

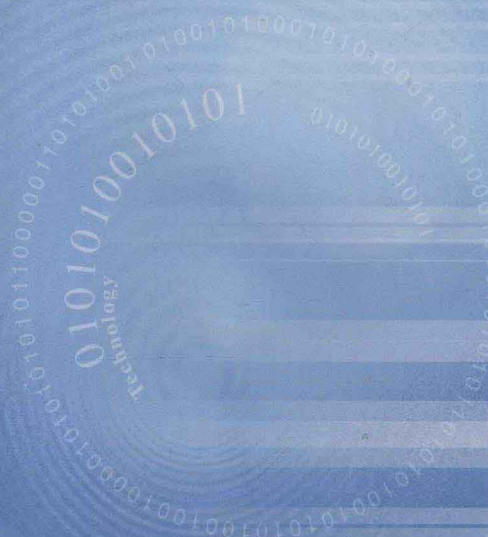
教育部人文社会科学课题(项目编号06JA870001)研究成果

数字 图书馆管理

Shuzi

Tushuguan Guanli

张文德 著



机械工业出版社
CHINA MACHINE PRESS



数字图书馆管理

张文德 著



机械工业出版社

内 容 提 要

数字图书馆自西方发达国家提出后，迅速向全球扩展，数字图书馆管理中的问题，如数字图书馆知识产权价值的评估；如何对数字图书馆知识产权问题进行诊断，并制定合理的策略，以及对数字图书馆知识产权价值进行合理评估，为知识的生产者和使用者在交易时提供一个可靠的依据，以促进经济的发展；在数字图书馆的管理中，如何处理知识产权权利人、网络服务提供商和作品使用者三方利益平衡，从而正确合理使用；数字图书馆的数字权限管理系统价值评估；数字图书馆的一些关键技术问题等，这些都是数字图书馆管理中应解决的问题。

本书注重理论与实践的结合，先后经过了五年多的时间进一步深化完成，有较强的创见性及指导性。

本书适用于图书馆的管理者、图书学、情报学专业、信息管理专业的大学生和研究生，以及关心数字图书馆管理的人士。

数字图书馆管理

机械工业出版社（北京市百万庄大街 22 号 邮政编码 10037）

责任编辑：李书芬 封面设计：张静

责任印制：王书来

北京兴华昌盛印刷有限公司印刷

2009 年 4 月第 1 版第 1 次印刷

184mm × 260mm · 13.5 印张 · 328 千字

标准书号：ISBN 978-7-111-05908-0

定价：26.00 元

前言

数字图书馆自西方发达国家提出后，迅速向全球扩展。数字图书馆建设中知识产权问题是一个非常新的研究课题，如数字图书馆知识产权价值的评估研究；如何对数字图书馆知识产权问题进行诊断，并制定合理的策略，以及对数字图书馆知识产权价值进行合理评估，为知识的生产者和使用者在交易时提供一个可靠的依据，以促进经济的发展；在数字图书馆的建设中，如何处理知识产权权利人、网络服务提供者和作品使用者三方利益平衡，从而正确合理使用；数字图书馆的数字权限管理系统价值评估等，这些都是数字图书馆建设中应解决的关键问题。

本书是教育部人文社会科学课题（项目编号 06JA870001）“数字图书馆知识产权评估与使用管理研究”资助下的研究成果，是在作者的博士论文的基础上，进一步深化完成的。本书突出实用性，通俗易懂，以便于广大读者阅读、理解及参考。

在本书写作过程中，福州大学情报学研究生陈振标、赵需要、曾丽霞、封俊莉、曾彩凤、康云萍、刘丹、王玲艳、陈俊林，计算机专业研究生杨传耀、于志敏、宋艳娟、朱丹红等为本书的完成做了许多有益的工作。可以说，本书是我们这个研究团队多年来的研究成果的总结。

囿于水平，限于时间，书中难免有不妥或错误之处，敬请匡谬补正，欢迎读者批评指正！

张文德
于福州大学

序

数字图书馆是一个分布式的图书馆群体，它通过宽带高速互联的计算机网络，把大量分布在一个地域或一个国家的众多信息资源单位节点组成联合体，把不同地理位置上及不同类型的信息按统一标准加以有效存储、管理并通过易于使用的方式提供给读者，超越空间和时间的约束，使读者在任何时候、任何地方都可以在网上远程跨库获取任何所需信息资源节点的信息，达到高度的资源共享。

数字图书馆的出现，极大程度上提高了人类知识与信息的利用效率。在数字图书馆时代，传播方式发生了巨大的变化，不仅是将传统图书馆的印刷、音像作品等资料数字化，而且是把数字化了的资料上传到网络，成为网上资源。它可以提供各种电子出版物、数据库、互联网上的各种信息、数字化文献等资料。作为一个具体提供服务的信息资源节点，其由两部分组成：一是现实节点馆藏，是本馆实际拥有的各类信息资源的总和。节点单位开发的电子图书、电子期刊、数据库等，存放在节点计算机系统中，节点对现存馆藏有更新、修改、利用和支配的权利。二是虚拟馆藏，是节点以外的通过互联网利用各种稳定的信息资源的总和。

