

应重视基于 P-W 的网络文献的引用和著录

■王 冰

收稿日期: 2013-03-21

修回日期: 2013-06-28

广东省医学学术交流中心(广东省医学情报研究所《广东医学》编辑部 510180 广州市越秀区惠福西路进步里2号之6 ,E-mail: wangbing71@126.com

摘要 互联网的普及与发展,学术研究查询收集资料更多的是通过网络进行。引用了网络文献在文后的参考文献表中要有著录,以体现网络信息真实的学术价值。目前网络文献的著录存在《国标》规定不详细,缺乏具体操作指引,使作者、编者、数据库工作者提取著录信息问题较多。为规范网络引文著录,建议:(1)信息服务商完善网页的标准化以使得著录数据容易获取;(2)加强《国标》网络引文著录的解读与研究,并形成规范以使作者及编辑著录时有据可依;(3)网络引文的可获得性及稳定性,不能仅仅依赖网络信息的长期保存,应充分利用编辑部或图书馆资源,建立一套被引网络信息的长期保存系统。

关键词 网络文献 网络引文 著录 规范 国家标准

随着互联网的普及和发展,我们进入了全新的数字时代,在阅读方式上,数字阅读率(包括电脑、手机及其他阅读器等媒介)大幅提升,电子图书将成为人们未来阅读方式的主流^[1]。电子技术的发展及普及利用,不仅仅使科技期刊实现了投稿、审稿、编辑加工、发排出版等编辑出版流程的网络化,更重要的是医学科研工作者已经从阅读纸质的期刊获取信息过渡到更多的通过网络平台收集检索医学文献。OCLC在2005年的一项对包括美国、加拿大、澳大利亚在内的6个国家3300个用户的调查发现,84%的用户使用在线搜索引擎查找信息^[2]。医学论文中网络文献的引用与著录已经成为我们期刊工作者必须面对的工作。现对纸质期刊中网络文献(Print-Web, P-W)的引用、著录探讨如下。

1 网络文献的定义及外延

对于网络文献不同的学者有不同的划分。郭伟等^[3]主张将第一载体为纸质的图书、报纸、期刊等从网络文献中分离出来。刘佳^[4]指出:网络文献不是指一切的网络信息,而是指相对稳定、能够在一定时间内反复访问的信息单元,包括网络报刊、网络数据库、网络图书数据集合,也包括HTTP协议传递的网页,用专用通讯线路访问的数据库等。我国新闻出版署1996年颁发的《电子出版物管理规定》指出“电子出版物是指将图、文、声、像等信息储

存在磁光电解质上,通过计算机或具有类似功能的设备阅读使用,用以表达思想、普及知识和积累文化,并可复制发行的大众传播媒体”。电子文献一是纸质文献以电子文献形式出现;二是直接就是电子文献,三是与纸质文献并不完全相同的网络版形式。目前何为网络文献还没有统一的定义。在互联网上,我们可以通过各大数据库如CNKI、维普中文、万方数据库检索到纸质医学期刊的电子版,广义来讲其第一载体是纸质的,但其最普遍的被获取方式是通过互联网,载体是电子设备。电子文献中引用纸质文献或网络文献等均不在本文中探讨,本文所讨论的网络文献特指第一载体非纸质的文献。讨论纸质期刊中网络文献(P-W)的引用情况。

2 适合作为引文的网络文献

学术论文中的参考文献是作者对所引文献的认可,是学术的延续与传承。网络文献作为引文是对其的学术认可。网络引文分析可作为评价网络信息及网站的条件之一。

网络文献大而繁杂,良莠不齐,形式多样,能作为科技论文引文出现的,多为各专科学会、协会、政府组织等专业的网站。《中华内科杂志》就设立了“网络信息导航”栏目,系统介绍专业网站,为作者阅读与引用网络文献提供了指引。专业性的学术网站的信息,如专业动态、专家学者、重要文献、统计

数据、法律法规、学术会议等这些学术信息,都经过了充分论证和推理的结论,有的是专科学会的专业指南,有的是政府部门的法律法规条文,都适合作为引文。政府机构的网站发布的调查报告或者公告也同样具有权威性。对于不明网页或者网页无法重现的文献因其真实性有疑问,在学术论文中尽量不要引用。对于专业论坛、BBS发表的论文,观点新争议大,学术论文引用前要对其数据进行重复及真实性检验,引用即意味着认可,即使是反证也应谨慎。目前在物理和计算机领域网络已经成为获取信息、科学发现和与同行保持联系的首选^[4]。甚至学术交流模式也从以期刊为中心向以适合网络的新模式发展。2006年菲尔茨奖得主之一——俄罗斯数学家佩雷尔曼,就因在互联网上发表3篇“庞加莱”猜想的关键论文而获奖。

