

# 数图工程 开创我国文化的未来

[作者] 徐文伯

[单位] 文化部

[摘要] 本文介绍了数图工程,开创我国文化的未来,包括挑战与机遇:启动中国数字图书馆工程的大背景、抢占与盘活:发展中国数字图书馆工程的大思路、基础与优势:培育中国数字图书馆工程的大环境 3 个方面的内容。

[关键词] 数图工程

以信息高速公路和多媒体技术为标志的信息技术革命,正在把即将迈入 21 世纪的人类社会带入到高科技经济时代。为了迎接新时代的挑战,一个由文化部牵头,中国国家图书馆(北京图书馆)、中国电信总公司、中国科学院、航天工业总公司、广播影视信息网络中心、清华大学、北京大学等单位联手合作,诸多专家学者共同参与的宏大文化工程——中国数字图书馆工程,已在积极筹备之中。为什么要组织这一工程,怎样建设好这一工程,我有几点想法。

## 一、挑战与机遇:启动中国数字图书馆工程的大背景

中国数字图书馆工程,可以简称为数字图书馆。所谓数字化,就是利用现代信息技术对传统介质的图像、文字和声音进行压缩处理,使其转化为数字信息,然后,通过计算机技术进行贮存,通过网络通讯技术进行传播、接收。数字化不仅同光导纤维和计算机一起成为 20 世中期以来信息技术革命的三大支柱之一,而且也同网络化一起构成 21 世纪高科技经济时代的基本特征。网络和数据库建设,是新时代各国信息和文化产业必须进行的基本建设。由于图书馆是人类知识的宝库,也将是信息高速公路上最大的信息源,因而图书馆的数字化就成为首选目标。

所谓数字图书馆就是对有高度价值的图像、文本、语音、音响、影像、影视、软件和科学数据等多媒体信息进行收集,组织规范性的加工,进行高质量保存和管理,实施知识增值,并提供在广域网上高速横向跨库连接的电子存取服务。同时还包括知识产权、存取权限、数据安全管理等。它的特点是:收藏数字化,操作电脑化,传递网络化,信息存贮自由化,资源共享化和结构连接化。数字图书馆的建设,不仅给图书馆业带来一场革命,而且也文化传播打开了新的电子时代的大门。如同工业经济离不开交通和能源一样,数字图书馆也是高科技经济的基础设施和必要条件。知识经济是建立在知识和信息的生产、存贮、使用和消费之上的,数字图书馆所收藏的各类信息对于知识经济的整个过程都是必不可少的。而数字图书馆的数字化馆藏凭借高新技术,可以经济、快速地传播,方便地被人们所利用,从而不断地激发人们的想像力和创造力,推动全民族文化素质的不断提高。因此可以说,数字图书馆是今后经济和文化的。

正因为如此,数字图书馆的建设已经成为当前评价一个国家信息基础水平的重要标志,也是近年来发达国家十分重视的基础性项目。美国于 1991 年率先开始研究数字图书馆,其后,英国、法国、日本、德国、意大利等国也相继开始投入巨资开发本国的数字图书馆。据不完全统计,美、英、法、日四国,仅其国家图书馆所完成的数字馆藏,其存储量最少的约为 500GB,最多的已超过 3000GB。他们已为迎接 21 世纪信息时代的到来,完成了巨量的准备工作。

不仅如此,这些发达国家在开发本国数字图书馆的同时,还相继提出了本国整体文化信息网的建设计划。1993 年美国提出“国家信息基础结构”(NII)行动计划,继而又提出建设“全球信息基础设施(GII)的主张;1994 年欧盟宣布在欧洲建立信息社会的计划,确定了欧洲信息社会应用领域;法国若斯潘政府提出的《为法国进入信息化社会做好准备》的实现社会信息化政府行动纲领中,将文化信息网的建设列为 6 个优先发展的主题之一,其要点涉及到文化领域的各个方面;加拿大政府于 1996 年 5 月公布了《建设信息社会:使加拿大进入 21 世纪》的行动计划,其后又专门成立了“数字化工作小组”,统筹全国的数字化工作;英国政府于