数字图书馆具有强大的信息传播与发布功能。数字图书馆的服务方式与传统图书馆有着重大的差别，它改变了传统的被动式服务为主动服务，可以通过网络随时发布和传播各种文献资源的信息，对读者进行“引导”或“导航”，向读者提供多种语言兼容的多媒体远程数字信息服务。

数字图书馆需要热心这一事业的人的探索，共同解决有关管理、技术等方面的问题，不同学科人将从不同领域对发展数字图书馆作出贡献。情报学的理论、方法、技术是建设数字图书馆重要支柱，同时，情报学也将吸收数字图书馆集成各个学科的养分进一步发展。福州大学张文德博士，担任中国图书馆学会理事、福建省高校数字图书馆专家组组长。在担任图书馆馆长中，主持了福州大学“211工程”的“九五”、“十五”、“十一五”的第一至第三期的数字图书馆建设，有丰富的实践经验，指导了20名计算机技术专业、软件专业、情报学专业的硕士研究生开展了这一方面的研究，还得到了教育部人文社会科学课题（项目编号06JA870001）“数字图书馆知识产权评估与使用管理研究”的支持。他勇于探索，在数字图书馆的新兴领域进行探索，发表了一系列的论文。他在中国科技信息所从事博士后研究期间，将其研究成果，撰写成书出版。面对这部凝聚作者多年心血的书稿，我很受鼓舞，也非常欣慰。该书创新性强，重点突出，观点鲜明。它的出版将对我国数字图书馆的建设有着重要的借鉴和指导意义，对我国数字图书馆的管理将起到促进作用。

中国科技信息研究所所长
研究员、博士生导师
贺德方

2009年2月4日

目 录

第1章 数字图书馆	(1)
1.1 数字图书馆的定义	(1)
1.2 数字图书馆的发展历程	(3)
1.3 数字图书馆的理论模型	(6)
1.4 数字图书馆的发展趋势	(6)
1.5 数字图书馆的知识产权问题	(8)
1.5.1 信息资源建设中的知识产权问题	(8)
1.5.2 信息资源传播中的知识产权问题	(9)
1.5.3 信息资源服务中的知识产权问题	(9)
参考文献	(12)
第2章 数字图书馆的合理使用与利益平衡	(13)
2.1 合理使用概述	(13)
2.1.1 合理使用的起源及概念	(13)
2.1.2 合理使用的特征	(13)
2.1.3 合理使用的判断标准	(14)
2.2 数字图书馆中的合理使用	(14)
2.2.1 数字图书馆合理使用存在的价值和必要性	(14)
2.2.2 版权保护对数字图书馆的影响	(15)
2.2.3 数字图书馆的合理使用制度	(16)
2.2.4 数字图书馆下著作权的合理使用	(23)
2.3 数字图书馆合理使用制度面临的挑战	(26)
2.3.1 版权人的权利呈扩张趋势	(26)
2.3.2 图书馆合理使用的范围相对缩小	(27)
2.4 数字图书馆的利益平衡	(27)
2.4.1 利益平衡概念	(30)
2.4.2 产品的属性和外部经济效应原理的运用	(30)
2.4.3 博弈论理论的运用	(35)
2.4.4 帕累托最优理论的运用	(36)
参考文献	(39)
第3章 数字图书馆知识产权保护机制	(41)
3.1 图书馆自主开发的各种数据库将依法得到保护	(41)
3.2 著作权许可使用制度	(44)
3.3 著作权集体管理组织	(45)

3.3.1 著作权集体管理的定义及特征	(45)
3.3.2 著作权集体管理模式对数字图书馆建设的影响	(46)
3.4 信息网络传播权	(47)
3.5 数字图书馆的知识产权保护的争议范畴	(49)
3.6 数字图书馆的知识产权保护	(51)
3.6.1 解决数字图书馆知识产权问题的思路	(51)
3.6.2 数字图书馆知识产权技术保护措施	(54)
参考文献	(56)
第4章 数字图书馆知识产权诊断与策略	(57)
4.1 作品数字化行为对数字图书馆建设的影响	(57)
4.2 数字图书馆侵权案例归类	(58)
4.2.1 数字图书馆侵权问题	(58)
4.2.2 全文数据库建设著作权问题	(61)
4.2.3 建立数据库权利保护问题	(62)
4.2.4 链接的法律问题	(63)
4.3 数字图书馆知识产权的问题诊断要点	(67)
4.4 面对数字图书馆的侵权应采取的保护对策	(69)
参考文献	(83)
第5章 数字图书馆评估	(85)
5.1 传统的评估方法	(85)
5.1.1 国外无形资产评估方法	(85)
5.1.2 国内著作权评估方法	(87)
5.2 传统评估方法在数字图书馆著作权评估中的应用探讨	(95)
5.2.1 重置成本法的应用	(95)
5.2.2 现行市价法的应用	(96)
5.2.3 收益法的应用	(97)
5.3 四种评估方法的对比研究	(97)
5.4 改进的收益法评估模型	(98)
5.4.1 AHP 方法的导入	(99)
5.4.2 利用 AHP 法构建数字图书馆下著作权评估指标体系	(101)
5.5 模糊评价法的使用	(102)
5.5.1 模糊评价法概述	(102)
5.5.2 模糊评价法进行纠偏	(104)
5.6 修正的收益法	(104)
5.6.1 收益法的缺陷	(104)
5.6.2 修正的收益法评估思路	(105)
5.6.3 具体评估步骤	(105)
5.7 选用收益分成法来评估数字图书馆下的著作权	(107)
5.7.1 收益分成率法的理论依据	(107)
5.7.2 影响收益分成率的因素	(108)

