

Introduction to Digital Library

数字图书馆 导论

夏立新等编著

湖北人民出版社

数字图书馆导论

夏立新 等 编著



湖北人民出版社

鄂新登字 01 号
图书在版编目(CIP)数据

数字图书馆导论/夏立新等编著.
武汉:湖北人民出版社,2004.5

ISBN 7-216-04019-8

I. 数…
II. 夏…
III. 数字图书馆
IV. G250.76

中国版本图书馆 CIP 数据核字(2004)第 080370 号

数字图书馆导论

夏立新 等编著

出版: 湖北人民出版社
发行:

地址:武汉市雄楚大街 268 号
邮编:430070

印刷:武汉中远印务有限公司

经销:湖北省新华书店

开本:787 毫米×1092 毫米 1/16

印张:21

字数:346 千字

插页:4

版次:2004 年 5 月第 1 版

印次:2004 年 5 月第 1 次印刷

定价:35.00 元

书号:ISBN 7-216-04019-8/G·1130

本社网址:<http://www.hbpp.com.cn>

内 容 简 介

本书系统阐述了数字图书馆的相关概念、发展背景、体系结构、技术架构、“馆藏”建设与维护、信息资源的组织技术、文本信息检索、多媒体信息检索、著作权保护问题、中外数字图书馆发展比较以及数字图书馆的应用等问题。

作为一本数字图书馆方面的著作,本书内容较为全面详实,适合于作为信息管理与信息专业、图书馆学专业的教材,同时也适合于作为因特网环境下从事信息资源的开发与利用的实际工作者的参考书。通过对本书内容的系统学习,读者能够较为系统地了解 and 掌握数字图书馆方面的基本知识、基本理论和基本技能,能够从事数字图书馆的规划、设计、建设和管理的具体实务。

前 言

数字图书馆是采用现代信息技术的数字信息资源系统,是下一代因特网上信息资源的管理模式,将从根本上改变目前因特网上信息分散不便使用的现状。作为知识经济的重要载体,数字图书馆是国家信息基础设施建设不可缺少的重要内容,也是新形式下国家间资源争夺的一个新领域。为此,世界各国高度重视和大力推进数字图书馆建设,一批有影响的实验性项目和工程建设项目纷纷涌现。目前,数字图书馆已成为评价一个国家信息基础水平的重要标志和本世纪各国文化科技竞争的焦点之一。在我国的学术界,伴随着数字图书馆概念的研究、国外数字图书馆理论研究和工程建设成果的翻译与评价、国内的技术攻关以及工程实施的不断深入,涌现了不少优秀的学术成果。本书是我们多年来开展数字图书馆相关问题的研究,以及近几年从事数字图书馆课程教学的经验总结。

数字图书馆实践的蓬勃发展亟需教育界培养和造就具备数字图书馆建设和管理专业素质的人才。为此,本书的撰写紧密结合了教学改革和课程建设的实际,努力反映数字图书馆方面的理论研究和实践发展的最新成果,体现内容体系的创新,在内容上注重基础性与创新性、系统性与选择性、经典性与前沿性的有机结合。与同类著作相比,本书的特色主要体现在:①从信息管理与信息系统的专业视角来研究和介绍数字图书馆,把数字图书馆看成因特网环境下履行数字化信息资源的采集、加工、组织、存储、检索和提供利用的信息系统,并以此为主线展开章节的安排和内容的介绍,力求系统介绍数字图书馆所涉及的基本知识、基本理论和基本技能,在内容的安排上注重基础性与创新性、系统性与选择性、经典性与前沿性的有机结合。②注重理论和实践的结合。在理论上

进行归纳总结的同时,结合典型的数字图书馆实例,介绍数字图书馆系统的体系结构、设计与开发工具。通过对本书内容的系统学习,读者既可以系统地掌握数字图书馆方面的理论和专门知识,又能具备一定的从事数字图书馆规划、设计、建设和管理的动手能力和实践能力。

