

●李秀云

网络环境下的知识组织研究^{*}

摘要 网络环境下的知识组织,是指对网上信息进行分析,明确其知识本质,把知识客体中的知识关联予以揭示,形成便于用户理解、查询的检索系统。它应遵守标准化、系统性、特色性、安全性等原则,有主题树、文件、数据库、知识重组、智能重组等方式。参考文献14。

关键词 网络环境 知识组织 理论方法 信息服务

分类号 G250.73

ABSTRACT Knowledge organization in the network environment is to analyze network information, clarify its knowledge nature, reveal knowledge correlation of knowledge objects and form easily understandable and searchable retrieval system. In this paper, the author proposes some principles and methods. 14 refs.

KEY WORDS Network environment. Knowledge organization. Theoretical methods. Information services

CLASS NUMBER G250.73

在网络环境下,人们的思维方式和工作方式也向知识的传播、运用和创新发生了转变,知识已成为真正的资本与首要的财富。

1 知识和知识组织

1.1 知识的涵义

知识是一个古老的哲学命题。古今中外,许多学者为此而付出了大量心血,迄今未获得统一的认识。《现代汉语词典》是这样解释的:“知识是人们在改造世界的实践中所获得的认识和经验的总和。”《辞海》则把知识定义为:“人们在实践中积累起来的经验,从本质上说,知识属认识范畴。”二者都把知识作为实践经验或实践活动的认知成果来看待。这是正确的。然而知识的定义还应该揭示认知活动中主体与客体的动态关系。根据认知的信息加工理论,人们的认知活动是在信息传输水平上发生着的主体与客体双向投入的过程,也是主体以本身的认知图式、观念体系和意会能力掌握客体的属性、结构、关系、本质规定的过程。这一过程淀积的成果就是知识。

刘植惠教授将知识定义为:“认知主体以其认知图式适应、同化被认知客体的信息内容,经整合重构而再现的观念化、符号化的有序信息组合。”^[1]

从上述关于知识的定义可见,知识是人的主观

世界对客观世界反映与认识的结晶。知识有两种存在形态,一是以人脑为主体的主观知识,二是记录于各种文献载体之上的客观知识。

知识产生是一个比较复杂的动态过程。它体现着主体与客体的双向投入,既包括适应、同化、意会,又包括编码、转译、重组等操作。对知识进行组织,其前提是首先要了解和掌握什么是知识及知识的结构形态。

1.2 知识组织理念

“知识组织”(knowledge organization)这一概念,并不是最近提出来的。早在1929年英国著名分类法专家布利斯(H. E. Bliss)就曾使用过这个概念^[2]。蒋永福认为:“知识组织是指对事物的本质及事物间的关系,进行揭示的有序结构,即知识的序化。通俗地说,知识组织是对知识客体所进行的诸如整理、加工、揭示、控制等一系列组织化过程及其方法。”^[3]王知津的文章把知识组织描述为:“知识组织是对知识进行整序和提供,既处理大量的现有知识,又能相对降低存贮知识的物理载体文献的盲目增长以免知识过于分散化。所以提供文献、评价科学文献和系统表述以生成新的便于利用和获取的有序化知识单元的处理系统即是知识组织。”^[4]田书格的文章指出:“知识组织是关于知识的组织与检索系统,是现代网络信息环境下获取知识与利用知识的所有手段、技术

^{*}本文系河南省哲学社会科学“十五”规划项目(2001FZH005)的研究成果。

与能力的总和。^[5]总之,知识组织就是关于知识组织的理论与方法。从知识的结构特点看,知识因子是知识组织的基本单位。一个概念、一个词语、一件事物都可以成为知识因子,知识关联是若干个知识因子间建立起来的特定联系。知识组织的基本原理是以固定的方式方法把知识客体中的知识因子和知识关联揭示出来,并把被揭示的知识按一定的方式编排序,以便于人们利用。本文所讨论的知识组织是对网络环境下客观知识的组织。

