

# OA 网站与商业网站论文高下载量对比分析及其与被引频次的相关性研究

## ——以《中国科技期刊研究》为例

方红玲<sup>1,2)</sup>

收稿日期:2013-05-15

修回日期:2013-06-03

1)河南省科技期刊研究中心,453003 河南省新乡市新乡医学院,E-mail:fanghongling@xxmu.edu.cn

2)新乡医学院期刊社《眼科新进展》编辑部,453003 河南省新乡市

**摘要** 以《中国科技期刊研究》OA 网站的下载量和商业网站的下载量、被引频次为数据源进行研究,探讨我国期刊 OA 网站与商业网站论文高下载量的差别及其与被引频次的相关性,结果显示:(1)《中国科技期刊研究》2007~2011 年 30 期论文的 OA 网站的整期下载量有 29 期大于论文篇均下载量,只有 2007 年第 5 期论文的 OA 网站整期下载量小于论文篇均下载量。(2)《中国科技期刊研究》2007~2011 年 OA 网站的下载量( $622.88 \pm 568.03$ )和高下载量论文的下载量( $1001.95 \pm 1173.88$ )都远高于商业网站对应下载量( $32.89 \pm 30.67$ )、( $78.20 \pm 40.08$ ) (均为  $P < 0.01$ )。(3)①《中国科技期刊研究》2007~2011 年各年度 OA 网站论文高下载量与被引频次不相关,2007~2009 年各年度 OA 网站论文高下载量与商业网站的下载量相关;②《中国科技期刊研究》2007~2010 年各年度商业网站论文高下载量与被引频次不相关,只有 2011 年的商业网站论文高下载量与被引频次相关;与 OA 网站 2007~2010 年论文的下载量相关;③《中国科技期刊研究》2007~2011 年各年度发表论文的被引频次相对于 OA 网站来说,与商业网站论文的下载量相关性较强。

**关键词** OA 网站 商业网站 高下载量论文 被引频次 相关性分析

开放存取(open access,OA)作为一种基于网络的出版模式,在我国的发展已有近 10 年的历史。OA 出版跨越了原始文献资源费用障碍,改变了传统学术交流的方式,加快了信息的传播速度,扩大了读者群。对于 OA 出版的形式<sup>[1-2]</sup>,编辑部常采用自建独立网站(简称 OA 网站)的方式,期刊通过 OA 网站实现 OA 出版对于加快期刊学术论文的下载、阅读和引用,以及提高期刊的显示度和引用率来说都具有积极的促进作用。但由于编辑部 OA 网站仅提供单一期刊文献的检索,对于想查阅大量相关学科文献的读者来说十分不便。

因此,很多期刊通过中国知网、万方数据资源系统和维普资讯的子网系统(中国学术期刊网、万方数字化期刊群和中文期刊全文数据库)提供印刷期刊电子版的网络化服务,但没有实现开放获取,读者需要付费才能浏览全文。我们把中国学术期刊网、万方数字化期刊群和中文期刊全文数据库称之为商业网站。对比 OA 网站和商业网站的论文下载量,可以了解国内读者获取文献的主要途径。

因此,我们以《中国科技期刊研究》OA 网站的下载量和商业网站的下载量、被引频次为数据源进行研究,进而探讨我国期刊 OA 网站与商业网站论文高下载量的区别及其与

被引频次的相关性。

## 1 资料与方法

### 1.1 资料来源

登陆中国知网(www.cnki.net),收集《中国科技期刊研究》2007~2011 年所有论文在商业网站的下载量、被引频次,登陆《中国科技期刊研究》网站(<http://cjstp.ijournals.cn/cjstp/ch/index.aspx>)收集《中国科技期刊研究》2007~2011 年所有论文在 OA 网站的下载量。依据文献的二八规律(帕累托法则)确定高下载量论文<sup>[3]</sup>:将论文按下载量降序排列,前 20% 的论文被认定为高下载量论文,和最后一篇高下载量论文有相同下载量的论文均被确定为高下载量论文。高被引论文的确定同以上方法。

本研究所有数据收集日期:2012 年 12 月 7 日—12 日。

### 1.2 统计学方法

将《中国科技期刊研究》2007~2011 年所有论文在 OA 网站的下载量和商业网站的下载量、被引频次输入 Excel 表格,比较两种网站下载量的差异,探讨被引频次与 OA 网站和商业网站下载量的相关性。

## 2 结果

### 2.1 《中国科技期刊研究》2007~2011年OA网站整期下载量与论文篇均下载量比较

《中国科技期刊研究》2007~2011年OA网站整期下载量与论文篇均下载量比较见表1。

由表1可以看出,《中国科技期刊研究》2007~2011年30期论文的OA网站整期下载量有29期大于论文篇均下载量,只有2007年第5期整期下载量小于论文篇均下载量。这说明由于文献的免费获取,读者下载论文时一般更愿意选择打包下载整期论文的资料保存备用。