这本著作的顺利出版,首先要感谢华中师范大学教务处的立项资助,感谢信息管理系的领导与老师的关心与支持。另外,还要感谢这部著作的合作者:华中师范大学的王伟军教授、熊回香副教授、刘刚老师,武汉大学的陆伟博士、金燕博士、刘玮博士、宋登汉馆员、杨立新高级工程师。这部著作的撰写分工如下:夏立新负责本著作写作大纲的设计、全书的统稿,执笔第1、4、6、9章,金燕、陆伟执笔第8章,金燕、刘刚执笔第11章,陆伟、刘玮执笔第2章,刘玮、王伟军执笔第10章,熊回香执笔第7章,杨立新执笔第3章,宋登汉、陆伟执笔第5章。高丹同学承担了本书部分章节的文字校对工作。没有我们这个团队成员的精诚合作,本著作也难以顺利完成。

在本著作的撰写过程中,我们广泛吸取了国内外有关数字图书馆理论与实践的研究成果,参考和引用了大量的相关文献及网上在线资料。借此书出版之际,作者谨向这些原文作者以及所有关心和支持本书的撰写与编辑出版的同志表示由衷的感谢!

数字图书馆的理论研究在不断深入,工程建设项目也不断涌现,本著作需要补充和完善的内容还很多,加之作者的水平有限,时间仓促,书中难免出现许多疏漏和错误,恳请专家和读者批评指正。我们期待着能有机会就数字图书馆理论研究与工程实践中的有关问题,与有志于数字图书馆领域研究的同仁进行交流与合作。

编著者

2004年3月

目 录

1 数字图书馆及其相关概念	1
1.1 数字图书馆的产生背景	1
1.2 数字图书馆的相关概念	7
1.3 数字图书馆的特征	17
1.4 数字图书馆的功能	21
1.5 数字图书馆与传统图书馆的业务比较	27
2 数字图书馆的体系结构	31
2.1 数字图书馆体系结构的发展	32
2.2 数字图书馆的信息结构	43
2.3 一个典型实例的分析	54
3 数字图书馆系统技术构架	61
3.1 数字图书馆技术体系结构	61
3.2 系统设计的基本原则与模式	67
3.3 系统的核心技术环节	69
4 数字图书馆的“馆藏”建设与维护	71
4.1 演进中的馆藏概念	71
4.2 数字图书馆“馆藏”构成分析	77
4.3 数字图书馆“馆藏”发展与维护	83
4.4 数字图书馆“馆藏”评价	92

5 数字图书馆的信息组织·····	98
5.1 概述·····	98
5.2 数字图书馆信息组织方法和技术·····	102
5.3 数字图书馆信息组织方案:MARC、元数据,还是XML/RDF·····	108
6 数字图书馆的文本信息检索·····	147
6.1 数字图书馆的文本信息检索基础·····	147
6.2 数字图书馆的文本索引方式·····	153
6.3 数字图书馆的信息检索性能评析·····	158
7 数字图书馆的多媒体信息检索·····	165
7.1 多媒体检索基础·····	165
7.2 基于内容的图像检索·····	169
7.3 基于内容的视频检索·····	180
7.4 基于内容的音频检索·····	189
7.5 多媒体信息检索的相关反馈技术·····	196
7.6 多媒体信息检索技术的局限与发展·····	198
8 数字图书馆的应用·····	202
8.1 公众服务·····	202
8.2 电子商务服务·····	212
8.3 远程教育·····	223
9 数字图书馆的用户研究·····	231
9.1 数字图书馆用户·····	231
9.2 数字图书馆的用户分析·····	236
9.3 数字图书馆的用户管理·····	242
9.4 数字图书馆的用户服务·····	247
10 数字图书馆建设中的著作权保护问题·····	256
10.1 数字图书馆的主体性质·····	257
10.2 数字图书馆的著作权的运行与管理·····	263
10.3 如何保护作品的著作权·····	267
10.4 对著作权制度的几点思考·····	279

11 中外数字图书馆发展比较·····	286
11.1 数字图书馆的研究里程·····	286
11.2 国外数字图书馆项目概况·····	294
11.3 中国数字图书馆发展现状·····	310
主要参考文献·····	322

