

●朱学芳 袁顺波 徐 强

我国数字图像信息资源应用现状及分析

摘 要 在对数字图像信息资源应用现状的调查与分析中,可看到其用户接纳程度是比较高的。数字图像信息资源已经成为人们学习、工作、生活中重要的信息资源。用户获取这种信息资源的途径与方式是多样化的。但由于数字图像信息资源本身具有比较复杂的特点,为更好地促进其应用,不仅需要研究信息处理技术,提高相关知识的普及率,还要加强数字图像信息资源的开发与管理研究。表1。图4。参考文献7。

关键词 数字图像信息资源 图像信息 资源管理 调查研究

分类号 G250.76

ABSTRACT In an analysis of a survey of the present status of the application of digital image information resources, we can see a high user acceptance. Digital image information resources have become important information resources in people's work, study and social life. User's approaches to obtain such information resources are diversified. Because of the complexity characteristics of digital image information resources, the authors recommend studying information processing technologies, promoting the popularity of related knowledge and strengthening the researches on the development and management of digital image information resources. 1 tabl. 4 figs. 7 refs.

KEY WORDS Digital image information resource. Image information. Resource management. Investigation.

CLASS NUMBER G250.76

发展数字信息资源及其技术是信息时代的要求,是知识经济发展的需要。在各种数字信息资源中,数字图像信息资源是其中一个重要组成部分,它以生动的形象、易于理解的表现形式,成为在公安刑侦、广播电视、新闻、商业、教育、科技、医学等众多领域应用的媒体形式之一。因此,非常有必要加强数字图像信息资源的应用研究。加强数字图像信息资源的应用研究,需要了解应用图像信息的领域有效应用相关的图像信息处理技术的方式,针对图像信息的获取与表示、存储与检索、识别与监测、网络化压缩传输等应用研究的特点、方法^[1],还需了解数字图像信息资源的应用现状,调查分析数字图像信息资源的应用情况,明确数字图像信息资源应用中所存在的问题,征求各层次用户对数字图像信息资源应用状况的客观评价,以促进数字图像信息资源在社会各领域中更好地应用,加快社会信息化进程。

为了了解数字图像信息资源的应用现状,了解用户对目前数字图像信息资源应用的评价,本次调查研究主要涉及数字图像信息资源与印刷型图像信息资源的关系,数字图像信息资源的应用状况和用户对当前数字图像信息资源应用相关状况的评价等问题。采取随机抽样调查的方法,发放纸质调查问卷1600

份,电子问卷500份。其中,纸质问卷回收1341份,回收率为83.81%,有效问卷1310份,有效率81.88%。电子问卷回收109份,回收率为21.80%。需要说明的是,由于电子问卷题量小回收数量相对较少,因此本次分析以纸质问卷的调查结果为主。

1 数字图像信息资源的优缺点

在用户眼中的数字图像信息资源有哪些优缺点?他们是否会因为其存在的优缺点而增加或减少其利用频率?一般说来,数字图像信息资源查找使用方便,内容丰富,多种格式、多种类型的数字图片均可获取,并且更新速度快,还可以通过网络等媒体快速传递,实现共享。通过调查可以看出,数字图像信息资源的这些优点已经得到了用户的广泛认可。有813位用户认为数字图像信息资源查找方便,占总数的62.06%;834位用户认可数字图像信息资源更新速度更快;801位用户认为数字图像信息资源内容丰富。与此同时,大多数用户认为与印刷型图像信息资源相比,数字图像信息资源存在着需要在电脑屏幕浏览,比较吃力、使用受硬件条件限制、需要下载特殊浏览器以及需要学习一定的专业知识等缺点。

表1 数字图像信息资源优缺点分布

优点	查找使用方便	共享性	内容丰富	更新速度快
认可用户数	813	773	801	834
比例(%)	62.06	59.01	61.15	63.66
缺点	电脑屏幕上浏览比较吃力	使用受硬件条件限制	需要下载特殊浏览器	需要学习检索方法
认可用户数	483	814	579	329
比例(%)	36.87	62.14	44.20	25.11