1997 年提出了“全国学习网”计划,使全国的大专院校、图书馆、博物馆与“全国学习网”连通,从而扩大整个社会获取知识及接受教育的机会和途径,达到创造一个网络化“知识社会”的目标。特别值得我们注意的是,目前在亚洲使用中文的国家和地区,也已积极行动起来。

新加坡政府在 1994 年提出了“2000 年图书馆发展计划”,准备建立一个无边界电子图书馆网络,把全新加坡的公共图书馆和 500 多个学术与专业数据库连接起来。其后,又于 1996 年 3 月,成立了一个“华文国际网络指导委员会”,其目标是要使新加坡成为一个华文因特网的发展中心。他们认为,一旦华文因特网在世界上占有举足轻重的地位,新加坡便可以在已有充分准备的状态下进入这个广阔的领域;我国台湾省也于 1994 年制订了建设“亚太智能信息服务中心”的计划,该计划投资额上百亿美元,准备用 6 年时间完成。

综观各国数字图书馆建设情况,可以得出如下几点认识:

其一,数字图书馆将是 21 世纪全球文化科技竞争的焦点之一。这种竞争既是科学技术的竞争,也是文化和意识形态的竞争,更是知识经济时代的市场竞争。由于美国以信息产业带动经济高速发展已成为不争的事实,因此西方各主要发达国家及许多发展中国家也都纷纷制订自己的信息社会发展计划,以求得在未来的竞争中立于不败之地。在网络时代,谁最先掌握了技术和资源库,谁就掌握了先机。这就是启动中国数字图书馆工程的重要性之所在。

其二,数字图书馆工程不仅是高科技项目,也是跨部门、跨行业的大文化工程,必须由政府出面来统一规划、组织和协调,并在资金和政策方面给予支持和保障。例如在 1995 年美国政府的蓝皮书中,数字图书馆被认定为“国家级挑战”,置于国家信息基础设施(NII)的高度上通盘考虑。这种政策上的倾斜,带动了美国科学界、产业界的高度重视,也带动了各种基金会在资金上的投入。

其三,数字图书馆工程已经获得了可靠的技术保障和可观的效益前景。以美国为代表的数字图书馆的建立和运行,十分有效地获得了信息资源的增值效益,在资源建设和知识创新方面取得明显的进展。同时,这些国家也明确了数字图书馆的基本构造,技术手段和运行方式,开发了相关技术和设备,并取得了十分宝贵的工程经验。

其四,不仅首先要开发数字图书馆,而且还要在此基础上,陆续把其它国家级的文化信息资源单位,如图书馆、档案馆、博物馆、文化艺术、音像影视、新闻出版、旅游、体育等有关的文化信息资源的精华,发展为数字式资源库,并用这些丰富的信息资源构成图书信息资源网,通过因特网向全球传播。

其五,在文化传播的电子时代,作为中华文明发源地的中国大陆,并不一定理所当然地会成为全球最大的中文信息中心。建设中国数字图书馆工程,实际上也就是建设中文因特网。这对于我们继承和弘扬中华文化,力争在未来的全球性竞争中取得主动权具有重要的社会和经济意义。目前,已经有一些国家和地区在关注中文因特网的建设,如果我们不牢牢抓住机遇,就势必要在中文信息方面失去主导地位,从而丧失巨大的社会和经济利益。这也就是启动中国数字图书馆工程的紧迫性之所在。

