



## 2 DOI 后缀构成的指导规则

因 DOI 前缀的构成规则已确定, 因此本文重点对 DOI 后缀构成的指导做分析。

(1) DOI 后缀应该尽量简明, 而且 DOI 号将在网络和纸质印刷品上公布, 以后的用户会不断地进行复制。

(2) 后缀在同一组前缀中是唯一的, 而且不分大小写。例如: 10.1006/abc 和 10.1006/ABC 是一样的。

(3) 后缀还应该反映一个连续的、逻辑的系统, 这样它能比较容易被承认和接受。在分配后缀时, 可以让它包括现有的正在使用的国内标识符。著书目录的元数据 (如期刊、卷册和页码) 可以用在后缀中, 因为这些元数据和 DOI 号一样都是持久性的。

(4) 后缀的结构应反映注册内容的层级信息或层次的间隔, 例如第一个节点用一个字符串来表示期刊标题, 接下来的节点用来表示期刊的年份和顺序。以学术出版社 (Academia Press) 的 DOI 号为例: DOI: 10.1006/jmbi.1998.2354。

(5) DOI 号的后缀应该是可以扩展的, 而且后缀中的每个节点都应该具有这样的特点。如果不久以后某篇已经注册的文章的部分内容, 如数字、图表、或可以补充的其他材料, 也可以补充到 DOI 系统中, 而后缀应该在以前的基础上进行扩展。还是以学术出版社 (Academia Press) 为例: DOI: 10.1006/jmbi.1998.2354.f002。

## 3 国外数字期刊唯一标识符发展现状

2000 年秋, 美国出版业的开放性电子书籍出版标准协会极力推荐采用 DOI, 并运用于电子图书的主要标识系统。目前, 国外越来越多的数字文献生产商正逐渐采用 DOI 来标识自己的内容实体。如 Elsevier、Springer 等大型出版商都已使用 DOI 对数字资源进行标识, 各自均形成了比较完整的 DOI 运行管理机制, DOI 的解析系统发展也更加成熟。现在 DOI 已经广泛地应用在期刊、学位论文、科学数据、会议记录、图书、甚至包括文献中的图、表等诸多科技信息资源领域中。

### 3.1 Elsevier

荷兰的爱思唯尔 (www.elsevier.com) 是一家经

营科学、技术和医学信息产品及出版服务的出版集团。通过与全球的科技与医学机构的合作, 公司每年出版 1000 多种期刊和 2000 多本新书, 以及一系列电子产品, 如 Science Direct、Scopus、Embase.com、SDOL 等。

SDOL (Science Direct Online) 是爱思唯尔公司的科技、医学期刊全文数据库产品, 包括 1800 多种期刊, 六百多万篇全文文献<sup>[2]</sup>

SDOL 采用 DOI 对其数字期刊进行唯一标识的格式为: “DOI: 10.1016/刊物类型. 刊物代码. 出版年. 出版月. 流水号” 例如: “Academic Radiology” 里一篇论文 “Comparison of Three Methods to Increase Knowledge About Breast Cancer and Breast Cancer Screening in Screening Mammography Patients” 的唯一标识符是: “10.1016/j.acra.2007.01.011”。其中: “10.1016” 为 Elsevier 的 DOI 统一前缀; “j” 为刊物类型; “acra” 为论文所载期刊《Academic Radiology》的代码; “2007.01” 为论文录用时间; “011” 为论文流水号<sup>[3]</sup>。

### 3.2 Springer

德国的施普林格出版集团 (www.springerlink.com) 是通过 Springer LINK 系统提供学术期刊及电子图书的在线服务。

Springer LINK 电子期刊包含 439 种学术期刊, 其中近 400 种为英文期刊。学科分为以下 11 个“在线图书馆”: 生命科学、医学、数学、化学、计算机科学、经济、法律、工程学、环境科学、地球科学、物理学与天文学, 是科研人员的重要信息来源<sup>[2]</sup>。

Springer 采用 DOI 对其数字期刊进行唯一标识的格式为: DOI: 10.1007/刊物类型. 期刊代码. 论文流水号. 校验码”。例如: “The European Journal of Health Economics” 里的一篇论文 “A decision model to compare health care costs of olanzapine and risperidone treatment for schizophrenia in Germany” 的唯一标识符是: “DOI: 10.1007/s10198 - 006 - 0347 - 0”。其中: “10.1007” 为 Springer 的 DOI 统一前缀; “s” 为刊物类型; “10198 - 006” 为期刊代码; “0347” 为论文流水号; “0” 为校验码<sup>[4]</sup>。

目前, 施普林格集团正在和中国的科学出版和高等教育出版社这两家优秀出版社进行合作。将建立一个名为“中国科学图书馆”的网络平台。到时候, 用户就能通过网络付费, 从网上可以下载中国

