

基于信息用户认知的图书馆“信息唤醒”

陆浩东

贺州学院图书馆 贺州 542800

〔摘要〕图书馆“信息沉睡”现象非常普遍,资源浪费也相当惊人。透过网络环境中的信息用户认知现状,分析用户信息认知的特点、认知信息活动的基本模式以及制约用户认知信息活动的主要因素,认为基于用户的信息检索、基于本体论的信息整合及智能信息推送服务是提高用户信息认知能力的有效途径。

〔关键词〕图书馆 用户认知 信息构建

〔分类号〕G250 73

Wake-up of the Library Information Based on the Information User's cognition

Lu Haodong

Hezhou University Library, Hezhou 542800

〔Abstract〕 This paper analyzes the characteristics of the user's information cognitive psychology and the basic mode of the cognitive activity from the present situation of the network information environment. It also analyzes the main factors which constraint the cognitive activities of users information. It discusses that the construction of user-based information retrieval, ontology-based information integration and intelligent information push service is the effective means to improve the users information of cognition.

〔Keywords〕 library user cognition information architecture

1 网络环境下图书馆资源与服务现状——“信息沉睡”现象

目前,图书馆普遍存在文献资源利用率偏低的现象,相当一部分资源被搁置。到2004年为止,我国大学图书馆藏书流通率平均低于40%,个别竟不足20%,也就是说有4亿册图书在睡觉,相当于全国公共图书馆的藏书总量,造成极大的馆藏资源、设备及人员的浪费^[1]。显然,图书馆“信息沉睡”现象十分普遍,存在大量信息一直不为用户所得,很多资源无法使用、增值和再造、共享,既阻碍人们信息利用水平的提高,也不利于建设信息化社会。如何更进一步开发利用图书馆信息资源,让这头“狮子”的清醒是图书馆信息化建设的当务之急^[2]。出现“信息沉睡”的主要原因是:

① 搜索资源问题,网络环境中图书馆的服务发生了巨大的变化,包括导航、检索、分类等都与用户的实际需求相差很大。现在的检索工具多是综合性的,其功能特点存在较大的差异,比如检索点、检索词的表达方式、逻辑组配方式、检出结果的显示方式、链接内容及发布方式等都存在较大差异。但检索软件的智能性不

够高,造成检索的准确性和效率低,检索障碍使用户误检漏检大量信息。图书馆传统的按学科进行分类组织的信息资源系统显得过时与落后了,它和人们一般信息认知不一致,远远不能满足人们的信息使用需求。而搜索引擎等网络信息检索工具具有简洁、易用、交互的特点,使得用户纷纷用它来代替传统图书馆服务。

② 信息内容问题,一方面人们学习和生活带来了极大的信息需求;另一方面用户在良莠不齐、无序的海量信息面前,面临着许多困难和障碍,有时甚至达到无所适从的地步。这些需要图书馆彻底转变观念与方式,一切以“读者为中心”、“以用户为本”来重新审视信息资源的组织与服务。本文将尝试从用户认知角度对图书馆“信息沉睡”问题进行探究。

2 网络环境下的信息用户认知问题分析

用户对自身需求的表达实际上是以自身当下知识结构欠缺的表达,这种表达首先是一个认知问题。用户的信息认知是经过选择、构造、补充三个环节而形成的。这三个环节都和息心理有着密切的联系,直接影响到用户对信息的获取。选择关系到用户获取信息

