

数字图书馆 信息资源建设

孙长怡 著

辽宁大学出版社

数字图书馆

信息资源建设

孙长怡 著

辽宁大学出版社

©孙长怡 2009

图书在版编目 (CIP) 数据

数字图书馆信息资源建设/孙长怡著. —沈阳: 辽宁大学出版社, 2009. 3

ISBN 978-7-5610-5752-0

I. 数… II. 孙… III. 数字图书馆—图书情报工作
IV. G250.76

中国版本图书馆 CIP 数据核字 (2009) 第 036268 号

责任编辑: 马 静
责任校对: 何 丽

封面设计: 建 平

辽 宁 大 学 出 版 社

地址: 沈阳市皇姑区崇山中路 66 号 邮政编码: 110036

联系电话: 024-86864613 网址: <http://press.lnu.edu.cn>

电子邮件: lnupress@vip.163.com

沈阳市市政二公司印刷厂印刷 辽宁大学出版社发行

幅面尺寸: 145mm×210mm

印张: 9.25

字数: 230 千字

2009 年 3 月第 1 版

2009 年 3 月第 1 次印刷

书号: ISBN 978-7-5610-5752-0

定价: 20.00 元

前 言

图书馆作为保存人类知识文化遗产和传播知识信息的部门,其在整个社会的政治、经济、文化活动中具有十分重要的作用。图书馆发展水平的高低,在一定程度上反映了整个社会的发展水平和文明程度,这一点已经成为人们的共识。

图书馆的建设和发展经历了农业社会、工业社会和信息社会。从图书馆本身的发展看,当前正处在由传统图书馆向数字图书馆的转变和发展时期。图书馆的重要性的信息社会信息对人们的重要性的双重作用,决定了数字图书馆建设在整个图书馆发展过程中具有十分重大的意义。

图书馆社会价值的体现,主要是通过图书馆信息资源建设来实现的。因此,信息资源对图书馆而言,其作用是不言而喻和显而易见的。数字图书馆和数字信息资源的建设对图书馆人而言,完全是一项新课题、新挑战。对数字图书馆信息资源建设进行科学分析和研究,对保证数字图书馆建设始终沿着正确方向发展具有保驾护航的作用。

令人欣慰的是,当前数字图书馆建设不断掀起新的高潮。但是,数字图书馆建设毕竟是一个新的课题,各图书馆在数字图书馆建设中,特别是数字图书馆信息资源建设中,对于数字图书馆信息资源的建设理论、建设原则、建设内容、建设方法等缺乏统一的认识。由于这种认识上的偏差会导致数字图书馆建设成果的不尽如

人意,造成人力、财力、信息资源的浪费。因此,正确的理解和掌握数字图书馆信息资源建设的理论和方法就显得尤为重要。

本书正是从数字图书馆信息资源建设的理论出发,力求对数字图书馆信息资源建设的理论、方法、内容、原则等方面加以系统的阐述,结合国外的经验和国内的实践,试图对数字图书馆信息资源建设进行一些有创建性的指导。但是,由于作者的水平有限,书中的内容和观点难免存在一定的疏漏,欢迎各位同仁批评指正。如果本书对数字图书馆的信息资源建设起到一点积极的作用,那么作者将会感到极大的欣慰。

作 者

2009 年 3 月 5 日

目 录

第一章 数字图书馆的信息资源	(1)
第一节 图书馆信息资源的含义	(1)
第二节 图书馆信息资源的构成	(8)
第三节 信息资源的分类	(13)
第四节 数字图书馆信息资源分析	(18)
第二章 数字图书馆信息资源建设	(23)
第一节 数字图书馆信息资源建设内含	(23)
第二节 信息资源建设的理论与模式	(34)
第三节 数字图书馆与传统图书馆的关系	(37)
第三章 馆藏建设的原则与构成	(44)
第一节 馆藏建设的基本原则	(44)
第二节 馆藏信息资源的整合	(50)
第三节 馆藏信息资源的结构体系	(57)
第四章 数字图书馆藏书建设	(63)
第一节 电子信息资源的选购	(64)
第二节 特色数据库建设	(82)
第三节 网络信息资源的建设	(96)
第四节 数字馆藏的安全管理	(110)
第五章 实体图书馆藏书建设	(118)
第一节 实体信息资源的收藏	(118)

