

文献信息化一瞥

刘 振 华

计算机技术的飞速发展是有目共睹的，电脑的普及已成为不可阻挡的大趋势。如果说几年前，大家经历了为电脑的购置、操作以及硬件的不断升级而辛苦奔忙之后，可以说今天我们已经有了足够等级的计算机硬件、越来越完善的软件系统和汉字平台了。由此引发了新的问题，就是我们用电脑来做什么？以各种形式出现的信息成为当今电脑界最为热门的话题。“信息社会”这个全新的概念，逐渐在全社会渗透开来。从信息的分类来看，它大致可分为现行信息和历史信息两个基本门类。所谓现行的信息，指当今社会上所发生的政治、经济、文化和艺术诸方面的新状况、新进展、新成就和新事件，它具有明显的趋时性和兼蓄性。所谓历史的信息，主要是指“文献”所涵盖的内容。历史信息较现行信息稳定并且数量巨大。美国的丹尼尔·伯斯丁先生（Daniel · Boston）出版了一本《感悟人生》，诉说了信息给人类生活带来的苦恼。他说，信息“把我们的头脑塞得满满的，把我们的知识和理解力挤了出去”，并进而得出结论，“书籍有长久价值”，只有“书能建立传统”等等（见内蒙人民出版社中文版）。显然，这位先生把新闻当成了信息，他的理解未免失之偏狭。广播、报纸、书籍，都可以传递信息，它们的区别，仅仅是不同的手段，不同的载体而已。关键不在形式，而在内容。信息应该是一个合乎实际的总称。人们不应该为所谓的信息爆炸而苦恼，应该苦恼的是如何更加有效地存储信息和利用信息。因此，信息作为一种客观存

在，不管它的形态如何，都面临着一个电脑化、数字化的前途，都将归于“0”和“1”。这个前途就是“信息化”过程。

显然，只拥有高级别的电脑，不等于搭上了信息快车，更不能认为自然地驶入了信息社会。快速的跑车，跑上高速路，是不是就算立身于信息社会了呢？显然不是。二十一世纪将进入信息社会，信息高速路已经应运而生。经历了近几年的运作，有识者发现，空车空路，谁跑得快，谁的损失就会越大。这样的高速车路还有什么意义呢？可以肯定地说信息社会的支柱是“车”（即电脑）和“路”（即信息高速公路、网络），其基础便是“货”——现行信息和历史信息的总和，而将这些天文数字的信息有机地组合起来的过程便是信息化。归根到底，社会的信息化运作一天比一天显得更加紧迫、更加重要了。我国作为一个文明古国，历史文献的信息化便显得越来越迫切、越来越重要。

其实很早就已经有有识之士在积极开展这方面工作了。像北京玮云科技发展公司制作的《古典文学大宝库》，其中包括了《全唐诗》、《水浒》、《红楼梦》、《三国演义》、《西游记》、“三言二拍”等，既有音乐，又有图画的光盘。东森电子科技有限公司出品的中国古典四大名著《水浒》、《红楼梦》、《西游记》、《三国演义》，里面配置了大量精致的彩色水墨国画、人物介绍及原著的改编文字数十万言和生动的旁白。此外，又有大百科出版社出版、北大方正集团公司多媒体部制作的《中国大百科全书》以及《中国烹饪》，北京大学出版社出版，北大方正集团制作的《龙的故乡——中华名胜》以及《孙子兵法》等等。其中都有许多珍贵的图形、优美的配乐和解说等，资料相当丰富。《唐诗三百首》更是人们关注的焦点，不断有不同版本的光盘问世，配乐、动画、朗读以及启迪心智的再创作等等，很是赏心悦目。尽管它们多为选本，也可能有这样那样的缺憾，但毕竟是一个在电脑上普及古典文献的良好开端。这些古代典籍光盘的面世，在很大程度上促进了我国电子出版物的健康发展，繁荣了我国多媒体电子市场，这个良好开

端，令人欣慰。但是，有些情况又不容乐观，主要是盗版现象风起，严重干扰了文献数据的开发和发展。其次恐怕要从已制作完成的产品上找原因了。上面所提到的这些电子出版物，大部分还属于一般性读物，主要是供一般读者观赏所用，有的只是一个简单平面的设计，无法让更多的读者对中国古代典籍有更深入、更准确的了解和更加便捷的利用。当然，就更谈不上成为研究者的工具，也就难怪会受到“中看不中用”的批评了（见1997年8月28日《光明日报》）。

