

# 简帛文献数字化述论

□ 河南新乡学院文学院 毛建军

**摘 要** 简帛文献数字化是简帛学发展的必然选择。简帛文献数字化以其特有的资源特点必将给简帛学研究和中国文化的传播带来全新的机遇。文章调查和分析国内外简帛文献全文数据库的建设情况,并给出思考和建议。

**关键词** 简帛 简帛全文数据库 古籍数字化

随着计算机技术的发展,简帛文献的研究手段和实践有了新的突破和发展趋势。其中尤以计算机技术与简帛文献研究的结合最为突出。计算机巨大的储存空间和强大的数据处理能力,给简帛文献的数字化带来了全新的机遇,简帛文献的保存和利用也迎来了数字化时代。简帛文献数字化就是采用计算机技术,将常见的文字或图形符号转化为数字符号的过程。简帛文献数字化是简帛学发展的必然选择。简帛文献数字化以其特有的资源特点必将给简帛学研究和中国文化的传播带来全新的机遇。其重要性和意义在于:一是,有利于简帛文献的抢救和保护。由于种种原因,我国出土的简帛文献正处于濒危阶段。简帛文献数据库以数字化的形式存贮和利用,是抢救简帛文献的最好形式。二是,有利于全方位开发利用简帛文献。简帛文献全文数据库是一个存储在一起的相关数据的集合,具有高度数据集成的特征。简帛文献全文数据库实现了检索自动化,可千百倍地提高检索速度。简帛文献的释文、分类、出处、关键词、摘要等都可成为检索区段,实现全文检索。三是,有利于促进古籍文献资源共建共享的形成。随着国际互联网的迅速发展,简帛文献实现了网络共享,可以在任意时间、任意空间传播,数字化简帛文献的网络共享大大提高了简帛文献的利用效率。

在短短的20年间,国内外利用计算机技术编制了大量简帛文献数字化资源。对简帛文献数字化资源作调查分析,既有利于指导下一步的简帛文献数字化工作,促进简帛文献数字资源的整合,又有利于社会各界充分利用简帛文献数字资源,充分挖掘简帛文献信息。图书馆和研究机构是简帛收藏和研究的主要机构,负有对简帛的保存、利用和研究的重要职责,因此,也是简帛文献数字化资源建设的主要单位。

日本收藏有大量的中文古籍,随着古籍数字化进程的开展,日本各大科研机构也开始将简帛文献数字化,并提供网络化服务。其中在简帛文献数字化资源的开发与建设方面,以早稻田大学和奈良文化财研究所最具成就。楚简数据库:早稻田大学文学部工藤研究室开发。数据库由简牍资料库和楚系文字资料库组成,内容包含楚帛书、鄂君启节、天星观楚简、信阳楚墓竹简、包山楚简、郭店楚简、曾侯乙墓竹简、望山楚简等中国南方9个地点新出土的帛书和竹简。具备良好的检索功能,使用者可选择特定文本与篇章或全部文献进行检索。

web 版本简帛数据库:奈良文化财研究所制作。奈良文化财研究所是日本文化遗产综合研究的机构,主要工作包括发掘遗迹、文物及考古科学的研究等。其中尤以木简典藏最为突出。奈良文化财研究所网站提供的数据库包括木简数据库、墨书土器数据库、文书数据库、遗迹数据库等。木简数据库可提供浏览检索、文字检索和全文影像阅读。

目前,港台地区已开发出许多简帛文献数字化资源,其开发建设单位主要为港台地区的研究机构和部分数字公司。其中以香港中文大学中国文化研究所古文献数据库中心开发制作的汉达古籍数据库检索系统——竹简帛书出土文献数据库和台湾“中研院”历史语言研究所开发制作的简帛数据库最具特色。汉达古籍数据库检索系统——竹简帛书出土文献数据库,香港中文大学中国文化研究所古文献数据库中心开发制作。系统包括甲骨文数据库、竹简帛书数据库、金文数据库、先秦两汉文献数据库、魏晋南北朝文献数据库以及类书数据库等。汉达竹简帛书出土文献数据库已收录十二种竹简帛书出土文献,约近140万字,数据库提供单字、字符串及句式检索。文献检索结果可以同时显示简帛图片和对照释文。

