

陈光祚 阮建海 臧国全

# 论个人数字图书馆

**摘要** 个人数字图书馆是数字图书馆的一种类型,是最贴近用户个性化需求的数字图书馆,是e时代的私人藏书楼。它具有资源配置合理、固化信息功能、费用低廉和简易性等特点。宣传普及个人数字图书馆的理论与方法将是图书馆用户教育的另一个热点。表1。

**关键词** 个人数字图书馆 数字图书馆 电子信息组织 网络信息管理

**分类号** G250.76

**ABSTRACT** As a kind of digital library, personal digital library is a personalized digital library and a private library in the electronic age. In this paper, the authors discuss details of such kind of digital library. 1 tabs.

**KEY WORDS** Personal digital library. Digital library. Electronic information organization. Network information management.

**CLASS NUMBER** G250.76

所谓个人数字图书馆,是指个人为了读书治学,在自己的计算机上采用免费或基本免费的全文数据库软件,将有关的网上信息和自创的数字化信息资源进行采集、存储,使之成为有组织的信息集合。个人数字图书馆是数字图书馆的一种类型,是最贴近用户个性化需求的数字图书馆。实际上,个人数字图书馆就是e时代的私人藏书楼。建立个人数字图书馆是广大网络用户的要求,普及个人数字图书馆建设的理论与方法,将是用户教育的另一热点。

## 1 营造e时代的私人藏书楼——个人数字图书馆

在古代,私人藏书楼(以下简称“私藏”)在传播文化、培养人才及保存文化典籍方面起了很大作用。随着公共图书馆的出现和发展,私藏的地位日益被取代。但是应该承认,私藏一直延绵至今,个人、学者都有规模不等的藏书,然而,私藏需要相当大的经济支持。私藏与图书馆是互相补充、相辅相成的。

今天社会已进入e时代,也就是电子信息和网络时代。整个因特网就其信息资源来讲是一个巨大无比的图书馆。网上的大量信息(论文、专题数据库、新闻等)可供用户下载。但是网上资源过于庞杂,所驻存的网站网页不稳定,内容质量参差不齐,检索使用不便,使得用户利用时力不从心,不得要领,无所适从。于是就出现了各种网络信息的检索工具和提供个性化服务的众多项目。

网上存在浩如烟海的数字化信息资源,并提供有数以千计的搜索引擎、网络导航等检索工具,并且许多图书馆和商业性机构在网上提供非营利性或营利性的数字图书馆服务,其信息加工的标准化程度高,查找方便,信息更新及时,上网即可获得个人所需信息。加之当前网上倡导与实行个性化服务,力求满足个人用户的特殊要求。鉴于此,或许有人会问,建立个人数字图书馆还有必要吗?我们的回答是肯定的。

(1)从信息资源配置和利用模式来看,个人数字图书馆的存在是必要的、合理的。历史上,公共图书馆的发展与普及,在任何时候都不可能完全取代和消除个人藏书的存在。在我国漫长的封建社会中,为数不少的私人藏书楼在保存文化典籍、传播文化、辅助培养人才等方面发挥了很大的作用。随着图书出版业的兴旺发达,出版的图书大部分是由个人购买,而非图书馆所购买。私人藏书是治学者的案头之书,是他们最接近的也是使用最频繁的信息来源。文献计量学的统计也表明,在作者所引文献中,个人藏书的被引率往往是最高的。当然,个人藏书在规模、支持费用等方面的局限性是显而易见的。要想获得全面的信息需求,必须依靠面向公众的信息源。在传统模式中,这种面向公众的信息源和信息服务主要就是各种类型的图书馆。它们的营运成本由社会机构(包括政府、团体、学校等)所承担。在e时代,传统印刷型的信息资源纷纷进行数

字化上网,再加上众多的报纸、杂志、图书都在网上出版发行,甚至传统的信件也通过 E-mail 收寄,因特网成了储存、传播人类有史以来最大的信息资源库的工具。人们从网上下载与自己工作、生活、研究相关的各种信息,日积月累,下载的信息多了,就需要管理。个人数字图书馆就是实现个人数字化信息收藏和对个人收藏的数字化信息进行管理的工具。