鉴于数字图像信息资源存在的这些优缺点,调查结果表明,用户的使用情况是,接近半数的用户不会因数字图像信息资源存在的缺点而减少其利用频率,47%的用户则有可能因此而减少数字图像信息资源的使用,只有极少数用户表示肯定会因此而减少使用。同时,30%的用户表示并不会因为数字图像信息资源的诸多优点而减少印刷型图像信息资源的使用,因为印刷型图像资源也有其固有的优点,如阅读方便、保存时间长、携带及应用方便等。而表示肯定会减少印刷型图像信息资源应用的用户人数为231人,占总数的近两成。据此可预测,在今后相当长一段时期内,数字图像信息资源和印刷型图像信息资源将会共存,二者都会得到广泛的应用。印刷型图像信息资源已经受到了数字图像信息资源的冲击,数字图像信息资源也将在应用方式的进化与完善的过程中逐渐占据主导地位。

2 数字图像信息资源应用现状

用户利用数字图像信息资源的相关情况是本次调查重点关注的问题,通过调查,数字图像信息资源应用有以下几个比较明显特点。

2.1 数字图像信息资源已经成为重要资源

首先,在用户利用的图像信息资源中,数字图像信息资源占据了主要部分,在调查数字图像信息资源占用户利用的图像信息资源的比例时,有40%的调查对象反映为50%~80%,有14%的用户则选择了80%以上,此外还有三成多的用户则选择了20%~50%;同时,40%的用户平均每周使用数字图像信息资源2~4次,314位用户使用4~7次/周,占总数的23.97%,使用7次以上/周的用户占总数的13%。

其次,用户获取数字图像信息资源的目的涉及到生活中的许多方面,有925名调查者反映在学习中应用数字图像信息资源,占总数的70.61%,在工作、生活和娱乐中应用数字图像信息资源的人数也分别达到了586、557和862人,比例分别为44.73%、42.52%和65.80%,网络调查的数据也显示,近8成

的用户应用数字图像信息资源的目的是为工作、生活、娱乐等,这也再次证实了数字图像信息资源已经渗透到了人们生活中,成为人们学习、工作、生活中的重要信息资源。

最后,用户可以从众多场所中获取、利用其所需的数字图像信息资源,调查数据显示,在1310名调查者中,在家即可获取、利用数字图像信息资源的用户达514名,比例高达近40%;在学校、工作场所、网吧、公共场所的用户分别为738、417、525、132,分别占总数的56.34%、31.83%、40.08%和10.08%。可获取场所的多样化无疑有利于用户更为方便、快捷地获取其所需资源,而这也正是数字图像信息资源得到广泛应用的体现和一大重要原因。

数字图像信息资源能够成为人们生活中的重要资源,究其原因,主要得益于以下几个方面的进步:第一,我国经济的快速发展,改革开放以来,我国经济实现了持续、快速、稳定增长,人民生活水平不断提高,经济的发展,为用户产生、利用数字图像信息资源提供了的各种物质条件和环境保证;第二,社会信息化水平大幅提升,社会信息化与数字图像信息资源的应用相辅相成、相互促进。第三,我国教育事业的发展,信息主体的素质对信息应用水平有着非常重要的影响。《中国互联网发展状况统计报告》也显示,网民普及率与受教育程度有直接的联系,学历越高,网民普及率越高^[2]。改革开放以来,我国教育投资、受教育的人数逐年增加,教育事业得到了较大的发展,大大提高了信息主体的素质,从而也促进了数字图像信息资源在各领域的广泛应用。第四,网络的普及,据CNNIC统计,截止到2006年12月,我国上网计算机总数为5940万台,域名总数为4,109,020个,网站总数为843,000个,全国网页总数为44.7亿个,其中29.5%为专门的图像网页,在多媒体网页中,gif格式的网页占87.9%,jpg格式的网页占10.9%。网民获取信息的主要途径依次为网络(85.0%)、电视(66.1%)、报纸(61.1%)、杂志(19.5%)等^[3],这些都表明互联网已经在我国得到了大范围的普及。网络的发展为数字图像信息资源的应用提供了直接条件。

2.2 用户获取数字图像信息资源的途径、方式多样化

获取途径是否方便快捷,获取方式是否多样,都会在一定程度上影响数字图像信息资源的应用。调查数据表明,用户可以通过多种途径来获取其所需的数字图像信息资源,其中,图像搜索引擎是用户的首

选,有804名用户选择通过图像搜索引擎获取数字图像,占总数的61.37%,紧接其后的获取途径分别是网站、数码相机、电脑上画图、网络论坛和扫描仪。这就表明,现阶段人们已经可以通过多种途径获取所需资源,这必将促使数字图像信息资源更进一步地渗透到人们生活中来。

