

大数据时代 DOI 的应用意义与中国科技期刊应用现状

姚戈 王淑华 王亨君

中国地质大学《地球科学》编辑部,430074,武汉

摘要 数字对象唯一标志符 DOI(digital object identifier),已成为云计算平台上电子文献样本存储和按需服务的最佳应用技术。本文介绍了 DOI 的发展历史与国际现状,分析了大数据时代期刊 DOI 的存在意义,特别是 2009 年 JCR 新增了 eigenfactor score 和 article influence score 2 个评价指标后,DOI 已成为评价指标的唯一依据。对 2013 年 300 种高校期刊进行了分析,发现其 DOI 应用比例严重偏低,仅占 17.6%,原因有 3 点:1) 期刊无独立网站和过刊回溯功能;2) 申请了 DOI,但没有针对期刊的文章信息进行注册或在文章上注明 DOI;3) 期刊网站改版或者网络链接(URL)的变更,没有及时重新注册 DOI 元数据。建议我国期刊建立数字化平台网站或与国内外大型出版集团合作,在拥有 DOI 号的同时,同步进行注册,使其成为期刊数字出版的“利器”,为广大受众服务,使国内与国外数字出版同步发展。

关键词 大数据时代;数字对象标志符;URL;高校期刊;数字化出版

Significance of DOI in the big data era and its current status in Chinese scientific journals // YAO Ge, WANG Shuhua, WANG Hengjun

Abstract Digital object identifier (DOI) has become one of the best technologies of electronic literature sample storage and applications on cloud computing platform. This paper introduces the history and current status of DOI, analyzes the significance of the journal DOI system in the big data era. DOI has become one of the key identifiers in the journal quality evaluation, in particular, two evaluation indicators, namely, eigenfactor score and article influence score have been added in JCR in 2009. The study examines 300 college journals in 2013 and the results show that the applications of DOI in college journals are at a seriously low level. Three reason cause this situation: 1) The journals are lack of independent websites and back issues cannot be back tracked. 2) The journals have applied for DOI, however the DOI has not been registered with the relevant journal's information or been clearly noted in the articles. 3) The websites or URL have been changed without reregistering or updating the DOI metadata. As mentioned above, some suggestions for college journals have been stated: college journals need to develop digital platform such as build websites or cooperate with international publishing groups. The DOI could be registered when journals have successfully applied for DOI, which will benefit the readers and make Chinese college journal keep up with foreign digital publishing.

Keywords big data era ; digital object identifier; uniform resource locator; college journals; digital publishing

Authors' address Editorial Office of Journal of Earth Science, China University of Geosciences, 430074, Wuhan, China

在 1986 年到 2007 年,全球数据以每年 23% 的比率递增。2000 年数字化存储的数据比例为 25%,随着数字技术的普及,到 2007 年数字化数据占绝对优势,达到 94%^[1],IDC 预计 2010—2020 年全球数据量将增加 50 倍,其中 80% 以上为结构化数据。国际出版商已经全面进入数字领域,信息技术被广泛运用到出版工作中的每一个环节,云计算极大提高了出版机构全面整合资源的能力,大数据储备愈加丰厚。国外几大著名出版商,例如培生、爱思唯尔、汤森路透和施普林格等集团都十分重视数字出版,均建立了数字化平台或云计算。为了保证各个数字化平台文献的跨平台云计算,需要一个统一的标准和技术,DOI(digital object identifier)就是专门为这个统一标准而设立的。