在网络环境下,浩如烟海的网络信息,比纸张原始信息更随意、更庞杂、更离散,因此,对网上信息资源进行知识组织已成为急待解决的课题。网络环境下的知识组织,是指对网上的各种有关概念、数据、事实、文献等,通过分析、明确信息中包含的知识本质,把知识客体中的知识关联进行揭示、整序,形成便于用户理解、查询的检索系统。

2 网络环境下知识组织原则

2.1 标准化原则

标准化是实现知识组织管理科学化和网络化的必要技术条件。如果各自为政,工作中的标准和格式不一致,开发的数据库就无法在网上实现共享。因此,图书馆的知识组织必须按标准化原则进行,解决数据的标准化问题。

2.2 系统性原则

网络环境下图书馆的信息资源由现实馆藏和虚拟馆藏两部分组成。现实馆藏是指本馆所拥有的数字化信息资源,虚拟馆藏是指本馆读者借助于计算机系统、通信网络等设备可以广泛使用(如检索、传输等)的网上电子信息资源的总和。只有系统地、连续地从馆内与馆外、国内与国外收集和组织的各种信息资源,才能不断充实和发展图书馆的现实馆藏和虚拟馆藏,发挥信息资源的效用。尤其是对于本馆特色资源的采集,更应力求保持系统、连贯和完整,系统性是知识组织的重要条件,时断时续、支离破碎、不成系统会大大降低知识的实用价值。

2.3 特色性原则

图书馆的知识组织应从网络整体出发进行资源的合理配置,要把数字图书馆建设纳入到整个地区、国家和全球信息网络中去加强各馆富有特色的数字化文献信息资源的开发,建立起各具特色的数字图书馆。只有能反映并提供馆藏特色资源的数字图书馆才能拥有众多的用户和求得自身的长远发展,并

显示其自身的优势。

2.4 安全性原则

由于目前计算机病毒、计算机黑客、软件炸弹与信息威慑、信息垃圾为害相当严重,给数字系统的安全性提出了挑战,因此,在网络环境下,知识组织必须采取一些必要的安全保障措施来保证图书馆网络与信息资源的正常运行。

3 网络环境下知识组织的方式方法

在网络环境下,知识文献这个概念的外延不断扩大,内涵不断加深。随着网上知识信息的激增及人类知识交流模式的改变,研究者们对知识组织的概念理解不一,从而对知识组织研究形成了各种各样的方式方法。

3.1 众多的研究者开始探讨传统分类法、主题法在网络环境下的转变

(1)传统分类法的知识系统性和标识语言的通用性以及族性检索能力和扩/缩功能,是其他情报检索语言所不具备的,也是无法取代的。因此,它在新信息环境下也必然会有一席之地^[6]。

(2)随着信息网络的发展,占据主导地位的分类法不可避免地会受到挑战,但与此同时也面临着新的发展机遇。分类法在信息网络中的应用主要表现在用于联机浏览检索、非文本信息的组织、超文本系统的管理,以及作为网络信息组织的通用工具并促进分类主题一体化^[7]。

(3)使用分类法组织网络分类目录并提供浏览方式进行查询,具有直观、信息质量高等优点^[8]。

(4)用主题词组织与揭示信息具有直接和直观的特点,在网络环境下有相当一部分网络信息资源浏览器与引擎都以主题词为组织与揭示信息的重要途径与方法。由此可见,分类法与主题法是各种网络信息资源最基本最主要的组织法与检索法^[9]。

3.2 网络环境下使用较为广泛的知识组织方式方法

(1)主题树方式。

主题树是一种网络信息资源的组织方式。它实际是提供一种界面机制,用户通过这个界面只与网络信息资源的主题目录进行交流,并通过主题目录间接地连结并使用多个实际的数据库^[10]。