郭店楚简数据库:由香港中文大学图书馆与香港中文大学中国语言及文学系张光裕教授共同制作。郭店楚简又称郭店楚墓竹简,是中国湖北省沙阳县纪山镇郭店一号楚墓内的竹简,竹简年代属战国中期偏晚。郭店楚简共804枚,其中有字的竹简有726枚,13,000余字,内容包括《老子》、《太一生水》、《缁衣》、《五行》、《尊德义》等十六篇道家及儒家著作。郭店楚简数据库可按竹简编号、篇目名称或相关字词检索,支持全文检索。

简帛金石数据库:台湾“中研院”历史语言研究所开发制作。历史语言研究所文物收藏非常丰富。历史语言研究所数据库处理对象主要包括秦简、楚简、金文、汉印、汉碑、汉镜、汉砖、墓券、墓牒等。简帛金石数据库目前已完成,收录60余种简帛金石资料及研究书目共3,401,684字的纯文字数据,以及居延汉简补编(计1,324简)等图文资料。系统支持布尔运算、多项目查询、词组查询、排除字符及自然语言查询。

先秦金文简牍词汇数据库:台湾“中研院”历史语言研究所开发制作,包括金文和简牍两个子数据库。简牍词汇数据库包括:词汇库查询、词汇库索引以及词汇库管理。词汇库查询可以自行输入关键词进行检索。中国大陆地区是简帛文献收藏最为丰富的地区。由于计算机技术发展略滞后于日本和港台地区,在简帛文献数据库开发建设上起步较晚,但后来居上。目前,中国大陆地区的简帛文献数据库的建设已取得丰硕成果。中国大陆地区参与简帛文献数据库开发建设的单位主要为科研机构以及数字化出版企业。

战国楚简帛文字典型形体检索系统:华东师范大学中国文字研究与应用中心研制。为顺应中国文字数字化发展趋势,华东师范大学中国文字研究与应用中心构筑了规模庞大并富于补缺中文信息化现状特色的中国文字数字资源,主要包括商周金文检索、战国楚简帛文字典型形体检索系统、花园庄东地甲骨检索等数据库系统。战国楚简帛文字典型形体检索系统所收录的材料范围包括《包山楚简》、《郭店楚墓竹简》、《九店楚简》、《曾侯乙墓》、《战国楚竹简》

# “从档”的艺术——和覃兆别教授《档案事业“三字经”》

□ 上海理工大学 章华明 于 喆

**摘 要** 覃兆别教授《档案事业“三字经”》一文是对档案事业的宏观解读。本文则从基层档案工作的具体实践出发,就如何做好档案工作提出了自己的见解,并将其归纳为四句话、二十个字。

**关键词** 档案工作 方法 艺术

随着“档人”<sup>[1]</sup>整体素质的逐步提升,越来越多的档案机构升室为馆,档案事业的地位也在逐渐提高,有力地推进了整个社会文化事业的发展。但我们又不能过于乐观,毕竟我国档案事业总体上尚处在发展期。即便在高校系统,教育部办公厅档案处处长蒯海波也不得不撰文呼唤档案工作“从边缘走向中心”<sup>[2]</sup>。在这样的条件下,要做好档案工作,一定要注重策划,注意方式方法,研究“从档”艺术。读了湖北大学覃兆别教授《档案事业“三字经”》一文,深以为是。但从基层档案工作者的角度,考虑到文章学术性较强,用典较多,普通读者难以领会其要旨,笔者试图从实践出发,结合自身体会,将从事档案工作的“艺术”概括为四句话、二十个字,以为对覃文的注解,希望对读者有所启发和帮助。

## 一、名人大事,做足文章

任何一个单位都有历史,都有自己的名人和曾经发生过的重大事件。以高校为例,就有历任主要领导、知名教授、杰出校友等名人,也都有所谓的大事要事,或多或少。这些都是档案工作的重要对象。另一方面,单位现任领导所关注的事情,无论是历史上已经发生的,还是正在发生的事情,多半是单位的中心工作或者重点工作,是大事,也是“档人”关注的重点,这就为“档人”在名人大事上做文章提