DOI 的中文称作“数字对象标志符”^[2],是由数字、字母和符号组合而成的字符串,被用来指向数字目标,例如数字期刊、数字图书、数字报告和论文等。每一个 DOI 命名都具有唯一性,在互联网上作为稳定和持久的链接为数字化文献服务。DOI 不像统一资源定位符(uniform resource locator, URL),它不会随着时间和空间的改变而改变,不会发生 Error 404 错误或空链接,如果数字期刊、图书、报告和论文等在互联网上转移到了新的位置,空间位置发生了变化,或者 URL 发生变化,DOI 永远不会发生改变,避免“信息孤岛”的产生。国际上现在有 10 家注册代理机构,分布在美国、欧洲、中国以及日本,其中美国(<http://www.crossref.org/>)是世界上最大的注册代理机构^[3],目前 DOI 号都是以 10 开始。DOI 命名需要遵循一定的规范^[4],DOI 号的后缀可以随着元数据信息的增多随时扩充。在国际学术界,对于学术研究和科研人员来说,DOI 作为稳定持久的对象标志符已经被广泛的所接受,用来进行文献的查询和标注。DOI 可以用在纸质印刷版上,也可以用于数字化出版方面。

1 DOI 在国外的发展现状

2000 年 6 月,美国 CrossRef 成为由国际 DOI 基金会(International DOI Foundation, IDF)授权的 DOI 唯一注册机构,它维护着一个 DOI 查询系统,从而保证

其他参与其中的出版商、图书馆、研究者或销售商能够实现内容的自动链接^[2]。CrossRef 的主要目的是利用 DOI 来实现不同出版商出版的在线学术资源之间高效而可靠的交叉链接。

Crossref 依据刊物或者出版商的年收入收取会员费,成为会员后,注册 DOI 信息时,仍然需要按照 1 美元/个进行收费。截至 2013 年 3 月,Crossref 会员机构共注册 DOI 号 6 079 万 9 986 个^[2](表 1),其中属于期刊的 DOI 号共 4 903 万 8 913 个,占 80.7%,说明 DOI 在期刊中的重要作用,是期刊国际化的重要组成部分。同时,书本进行 DOI 注册也在快速发展,较 2009 年 12 月增长了 218%。

表 1 2009-12 和 2013-03 CrossRef 会员机构 DOI 号注册分类情况

种类	2013-03		2009-12		增长/%
	数量	百分比/%	数量	百分比/%	
期刊文献	49 038 913	80.7	34 641 674	88.2	41.6
书本	6 476 927	10.6	2 036 321	5.2	218
会议论文	3 242 853	5.4	2 114 243	5.4	53.4
总计	58 758 693	96.7	38 792 238	98.8	

DOI 的迅猛发展,使著名互联网公司谷歌也和 CrossRef 进行合作。通过谷歌(www.google.com)搜索引擎,输入 DOI,也可快速定位到响应页面查询文献。

2 DOI 的应用意义

在数字出版时代,人们查询学术信息无时无刻离不开网络,DOI 逐渐成为学术检索之间的桥梁,有利于提高期刊的影响力和被引频次,它的应用将越来越广泛。

1) 跨平台检索。期刊给文章在 DOI 平台注册了 DOI 号后,可以通过 <http://dx.doi.org/> 网站快速定位到文章所在网址,检索查询极其方便,实现了资源的网络快捷共享,方便读者和专家,提高期刊影响。国际权威期刊和数据库,如爱思唯尔、施普林格等集团很早就进行了 DOI 号的注册;一些科学数据也进行 DOI 注册,在之后的数据研究中可以被引用或者进行检验。例如美国《American Chemical Society》期刊注册 DOI 号达到了 122 万 826 个。

2) DOI 和优先数字出版相结合。优先数字出版可以大大提高出版的时效性,提前几周乃至几个月快速刊出论文,如果同时针对数字出版的论文注册 DOI,印刷出版前将论文发布至期刊自身网站或合作检索系统,这样文章即可以快速的传递最新信息,而且更有利于读者通过 DOI 进行检索浏览。有些期刊进行了优先出版并使用了 DOI 号,如《西安交通大学学报》《地球科学》《山东大学学报(理学版)》等。

3) 期刊可在二次文献中标注 DOI 号。可以跨平台、快速便捷关联到相关的文献,进行查询,方便文献的快速获取。比如《地球科学》的文献中均标注了 DOI 号,使读者跨平台获得二次文献更加快捷,但应注意针对文献标注 DOI 格式的需要进行统一^[5-6]。

