

·实践平台·

中文古籍数字化的现状与意义

刘伟红 (哈尔滨师范大学图书馆 黑龙江哈尔滨 150025)

摘 要: 古籍数字化是伴随信息技术的进步而发展起来的一个新领域。它与古籍整理、文献保护、文化传承紧密相关。文章从古籍数字化的概念谈起,详细介绍了古籍数字化的现状,旨在阐述古籍数字化的意义。

关键词: 古籍数字化 现状 意义

中图分类号: G256; G250.74

文献标识码: A

文章编号: 1003-6938(2009)04-0134-04

Present Situation and Significance of Chinese Ancient Book Digitization

Liu Weihong (Library of Harbin Normal University, Harbin, Heilongjiang, 150025)

Abstract: The ancient book digitization is a new domain which develops by following the information technology progress. It is closely relevant with ancient book reorganization, literature protection and cultural inheritance. This article mentions from the concept of the ancient book digitization, introduces the status of digital books in detail, which is for the purpose of elaborating the significance of ancient book digitization.

Key words: ancient book digitization; present situation; significance

CLC number: G256; G250.74

Document code: A

Article ID: 1003-6938(2009)04-0134-04

1 概念

古籍整理,事关文化传承,需要古为今用、推陈出新,既要继承传统又要超越传统。我们要用现代理念对传统文化进行再认识,用现代手段对传统文化进行再表现,用现代研究成果对传统文化进行补充。随着社会的进步和科学技术的发展,古籍数字化已经显露出信息传播和利用上的极大优势。

古籍数字化是从利用和保护古籍的目的出发,采用计算机技术,将常见的语言文字或图形符号转化为能被计算机识别的数字符号,从而制成古籍文献书目数据库和古籍全文数据库,用以揭示古籍文献信息资源的一项系统工作。^[1]简言之,就是利用现代信息技术对古代文献进行整理与研究,并以电子数据的方式提供使用。数字化古籍检索方便,易于保存,有传统的纸质出版物无法替代的优势,深受学术界和用户的欢迎。中国古籍的数字化自上世纪80年代兴起后,取得了一系列重要成果。

2 现状

2.1 中国香港、台湾地区中文古籍数字化现状

为了传承中国传统文化,中国香港、台湾和中国大陆三地区的古籍数字化形成了互为补充、共同发展的局面。20世纪80年代,台湾中央研究院开始了“史籍自动化计划”,建立了廿五史全文资料库。经过20多年的积累和发展,现在全球因特网上,由台湾地区整理的中国古典文献资料已有数亿字,而且检索功能齐全,媒体资料多样。香港的古籍数字化工作经过20余年的发展,也取得了一定的成就。在香港,从事古籍数字化工程的中坚力量是香港中文大学中国文化研究所下属的“汉达古文献数据库中心”。该中心成立于1988年,是香港中文大学中国文化研究所“古文献数据库研究计划”的专门出版单位,其研究和开发的目的在于将中国古代全部传世及出土文献加以校订、整理,并收入计算机数据库,然后通过各种媒体出版,从而为研究工作者、教育界以及大众提供重要的学术工具与文化宝库。^[2]

2.2 大陆中文古籍数字化的现状

与港台相比,大陆的古籍数字化工作起步稍晚,但发展迅速,大有后来居上之势。北京书同文数字化技术有限公司制作了文渊阁《四库全书》(与迪志文化出版有限公司合作,2000年完成)、《四部丛刊》(2001年完成)、《历代石刻史料汇编》(2004年完成)、《十通》(2004