第二节	实体馆藏信息资源的采集途径	(127)
第三节	馆藏信息资源的有序化管理	(134)
第六章	数字馆藏资源的存储与组织	(143)
第一节	存储体的功能	(143)
第二节	电子书库构建与网络存储	(154)
第三节	信息资源组织管理	(169)
第四节	馆藏信息资源时效性控制	(182)
第七章	数字馆藏信息资源评价体系	(190)
第一节	网络环境下馆藏评价内容与沿革	(190)
第二节	数字馆藏的质量评价标准	(198)
第三节	网络信息资源评价的理论与实践	(202)
第四节	数字图书馆馆藏评价方法	(207)
第五节	网络信息资源安全管理	(216)
第八章	信息资源共建共享共管	(225)
第一节	信息资源共建共享共管的内涵	(225)
第二节	我国信息资源共建共享共管的理论与实践 ...	(234)
第三节	信息资源共建共享共管发展趋势	(260)
第九章	信息资源的知识产权保护	(270)
第一节	国内外数字化信息资源知识产权保护现状 ...	(271)
第二节	我国数字化信息资源知识产权保护的产业运行机制	(274)
第三节	数字化信息资源知识产权保护的法律法规体系	(277)
第四节	信息资源知识产权保护的技术保障	(282)
参考文献		(286)
后记		(290)

第一章 数字图书馆的信息资源

信息同能源、材料并列为当今世界三大资源。信息资源广泛存在于经济、社会各个领域和部门。是各种事物形态、内在规律、和其他事物联系等各种条件、关系的反映。随着社会的不断发展,信息资源对国家和民族的发展,对人们工作、生活至关重要,成为国民经济和社会发展的重大战略资源。它的开发和利用是整个信息化体系的核心内容。信息是一种重要的资源,这已成为人们的共识。对信息资源进行组织、管理、建设、开发成为一门新兴学科。信息资源建设既是图书馆学的一门分支学科,又是信息资源管理学科的重要分支领域。

在 2000 年 8 月 21 日召开的第十六届世界计算机大会上,江泽民同志第一次系统地阐述了新经济时代我国的国家战略。江泽民同志特别强调“信息资源已经成为与特质资源同等重要的资源,其重要作用正在与日俱增。”江泽民同志的讲话表明信息资源已进入国家最高决策者的视野,信息资源已被列入国家战略资源范畴,信息资源必将在我国今后的经济和社会发展中发挥更大作用。

第一节 图书馆信息资源的含义

信息资源是图书馆信息资源建设领域的基本概念之一,也是信息资源建设学科研究的逻辑起点。信息资源发展经历了三个阶

段:①传统管理阶段:20世纪50~70年代,以图书馆、情报所为代表的文字信息资源管理;②信息管理阶段:20世纪70年代末至20世纪末,以计算机应用和数据处理为典型代表;③信息资源管理阶段:21世纪初至未来20年,以网络平台、海量数据库、信息处理技术为代表,信息交换、信息共享、信息应用为内容,视信息资源为主要经济资源进行管理的信息资源管理。

然而时至今日,对什么是信息资源仍是众说纷纭,莫衷一是,国内国外都未形成一致看法。因此,对这一概念进行深入地研究,并给出一个较为科学的定义是十分必要的。

一、“信息资源”定义的综述

“信息资源”这一概念最早是由国外学者提出来的,时间大约在20世纪60年代末、70年代初。冯承柏、王崇德两位先生曾对国际著名的检索工具——《图书馆学和情报学文摘》(library and Information Science Abstract LISA)的光盘版(1970~1994)进行了检索,发现最早以“信息资源”为题的论文是奥罗尔科(J. O. Bourke)写的《加拿大的信息资源》(Information, Resources in Canada)。该文刊载于《专业图书馆》(Speciallibrary),1970年2月第61卷第2期第59~65页)。此后,以Information Resources为标题的论著逐渐增多,从1970~1994年共有48篇,其中不乏对信息资源的描述、解释和定义。

我国对信息资源的概念及相关问题的研究始于20世纪80年代。1985年孟广均先生著文指出:“我国的信息资源很多,经济的、科学的、技术的、政治的、文化的、教育的、军事的……现在国外普遍认为,没有控制、没有组织的信息不成为一种资源,因此都加强了对信息的管理。”此后,一些学者开始对信息资源进行定义和研究。1996年,乌家培先生撰文指出:“对信息资源有两种理解,一种是狭义的理解,即指信息内容本身。另一种是广义的理解,指

的是除信息内容本身外,还包括与其紧密相联的信息设备、信息人员、信息系统、信息网络等。”这一定义在国内比较流行,但一些学者也提出了质疑。

我们将国内外关于信息资源的定义归并整理,去同存异,则关于信息资源的解释大体可概括为七种:

(一)信息资源 = 信息源

这种观点将“信息资源”与“信息源”这两个根本不同的概念混为一谈,显然是不正确的。信息源是指信息的出处或来源。信息的来源多种多样,可来自个人、单位、机构、团体,也可来源于文献、实物,还可来源于市场、会议等场所。信息源中有些可称为信息资源(如文献),但大多数不能称其为信息资源。按照信息论的观点,万物皆信源,若将信息资源等同于信息源,岂不要得出“万物皆为信息资源”的荒谬结论吗?