的各种专业书籍。整个自然科学范畴都将囊括其中,如物理、化学、电子、材料管理学、医药学。同时这个平台还将涉足历史、法律、心理学和经济学等领域。

#### 4 国内数字期刊唯一标识符发展现状

同国外相比,当前国内3家大型的数字文献生产商清华同方、万方数据、重庆维普在其内部都不同程度使用了唯一标识符,如万方数据、清华同方、重庆维普均对其期刊自制定了唯一标识符规范。除了自定义的唯一标识符规范以外,目前还没有成熟的唯一表示符解析体系。这三家的具体标识符规范如下。

##### 4.1 清华同方

清华同方出版的《中国学术期刊(光盘版)》(CAJ-CD)是我国第一部以电子期刊方式连续出版的大型集成化学术期刊全文数据库。其使用的唯一标识符比较规范,于1998年12月通过由新闻出版署主持的专家评审会的审定。随后新闻出版署发布了“关于印发《〈中国学术期刊(光盘版)〉检索与评价数据规范(试行)》的通知”(新出音[1999]17号),从1999年2月1日起在全国近3500种人编期刊中试行。数据规范中主要包括刊名、刊号、文献标识码、题名、作者、文章编号等字段<sup>[1]</sup>。

其中涉及到唯一标识的是“文章编号”字段,它被定义为:“为便于期刊文章的检索查询、全文信息索取和远程传送以及著作权管理,凡具有文献标识码的文章均可标识一个数字化的文章编号。该编号在全世界范围内是该篇文章的唯一标识”<sup>[5]</sup>。

清华同方自定义的数字期刊唯一标识符,自称为DOI。其唯一标识符的结构为:“CNKI:ISSN:XXXX-XXXX.0.YYYY.NN.ZZZ”。其中:“XXXX-XXXX”为论文所在期刊的国际标准刊号ISSN;“YYYY”为文章所在期刊的出版年;“0”为文献类型;“NN”为文章所在期刊的期次;“ZZZ”为流水号。例如:发表在《海洋环境科学》2006年第S1期的文章《石油污染对牡蛎超显微结构毒性效应的研究》;CNKI给它的DOI是“CNKI:ISSN:1007-6336.0.2006-S1-002”。其中:“1007-6336”为论文所在期刊的国际标准刊号ISSN;“0”为文献类型;“2006”为文章所在期刊的出版年;“S1”为

文章所在期刊的期次;“002”为流水号<sup>[6]</sup>。

此唯一标识的优点是唯一标识符比较规范,已初步为今后作出规范的DOI标识工作打下了坚实的基础,在唯一标识的基础上运用DOI标准就可以实现与外部信息资源的链接。

##### 4.2 万方数据

万方数据在数字期刊加工时为期刊中每篇论文分配了唯一代码,该代码作为数字期刊论文的唯一标识符。其结构为:“期刊代码/论文出版年/论文所载期刊的期号/流水号”。其中:“期刊代码”是万方数据标识期刊的代码,取期刊名称的首个汉字的汉语拼音字母组成。例如:《安徽林业》2006年第3期中的一篇论文:《浅析国有林场可持续发展之策略》。它的唯一标识符是:“ahly200603015”。其中:ahly是《安徽林业》的期刊代码;200603是论文录用时间;015是论文的流水号。

此唯一标识的优点是每篇期刊论文都有唯一的标识符。不足是没有建立与DOI标识规范及相应的DOI解析系统,只能在内部使用。不能和外界动态的链接,实现资源共享。

##### 4.3 重庆维普

重庆维普的中文科技期刊数据库使用《〈中国学术期刊(光盘版)〉检索与评价数据规范(试行)》中定义的“文章编号”对文章进行唯一标识。期刊是用文章编号进行唯一标识的,即:国际标准刊号、出版年、期号及文章的篇首页码和页数等5段共20位数字组成。其结构为:“XXXX-XXXX(YYYY)NN-PPPP-CC”。其中:“XXXX-XXXX”:文章所在期刊的国际标准刊号ISSN;“YYYY”:文章所在期刊的出版年;“NN”:文章所在期刊的期次;“PPPP”:文章首页所在期刊页码;“CC”:文章页数。例如:《安徽医科大学学报》2006年1期《哮喘模型大鼠支气管平滑肌肌球蛋白轻链激酶的表达》重庆维普将其唯一标识符定义为:“1000-1492(2006)01-0001-03”。其中:“1000-1492”为文章所在期刊的国际标准刊号ISSN;“2006”为论文所在期刊的出版年;“01”为论文所在期刊的期次;“0001”为论文首页所在期刊页码;“03”为论文章页数<sup>[7]</sup>。

