

数字图书馆 发展研究

Research

on Digital Library Development

王芬林 吴 晓 著

國家圖書館出版社

数字图书馆发展研究

王芬林 吴 晓 著

圖 國家圖書館出版社

图书在版编目(CIP)数据

数字图书馆发展研究 / 王芬林, 吴晓著. —北京: 国家图书馆出版社, 2012. 7

ISBN 978-7-5013-4808-4

I. ①数… II. ①王… ②吴… III. ①数字图书馆—发展—研究 IV. ①G250.76

中国版本图书馆 CIP 数据核字 (2012) 第 140154 号

责任编辑: 高爽

书名 数字图书馆发展研究

著者 王芬林 吴 晓 著

出版 国家图书馆出版社(原北京图书馆出版社)

(100034 北京市西城区文津街 7 号)

发行 010-66139745 66151313 66175620 66126153

66174391 (传真) 66126156 (门市部)

E-mail btsfxb@nlc.gov.cn (邮购)

Website www.nlcpress.com→投稿中心

经销 新华书店

印刷 北京佳顺印务有限公司

开本 710×1000(毫米) 1/16

印张 16.75

字数 300 千字

版次 2012 年 7 月第 1 版 2012 年 7 月第 1 次印刷

书号 ISBN 978-7-5013-4808-4

定价 62.00 元

对数字图书馆的系统思考与研究

——写在王芬林系列成果出版之际

七年理学教育背景,十年科技情报从业经历,十二年数字图书馆建设实践,期间直接参与了全国文化信息资源共享工程从无到有发展跨越的全过程,长期的知识积累、工作实践和思考研究,终于结出硕果,王芬林女士将她的两部系列专著《数字图书馆发展研究》和《数字图书馆实践思考——文化共享工程的发展与创新之路》摆到了世人面前。

由于是校友,也因为近年来承担了与文化共享工程相关的多项课题,我对王芬林女士是熟悉的。十六岁被北京大学录取,称得上是同龄人中的佼佼者;二十二年的专业实践,芬林女士在工作中积累,在实践中思考,并随时把实践和思考的结果记录下来,积少成多,集腋成裘,自然是不发则已,一发便是成体系的系列化成果。她的思考和研究不是偶发的、随意的,而是有计划、有步骤,从梳理数字图书馆的发展开始,到思考数字图书馆的建设实践,相信将来在该领域还会有进一步的拓展和延伸。

《数字图书馆发展研究》是芬林女士系列成果中的基础性著述,阐述了她对数字图书馆基本概念、基本原理、发展历程、国内外现状的理解和认识,体现了她在数字图书馆方面的深厚积累。介绍数字图书馆基本知识和原理的书很多,作为一本专著,“规定动作”在所难免。在浏览书稿时,我特别关注了作者对数字图书馆建设的独到思考。数字图书馆成功的要素是什么?怎样对数字图书馆进行评价?这是作者重点思考的问题。“认识决定成败”,数字图书馆不是传统图书馆信息化的代名词,不能把数字图书馆理解为一个信息化工程、项目或一个数字资源库。数字图书馆服务以网络为依托,应用的是网络技术,使用的是网络资源,参与的是网络群体,服务的是网络社会,实质上是一个整合、组织和利用各类网络资源的知识服务系统。作者提出的如上思考,给了数字图书馆以更多的人文关怀和社会性解读,新颖独到,令人深思,说明作者在理解和认识的过程中,同时也在思考、在批判、在升华。作者还较为系统地梳理总结了美国、英国、法国、日本、俄罗斯等主要国家的数字图书馆发展状况,一方面帮助读者了解数字图书馆建设在全球的发展;另一方