5.7.3 确定收益提成率的方法	(108)
5.7.4 运用专家打分法确定提成率	(110)
5.8 收益分成法评估数字图书馆著作权	(110)
5.8.1 利用 AHP 法构建数字图书馆著作权评估指标体系	(110)
5.8.2 数字图书馆著作权的评估思路	(112)
参考文献	(113)
第6章 数字图书馆著作权评估的实例分析	(115)
6.1 修正的收益法案例	(115)
6.1.1 星星数字图书馆简介	(115)
6.1.2 星星数字图书馆现行版权费用支付模式	(115)
6.1.3 现行版权费用支付模式的问题	(116)
6.2 案例具体评估过程	(117)
6.2.1 案例材料	(117)
6.2.2 根据收益法计算评估初值	(117)
6.2.3 构建评估因素体系并根据 AHP 法计算各因素指标权重	(117)
6.2.4 根据模糊评价法进行双权数综合评价得出纠偏系数	(120)
6.2.5 数字图书馆环境下著作权的公允评估价值	(122)
6.3 具体评估过程	(122)
6.3.1 案例材料	(122)
6.3.2 构建评估因素体系并根据 AHP 法计算各因素指标权重	(122)
6.3.3 确定收益分成法中的提成率	(125)
6.3.4 数字图书馆著作权的评估价值	(126)
参考文献	(126)
第7章 数字图书馆关键技术	(127)
7.1 信息抽取技术	(127)
7.1.1 信息抽取相关技术	(127)
7.1.2 基于规则的信息抽取	(127)
7.1.3 实例:机构信息抽取系统	(130)
7.2 基于内容的检索技术	(131)
7.2.1 CBR 技术分类	(131)
7.2.2 CBR 关键技术:	(132)
7.2.3 CBR 在数字图书馆中的应用	(133)
7.2.4 实例:QBIC 系统简介	(133)
7.3 海量存储	(134)
7.3.1 NAS	(134)
7.3.2 SAN	(135)
7.3.3 iSCSI	(136)
7.3.4 InfiniBand	(137)
7.3.5 网络存储	(138)
7.3.6 实例:国图存储局域网一期工程	(138)

7.4 Web 2.0 技术	(139)
7.4.1 图书馆 2.0——Web 2.0 在图书馆中的应用	(139)
7.4.2 图书馆 2.0 与用户体验	(140)
7.4.3 实例:亚德现代图书馆管理系统	(140)
7.5 网格技术	(142)
7.5.1 网格与数字图书馆	(142)
7.5.2 构建 DLG 的关键问题	(142)
7.5.3 实例:DILIGENT——基于网格的虚拟数字图书馆体系结构	(145)
参考文献	(146)
第 8 章 数字图书馆的数字权限管理	(149)
8.1 数字权限管理技术概述	(149)
8.1.1 数字权限管理技术的组成	(149)
8.1.2 数字权限管理技术体系	(151)
8.1.3 实例:书生公司 SureDRM 版权保护系统	(152)
8.2 流媒体技术的概述	(153)
8.2.1 流媒体技术的定义	(153)
8.2.2 流媒体技术的特点	(153)
8.2.3 流媒体的发展现状	(154)
8.3 流媒体的相关技术	(154)
8.3.1 流媒体的关键技术	(154)
8.3.2 流媒体的传输方式	(155)
8.3.3 流媒体技术原理分析	(156)
8.3.4 流媒体的播放方式	(157)
8.4 主流流媒体 DRM 的应用实例	(158)
8.4.1 微软 WMRM 的结构	(158)
8.4.2 Windows Media DRM 的流程分析	(160)
8.4.3 Windows Media DRM 的工作原理分析	(161)
8.4.4 系统的特点	(162)
8.4.5 系统的安全性分析	(163)
8.5 数字图书馆的数字权利管理	(163)
8.5.1 数字权限管理的未来发展应注意的问题	(167)
8.5.2 数字权限管理全面测量与评估思路	(168)
8.5.3 数字权限管理的系统价值评估	(174)
参考文献	(177)
第 9 章 基于动态多层加密机制的网络 DRM 系统	(180)
9.1 DRM 框架的功能结构	(180)
9.1.1 功能体系结构	(180)
9.1.2 信息体系结构	(181)
9.2 DRM 的典型流程分析	(182)
9.3 网络中数字权限管理的需求分析	(183)