基金项目:本文系黑龙江省社会科学基金艺术学项目(省艺科立字(2008)76号)研究成果之一。

收稿日期:2009-02-20,责任编辑:宋戈

年完成)、《大清五部会典》(分别编撰于康熙、雍正、乾隆、嘉庆、光绪五朝)、《大清历朝实录》等电子文献。电子版《四库全书》把原文转化为电子字符,有全文(逐字)、分类(经、史、子、集)、书名、著者四种检索模式,每种模式下还可以进行“与”(同现一卷)、“或”(可只现其中一项)、“非”(排除其中一部分)三种高级检索,电子字符可与图版进行对照。《四库全书》的数字化是古籍数字化进程中具有里程碑意义的一件大事,为古籍的数字化制作树立了标杆,积累了经验。《四部丛刊》除了具有与《四库全书》相同的检索功能外,还提供摘要、笔记、纪元换算及简、繁、异体字相互关联查询的功能。《历代石刻史料汇编》全文版采用当代数字化最新技术制作,中、日、韩汉字大字符集文字平台,也有高级检索功能。1998年成立的北京爱如生数字化技术研究中心是大陆很有实力的古籍数字化专业公司,它制作完成的数字化古籍总数在3万种以上,是目前大陆制作数字化古籍最多的公司。爱如生有一个庞大的古籍数字化规划,其网站上公布的古籍数据库有:

(1)大型数据库:

中国基本古籍库 中国经典库 中国方志库 中国谱牒库 中国丛书库 中国金石库 中国俗文库

(2)数字古典:

敦煌遗珍 明清实录 永乐大典 道教全书 宋会要辑稿 辑佚书汇编 古今图书集成 清帝朱批奏折 历代笔记汇纂 增订四部备要 全四库 古版画

(3)系列数据库:

①别集丛编系列:汉魏六朝人别集丛编 唐五代人别集丛编 宋人别集丛编 金元人别集丛编 明人别集丛编 清人别集丛编 民初人别集丛编

②国学要籍系列:易学要籍 诗经学要籍 尚书学要籍 三礼学要籍 春秋学要籍 四书学要籍 小学要籍

③断代史料系列:秦汉史料库 六朝史料库 唐五代史料库 宋辽金史料库 蒙元史料库 明代史料库 清代史料库

④古典大观系列:古典散文大观 古典骈赋大观 古典诗歌大观 古典词曲大观 古典戏剧大观 古典小说大观 古典评论大观

⑤地方文献系列:山东文献 山西文献 河北文献 河南文献 湖北文献 湖南文献 安徽文献 江西文献 江苏文献 浙江文献 福建文献 广东文献 四川文献 云南文献 陕西文献 台湾文献 广西文献 贵州文献 甘肃文献 辽宁文献 上海文献 北京文献

⑥诸书集成系列:兵书集成 官箴书集成 法律书集成 典制书集成 邦计书集成 禅宗书集成 医

书集成 农书集成 天算书集成 水利书集成 日用书集成 术数书集成 赏鉴书集成 类书集成 辞书集成 目录书集成 博物书集成 清真书集成 艺术书集成

这些数据库有些已经完成,有些正在进行。

2003年,北京国学时代文化传播有限公司与商务印书馆联合推出“中国历代基本典籍库”大型数据库系列光盘。每部书都配有提要,便于读者对使用的典籍有一个基本的了解。国学公司还制作了《文献目录典》、《地理文献典》(收历代地理著作90种)、《中国古代小说典》等专题光盘,这些光盘均具有全文检索、复制及打印的功能。国学公司还从普及的角度推出了《国学备览》、《国学备要》等价廉物美的电子产品,具有全文检索、打印、复制等功能,并内置了联机字典。此外,国学公司又推出了《国学U盘智能图书馆》系列U盘图书,U盘图书配有图片和音乐,还有真人发音字典及人名、书名、地名、帝王年号等专题知识库,数据采用Unicode编码,可在各种版本的Windows下使用。

我国现存最大的古代类书《古今图书集成》已有电子版,该电子版由广西金海湾电子音像出版社和广西师范大学出版社于1999年联合出版,共27张光盘,只是原文图版。为了便于检索,另编有索引数据库,该索引数据库分为“经纬目录”和“索引目录”两大部分,经纬目录和索引目录均提供现代术语与古代术语的对应转换功能和模糊检索功能。