面也体现了作者的开阔视野,全面、系统、客观地审视了中国数字图书馆发展的水平和能力。

《数字图书馆实践思考——文化共享工程的发展与创新之路》,则把文化共享工程作为数字图书馆的一个实践案例,探讨数字图书馆建设中遇到的实际问题。文化共享工程和数字图书馆的交汇点在哪里?作者的一个基本观点回答了这一问题:文化共享工程是数字图书馆在文化信息资源服务中的实际应用,是服务型数字图书馆。这意味着什么?意味着文化共享工程不是一个立足于突破数字图书馆前沿技术的项目,而是一个以提供资源、服务基层、服务群众为目标的工程。因此,文化共享工程建设重点围绕资源建设和提供而展开:怎样规划和建设资源,怎样传输资源,怎样利用资源开展服务,怎样保证资源利用的合法性,怎样评价资源利用和服务开展的效果,等等。这些内容是该书探讨的主要内容。十多年来,芬林女士亲历了文化共享工程从筹备到启动到今天实现全国广覆盖的全过程。书中总结的经验、概述的模式、剖析的问题、提出的对策,都明显地体现了她在实践中总结反思、在领悟中提升发展、在理论指导下升华完善的特点。

由于职业和研究方向的缘故,我一直比较关注文化共享工程在基层公共图书馆建设和发展中发挥的重要作用。在许多场合我表达过一个观点:文化共享工程带动和促进我国基层公共图书馆数字化、网络化事业实现跨越式发展。十多年前,我国基层,特别是中西部经济欠发达地区的基层公共图书馆,数字化、网络化几乎是一片空白。文化共享工程的资金、设备、资源、技术、人才培养,让我国基层公共图书馆迅速具备了数字资源提供能力和远程服务能力,图书馆的现代化水平迅速提高。今天,公共图书馆的数字化、网络化水平在所有公共文化服务机构中走在前列,与这些年来文化共享工程的实施以基层公共图书馆为主要依托密切相关。作者在书中总结概括的文化共享工程数字资源建设与整合的方式、数字资源的传输手段和多样化的地方传输模式、数字资源的服务模式和服务成效、数字资源建设的知识产权保护实践、文化共享工程的绩效考核体系,都是以基层公共图书馆为主要阵地创造和实现的。这又从另一角度印证了作者提出的基本观点:文化共享工程就是服务型的数字图书馆。

芬林女士在书中对文化共享工程建设服务型数字图书馆的实践作了全面的总结、分析、提炼和探究,对中国其他文化惠民工程、信息化工程以及他国的信息化项目建设均具有借鉴意义。文化共享工程是中国政府在信息时代缩小数字鸿沟、消除知识差距、保障人民基本知识和信息权利的创举。放眼当今世界,各国政府面对信息时代的挑战,不约而同地都在实施一些政府主导或引导的信息化项目,以期

缩小数字鸿沟,弥合知识差距。发达国家如美国的“社区科技中心计划(Community Technology Centers Program)”、英国的“学习资料数字化和学习型社区网络(Digitisation for Learning Materials and the Community Grids for Learning)”；发展中国家如印度的“数字农村(Digital Village)”和“墙上之窗(A Hole in the Wall)”项目、智利的“图书馆网络(Biblioredes: Abretu Mundo)”项目、南非的“数字之门(Digital Doorway)”项目等。但像中国政府实施文化共享工程这样大规模的政府项目,在全国范围内把计算机公共接入点大规模地延伸到社区乡村,通过计算机网络把知识和信息送到老百姓身边,如此创举在当今世界堪称举世无双。因此,以国际视野来研究文化共享工程,向世界推广文化共享工程在保障公众知识和信息权利方面取得的实际成效,也应是未来文化共享工程研究的重要任务之一。

党的十七届六中全会提出了建设社会主义文化强国的宏伟目标。基本建立覆盖全社会的公共文化服务体系、努力实现基本公共文化服务均等化,是文化强国战略对公共文化提出的明确要求。覆盖全社会的公共文化服务体系,必须以互为补充、相辅相成的固定设施体系、流动服务体系和数字传播体系为载体。文化共享工程作为我国的一项重大文化惠民工程,在公共数字文化体系建设中,如何发挥主导作用真正实现文化信息资源的共建、共享和共用;“云技术”怎样为文化共享工程提供更好的技术支撑;等等,都是新时期文化共享工程面临的新任务,也是迫切需要进一步研究的新课题。希望能不断看到芬林女士新的研究成果。