9.3.1 传统的版权保护方式的不足	(183)
9.3.2 网络中流媒体 DRM 的要求	(184)
9.4 基于动态多层加密机制的流媒体 DRM	(184)
9.4.1 流媒体 DRM 的功能结构分析	(184)
9.4.2 流媒体 DRM 的框架	(185)
9.4.3 系统流程分析	(186)
9.4.4 流媒体 DRM 模块的组成	(186)
9.5 流媒体 DRM 的子系统功能	(187)
9.5.1 打包器	(187)
9.5.2 流媒体服务器	(187)
9.5.2 安全播放器	(188)
9.6 系统模块的具体工作	(189)
9.6.1 服务端	(189)
9.6.2 系统客户端	(191)
9.7 系统的关键技术分析	(191)
9.7.1 动态多层加密机制原理	(191)
9.7.2 多层密钥技术分析	(192)
9.7.3 ECM 和 EMM 的基本结构	(194)
9.7.4 密钥分配	(195)
9.7.5 RSA 在智能卡中的实现	(195)
9.7.6 身份认证	(196)
参考文献	(197)
附 录	(198)
附录 A 数字图书馆的著作权价值评估	(198)
附录 B 专家评分表	(201)
附录 C 问卷调查表	(202)

第1章 数字图书馆

随着计算机技术、通信技术和网络技术的迅速发展,信息高速公路的建设与利用为大规模的信息系统、图书馆系统的发展提供了环境和条件。目前,网络信息管理技术、数字化处理技术和数字式信息资源建设已成为国际竞争的焦点,各国都为此投入了相当的实力进行研究和开发。“数字图书馆”这一新概念、新模式应运而生,并被视为21世纪信息产业主要的发展方向之一^[1]。

“数字图书馆”一词由英文 Digital Library 翻译而来。与数字图书馆的蓬勃发展相比,多少有些尴尬的是,到目前为止,学界对什么是数字图书馆仍没有一个一致的定义。因此,本章首先讨论什么是图书馆,然后再叙述数字图书馆的特点及其所面临的挑战。

1.1 数字图书馆的定义

目前,关于数字图书馆的定义,学界众说纷纭,没有一致的定义。下面先列举其中一些较有代表性的观点,然后再给出我们自己的观点。

大英图书馆从传统图书馆的观点出发,认为:“数字图书馆就是利用数字技术获取、存储、存取、发布信息的图书馆。”

美国数字图书馆联盟(DLF)从组织的角度认为:“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。”

中国科学院高文等人从技术的角度将数字图书馆定义为:“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。”

中国科学院文献情报中心孙坦则从社会需求和技术条件的角度分析,认为:“数字图书馆的核心和本质是利用现代信息技术,以计算机网络为基础平台,构建一个有利于产生影响新知识的资源、工具和合作环境。这种作为环境的数字图书馆不仅仅局限于网络数字信息资源的开放利用,更是一个促进信息获取、传递、交流的知识网络。”

上海图书馆刘炜等人从信息的角度出发,认为:“数字图书馆是社会信息基础机构中信息资源的基本组织形式,这一形式满足分布式面向对象的信息查询的需要。其中,分布式是指跨地区、跨物理形态的查询;面向对象是指直接获取一次文献而不是获取一次文献的线索。”

南开大学李培等人从数字图书馆的特征出发,认为:“数字图书馆是同时具备数字资源、网络服务和特色技术三大特征的图书馆。”

而美国总统信息技术建议委员会数字图书馆小组则将数字图书馆定位为:“数字化文本、文档、图像、声音、科学数据和计算机软件的网络集合,将成为当今互联网的核心服务,并将成为未来可供全人类使用的知识仓库。”

关于数字图书馆的定义还有:

中国数字图书馆工程领导小组组长徐文伯^[2]认为:“所谓数字图书馆,就是利用现代信息技术对有高度价值的图像、文本、语音、音响、影像、影视、软件和科学数据库等多媒体信息进行收集,组织规范性的加工和压缩处理,使其转化为数字信息,然后通过计算机技术进行高质量保存和管理,实施知识增值,并通过网络通信技术进行高效、经济地传播、接收,使用户可以在任何时间、任何地点,都能从网上得到各种服务,成为国家的知识基础设施。”从这个观点出发,他认为数字图书馆不是传统图书馆的数字化,但传统图书馆的数字化是数字图书馆重要的信息来源和组成部分^[3]。数字图书馆不是图书馆的专项,信息资源的开发是全社会的事,传统图书馆只不过在占有图书资料方面占有先机。以IBM公司为代表的数字图书馆,是技术解决方案;首都图书馆等单位的数字图书馆是图书信息管理自动化和馆藏数字化的结合;超星公司的数字图书馆是基于网络的信息资源库;863数字图书馆演示系统和中央党校数字图书馆是基于局域网和宽带高速网的信息服务。除了IBM纯技术解决方案,其他数字图书馆模式都是数字图书馆的不同存在形式。但无论何种形式,它都是基于网络平台,并以网站的形式表现出来。从这个角度来看,数字图书馆则是网站上支持多种搜索功能的强大的海量数据库。