此外,用户在获取数字图像信息资源时,主要是通过关键词进行检索,在1310名调查对象中,有995名用户选择了通过关键词进行检索,占总数的75.95%,此外还有一部分用户选择通过图像的内容特征(如颜色、纹理等)和综合途径进行检索。究其原因,是目前为广大用户所熟悉的数字图像检索工具主要是图像搜索引擎,而现有的知名图像搜索引擎如Google、Ditto、百度等提供的都是基于关键词的检索服务。

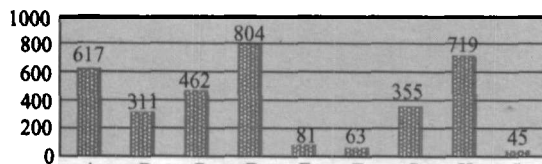


图1 数字图像信息资源获取途径

A:数码相机,B:扫描仪,C:电脑上画图,
D:搜索引擎,E:自建图像数据库,
F:购买的数据库,G:BBS,H:各类网站,I:其他

2.3 用户普遍具备了一定的数字图像信息处理能力

信息能力素质是指信息技术应用能力,信息查寻、获取能力,信息组织、加工、分析能力^[4]。越来越多的研究表明,对社会信息应用水平具有最显著影响的,并不是物质方面的因素,而是劳动者的信息素质。“收入与技术利用的关系是稳定的,但教育与技术利用的关系却水涨船高。在决定技术的采用上,费用将变得相对不重要。”^[5]因此,用户的信息处理能力与信息的应用水平息息相关。

大多数用户具备了一定的数字图像信息处理能

力,有近60%的用户使用过一种数字图像图形应用软件处理数字图像信息。近一半的用户拥有数码相机,30%多的用户在家或单位拥有扫描仪,拥有打印机及其他设备的用户也不在少数;此外,有少数单位已经建成或正在建设专门的数字图像管理信息系统。而在网络调查中,83.49%的用户曾使用过数字图像图形处理软件,超过半数的用户拥有相应的数字图像输入/输出设备。我们知道,图像信息处理能力的培养,图像信息技术的掌握与应用都需要一定时间的学习,其过程一般是从对相关基础知识的了解入门,使用一些图像图形工具软件,更深层次是能够研究出解决有关应用领域的实际图像处理问题的方法和实用系统^[6]。因此调查表明,用户并不是被动地接受,而是以一种积极态势参与到数字图像信息资源管理中来,他们既是数字图像信息资源的消费者,也是其创造者和处理加工者,这种良性的互动有利于数字图像应用水平的提高。

2.4 相关知识的普及率有待提高

普及数字图像信息资源的相关知识,将有助于用户更为客观、全面地了解数字图像信息资源,增强用户对利用数字图像信息资源的兴趣,进一步培养用户的接纳度和满意度。同时,还可以发掘、培养潜在的用户。因此,提高数字图像信息资源相关知识的普及程度能推动数字图像信息资源在更大范围内的应用。在本次调查中,主要设计了数字图像信息资源的相关标准和存储格式等相关的问题,结果表明有一半的用户对数字图像信息资源的相关标准如SVG、MPEG-7、VRML等都不了解,26.41%的用户对数字图像信息资源的存储格式如JPEG、TIFF、BMP等都不了解,具体如图2所示。这就说明数字图像信息资源相关知识的普及并不理想,有待提高。分析其原因,主要是因为数字图像资源相对于文字来说本身就更为复杂,并且目前缺乏一个统一的机构对数字图像信息资源进行广泛的宣传、推广,而在各高校的信息知识教育中也大多未有涉及到图像信息资源的相关知识。