北京大学教授 李国新
2012年7月

目 录

1 数字图书馆的定义	1
1.1 最早关于数字图书馆的描述	1
1.2 数字图书馆概念的相关叙述	4
1.3 数字图书馆概念的理解	7
2 信息革命的影响	18
2.1 对社会组织形态的影响	20
2.2 对组织结构的影响	21
2.3 对知识结构的影响	22
2.4 对文化传播的影响	24
2.5 对传统文化的影响	25
2.6 对图书馆的影响	26
3 数字化的机遇与挑战	28
3.1 从胶片业的发展看数字化的影响	28
3.2 电视领域的数字化影响	31
3.3 数字化冲击下的出版业	34
3.4 数字图书馆的机遇与挑战	36
4 从信息技术的发展变化理解数字图书馆	41
4.1 信息技术的发展	41
4.2 信息技术发展的阶段性与数字图书馆的发展	44
4.3 信息技术处理内容的阶段性与数字图书馆的发展	47
4.4 信息技术指导思想的阶段性与数字图书馆的发展	51

5 用户的数字图书馆	56
5.1 谁为用户	56
5.2 总体需求	59
5.3 应用环境	62
5.4 服务功能	67
5.5 数字阅读	75
5.6 用户终端	77
6 业务角度	83
6.1 不同社会条件下的定位	83
6.2 数字资源建设	87
6.3 数字资源服务	100
6.4 标准体系	108
6.5 数字资源版权	112
7 系统角度	115
7.1 系统架构	116
7.2 软件体系	125
7.3 硬件架构	139
7.4 新技术的意义和作用	165
8 关于数字图书馆建设的思考	173
8.1 数字图书馆成功的内在要素	173
8.2 数字图书馆评价指标	179
8.3 关于数字图书馆建设的思考	184
9 数字图书馆的发展	187
9.1 国外数字图书馆的发展	187
9.2 我国数字图书馆的发展	205

10 数字图书馆建设方案实例	222
10.1 国家数字图书馆工程	222
10.2 文化共享工程	231
主要参考文献	243
附 件	246

1 数字图书馆的定义

随着现代信息技术的飞速发展,以及图书馆和其他相关信息服务机构在数字化服务方面的长期积累和投入,数字图书馆的理念已逐步深入人心,为普通大众所接受。但是,关于什么是数字图书馆,它涉及多大的范围和领域,具体提供什么样的服务,这些问题却因人而异,有着不同的答案。这一方面是由于数字图书馆这一概念还没有形成普遍认可的定义,特别是它所具备的显著的跨学科特征,使得其在所跨学科的每一领域内都有自己具体的特点、要求和问题,从而使得这一定义本身就很难准确限定;另一方面,则是由于数字图书馆的理论定义与工程建设中的实际情况往往又存在一些差异和冲突,使得人们在现实生活中往往陷入困惑。因此,我们将在引入多种数字图书馆概念定义的同时,通过多种角度来解释和说明数字图书馆的概念,为读者理解数字图书馆在信息化背景下的发展奠定基础。

1.1 最早关于数字图书馆的描述

现今公认的最早的关于数字图书馆的描述是范内瓦·布什(Vannevar Bush)于1945年7月在《大西洋月刊》所发表的《诚如所思》(As We May Think)中提到的Memex^[1]。文中通过对Memex的功能性描述,反映了20世纪50年代人类对于数字图书馆的理想设计。其中关于Memex说明的章节内容如下:

Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, "memex" will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.

It consists of a desk, and while it can presumably be operated from a distance, it is primarily the piece of furniture at which he works. On the top are slanting translucent screens, on which material can be projected for convenient reading. There is a keyboard, and sets of buttons and levers. Otherwise it looks like an ordinary desk.

In one end is the stored material. The matter of bulk is well taken care of by improved microfilm. Only a small part of the interior of the memex is devoted to storage, the rest to mechanism. Yet if the user inserted 5000 pages of material a day it would take him hundreds of years to fill the repository, so he can be profligate and enter material freely.

