research LAN (local area net). This project, which can be realized through authoritative editions and multiple stratification of culture sharing, is to achieve the sharing of bibliographies, literature, achievements and retrieval software. The fund of this project can be operated with the government subsidy and interuniversity reciprocal funding model.

**Key words:** ancient books; digitization; culture sharing