美国十大数图专家之一 William Y. Arms 认为:数字图书馆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。

笼统的“数字图书馆”包括两种^[4]:数字图书馆和数字化图书馆。前者是指提供电子图书的治理和服务系统平台,后者主要是把图书馆中的纸质图书扫描,OCR或者录入校正的作业平台;二者都含有图书馆的基本业务流程,例如,编目。一般来讲,数字化图书馆图书的分类是采用中图法,而数字图书馆的分类在采用中图法的同时还会兼顾其他分类法;在形式上,数字图书馆对采用图书服务网站的方式对外服务;而数字化图书馆是为图书馆服务,一般是一个由纸质到电子档的服务过程。

数字图书馆^[5]是用数字技术处理和存储各种图文并茂文献的图书馆,实质上是一种多媒体制作的分布式信息系统。它把各种不同载体、不同地理位置的信息资源用数字技术存储,以便于跨越区域、面向对象的网络查询和传播。它涉及信息资源加工、存储、检索、传输和利用的全过程。

数字图书馆^[6]是一门全新的科学技术,也是一项全新的社会事业。简而言之,就是一种拥有多种媒体内容的数字化信息资源,能为用户方便、快捷地提供信息的高水平服务机制。虽然称之为“馆”,但并不是图书馆实体:它对应于各种公共信息管理与传播的现实社会活动,表现为种种新型信息资源组织和信息传播服务。它借鉴图书馆的资源组织模式、借助计算机网络通信等高新技术,以普遍存取人类知识为目标,创造性地运用知识分类和精准检索手段,有效地进行信息整序,使人们获取信息消费不受空间限制,很大程度上也不受时间限制。其服务是以知识概念引导的方式,将文字、图像、声音等数字化信息,通过互联网传输,从而做到信息资源共享。每位拥有任何计算机终端的用户只要通过联网,登录相关数字图书馆的网站,都可以在任何时间、任何地点方便、快捷地享用世界上任何一个“信息空间”的数字化信息资源。

通俗地说,数字图书馆就是虚拟的、没有时空限制的图书馆,是基于网络环境下共建共享的可扩展的知识网络系统,是超大规模的、分布式的,便于使用的,可以实现跨库无缝链接与智能检索的知识中心。数字图书馆既是完整的知识定位系统,又是面向未来互联

网发展的信息管理模式,可以广泛地应用于社会文化、终身教育、大众媒介、商业咨询、电子政务等一切社会组织的公众信息传播。

数字图书馆^[7]是运行在高速宽带网络上、分布式超大规模、可跨库检索的海量数字化信息资源库群。数字图书馆最大限度地利用各地已经搭建的软、硬件环境,对信息资源进行接收、存储和服务。它需要拥有超大容量的分布式资源数据库群,要为用户提供可靠的信息就需要配置高性能的软、硬件基础设施。通过网络传输的信息资源是按统一标准规范加工整合过的数字化信息资源,它包括:图书、戏曲、音乐、舞蹈、电影、文化旅游、农业科技、科普知识、法律知识和文化新闻等。它超越了人们意识中图书的概念。

此外,美国研究图书馆协会(ARL)^[8]通过归纳了20世纪90年代前期,流行的数字图书馆的各种定义中具有共性要素,得出数字图书馆的五要素:①数字图书馆不是一个单一实体;②数字图书馆需要链接许多信息资源的技术;③多个数字图书馆及信息机构之间的链接对最终用户透明;④全球范围存取数字图书馆与信息服务是一个目标;⑤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。

数字图书馆是利用网络来提供服务的、一个经济性的、拥有许多数字资源的图书馆;而无经济性数字图书馆,即通常所说的公益性图书馆,应认为是该数字图书馆的延伸体。

我们认为,数字图书馆是一种社会工程。它以为全民提供高效、便捷、低成本方式获取所需信息和知识为目标,以数字资源生成、组织、存储、检索、表现、传递及知识化为生命周期,并由此涉及了一系列相关的计算机、网络技术以及社会伦理道德。

1.2 数字图书馆的发展历程

数字图书馆是图书馆学和信息科学结合的产物。它经历了3个阶段,即以MARC为代表的图书馆的计算机管理阶段;以Dialog为代表的网络化电子文献服务阶段;以American Memory、CDL等为代表的数字图书馆阶段,而数字图书馆的由来可以追溯到20世纪40年代。

1945年,美国著名科学技术管理学家布什(V. Bush)在《大西洋月刊》上发表了《诚如我们想象的那样》一文。文中他首次提出将传统的图书馆馆藏文献的储存、查找机制与计算机结合起来,构思并描述了他所设想的一种Memex装备——一种机械化的个人文档与图书馆,它拥有存储书、记录和通信功能,是一种台式个人文献工作系统。这被公认为计算机辅助检索的先驱。