(2)个人数字图书馆实现了真正意义上的个性化服务。因特网是一个超大规模的信息资源库,为了有效利用这个信息资源库,出现了众多的检索工具。但是,包括数字图书馆在内的众多的因特网信息资源检索工具提供的大都是一个范围较大的、比较粗放的工具。目前网上提供的各种个性化服务从其原理上讲仍是基于机械性的匹配,只能说是处于个性化服务的初级阶段,难以满足个人的真正需求。个人数字图书馆的建立才是提供真正个性化服务的有效途径。

(3)从收录信息资源的类型来说,个人数字图书馆也比面向公众的数字图书馆广泛。个人数字图书馆除网上下载文献外,可收录个人文章与著作的电子文本、自己的 E-mail 文件等。在一个库内,所收录的文件类型和格式可以是多种多样的,如文本文件、HTML 文件、DOC (WORD) 文件、WPS 文件、图形,乃至动态图像、音乐等文件。个人数字图书馆的收录文件之所以多类型、多格式、多编码、多文种,是因为它采自网上不同的网站和数据库,但它们都同属某一主题,可以混合建库。因而数据库结构显得杂乱、不够标准化。但这种做法正是个人数字图书馆的优点。例如以文本文件作为存储的主要形式是它体积小,可以进一步作文字编辑处理。而以 HTML 格式存储,则可以保留必要的图像图形,并保留进一步进行链接的功能。DOC 文件是学者们产生自己文件、彼此交换文件的常用形式。而面向公众的数字图书馆则收录范围较广,收录文献类型较为单一,不同类型的文献往往分别建库,使用时得分别检索浏览。所以,现在已经构建和正在建设的数字图书馆无法实现个人数字图书馆的这种功能。

(4)个人数字图书馆具有固化信息的功能。它一旦建立,就相对独立,不受网络通讯的影响,不受网站及其资源的更新的影响,也不受有关软件变更的影响。个人图书馆中的信息资源是个人“到手了”

的资源,因而具有“实在性”,是一种“实藏”。目前因特网上广泛存在的虚拟图书馆也无法取代个人数字图书馆。

(5)面向公众的数字图书馆一般由公共图书馆、高校图书馆、科研部门图书馆、商业性机构支持经营。它们一般是将原来所负担的为公众提供信息服务的职能在网上得以实现,数据库构建复杂,要求有专职人员,成本甚高。而支撑个人数字图书馆的软件功能相对简单,对收录的信息一般不进行著录、标引等预处理,检索时以对全文关键词扫描为主。构建个人图书馆可由用户自己一人完成。

目前看来,建立个人数字图书馆也是可行的。

首先,从可采集的信息资源来看,供个人数字图书馆采集的信息资源已具相当规模,并呈指数增长趋势。个人数字图书馆的信息来源除主人的数字化稿件、家庭像册、导入的外部文件外,大量的的是对网页的抓取和拖放。据估计,目前因特网上的网页数量已达数十亿页,并呈快速增长的趋势,这为个人数字图书馆提供了可采集信息的保障。

其次,从建设成本和运行费用来看,建立和运行个人数字图书馆的费用是很低廉的。在已具备微机和上网条件的情况下,支撑个人数字图书馆的软件大都是免费的,可在网上直接下载。个人数字图书馆的运行费包括抓取和拖放网页时的上网费。

再次,从软件技术上讲,能够胜任个人数字图书馆主要功能的软件已基本成熟。

综上所述,e时代的私人藏书将会有很大发展,建立个人数字图书馆的时代已经到来。

## 2 个人数字图书馆与面向公众服务的数字图书馆之比较

关于个人数字图书馆较之面向公众服务的数字图书馆的特点及其两者的比较可见表1。

## 3 个人数字图书馆软件的评价指标体系

一个理想的个人数字图书馆软件应具备下述功能:

(1)信息的组织方式。指个人数字图书馆对导入信息及个人自产信息的组织方式。包括:建新库;建新文件夹;独立存库;对导入文件自动建立索引;对记录的编辑功能,包括增加、删除、修改内容,字体大小及颜色的设置、插入图片。

(2)可导入的文件类型。导入广泛类型的文件是

表 1 个人数字图书馆与面向公众服务的数字图书馆之比较

| 比较项目    | 个人数字图书馆   | 面向公众的数字图书馆                                   |
|---------|---|--|
| 服务对象    | 面向个人提供信息存储、检索、输出等服务。服务面窄。对象明确。  | 面向公众提供信息服务。服务面广,有一定的综合性。                     |
| 收录范围    | 收录信息的范围专深,贴近个人用户的需求。  | 收录的范围较广,可供不同个人用户选择。                          |
| 收录的文献类型 | 类型较广,主要是网上下载的文本资料、网页、图像、DOC 文件、WPS 文件、CAJ 文件、音乐等, E-mail, 以及个人自产的电子文件,如作者稿件与出版发表文章的电子文本,家庭像册等。上述不同类型的信息可以综合建库。<br>收录的文献以全文为主。 | 文献类型较为单一。不同类型的文献往往分别建库。<br>收录的文献可以是书目、文献全文等。 |
| 信息加工    | 力求简单,一般不标引,不追求著录的标准化。   | 深度大,往往有词表支持。强调著录、标引的标准化。                     |
| 信息检索    | 以自然语言为主,以中文顺序扫描为主要的检索手段,辅之以字段检索(文献题名或正文),一般对用户的检索知识要求较低,易使用。  | 由于有人工语言的支持,检索效率较高。要求用户具备一定的检索知识。             |
| 检索结果输出  | 要求支持检索结果的多种输出功能,如检索结果以字顺排列、相关度排列、进度先后排列。同时要求检索结果以文本、HTML、DOC 等格式输出,以供用户不同需求的利用,并能建立专题文献汇编、调入自己写作的文章等。                         | 检索结果一般仅供网上在线浏览,少有下载功能,以保护数据的产权。              |
| 使用方式    | 除采集网上资料时需上网进行外,一般以脱机方式进行检索、浏览、打印及其他形式的输出。因而节约上网连接的时间与费用。  | 每次利用都必须进入网络,较为费时,通讯费用较高。                     |
| 软件购买费用  | 个人数字图书馆的软件一般可从网上免费下载。   | 必须购买、租用或自编数字图书馆的软件,费用较高。                     |
| 软件维护费用  | 较低。   | 较高。  |
| 收藏形式    | 主要由下载资料、接收的 E-mail 及个人自产电子文本为主,可以不依靠网络,因而也不受网络更新的影响。实际上是个人长期反复利用的“实藏”。  | 资料都在网上,且一般不能下载,对用户来说是虚拟藏书。                   |

评价个人数字图书馆的重要指标之一。包括:文本文件,格式有 TXT,DOC,WPS,RTF;网页文件,格式有 HTML,HTM,ASP,XML,MHT,SHTML;压缩文件,格式有 ZIP,ARJ,RAR,ACE;多媒体文件,格式有 BMP,GIF,JPG,DPEG,PCX,TIF,SWF,PDF,AVI,RM,MP3,DAT 等;其他类型的文件,格式有 EML(MS OUTLOOK 接收的 E-mail 格式),INI,ICO,CAJ。

(3) 导入功能:包括:导入文件夹中的所有文件和选定单个文件导入;目录捆绑功能;对因特网网页的在线抓取功能;对网页内容的在线拖放功能,包括

网页中文字和图片的拖放;导入时的自动压缩功能。

(4) 检索功能。包括:组合检索;单词检索;布尔逻辑检索;题名限定检索;全文检索(检索词可是文献中的任何字符串);对 DOC 文件直接检索,而不是调出 WORD 文档后再利用 WORD 软件中“编辑”菜单的“查找”项进行关键词检索;对 CAJ 文件直接检索;是否用词表;检索速度。