开发数字图书馆工程不仅是世界潮流和信息技术革命所带来的挑战,也是在新的历史条

件下中华文化大发展千载难逢的历史契机。因此，建设中国数字图书馆工程，将会发挥以下积极作用：

一是为实施科教兴国战略和提高全民族素质提供强有力的文化基础支撑。数字图书馆工程将会根本改变我国文化信息资源保存、管理、传播、使用的传统方式和手段，克服我国文化信息资源得不到有效利用和共享的弊病，为知识创新和两个文明建设营造一个汲取文化信息的良好环境。特别对因信息不畅和文化落后而导致经济不发达的老、少、边、穷地区的发展，将产生巨大的推动作用。在信息时代，终身教育是提高公众素质的极有力的手段，而中国数字图书馆则会成为实施终身教育的大课堂。无论是儿童还是老人，也无论其身处何方，只要连通数字

图书馆的网络系统，都能方便地使用丰富多彩的文化信息资源。

二是可以用现代信息技术把中华民族的文化精华和当代文化教育、经济建设和科学研究成就等各方面的重大进展全面的展示出来，并通过因特网传送到世界各地，扩大中华文化在全世

界的影响，为人类的文明进步和发展作出应有的贡献。

三是中国数字图书馆工程是跨部门、跨行业和跨世纪的大型高新技术项目，它的启动必将带动相关产业，特别是信息产业和文化产业的蓬勃发展，并通过知识的有效传播，最终关联到各行各业，从而产生巨大的经济效益和社会效益。这也正是启动中国数字图书馆工程的必要性之所在。

## 二、抢占与盘活：发展中国数字图书馆工程的大思路

发端于 20 世纪中期的一系列高新技术革命，特别是信息技术革命，极大地促进了全球范围的产业变革。国家经济的支柱产业，正在由服装纺织等劳动密集型产业和冶金、化工、汽车等资本密集型产业向宇航、软件、微电子等知识技术密集型产业转移。另一方面则是与信息产业相关的电子计算机、通信设备、软件生产、激光技术和生物工程等一大批知识技术密集型产业得到蓬勃发展。信息产业在经过优化调整的产业结构中将成为占主导地位的产业。发展中国数字图书馆工程，就是要以此为龙头带动我国的信息产业和相关行业，积极迎接新世纪的挑战，把握住下一轮世界性产业结构调整为契机，使自己迅速发展壮大，从而在国际经济一体化和信息高速公路全球化的发展趋势中，占据自己应有的位置。

其实，产业结构的调整。即所谓产业“朝阳—夕阳”周期的变换，更深层次的因素是知识创新和技术进步。当一个产业的知识创新和技术进步刚开始发展的时候，这个产业就会处在快速上升的阶段，成为一个朝阳产业；反之则会成为一个夕阳产业。因此，产业进步实际上就是知识技术的进步，产业的停滞也就是知识技术的停滞。对于朝阳产业，为维持其强劲的发展势头，就必须不断进行知识创新和技术进步以加速发展。历史和现实都已经证明，政治和军事的力量取决于经济 and 产业的发展，经济 and 产业的发展又取决于科技的进步和突破，科技的进步和突破的关键又在于知识信息的及时获取和有效利用。因此，发展中国数字图书馆工程，要立足于紧紧跟上世界科学技术发展的步伐，立足于对知识信息的及时获取、深层开发和有效利用。

要能为我国的知识创新和技术进步提供强有力的信息保障，从而成为国家创新体系总体格局的重要组成部分。

纵观人类的发展史，任何一个时代的文化强盛都表现了这个时代国家的强盛。现在西方发达国家数字图书馆的主要内容就是将其历史上的优秀文化整理上网，力图在因特网上宣扬西方的文明与历史，展现西方的优秀文化。有关统计资料表明，目前在因特网的信息中，90%以上是英文，中文信息少得可怜，而这有限的中文信息还大多是港台制作的。另外，在

整个因特网的信息输入和输出流量中,我国目前分别仅占 0.01%和 0.05%,这与我国在国际上的地位和影响极不相称,也与中华文明在世界文明中的地位和影响极不相称。发展中国数字图书馆工程,就是要积极组织图书馆和有关机构的中文信息资源上网,使我国悠久的历史、灿烂的文化以及社会主义建设的方针、政策、成就和文化教育、科学技术等方面的重大进展,通过因特网主动地传送到世界各地,占领因特网上中文信息资源领域,扩大中文信息资源的利用和影响。