此唯一标识的优点是唯一标识符比较规范。缺点是没有建立与DOI标识相应的DOI解析系统,只能在内部使用。

从上述三家数字期刊生产商的标识规范来看,唯一标识符的应用仍处于一个初级阶段。

①它们自定义的唯一标识符作用仅仅是数字对象的唯一编号,唯一标识符的制定和使用缺少规范,各个数字文献生产商使用的唯一标识符各异,没有形成统一的规范标准。

②唯一标识符仅在各自资源范围内发挥作用,一旦脱离各自的数据库,其唯一标识符便不能揭示文献的任何特征。

③唯一标识符的应用层次比较低,其作用仅限于对内部数字对象的标识,与唯一标识符应用相关的解析系统和管理机制也没有建立起来,无法实现各个数字文献生产商的资源共享了。这将会造成国内数字文献生产商时间、资金的浪费,资源的重复建设。

④上述三个各数字期刊生产商各自为战,未构建成中文资源统一规范标准,目前无法与国外数字文献生产商实现资源共享。

## 5 建立中文 DOI 标识规范和中文 DOI 解析系统的意义

为了推动中文数字资源的 DOI 标识规范工作,最近,国际 DOI 基金会授权由中国科技信息研究所(简称中信所)及北京万方数据股份有限公司(简称万方数据)共同成立中文 DOI 注册机构。中信息所及万方数据成为唯一的中文 DOI 注册机构,它们将共同管理和运行中国 DOI 注册机构,同时建立一个中文 DOI 应用平台与门户网站,它类似于 Cross-Ref 这样的跨出版商的西文参考文献链接系统。这些工作的开展,对于中文数字资源的建设与发展具有重大意义。

(1) 促进国内中文数字化资源 DOI 唯一标识符规范及相应中文 DOI 解析系统的建设工作。

(2) 通过建立一个 DOI 中文应用平台与门户网站,国内各数字文献生产商可以成为此网站的成员,每个成员只要向 RA 提交并更新数据即可保证链接的有效性。

(3) 通过建立一个 DOI 中文应用平台与门户网站,可以将不同数字期刊出版商之间的资源有效地链接起来,实现资源共享,避免资源的重复建设。各数字文献生产商之间形成互补优势,提高自身的服务质量和品牌竞争力。

(4) 通过建立一个 DOI 中文应用平台与门户网

站,广大用户可以直接从中文 DOI 系统上检索、获取相关文献。中文 DOI 系统本身不提供全文文献,它会将检索结果链接到数据库提供者或出版者提供服务的网页,用户可以根据自身的需要向数据库提供者或出版者索要全文文献。

(5) 可以为广大数字文献生产商提供中文 DOI 命名及应用相关的增值服务。

(6) 实现与国际规范接轨,与国外数字文献生产商的资源共享。

## 6 结 语

开展中文 DOI 唯一标识符规范与中文 DOI 标识相应的 DOI 解析系统工作,是中文数字信息资源建设的重要组成部分,是实现国内外各数字文献生产商资源共享的前提,是推动各数字文献生产商增值服务的保障。它将推动我国数字信息资源进入良性发展阶段,DOI 今后将会在网络上不同的有版权归属的信息资源领域将得到空前的发展。

## 参考文献

- 1 我国数字图书馆标准规范建设:数据对象唯一标识符的现状与发展[R].科技部科技基础性工作专项资助重大项目报告,2003.2.
- 2 我国数字图书馆标准规范建设:主要数字文献生产商唯一标识符应用现状[R].科技部科技基础性工作专项资助重大项目报告,2006.8.
- 3 荷兰爱思唯尔出版集团网站[EB/OL].[http://www.science-direct.com/science?\\_ob=ArticleURL&\\_udi=B75BK-4NH7KK1-8&\\_user=10&\\_coverDate=05%2F31%2F2007&\\_rdoc=1&\\_fmt=&\\_orig=search&\\_sort=d&view=c&\\_acct=C000050221&\\_version=1&\\_urlVersion=0&\\_userid=10&md5=efce1b0446d21fb7c76d728516c3e81a](http://www.science-direct.com/science?_ob=ArticleURL&_udi=B75BK-4NH7KK1-8&_user=10&_coverDate=05%2F31%2F2007&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=efce1b0446d21fb7c76d728516c3e81a),2007-3-6.
- 4 德国施普林格出版集团网站[EB/OL].<http://www.springerlink.com/content/?k=DOI%3a10.1007%2fa10198-006-0347-0>,2007-3-12.
- 5 毛军,孟连生,镇锡惠,倪金松,王燕,胡铁军,刘燕飞.试论我国数字资源唯一标识符发展战略[J].现代图书情报,2005,(2):1-4.
- 6 中国知网网站[EB/OL].<http://lsg.cnki.net/grid20/detail.aspx?QueryID=18&CurRec=2>,2007-3-28.
- 7 维普资讯网站[EB/OL].<http://www.cqvip.com/content/citation.dll?id=21194360&SUID=EGBNBFHDNCBCOPIBOLGEMCBCDOIPNFP#>,2007-3-28.

(责任编辑:孙晓明)