此后《人民日报》五十年的光盘上市，它以该报的完全的版面资料见长，并配以每篇文章的篇名、作者索引，可以基本满足收藏者和读者一般查询的需求。美中不足的是，尚不能提供全方位的检索功能。《大百科全书》光盘，以其极有价值的全面信息资料取胜，但对全方位乃至触发式的参考检索方式，使用者会有更高的要求。至于怎样使我国的古典文献及时妥善地深入计算机，将成为我国古老文化传统发扬光大的当务之急。对此，我国的计算机界和文献界，理应当仁不让。要进行这项事业，首先要有统一规划，不能盲目重复上马。在统一规划的前提下，仍须审慎选题、精细制作，并留有充分的升级余地。对历史文献的保存和使用，是这个总题目下的两个基本目标。可喜的是，山东汇文科技开发中心的《四库全书》电子版已有部分问世。它以三万多种古籍为目标，逐步扫描，以图形方式输入计算机，并设置了基本的检索功能。如果能在制作过程中，逐步增强检索功能，并吸取其他版本《四库全书》的长处，它将成为该书的最佳版本。最近，欣闻商务印书馆国际有限公司出版了中国广澳开发集团公司中文电脑数据中心开发研制的《百衲本二十四史》，笔者有幸了解了这部电子版古籍的制作情况，很受鼓舞。

首先这部电子版的《二十四史》的底本选择极为精良。他们所用的“百衲本”，是一部举世公认的古籍善本，是本世纪我国出版界的重要成果之一。这部将近十四万页面的浩大通史，全部以

高密度扫描送入计算机，不仅起到了保存古籍的作用，而且还可以让众多的读者方便地看到这些稀世珍本的原貌。“只收入代表我国璀璨文明的善本珍本；非实用完本，不足以供研讨、阅读、查验、宝藏之用者，一概不予收录。”这便是他们选择版本的原则。这在古典文献方面，无疑是正确的。

这部电子书容量大，文字数据准确完整。它汲取了此前科技研究的优秀成果，根据原书的不同情况，以全新的索引查验系统为经络，用大量优秀的图像和文字数据构成。作为技术关键的诸多索引，像人名、地名、书名，都与读者有友好的界面，有周到完整的设计，而且又精确、快速，简单易行，读者使用起来犹如轻车捷径。这套光盘的出版，说明古典文献的信息化，可以涵盖传统书籍的诸多优势，可以满足使用者的一般阅读、欣赏、科学研究和珍藏传世的多方面要求，同影印本、排印本相比，具有不可比拟的优越性。它不仅大大提高图书的品位和使用效率，而且将改变已延续了千百年的“书”和“读”的概念，改变书籍编辑、印刷和出版的传统格局。变化之大、进步之快，让人高兴。

更重要的是，在阅读和检索这部《二十四史》时，读者可以不懂、不会、不学那些令人头痛的中文输入法。换句话说，使用这部《二十四史》，根本就不必输入汉字，可以不具备高深的电脑操作技术，只需使用鼠标，在图面上轻轻一点，就会出现《二十四史》某卷某页的原貌，在检索人名、地名时，还可以同时得到标点本的相应位置。至于界面的放大或缩小以及原件的复制，则悉听尊便。随处设置的“帮助”功能，是协助读者顺利阅读和利用这套光盘的好帮手。作为读者，只有亲手操作，身临其境，才能体会到这种意想不到的便捷，才能领略研制者们的精心设计和良苦用心。

据说，这项成果的开发研制，由多名富有学养的老专家和有志此道的年轻人完成，中国广澳开发集团公司一直充满信心地支持着这项事业的发展。经过他们多年的苦心钻研，已经把一些著

名的古代典籍，区别各种不同的情况，逐步妥善地送入了计算机。

平心而论，上述这些项目，既是一种尝试，也是一种创举。笔者希望，这部经多年研究的优秀成果的问世，能够为历史文献的信息化提供某种启迪和参考，能够得到广大读者的品评和认可。

据悉，文献信息化的工作，资金投入数额巨大，回收速度缓慢，现行体制又是归入出版社出版的程序，据多数制作单位估算，发行单位占有销售份额的 35%，出版单位占有销售份额的 35%，制作单位仅占 30%，其中包括底本费、稿费、软件制作等等。可以说出力最大、出钱最多者，所承担的风险也最大，而获利却微乎其微。这还没有考虑盗版的猖獗和产品面世后的技术咨询、售后服务以及升级换代等问题。这种状况如不及时改变，那么信息化产业的再生产将滑入恶性循环，后果也是可想而知的。

文献的信息化将有一个美好的前景，也是充满了各种困难的新的边缘学科。对于信息社会，这部分数据占有它特殊的不可或缺的位置。只有现在和未来的民族，则是一个缺乏文化和精神的人群。有了传统，那可就不一样了。