网络引文应具有可回溯性。网络文献经过一段时间再次查找很多无法找到原文。这也是影响网络文献作为引文的关键问题。网络信息频繁变化的URL地址,产生了链接衰减或消失,促使产生DOI及URN及基于此出现的DOI-X、PURL、Handle。DOI及URN能解决由于URL变动而产生的“linkrot”,但对删除造成的“linkrot”,无能为力。李孟臣等^[5]通过研究2000~2003年网络引文的引文有效率、年度衰减指数、有效性半衰期对Web资源是否合适引用进行探讨,认为虽然目前互联网信息引用数目越来越多,但从治学严谨和论文价值实现等多方面考虑,只有通过其他辅助机制保证Web引文的半衰期远大于论文的引用半衰期后,网络文献才更合适大规模引用。段宇锋等^[6]通过对《中国图书馆学报》1995~2008年所载网络引文为对象发现“linkrot”的发生率和与时间有关。

我们建议作者在引用网络文献的同时,应当保留引用时的网络信息及内容,并提供给编辑部,以便于进行系统的保存。

3 网络引文发展趋势

电子文献的优势是传播速度快,范围广,信息存储容量大,具有通用性和易于复制性,检索方便,便于开展研究。与传统文献比较其具有方便修改、巨量增长、批量检索方便、更新快等。在Web2.0时代,用户产生内容(UGC),个人除了是互联网的使用者,也是网络内容的主动传播者和作者,如目前的博客、BBS等,网络信息更容易发布。而传统文献

一经发表,其内容不会再发生改变,而互联网信息其内容及网址均可发生变化。目前网络用户激增,截至2012年12月底,我国网民规模达到5.64亿,全年共计新增网民5090万人。互联网普及率为42.1%,较2011年底提升3.8%。微博用户规模为3.09亿,较2011年底增长了5873万,网民中的微博用户比例达到54.7%。手机微博用户规模2.02亿,占所有微博用户的65.6%,接近总体人数三分之一^[7]。网站数268万,国际出口宽带数,1899792Mbps,IPv43.31亿,域名数1341万^[8]。2011年,我国18~70周岁国民上网率为54.9%,比2010年的49.9%增加了5个百分点。有超过半数(50.6%)的国民通过电脑上网,有25.4%的国民通过手机上网。手机上网率的增幅明显,与2010年的16.6%相比,增幅高达53.0%^[9]。网络信息的发布者、使用者都以巨量的速度增长,图书馆越来越重视电子图书馆及数据库的建设。同时网络中如:CNKI、维普中文、万方数据库等大型的专业数据库、专业性的学术网站也逐渐的建立并完善。网络文献或其他电子资源被科技工作者越来越多的引用。薛卫双等^[10]研究指出,《图书情报工作》近10年的网络引文量在总体上呈现递增的趋势,由2000年的125篇上升到2009年的1507篇,增幅达12倍。科技期刊论文中的网络引文已经成为不可逆转的趋势。

4 网络文献的引用和著录

4.1 网络文献引用的问题

网络文献引用最大的问题是论文引用了网络文献却没有在文后参考文献表中著录。医学论文中出现URL地址的不少。如内文描述“根据周期素D2序列及Target Scan(<http://www.targetscan.org/>)软件预测结果”“根据miRbase(<http://www.mirbase.org/>)得到Rno-miR-145成熟体序列”。“获得TBX1基金外显子序列(<http://www.ensembl.org/>)”等。很多基础研究类论文都涉及到DNA数据库如美国的GenBank、欧洲的EMBL和日本的DDBJ,但均未将上述数据库列为参考文献。有的作者对于网络公告数据直接给出而不给予出处,如WHO网站,或政府网站公告的数据等;有的作者将专业网站中获得的行业标准,以纸质的版本简单著录等等。在编辑加工的过程中我们让作者补充参考文献时,作者能给编辑部的是专业网站上的信

息。网络引文是网络资源应用于学术研究的直接表现,是专业技术人员对网上学术资源的一种变相的同行评议,是对网络资源学术价值的认可。因此论文中引用了网络文献的信息,应当也必须在参考文献表中著录。