1948年,美国数学家维纳(N. Wiener)指出电子计算机将能够在图书馆运作中大显身手。

1965年,Licklider在他的著作《未来的图书馆》中首次提出了在世界范围内收集出版的书籍,并把它们放到互联的计算机上进行共享的设想。

1978年,美国著名图书馆学家兰卡斯特(F. W. Laneas)发表了《走向无纸的时代》和《电子时代的图书馆员》两部著作,第一次向世人详细描述了电子图书馆的前景。

1979年,英国不列颠图书馆的哈利(A. J. Harley)提出了“虚拟图书馆”的概念。

1988年,美国国家科学基金会(NSF)的伍尔夫(W. Wulf)撰写国际合作白皮书,正式提出了数字图书馆的概念。

20世纪90年代后,数字图书馆进入高速发展阶段,亚欧美洲中的许多国家分别推出

了自己的数字图书馆计划。

1994年9月,美国国家科学基金会正式公布了一项为期4年、投入2440万美元的“数字图书馆首创计划”。在该计划的影响下,美国大学和学院图书馆率先进行数字化建设,如以斯坦福大学、卡内基梅隆大学等6所大学为主的美国数字图书馆研究基地和数字图书馆模型建立的基地。以美国国会图书馆、美国保存及存取委员会、哥伦比亚大学、哈佛大学、国家档案记录管理委员会、纽约公共图书馆、普林斯顿大学、耶鲁大学等为主体提出了“全国数字图书馆联合协议”。同年10月,美国国会图书馆推出数字化项目,使该馆馆藏逐步实现数字化,并领导与协调全国的公共图书馆、研究图书馆,将其收藏的图书、绘画、手稿、照片等转换成高清晰度的数字化图像并存储起来,通过互联网供公众使用。由此展开了整体意义上的数字图书馆建设。

1993年,英国发起“国家图书馆的存储创新协议”,以信息存储方式的变革为出发点,共涉及20个项目,目前大部分已经完成,并在互联网或馆域网上提供服务。

1994年,新加坡政府从整体的国家信息化的角度来考虑,提出了国家“智能岛”计划。该计划的目标是将新加坡所有公共图书馆和500多个学术机构和软件业等连接,实现真正意义上的国家网络化信息中心。

1994年,日本国会图书馆开始实验建设电子图书馆,目标是将其关西新馆建成日本最大的数字图书馆,使之成为日本文献和信息提供中心、亚洲资源信息中心。

1995年3月,新建的法国国家图书馆致力于将该馆收藏的100万册图书数字化,形成两个大型数据库,包括约200万条书目数据。该馆特别重视“文化遗产”数字化,将该馆收藏的艺术精品及分散在法国各地的古书艺术插图用彩色高分辨率扫描仪录入光盘。

1995年5月29日,法、日、美、英、加、德、意7国的国家图书馆在法国成立了G7全球数字图书馆集团。后来,俄罗斯也参加了这个全球信息社会电子图书馆项目。该项目的目标是从现存的数字化项目中组织一个大型的人类知识的虚拟馆藏,通过网络为广大公众服务。

1996年8月,德国联邦内阁会议正式通过了德国1996—2000年信息技术发展计划,中心内容是建立全球电子图书馆和开展电子出版业服务。

1997年,中国“973计划”之“试验型数字图书馆”项目立项,2001年5月鉴定验收。该项目创建了一类多馆合作的网络内容资源建设和共享体系,实现了一个基于分布环境的、以藏品建设为基础的数字图书馆应用系统。该系统从功能上覆盖了对内容资源从采集加工、处理、储藏组织、管理调度、资源发布、用户利用等全过程,并支持分布式网络环境下多馆合作资源共建共享的模式。该项目完成了系统构架的研究,创建开发了数字资源加工系统、调度系统、资源发布系统和用户界面等。应用上述系统,该项目完成了900GB多媒体资源的建设与发布,在互联网环境下是用户一侧实现了透明的页面级的无缝跨库检索与连接。同时,该项目初步实现了一个包括旅游、名人、军事、大百科等数字资源库的数字图书馆系统,并为这些资源库设计了公共的元数据集合,建立了相应的元数据集合和元数据库。

1998年,美国开始数字图书馆首创计划第二期工程(DLI-2)^[9]。DLI-2更强调与数字图书馆相关的各种问题研究,强调以人为中心、以系统为中心、以内容和藏品为中心,扩大学科研究范围,探讨与数字图书馆建设相关的经济、社会、法律和政策等问题,如信息共享格式、信息共享的国际标准、数字图书馆网站的可靠性和稳定性等。它大大拓

展了数字图书馆的研究与开发,增加了音频、软件、视频等媒体藏品;扩充了民间文学、流行歌曲、文学手稿、Web 资料等新资源形态;增添了考古学、经典著作、社会科学、音乐、医学、生命科学等多个学科领域。