4) DOI 有利于版权管理。DOI 的唯一性和永久性等特性,保证了其链接原始数据的版权。通过 DOI 注册元数据,可使用户登录到数字内容的所有者服务器、版权在线购买服务器或其他版权解决机制的服务器,从而较好地解决了版权问题^[7]。

在国内,2007 年 7 月,中国科技信息所联合万方数据成为中文 DOI 注册代理机构,成为在国际 DOI 基金会(IDF)全球的 10 家注册机构之一。2013 年开始为中国科技期刊引证报告收录的核心刊物以及 OA 刊物开展 DOI 免费注册,这将为 DOI 在中国期刊的普及打下坚实基础^[8]。截止到目前,我国经过中文 DOI 系统统计的中文注册数量已经达到全球第二,突破了 1 700 万,其中期刊 DOI 注册数超过 1 500 万^[8]。

国外出版集团如施普林格、爱思唯尔也可为合作外文期刊提供并注册 DOI。汤森路透集团于 2009 年 1 月推出 JCR 的增强功能,新增了评价指标 eigenfactor score 和 article influence score^[9]。这 2 个指标均以对 DOI 的解析量作为引用次数、社会网络链接、媒体报道、评论以及读者评分的依据。

3 我国高校期刊使用 DOI 情况及讨论

3.1 样本期刊选择 本次研究对象是 300 种高校期刊,其中:社会科学类 101 种,自然科学类 199 种;自然科学类中综合高校(“211”院校)学报 54 种,专业期刊 145 种。根据中国学术期刊综合引证报告 2009 年版提供的高校期刊名称,通过互联网搜索引擎谷歌、百度和中国知网,对这 300 种期刊最新一期出版的文章 DOI 使用情况进行调查研究。

3.2 样本期刊使用 DOI 情况 对 300 种高校期刊 2013 年最新的文章进行调查发现,仅有 53 种期刊使用了 DOI,占 17.6%。高校自然科学类 199 种期刊有 50 种使用了 DOI,占 25.1%,其中:综合高校的 54 种期刊里有 24 种使用了 DOI,占 44.4%;而 145 种专业期刊仅有 26 种使用了 DOI,占 17.3%。而 101 种社会科学类期刊仅有《武汉理工大学学报(社会科学版)》等 3 家使用了 DOI,占 3.0%。这说明“211”学校期刊数字出版发展相对较快,对 DOI 等数字出版新技术和新标准较为重视;而社科期刊的数字出版较为落后。

3.3 结果分析 在 300 种高校期刊中使用 DOI 的仅占 17.6%,说明高校期刊使用 DOI 比例还是很低,其

中大部分是和中文 DOI 机构合作进行注册,只有极少数期刊自己申请注册 DOI。而在中国,学术网络出版总库共有 8 014 种正式出版的期刊^[10]。造成 DOI 使用比例较低的原因主要有以下几个:

1)注册和使用 DOI 必须在网络上有相应的链接,可以用来跨平台检索,因此需要期刊设立独立网站的网站或者进行过刊回溯。有 85% 的高校期刊以各种形式在网上发布自己的信息,如全文或题目和摘要等^[11],但是根据针对其中 600 种高校期刊调查发现,423 种有独立期刊网站中,仅 255 种期刊有过刊检索,占 60.3%,大部分只有标题和作者姓名,拥有独立网站的期刊较少;其中有 68 种期刊依靠国内大型数字出版平台网站,例如 CNKI 数据平台、万方知网和龙源科技。这些大型数字平台网站虽然技术条件成熟,但是发布数字期刊相对滞后,受平台制约,检索注册 DOI 有所不便。有些期刊虽然有 DOI,但并不能实现真正意义的跨平台检索。

2)使用 DOI 进行跨平台检索,需要对 DOI 元数据进行注册,建立 DOI 响应页面和包括出版者信息以及 DOI 在内的元数据信息(标题、作者、文章题目、URL 和参考文献等)的 XML 文件^[12],并且 XML 文件必须是基于 CrossRef 规定的语法格式。有些期刊虽然有 DOI,但因各种原因没有针对期刊的文章信息进行注册或在文章上注明 DOI,致使 DOI 成为摆设。