1 数字图书馆及其相关概念

1.1 数字图书馆的产生背景

20 世纪末,随着计算机技术、数字存储技术、现代通讯技术、网络技术的飞速发展与有机结合,数字化信息广泛存在于我们身边。象其他格式的信息一样,数字信息在整个生命周期内运动,它们被产生、编辑、描述、标引、传播、获得、使用、注释、修订和再产生。而当我们谈论图书馆时,首先想到的是雄伟的建筑,汗牛充栋的书库和宽敞明亮的阅览室。但在数字化信息的冲击下,一种新型图书馆——数字图书馆应运而生,它打破了人们对图书馆传统认识的时空观念,在基于计算机网络特别是因特网的多媒体信息管理系统下,数字图书馆的信息不再存放在一栋物理建筑里,而是分布在世界各地的网络系统中,呈现出一种虚拟状态。

我们认为,如果从社会、经济、技术发展大背景的角度来分析数字图书馆产生与发展的影响因素,则会进一步加深我们对数字图书馆的了解,真正认识其存在与发展的本质。

1.1.1 数字图书馆是现代信息技术发展与人类信息知识需求发展相结合的产物,人类社会中任何一种事物的产生与发展都与人类生产生活的需求分不开。

人类文明自产生发展以来,绝大多数是通过文献记录下来并流传后人的。图书馆作为文献最重要的保存、利用场所之一,通过对文献的开发利用,传播信息和知识,积累、继承并创造出新的知识,因而成为人类文明发展的重要推动者。随着现代信息技术在文献和图书馆中的应用,这种推动作用将越来越巨大。

回顾 20 世纪的 100 年,冲击和影响人类社会各个方面幅度最大、速度最快、震撼最为强烈的,莫过于延续至今仍在飞速发展的信息技术。其影响涉及社会生活的各个方面,图书馆也不例外。从历史上看,信息技术的发展水平及其在图书馆领域的应用水平和程度在很大程度上被视为图书馆发展阶段的重要依据。

从总体上看,20 世纪 60 年代以前的图书馆是以收藏印刷品为主,通过卡片式或书本式目录反映馆藏,通过检索刊物反映本馆可能没有收藏但确实已经存在的文献,图书馆业务工作的开展基本上以手工为主。

20 世纪 60 年代以后,以计算机、通讯以及网络为核心的现代信息技术开始全面渗透进图书馆领域,将图书馆推向自动化发展阶段。1960 年,美国人尼诺提出编印目录卡片的机械化装置,被认为是最初的机读目录。1962 年,美国加利福尼亚大学和南伊利诺斯大学分别应用计算机进行期刊管理和流通管理。1964 年,美国国家医学图书馆用计算机编制《医学文摘》,并生产 MEDLARS 数据库磁带。1966 年,美国国会图书馆开始了 MARC 计划(Machine Readable Catalogue),特别是 1969 年开始正式发行 MARCII 机读目录,为实现图书馆大部分业务工作的自动化奠定了基础,为图书馆资源共享创造了条件。

20 世纪 70 年代,计算机与通讯技术相结合,促进了图书馆联机系统和网络化的发展,出现了一批联机编目网络和一些商业性联机检索系统,如 OCLC(Ohio College Library Centre)易名为“联机计算机图书馆中心”(Online Computer Library Centre),并发展成为全球最大的书目数据库中心,以及 Dialog、ORBIT、MEDLARS 等联机检索系统先后建成并投入商业运行。

20 世纪 80 年代,图书馆自动化的显著特点是图书馆联机系统进一步发展与各馆建立的自动化集成系统联成网络。各馆普遍实现了采购、编目、流通、连续出版物管理、检索、馆际互借以及行政管理等方面的自动化,建立了机读目录数据库和二次文献数据库。用户通过联机公共存储目录(OPAC)查询馆藏书目,通过联机检索终端查找二次文献信息。