(二)信息资源 = 文献资源

这种观点缩小了信息资源的外延。因为文献资源只是信息资源中的一部分,而非全部的信息资源。

(三)信息资源 = 数据

数据有广义和狭义之分。广义的数据多用于计算机和通讯领域,指形成信息的原材料;狭义的数据则是与文字、图形等相对的一种高度浓缩的信息。将信息资源等同于广义的数据显然扩大了其外延;而将信息资源等同于狭义的数据则又缩小了其外延。

(四)信息资源 = 信息

信息有有效、无效、负效之分。作为资源必须有用,有用性是资源的基本属性。从这个意义上说,那些无效的甚至有害的信息是不能作为信息资源的,将信息资源等同于全部信息,同样扩大了信息资源的外延。

(五)信息资源 = 支持工具(人员、技术、设施等)

信息资源的核心是信息,信息人员是收集、开发、传播信息的

主体,信息技术和信息设施则是获取、传播、开发、利用信息的方法、手段、工具、条件。舍弃信息而将信息支持工具作为信息资源无疑是一种本末倒置。

(六)信息资源 = 信息 + 支持工具

将支持工具纳入信息资源的范畴,有悖于常理,显得牵强附会。因为人们在说信息资源时是不将其开发人员、开发条件包括在内的。例如,一个国家的石油资源,往往是指该国的石油储量,似乎并不包括其石油工人、开采技术和开采设施。可见,将支持工具包括在信息资源的范畴中,同样扩大了信息资源的外延。

(七)信息资源 = 部分信息的集合

这种观点比较有道理,但关键在于是哪一部分信息的集合。国外学者列维(K. B. Levitan)认为“是一系列已经制度化了的、为一个或多个用户集团反复使用的信息”;而国内学者霍国庆则认为“是可以利用的信息的集合”。应该说,这些表述已接近信息资源的本质,但还需要进一步地推敲和修正。

从以上的介绍、分析,可以得出以下几点结论:

1. 关于“信息资源”的概念,无论是国内还是国外都还未形成统一的认识,说明“信息资源”还是一个正在发展着的概念。
2. 由于“信息资源”的概念来自于外国,因而我国早期对信息资源概念的认识受国外、特别是美国的影响较大。现今比较流行的所谓广义、狭义之说,实质上是对国外观点的综合概括。
3. 国内关于信息资源概念的认识正在逐渐深化,总的趋势是从广义、狭义之说向狭义方向突破,把握这一点对于准确定义信息资源尤为重要。
4. 美国学者关于信息资源的认识,在美国同行看来也许十分自然,但在我们看来却有着“泛化”倾向。究其原因,在于中美对“信息”和“资源”的理解有着较大的差异。如果盲目照搬,很可能造成理论体系上的一系列漏洞。

信息资源是由“信息”和“资源”两个概念复合而成的,要定义信息资源,有必要先厘清信息和资源的含义,特别要弄清中外在这两个概念理解上的差异。

二、信息与资源

(一) 信息

自 20 世纪 40 年代申农和维纳分别提出信息论和控制论以来,人们对信息的认识可谓众说纷纭,不同领域的专家学者从不同的学科角度作出了不同的解释。哲学家认为信息是认识论的一部分;物理学家认为信息是熵的理论;通讯工作者把它看成是不定度的描述;数学家认为它是概率论的发展;情报学家则认为信息是生物及自控系统与外界交换的一切内容。据 1975 年统计,公开发表有关信息的定义达 39 种,到 1986 年上半年已达 85 种,估计目前已逾百种,说明信息是一个多维和多层次含义的概念。

也许正是由于上述原因,近年来我国信息学者多从哲学的角度来理解信息,认为信息是事物运动的状态与方式。信息不是事物本身,但它反映了事物的特征与特性。不同的事物有不同的特征,并在不同的条件下发生变化,这种特征与变化就是信息。人们正是通过获取和识别这些信息来认识不同事物的。

与我国不同,美国学者则更多地从管理学和计算机应用相结合的角度来研究信息资源及其问题。

根据孟广均等人的研究,美国信息资源管理领域对信息的理解是高度统一的,他们所理解的“信息”类似于我们所理解的“信息资源”,认为信息是有意义的数据,是经过收集、记录、处理和存储的可供检索的事实与数据。