由华东师范大学中国文字研究与应用中心研制的《金文语料库》已于2003年由广西教育出版社正式出版,该软件检索功能强大,能够满足多方面的检索需求。华东师范大学中国文字研究与应用中心还研制了《战国楚文字数字化处理系统》(上海教育出版社2003年出版)、《说文全文检索系统》(包含《玉篇》和《万象名义》,广州南方日报社2004年出版),也是检索功能非常强大的电子古籍。

除此之外,至今有10多所高校和研究机构根据教学和研究工作需要,也进行了古籍数字化项目研究,建立了专门的数据库,如:中国社会科学院的《全唐诗》、《先秦两汉魏晋南北朝诗》、《全上古三代秦汉三国六朝文》、《十三经》、《全唐文》、《诸子集成》等,北京大学的《全唐诗》、《全宋诗》,南京师范大学的《全唐五代宋词》,河南大学的《南宋主要历史文献》等数据库。各大型图书馆则根据馆藏特色,进行了古籍数字化的规划和实施,如:中国国家图书馆的《碑帖菁华》、《敦煌遗珍》等特色资源库,上海图书馆的《中国古籍本查阅系统》等。

人类文明的每一次进步,都伴随着技术的进步,古籍数字化将传统文化与现代信息技术结合,在社会分

工日益明晰的今天,也顺应了人全方面发展的需求,大而言之,它肩负着传承中华文明的重要使命;小而言之,它为学术提供研究利器,满足新时代的阅读古籍的多方面、多层次需求。〔3〕

2.3 欧美地区中文古籍数字化概述

美国图书馆研究学会由北美36家图书馆组成,其宗旨是通过计算机数据库的联网实现馆际资源的共享。目前数据库有25000条数据,并在不断扩充之中。哈佛燕京图书馆“线装古籍计算机检索系统”是查询包括16000余部普通线装书、4000余部善本、3800余种方志、1000余种家谱的最好检索系统,现在已全部编目完毕。哈佛燕京图书馆还与加拿大麦基尔大学东亚系合作,将中国明清时期的妇女著作进行了数字化并上传到网络,目前已完成200余部古籍的数字化工作,该数据库可进行浏览阅读、目录检索和关键词检索。“Online Catalog”为美国国会图书馆的书目检索系统。美国国会图书馆是美国收藏中文书籍最丰富的地方,其中中文古籍754300册,地方志4000余种,中文善本1200种,还包括罕见的珍本图籍。“Online Catalog”于1995年投入使用后经过多次扩充和更新,使用者需以汉语拼音输入检索条件,数据结果以英文呈现。目前美国国会图书馆已成功地将21万件馆藏中文古籍予以数字化处理。大英图书馆典藏中国珍贵文献和古籍60000多种。大英图书馆于1980年开始采用MARC磁带建立“古版书简明标题目录”,它是全球该类文献最大的联机数据库,大英图书馆“中文图书联合检索”包括大英图书馆与牛津大学、剑桥大学等六所大学、研究所的图书馆所收藏的中文书籍,含中文古籍。使用者可直接输入中文或汉语拼音来查询书目,所得信息也以中文显示。

3 意义

3.1 古籍数字化有利于古籍的保护

利用古籍数字化资源能有效地降低纸质古籍的流通量,减少人为的使用性损伤机率,使历经岁月沧桑的古籍得到真正的“休息”和“保护”〔4〕通过复制转存,还可以相对真实、永久地保存古籍原貌。世界上有许多伟大的民族,也创造了同样辉煌灿烂的文明,但是都没有流传下来,只成为了历史的遗迹,其重要原因就是文明载体缺乏连续性,无法流传。因此,加强古籍数字化是古籍再生性保护的有效措施。

3.2 古籍数字化是延续民族文化的重要手段

一个民族的文化精神可以说是一个民族生存下去的理由和灵魂,对传统的扬弃,决定着这个民族文化精神的统一性、传承性和创新性。而对传统文化进行扬弃的前提是首先要了解传统文化,入乎其内,方可出乎其