从图书馆采用现代信息技术的历程看,图书馆一直强调内部业务工作的自动化和检索工具编制与使用的自动化。但由于受到计算机运算速度和存储容量等的限制,图书馆当时难以实现对文献本身存储与检索的计算机化。

到 20 世纪 90 年代,高密度存储技术、多媒体技术、通信技术、网络技术现代信息技术的高度发展、不断突破及其有机结合,使图书馆已可能对文献本身进行管理。从此,图书馆逐步迈进新的阶段——数字图

书馆。

而图书馆所收藏的文献从产生以来,也一直根据社会需求的发展而不断地应用新的信息技术进行保存与管理。从最早的山石、金属器皿、甲骨,到后来的竹简、丝帛,再到后来的纸张,进入 20 世纪后有了唱片、录音带、录像带,特别是产生了数字技术后有了 CD-ROM 光盘、DVD 光盘等,包括网络与数据库管理,以及可以预见到的数字图书馆系统等。可见文献资源的生产与管理总是走在新技术应用的前沿。进入 21 世纪,数字化的文献资源已成为社会需求与应用的主体。相应的,文献资源数字化生产与管理技术及其解决方案便成为重要的发展战略性课题。

社会公众对于图书馆的要求,主要是从馆藏文献中获取系统的、成型的信息——知识。即便是在欧美经济文化最为发达的国家和地区,公众的需求也主要表现为系统知识的获取。大量实际调查与研究表明,对于科研人员、企业决策人员、行政管理人员等来说,处于当代世界的激烈竞争氛围中,第一时间获取最新信息自然是非常重要的。但是对社会的绝大多数成员来说,情况却明显不同。社会环境及其个人生活的种种变化,随时可能要求他们涉足一个新的领域,此时他们特别需要的是这一领域最新的动态和相关知识。因此,图书馆将是他们获取不同领域知识的最佳场所,图书馆已开始成为终身学习环境的营造者。

进入新世纪,置身于人类社会全面信息化、知识化趋势下的文献资源生产、管理与利用,必然以数字化、网络化为其主要特征,以方便高效地满足人们知识获取与学习需要为目标。因此,新技术的发展与人类信息、知识需求发展相结合的结果,必然产生数字图书馆这一新型信息与知识的处理系统。可以说,数字图书馆就是记录人类文明的载体——文献在最新的相关技术应用于其中后,能够满足人类对知识、信息最新需要的必然产物。

1.1.2 从技术、需求方式、经济等具体角度讲,数字图书馆的产生及其持续发展曾经并仍将受以下几点主要因素的影响:

1. 新战略资源的无限增长与有效利用的矛盾亟需解决

自 20 世纪 60 年代以来,人类在生产实践中逐渐认识到,资源不仅有各种物质形态,也包括知识、经验、技术等非物质的信息形态。前者包括物质资源和能源资源,后者可称为第三资源即信息资源。这是一种战略资源,现已成为知识经济时代决定社会发展进步的决定性力量。但是,由于现代经济活动的发展为信息的生成、传递、存储和积累提供了条件,各种形态的信息以突飞猛进的速度增长并迅速积累起来,很快就达到了

一个非常庞大的基数。据统计,20世纪40年代以来所产生和积累的信息量超过了这以前人类有史以来所有的信息量之和。根据《自然》杂志1999年8月发表的研究结果,WWW上有近乎80亿页的可公开存取的信息^①;Web的规模仍在呈指数增长,根据另一项估算,过去两年,Web的规模扩大了3倍^②。由此可见,网上信息资源的发展势头已远远超过当年引发“情报危机”时文献的增长,正在出现因特网环境下的“信息爆炸”。

