

王心裁 吕元智

当前数字图书馆信息组织中存在的问题^{*}

摘要 数字图书馆信息组织的对象是数字化信息资源。它包括数字化特种馆藏、商用光盘数据库和联机数据库、网络信息资源等。数字图书馆信息组织工作中存在的问题有:具体操作中的问题(组织对象难以选择,具体信息特征难以揭示等);信息组织的要求问题;信息组织的方式方法以及信息组织的标准化与协调问题等。参考文献 13。

关键词 数字图书馆 信息组织 问题

分类号 G250.76

ABSTRACT The objects of the information organization of digital library are digital information resources, including digitized special collections, commercial CD-ROM and online databases, network information resources, etc. In the information organization of digital library, there are problems of practical operations (it is hard to select objects to be organized and to describe specific information characteristics), requirements, patterns and methods, and standardization and coordination. 13 refs.

KEY WORDS Digital library. Information organization. Problem.

CLASS NUMBER G250.76

1999年11月12日,国家863项目组与首都图书馆数字图书馆示范工程签约标志着我国数字图书馆建设正式进入实施阶段^[1]。两年多来,我国数字图书馆的建设取得了许多可喜的成绩,也存在着一些不容忽视的问题。为了加快我国数字图书馆的建设进程,本文对数字图书馆的信息组织问题进行探讨,以供数字图书馆建设参考。

1 数字图书馆信息组织对象引发的问题

数字图书馆信息组织的对象是数字化信息资源,它包括数字化特种馆藏、商用光盘数据库和联机数据库、网络信息资源等^[2]。与传统信息资源相比,数字化信息资源在范围、价值、存在形式、更新方式等方面有很大不同,从而给数字图书馆的信息组织带来了一系列问题。

1.1 信息范围与选择

目前,各种图书馆联机馆藏、电子出版物(网络报纸、期刊、图书等)、商用信息数据库等大量涌现,网络信息资源更是以前所未有的速度增长。据统计,早在1997年2月,Lycos的WWW网页已超过34亿个。因特网每天发布14万件新信息(总量约450MB),全网提供的信息总量逾20TB^[3]。数字图

书馆的内容主要包括馆藏资源数字化和社会资源馆藏化两个方面,它不可能涵盖全部社会信息资源,只可能对部分社会信息资源进行组织^[4]。这就使数字图书馆的信息组织面临一个信息范围选择问题。哪些信息是数字图书馆该组织的?哪些信息是数字图书馆可以获取的?我们要根据实际需要,综合考虑多方面因素,如行业需求、用户知识结构层次等,认真选择和确定信息范围,选择最有效、最有使用价值的信息,并对其进行有效组织。否则数字图书馆就会淹没在信息的汪洋中,变得无所作为。

1.2 信息价值与鉴别

数字图书馆的首要目标是为社会公众提供有价值的信息。但信息的价值是很难区分的。有些信息对有些人或工作没有价值,但对另外一些人或工作就有价值,有些信息现在没有价值,但在将来可能有价值。日益复杂的社会信息需求和信息价值的难以确定,给数字图书馆的信息组织提出了难题。它要求数字图书馆的工作人员对数字化信息不仅要有现实价值的鉴别能力,还要有未来价值的预测能力。

另外,数字图书馆的信息组织对象——数字化信息是易变的。它与载体相分离,具有流动性、易删改性等特点。数字化信息的这种特性,给数字化信

^{*}本文为武汉大学学术骨干项目“数字图书馆信息资源的组织与开发研究”论文。

息尤其是网络信息的获取和迁移提供了便利,但也给信息的修改、复制提供了方便和条件,给数字图书馆的信息组织与安全带来不利,使组织的信息不准确或失真。这就要求信息组织人员有较高的信息鉴别能力,尤其是要熟悉元数据的有关知识,通过元数据的内容,了解和控制数字化信息的来源、修改等情况,确保信息的可靠性和权威性。