在知识进步和技术革命的推动下,信息服务产业取得了迅猛发展,在发达国家中已经成为国民经济的支柱产业。由数字图书馆而实现的电子信息服务,同以往传统意义上的信息服务最大的不同在于无空间限制,形成了无房屋限制,无地域限制,甚至无国界限制的全球服务体系。

因此,谁掌握了信息源并使之获得效益,谁就具备了信息服务市场的竞争能力。中国一方面是信息源的博大宝库,另一方面又是巨大的信息服务市场。中国电子信息服务市场化的主要竞争对手不是来自于国内,而是来自于国际。率先拥有中国的信息源,占领中国的信息市场,进而实现国际间信息源的交流和市场的运营,创造巨大的经济效益,理应成为发展中国数字图书馆工程的市场战略原则。

由此可见,在发展中国数字图书馆工程的大思路中,所谓“抢占”至少有四个层次的含义:

(一)在 21 世纪世界政治、经济、文化大格局中,抢占产业结构调整的先机;(二)在 21 世纪世界政治、经济、文化大格局中,抢占获取知识信息的制高点;(三)在 21 世纪世界政治、经济、文化大格局中,抢占因特网上中文信息资源领域;(四)在 21 世纪世界政治、经济、文化大格局中,抢占中国的信息服务市场。

如果说我们用“抢占”突出强调了发展中国数字图书馆工程的时效性和竞争性;那么我们用“盘活”则能概括出发展中国数字图书馆工程的拓展性和效益性。

中国数字图书馆工程的目标,是建立起一个跨地区、跨行业的巨大文化信息资源网络,使之成为我国的“国家信息基础设施”;全面地收集图书馆、博物馆、纪念馆、新闻出版机构、艺术团体、音像影视、体育、旅游等单位的文化资源信息,将其中的精品内容进行数字化与深加工,建设起一批以数字化图书馆、数字化博物馆、数字化影视中心等为代表的资源库,形成巨大的知识宝库;依托国家骨干通信网,建立完整的信息资源网络体系,针对不同需求的用户,提供全面、灵活的网络连接方式,为用户提供对各种资源库的快速查询与检索;开发智能化中文用户界面和廉价的用户接入设备,普及网络的使用,使用户可以方便地、远距离地、象收看电视一样获取网上的资源,使文化信息资源得以最大限度地利用。

资源库建设是数字图书馆建设的核心。其基本思路是:先建设急需急用和容易的,先易后难,从小到大。出发点是大文化的角度,顶层设计涵盖整个文化建设。初步设想,注册“中华文化信息网”,将现有的网上资源进行有效的编排后,搭载到该主页上,使之成为政府发布文化信息的主页,同时,作为数字图书馆运行数字化产品的窗口,它将成为实现资源共享的首选网站。

以中国数字图书馆工程为例,该工程将完成中华文化史资源库、中华人民共和国资源库、中国共产党历史资源库,中国发明创造资源库、中国法制资源库、中国国情资源库、中国教育资源库、中国民族文化资源库、中国图书馆馆藏珍品资源库,中国博物馆馆藏珍品资源库、中国宗教资源库、中国医药学资源库、中国当代名人资源库、中国旅游资源库、中国艺术资源库、中国经济信息资源库、中国软件资源库、科技资源库以及面向青少年的一个百科全书式的知识宝库资源库的建设。这些资源库和以后其它根据需求开发和建设的资源库,统一在“中华文化信息网”上进行信息传播,将会十分有效地把我国几千年来文化积累转化为丰厚的文化资产,对我国社会进步、经济发展和知识创新产生难以估量的巨大推动作用。中

