

社会保障管理问题丛书

SHEHUIBAOZHANGUANLIWENTICONGSHU

数字图书馆知识管理

肖 芃 / 著



 吉林大学出版社

数字图书馆知识管理

肖 芃/著

江苏工业学院图书馆
藏书章

吉林大学出版社

数字图书馆知识管理

肖 芃 著

责任编辑、责任校对:陈颂琴

封面设计:张沐沉

吉林大学出版社出版
(长春市明德路3号)

吉林大学出版社发行
长春市永昌福利印刷厂印刷

开本:850×1168 毫米 1/32
印张:6.5 插页 4
字数:160 千字

2004 年 12 月第 1 版
2004 年 12 月第 1 次印刷
印数:1—500 册

ISBN 7-5601-2716-9/C·87

定价:18.00 元

目 录

第一章 数字图书馆的基本问题

- 第一节 什么是数字图书馆..... 1
- 第二节 数字图书馆的社会功能及发展前景 17
- 第三节 建设数字图书馆的社会意义 28

第二章 数字图书馆知识管理的引入

- 第一节 知识管理的基本问题 39
- 第二节 传统的信息管理面临挑战 49
- 第三节 从信息管理到知识管理 75

第三章 数字图书馆知识管理的主要内容

- 第一节 数字图书馆知识管理的必要性和可行性
..... 101
- 第二节 数字图书馆知识管理的内容..... 111
- 第三节 数字图书馆知识管理的特征..... 127

第四节	数字图书馆知识服务的功能与特点·····	129
-----	----------------------	-----

第四章 数字图书馆知识管理实施的现实条件和策略

第一节	数字信息源是数字图书馆知识管理的核心 ·····	138
第二节	数字图书馆的网站建设·····	151
第三节	知识管理在数字图书馆中的实施·····	162
第四节	我国数字图书馆工程中有待解决的问题 ·····	173
第五节	数字图书馆知识管理与人才管理·····	187

第一章 数字图书馆的基本问题

第一节 什么是数字图书馆

一、数字图书馆的定义

数字图书馆是伴随计算机技术、网络技术而发展起来的新事物。正如我们所看到的计算机和网络正在不断发展、不断升级换代一样,数字图书馆也处在生长期中,还没有完全定型,其突出的表现就是在理论上有关数字图书馆的定义可谓见仁见智,众说纷纭。刘炜等著的《数字图书馆引论》一书中有这样一段话:“我们整理过的定义就有近百种之多,几乎每个研究项目都有一个自己的界定,有多少著作就有多少定义。”这反映了人们正在积极地思考问题,是十分可喜的现象。

在众多关于数字图书馆的定义中,我们发现大多数定义都是列举式的、解释式的,例如:“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搜集,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。”

又如：“数字图书馆建设是以统一的标准和规范为基础，以数字化的各种信息为底层，以分布式海量资源库群为支撑，以智能检索技术为手段，以电子商务为管理方式，以宽带高速网络为传输通道，将丰富多彩的多媒体信息传递到千家万户。它涉及数字信息资源的生产、加工、存储、检索、传递、保护、利用、归档、剔除等全过程。”

我们认为数字图书馆的定义应该也可以更简洁、更概括、更抽象，用简明的语言反映出数字图书馆的本质。

在试图给出关于数字图书馆定义的时候，我们注意到美国数字图书馆特别小组在《数字图书馆：对人类知识的普遍访问》报告中的一段话：“所有人在任何时间任何地点都可以用任何连接互联网的数字设备来访问所有人类知识”。这段话不是对数字图书馆进行定义，但比较简洁地表达了人们对于数字图书馆功能的一种“理想目标”，并且指出了数字图书馆的几个基本点：一是向用户提供信息存取服务，二是这种服务是以数字化信息为基本内容，三是这种服务是通过网络和连接网络的数字设备实现的，四是用户在其中扮演着主动角色。我们进一步研究认为，对这个“理想目标”还可以进一步完善。

参考众多的关于数字图书馆的定义、解释，结合数字图书馆所要做的主要工作、主要工作方式及实现这些工作所依赖的条件，以及我们对数字图书馆功能的设想，数字图书馆可以简明地定义为：

数字图书馆是运用信息技术以数据库为基本方式组织数字信息并通过互联网络向用户提供存取服务的信息系统。

这一定义包含了以下几层意思:

1. 说明了数字图书馆的主要工作内容的两个层面: 进行数字信息的组织和提供服务;
2. 说明了数字图书馆组织数字信息的基本方式是建立数据库;
3. 说明了数字图书馆的技术基础是信息技术 (主要是计算机技术、通信技术和网络技术);
4. 说明了数字图书馆的依存条件是互联网络;
5. 说明了数字图书馆的服务对象是用户或者说用户在数字图书馆中的地位;
6. 说明了服务方式是具有交互性的存取;
7. 点明了数字图书馆是一个具有复杂内在结构的信息系统;
8. 隐含了数字图书馆与社会环境的交互关系。

二、数字图书馆与传统图书馆的关系

在讨论数字图书馆的时候, 不能回避的一个问题就是数字图书馆与传统图书馆的关系。基本上有两种观点。一种观点认为, “数字图书馆是正在成长中的信息技术和信息服务方式, 目前世界上大部分数字图书馆是在传统图书馆基础上发展起来的, 它部分沿袭了图书馆的业务框架, 大量应用自动化、数字化和网络化技术, 改变业务流程, 进行服务创新。数字图书馆推动了资源建设、文献整理、信息服务、资源共享等领域的巨大变革, 并且引入了知识导航、电子商务和信息营销的新观念和新技术。我们把这些变化结果综合起来看, 就可以勾勒出一个数字图书馆的雏形, 因此, 数字图书馆既与传统图书馆紧密结合, 又促进传统图书馆的变革, 提升传统图书馆的服务能力

和服务质量。”另一种观点认为,“数字图书馆植根于现代信息技术的土壤,是一种网络环境下的全新的图书馆形态,具有与传统图书馆完全不同的理念追求、运作方式和管理模式。”

从形式上看、从方法论上看,前一种观点是求同,后一种观点是求异。事实上问题不是这么简单。前一种观点,描述了一种现存现象,但对现象背后的原因与根据缺乏深刻的分析与研究。与此相对应的是,后一种观点就登堂入室抓住了数字图书馆与传统图书馆的要害,从而得出完全不同的结论。

关于数字图书馆是对传统图书馆的革命性变革问题,试举一例予以说明。

分类法是图书馆的核心问题。为了解决图书组织,图书馆专家们设计不同类型的分类方法,成为图书馆最值得自豪的标志性成果。这一成果现在却遭到了严厉的冲击,或者说正是在这个关键问题上,数字图书馆与传统图书馆是相互区别的。冲击的第一步是互联网上索引网站毫不理会图书馆的分类法,而是根据网民的方便与习惯自创一套分类体系,并得到网民的认可;一些大型数据库如中国学术期刊网数据库也是自创分类体系。冲击的第二步是 Dublin Core Element Set 等元数据集以崭新的著录格式代替了传统图书馆的分类法为基础的著录格式。Dublin Core Element Set 在分类方面集知识分类与主题分类为一体,并具有简单明了的优点,不需要太专业的训练即可使用。Dublin Core Element Set 的提出一是基于现代技术的支持,二是基于需要处理的信息量呈指数增长,三是面向用户使用(用户不是也不可能是图书馆专家)。如果说图书馆界对冲击的第一步还愤愤不平的话,对冲击的第二步却是心悦诚服的。虽然到目前为止,完善的数字图书馆分类体系还没有形成,元数据格式也在研拟过程中,但人们已经不再怀疑,在最核心的分

类与著录问题上, 数字图书馆与传统图书馆是迥然相异的。

数字图书馆与传统图书馆的区别是多方面的。下表是数字图书馆与传统图书馆在工作内容方面的区别简表。

数字图书馆与传统图书馆工作内容和模式对比简表

	工作中心	馆藏基础	工作方式	馆藏加工	服务对象	检索手段	服务方式
传统图书馆	馆藏	印刷型	手工作业	不加工	到馆读者	手工检索卡片	手工服务
数字图书馆	用户	数字信息资源	对文献内容进行自动化加工	加工, 并使馆藏具有增值效应	全球用户	对文献内容进行智能检索	同上服务