供了前提和可能。“档人”当积极搭台,为单位领导参与档案工作创造条件。长此以往,档案工作就有可能因为领导的关心、重视,走上健康发展的轨道。但如何实践呢?首先,“档人”要尽可能地对现有线索进行梳理,顺藤摸瓜,广泛收集资料、采访和调研。在此基础上,条件成熟的写成独立的文章,投稿至《档案春秋》、《上海滩》、《世纪》等类似风格的历史文化期刊或设有相关栏目的档案期刊。此外,要联系实际,结合重大时间节点,巧用档案,以收集的素材为基础,组织相关的座谈会、研讨会,并邀请相关人物参加。这样一来,活动的影响力就大了,档案工作的重要性就得到了表现和说明的机会。如果能形成机制,有序地向前推进,档案工作的环境就会得到逐步改善,客观上也就起到了服务中心的作用,进而融入中心。当然,这些素材也可成为相关企业史、校史教育讲座的内容,还可以整合在网站上加以宣传,甚至在适当的时候推出专题展览。也就是说,遇到这样的素材,不能浅尝辄止,蜻蜓点水,相反一定要穷尽所有,一网打尽,并且不断变化组合,将名人大事的效应“利用”到极致,发挥到极致。这样,名人才能更“有名”,大事才能更“大”,“名人大事”也才能为档案事业服务,为档案事业呐喊助威。

## 二、审时度势,以研促管

编》、《上海博物馆藏战国楚竹书》以及《葛陵新蔡楚墓》等材料。系统目前提供 GBK 编码范围内通用字头的检索。

北京时代瀚堂科技有限公司制作的瀚堂出土文献库含甲骨文、金文、简帛、印章、钱币和石刻等分库,是目前唯一的集成性的中国出土实物文献数据库。数据库图片清晰,可全文检索,支持复制、打印功能,并附有联机字典。此外,随着互联网的普及,国内各大博物馆还开通了馆藏精品数据库查询,如国家博物馆馆藏精品查询,故宫博物院馆藏精品查询均可进行甲骨文、铭刻、简帛、法书、碑帖等图谱的检索查询,并附有图谱在线功能。从这些精品在线查询所具备的检索功能来看,也应视为简帛文献数据库。目前,全国已经有 60 多家博物馆建立了网站,有的还设置了专门的数字博物馆,读者只需注册,即可阅读或下载大量简帛文献资料。

简帛文献数字化是对已存文献的再现和加工,是 21 世纪简帛文献整理和研究的发展方向。尽管简帛文献数字化资源的建设已经取得了相当大的成绩,但还存在一些值得思考的问题。一是资源重复建设问题。由于目前还没有一个统一的组织和机构去管理和协调简帛文献数字化资源的建设,加之简帛收藏单位和研究单位又十分复杂,因此,国内外重复开发和建设了大小不一的数据库,造成巨大的人力、物力和财力的浪费。这种重复开发不仅给开发单位带来了巨大损失,而且也增加了读者检索和利用上的困难。简帛文献数字化资源的建设是一项系统工程,需要一个全国性协调机构或组织来

全面实施数据库的规划,以避免重复开发和资源浪费。二是数据库兼容问题。由于古籍数字化制作的主体和各自的利益诉求不同,所制作的古籍数据库常常是封闭的,在技术上很难与其他数据库融为一体。至今简帛文献数据库还没有一个统一的数据库著录和运行系统标准。简帛文献数字化资源建设必须依照相应的标准和规范进行建设。规范化的简帛文献数字化资源必须是在相同软件支持下做出的在格式、内容、标引、著录以及字体等方面都按国家标准做出的一致的数据库。三是资源共享问题。简帛文献数字化资源的利用在资源共享上还存在资源封闭和利益保护的弊端。从目前已建成的简帛文献数字化资源来看,在使用方面设立了许多障碍,如限制局域网使用,给读者带来了诸多不便。因此,实现简帛文献数字化资源的共享势在必行。

需要指出的是,目前简帛文献数字化资源建设最迫切的任务是做好资源整合工作。从已开发的简帛文献全文数据库来看,简帛文献数字化资源已相当可观。资源的剧增问题必然导致资源查找不便和难以系统呈现问题,因此需要对于网络上的数字化资源进行导航。因此,整合利用好这些资源是件很重要的工作。

## 参考文献

- [1]毛建军.古籍数字化与人物图像数据库资源建设[J].档案与建设,2009(1).
  - [2]赵保英,苏沫.论古籍数字化资源的协调建设[J].兰台世界,2010(8).
  - [3]童顺荣.古籍数字化相关问题的开放思考[J].兰台世界,2009(18).
- ★基金项目 2010 年教育部人文社会科学研究青年基金项目《古籍电子文献研究》(项目编号:10YJC870027). (作者毛建军,新乡学院文学院副教授,博士。)