(5) 检索结果的输出功能。包括:将检索结果的题名一次全部列出显示;对检索结果命中的每篇文献有字频统计功能;检索结果的排序,包括按相关度、按字顺、按入库时的时序选择排序;对命中的

赵继海 王效良

## 正确处理数字图书馆宏观决策中的八个关系

**摘要** 为了正确把握数字图书馆的建设方向,避免在项目和技术选择上发生重大决策失误,要认清和恰当处理一些关系问题:数字图书馆与传统图书馆、技术先进性与适用性、数字资源建设与整合、业务社会化与个性化、项目建设与用户服务、馆际协作与资源共享、计算机人才与馆员队伍建设、数字化建设与体制创新。参考文献5。

**关键词** 数字图书馆 图书馆建设 宏观决策

**分类号** G250.76

**ABSTRACT** In this paper, the authors think that to develop digital library in a correct direction and to prevent major mistakes in technological decision making, we should properly treat the following relationships: digital library and traditional library, technological advancedness and suitability, development and integration of digital resources, workflow socialization and personalization, project development and user service, interlibrary cooperation and resources sharing, computer professionals and library staff development, and digitization development and system innovation. 5 refs.

**KEY WORDS** Digital library. Library development. Macroscopic decision making.

**CLASS NUMBER** G250.76

近年来,数字图书馆已经成为图书馆学研究最热门的领域。这些研究主要集中在数字图书馆的技术层面,特别是集中在数字化生成技术、数据存储技术、Web 信息检索技

术和数字图书馆整体解决方案及其技术平台的开发应用。这是因为数字图书馆之产生,根本缘由便是技术进步的结果。但是,数字图书馆建设是一项非常复杂的系统工程,涉

WORD 文档及 CAJ 文件进行全文浏览,无需调用专用浏览器;对命中文献中的检索词有反象显示功能,包括对题名和正文中的检索词都进行反象显示,并直接指向反象显示的第一处;检索结果显示时无乱码;检索结果显示时,显示窗口可自动覆盖原文件;检索结果的打印功能;输出时的格式转换功能,包括可转换为 TXT,DOC 及 HTML 格式。

(6)软件的辅助性能。包括:汉字内码自动转换功能,包括 GB, BIG5 及 HZ 三种汉字编码之间的自动转换;软件的易用性,包括联机帮助,界面设计符合心理学的要求,使用过程中有向导;软件的可持续更新保障;软件的购买费用;软件的保护个人隐私功能,包括可设置系统口令、数据库口令、数据库中文件的口令等,并提供更改口令的操作;自动备份功能;配套资料,包括用户手册、系统说明书;与面向公众的数字图书馆相比,个人数字图书馆可管理的数据库容量要小得多,无需超大规模的数据库支撑,但容量要适中;软件大小;安装个人数字图书馆软件时,如果自带插件,就比较方便,否则,就要搜索所需插件,给用户带来麻烦;软件的稳定性。

### 4 宣传普及个人数字图书馆的理论与方法

宣传普及个人数字图书馆的理论与方法是图书馆把开发信息资源的又一把钥匙交给了用户。如果说《文献检索》课的开设是中国图书馆进行用户教育的重要一步,那么,把建立个人数字图书馆的理论与方法教给用户将使用户教育迈上一个新的台阶。

即将兴起的个人数字图书馆将打破现有的信息服务的格局,为图书馆的信息服务增加一个全新内容。图书馆应热情支持并结合图书情报学的理论,大力宣传与普及个人数字图书馆的理论和方法,丰富教育内容,使用户教育迈上一个新台阶。

陈光祚 武汉大学信息管理学院教授、博士生导师。  
通讯地址:武汉市。邮编 430072。

阮建海 西南师范大学图书馆副馆长、副教授。武汉大学信息管理学院博士生。通讯地址同上。

臧国全 郑州大学信息管理系副教授,武汉大学信息管理学院博士生。通讯地址同上。

(来稿时间:2001-11-19)