3)期刊网站改版或者网络链接 URL 的变更,需要重新注册 DOI 元数据,所以正式在期刊文章上使用 DOI 的期刊较少。

4 结束语

建立 DOI 平台,可以实现数字出版界的沟通与合作,实现不同信息服务服务商和不同平台的无缝结合,有助于实现国内和国外、中外文信息的融合,为连接学术研究不同范畴和领域提供了可能性,延伸了出版行业中每个个体的链接空间和合作空间^[13];同时,其他技术和 DOI 的结合更加有利于 CrossRef 向更深的层次发展:因此,建立数字化平台网站或者与国内外大型

出版集团合作,在拥有 DOI 的同时同步进行注册,使其成为期刊数字出版的“利器”,真正为广大受众服务,实现国内与国外数字出版同步发展。

数字时代是信息竞争的时代,只有把握住信息的脉搏,才能更有利于期刊的进一步发展,才能适应当前数字出版的新形势,向国际化迈进。

5 参考文献

- [1] 全球数据存储总量达 295EB[EB/OL]. [2011-02-14]. http://www.edu.cn/sjc_9768/20110214/t20110214_576661.shtml
- [2] Crossref 中文官网[EB/OL]. [2013-05-07]. <http://www.crossref.org.cn>
- [3] 维基百科. DOI[EB/OL]. [2013-05-07]. <http://zh.wikipedia.org/zh-cn/Doi>
- [4] CrossRef. Establishing a DOI suffix pattern [EB/OL]. [2012-12-30]. http://help.crossref.org/#establishing_a_doi_suffix_pattern
- [5] 宋静,倪东鸿. 如何规范著录带有 DOI 的文后参考文献[J]. 编辑学报,2012,24(1):49-50
- [6] 廖京辉,董秀玥. 谈有效利用数字对象唯一标志后缀部分[J]. 编辑学报,2013,25(1):41-42
- [7] 洪瑞. DOI 对期刊出版及传播的影响[J]. 新闻传播 2012,28(6):15-16
- [8] 中文 DOI. DOI 注册推新政[EB/OL]. [2012-12-19]. <http://www.chinadoi.cn/portal/newsAction!detail.action?id=91>
- [9] 陈少华. 2008 北美数字出版的发展概况及趋势[J]. 出版广角,2009(4):23-27
- [10] 中国知网. 期刊[EB/OL]. [2013-08-08]. http://epub.cnki.net/KNS/oldnavi/n_Navi.aspx?NaviID=100
- [11] 王亨君,王淑华,姚戈,等. 科技期刊国际化研究[M]. 武汉:中国地质大学出版社,2011
- [12] CrossRef. Deposit basics [EB/OL]. [2012-12-30]. http://help.crossref.org/#deposit_basics
- [13] 赵蕴华. 中文 DOI 应用现状及建设构想[J]. 情报理论与实践,2007,31(6):855-857

(2013-07-06 收稿;2013-09-15 修回)

压力、压强的量符号应为“p”

2001 年出版的《法定计量单位在医学上的应用》(第 3 版)中,编写者参考国外有关资料,以表格形式给出了血流动力学一些压力量符号的下标写法,这很有导向意义,也是值得提倡的。例如:动脉收缩压(arterial blood pressure, systolic)的量符号为 $P_{a,\text{sys}}$,肺动脉压(pulmonary artery pressure)的量符号为 P_{PA} ,左心室压(left ventricular pressure)的量符号为 P_{LV} 。然而遗憾的是这里将压力的主符号错写成了“P”。依据 GB 3102.3—1993,压力、压强的量符号为“p”,因此,上述 3 个量符号的正确写法应为 $p_{a,\text{sys}}$ 、 p_{PA} 、 p_{LV} 。

(诸仁)