

## Web 引文数量探析

张翠英 安美荣 王建芳 姜仁珍

(山西大学信息管理系, 太原 030006)

**摘要** 通过对我国图书情报学科的 7 种核心期刊近 5 年来发表论文的参考文献中的网络引文的数量分析, 本文揭示了网络信息资源对学术交流活动的影响和学者们利用网络信息资源的状况。并与国外的网络引文研究进行了比较。

**关键词** 网络引文 网络信息资源 数量分析 学术交流

## Quantitative Exploration of Web Citation

Zhang Cuiying, An Meirong, Wang Jianfang and Jiang Renzhen

(Department of Information Management of Shanxi University, Taiyuan 030006)

**Abstract** Through the quantitative analysis of Web citation of the articles published in the seven core journals of library and information science from 1998 to 2002, the paper indicates how great impact does the Web resource have on scholarly communication and to what extent are scholarly searchers using Web resource in their research. Also, the author compares the current study with foreign country's Web citation research.

**Keywords** Web citation, Web information resources, quantitative analysis, scholarly communication.

### 1 引 言

当今, 伴随着 Internet 的飞速发展和网络信息技术、通信技术的日益更新, 电子化、数字化、网络化已成为一种势不可挡的潮流。网络信息资源以其快速、便捷、高效、内容丰富的优点而得到众多学者的青睐。最近几年, 在国内的许多期刊上都出现了将网络信息资源作为引文的现象(简称为网络引文或 Web Citation, 或 Internet-based Citation, 或 Web-based citation.)。

引文分析是指利用各种数学及统计学的方法和比较, 归纳, 抽象、概括等逻辑方法, 对科学期刊、论

文、著者等各种对象的引用或被引用现象进行分析, 以便揭示其数量特征和内在规律的一种文献计量方法<sup>[1]</sup>。它可以揭示科学文献的相互引证关系。科学文献的引用与被引用, 从文献使用的角度反映了科学与技术之间、学科之间的相互联系。传统的引文往往是引用期刊或书目的文献或载文。

近年来, 国外的学者已经致力于对 Internet 作为一种信息资源的研究, 并且, 他们也关注著者引用网页或网站上的信息来进行学术研究的频率<sup>[2]</sup>。Stephen P. Harter 和 Hak Joon Kim 于 1996 年, 系统地研究了电子资源对于学术交流的影响。他们对发表在电子期刊中的 279 篇论文中的 4317 个引文进行了分析, 结果表明, 只有 1.9% 的参考文献引用了网络

收稿日期: 2003 年 11 月 25 日

作者简介: 张翠英, 女, 1949 年生, 山西大学信息管理系教授, 硕士研究生导师。安美荣, 女, 1975 年生, 山西大学信息管理系 2002 级硕士研究生。王建芳, 女, 1978 年生, 山西大学信息管理系 2001 级硕士研究生。姜仁珍, 女, 1971 年生, 山西大学信息管理系 2002 级硕士研究生。山东莱阳农学院图书馆馆员。

信息资源<sup>[3]</sup>。此后, Yin Zhang 分析了 10 种图书馆学和情报学的杂志, 发现, 从 1994 年到 1996 年, 网络引文只占总引文的 1.13%。在随后的研究中, Yin Zhang 发现, 从 1991—1998 年, 印刷型期刊引用网络信息资源的比例从 0.2% 增到 5.2%, 而且, 含网络引文的文章数的比例从 1.8% 上升到 33.9%<sup>[4]</sup>。2002 年, Mary F. Casserly 比较系统地、全面地从网络引文的可获得性进行了研究<sup>[5]</sup>。2002 年, Susan Davis Herring 通过引文分析研究了在学术电子期刊中利用网络资源的状况<sup>[6]</sup>。这些研究表明, 国外的许多学者已经在深入探讨学者们利用网络信息资源的状况, 并卓有成效。

和国外的网络引文研究相比较, 我国的学者对这方面的研究起步较晚。迄今, 笔者尚未见到有关这方面的比较系统的研究。借鉴国外的研究经验与方法, 本文通过对印刷型期刊学术论文的参考文献中的网络引文的数量进行分析, 以探讨学者们对网络资源的利用状况及网络资源对学者的学术交流及科研的影响程度。毕竟, 正如 Stephen P. Harter 所认为的, 在评价学术论著的影响时, 引文分析仍是一种很有效的工具<sup>[7]</sup>。