(2)文件方式。

文件是一个有符号名的一组机关的逻辑记录的集合。一个文件可以包括数据、程序或字符,是计算

机保存处理结果的基本单位^[11]。

(3) 数据库方式。

数据库是在计算机存储设备上合理存放的相互关联的数据集合。

(4) 知识重组。

知识重组是知识组织的一种重要方式,其原理是对原有知识的信息因子或知识关联进行语法结构上的重新组合,其结果产生另一形式的知识产品^[12]。

此外,知识组织的具体方式方法还有:代码法、时序法、地序法、序号法、加权法^[13];引用组织法、文摘组织法、索引组织法等。

(5) 智能重组。

所谓智能重组是指图书馆人利用自身在知识组织方面的素养,在充分了解网络信息资源分布的前提下,根据读者需求,对特定范围的网络信息进行查询、下载、分类、提炼、编辑、加工和输出的过程。笔者认为,智能重组是对网络信息资源进行知识组织的有效方法^[14]。根据所投入的智力程度不同,可将智能重组分为三个服务方式:

原装信息服务。

网上存有大量事实性、数据性、参考性、凭证性的信息,这些信息的知识含量很高,几乎不需要任何二次加工就能利用,对许多读者来说具有极大的使用或实用价值。

重装信息服务。

顾名思义,这一层次的服务不仅要为读者搜寻到特定的信息,而且还要根据读者的要求对所查信息进行“重新包装”。这种重装包括两个方面,一是要对查到的信息进行取舍和编辑,二是要对信息进行提炼,从信息中结晶出知识,增加信息产品的知识含量。

方案信息服务。

这是智力投入最大的高层次服务,它要求我们要对所查到的网络信息进行智力开发,通过分析、比较,不仅要为读者提炼出蕴涵在信息中的各种知识,还要根据读者的现实需求将精炼出来的知识形成一个或多个方案给读者,使读者能够直接将知识的潜在价值转化为现实效益。如俊达石化公司拥有 TH 石油化工液位界面温度测控系统的专利,希望从国内外找到石油化工系统的合作单位,要求为公司查询。我们通过 YAHOO 的搜索引擎找到了几十家国内外厂家的网址,从中选择了几家内容较丰富的网页,并通过这些网页上的相关链接找到了更多

相关网址。我们首先将这些资料编辑整理成册提供给该公司;其次根据公司推荐为他们从中确定三家合作单位供挑选,通过 E-mail,索取资料,反复权衡;最后确定了三个方案。公司采纳了我们的方案,与加拿大的一家公司达成合作意向。所有这一切,都是通过网络并利用信息搜索优势和知识重组来完成的,获得了读者好评。

笔者认为智能重组的基本步骤为:

(1) 确立主题。

网络信息资源丰富多样,如果不确定一个搜索范围,全面出击必将一无所获。一般来说,主题的确立主要有两方面影响因子:一是读者的实际需求,即读者向图书馆提出某一专题的检索要求,信息服务人员根据读者的需求进行网络信息检索;二是图书馆要根据本馆的服务对象和馆藏特色确定“网藏”的收集重点,定期从网上下载有关专题的资料进行加工整理,以备读者检索。

(2) 资源检索。

在明确收集的主题以后,应广泛利用各种搜索引擎对拥有该主题信息资源的网站进行浏览考察,依据所含信息的重要程序、更新频率、知识含量、网站开发者的权威性确定一部分核心网址。鉴于许多网页上的“友情链接”具有像“引文索引”一样的发散功效,我们还可以通过这种链接聚集一部分非核心网址,然后将这些网站分类整理后收藏在文件夹里。

(3) 信息遴选。

不要以为定期从核心和非核心网站上下载专题信息只是一件操作性的工作,它实际上也是一门学问。因此,即使是在事先确定的栏目范围内,信息的专指度和知识含量也各不相同,如果不加取舍地全部下载,势必造成二次信息污染。因此,获取信息必须按照确立的主题遴选,要起到“知识过滤器”的作用。这需经过三个过程:一是模式识别,看该知识是否正确,是否与已掌握的知识一致?二是决策分析,看该知识是否揭示了新的规律?三是知识获取,将最确切的信息筛选进来,利用智能重组的方法为科研课题重装出知识精华。