(二) 资源

“资源”一词,根据《现代汉语词典》解释,是“生产资料或生活资料的天然来源”;《辞海》解释为:“资财的来源。一般指天然资

源”。可见,中文里的“资源”主要指物质资源或自然资源,且主要指其来源。俄文里的“资源”一词为 *ресурсы*,亦指财富的来源。英文中的“资源”一词为 *re. source*,它由 *re* 和 *source* 组成,前缀 *re* 有“再”的意思,*source* 表示“来源”。可见,资源在各民族语言里词义相通,只是英文中的“资源”在外延上要宽泛一些。据冯承柏、王崇德考证,*resour* 既可指物理世界的资源,又可指信息的、智力的和精神方面的东西,还可兼指二者;*resource* 的初始含义是资源,后来又演化出手段、方法、谋略、能力之意。美国的一些同行可能是从 *resource* 的语义出发去认识信息资源的,所以才得出“信息资源 = 信息 + 支持工具”的结论。

可见,如果我们顾名思义,简单套用国内外对“资源”的语义解释,就会得出“信息资源就是信息的来源”,或“信息与信息支持工具之集合”的推论。这显然是不正确的。

学术界是如何理解和使用“资源”一词的呢?对资源有独到研究的资源经济学是这样认为的:一是自然资源是自然物质、自然能量或自然条件中的一部分,而不是其全部。二是自然资源是指对人类有用的或能被人类利用的自然物质、自然能量和自然条件,它可作为生产的投入,产生经济效益或能满足人类的生活需要。三是自然资源还必须具有一定的丰度和凝聚度,是被人类利用的物质和能量的总和。四是自然资源的获得离不开人类的参与,自然资源转化为经济资源的先决条件是经过人类的劳动投入和改造。

资源经济学界对自然资源的上述认识,应当成为我们定义“信息资源”的“他山之石”。

三、“信息资源”的定义

根据前面我们对“信息资源”诸定义的评述以及对“信息”、“资源”的解析,可以得出下述几点结论:

(一) 信息资源是信息的集合

俗话“独木难成林”。一棵树构不成森林资源,一滴水构不成水资源,一滴原油也构不成石油资源。同样,一条信息或几条信息也构不成信息资源。只有当信息达到一定的丰度和凝聚度时,才能成为信息资源。从这个意义上说,信息资源应是多种多样信息的总和或集合。

(二) 信息资源是经过人类选择、获取的有用信息的集合

信息资源并不是所有信息的简单相加和堆砌,而是经过人类选择的、对人类有用或能满足人类需求的那部分信息的总和或集合。有用性是一切资源的本质属性,信息资源也不能例外。从信息海洋中挑选出有用信息,并将之与无用信息区分开来,正是信息管理人员的基本任务之一。

(三) 信息资源是经过人类组织、序化的信息的集合

与非信息资源相比,信息资源最显著的特征就是有序性。对水资源、石油资源、矿产资源等自然资源来说,无所谓有序、无序,只要具备一定的丰度和凝聚度,值得人们开采、获取即可。信息资源却不然,无序的信息不仅无法利用,还会造成信息通道的“栓塞”,阻碍信息的传播、交流、开发和利用。因此,组织、序化的信息才能成为信息资源,而没有控制的、未经组织的信息将不再成为一种资源。

信息资源是人类在生产实践,科学实验和社会生活中创造和累积的可利用的信息内容的总和。其中,“可利用性”是资源的本质属性,因为并非人类所创造的全部信息内容都具有可利用性,如虚假信息,错误信息,片断信息等就不具备可利用价值。根据“可利用性”进行分析,自人类产生以来所累积的可利用的信息内容总和属于信息资源存量,不同国家、地区,组织或人类个体所拥有的信息资源存量是他们进一步创造信息资源的重要条件之一。特定国家、地区、组织或人类个体在特定时期内所创造的可利用的信

息内容总和则属于信息资源增量,信息资源增量更多地取决于特定国家、地区、组织或人体的信息资源生产能力;需要指出的是,特定国家、地区、组织或人类个体在特定时期内实际的信息资源增量包括他们自己所创造的信息资源,也包括他们所获得的其他国家、地区、组织或人类个体所创造的信息资源。所以信息资源增量在一定意义上也代表他们信息资源的消费能力。

综上所述,笔者将信息资源定义如下:信息资源是经过人类系统收集加工、有序化组织的有用信息的集合体。

第二节 图书馆信息资源的构成

一、信息资源的构成

关于信息资源构成的研究,国内外还没有予以高度重视,专题性论著比较鲜见。我们只能从散见于一些论著中的片言只语、篇章段落,乃至信息资源概念的描述来窥探人们的信息资源构成观念,理出这一问题的研究脉络。