1.3 信息更新

由于现代数字化信息的易变性和开放性,数字图书馆信息组织对象大部分都处在不断地更新变化过程中。就拿电子出版物来说,一些电子出版物,每隔一段时间,就要更换新的内容,往往是新出版的内容覆盖了以前出版的内容。而网络信息资源,则存在着网址更改及网页更新等问题。这样很容易使数字图书馆的信息组织出现重复,导致数据库中出现重复、冗余的数字记录,由于网址更改、网页更新而出现找不到所组织的资源的情况。数字图书馆在信息组织时要采取行之有效措施,控制与处理冗余信息及被更新信息。

2 数字图书馆信息组织工作的具体问题

2.1 信息组织具体操作中的问题

2.1.1 具体组织对象难以选择

数字化信息与传统的文献型信息不同。传统的印刷型文献可以以一个整体(如本、卷、册等)为单位进行信息组织,其物质形态决定了文献组成部分之间的关系(如页数、章节数等)。而数字化信息,由于可压缩和可被高密度存储,往往许多不同主题的信息同时存储在同一载体上,如同一光盘、同一硬盘等。网络信息的情况更为复杂,它们往往是以一系列相关信息同时汇集于某一相同主题或主页下,或是以构成某一个大型数据库中的一部分等方式存在,且人们通常不知这些信息储存在什么载体上。数字化信息的这些特征虽然给数字图书馆节省空间和保管费用提供了便利,但也给信息组织带来了困难。数字化信息组织是以整体(光盘、硬盘、某网页等)为对象,还是以具体的某个数字化信息为对象?如是以一张光盘为对象,还是存储在光盘中的一本图书为对象^[5]?如此等等,是数字图书馆信息组织难以处理的问题。

2.1.2 具体信息特征难以揭示

印刷型文献由于有比较通用和成熟的标准与规范,人们通常选择版权页或题名页作为文献著录的基本依据。但是数字化信息由于生产者众多,目前还没有统一的规范和标准,因此作为著录依据的信

息特征不全,从而难以准确揭示。1992年OCLC的编目试验表明,在因特网上的300个独立文件的书目记录中题名出现频率最高,为93%,出版数据出现频率为73%,作者的出现频率仅为18%^[6]。对数字化信息生产进行规范是当前数字图书馆建设的关键问题之一,也是数字信息组织的基础问题。

此外,数字化信息是按特定格式存储的,使用不同存储格式,必然导致信息的格式和篇幅不同,也会给著录带来麻烦。

2.2 信息组织的要求问题

(1)与利用印刷型文献信息不同,利用数字化文献或信息需要借助相应的软硬件设备,并且不同格式的数字化文献对所依存的硬件和软件都有不同的要求。数字化信息的地址必须绝对精确,因为任何微小的错误都会造成地址的错误和无效。这就要求信息组织工作更加细致精确,数字化信息组织不仅要准确揭示信息的各项基本特征,如题名、著者、来源等,而且还要注明该信息利用时所需要的软硬件环境。

(2)随着社会的飞速发展,技术的淘汰也越来越快。借助原来技术组织的信息是否能在新技术环境中发挥作用是一个不容忽视的问题。

(3)数字图书馆的信息组织,其实是一项主观作用于客观的活动。在信息组织过程中,工作人员会自觉或不自觉地带有一定的主观倾向。这种主观倾向与信息组织的客观性要求是矛盾的。如何尽量客观地组织信息?信息组织人员应不断提高信息素质,尽量客观地组织信息。

2.3 信息组织的方式方法问题

信息组织的方法很多,主要有语法信息组织法(字顺组织法、代码组织法、地序组织法、时序组织法等)、语义信息组织法(元素结构组织法、逻辑组织法、分类组织法、主题组织法等)和语用信息组织法(权值组织法、概率组织法、个性组织法等)等^[7]。信息组织方式有文件方式、数据库方式和超媒体方式等多种方式^[8]。

在所有的组织方法中,分类法与主题法是两种最根本的方法。分类法适用于族性检索,主题法适用于特性检索。但组织网络资源,仅用分类方法或主题方法难以有好的效果。分类主题相结合的方法是今后的发展方向。分类法可以给予信息资源一个科学的分类系统,而主题法则使资源中所包含的信息能得到充分揭示。但目前还没有规范统一的组织数字信息资源的分类主题相结合的方法。