4.2 网络引文著录信息的获取

4.2.1 作者自己抽取数据著录

2005年10月1日正式执行的中华人民共和国国家标准 GB7714-05《文后参考文献著录规则》规定电子文献可作为引文并给出著录方向,但偏于简单。具体的著录中存在很多的问题不能靠《规则》解决。如纸质文献的电子文献按电子文献著录还是按纸质文献著录?如何著录印刷前的在线论文?(其在印刷时可能内容及格式均发生变化),如何抽取著录信息?国内缺乏指引。APA(美国心理协会)规定 URL 必须直接指向具体的引用处,URL 的书写,空格、转行、斜线都会使地址产生错误,必须准确的书写印刷。对于网络中无期刊性文件的引用,无确定作者、无时间信息,某部门、政府、组织网站上公布的官方非官方文件、技术报告,可采用文献名作为开头,也可在 URL 前著录其组织名称。这就解决了我们关于有些网络文献责任者的认定。《中国学术期刊(光盘版)检索与评价数据规范》也有相关的叙述。马爱芳^[11]调查发现,很多的中文科技核心期刊的网络引文采用《规范》来进行网络引文的著录,而其与2005年《国标》不同的是日期著录规则,其著录在地址后,而《国标》著录在地址前。《规范》的著录日期存在问题,发表日期,更新日期,引用日期,标识不明。《国标》《规范》的统一使用,辅以外国的各种电子文献的著录规范就可以解决我们期刊编辑目前著录存在的认识问题。期刊编辑所要做的是广泛的介绍著录规范,详细的制定执行的具体细则,同时向作者进行推广,提升作者提取网络文献引用数据元素的能力。国内期刊界应加强《国标》有关网络引文著录的解读与研究,并形成行业规范以使作者及编辑部著录时有据可依。

4.2.2 自动获取各著录数据元素

ALA(美国图书馆协会),ACM(美国计算机协会)以及一些数据库商提供固定的格式方便用户对电子文献的引用。

(1) 网页中电子文献的自动引用,ALA 在它网页的左下部设置了“cite this page”按钮,提供 ALA 格式的固定引用格式,通过复制黏贴即可列在参考

文献表中。

(2) 数据库中电子文献的自动引用。

(3) 提出利用语义 WEB 元数据实现电子文献引用的自动获取,网络文献不像纸质文献一样具有规范的格式,有时从网页上根本无法获得著录的相关信息。第一步分析电子文献,在 HTML 语言中每个元标签都是以 < meta > 开始,方便程序查找元标签起始处,所有的元数据都在 < head > 处当程序遇到 < /head > 或 < body > 就可结束。电子文献需要自动抽取的主要 DC 元数据有:文献作者(DC: creator),文献名称(DC: title),文献类型(DC: Format),文献编写日期(DC: date)和网络电子文献的 URL 地址(DC: Identifier)等几项。元数据的规范化,电子文献引用的标准化,越来越多的元数据标签应用于电子资源中,信息服务商完善网页的标准化以使得著录数据容易获取,一定会加速对电子资源书目信息自动引用的进程。

4.3 网络引文著录中存在的问题

一条著录规范的网络引文应包括网络资源的著者、题名、日期、来源四个字段。来源字段区别与普通引文著录,其 URL 包含丰富的信息,从中可以发现网络资源对应的网络协议、域名、服务器文件夹名称、端口号、文件类型等有用的信息,著录中存在的问题:

- (1) 题目的载体类型及电子标识符不规范,如只著录了文献类型 EB,没有著录载体类型 OL。
- (2) 责任者项不明确。
- (3) 不著录责任者或题目项,直接给 URL 地址。
- (4) 网址著录没有到具体到文件。
- (5) 没有给出引用日期。
- (6) 对网络文献的文献类型判断失误等。

综上所述,网络文献作为引文在学术期刊中出现并逐年增多是不争的事实,其著录存在许多的问题,规范著录我们建议:

- (1) 需要信息服务商完善网页的标准化以使得著录数据容易获取;
- (2) 需要同时加强《国标》中网络引文著录的解读与研究,并形成规范以使编辑部著录时有据可依;
- (3) 网络引文的可获得及稳定性,不能依赖网络信息的长期保存,应充分利用编辑部或图书馆资源,建立一套被引网络信息的长期保存系统才是有效的促进网络信息利用的方法。