信息资源是一个相当广泛的概念,对于它的构成可以有各种不同的角度。从开发和利用工作的需要来看,主要可以从主体和载体两个方面去考虑。

一是从主体角度。所谓主体角度,就是从拥有者的角度对信息资源进行区分。按此角度全社会的信息资源大致可以分为政府信息、企业信息、个人信息等。在目前的情况下,我国正处在计划经济向市场经济过渡的过程之中,各级政府部门仍然是信息资源的最大拥有者,掌握着大量的信息资源。无论从数量来说,还是从重要程度来说,政府信息都是整个信息资源的重要组成部分,其中包括经济、科技、教育、文化、卫生等领域的信息资源。政府信息资源对于政府更好地履行经济调节、市场监管、社会管理和公共服务

职能具有重要的意义,因而是当前信息资源开发利用的重点。

二是从载体的角度。所谓载体角度,是从载有信息的介质对信息资源进行区分。按照这一角度,信息资源可以大致分为传统载体的信息资源和电子载体的信息资源。前者主要是指以纸张和实物为载体的信息资源。由于我国具有广袤的国土、众多的人口、悠久的历史、灿烂的文化,在政治、经济、文化、军事、科研、教育、卫生等领域积累的传统信息资源十分丰富。后者则包括以广播、电视、电视会议、图文电视、各种不同类型的数据库以及网络等为载体的各种信息资源。这些信息资源是在计算机技术、通信技术和高密度存贮技术迅速发展,并在各个领域里得到广泛应用的背景下形成的,产生的时间还不长,但是发展迅速。目前,电子技术在信息的存储和传播信息的一统天下,代表了信息资源管理的长远发展的方向。

目前,国内外对信息资源构成的认识不尽一致,但综合起来有两种比较典型的观点:一种是狭义的理解,认为信息资源是指人类社会经济活动中经过加工处理有序化并大量积累后的有用信息的集合,如科学技术信息、政策法规信息、社会发展信息、市场信息、金融信息等,都是信息资源的构成要素。另一种观点是广义的理解,认为信息资源是人类社会信息活动中积累起来的信息、信息生产者、信息技术等信息活动要素的集合。也就是说,信息资源包括下述几个部分:①人类社会经济活动中经过加工处理有序化并大量积累后的有用信息的集合;②为某种目的而生产有用信息的信息生产者的集合;③加工、处理和传递有用信息的信息技术的集合;④其他信息活动要素(如信息设备、信息活动经费等)的集合。这两种观点争论的焦点在于信息资源的构成是窄化(即纯“信息”化)还是泛化,在于除信息之外的信息活动要素是否为信息资源的组成部分。其实质是信息资源概念的广狭之争在信息资源构成问题上的反映。

(一) 信息资源的构成要与信息资源的定义相一致

目前,国内许多论著在信息资源的认定上持广义之说,而在阐释信息资源的构成时却又仅限于狭义的信息资源。

(二) 应将信息资源的构成要素与其开发利用条件区分开来

信息资源之所以重要,主要是因为其中所蕴含的有用信息具有十分重要的经济功能,如能帮助人们进行选择决策,消除经济活动中的不确定性,减少经济活动中其他资源的损耗,降低成本、节省开支等。而信息生产者、信息技术、信息设施、经费等信息活动要素只不过是信息资源开发利用的必备条件。没有信息要素的存在,其他信息活动要素都没有存在的意义。当然,信息要素价值的实现离不开信息生产者、信息人员、技术、设施、经费等信息活动要素的综合作用,但我们不能因此就将之看做是信息资源的构成部分。正如我们不能因为开发利用海洋资源需要专门的人员、技术、设施、资金等就将这些也作为海洋资源的构成要素一样。

(三) 应将信息资源的构成与信息资源的类型区别开来

构成是指事物的组成单元,类型则是“按事物的共同性质、特点而形成的类别”(《辞海》缩印本 1979 年版第 1931 页)。如汽车是由发动机、底盘、车身三部分组成的,这是其构成,而其类型则可依用途、性能、颜色等多种标准做不同的区分。但不管哪种类型的汽车,其构成都是一样的,都是由发动机、底盘、车身组成的。从这个意义上说,信息资源的构成就是从千差万别的各种信息资源中抽象出来的,所有信息资源都具备的基本“构件”。

根据信息资源的定义以及上面的讨论,我们认为信息资源从总体上说就是由同质的单一体——信息构成的。当然,这里的信息是有条件的,即是由人类选取、组织、序化的有用的那部分信息。除此之外,别无他物。