2.4 信息组织的标准化与协调问题

数字图书馆是一个庞大的信息系统,它具有分布的、大规模的和有组织的数据库和知识库,用户或用户团体可对系统内的数据库和知识库进行一致性访问,获取自己所需的最终信息^[9]。为了便于用户获取有效信息,进行信息共享,数字图书馆在进行信息组织时,就要有一个统一的标准,以便使组织出来的信息在各数据库、各网络平台之间自由流动。标准化是数字图书馆信息组织的关键。但从目前全国各高校图书馆的数字化建设来看,其情况不容乐观,各图书馆信息的组织仍在采用自己原用的标准和方法。这对数字图书馆的建设和发展极为不利。值得庆幸的是,我国正在着手数字化信息组织的标准化研究。中文资源的元数据标准正在制订。广东省中山图书馆已制订了《数字式中文全文文献通用格式》。北京大学数字图书馆研究所、北京大学图书馆数字图书馆工程也正在进行“中文 Metadata 标准研究”等研究。这些必然为数字图书馆的信息标准化建设奠定基础。

数字图书馆是一个浩大的工程,它的建设需要有一个部门对其进行规划、协调。但从总体上看,我国数字图书馆的建设仍处在探索阶段,相关的协调机构和措施还没有出台,数字图书馆的信息建设总体上存在缺乏整体性、协作性、重复建设等问题^[10]。

3 数字图书馆信息组织的其他相关问题

3.1 社会信息意识问题

随着知识经济的发展,人们逐步注意到信息和知识在知识经济中的地位和作用。但是由于我国国民经济发展不平衡,在许多地区行业,信息的作用和地位并没有得到重视或完全重视,信息和知识作为第三资源的观念还没有深入“民心”。社会信息意识淡薄是阻碍我国数字图书馆发展的重要障碍,也是影响数字图书馆信息建设与发展的一个关键因素。提高社会信息意识是顺利进行数字图书馆信息资源建设的关键和保障。

3.2 技术问题

在数字图书馆建设和发展过程中,有许多技术问题需要解决,如信息的存储、检索、自动分类、自动标引、信息安全、信息标准化处理、数据采掘(Data Mining)、智能 Agent 等技术^[11]。这些技术要求数字图书馆的信息组织人员具有较高的综合素质,他们不仅要熟练掌握基本的计算机操作技能,而且要具

备较高的信息处理能力。

3.3 人才问题

人才是事业成功之本。数字图书馆的信息组织,不仅仅是一个信息组织问题,还涉及许多如管理、技术等方面的问题,这些问题的解决都需要人才。数字图书馆建设也可以说是一个人才工程问题。但是我国数字图书馆建设所需的人才不足,其中一流人才严重匮乏,并且现有的一流人才也因种种原因而不断流失^[12]。

3.4 国家信息基础环境问题

从我国现状来看,信息基础环境不容乐观。整个信息资源布局也不合理,在沿海一些发达地区和城市,信息布局比较好,但在不发达地区,尤其是在广大农村,信息资源相对短缺,存在着严重的信息贫富不均的现象和“数字鸿沟”(Digital Divide)问题^[13]。

3.5 资金问题

在数字图书馆的探索与建设过程中需要花费大量资金。我国是一个发展中国家,数字图书馆的信息资源建设难免遇到资金短缺问题。资金短缺是制约我国数字图书馆发展的一个最基本问题。

3.6 知识产权问题

数字图书馆的信息组织,尤其是对网络信息的组织,会涉及到许多法律问题,其中知识产权问题是一个最为紧要的问题。如哪些是公开的信息?哪些是受知识产权保护的信息?哪些信息可以自由获取?哪些信息需要许可才能获取?哪些信息是共享的?等等。国际上已有一系列相关的公约来解决这些问题,也制定了一些相关政策和法规,如《中华人民共和国著作权法》、《计算机软件保护条例》等,但并不完善,仍存在一些漏洞。数字图书馆的信息组织人员,不仅要熟悉数字化信息尤其是网络信息的知识产权交易(如转让、许可等),更要熟悉相关的知识产权法,否则,会引来产权纠纷。