外。数字化工程的同时,我们对传统文化进行全面整理,也是对传统文化的再学习。

中华五千年优秀文化遗留下来卷帙浩繁的古籍文献资料,但由于人为、自然等许多原因使许多古籍消失了,而今流传下来并封藏于高阁之中的古籍文献资料正在呼唤人们重新发现和认识它的价值。古籍数字化是当代技术进步为古籍整理工作提供的一个新途径,我们可以利用现代技术手段,对古籍进行整理和挖掘,发现其中对现代社会的发展有益的部分,为世界人民的和谐共生服务。文化的传承与保护只有与技术的进步保持同步,才能世代相传。〔5〕

3.3 古籍数字化为学术研究开辟了一条新途径

具体而言,古籍数字化改变了学者查阅图书的方式。学者可以用移动硬盘或类似的便携式储存器随身携带一个中型图书馆,可以随时随地查阅互联网上丰富的资源这样的阅读方式具有划时代的意义。眼界的极大开阔必然带来学术研究方法和研究范式的进步,进而推动科学研究的发展。古籍数字化还可以帮助学者排除冗余信息的干扰,更快地深入研究对象和研究资料的内部借助便捷的检索技术,学者们可以在最短的时间内找到所需要的信息,而无需因阅读那些与研究课题无关的书籍而浪费大量时间和精力。〔6〕

通过古籍数字化,建设大型古籍数据库,提供贴近现代学术的典籍资源,推进传统文化的现代化进程,为学术研究提供新的动力。将古籍原典、历朝历代的整理成果数字化,形成一个庞大精深的立体知识体系,使数据库建设成为连续性的、开放式的、与文化遗产和学术研究息息相关的一项工作,实现海量信息查询、字词频分析,用字量统计等古人难以实现的功能,为学术研究拓展一片新领域。

通过数字化工程,为全球华人,为学术界提供良工利器,提供更贴近现代学术的典籍资源,与学术界互动,推进传统文化的现代化进程。通过对传统文化的深入了解与接触,能够增强我们的民族自尊心和对本族文化的自信心,取其精华弃其糟粕,弥合文化的断层,增强民族凝聚力,使中国传统文化获得健康的发展,推进社会的稳定与和谐进步。

3.4 古籍数字化为中国传统文化走向世界提供了必要条件

随着互联网技术的发展,网络传播不受时空限制、信息量大、传播速度快的特点,使数字信息成为最重要的文化资源。互联网上的数字图书馆,将使人类的精神文化宝库跨越时空,在全人类面前实现资源共享。中华古籍数字化,让古老的中国文化借助现代信息技术再现辉煌,让全世界人们感受到东方文化恒久弥新的魅

力,也可以使全球范围的炎黄子孙能够方便、快捷地学习了解本民族的文化。^[7]

3.5 古籍数字化将使国际性合作项目成为未来发展目标

当今社会,只有数字化产品才能迅速、准确地进行传递和交流,为互联网信息平台提供有价值的信息资源,使传统文化在信息传播的深度、广度、速度上有一个质的飞跃。所以古籍数字化已不再是一个国内的小项目,我们完全可以进行政府、各机构组织及国内、国际间的通力合作,这对解决古籍数字化的资金、技术瓶颈无疑是有益的,同时使我国数千年来宝贵文化遗产进一步的发扬光大,耀眼世界。^[8]

五千年的存世古籍,负载着厚重的中华文明,凝聚着民族智慧,是祖先留给我们的一笔庞大的精神遗产。传承文明是时代赋予我们每一个古籍工作者义不容辞的责任,古籍数字化任重而道远。

参考文献:

[1]毛建军.古籍数字化的概念与内涵[J].图书馆理论

与实践,2007(4):82-83.

[2]孟忻.古籍数字化的现状与发展方向[J].中国索引,2008(1):39.

[3]毛建军.欧美地区中文古籍数字化概述[J].数字与缩微影像,2008(1):36-37.