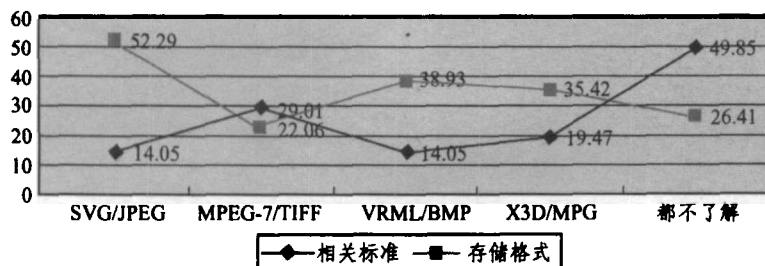


图2 数字图像信息资源相关知识普及率

3 数字图像信息资源管理评价

数字图像信息资源的应用是数字图像信息资源生命周期中的重要一环,同时,数字图像信息资源的应用又与其创建产生、采集、组织、发布、应用等其他管理环节相辅相成,相互影响。因此,研究数字图像信息资源的应用现状,就有必要了解数字图像信息资源管理中众多环节的现状。为此,本文采用利克特(Likert)分量表来调查用户对数字图像信息资源管理的评价,要求调查对象表明他是“强烈赞同”、“赞同”、“反对”、“强烈反对”或“未决定”,也可以根据需要简化或增加等级^[7]。在本次调查中也分为5个等级,分别是“非常同意”、“比较同意”、“基本同意”、“比较不同意”、“非常不同意”。主要从用户总体满意度,数字图像信息资源的分类、标引,数字图像信息资源的获取便捷程度,相关知识产权问题以及数字图像信息资源重要程度等角度出发,得出基于用户满意度的数字图像信息资源管理评价,相关数据见图3,图4。

调查数据表明,“用户的学习、科研、工作和生活当中,数字图像信息资源起着非常重要的作用(A1)”,对此非常同意的用户608人,占总数的46.41%,而对此非常不同意的28人,只占总数的2.14%。同时,该项总得分为5433分,平均得分为4.15,说明用户也非常认同在他们的生活当中,数字图像信息资源是起着重要作用的。

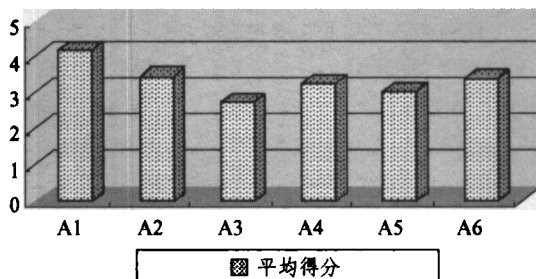


图3 A1-A6 平均得分

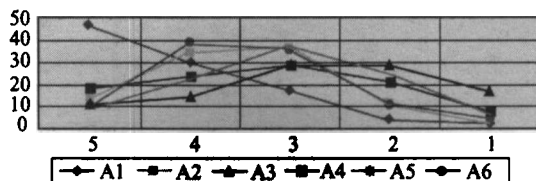


图4 A1-A6 中各等级人数比例示意

对“用户能够方便、快速、准确地获取所需数字图像信息资源(A2)”的说法,非常同意的有181人,

非常不同意的有0人,总得分为4527,平均得分为3.45。这表明用户基本上能够方便获取所需的数字图像信息资源,但也存在着一定的问题。这主要是因为目前用户广泛应用的图像检索工具主要是基于关键词的图像检索,这种检索方式虽然简单、比较容易实现,但是需要人工标引,查全率相对较低,并且存在不精确甚至错误的描述文字误导图片索引的可能,同时目前图像标注缺乏统一标准,客观性较差,容易产生“答非所问”的检索结果。

对“检索、利用所需的数字图像信息资源时,并无不健康图片(A3)”的陈述,138人表示非常同意,221人则非常不同意,基本同意的人数为379人,占总数的28.93%。总得分为3580分,平均得分为2.73。这说明用户在获取数字图像信息资源时,可能检索出一些不健康的图片,这主要是因为网络中各种信息鱼龙混杂、泥沙俱下,而各种检索工具一般都采用机器自动搜集、分类、标引数字图像信息资源,因而导致了这一现象发生。

对于“通过搜索引擎或其他网站,获取、利用数字图像信息资源,并不侵犯其知识产权(A4)”的陈述,对此非常同意的235人,非常不同意的105人。该项总得分4223分,平均得分3.22。这说明现在已有一部分用户意识到利用数字图像信息资源时应该保护资源所有者的利益,而知识产权的有效保护也正是利用数字图像信息资源的前提。

对于“现有图像数据库中的图像信息分类、标引准确、合理、规范(A5)”的陈述,从非常同意到非常不同意的人数分别为111、300、502、325和72人,比例分别为8.47%、22.90%、38.32%、24.81%和5.50%。总得分为3983,平均得分为3.04。这说明用户对现有图像数据库中图像信息分析、标引并不是十分满意,现有的分类、标引技术还有待进一步完善。

对于“总体而言,对目前可供应用的数字图像信息资源很满意(A6)”的陈述,非常同意的126人,非常不同意的62人,509人选择了比较同意。总得分4425,平均得分3.78。说明目前用户对数字图像信息资源管理总体状况还是比较满意的,但由于现在数字图像信息分类、标引、检索都还有尚待完善的地方,因此,用户的满意度并不是非常高。

4 结论

通过这次数字图像信息资源应用现状调查及分析可以看到,用户已经比较充分地认(下转第74页)

