

朱 岩

## 《四库全书》电子版问世的启迪

**摘 要** 1999年春,《文渊阁四库全书》电子版问世。这是中国传统文化信息资源实现数字化取得突破性进展的重要标志。它将推动中国电子出版物的出版和电子图书馆的建立与发展,并积累了信息资源拥有者、资金拥有者和高新技术拥有者三者结合发展的经验。表1。

**关键词** 四库全书 电子版 数字化 电子图书馆

**分类号** G255

**ABSTRACT** In Spring 1999, an electronic version of *Si Ku Quan Shu* (*Complete Library of the Four Branches of Literature*) was published. It is a landmark of digitisation of information resources of traditional Chinese culture and will give impetus to the publishing of electronic publications and the creation and development of electronic libraries in China. Besides, it is also a model integrating information resources owners, capital investors and high-tech developers. Table 1.

**KEY WORDS** *Si Ku Quan Shu* (*Complete Library of the four Branches of Literature*). Electronic version. Digitisation. Electronic library.

**CLASS NUMBER** G255

1999年春,《文渊阁四库全书》电子版(原文与标题检索版)问世,并装载在国家图书馆的网上向读者提供服务。原文与全文检索版年内也将完成。这不仅是我国出版界由传统印刷出版向现代化电子出版方式转变的一件大事,而且是我国传统文化信息资源实现数字化参与世界范围共享取得突破性进展的重要标志。

### 1 电子版不是纸张版翻印

电子版《四库全书》以当代最先进的信息技术,在保持原书真迹的基础上,将书中一切具有检索意义的书名、作者、类目、标题乃至全文中的字、词、语全部数字化,同时还挂接了扩充的《中华古汉语词典》《四库大辞典》《四库全书简明目录》等电子工具书,从而给读者提供了手工查阅难以比拟或从未有过的快捷有效的检索、统计、整理和编辑功能,有利于读者开展更广泛、更深入的研究。

一部《四库全书》原书大约需要 $500\text{m}^2$ 以上的存放空间,影印本也要 $100\text{m}^2$ 空间,而电子版只需

手提箱大小的机箱(百G硬盘)就可存储。过去,一部书在一个时间里只能供一个读者阅读,现在电子版《四库》通过网络可供多个读者同时使用。电子版实现了书名、著者、类目和全文中的字、词、语等多途径检索。读者可以进行单项检索,也可以多项信息组配检索。检索时,可以输入检索的全部信息,也可以输入其中部分信息,即“任意一致”检索。如果读者对输入汉字的一些方法不熟悉,还可以选用电脑提示的拼音、或笔画、或部首3种不用学习就可掌握的简单方法即可确定检索文字。查询效率也是传统方法无法比拟的。一般情况下,电子版一次检索均可在5秒内获得结果。阅读时,可随意在电脑上作阅读笔记,可摘录检索到的有关内容,或编排下载、打印。多种电子工具书又与《四库》挂接,给读者带来更多方便。例如,你对某些字义不明确,可以点击“古汉语字典”;对某个作者或某部作品不熟悉,可点击《四库大词典》,或《四库简明目录》,有关的解释或说明会立即显示其旁。如果对解释不明白,还可从解释说明中的书名、作者再去查检《四库》。

过去,人们查阅图书,一般是先通过目录找

书,查到后再查书中目次,据目次直接顺序翻阅。而电子版《四库》,读者不仅可以在瞬间查到所要图书,而且可以直接深入文献中查到某局部。不只如此,还可以把与一个或与多个主题相关的专著、专著中与主题相关的篇章、段落查到。现代检索技术,不仅能在一部文献的纵深方向搜寻,也可以在多部文献之间横向搜寻,并进行精确的统计、整理。这种能力极大地启发和扩展了学者研究的广度与深度,帮助其快出成果、多出成果。

表1显示的是利用大家熟悉的词语在《四库全书》全文版经部查检的结果,这些精确的结果都是在瞬间完成的,是过去手工查询办不到的。

表1 检索事例(经部全文检索)

检索词语	检索命中处数
法 治	322
法 制	598
生 态	0
弹 劾	6
经 济	137
革 命	561
炎 黄	12
书 同 文	55
车 同 轨	54
李白与杜甫(一卷内文中涉及二者)	6
荆轲与高渐离(一卷内文中涉及二者)	9
风马牛不相及	32
鱼和熊掌不可兼得	91

综上所述,《四库全书》电子版不是纸张本简单的翻版,而是具有丰富研究功能与工具的中华优秀传统文化信息宝库。

## 2 有利于对中华文化的认知与研究

《四库全书》是迄今为止人类历史上规模最大的一部百科全书,它对我国古代文化主体的儒家学说以及历史、人文、科技、众多学科的历代典籍进行了成功地整理与总结,它是具有重大历史价值的中华传统文化宝库。

《四库全书》电子版的出版、发行,使《四库》从过去广大读者难以利用的“镇馆之宝”走向百姓的

“案头”,使我国成系列、大规模、具有代表性的文献资源数字化,并可进入因特网,参与全球信息资源共享。这不仅有利于更多的国内外研究中国的学者、专家开展更深入、系统、有效研究,也有利于更多的人利用现代化手段学习中国古籍,了解中国传统文化。