Most of the memex contents are purchased on microfilm ready for insertion. Books of all sorts, pictures, current periodicals, newspapers, are thus obtained and dropped into place. Business correspondence takes the same path. And there is provision for direct entry. On the top of the memex is a transparent platen. On this are placed longhand notes, photographs, memoranda, all sort of things. When one is in place, the depression of a lever causes it to be photographed onto the next blank space in a section of the memex film, dry photography being employed.

There is, of course, provision for consultation of the record by the usual scheme of indexing. If the user wishes to consult a certain book, he taps its code on the keyboard, and the title page of the book promptly appears before him, projected onto one of his viewing positions. Frequently-used codes are mnemonic, so that he seldom consults his code book; but when he does, a single tap of a key projects it for his use. Moreover, he has supplemental levers. On deflecting one of these levers to the right he runs through the book before him, each page in turn being projected at a speed which just allows a recognizing glance at each. If he deflects it further to the right, he steps through the book 10 pages at a time; still further at 100 pages at a time. Deflection to the left gives him the same control backwards.

A special button transfers him immediately to the first page of the index. Any given book of his library can thus be called up and consulted with far greater facility than if it were taken from a shelf. As he has several projection positions, he can leave one item in position while he calls up another. He can add marginal notes and comments, taking advantage of one possible type of dry photography, and it could even be arranged so that he can do this by a stylus scheme, such as is now employed in the telautograph seen in railroad waiting

rooms, just as though he had the physical page before him. ①

上述文字的大致意思是：

考虑一下未来个人使用的设备，它将是一个机械化的个人图书馆。它需要一个名字提醒人们注意，“MEMEX”就可以。MEMEX 是这样一种机械化设备，人们可以在其中存储他所有的书、记录和信件，同时可以很高的速度和极强的灵活性完成检索。作为辅助设备，它是人脑的无限扩大。

它由一张桌子构成，大概可以远距离操作。桌子上有一个倾斜的半透明的屏幕，资料可以投影到上面进行阅读。还有一个键盘，一系列按钮和把手。除此以外，它就是一张普通的桌子。

它的一端是存储的材料。体积问题由于采用了改进的缩微胶片而得到很好的解决。MEMEX 内部只有很少一部分用于存储，其他的部分都是机械。即使用户每天塞进去 5000 页材料，也需要几百年才能把它的仓库填满。所以使用过程中完全可以挥霍一点，自由地填进去各种东西。

MEMEX 的大部分内容都是直接买来的，可以直接插入。各种书籍、图片、现行的期刊、报纸都可以这样得到并插入其中进行阅读。商业信函也可以用同样的方法，并且还备有直接输入的设备。桌子上有一个透明的平板，可以较长期地存放注释、注解、照片、备忘录等各种各样的材料。把材料放在上面，按下一个按钮就可以对它进行照相，并存储到 MEMEX 的胶片库中下一个空白位置——这里就要用到干式照相。

这里当然也有用于咨询的记录，它们来自于日常的索引工作。如果用户想从一本书中查找资料，他只需在键盘上轻轻敲出这本书的代码，这本书的封面就会马上投射到他面前的一个浏览位置上。系统会记住经常使用的书的代码，这样他就不需要经常去翻代码本了。如果真的要查代码本，只需轻轻敲一下键盘，这些代码就会投射出来供他查找。此外，他还有一个额外的游标可以使用。把这个游标拉到右边，这本书就在他面前一页页地投射出来，速度正好可以让人清晰地浏览每一页。再向右拉一点，浏览的速度就会变成每次 10 页，再搬过去一点还能以每次 100 页的速度浏览。向左搬这个游标则能向后翻页。

还有一个键可以让用户直接回到索引的第一页。用户的图书馆中的每一本书都可以用这种方式调出来查看，并且要比从书架上查找方便得

① 引自 As We May Think 的第六章，关于 memex 的内容主要参考第六及第七章。

多。由于有多个投影位置,因此用户可以在调另一本书时把第一本书留在桌面上。利用干式照相的优势,用户可以在书边上加上注解和评论;当然也可以重新设计,比如可以像我们在火车站候车室里见到的自动电报机一样,使用刻写的方案,就如同面对真实的纸张一样。