的适用性,构造意味着信息的整体性,补充预示着信息的完整性。面对纷繁冗杂的网络信息,用户很多时候都无从下手或难以选择。2005年OCLC推出的《对图书馆与信息资源的认知:给OCLC成员的报告》^[3]显示,基于Web的超文本信息环境的形成与发展,使得可利用的潜在认知信息资源的数量和质量大为增长和提高;传统认知信息活动的固定逻辑结构发生随意性变化;认知信息活动难以收敛;认知信息目标出现快速而多维度的转变;用户获取信息的活动面临更多、更复杂的认知选择。这些变化启动了Web用户许多以往未加运用的认知潜能。同时,由于网络信息超载、信息迷航和信息疲劳等问题也可能引起用户利用信息时的认知失调^[4]。网络信息的查寻和利用是用户基于自己的知识储备和认知能力,与检索系统交互作用,以解决当前问题为目的,以完善、改变知识结构为结果的过程。这一过程始于知识结构不充分导致的情报需求,结束于知识结构暂时充足状态。检索系统和信息服务的任务是助成用户这一转换过程的实现和完成。因此,图书馆必须关注用户信息吸收和利用效果。用认知的观点去看用户信息需求和用户行为,用构建用户认知的视角审视信息服务。了解用户的知识结构和认知特征,扩展通向相关信息资源的路径支持,使检索系统发挥综合效应。

2.1 用户认知信息活动的特点

从认知角度来看,人的知识结构对外来信息的依附、同化、选择、建构和实现社会化的过程就是知识的进化。通常用户认知具有如下显著特点:①动态性,信息认知活动本身是动态的、随着认知活动过程的展开而不断调整、变化;②多维性,源于人类社会实践活动的多样性,如知识创新、信息浏览等;③复杂性,主要体现在用户本身的心理活动、社会活动和信息形式的复杂性;④整体性,即认知系统内部各认知问题之间的相互制约性。用户已有的认知结构与信息互动的过程中,一方面用户的认知结构不断扩展、分化和认知重组,认知结构得以更新和优化;另一方面用户认知结构对信息具有选择性吸收。这表明信息服务效果与用户的认知结构有着紧密的关系。如果用户拒绝选择或拒绝授受,其服务效果为零。要获得良好的信息服务效果有两个办法:一是提供适应用户认知结构的信息;二是激活和构建用户的认知结构。基于用户认知特征、结构和模式的新视角对网络信息资源、主观需求和世界的再认识,不仅能改变图书情报的工作方式,更将催生出新的信息组织生产方式:以属性的提炼和整合

为中心的资源需求组合工艺^[5]。

2.2 用户认知信息活动的基本模式

认知过程从广义上说包括信息获取的过程和信息处理的过程(即思维过程),狭义上理解指主观信息产生(即信息再生)的过程。随着信息技术的飞速发展,信息资源的内容、形式以及组织的纷繁复杂,已经远远超出用户所具有的认知能力范围。用户认知信息活动的固定逻辑结构将发生巨大变化,用户信息活动面临多样性的认知选择。面对复杂性带来的用户认知问题,要在了解用户认知过程的基本模式的基础上构建与之相适应的数字信息组织,如图1所示:

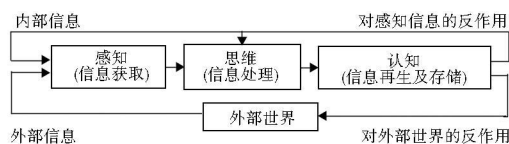


图 1 用户认知过程的基本模式

在复杂的数字信息环境中,形成用户认知信息活动基本模式的三个维度之间的关系表现如下^[6]:①结构反应,即用户对信息环境提供的信息组织方式的反应,亦称信息检索;②目标指向,指反应了用户认知信息活动中对信息需求表述方式的差异;③交互方法,它反映了用户在超文本信息环境获取信息时的认知风格。上述三个方面彼此之间的关系不仅直接影响用户对数字环境的认知能力,而且对用户怎样降低认识成本,如何优化由结构化、可视化技术带来的认知优势具有重要意义。

2.3 制约用户认知信息活动的因素

信息时代的用户面对网络环境中纷繁复杂的信息,制约用户认知信息活动行为不乏有以下几个方面因素:①网络信息的复杂性,由于用户的认知资源的有限性,当面对无序的、良莠不齐的网络信息资源时,用户认知有用信息容量会因为无用信息的侵入而感到有用信息的不足;②信息组织类目过于繁琐,通常用户认知信息组织类目层次容量是5级到6级,如果超过这个层级,用户往往容易产生混乱心理,而且类目层次的增加,难免会信息重叠交叉出现的现象;③用户的信息素养问题,用户自身的知识与知识结构决定了其对信息认知能力高低;④用户对信息的关注程度,用户对于信息的注意取决于当前的唤起水平。在因特网上搜索信息时用户的焦虑、恐惧等情绪,也会影响用户当前的唤起水平,唤起水平的高低直接影响用户当前的注意行为,从而影响用户对信息资源的认知容量和