[4]黄显功.上海图书馆的家谱数字化资源服务[J].图书馆学刊,2008(1):4-6.

[5][7]尹小林.古籍数字化的使命与前景[EB/OL].[2008-05-19].<http://www.douban.com/group/topic/3230116/>.

[6]郑永晓.古籍数字化对学术的影响及其发展方向[EB/OL].[2006-11-05].<http://www.literature.org.cn/article.asp?ID=11591>.

[8]郝淑东等.古籍数字化的发展概述[J].情报探索,2007(7):115-116.

作者简介:刘伟红(1969-),女,哈尔滨师范大学图书馆副研究员,研究方向:古籍整理。

(上接第93页) 动创新的要求比较高。只有俱乐部活动内容有特色、具有吸引力,才能吸引会员更多参加。

5.4 开展丰富的个性化服务

个性化服务是图书馆在网络和数字信息环境中,提高服务质量,赢得发展机遇的重要选择,个性化服务也是俱乐部生存的必要条件。图书馆个性化服务是指图书馆在数字信息环境中,主要利用网络和信息技术,获取并分析各用户的信息使用习惯、偏好、背景和要求,从而为用户提供充分满足其个体信息需要的一种集成性信息服务。图书馆个性化服务与个别服务或重点服务、定题服务、定制服务、特色服务都有一定的区别,它有不同于图书馆其他服务的多个特点。为会员量身订做个性化服务,给会员提供各种方便,满足会员被尊重的需求,显示其社会地位,这就是会员加入俱乐部的目的。而“书苑天香”俱乐部目前在这方面做得比较弱。笔者建议:(1)从会员入会之日起,就要做个调查,他参加俱乐部是希望得到什么样的服务。(2)为会员提供个性化信息服务。个性化信息服务是根据用户个人的具体需要,提供高质量的、系统的信息服务。图书馆可以调查会员的信息需求,对各种信息资源进行过滤,得到会员所需要的、个性化的信息资源精品。然后利用电子邮件等方式传送给会员。这种服务是持续性的,要不断跟踪相关课题的最新动态,定期向会员通报。同时,还要跟踪会员的使用情况,分析其信息需求的变化,优化检索过程和检索结果,逐渐形成会员个人的信息资源系统,它不仅限于为会员找信息,更倾向于为信息找会员,因此更具有个性化特征,也更容易为会员接受。

5.5 加大宣传力度

多年的图书馆工作实践告诉我们,要做好图书馆宣传工作并不是一件轻而易举的事情,需要我们付出艰辛的劳动。要让社会各界都来关心和支持“书苑天香”俱乐部,使社会上更多的人了解俱乐部、加入俱乐部,不仅需要我们脚踏实地做好各项服务工作,更需要我们做好宣传工作。(1)利用社会各个方面、各种机会对俱乐部进行宣传,宣传方式也应该是多维度进行,既有宣传册、广告版,也有网络、报纸等利用新闻媒体进行宣传。(2)利用新闻媒体和图书馆网站及时报道俱乐部开展的活动和取得的成绩,扩大俱乐部在社会的影响力。(3)充分发挥俱乐部会员的宣传作用。通过俱乐部会员现身说法宣传俱乐部,会具有更大的说服力。

为企业提供信息服务、支持地方经济发展,是图书馆工作的重要职能之一。“书苑天香”俱乐部是图书馆企业信息服务中的创新之举,可以起到为企业提供有针对性的信息服务,拉近企业与图书馆的距离,培养企业的社会责任感,利用企业的资源为图书馆发展提供资助等等作用,因此,“书苑天香”既是图书馆重要的服务延伸,又能充分实现合作共赢。

参考文献:

[1]中共苏州市委宣传文化工作领导小组.“文化苏州”行动计划.2004年6月.

[2]查磊.引入民间资本,完善苏州中小企业融资市场基础建设[EB/OL].[2009-01-16].<http://www.qnsh.com.cn>.

作者简介:蒋蕾静(1978-),女,苏州图书馆情报部馆员。