从今天的角度来看 Memex 显得有些原始,并且从某些方面来看,还有些不尽合理之处,而且甚至不是一个数字化的系统——它是基于微缩胶片的,信息根本没有数字化。但是它所提出的数据压缩、信息加工保存、快速检索、屏幕与键盘结合的操作形态等对于日后的计算机及数字图书馆的发展都起到了重要的指导作用,因此被认为是数字图书馆最早的理论描述。

至于数字图书馆(Digital Library)一词的出现,则要推迟到 Memex 提出的 40 多年之后。现在较为认可的关于 Digital Library 最早的书面材料是 Robert E. Kahn 和 Vinton G. Cerf 在 1988 年 3 月发表的“The Digital Library Project Volume I: The World of Knowbots, (DRAFT): An Open Architecture For a Digital Library System and a Plan For Its Development”。作者在文中对于数字图书馆的技术架构、应用和实施都做了详细的论述,是现代数字图书馆理论研究的最早文献之一。

1.2 数字图书馆概念的相关叙述

随着信息技术的快速发展,人们对于数字图书馆有了更为深入和广泛的认识与理解,并且由于出发点和落脚点的差异,对于数字图书馆的定义形成了许多不同说法,可以说是仁者见仁、智者见智。这里我们并不试图整合这些定义或明确支持其中的某一观点,而是摘录其中部分较有代表性的定义作为参考。

IFLA(国际图书馆协会联合会)关于数字图书馆的定义是:数字图书馆是高质量数字化馆藏的在线汇集,依据国际普遍接受的馆藏发展原则制作、收藏和管理,以协调统一和可持续的方式开放馆藏,并辅以必要的服务,使读者能够借阅和使用其资源^[2]。

原文: A digital library is an online collection of digital objects, of assured quality, that are created or collected and managed according to internationally accepted principles for collection development and made accessible in a coherent and sustainable manner, supported by services necessary to allow users to retrieve and exploit the resources.

2001 年美国 总统 信息技术 咨询 委员会 (PITAC, President's Information Technology Advisory Committee) 报告《数字图书馆: 获取人类知识的通用途径》中定

义:“数字图书馆:获取人类知识的通用途径。所有公民在任何时间、任何地点都可以使用与互联网连接的数字设备,搜寻到所有人类知识。通过利用互联网,他们可以访问到由传统图书馆、博物馆、档案馆、大学院校、政府机构、专门组织,甚至世界各地的个人所创建的数字藏品。这些新的图书馆提供的是传统图书馆、博物馆、档案馆的馆藏资料的数字版本,其中包括文本、文件、视频、声音及图像。它们所提供的强大技术实现能力,使用户能够改善其查询功能,对查询结果进行分析,并且改变信息的形式以便交互。高速网络使各个不同数字图书馆群的用户能够协同工作,对其各种发现相互进行交流,并使用仿真环境、科学遥感仪器、流式音频和视频。不管数字信息存放的物理位置在什么地方,先进的搜索软件都能找到,并及时提供给用户。在这样的美好前景中,任何教室、任何群体和个人都会与世界最大的知识资源近在咫尺。”

原文: All citizens anywhere anytime can use any Internet-connected digital device to search all of human knowledge. Via the Internet, they can access knowledge in digital collections created by traditional libraries, museums, archives, universities, government agencies, specialized organizations, and even individuals around the world. These new libraries offer digital versions of traditional library, museum, and archive holdings, including text, documents, video, sound, and images. But they also provide powerful new technological capabilities that enable users to refine their inquiries, analyze the results, and change the form of the information to interact with it, such as by turning statistical data into a graph and comparing it with other graphs, creating animated maps of wind currents over time, or exploring the shapes of molecules.