如此海量的信息增长与有效的信息管理和获取利用两者间必然存在着尖锐的矛盾。随着信息量的增加,用户需求水平的提高,矛盾的尖锐程度几乎达到无法调和的地步。即使使用最强大的搜索引擎,当用户提出一个具体问题时,返回的结果仍是大量而繁杂无序的。这样,用户只能在众多的网站和网页中换来换去,不知不觉中时间很快就过去了,真正准确、需要的结果却没有得到。浪费了时间不说,用户反而更加迷茫了,因为他(她)的需求没有在最短时间内得到满足。由此可见,当因特网神奇般地通过数以万计的计算机的连接,把整个地球联系在一起的时候,人们在学习、生活和工作中享受了由此带来的便利。而另一方面,网上信息的急剧增加与用户对信息的个性化需求之间的矛盾不但没有得到有效克服,反倒愈加突出,人们似乎又回到“情报爆炸”时代所面临的同样困境。虽然因特网使我们得到的信息比以前任何时候多,但并不意味着我们总能得到我们想要的信息。

信息利用者,特别是知识型信息的利用者,已成为信息海洋中无助的孤独者,迫切需要一个真正的知识导航系统。毫无疑问,这个真正的知识导航系统应该就是数字图书馆。

2. 科学技术的发展为数字图书馆提供了实现手段

数字图书馆是一个宽带多媒体网络和海量信息管理系统,它所面对的管理对象、存储对象和技术领域远远超出了传统图书馆的范围。存储对象包括数字化的图书、音像、新闻、美术作品、照片、雕塑、电影、软件、电子出版物、因特网内容、卫星数据、气象数据、地理数据、政府文件等各种各样的人文与科学数据,因而牵涉到大量的技术性管理问题。由于计算机技术、存储技术、网络技术、信息处理技术的不断发展,使得计算机性能不断提高。根据摩尔定律,计算机计算能力每18个月翻一番,即每

① Steve Lawrence and C. Lee Giles. Accessibility of Information on the Web. Nature 400 (Jul. 8, 1999), 107

② OCLC Office of Research. June 1999 Web Statistics. Web Characterization Project. <http://www.oclc.org/oclc/research/projects/webstats/statistics.htm> (访问时间: Jan. 2, 2001)

5 年增长 10 倍。但现代信息技术的发展已突破了摩尔定律,如计算机的三维图形性能、广域网的带宽、光纤传输容量、存储器的容量、信息处理技术、智能处理技术等都在以超常的速度发展,特别是这些技术在应用领域的广泛普及,为数字图书馆的建设与发展提供了现实技术条件和应用条件。

3. 人们需要一个真正的文献信息资源集成共享系统

对通过因特网来查找图书馆藏书和专题数据库资料的读者来说,他们目前还不能不关心从哪里得到所需的内容,而必须不断地做出判断和进行繁琐的操作,即到某个图书馆的数据库中查找藏书信息,或者用去较多的时间等待自己成为新的一家图书馆数据库的登记用户,以便使用各馆并不免费提供的内容。目前在图书馆系统内普遍存在着馆藏资源不能共享的问题,文献或数据库的严重重复与缺藏的问题并存。

事实上,即使规模再大的图书馆也无法将世界上所有的出版物收集齐全,而且也无法包揽所有文献信息的处理工作,更无法满足众多个性化的服务需求。在信息资源无限扩大的今天,如果一个图书馆不去借用外力,不去享用已有社会劳动成果,而只是一味地闷头工作的话,是难以完成自己使命的。因此无论是从读者需求的角度来看,还是从图书馆业务管理工作的角度来看,通过数字图书馆系统来实现各馆传统文献和数字化文献资源的共享是一个必然。而且,这种共享系统将逐步扩展到图书馆领域之外,与包括档案馆、出版社、新闻媒体、博物馆在内的资源拥有者共同组成。因此,数字图书馆系统必然是一个资源集成管理系统。这个系统既要管理传统文献,又要管理与传统文献在内容上紧密联系的数字化文献,而且不分类型,不分内容,这样才能实现真正意义上的资源共享。

4. 人们需要一个全新的知识型信息的管理与获取平台

信息利用是极易迷失方向的。很多不全面甚至是错误的信息被利用的例子屡见不鲜。而现有的因特网系统并未能从根本上解决这一问题。那么如何避免这样的现象呢?建立一个信息过滤系统,甚至是一个知识的生产系统是必然的选择。从前面的分析看,数字图书馆就应该是这样一个信息过滤系统。因此,建立一个可供全人类进行知识的管理、生产与获取利用的平台,应该是数字图书馆建设的目标。当前各地建设数字图书馆的热情很高,有的将一个馆的部分资料数字化了,就称自己是数字图书馆;有的公司将一些图书数字化了可提供网络阅读,也称是建成了数字图书馆,甚至有的建立了一些可进行检索的网页,也称建成了数字图书馆系统,这些都是真的数字图书馆吗?