### 2.2 《中国科技期刊研究》2007~2011年OA网站与商业网站下载量比较

《中国科技期刊研究》2007~2011年OA网站与商业网站下载量比较见表2。

由表2可以看出,《中国科技期刊研究》2007~2011年OA网站的下载量为 $622.88 \pm 568.03$ ,高下载量论文下载量为 $1001.95 \pm 1173.88$ ,都远高于商业网站对应下载量 $32.89 \pm 30.67$ 和 $78.20 \pm 40.08$ (均为 $P < 0.01$ )。这说明国内学者查阅文献时一般首选OA网站下载论文,只有少部分读者通过商业网站付费下载论文。

### 2.3 《中国科技期刊研究》2007~2011年论文的被引频次与两种网站下载量的关系

《中国科技期刊研究》2007~2011年全部论文的篇均被引频次和高下载量论文的篇均被引频次见表3。

由表3可以看出,高下载量论文篇均被引频次远高于所有论文篇均被引频次。为了进一步探讨我国期刊论文被引频次与OA网站、商业网站高下载量论文相关性,将《中国科技期刊研究》2007~2011年发表论文高下载量与被引频次的相关性进行分析,结果见表4。

表1 《中国科技期刊研究》2007~2011年OA网站整期下载量与论文篇均下载量比较

期次	2007年			2008年			2009年			2010年			2011年		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
1	1093	522.1	0	534	500.4	14	839	673.0	3	2085	813.1	0	800	555.4	1
2	823	572.5	1	703	494.2	1	1228	715.9	2	1532	1461.9	1	762	451.4	1
3	1109	634.7	4	859	520.8	0	1115	892.8	2	1169	824.4	4	786	451.5	0
4	973	576.4	2	926	551.0	0	1799	939.5	2	1059	679.6	0	773	412.8	0
5	656	731.6	39	871	586.2	2	1605	810.7	2	3470	481.3	0	552	394.4	2
6	809	527.8	0	854	607.7	4	1883	767.4	1	870	510.8	0	1645	255.5	0

注:a:该期论文整期下载量,b:该期论文篇均下载量,c:b>a的论文数

表2 《中国科技期刊研究》2007~2011年OA网站与商业网站下载量比较(次)

年份	OA网站		商业网站	
	平均下载量	高下载量论文平均下载量	平均下载量	高下载量论文平均下载量
2007	$594.33 \pm 178.81$	$836.76 \pm 263.03$	$22.96 \pm 18.52$	$52.04 \pm 21.18$
2008	$544.86 \pm 111.78$	$700.74 \pm 133.58$	$30.17 \pm 26.43$	$68.21 \pm 36.67$
2009	$806.86 \pm 344.11$	$1193.82 \pm 610.51$	$25.77 \pm 26.03$	$67.05 \pm 28.33$
2010	$756.09 \pm 1254.04$	$1360.63 \pm 2721.58$	$32.15 \pm 22.42$	$68.51 \pm 23.46$
2011	$415.34 \pm 124.47$	$588.35 \pm 68.64$	$56.18 \pm 43.33$	$117.40 \pm 61.78$
合计	$622.88 \pm 568.03$	$1001.95 \pm 1173.88$	$32.89 \pm 30.67$	$78.20 \pm 40.08$

表3 《中国科技期刊研究》2007~2011年发表论文的被引频次

	2007年	2008年	2009年	2010年	2011年	均值
所有论文篇均被引频次	$0.83 \pm 1.33$	$1.10 \pm 1.69$	$1.08 \pm 1.93$	$1.66 \pm 2.19$	$1.37 \pm 2.23$	$1.18 \pm 1.90$
高下载量论文篇均被引频次	$2.94 \pm 1.53$	$3.64 \pm 2.18$	$3.95 \pm 2.66$	$5.06 \pm 2.57$	$3.64 \pm 2.81$	$4.00 \pm 2.53$