Very-high-speed networks enable groups of digital library users to work collaboratively, communicate with each other about their findings, and use simulation environments, remote scientific instruments, and streaming audio and video. No matter where the digital information resides physically, sophisticated search software can find it and present it to the user. In this vision, no classroom, group, or person is ever isolated from the world's greatest knowledge resources. ^[3]

Gary Marchionini 在其为《图书情报学百科全书》所写的“数字图书馆研究与发展”中指出,“数字图书馆是在不同群体中有不同意义的概念。对于工程和计算机科学群体而言,数字图书馆是一个管理结构化的多媒体数据的新型分布式数据库

服务设施的隐喻。对于政治与商业群体而言,这个词代表一种新的世界情报资源与服务的市场。对未来派群体而言,数字图书馆代表着韦尔斯的‘世界大脑’^①的表现形式”。

原文: Digital library is a concept that has different meanings in different communities. To the engineering and computer science community, digital library is a metaphor for the new kinds of distributed data base services that manage unstructured multimedia data. To the political and business communities, the term represents a new marketplace for the world's information resources and services. To futurist communities, digital libraries represent the manifestation of Wells' World Brain. The perspective taken here is rooted in an information science tradition. ^[4]

数字图书馆的概念并不仅仅是一个有着信息管理工具的数字收藏的等价词,数字图书馆更是一个环境,它将收藏、服务和人带到一起以支持数据、信息,乃至知识的全部流程,包括从创造、传播、使用、到保存的全过程^②。

美国研究图书馆协会(ARL)^③归纳了流行的数字图书馆的各种定义中具有共性的五个要素:数字图书馆不是一个单一实体;数字图书馆需要链接许多信息资源的技术;多个数字图书馆及信息机构之间的链接对最终用户透明;全球范围存取数字图书馆与信息服务是一个目标;数字图书馆的收藏并不局限于文献的数字化替代品,还扩展到不能以印刷形式表示或传播的数字化人造品。

美国数字图书馆联盟(DLF)的定义是:数字图书馆是一个拥有专业人员等相关资源的组织,该组织对数字式资源进行挑选、组织、提供智能化存取、翻译、传播、保持其完整性和永存性等工作,从而使得这些数字式资源能够快速且经济地被特定的用户或群体所利用^④。

康奈尔大学的名誉教授 William Y. Arms 在其专著 *Digital Libraries* 中对于数字图书馆的定义是:具有服务功能的整理过的信息收藏,其中信息以数字化格式存储并可通过网络存取。该定义的关键在于信息是整理过的。

① 世界大脑是英国科幻作家韦尔斯(H. G. Wells)在1938年的作品,描述了一个全新的、自由的、综合的、权威的和永久的“世界百科全书”,也就是世界大脑,它能够帮助公众最有效地获得全球统一的信息资源,并进而推进世界和平。相关资料可参看 World Brain, 或 Wikipedia 的世界大脑。http://en.wikipedia.org/wiki/World_Brain

② NSF“分布式知识工作环境”专题讨论会,1997

③ ARL(Association of Research Libraries)是一个在涉及美国以及加拿大的非营利性的研究型图书馆组织。相关内容可参考:<http://www.arl.org/>

④ 上述定义引自:李培. 数字图书馆原理及应用. 北京:高等教育出版社,2004

原文:a managed collection of distributed digital objects (contents), and services (functionality) associated with the discovery , storage , retrieval , and preservation of those objects in a context of networks , specifically the Internet. ①

中国工程院院士高文关于数字图书馆的定义是:数字图书馆是以电子方式存储海量的多媒体信息并能对这些信息资源进行高效的操作,如插入、删除、修改、检索、提供访问接口的信息保护等。并且具有三个核心定位:数字图书馆应该是中国的一个国家数字文化平台;还应该是一个国家数字教育平台;也是一个国家数字资源中心^[5]。

另外,国家图书馆还专门针对自身建设的国家数字图书馆工程给出了对应的定义:“国家数字图书馆是网络环境和数字环境下图书馆新的发展形态,是通过计算机技术、网络通信技术、多媒体技术等高新技术,搭建基于网络环境的资源建设系统平台,实现传统文献的数字化,及各类数字资源的组织、整合、管理和保存;是提供公共文化服务的全媒体服务平台;是面向专业机构和决策机构提供专业信息服务的学术性服务平台;是展示优秀中华文化、提升我国文化软实力的文化交流平台。”^[6]