对信息的注意度。诚然,用户的认知结构既非先天形成,也非客观环境单纯作用的结果,而是在认知主体与客体相互作用的过程中由主体主动构建而成。图书馆的知识组织产品——检索系统的设计,不应只以“科学”、“规范”为依据,而以对应用户的需求表达为根本依据。也就是说,不应只以注重知识组织的技术方法问题,更应重视如何识别用户需求的认知问题^[7]。鉴于上述情况,以人为本,从信息用户认知的角度实施图书馆“信息唤醒”行动方案,以求探索适应用户认知结构的信息组织体系的途径。

3 图书馆的“信息唤醒”行动——基于用户认知特性的信息组织构建

传统的情报检索的弊端是重视系统的经济和技术效率,轻视用户效率,从而限制了系统的效率。在情报检索系统的设计开发实践中也普遍存在着过于重视程序设计和设备的技术问题,在一定程度上忽视了情报检索中的用户因素,从而导致许多检索系统在使用中不受用户的欢迎,检索效果远远低于用户的期望,使开发出来的产品与用户及图书情报学界的期望相去甚远。信息检索机制存在“忠实表达”、“表达差异”和“词汇孤岛”三个深层次的问题,缺乏知识处理和理解能力,基于关键词的信息检索技术已不能满足用户在语义上和知识上的需求,造成用户检索困难的根本和关键是要把信息检索从传统的基于关键词层面提高基于知识或概念层面。语义检索把信息检索与人工智能技术、自然语言技术相结合,立足于对信息进行语义层次上的分析和理解,能有效地理解并挖掘用户的信息请求。信息时代一个是信息,一个是信息技术。信息技术服从于信息,信息技术为信息服务。如何将分散异构的信息资源进行基于用户认知的组织与整合,以充分发挥资源的整体效用,最大限度地满足用户的信息需求,如何唤醒“沉睡信息”呢?

3.1 基于用户认知的信息检索

信息检索是面向用户需求,利用相应的技术和方法汇集、存储信息,以方便用户索取的形式所提供的信息查询和获取的服务。从 DDC 以来的现代图书分类到当前的元数据、知识地图、知识本体等,这些西方式的知识组织模式通过改进文献(知识单元)标引和描述中自然语言的演算方法,以努力达到明确化、形式化和规范化的要求。作为理性逻辑突出的反映,它们忽略了一个根本的事实:人类认知的优点。比如“Ben-

zene”这个词,对化学师而言是“笨”,但对消防人员而言,则意味着另一个意思。因此,当知识组织取决于以技术形式出现的若干精确规则时,也与用户基于自然语言的习惯性提问渐行渐远。古代书目认为知识组织更多地取决于认知主体的认知背景、心理结构和思维方式,凸显的不是科学理性而是人文信念,这对于纠偏现代知识组织的一般理论、方法和原则不无启迪^[8]。

目前图书馆的网站大都采用分类体系,但各网站的分类体系却各自为政,缺乏统一性。为此建立两个入口:一个按照学科体系分类;另一个按照自己网站的特性分类。前者不仅符合用户逻辑思维的习惯,而且也满足具有一定知识体系的用户的认知需要的方式;后者则既符合一般用户的检索需要,又在一定程度上满足了用户的个性化信息需求心理。信息检索过程符合用户的认知心理,用户在提出问题后,信息检索系统将用户的提问加工成适合于计算机检索的语法形式,通过建立问题输入模式的方式帮助用户提出正确的问题。例如规定提问的关键词、类型、逻辑关系等几个方面引导用户按照标准模式提问,这样可以大大地提高检索的效率和准确性。此后,根据检出的结果进一步评价,看其是否符合用户检索目标,能否满足用户的需要。如果不符合要求,要重新对用户的提问加工并进行检索,直到符合用户思维习惯,在一定程度上满足用户个性化信息需求为止。这种理想的情报检索系统应能够在用户为实现知识状态转换,完成与系统的各种交互作用任务过程中,对不同性质、类型知识的使用和需要,以及不同认知能力的运用,提供针对辅助、支持,具有高度适应能力。