表4 《中国科技期刊研究》2007~2011年发表论文高下载量与被引频次的相关性分析

年份	OA 网站论文高下载量				商业网站论文高下载量				论文的被引频次			
	与被引频次		与商业网站下载量		与被引频次		与 OA 网站下载量		与 OA 网站下载量		与商业网站下载量	
	r	p	r	p	r	p	r	p	r	p	r	p
2007	-0.013	0.914	0.454	0.000	-0.053	0.672	0.292	0.018	-0.102	0.420	0.184	0.139
2008	0.220	0.097	0.446	0.000	0.159	0.233	0.517	0.010	0.255	0.053	0.276	0.036
2009	-0.206	0.109	0.363	0.004	0.232	0.069	0.528	0.000	0.020	0.877	0.341	0.007
2010	-0.052	0.725	-0.033	0.824	0.162	0.267	0.399	0.005	-0.055	0.706	0.299	0.037
2011	0.251	0.064	0.198	0.147	0.536	0.000	0.250	0.066	0.376	0.005	0.616	0.000
合计	-0.042	0.478	0.064	0.278	0.362	0.000	0.158	0.007	0.154	0.009	0.456	0.000

由表4可以看出:(1)《中国科技期刊研究》2007~2011年各年度OA网站论文高下载量与被引频次不相关,《中国科技期刊研究》2007~2009年各年度OA网站论文高下载量与商业网站同年度的下载量相关。(2)《中国科技期刊研究》2007~2010年各年度商业网站论文高下载量与被引频次不相关,只有2011年商业网站论文高下载量与被引频次相关;与OA网站2007~2010年各年度论文下载量相关;(3)《中

国科技期刊研究》2008~2011年各年度论文的被引频次与商业网站论文的下载量均呈正相关(均为 $P < 0.05$ ),只有2011年论文的被引频次与OA网站当年下载量呈正相关。

#### 2.4 《中国科技期刊研究》2007~2011年论文被引频次≥10的论文

《中国科技期刊研究》2007~2011年论文被引频次≥10的论文见表5。

表5 《中国科技期刊研究》2007~2011年被引频次≥10的论文

作者	题名	被引频次	发表时间	OA 网站下载量	商业网站下载量
汪新红	优先数字出版是提高学术期刊出版速度的一种新模式	18	2011年	782	230
程维红	我国科技期刊由传统出版向数字出版转型的对策建议	13	2011年	583	479
刘飏	国外科技期刊网站的调查与思考	13	2009年	823	104
任胜利	特征因子(Eigenfactor):基于引证网络分析期刊和论文的重要性	13	2009年	921	93
程维红	中国科协所属科技期刊的网络化建设进展	13	2010年	796	48
曾建勋	基于知识链接的科技期刊数字化出版策略	12	2011年	641	169
韩维栋	挖掘高被引论文有利于提高科技期刊的学术影响力	12	2010年	676	90
程维红	国外科技期刊的在线出版——基于对国际性出版商和知名科技社团网络平台的分析	11	2008年	1053	117
杨彩霞	《中国科技期刊研究》1998~2007年论文被引分析	11	2009年	570	50
袁桂清	中国科学技术协会期刊品牌建设与发展对策研究	11	2010年	797	39
张建合	《中国科技期刊研究》高被引论文特征分析	10	2011年	480	154
史春杨	《中国科技期刊研究》2003~2005年载文、作者及引文分析	10	2007年	466	67
程维红	中国科协科技期刊网络化出版现状	10	2008年	514	45
方红玲	2003~2008年眼科学高被引论文统计分析	10	2010年	792	35

由表5可以看出:《中国科技期刊研究》2007~2011年论文已有14篇论文的被引频次≥10次,最高被引频次为18次,发表时间为2011年,其中程维红有4篇论文被引频次≥10次。

### 3 分析与讨论

OA自提出以来,国内外学者已从不同的角度和层面进行了很多系统深入的研究,提出了许多切实可行的方案或模型,解决了许多问题,还有许多尚未解决的问题,这些都为后

续研究OA出版提供了大量丰富的文献信息,也为推动OA运动的发展奠定了一定的基础。目前,国内学者针对我国OA期刊出版模式或OA文献共享使用效果方面的相关研究尚未见报道<sup>[4]</sup>。因此,本研究以《中国科技期刊研究》OA网站的下载量和商业网站的下载量、被引频次为数据源进行研究,探讨我国期刊OA网站与商业网站论文高下载量的差别及其与被引频次的相关性。