数字图书馆是新形态的图书馆, 但为什么在许多方面还与传统图书馆相似呢? 有专家指出, 数字图书馆与传统图书馆“在概念上没有必然联系, 但是由于目前资源主要集中在图书馆, 因此必然要借鉴图书馆的已有理论和信息管理。”

三、数字图书馆的结构模型

根据前述对数字图书馆的理解, 数字图书馆作为信息服务系统, 其结构框架可以由 9 个子系统组成的。下面分别论述这 9 个子系统。

1. 多类型用户

用户是数字图书馆的服务对象, 为用户服务是数字图书馆建设的起点和归宿。为什么要建数字图书馆? 因为用户需要它, 它可以为用户提供标准更高质量更好的信息服务。怎样评价数字图书馆? 用户的满意度是客观标准, 舍此无他。因此,

传统图书馆所奉行的“读者第一”的原则在数字图书馆中得到了很好的继承和发扬。在这里，用户是中心，而且必须是中心。满足用户的信息需求是数字图书馆的第一信条。

同传统图书馆系统相比，在数字图书馆系统中的用户具有更加重要的地位、发挥着更加重要的作用。他们不仅是信息的接受者，更是信息的创造者。就作为信息接受者的角度来说，他们不再是被动地接受信息，而是更加主动地发现信息。用户的存在是数字图书馆存在的客观基础，忽视用户的想法是不可接受的。在数字图书馆中要更加重视用户、更细致地分析用户需求。不仅如此，数字图书馆还要注意每一个用户，满足每一个用户的信息需求，更提倡个性化服务。当然也只有在数字图书馆的条件下，才能做到这一点。

用户既然是中心，就必须认识用户，才能把握用户的信息需求。依据不同的标准可以将用户划分为不同的类型、不同的层次，即不同的群体。而不同的群体的信息需求是相互区别的。首先，可以把用户区分为团体用户与个人用户。由于团体用户经常要落实到个人头上，所以在称用户时一般指个人用户。在个人用户中以年龄段为标准，可以划分为儿童、青年、中年、老年，就以娱乐欣赏需求来说，每一个年龄段的差别都是十分明显的。如果以职业来区别用户，则可以区分为工人、农民、军人、公务员，等等，其间的信息需求差别是不言而喻的。只有对用户进行科学的分析，把握其间不同的信息需求，才能做好信息服务工作。

2. 数字化资源

一本本的图书是传统图书馆赖以存在的“物质”基础。在数字图书馆系统中，以数字方式存在的信息资源是数字图书馆存在的“物质”基础。如果没有相当量的数据库作支撑，那就

不能称之为数字图书馆。数字图书馆“大小”的衡量标准主要是其所拥有数字信息资源的多少。因此,有没有数据库是数字图书馆的质的规定,数据库中信息资源的多少是数字图书馆量的规定。

数字图书馆资源必须是以数字化方式存在。传统图书馆的基础是各种实物资料,数字图书馆的存在方式是信息脱离原有的实物形态以数字方式存在。不论是书籍、期刊、报纸、手稿、照片、地图、录音、录像、电影,都要经过数字化处理,以数字化方式存储,才能成为数字图书馆的现实“馆藏”。

数字图书馆的数字化资源还有数字化程度的区别。按数字化程度分:

- (1)标引性资源:仅有标引信息入库,资源本身未数字化;
- (2)数字化资源:经数字化手段转换成的资源,如扫描图书、经 OCR 但未校对的资源;
- (3)数据化资源:按照规定标准(如 XML)格式化并归档的资源;
- (4)信息化资源:在数据化基础上,针对用户(或用户群)抽取信息,从而实现增值,简单地说,是附加了相关信息的资源;
- (5)知识化资源:在信息化基础上,抽取知识,从而进一步增值,简单地说,是附加了相关知识的资源;

数字化了的资源必须要有科学的合理的组织。数字图书馆和传统图书馆从事的工作实际上都是进行知识组织以方便用户。区别在于传统图书馆是以书籍、期刊的册(或种)为单元进行组织,而数字图书馆舍弃了知识赖以存在的书籍、期刊、报纸的外形,以知识点为单元对象进行组织。