尽管信息利用容易迷失方向,但很多人并没有意识到信息处理工作是更易迷失方向的。数字图书馆就是最高级的信息处理工作。如何才能不迷失方向呢?明确建设一个知识型信息的管理与获取平台的目标是非常必要的。诺贝尔奖获得者 H. 西蒙在他的名著《决策管理新科学》中认真地分析了 20 世纪 60 年代美国企业所搞的一批计算机信息系统都失败的现象。其表现是,这些原打算为企业高层领导服务的系统,实际上领导很少问津。H. 西蒙分析原因,原来是这些系统把方向搞错了。他们都把目标放在为领导提供更多的详细资料上,而在信息爆炸的社会,领导并不需要更多的资料,他们的信息负荷已经过载了。高层领导真正需要的是筛选信息的系统,而不是增加资料的系统。让我们就这个例子冷静地思考一下,在今天的数字图书馆研究和建设中,领导者和技术人员关注最多的是计算机技术、存储技术、网络技术、海量资源建设、技术标准等,用户的需求研究和资源内容研究明显被忽视。在这两项还没有清楚时,大量的资源已经被数字化了并开始提供服务了。这些系统是一个筛选信息的系统吗?我想今天的用户,就像 20 世纪 60 年代美国的企业领导者们一样,他们是会最终做出理智的判断的。同样清醒的数字图书馆设计与建设者,最终也会做出明智的选择。

5. 数字图书馆作为一个新的经济体有其不断发展的内在动力

现在很多人自觉或不自觉地将数字图书馆视为公益型的基础信息设施,这在其建设的初始阶段无疑是正确的。然而,数字图书馆最终将成为一个不断发展的经济体是必然的,这是由其与经济的密切关系所决定的。

首先,数字图书馆的建设需要信息技术厂商广泛参与。无论是计算机设备、存储设备、网络设备、数字化设备、通用软件的应用,还是应用软件开发与资源有偿使用等,都使得数字图书馆的建设引来商机无限。其次,从数字图书馆的运行本身来说,数字图书馆已经不可能是传统图书馆时代以公益性服务为主的非盈利性机构。数字图书馆中所管理的知识性资源是知识经济时代获取利润的源泉,谁掌握了这些知识性信息,谁就有可能创造出更多的精神和物质财富。因此,数字图书馆的一个主要功能,就是不断地发掘其中知识信息资源及其相关要素的经济功能,获取合法的经济收益。再次,数字图书馆可以实现有关文献资源机构的资源共享。这样可以有效解决文献资源(含数字化资源)的重复购置。而新的数字化资源进入系统后,其他机构就可以充分利用这部分资源,从而节省大量的人力、物力,提高工作效率,减少资金投入,这也是最大的效益。第四,在资源生产者的知识创造与所有权利真正得到法律的保

护,并通过数字图书馆系统得到技术保护和经济收益时,一个知识增值的链条就建立起来了。

知识的创造者会由于数字图书馆系统给其带来经济收益,而信任并支持数字图书馆的资源不断丰富发展,因此愿意提供自己所拥有知识产权的作品。而数字图书馆的用户,则因此得到更多真正需要的知识产品。这不正是一个以知识为产品的有效运转的经济体吗?知识创造者从数字图书馆系统中能够获得源源不断的经济收益,并将随着数字图书馆系统的完善而逐渐增大。