国外The Open Citation Project<sup>[5]</sup>汇聚大量学者开展的OA期刊及其论文的影响评价方面研究的数百篇文献资料,

研究内容主要包括开放访问与影响的关系、开放与引用的关系、下载与引用的关系,网络评价工具、引用次数、影响因子、下载次数等,研究的方法有问卷调查、网络评价、数据库统计评价、比较研究等。研究的结论基本都一样,即 OA 可增加期刊发表论文的显示度、引用次数,并扩大影响力。而基于 2002 年 JCR 中被收录的 192 种 OA 期刊,McVeigh<sup>[6]</sup>首次系统研究了 OA 期刊的文献计量学指标,结果表明:OA 期刊具有快速传播和快速引用的优势。基于 2008 年 JCR 数据研究表明,OA 期刊的影响因子位于前 50% 的比例为 38.62%<sup>[7]</sup>,与基于 2002 年 JCR OA 期刊的影响因子相比,影响因子有较大幅度的提高,再次证实了 OA 期刊可以扩大显示度、引用率和影响力。本研究发现,《中国科技期刊研究》2007~2011 年 OA 网站论文的篇均下载量(622.88 ± 568.03)和高下载量论文的篇均下载量(1001.95 ± 1173.88)都远高于商业网站对应下载量 32.89 ± 30.67 和 78.20 ± 40.08(均为  $P < 0.01$ ),说明国内学者收集相关专业文献时一般首选 OA 网站下载论文,只有少部分可能不知道可免费获取文献的读者通过商业网站付费下载论文。这表明,我国期刊的 OA 出版也具有扩大显示度和快速传播文献的优势。

论文的下下载量和论文的被引频次一样,均为评价论学术影响力的重要指标。而从理论上讲,实用性越强、学术水平越高的论文,越容易受到读者的关注,具体表现在论文的频繁被引用和被下载<sup>[8]</sup>。通常情况下,同一篇论文的下下载量和被引频次是一致的。

本研究发现:(1)《中国科技期刊研究》2007~2011 年各年度 OA 网站论文高下载量与被引频次不相关,《中国科技期刊研究》2007~2009 年各年度 OA 网站论文高下载量与商业网站的下下载量相关。

(2)《中国科技期刊研究》2007~2010 年商业网站论文高下载量与被引频次不相关,只有 2011 年商业网站论文高下载量与被引频次相关;与 2007~2010 年各年度 OA 网站论文的下下载量相关。

(3)《中国科技期刊研究》2008~2011 年各年度论文的被引频次与商业网站论文的下下载量均呈正相关,只有 2011 年论文的被引频次与 OA 网站当年下下载量呈正相关。这表明,相对来说我国期刊论文的被引频次与商业网站论文的下下载量相关性较强。这可能是由于 OA 网站的部分高下下载量论文由于作者本身多次下下载导致,或者虽然论文本身学术水

平一般,但实用性较强,对国内同行编辑具有一定的指导意义,因而被大量反复下下载,但论文的被引次数并不高,也即是刘雪立等<sup>[8]</sup>提出的“量引背离现象”。

《中国科技期刊研究》2007~2011 年论文已有 14 篇论文的被引频次 ≥ 10 次,最高被引频次为 18 次,发表时间为 2011 年,其中程维红有 4 篇论文的被引频次 ≥ 10 次。优质的作者队伍是刊物生存和持续快速发展的关键因素。因此,《中国科技期刊研究》应在作者队伍建设方面积极挖掘、选拔和培养,在周围形成一支高产出、高被引的作者团队<sup>[9]</sup>。

**致谢** 衷心感谢刘雪立编审在论文选题方向和论文修改方面的指导和帮助!

#### 参考文献

- 1 程维红,任胜利,王应宽等.中国科协科技期刊开放存取出版现状.中国科技期刊研究,2008,19(4):554-560
- 2 秦小川,刘雪立.我国学术期刊开放存取现状.中国科技期刊研究,2011,22(4):485-489
- 3 刘雪立,方红玲,苗媛等.我国 5 种眼科学核心期刊 2004~2008 年高下载量论文统计与分析.中国科技期刊研究,2010,21(4):459-462
- 4 王应宽,吴卓晶,程维红等.国内外开放存取期刊研究进展综述与发展动态分析.中国科技期刊研究,2012,23(5):715-724
- 5 The Open Citation Project-Reference Linking and Citation Analysis for Open Archives. The effect of open access and downloads ('hits') on citation impact: a bibliography of studies. [2012-03-01]. <http://opcit.eprints.org/oacitation-biblio.html>
- 6 McVeigh ME. (2004). Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns Acitation study from Thomson Scientific, Thomson Scientific. [2012-03-01]. <http://science.thomsonreuters.com/m/pdfs/openaccesscitations2.pdf>
- 7 Giglia E. (2010). The Impact Factor of Open Access journals: data and trends. ELPUB 2010 International Conference on Electronic Publishing, Helsinki (Finland), 16-18 June 2010. [2012-03-01]. <http://dhanen.shh.fi/dspace/bitstream/10227/599/72/2giglia.pdf>
- 8 刘雪立,方红玲,苗媛等.五种综合性眼科学期刊论文下下载量和被引量的关系及部分论文的量引背离现象.中国科技期刊研究,2010,21(5):629-632
- 9 张建合,任长江.《中国科技期刊研究》高被引论文特征分析.中国科技期刊研究,2011,22(2):207-210