政务文件进行受控标引仍非常必要。

(2)自动标引是发展趋势,应当尽早上马。手工标引费时费力,自动标引的效率大大优于手工标引。对比测试表明,自动标引与自由标引的相符度为76.06%,自动标引的质量接近自由标引。随着网络环境下信息资源的飞速增长,完全依靠手工方式对信息资源进行标引显然是不可能的。“由于文献数量和处理时间方面的要求,可以在一定程度上容忍比较粗放的标引处理”。因此在网络环境中,“应当对标引的要求进行相应的调整,降低对标引准确性的要求”^[10]。文献信息部门在条件具备时,应当尽早采用自动标引。目前国内一些文献信息部门,如中国医科院信息研究所和中国电信集团上海黄页信息公司等均已转向机器标引;一直使用受控标引的上海图书馆《全国报刊索引》编辑部,从2000年开始与南京农业大学信息管理系合作开发了“《全国报刊索引》数据库自动标引与自动分类系统”,该系统已经投入使用。

面对目前自动标引效果仍不能令人满意的现状,建议在自动标引系统投入使用的初期对机标结果进行人工判别(相当于标引终审),以确保标引质量。我们相信,随着自动标引及相关技术研究的深入,自动标引的质量将会不断提高,逐步达到令人满意的效果。

参考文献

- 1 邹刚,刘洋等.面向 Internet 的中文新词语检测.中文信息学报,2004(6)
- 2,4,8 马张华,侯汉清.文献分类法主题法导论.北京:北京图书馆出版社,1999
- 3 苏新宁.汉语文献自动标引综析.情报学报,1993(4)
- 5 仲云云,侯汉清,薛鹏军.网页自动标引方案的优选及标引性能的测评.情报科学,2002(10)
- 6 王兰成.基于 EMM 中文抽词算法的 XMARC 主题信息挖掘.情报学报,2005(1)
- 7 周小磊,侯汉清.书目数据库与引文数据库标引质量的测评.图书馆理论与实践,2003(1)
- 9 周全明.六十年代以来我国标引深度研究综述.情报学报,1994(6)
- 10 马张华.论自动标引的实际应用.图书情报工作,2003(2)

刘 竟 南京农业大学博士研究生.通讯地址:南京童卫路6号.邮编210095。

朱书梅 南京农业大学硕士研究生.通讯地址同上。

侯汉清 南京农业大学信息科学技术学院教授,博士生导师.通讯地址同上。

(来稿时间:2007-04-10)

(上接第59页)识到数字图像信息资源的各种优点,数字图像信息资源的用户接纳程度比较高;数字图像信息资源已经成为了人们学习、工作、生活中重要的信息资源,用户获取数字图像信息资源的途径、方式呈现多样化的特征,用户普遍具备了一定的数字图像信息处理能力。但由于数字图像信息资源本身的特点比较复杂,为了更好地促进数字图像信息资源的应用,不仅需要对应应用图像信息处理技术的领域进行研究,还需要提高相关知识的普及率,加强数字图像信息资源开发与管理研究。

参考文献

- 1 朱学芳.计算机图像信息资源管理研究.现代图书情报技术,2004(12)
- 2 中国互联网络信息中心.中国互联网络发展状况统计报告(2006年1月).[2007-04-10].<http://www.cnnic.net.cn/imagee/2006/download/2006011701.pdf>
- 3 中国互联网络信息中心.中国互联网络发展状况统计报告(2007年1月).[2006-04-10].<http://www.cnnic.net.cn/index/0E/00/11/index.htm>

[net.cn/index/0E/00/11/index.htm](http://www.cnnic.net.cn/index/0E/00/11/index.htm)

- 4 Shannon Williams. Information Literacy Instruction for Educators: Professional Knowledge for an Information Age. Library Journal. 2004, 129(20): 174
- 5 [美]安德鲁·利,罗伯特·阿特金森.冷眼看数字鸿沟.见:吴士余,曹荣湘.解读数字鸿沟.上海:上海三联书店,2003:69
- 6 朱学芳,智文广.计算机图像处理导论.北京:科学技术文献出版社,2003:1
- 7 心理学名人词典:R·利克特.[2007-04-10].<http://www.whpsy.com/person/1/Likert.R>

朱学芳 南京大学信息管理系教授,博士生导师.通讯地址:南京大学.邮编210093。

袁顺波 嘉兴学院教师.通讯地址:浙江嘉兴.邮编314001。

徐 强 南京大学信息管理系博士生.通讯地址:南京大学.邮编210093。

(来稿时间:2007-05-23)