数字图书馆的目的是直接提供读者所需的最终信息,而不

只是二次文献（获得文献的线索），但数字图书馆并不排斥书目数据、索引文摘等二次文献，因而它们也是数字图书馆的组成部分。

3. 立体化网络

在数字图书馆内，数字信息与用户之间是如何连接的呢？是数字通信网络。高速的数字通信网络是数字图书馆的存在基础。数字图书馆依附于网络而存在，其对内的业务组织和对外的服务都是以网络为载体，得益于网络也受制于网络，离开了网络，数字图书馆的功能就不能实现，其优越性也难以表现。

数字图书馆的网络首先是内部局域网，其次是与国际互联网接通。数字图书馆内部本身由局域网构成，一般是高速主干联接若干服务器及工作站，外部通过若干广域网服务器面向浩瀚的国际互联网。

数字图书馆赖以存在的网络应当是立体化的，才能满足用户在“任何地点”都能上网的要求。在电信网、数据通讯网、电视网覆盖的范围以外，应该还有支持移动通讯的卫星、微波网给予支持，方能使持移动终端设备的用户能在野外上网，获取所需信息。

4. 高可靠硬件

数字图书馆当然不能缺少各种各样的硬件设备，如信息转换设备、信息处理设备、信息存储设备、信息交换设备、信息接收设备等等，以及存放这些设备的场所。这些都是数字图书馆不可缺少的基本物质条件。如果这些基本物质条件不具备，要进行数字图书馆建设是无从说起的。所以购置这些设备是进行数字图书馆建设最起码要付出的投资，有了这个投资才有开展工作的起码条件。

数字图书馆所赖以存在的硬件设备在性能上要求不仅速度

要快而且要有高可靠性。数字图书馆的运转应该是一年 365 天、一天 24 小时、1 小时 60 分钟全天候不间断提供服务,才能满足用户“在任何时间”的信息需求。要实现全天候不间断服务,设备的可靠性要求是越高越好。设备运行要稳定、少出故障尤其是不可恢复性故障。

5. 智能化软件

在数字图书馆系统中软件的地位尤其重要。虽然各种系统软件、应用软件是使各种计算机系统、网络系统得以正常运行的东西。没有它们计算机系统就不能运转,成为没有灵魂的“菩萨”。有了它们,各种硬件设备才获得了“生命”,得以运转起来。尽管软件是硬件的灵魂和生命,但它们不能独立存在而只能依附于各种硬件设备。由于这个原因,致使有的人只注意到软件的“依附性”而忽视了其“独立性”,没有给予软件以应有的地位。

在数字图书馆系统中运行的软件是多种多样的,有的软件同其他系统运行的软件没有更多的要求,但在有些方面却有着特别的要求。这就是要求更高的智能化水平。就以检索查找软件来说,只有高智能化方能满足用户的需求。例如在回应用户查询时,能接受用户的“自然语言”。能对用户的查询请求进行归纳、分析,等等。现在使用的支持全文自由词检索的搜索引擎在给我们带来很大的方便的同时,我们也更加感到搜索引擎智能化水平亟待提高。面对海量信息,用户的要求是越来越高。数字图书馆必须根据用户的需求来提供服务,用户需要什么就提供什么,而且是无遗漏的,也就是既准确又全面,用图书馆原有的术语就是既有较高的查准率也有较高的查全率。如果用户不需要的就不应该出现,以节约用户的时间与精力。从这一点出发来设计的搜索引擎,就应该超出基于字词检索的水

平，达到基于知识或基于内容检索的水平。这仅是软件中的一个小的方面，在数字图书馆中软件本身又是一个复杂的“系统”！

6. 多样化终端

作为用户与机器系统交互的中介就是用户终端。用户终端包括硬件和软件两个方面。硬件方面就是显示器，软件方面就是用户人机界面（简称用户界面）。

数字图书馆作为新形态的工具系统同其他新事物一样被人们接受需要一个过程，这是可以理解的。在这个过程中，数字图书馆系统被人责难、怀疑最多的是用户终端：嫌它有线，不能离线使用；嫌它笨重，不便携带；嫌它界面难看，不符合阅读习惯；等等。直到 2001 年还有文章说：“即使今后普及掌上电脑，可以随身携带阅读，但其分辨率不佳，屏幕闪动，视角有限，不仅容易疲劳，且难免辐射。绝大多数读者，特别是中老年人只是在不得已的情况下才通过电脑查阅资料。”随着技术的进步，用户终端在不断改进，人们的责难一步步在解决，但我们必须指出的是，用户终端的改进空间还很大。可以说，用户终端的发展水平与数字图书馆的发展有着直接的关系。数字图书馆的发展需要开发多样化的用户终端设备。根据用户在不同场合不同环境的不同需要开发不同类型的终端。这些用户终端还要有较高的性能要求，如符合环保的、卫生的要求，不对人的视力产生不良影响。