## 2 数据源的采集

本文以我国图书馆学、情报学的 7 种核心期刊《中国图书馆学报》、《大学图书馆学报》、《图书馆论坛》、《情报理论与实践》、《情报资料工作》、《情报学报》、《图书情报工作》作为研究样本, 对其近 5 年间 (1998—2002 年) 所发表的学术论文的参考文献和网络引文 (以下简称 Web 引文) 进行了统计, 归纳、分

析, 共统计了 5852 篇文章, 37010 篇引文, 其中, 含 Web 引文的文章数为 918 篇, Web 引文为 3557 篇。在统计过程中, 由于人为因素的影响, 统计数据也许会有些小偏差。另外, 在统计过程中, 期刊中的新闻通告、快报、贺词等不包括在内。由于篇幅有限, 本文只从 Web 引文的数量这个角度进行分析。

## 3 Web 引文的数量分析

### 3.1 Web 引文量的分析

引文总量以 C 表示, Web 引文量以 W 表示。W/C 表示 Web 引文量占引文总量的比例。表 1 以各刊总计中 W/C 的比值大小排序。

从表 1 可以看出, 从 1998—2002 年, 各刊的 Web 引文量在逐年增加。而且, 我们也可以得出一个很乐观的预测, 即, 今后, Web 引文量仍有继续增长的趋势。对期刊而言, 各刊 Web 引文量增长的幅度不同, 《大学图书馆学报》、《情报理论与实践》、《图书情报工作》3 种期刊的 W/C 的比值较大, 在这 5 年内所引用的 Web 引文最多。《图书馆论坛》所引的 Web 引文最少。从总体来看, W/C 的平均比值偏低, 只有 9.6%。也就是说, 在所引用的 100 个引文中, 只有不到 10 个 Web 引文。

这从侧面反映出, 比起传统的印刷型文献, 网络信息资源对学术交流活动的影 响仍很小。这与国外学者对网络引文所作研究的结果相同。究其原因, 笔者认为, 这与网络信息资源自身的特点有关: (1) 网上所发布的文献不需经过严格的筛选、审查, 具有很大的自由性和随意性, 因此良莠不齐, 可靠性较

表 1 5 年间各期刊的引文总量与 Web 引文量统计表

时 间 刊 名	1998	1999	2000	2001	2002	总 计	
	[W]C	[W]C	[W]C	[W]C	[W]C	[W]C	W/C%
大学图书馆学报	[10]909	[37]695	[134]824	[183]964	[195]866	[559]4258	13.1
情报理论与实践	[15]634	[27]820	[159]979	[194]956	[157]1033	[552]4422	12.5
图书情报工作	[33]1441	[85]1493	[210]2512	[347]2210	[512]3023	[1187]10679	11.1
情报学报	[20]435	[38]600	[96]771	[102]878	[181]1302	[437]3986	11.0
中国图书馆学报	[18]1051	[72]968	[80]1207	[114]1230	[205]1322	[489]5778	8.5
情报资料工作	[1]418	[13]425	[19]427	[68]1039	[111]1070	[212]3379	6.3
图书馆论坛	[0]477	[9]606	[11]794	[22]1032	[79]1599	[121]4508	2.7
合计	[97]5365	[281]5610	[709]7514	[1030]8309	[1440]10215	[3557]37010	9.6
W/C (%)	1.8	5.0	9.4	12.4	14.1	9.6	

表 2 各期刊 5 年间的总发文量与含 Web 引文的文章数的统计表

时 间 刊 名	1998	1999	2000	2001	2002	总 计	
	[H]P	[H]P	[H]P	[H]P	[H]P	[H]P	[H]P/%
情报学报	[4]79	[13]87	[27]105	[42]112	[47]133	[133]516	25.8
情报理论与实践	[5]134	[10]151	[41]157	[43]161	[54]159	[153]762	20.1
图书情报工作	[6]242	[16]231	[53]310	[83]317	[119]335	[277]1435	19.3
大学图书馆学报	[5]203	[13]143	[26]120	[40]140	[53]140	[137]746	18.4
中国图书馆学报	[6]112	[10]114	[23]135	[32]144	[26]141	[97]646	15.0
情报资料工作	[1]82	[4]94	[10]96	[19]162	[43]162	[77]596	12.9
图书馆论坛	[0]185	[4]225	[5]201	[12]242	[23]298	[44]1151	3.8
合计	[27]1037	[70]1045	[185]1124	[271]1278	[365]1368	[918]5852	15.7
H/P(%)	2.6	6.7	16.5	21.2	26.7	15.7	