《四库全书》成书以来,学术界对清廷的编纂,对《四库》本身一直有不同的评价与争议,有人称之为“四库学”。《四库》电子版的问世,能使更多的人接触到此书,因而可以使更多的学者参与研究。而这种研究可以借助电子版进行准确、充分的定量统计与分析,因而才有可能获得切合实际的定性结论。像清廷在纂修《四库》时对一些典籍的删、改、取、舍,以及由此体现的清人文化史观,读者利用电子版则可能得到准确的统计数据,予以恰当评价。

学者研究的课题也绝不仅限于《四库》本身,电子版可以帮助学者开辟更多更广的专题研究。

## 3 中文信息处理技术的重要突破

《文渊阁四库全书》电子版成功出版,标志着中文信息处理技术实用化有了重大突破。该电子版是科技含量极高的产品。如:国际标准大字符集(ISO/IEC 10646 UCS,亦称UNICODE,含3万多个中、日、韩古今汉字)文字平台的投入使用,基于超大量和种类繁多的多特定人手写文字光学识别、编码转换、联机无纸快速准确的校对软件技术等,均属于首创。跨平台与跨语境(Single Data/Single Binary)全球版产品制作技术,多途径模糊与精确以及布尔逻辑组配的全文检索技术,全文与多种字典、词典工具书挂接以及多种文字关联查询的规范控制等一系列高新技术,正是该电子版的研究开发使其完善与实用,同时还带出一批新的研究课题。这些突破不仅意味着我国近代文献可以实施计算机处理,而且意味着古代文献也可以实施计算机处理,这必将加快我国中文信息数字化进程。

## 4 推动我国电子图书馆的建立与发展

《四库全书》电子版研制成功,将推动我国电子出版物的出版,加速我国电子图书馆的建立与

许 平

## 试论邓小平理论对图书馆事业的指导意义

**摘 要** 学习邓小平理论,要从科教兴国的高度认识发展图书馆事业的重要性和紧迫性,把“面向现代化,面向世界,面向未来”作为图书馆事业改革和发展的战略指导方针,把图书馆事业的改革和发展建立在科学的基础上。参考文献 14。

**关键词** 邓小平理论 图书馆 改革和发展

**分类号** G258.42

**ABSTRACT** The author thinks that we should use Deng's speech "Towards modernisation, towards the world, towards the future" as a guideline for the reform and development of Chinese librarianship. 14 refs

**KEY WORDS** Deng Xiaoping's Theory. Library. Reform and development

**CLASS NUMBER** G258.42

江泽民同志在党的十五大报告中指出:“在社会主义改革开放和现代化建设的新时期,在跨越世纪的新征途上,一定要高举邓小平理论的伟大旗帜,用邓小平理论来指导我们整个事业和各项工作。”<sup>[1]</sup>在图书信息战线贯彻十五大精神,就是要用邓小平理论来指导图书馆事业。笔者通过学习邓小平理论的基本原理和他关于教育、科技、文化的许多精辟论断,联系图书馆改革和发展的实际,深深体会到邓小平理论对于图书馆事业具有重要的指导意义。

发展。

在我国,电子图书馆的建设与发展日益引起重视。这件事涉及两个方面:一方面是电子图书馆所需的网络通信环境、计算机软硬件手段的发展与配备。另一方面也是最重要的一个方面,即信息资源的数字化,而这也是最费时费力的。《四库全书》电子版的问世恰恰带来了一个好的开头。因为它实现的信息量最大,所处理的文字包括图形种类多而且复杂,它的OCR技术具有代表性而且解决问题的难度大、水平高,这些技术难关的攻破为更多的单位开发信息资源扫清了障碍。

我以为,更可贵的是《四库》电子版研制运作过程积累的经验:信息资源拥有者、资金拥有者、高新

### 1 学习邓小平同志关于发展科技和教育的论述,从科教兴国的高度深刻认识发展图书馆事业的重要性和紧迫性

在《邓小平文选》第二、三卷中,有许多的内容是谈科技和教育。实现社会主义现代化,科技是关键,教育是基础,这是邓小平同志关于科技和教育的核心思想。他指出:“科学技术是第一生产力。”<sup>[2]</sup>“四个现代化,关键是科学技术现代化。”<sup>[3]</sup>“科学技

术拥有者三者结合,走合作之路,加快我国信息资源数字化开发。这一经验十分重要。因为较长时间以来,常有一些人总想什么都自己干,什么都想自己先得到,其结果是或资金不足或技术不具备,迟迟干不出来。许多事实说明,手工作坊式的思维与方式是建设不了中国的数字图书馆的,只有社会化合作,资源与收益共享,才是成功之路。

朱 岩 国家图书馆副研究馆员,从事自动化、汉字信息处理研究工作。成果多次获奖。通讯地址:北京白石桥路39号。邮编100081。

(来稿时间:1999-05-11。编发者:翟凤岐)