1998 年,世界银行提出国家知识基础设施 (China National Knowledge Infrastructure, CNKI) 的概念^[10]。1999 年 6 月,清华大学、清华同方发起并创建了 CNKI 工程。CNKI 工程的具体目标:一是大规模集成整合知识信息资源,整体提高资源的综合和增值利用价值;二是建设知识资源互联网传播扩散与增值服务平台,为全社会提供资源共享、数字化学习、知识创新信息化条件;三是建设知识资源的深度开发利用平台,为社会各方面提供知识管理与知识服务的信息化手段;四是知识资源生产出版部门创造互联网出版发行的市场环境商业机制,大力促进文化出版事业、产业的现代化建设与跨越式发展。经过多年努力,CNKI 工程集团采用自主开发并具有国际领先水平的数字图书馆技术,建成了世界上全文信息量规模最大的“CNKI 数字图书馆”,正式启动建设《中国知识资源总库》及 CNKI 网格资源共享平台。通过产业化运作,为全社会知识资源高效共享提供最丰富的知识信息资源、最有效的知识传播与数字化学习平台。

1999 年 2 月,俄罗斯联邦政府批准俄罗斯科学部、文化部、教育部等 9 个主管部门参加《俄罗斯电子图书馆》部际规划,年总预算大约 2 亿卢布(相当于 620 万美元)。

2000 年 2 月 1 日,韩国总统金大中在国会会议上作出建立“国立数字图书馆”的指示,拉开了韩国数字图书馆建设的序幕。

2000 年 4 月,中国数字图书馆工程启动^[11]。中国数字图书馆工程按照国际上数字图书馆的主流技术方案,采取自主开发与引进相结合的方式,建设分布式、可扩充的、具有自主知识产权的中国数字图书馆系统。它组织全国有关单位,联合建设超大规模资源库群,建设知识网络,使其不但为科技、文化、教育服务,而且在提高国民素质和知识经济中发挥着重要作用。

2002 年 9 月,欧盟建立 GRACE 项目^[12,13]。GRACE 项目是一个关于网格信息检索的项目,其目的是用可靠、廉价的一站式方式为用户提供海量资源访问服务^[14]。该项目的特色在于融合了本体概念和语义互用。它不使用集中式的数据库,而是在每个网格节点内索引文档,并使用和存储位置邻近的计算资源。这种分布式的应用比以前的集中式数据库更加规模化,并且可以为不同的数据源提供更多的个性化的搜索选项。同时,该项目使用数据网格技术及本体及语义概念,对语义索引、指定实体的抽取、推理检索、语义地图进行研究,以期开发出基于网格的支持多种语言能力的分布式搜索和分类引擎,并创建功能强大的网格工具组。

2003 年 12 月,澳大利亚国家图书馆^[15]根据 Kinetica Futures Group 于 2003 年进行的澳大利亚图书馆系统选择的调查结果,宣布进行为期两年的 Kinetica 改进项目 (Kinetica Redevelopment Project),将书目数据管理系统和资源发现系统进行整合,建立 TeraText 平台支持大型书目数据库,以满足对图书馆检索和数据服务功能的需求。2005 年 11 月 30 日,Kinetica 系统改进项目完成。该系统重新命名为 Libraries Australia,通过传递 MARC 记录和提供文档传送服务支持澳大利亚图书馆整体合作和资源共享,系统实现 24 小时连续服务。

2004 年 1 月,美国 Old Dominion 大学成立 DLG 项目^[16,17]。DLG 项目的目的是在原有的 OAI 系统的基础上,通过网格技术来降低数字图书馆的整体运行成本,同时增强服务

的可用性和服务质量,解决现有数字图书馆因使用不同互操作技术,导致接口难以统一的问题。DLG项目主要研究和探索下列问题的解决方案:使现有的OAI数据提供者能适用于网格环境的插件;使现有的发布工具,如DSpace、Kepler能在网格环境下使用;网格环境下的数据采集策略;制定针对数据提供者的强制性标准;能够表达被采集的元数据用于什么目的,在什么约束条件下被谁使用的信息描述规范。

2004年9月,欧盟拨款630万欧元启动DILIGENT原型项目^[18]。DILIGENT的目标是创建一个先进的试验平台,使动态虚拟的组织成员能够以安全、协调、动态和低成本高效益的方式开展合作,存取共享的知识。事实上,启动DILIGENT项目是为了创建一个由网格技术支持的先进数字图书馆基础架构。该架构允许来自动态虚拟组织的成员在共享计算资源、存储资源、多媒体资源、多类型内容资源,以及多类型应用资源的基础上建立一个临时的虚拟数字图书馆(Virtual Digital Library, VDL)。DILIGENT原型项目的建设目标表明,欧洲数字图书馆的建设正朝着与网格技术相结合、面向e-Science需求的虚拟数字图书馆迈进。