差。董小英在其一项调查中显示,当问及访问中英文网站的差距时,59.3%的用户认为我国网站提供的信息质量不高<sup>[8]</sup>; (2)在利用搜索引擎检索所需信息时,往往会检到一些与所研究的课题无关的信息,再进行二次检索过滤时,需花费大量的时间与精力,还有资金; (3)网络信息资源的权威性较差。

### 3.2 含 Web 引文文章的分析

表 2 中,以 P 表示年总发文量,以 H 表示含 Web 引文的文章数。H/P 表示含 Web 引文的文章数在发文量中所占的比例。以各刊总计中 H/P 的比值大小排序。

分析含 Web 引文文章数的多少,可以直观地反映出著者利用网络信息资源的意识的高低。从表 2 可以看出,5 年间,含 Web 引文的文章数也在逐年增加,在总发文量中所占的比例也在逐年增加,从 1998 年的 2.6% 增加到 2002 年 26.7%,增加了将近 10 倍。与前面所谈到的 Web 引文量的增长趋势相一致。这表明著者利用网络信息资源的意识正在逐步提高。但从总的来看,在所统计的 5852 篇文章中,有 918 篇文章引用了 Web 引文,所占比例为 15.7%。而国外的 Yin Zhang 于 1998 年的研究得

出,含 Web 引文的文章数所占的比例为 33.9<sup>[9]</sup>。Susan Davis Herring 在 2002 年的一项网络引文研究中,共统计了 175 篇文章,发现有 97 篇文章引用了网络引文,其比例为 55.4%<sup>[10]</sup>。与国外的研究结果相比,我国的学者利用网络信息资源的意识仍很低。一方面,由于我国计算机网络技术的发展仍然滞后于国外的发展;另一方面,各学者掌握计算机技术及网络技术的程度不一,从而对网络信息资源的利用也就不一。

就期刊而言,《情报学报》、《情报理论与实践》、《图书情报工作》、《大学图书馆学报》这 4 种期刊中含 Web 引文的文章在总发文中所占的比例较高。我们也可以看出,在情报学类的期刊上发表文章的著者比在图书馆学类的期刊上发表文章的著者更易使用网络信息资源。由于情报学类期刊上刊登与计算机或其他相关技术的文章较多,著者在网上易查找到最新的国内及国外信息,因而著者对网络信息资源的利用也就较高。

### 3.3 Web 引文率的分析

期刊引文率 = 该刊中参考文献数/载文量。参照此定义,期刊网络引文率 = 该刊中 Web 引文数/

表 3 5 年间各刊的发文量与 Web 引文量的统计表

	1998		1999		2000		2001		2002		总 计	
	发文量	Web 引文量	发文量	Web 引文量	发文量	Web 引文量	发文量	Web 引文量	发文量	Web 引文量	发文量	Web 引文量
共计	1037	97	1045	281	1124	709	1278	1030	1368	1440	5852	3557
Web 引文率	0.09		0.27		0.63		0.81		1.05		0.61	

载文量。表 3 的数据表明,在这 5 年间,Web 引文率有明显的提高,从 1998 年的 0.09% 增至 2002 年的 1.05%。平均 Web 引文率为 0.61%。也就是说,平均每篇文章引用还不到一篇网络文献。而 Mary F. Casserly 在 2003 年的研究中,发现每篇文章平均引用 2.5 篇网络文献<sup>[11]</sup>。与国外相比,我国图书情报学期刊的网络引文率较低。这从另一个方面表明,我国学者对网络信息资源的利用较低。

表 4 网络引文的频次分布表

频次	含 Web 引文章数	频次	含 Web 引文章数	频次	含 Web 引文章数
1	313	11	3	23	1
2	168	12	14	25	3
3	120	13	8	27	2
4	74	14	6	30	1
5	59	15	2	33	1
6	44	16	1	37	1
7	27	17	5	87	1
8	25	18	3	0	4934
9	19	20	2	合计	5852
10	12	21	3		

3.4 Web 引文的频次分布

对网络资源的利用因人而异。不同的学者,出于其对网络资源的利用能力、认可程度以及其所研究领域性质的不同,都将决定他们将网络资源作为引文的不同。统计数据表明,在所研究的 5852 篇文章中,有 4934 篇文章没有网络引文,其余的 918 篇文章的参考文献中分布了 5852 篇网络引文,其分布如图 1 所示。