WEB 界面的出现，已经使人机界面大大改善，但这还是初步的，还需要更友好的人机界面，以增加用户的亲近感，使人机交互更容易进行；还需要更智能的人机界面，能“理解”用户各种问题，满足用户较为复杂的信息需求。

7. 高效率管理

数字图书馆的管理是指运用现代化的管理方法和理念,借助现代化的科学技术,根据数字图书馆信息管理过程中的基本规律,运用数字图书馆的管理职能、管理方法与手段,充分调动、激活组成数字图书馆系统的诸要素,加强子系统及诸要素之间的联系,优化数字图书馆系统的运作过程,使数字图书馆充分实现其功能。它包括的内容十分丰富,涉及信息活动的各种要素,如信息、人员、技术、机构等进行管理,实现各种资源的合理配置,满足社会对信息需求的过程。随着技术的进步,信息管理向自动化、网络化、集约化方向发展。所谓信息管理的集约化,是指将信息系统从孤立、分散、局部地解决问题的方式变为系统、整体、全局性地解决问题的方式。近年来,知识管理逐渐成为主流,数字图书馆的管理也随之进入了知识管理阶段。

网络管理是数字图书馆管理的重要内容。网络管理是保证网络高效、可靠、经济和安全运行的重要基础,其主要目的是为了保障网络运作的品质,如维持网络传送频率,降低传送错误率,确保网络安全等。管理的内容大致包括:(1)系统管理。随时掌握网络内任何设备的增减与变动,管理所有网络设备的设置参数。当故障发生时,管理人员得以重设或改变网络设备的参数,维持网络的正常运作。(2)故障管理。为确保网络系统的高稳定性,在网络出现问题时,必须及时察觉问题的所在。它包含所有节点运作状态,故障记录的追踪与检查及平常可对各种通讯协议的测试。(3)效率管理。效率管理在于评估网络系统的运作,统计网络资源的运用及各种通讯协议的传输量等,更可提供未来网络提升或更新规划的依据。(4)安全管理。为防范不被授权的用户擅自使用网络资源,以及用户蓄意

破坏网络系统的安全,要随时做好安全措施,如合法的设备存取控制与加密等。(5) 计量管理。了解网络使用时间,能针对各个局部网络做使用量统计。

既可作为使用网络计费的依据,更可作为日后网络升级或更新规划的参考。(6) 信息管理。网络上的信息分成两部分,一是由管理员放置的,二是用户放置的。对这些信息要加以管理。传统的网络管理一般是集中式管理,但是随着网络规模的扩大,集中式管理已经不能适应了。新一代的网络管理要求具备4个方面的能力,即可控制性更强、复杂性减少、更快和更独立的问题解析能力以及信息的可访问性。于是出现了分布式管理等新的管理模式。

8. 通用化标准

没有规矩不成方圆。标准是对重复性事物和概念所做的统一规定。它以科学、技术和实验的综合成果为基础,经有关方面协商一致,由权威部门批准,以特定形式发布,作为共同遵守的准则和依据。因特网在短短的时间内能够发展到现在的规模,得力于因特网的开发者们创制了一套简单易行的标准规范并被人们所遵循。在数字图书馆建设的各个环节、各个方面都有一系列的标准与规范需要遵守。如果不遵守 TCP/IP 协议,局域网就联不上因特网;如果不采用 Z39.50 标准,数据库就不能被本系统以外的用户访问,成为网海中的信息孤岛;相应地也不能访问其他遵循 Z39.50 标准的数据库。遵循共同的标准规范是不同网络间实现互操作的必备条件,所以,在数字图书馆建设过程中,相关的标准规范是必须遵循的。

数字图书馆是开放的系统而不是封闭的系统,它所面对的是整个因特网用户。这就要求它采用标准时,要尽可能地采用国际标准,或国际通用规范,大多数人的习惯用法,总之是要