1.3 数字图书馆的理论模型

从狭义上讲,数字图书馆就是数字化的信息资源库。它应有以下几个特性:

1. 资源数字化

资源数字化是数字化图书馆的首要特征和必备条件。数字化图书馆的数据、文本、语音、图形图像和视频信息等所有信息资源都是经过计算机处理的数字化信息,以便提供更快捷、方便的服务。

2. 流通网络化

传统图书馆向读者提供的是直接面对面交流的流通服务,而数字图书馆使读者获取的查询、检索、预约、下载等一切的服务几乎都是通过计算机网络完成的。

3. 资源共享化

本地或者远程用户可以对图书馆提供开放的、有组织的信息资源库进行访问,获取自己需要的最终信息。同样的个体信息源,如一篇文章或者一个视频作品,可以提供给不同地点的几十个,甚至是几百个以上的用户同时利用,实现了真正意义上的资源共享^[21]。

4. 用户参与化

数字图书馆工程是一个社会工程,它的建设不能仅依靠某个单位或组织,它需要全社会的参与,特别是广大用户的参与。

数字图书馆的理论模型如图1-1所示。

1.4 数字图书馆的发展趋势

随着计算机和网络技术的研究和发展,数字图书馆正在从基于信息的处理和简单的人机界面逐步向基于知识的处理和广泛的机器之间的理解发展,从而使人们能够利用计算机和网络更大范围地拓展智力活动的的能力,在所有需要交流、传播、存储和利用知识的领域,包括电子商务、教育、远程医疗等,发挥极其重要的作用。

从全球的数字图书馆的发展来看,除了已经实现的资源数字化、服务网络化之外,数字图书馆还将朝着以下几个方向发展:

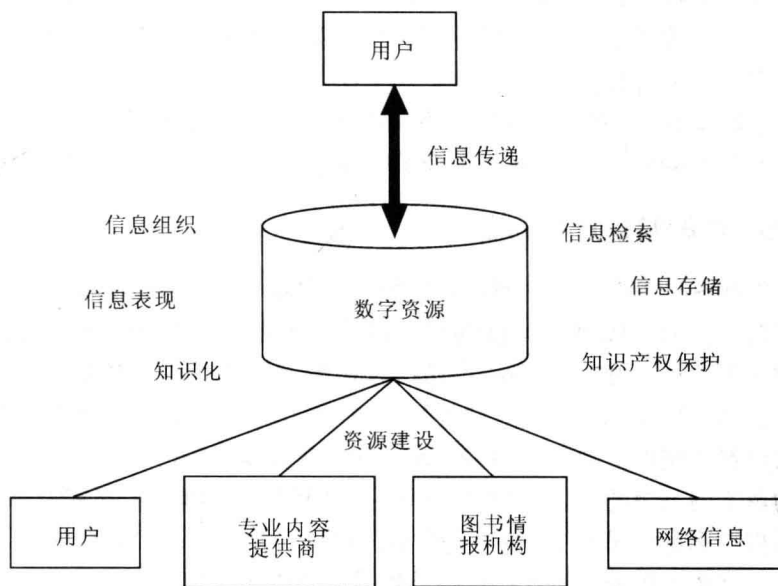


图 1-1 数字图书馆的理论模型

1. 体系结构动态配置化

为了让任何人在任何时间、任何地点都能以一种友好的、有效的、高效的方式访问人类知识，数字图书馆体系结构将由集成的、集中控制的架构转向可动态配置的、自适应的数字图书馆联盟架构，面向服务的体系结构（SOA）、网格（Grid）技术以及对等网（P2P）技术为这种转变提供了技术基础。

2. 检索技术智能化

随着自然语言检索、基于内容检索、跨语言检索等技术的发展，数字图书馆检索将实现智能化检索，检索结果也有望知识化。

3. 服务途径多样化

随着信息技术的发展，网络终端也逐渐多样化。未来的数字图书馆将为用户提供多种终端服务，包括各种有线和无线终端、移动和固定终端等。

4. 用户体验 2.0 化

Web 2.0 的兴起，不仅让用户获得了更好的网络体验，更让用户获得了参与资源建设的机，从而极大地激发了网络草根的积极性、创造性，也增加了用户黏性。在 Web 2.0 浪潮的影响下，以网络服务为主的数字图书馆也开始借鉴 2.0 思想、应用 2.0 技术，提升自身服务意识和水平。可以预见，2.0 化的数字图书馆一定会为用户提供优异的使用体验。

5. 用户阅读无纸化

目前，海量的数字化内容只能通过计算机、液晶显示屏表现出来，这与用户的阅读习惯相悖。但随着电子纸技术的成熟，用户能够在保持传统纸质阅读习惯的前提下阅读多媒体化的数字内容，无纸化阅读将不是梦想。

6. 资源建设全球化和社会化

数字图书馆工程涉及的内容很广，仅依靠某个单位或组织的能力显然是无法完成的。互联网的无国界特性又使合作的空间进一步扩大，应当考虑在全球范围内实现人才、技