国数字图书馆工程的实施,还会在资源有偿服务、人力物力节省等方面产生可观的直接经济效益。数字图书馆将最终实现全国各类图书馆、情报中心、博物馆等机构的资源共享。既可以有效解决图书、期刊特别是海外书刊的重复购置问题,又可以最大限度地扩大图书、期刊品种;在资源共享系统里,只要有一家图书馆对一种书刊进行一次采编加工,其它图书馆就可以充分利用这条数据,从而节省大量人力,提高工作效率;可以在国内外发行个人网络阅读磁卡,开展网上购书,网上广告等电子商务业务。中国数字图书馆工程的技术研究与技术开发将为我国软件系统研究开拓一个最好的直接应用场所,还将带动我国的计算机业、有线电视网和电信业等相关产业的快速发展,形成国民经济新的增长点。

### 三、基础与优势:培育中国数字图书馆工程的大环境

数字图书馆是一个集计算机技术、网络技术、通讯技术和多媒体技术等为一体的高新技术项目,它的建设必须依靠国家有关基础设施,依靠市场上成熟和实用的高新技术及相应的专业人才和管理人才的相互配合。目前我国已经基本具备了建设数字图书馆的条件。

首先是先期工作已取得初步成果。中国国家图书馆(北京图书馆),从1995年起就安排专人负责跟踪掌握国外关于研制数字图书馆的发展动态,1997年由国家图书馆牵头,与上海图书馆、南京图书馆、中山图书馆、深圳图书馆和辽宁图书馆联合开展了在国家计委立项的“中国试验型数字图书”项目的试点工作,在数字信息资源库设计、专用软件工具和检索标准化等方面取得了十分宝贵的初步成果,为实施全国范围的“中国数字图书馆工程”项目奠定了基础。

其次是相关技术已获得基本保障。中国科学院、北京大学、清华大学等在相关的技术领域开展了研究工作,特别是在海量数据库软件的引进开发和数据仓库技术开发等方面积累了成功经验,并培养了一批人才。这些单位所进行的国际间技术合作和引进技术运用等工作,达到了国际同期水平,为实施数字图书馆工程提供了技术支持。

其三是网络基础设施建设已初具规模。我国四个互联网已经相继实现了国际联网和国内互连。

中国互连网(China Net)已经覆盖了31个省市;中国科学技术网(CST Net)实现了科学院百所连网;中国教育与科研计算机网络(CER Net)连接了300多所大学;中国金桥信息网(China GBN)也在24个省市建立了站点。中国数字图书馆工程的建设与应用可以充分利用这些已有的网络平台。另外,我国的有线电视网(CATV)覆盖面广,现已有7800万用户,是全球最大的有线电视网。对其加以适当的技术改造,将可通过有线电视网将数字信息传递到千家万户,形成庞大的用户群。在我国电信网和有线网尚未达到的一些边远地区,可以通过卫星网作为传输手段。三网并用,特别是初步形成的宽带、双向传输网络,为中国数字图书馆工程提供了可选择的、有效的、国际间的互联方式。

其四是已拥有较为丰富的电子信息源。近年来,我国电子出版物飞速发展,全国各出版机构已出版了3000多种电子出版物,国内各种类型的数据库建设也取得可观的成绩。电子版图书、报纸和期刊的出现与发展以及国内出版发行的录音、录像等视听资料的日益丰富,为数字图书馆的建设积累了较为丰富的源数据。

我们不仅要认清我国现在已经具备的建设数字图书馆的基本条件,而且要看到自己独特的优势。

改革开放20年来中国经济保持了持续高速增长,物质基础大为雄厚,文化教育空前普及,科学技术全面进步,全社会正进入对文化需求越来越强烈的小康社会发展阶段,党和政府也适时地提出了加强文化建设的总体要求。江泽民同志在党的十五大报告中指出:“社会