3.2 基于 Ontology 的信息跨库整合

信息时代,用户最希望的是如何用最少的时间获得更多、更准确的信息,最关心的是信息的整合和信息的甄别。正如俄罗斯国家图书馆馆长维克多·费多洛夫所说的:“现在的问题不在于阅读信息,而在找到自己所需要的信息,比如上网之后可以找到关于任何问题的答案,可是永远不能确认找到的正确答案,也不知道找到的答案具有怎样的代表性,而图书馆实际上能够解决这方面的问题,也就是说能够保证所需资料的完整性以及代表性”^[9]。数据库的不断增多,各数据库的格式各异,彼此不能互相存取,建立统一的数据存取接口的方案就显得切实可行。

本体(Ontology)的发展为这个接口的实现提供了基础。它作为一种能在语义和知识层面上描述信息系统的概念模型建模工具,是“对某一领域内可共享的、

通用的概念的形式化的规范描述”^[10]。Ontology能使用户在概念层次上设计领域模型,支持多重继承,并对新数据进行一致性检查,可扩展。同时,它后台支持数据库存储、开放源码、模块清晰,提供完全的 API 接口,有助于用户对词汇语义的认知^[11]。Ontology的作用主要有两方面:①澄清了领域知识的结构,为知识表示打好基础;可以重用,避免重复的领域知识分析。②统一的术语和概念使知识共享成为可能^[12]。开发 Ontology 这样一个接口,其实质是建立各个数据库之间的联系,采用一个相同格式的接口建立不同格式的数据库之间的链接,实现语义化 Web 服务中对知识的共享和重用。人们对本体模型的期望是在语义层上描述实体内容以及实体间的关系,对目前的语义化 Web 服务的应用领域比较广泛的需求是采用本体模型建立精确化、智能化的信息检索方案。

在进一步改进 Ontology 提供合作开发支持的基础上,把它应用于包括知识重组、知识表示和知识记忆三个环节的知识组织、知识检索和知识服务,可以将某个或多个特定领域的概念和术语规范化,通过机读元数据自动建立书目数据库;通过语义对网页或文件进行自动标引和注释;通过语义聚类将信息资源中的相关主题进行分类,从而实现数字图书馆的定题服务;也可以对网络站点进行分类和导航,丰富数字图书馆的链接。它增加了检索结果的准确度,使用户避免了大量无用信息的人工筛选,能使检索结果在较大程度上符合用户的真正需求。因此,构建基于 Ontology 的信息跨库整合对消除信息用户检索障碍,提高检索效率,最大限度利用网络信息资源大开了方便之门。

3.3 面向用户的智能信息推送服务

在网络化社会,人们猎取知识和信息的途径逐渐多样化,图书馆不再是唯一获取信息来源的地方。图书馆直面这种冲击,改进多种服务,吸引更多的人利用图书馆势在必行。图书馆的既定目标是争取一切能够利用图书馆的社会成员成为图书馆的用户。不同的用户个体对同一信息对象的注意程度是存在一定的差别的,“推送”技术恰好可以根据用户的需求,将网络信息查询与收集有机地结合起来,建立由面向用户的服务模块和面向主题信息搜索加工模块组成的新型信息服务系统。有目的地按时将用户感兴趣的急需的信息主动发送给用户。在目前“推送”技术的基础上融入人工智能、知识发现技术及数据库技术,从而形成“智能信息推送”(IPP)技术,是“push-pull”技术的一个发展方向。它是将“推送”和“拉取”相结合的