参考文献

- 1 王纯. 中国数字图书馆建设、数字化信息资源开发及网络建设的现状. 中国图书馆学报, 2000(4)
- 2 王胜利. 21世纪图书馆的研究热点——数字图书馆. 情报杂志, 2002(2)
- 3 倪莉. 网络环境下信息组织面临的新问题. 情报理论与实践, 2001(4)
- 4 林海青. 数字图书馆信息组织. 中国图书馆学报, 2001(1)
- 5.6 周宁. 信息组织. 武汉:武汉大学出版社, 2001

贾君枝

市场环境中网络信息资源配置的影响因素*

摘要 网站信息资源的组织方式、网络用户需求、网络服务机构的能力是影响网络信息配置的主要因素。增强信息化意识、缩小信息化差距、提高信息化水平,是优化网络信息资源配置的关键。参考文献 12。

关键词 网络信息资源 资源配置 影响因素 信息市场 用户需求

分类号 G250.73

ABSTRACT The author thinks that the organization patterns of website information resources, the needs of network users and the capability of network service providers are major factors to impact network information allocation. To optimize network information resource allocation, we should promote the awareness of informatization, eliminate informatization gaps and bring the informatization to a higher level. 12 refs.

KEY WORDS Network information resource. Resource allocation. Impact factor. Information market. User needs.

CLASS NUMBER G250.73

网络信息资源配置的目的在于保证网络资源的投资主体能够在资源配置中以等量投入或尽可能少的投入获得最大化的产出,把各种信息资源有效地分配于各种用途,使有限的资源生产出更多符合社会需要的服务和产品。本文主要是对影响网络信息资源配置的因素进行分析研究,以利于提高信息资源配置水平。

1 影响网络资源的配置因素

市场条件下,网络资源的提供方及需求方是市场的主体及客体。现有网络资源的提供者主要表现为各种类型的网站,他们根据自身的市场目标定位来生产、组织网络资源,其目的是最大程度地实现经济效

益最大化或者社会效益最大化。网络资源的需求方主要为网络用户,他们旨在以最小成本或零成本方式获得自己所需的资源。其网络资源的需求状况及使用的方式很大程度上影响着网站的资源组织方式及其资源配置情况。因此,提供方和需求方二者之间的相互作用构成了网络资源总的配置格局。但网站作为经营主体而言,其网络资源的大部分来源于网络信息服务机构的产品,在一定程度上它的网络资源分布状况取决于网络信息服务机构的生产能力及发展潜力,因此我们需要将这方面的因素考虑进去。下面主要就这三方面的影响进行详细阐述。

1.1 网站信息资源的组织方式是影响资源配置的首要因素

- 7 霍国庆. 论信息组织. 情报资料工作, 1997(6)
- 8 张俊. 略论网络信息资源的组织. 图书情报知识, 1998(2)
- 9 李慧. 数据库技术与数字图书馆中的信息组织. 现代图书情报技术, 2001(5)
- 10 王文静, 毛晓蓉. 我国图书馆数字化建设有关问题的思考. 情报杂志, 2001(1)
- 11 张天俊, 张银犬等. 知识管理与数字图书馆资源的优化配置. 情报学报, 2001(12)

- 12 牛雅杰, 陈桂珍. 论图书馆数字化面临的十大难题. 情报杂志, 2002(3)
- 13 孙国锋, 苏竣. 电子政府——政府的深刻变革. 科学与科学技术管理, 2001(9)

王心裁 武汉大学信息资源研究中心副教授, 博士.
通讯地址: 武昌珞珈山. 邮编 430072。

吕元智 武汉大学信息管理学院研究生. 通讯地址同上。
(来稿时间: 2002-07-23)

*本文系教育部重大项目“网络信息资源组织与开发模式”(2002ZDXM870001)的研究成果之一。