由图 1 可看出,Web 引文在论文中的分布呈集中分散趋势。极个别的论文集中了大量的网络文献,而一些网络引文分散在大量的文献中。

注:这里的频次是指每篇文章所引用的 Web 引文的个数。如频次为 2,含 Web 引文文章数为 168,就是指有 168 篇文章每篇文章所引用的 Web 引文数为 2。

3.5 Web 引文的来源分析

就传统的印刷型引文而言,引文来源是指引文

所在的期刊来源。对 Web 引文来源而言,就是指 Web 引文的来源地址及其所属的类型。这里,笔者按其所属的域名类型分为 7 大类。具体的分布见表 5。

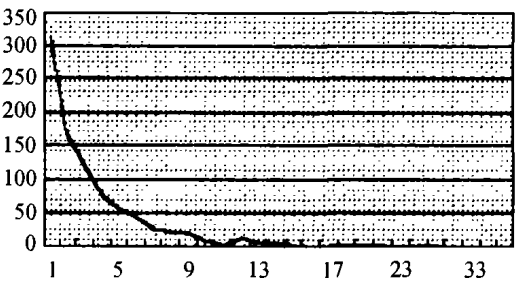


图 1 含 Web 引文文章数的分布图

表 5 Web 引文所属的域名类型及数量表

域名类型	Web 引文数	所占比例 %
Com = commercial	992	27.9
Edu = education	751	21.1
Org = organization	717	20.2
Gov = government	299	8.4
Ac = academic	220	6.2
Net = network	166	4.7
Other	412	11.6
共计	3557	100

从表 5 中可以看出,在所统计的 3557 个 Web 引文中,从域名为 com 的网站中引用的 Web 文献有 992 个,所占的比例最高,达到 27.9%。域名为 edu 的次之,所占比例为 21.1%。域名为 org 的随后,所占比例为 20.2%。从域名为 gov、ac、net 这 3 类网站中所引用的 Web 文献较少。由此看来,商业机构、教育部门和一些组织机构所提供的网络信息资源更易为图书情报学科的学者们所用,在学术交流活动中有较大的影响作用。

4 结束语

以上对 Web 引文的统计分析表明,虽然 Internet 的发展极大地提高了学者们发现和利用网络信息资源的能力,但总体而言,他们在进行学术交流活动时,对网络信息资源的利用仍较低。面对这种情况,笔者认为,今后如何提高网络信息资源的质量、如何提高检索的准确性或如何提高搜索引擎的专指度从

而提高网络信息资源的利用率是大家需关注的问题。

本文只从图书情报学的角度选取了研究样本,因此所得出的结论未免有“管窥之见”之嫌,但鉴于数据的数量,我们仍能从中略见端倪。目前,网络引文研究还是一个有待探究的领域,需要更多的、对之有兴趣的同仁来共同探讨。通过研究学者们如何利用网络信息资源,图书馆员和其他从事信息工作的人员就能更好地为他们的用户提供高质量的服务。

### 参 考 文 献

- 1 邱均平编著.文献计量学.科学技术文献出版社,1988
- 2 Mary F. Casserly and James E. Bird. Web Citation Availability: Analysis and Implication for Scholarship. *College & Research Libraries*, 2003(4): 300 ~ 317
- 3 同[2]
- 4 同[2]
- 5 同[2]
- 6 Susan Davis Herrning. Use of Electronic Resources in Scholarly Electronic Journals: A Citation Analysis, *College & Research Libraries*, 2002(4): 334 ~ 340
- 7 Stephen P. Harter, Hak Joon Kim. The Impact of Electronic Journals on Scholarly Communication: A Citation Analysis. In: *The Public-Access Computer Systems Review* 7, 1996 (5). Available online from <http://info.lib.uh.edu/pr/v7/n5/hart7n5.html>
- 8 董小英等.中国学术用户对互联网信息的利用及其评价. *图书情报工作*, 2002(10): 29 ~ 38
- 9 同[2]
- 10 同[6]
- 11 同[2]
- 12 Yin Zhang. The impact of Internet-based Electronic Resource on Formal Scholarly Communication in the Area of Library and Information Science: a Citation Analysis. *Journal of Information Science*, 1998(4): 241 ~ 254
- 13 Yin Zhang. Scholarly use of Internet-based electronic resources: A survey report. *Library Trends*, 1999(4): 746 ~ 770
- 14 王恺荣.刍论网络引文文献对评价核心期刊的影响. *情报科学*, 2003(3): 269 ~ 271

(责任编辑 芮国章)