智能信息技术,将会是未来网上信息获取技术的一个重要发展方向^[13]。图书馆利用 IPP 技术将信息查询与收集有机结合起来,解决图书馆供需信息匹配。将经过排序和过滤的信息推送给用户。使图书馆主动地搜集、处理信息并形成决策;将主动采集的信息进行深层加工处理,形成结论或报告,主动挖掘和积累新的知识;将用户需求的内容或动态生成的结果一并推送给用户,建立用户与服务器相关数据库的动态连接,使图书馆充当一个积极主动提供决策信息的智囊团,促使图书馆从“以信息为中心”走向“以用户、利用为中心”的知识创新服务。

4 结 语

网络信息的数量庞大和质量优劣一直是用户检索与利用信息资源的障碍,用户难以确认信息的真实性和可靠性,加之网络的虚拟性,站点的不稳定性和信息发布的自由性,直接影响对信息的可信度。在网络环境下,用户需求及其认知结构都得到了扩展:用户范围的扩展,人机直接交互,用户信息行为更加主动,用户获取信息方式多样,并且用户在信息需求、搜索、检索、利用过程中的认知心理、认知行为、认知能力都呈现出了新的特点,唤醒用户的信息化意识及使用信息技术来处理信息资源的意识,并且养成一种习惯。

近年来我国图书馆的藏书越来越多,服务的功能越来越全,知识、信息、技术全面提升,但却忽略了以人为本,比如忽略用户认知心理、认识能力等,图书馆的门槛被人为加高。构建知识服务的战略平台,在服务平台的前端组织知识咨询、参考服务,在平台的后端组织数据挖掘、知识发现与组织业务,在平台管理层推进知识搜集、开发和利用管理一体化,以用户需求为中心构建服务体系。图书馆与众不同的品牌仍然是信息资源和信息服务。信息用户认为图书馆是一个普通的寻找信息的地方、一个组织,一个拥有众多新老用户的实体。从本质上,图书馆是一个全球化的品牌;一个被怀旧情感所主导、并被日常经历所强化的品牌。这种全球性的、怀旧的认知应当使图书馆界有足够的理由得到关注,但这也为优化图书馆的价值、创造变化提供了一个坚实的基础。

参考文献:

- [1] 大学图书馆该不该向社会开放? 各界人士直抒己见. [2008-12-20]. http://heilongjiang.northeast.cn/system/2007/07/27/050910556_shtm1 (下转第 116 页)

或者根据 xml语法规则增加语种编码声明: <? xml version="1.0" encoding="gb2312"? >。

4.3 主题(Theme)创建及安装技巧

新建主题可以通过修改已有样式模板来快速完成。以新建主题 CAUIR 为例,主要过程如下:

4.3.1 建立 CAUIR 文件夹 把 /dspace/webapps/xmlui/themes 目录下的 template 文件夹另存为 CAUIR 并且其下的 xsl 文件必须与其同名: CAUIR_xsl

4.3.2 修改 CAUIR 目录下的 sitemap.xml 文件 ①把全局变量标签 <global-variables> 下的 theme-path 修改为 CAUIR ②把 theme-name 修改为 CAUIR institutional repository theme(随意命名); ③把 <map transform src="template_xsl"/> 注释掉,添加 <map transform src="CAUIR_xsl"/>。

4.3.3 编辑 CAUIR_xsl 覆盖 dr2html_xsl dr2html_xsl 文件中具有所有输入数据的初始样式,用 CAUIR_xsl 首先调用 dr2html_xsl 然后在它的基础上对某些部分进行样式覆盖。

4.3.4 style.css 文件编写调试 要充分表现界面的个性化风格,还需要对 CAUIR 文件夹下的 style.css 进行修改,并且针对不同浏览器进行兼容性调试。

4.3.5 安装 CAUIR 打开 xmlui.xconf 文件,在 <themes> 下注释掉正在起作用的主题引用行,添加新主题名称与路径 <theme name="CAUIR institutional repository theme" regex=".*" path="CAUIR"/>, 保