(下转第 231 页)

献计量学研究,并发挥网络上的链接功效等等^[9]。参考文献是否规范,不仅是作者严谨学术精神的体现,而且是评价论文学术水平的一个重要依据。准确、适宜的参考文献的引用至关重要,如实地呈现参考文献,不仅表明作者对他人劳动的尊重与承认,也是对他人研究成果实事求是的科学态度。文章仅从三个方面例举了现有环境下电子文献著录存在的不足,结合国际上电子文献的著录情况提出了一些建议。随着信息技术的进步,新媒体涵盖的范围将越来越广,突显电子文献著录格式的不足的地方也将越来越多,由此引发的争议也会越来越多,从而引导人们不断去探索,如此往复,逐渐完善对电子文献的著录。作为学术规范的形成和完善,我们应当尊重客观规律的变化,在一些尚存争议的地方,可以允许大家在实践中摸索探讨,当一种著录格式为作者、编者普遍欢迎和接受时,就说明他的科学性、简洁性和对著作权的保护性体现得最好,任何国标性质的规范条款都应在大家普遍能接受时推出更新为好。

作为学术期刊编辑,应更加重视文后参考文献的规范化标注,本着对读者和作者负责的精神,严格

按照电子文献的实际形式进行审稿、编辑、校对,以维护参考文献的真实性、准确性,并普及国家标准知识,提高作者的引用规范意识,帮助作者养成正确良好的引用习惯,使学术期刊在新媒体环境下发挥更好作用。

参考文献

- 1 马爱芳,王宝英,李瑞芳,吕秀林.有关电子文献著录情况的调查与建议.编辑学报,2007,19(5):269-270
- 2 常平.关于学报参考文献中电子文献著录的探索.武警学院学报,2004,(5):73-74
- 3 全国信息与文献标准化技术委员会第6分委员会.GB/T7714-2005文后参考文献著录规则[S].北京:中国标准出版社,2005:1
- 4 申慧.关于电子文献著录问题的探讨.现代情报,2005,(2):18-20
- 5 赵文青.对我国应用型本科院校发展战略的思考——潘懋元先生访谈录.高校教育管理,2014,(1):6-9
- 6 杨玉香.中国学术期刊有线数字出版的现状与展望.河北医学,2012,(8):1189-1190
- 7 苏振华,杨振和.OA期刊资源发展的现状分析.经济研究导刊,2010,(36):219-222
- 8 李明.纸质文献与电子文献的传播影响因素之比较.河南图书馆学刊,2011,(1):132-134
- 9 方秀菊.对参考文献作用及规范化著录的探讨.浙江科技学院学报,2003,(4):250-253

(上接第221页)

参考文献

- 1 张鹏.图书馆数字化进程中的电子文献发展探讨.河南图书馆学刊,2011,31(5):40-42
- 2 CHEN C,SUN K,WU G et al. The impact of internet resources on scholarly communication: a citation analysis. *Scientometrics*, 2009, 81(2):459-474
- 3 郭伟,陈丽华.网络文献概念的重新定义及范围界定.吉林医药学院学报,2009,30(4):230
- 4 刘佳.网络文献的界定和评价.情报科学,2003,(9):994-997
- 5 李孟臣,支茵,金坤.Web引文可引证性分析.现代图书情报技术,2005,9(总128):70-71
- 6 段宇锋,赵燕芳,刘璇等.对学术文献引用网络信息的反思.图书情报工作,2010,54(8):51-54
- 7 中国互联网络信息中心.CNNIC发布第31次《中国互联网络发展状况统计报告》[R/OL]. [2013-01-25]. http://www.cnnic.net.cn/gywm/xwzx/rdxw/2012nrd/201301/t20130115_38507.htm
- 8 中国互联网络信息中心基础数据 [EB/OL]. [2013-01-25]. <http://www.cnnic.net.cn/hlwfzyj/>
- 9 中国出版网.“第九次全国国民阅读调查”初步成果发布[R/OL]. (2012-04-20) [2013-01-28]. http://www.chuban.cc/tpxw/201204/t20120420_105469.html
- 10 薛卫双,郑春厚,王娟.《图书情报工作》2000~2009年网络引文衰减规律实证研究.图书馆界,2012,(1):56
- 11 马爱芳.网络文献著录项中日期著录的规范化建议.中国科技期刊研究,2006,17(1):148