以上4种获取合法经济收益或效益的动力,必然促使数字图书馆的建设快速发展,使其作为一个知识经济时代的新型经济体,在市场中不断发展完善,满足社会与经济发展的需要。这也就是世界上数字图书馆建设浪潮为什么会一浪高过一浪,技术创新与建设热度不减,而且数字图书馆管理方绝大部分具有公司背景,甚至有的就是由公司来建设的原因了。如:中国数字图书馆有限责任公司、书生之家数字图书馆、超星数字图书馆等。而且很多真正成功的信息资源服务公司首先在于其企业化管理背景,如慧聪国际资讯有限公司、万方数据集团公司、清华同方光盘股份有限公司等,而并不在于其原来是否是文献和相关资源的所有者。这说明面向市场经济的体制、运行机制是核心。如果数字图书馆的建设与发展不能深刻地认识到并解决好这一问题,仍沿用传统图书馆的管理体制和运行机制,是不可能得到良性发展的。

1.2 数字图书馆的相关概念

1.2.1 数字图书馆概念

1. 数字图书馆概念的多种理解

从目前数字图书馆在国内外的发展和研究情况看,数字图书馆涉及到众多的领域,其中有很大一部分研究工作是由图书馆界以外的人士进行的,如计算机科学家、软件工程专家、专家系统研究人员、图像处理和语言处理专家、网络专业人员、教育界人士、知识产权专家等等。从中可以看出,数字图书馆这一课题实际上是一个与信息技术密切相关的跨学科、多层面的新兴领域。每个领域的探索都从不同的角度丰富了数字图书馆研究的内涵,而且对数字图书馆问题,甚至字眼本身都有不同的理解。

对于一般用户而言,数字图书馆只不过意味着传统图书馆的计算机化;对于当代 WWW 的用户而言,他们感觉效率、组织、功能和可用性日益增强的 WWW 就是数字图书馆,从这个意义上讲,WWW 是全球最大的数字图书馆;对于计算机专业人员而言,数字图书馆只不过是一个分布式的基于文本和图像的信息系统、分布式信息服务的一个集合,是一个网络化的多媒体信息系统或者称为一种分布式的互联信息空间;教育界人士则把数字图书馆视为融合正式学习、非正式学习、职业性学习的学习支持手段和工具,甚至认为数字图书馆就是数字化学校,将专门技巧、专业知识以及综合浏览等进行组装以用于创新性发展、自我指导和非正式学习;而对于图书情报学研究者来说,数字图书馆将以一种新的方式执行图书馆的功能,包含新型的信息资源、新的采访和馆藏发展方式(更多的资源共享和电子订购服务)、新的存储和保存方法、新的分类和索引方式、与用户新的交互模式、对电子系统和网络更多的依赖性,以及图书馆在人员智力构成、组织、经济等方面的显著变化;对于信息管理与信息系统专业^①的研究人员来说,数字图书馆只是信息管理系统中的一种,正如工资人事管理系统、银行的电子支付系统、会计中的电算化系统、商务活动中的电子商务系统等,它用于图书馆数字化信息的采集、加工、存储、提供服务等业务工作。

现将有影响的关于数字图书馆概念的描述和提法归纳如下,以期通过共同的探索,得出一个绝大多数人能够接受的对实际工作有指导意义的关于数字图书馆的概念。

1990 年美国密执安大学的研究人员首次提出了“数字图书馆”的概念,并将其定义为:数字图书馆是若干联合机构(Federated Structure)的总称,它使人们能够智能地(Intellectually)和实实在在地(Physically)存取全球网络上以多媒体数字化格式存在的、为数巨大的,且仍在不断增多的信息。

1995 年一个题为“交互作用,定标及数字图书馆研究议事日程”的数字图书馆专题研讨会给数字图书馆作了如下定义:“一种多媒体数据与把数据表现为信息与知识信息管理方法的两者的有序集合”。一年后此定义又修改为:数字图书馆是一种多媒体数据与信息管理办法的有序结合,这种管理办法把数据表现为一种对不同社会环境的人们有用的信息

^① 信息管理与信息系统专业是 1998 年教育部在修订高等教育本科专业目录的过程中,在原科技情报、信息学、经济信息、管理信息系统等专业的基础上,调整合并设立的新专业。