1.3 数字图书馆概念的理解

由于描述者角度、观点和方法的差异,上述数字图书馆的概念和定义也各有不同。有的倾向于将数字图书馆看作是一个宏观的信息聚合和服务体,如美国总统信息技术咨询委员会(PITAC)的报告;有的倾向于将其看作是一个具体的功能实体,如IFLA的定义;而有的则更加明确地界定为是现有服务体的延伸,如国家图书馆关于国家数字图书馆的定义。

但是无论哪种描述,都明确地显示出数字图书馆绝不仅仅是传统图书馆的数字化,而是在新的时代、新的背景下,全新的、信息化、数字化、网络化的知识管理和和服务体。不过,这种理念性的说明在现实中还是会遇到一些问题。

首先,现代的图书馆工程建设中往往包含了大量的信息化建设工作,因此在建设方案中往往将图书馆建设与数字图书馆建设并提,最典型的的就是国家图书馆二期工程和数字图书馆工程的建设。这就使得普通读者时常难以区分两者的关系与差别,因而常常认为数字图书馆是图书馆的一个功能组成,或者干脆将图书馆建筑信息化和业务流程的自动化看做是数字图书馆。

① Digital Libraries 的电子版可以参考:William Y. Arms 的个人主页. <http://www.cs.cornell.edu/wya/DigLib/>

其次,由图书馆建设的数字图书馆工程往往既包括馆内环境建设,又包括数字资源服务,还包含馆内传统业务信息化改造等工作;而由网络信息服务商(如 Google 等)建设的数字图书馆项目通常只有数字资源服务,而不存在场地和场馆信息化问题。即使都是数字资源服务,图书馆自主建设的数字图书馆和网络信息服务商提供的数字资源服务在服务内容、方式上也往往存在相当的差异。而这些都很难直接利用上述的数字图书馆的定义来区分。

数字图书馆的建设和发展本身就是一个循序渐进、逐步实现的过程,对于它的理解和认识也必然是一个逐步变化和完善的过程,并与当时社会经济环境、技术条件和人类认识水平直接相关。因此,在这个过程中就会出现一些阶段性的相关定义,如自动化图书馆,数字化图书馆,以及最新的云图书馆等。因此,下面我们将引入几个相关概念,来深入理解数字图书馆的内涵。

1.3.1 模拟与数字

数字图书馆产生的前提是数字技术的出现。而所谓的数字技术是与模拟技术相对应的概念。关于数字技术与模拟技术的关系,对于大多数民众来说,最容易理解的可能就是模拟电视和数字电视。从生活角度来说,模拟电视与数字电视的差别主要是通过增加一台数字机顶盒,增加了几十个甚至上百个频道,并具备了一定的交互功能。从技术角度来说,数字电视是指从节目的采集、录制到发射、传输、接收等所有环节中,都采用数字电视信号或对数字电视信号采用数字处理和调制的方法的电视系统。与之对应的模拟电视就是在上述环节中采用模拟技术的电视系统。总之,数字电视所使用的全部是数字数据,模拟电视使用的全部是模拟数据。

关于模拟数据,也称为模拟量,是相对于数字数据而言的,指的是取值范围是连续的变量或者数值。例如温度、压力,以及常见的电话、无线电和电视广播中的声音和图像。而书籍、报刊所承载的文字、图表、图片从本质上说也是一种模拟数据。

数字数据是模拟数据经量化处理后得到的离散的值,典型的就计算机中用二进制代码表示的字符、图形、音频与视频数据。

对这两个概念更宽泛的理解就是,数据在数字化处理以前,全部可以认为是模拟的。数字化完成后就形成了数字数据,在今天的计算机里就成为 0 和 1 组成的离散代码,而且模拟数据和数字数据通过数/模转换技术能够实现相互转换。由于数字数据和模拟数据可以对应并相互转换。因此,理论上凡是模拟数据都可以用数字数据取代,反之也是如此。一个数据从采集、获取、保存、传输、复制、检索、再现等各个阶段都可以采用模拟技术或数字技术。也就是整个信息的所有状态都可以是模拟形式的,也可以是数字形式的。