主义现代化应该有繁荣的经济、也应该有繁荣的文化”，“面对科学技术迅猛发展和综合国力剧烈竞争，面对世界范围各种思想文化相互激荡，面对小康社会人民群众日益增长的文化需求，全党必须从社会主义事业兴旺发达和民族振兴的高度，充分认识文化建设的重要性和紧迫性。”朱容基总理和李岚清副总理在国家科教领导小组会议上的讲话，都强调了要根据国民经济发展的需要和已经具备的条件进行整体规划，集中力量重点突破，形成国家创新体系的总体建设格局，并推动和组织实施。所有这一切，都将成为中国数字图书馆工程建设强大而持久的动力。

中华文化以其悠久的历史积存，深厚的文化底蕴和众多的使用人群而在世界文明中独具魅力。在国际经济日益走向一体化和信息高速公路全球化的今天，随着汉字输入技术的突破，汉语，汉文化同英语、英文化必将成为文化传播和文化教育的东、西方两极。这一大的国际文化格局，为中华文化在电子信息时代的发展提供了广阔的空间。

社会主义制度的优越性之一就是能够集中力量办大事，而数字图书馆工程正是这样一件资金需求大、技术要求高、涉及领域广的大事，必须要多方共同协作才能完成。中国的数字图书馆工程，将由主管国家文化事业的文化部牵头，协作单位也都各俱优势。

中国国家图书馆（北京图书馆）馆藏文献资料达 2100 万册（件），居世界第五。而且还以典藏古籍善本闻名，是世界收藏中文图书最丰富的图书馆。115 种文字的外国文献资料占馆藏的 45%，是国内外国文献最大的藏家。国家图书馆已基本实现了图书采选、编目、读者服务的计算机管理，已建成建国以来的万余条国家书目数据库，实现了与中国互联网、中国教育与科研网和国际互联网的连接，网上提供了书目型数据 500 余万条，全文数据 300 余万页。”

中国科学院是国家科研中心，已建成的总数据量达 283GB 的科学数据库，是我国领域最广泛、数据量最大的网上数据库。中国科学院在手工文字和语音识别技术、图书信息语音查询技术、文本图象处理技术，网络安全和保密技术、身份鉴别和签章技术、多媒体通信技术等方面的众多科技成果，可为数字图书馆工程提供技术支持。

航天工业总公司是我国运载火箭、卫星系统工程控制管理、信息加工、光机电多种技术集合的高科技科研生产企业。近年来，在国家信息工程（金税工程、金卡工程、金关工程）和网络工程等方面，积累了许多实践经验，并取得丰硕成果。

广播电影电视总局广播影视信息网络中心，承担着全国有线电视网的建设和运行。目前已经建设完成连接全国 12 个省市，近 9000 公里的光缆工程，一期工程的联网工作已经接近完成，有线电视的服务已进入 7800 万户家庭，使中国的有线电视网超过美国成为世界第一大规模的网络。网络中心还成功地开办了数字广播业务、图文电视、数字高清晰度电视和高速多媒体业务，是我国有线电视领域的牵头单位。

清华大学是我国在信息科学与技术方面具有很强实力的教育和科研单位。其计算机科学与在图像识别、声象识别、机器人研究等方面取得了许多成果。清华大学信息网络中心的主要研究方向是计算机网络及其应用，其中包括数字图书馆的关键技术。

北京大学计算机系多年来承担了多项国家重点科技攻关项目，可以在已有科研成果基础上，提供多项技术支持：如软件开发支撑技术，数据仓库技术，网络与应用技术，信息安全技术，基于内容的多媒体检索技术等等。

总之，在中央的关怀重视下，在国家主管部门的领导下，通过全国范围专业人才、技术、设施和资源的大协作，集成国内外先进技术和设备，统筹规划，总体设计，突出重点，分布实施，我国的数字图书馆就一定能顺利完成并达到国际先进水平。数字图书馆必将为中华文化的大发展提供新的契机，让我们共同努力，去开辟文化繁荣和民族振兴的光辉前景。