〔作者简介〕杨国栋,男,1985年生,硕士研究生。

李晨英,女,1963年生,副研究馆员,硕士生导师,发表论文 20 余篇。

韩明杰,男,1960年生,研究馆员,副馆长,硕士生导师,发表论文 30 余篇。

臧琳,女,1984年生,硕士研究生。

存并重新启动 tomcat 服务。

基于 xml 技术的新一代 DSpace 用户界面不仅出色的完成了基本的浏览、登录、管理等功能,而且能够对不同社区页面实现深入的个性化,混搭元数据与其他多媒体材料,实现扩展功能的共享等。Manakin 能够实现的不仅限于此,更多的功能有待我们深入研究。

参考文献:

- [1] University of Nottingham, UK. The directory of open access repositories - OpenDOAR [2009-04-13]. <http://www.opendoar.org/>
- [2] University of Southampton, UK. Registry of open access repositories (ROAR). [2009-04-13]. <http://mar.eprints.org/index.php>
- [3] 奉国和. 开源软件与图书馆知识管理探讨. 图书馆论坛, 2008(4), 58-61
- [4] The DSpace Foundation. DSpace 1.5.1 Manual [2009-04-13]. http://www.dspace.org/1_5_1Documentation/
- [5] Sun Developer Network. Java server pages overview. [2009-04-13]. <http://java.sun.com/products/jsp/overview.html>
- [6] developerWorks 中国. Cocoon 2: 构建数据库驱动的站点. [2009-04-13]. http://www.ibm.com/developerworks/cn/views/xml/tutorials.jsp?cv_doc_id=84942
- [7] 金望正, 李莹, 徐江浩, 等. 面向方面编程技术研究. 计算机应用与软件, 2005(8), 42-45.
- [8] Scott P. Manakin developer's guide [2009-04-13]. <http://dev.ohiolink.edu/bitstream/handle/123456789/28/ManakinDeveloperGuide.pdf;sequence=1>
- [9] Maslov A, Green C, Mical A, et al. DRI schema reference [2009-04-13]. <http://ditamu.edu/projects/xmlui/schemaReference>

(上接第 49 页)

- [2] 邹清丽. “唤醒沉睡信息”是信息化建设当务之急. [2008-12-10]. http://industry.ceilnet.com/art/35/20060510/549733_1.html
- [3] 对图书馆与信息资源的认知: 给 OCLC 成员的报告. [2008-12-20] <http://www.oclc.org/report/2005perceptions.htm>
- [4] 胡昌平, 向菲, 周永红. 面向用户的信息资源整合与服务. 武汉: 武汉大学出版社, 2007: 69
- [5] 喻峰. 知识创新和情报学初探. 图书情报工作, 2001(3): 39-44
- [6] 毕强, 宋绍成, 杨达. 超文本信息环境用户认知活动的模式及其影响因素. 图书情报工作, 2003 47(6): 77-79
- [7] 付雅慧. 我国知识组织论研究的回顾与展望. 大学图书馆学报,

〔作者简介〕陆浩东,女,1968年生,副研究馆员,硕士,发表论文 15 篇。

- 2009(3): 44-48
- [8] 李满花. 图书馆学研究本土化何以可能: 基于中国古代图书馆学思想价值的可行性思考. 图书情报知识, 2008(1): 41-44.
- [9] 石剑锋. 图书馆不会被网络动摇. [2008-12-20]. <http://www.enorth.com.cn>
- [10] Ontology 是什么?. [2008-11-20]. <http://www.Ks1StanfordEdu/people/dln/papers/ontology101/ontology101-noy>
- [11] 杜小勇. 语义 Web 与本体研究综述. 计算机应用, 2004(10): 15-16
- [12] 邢维慧, 袁建敏. 用户信息服务的认知心理分析. 情报科学, 2004(11): 1405-1408.
- [13] 计算机世界网. [2008-12-20]. <http://www.2cow.com.cn/01/0145/b/0145b02-3.asp>