(4) 编辑输出。

将网上信息下载到本机之后,就进入到了最重要的知识整理阶段。一方面,对精练出来的知识信息再按照事先制订的规则或其本身的内在联系进行合理地编排,如果需要,还应赋予其一定的检索功能。另一方面,要挖掘各自信息之间的内在联系,依

●张李义

MPEG-7 与数字图书馆的信息检索^{*}

摘 要 MPEG-7 称为“多媒体内容描述接口”,主要由描述器,描述方案和描述定义语言组成。MPEG-7 与数字图书馆之间具有许多协同性,MPEG-7 中定义的多媒体内容描述将使数字图书馆的信息检索具有更大的效率,也可以较好地解决目前不同数字图书馆中的互操作性和没有统一标准的问题。图 2。参考文献 9。

关键词 MPEG-7 数字图书馆 多媒体 信息检索

分类号 G250.72

ABSTRACT MPEG-7 is called “multimedia content description interface”,which consists of descriptors, description schema and description definition language. It is closely related to digital library, can be use to enhance the efficiency of information retrieval of digital library, and can solve the present problems in standardization and interoperability. 2 figs. 9 refs.

KEY WORDS MPEG-7. Digital library. Multimedia. Information retrieval.

CLASS NUMBER G250.72

数字图书馆(Digital Library, Dlib)是全球信息高速公路上信息资源的基本组织形式,用于满足分布式面向对象的信息检索需要,这里的“面向对象”指

的是不仅要查到文献信息线索,而且要获得所需的文字、图像、音频等原始信息,这决定了多媒体信息是数字图书馆的一种主要信息表现形式。

据信息在内容上、时间上、逻辑上的相关程度组织知识信息,形成一个有机的知识体系。

(5)增值服务。

对经过加工整理的网络知识信息进行分析研究,通过比较判断,为读者提供具有更多知识含量的方案信息服务。反复多次地重复上述 1~4 个步骤,最后才能达到信息价值的升华。这就属于更高层次的增值信息服务了。

知识和信息的产生都是随机的、无序的,是不可控的,人们要想获取有用的知识信息十分困难。因此,图书馆人还应对有关网络信息的知识组织和服务的新技术如元数据集(Dublin Core)的数据挖掘、自动摘要、在线翻译等研究领域不断探索,加强知识组织,加速知识信息的转化和利用,以服务于读者,推动精神和物质文明建设。

参考文献

1 刘植惠. 知识经济中知识的界定和分类及其对情报科学

的影响. 情报学报,2000(2)

2 贾同兴. 知识组织的进步. 国外情报科学,1996(2)

3,12 蒋永福,付小红. 知识组织论图书情报学的理论基础. 图书馆建设,2000(4)

4 王知津. 从情报组织到知识组织. 情报学报,1998(3)

5,13 田书格. 知识组织浅论. 图书馆理论与实践,1999(3)

6 王忠红. 网络信息环境下的传统分类法. 图书情报工作,1999(2)

7 洪漪,梁树柏. 分类法在信息网络中的应用. 情报学报,1998(1)

8 强自力. 网络分类目录及其分类法. 大学图书馆学报,1999(4)

9 康桂英. 分类法与主题法在网络信息资源组织与揭示中的应用. 图书馆建设,1999(3)

10,11 严怡民主编. 信息系统理论与实践. 武汉:武汉大学出版社,1999

14

李秀云 郑州大学图书馆期刊部主任,副研究馆员。
通讯地址:河南郑州。邮编 450002。

(来稿时间:2002-03-04)

^{*} 本文得到教育部重点研究基地重大项目“信息可视化与知识检索”资助。