

●赵蕴华 (中国科学技术信息研究所, 北京 100038)

中文 DOI 应用现状及建设构想

摘要: 本文阐释了 DOI 产生的背景以及现阶段 DOI 在国外应用的状况, 详细介绍了西文 DOI 应用平台 CrossRef 的工作原理、运营机制及主要功能。从信息资源共享的角度阐述了目前中文 DOI 存在的主要问题, 以及建立中文 DOI 应用平台的必要性、目的和意义。

关键词: 数字对象唯一标识; 唯一标识符; 数字期刊

Abstract: This article explains the background of Digital Object Identifier (DOI) and its application status abroad. It gives a detailed description of the working principles, operating mechanism and main functions of CrossRef. Furthermore, the article discusses the major problems of Chinese DOI from the perspectives of information resources sharing, and the necessity, purpose and significance of constructing the Chinese DOI application platform.

Keywords: DOI; unique identifier; digital periodical

数字对象唯一标识 (Digital Object Identifier, DOI) 是针对数字资源的永久性标识符。1994 年美国出版协会 (American Associate of Publishing, AAP) 为探讨互联网上的著作权保护问题, 成立非盈利性组织国际 DOI 基金会 (International DOI Foundation, IDF), 开始探讨防止著作出版物被非法使用所需的技术, IDF 在美国国家创新研究组织 (The Corporation for National Research Initiatives, CNRI) 的配合下, 制定了 DOI 标准和相应的解析系统 (Handle System)。

目前, IDF 拥有 7 个 DOI 注册代理机构 (Register Agent, RA), 上千万个已经分配并解析的 DOI 号码在美国、欧洲和澳大利亚以及一些非英语国家的各 DOI 代理注册机构注册, 其应用范围已从科技领域扩展到了政府部门领域。

DOI 实现了动态的持久链接。如果文献的 URL 发生变动, 成员出版商只要向 RA 提交并更新数据即可保证链接的有效性; 其次是提供一站式服务, 即各出版商通过 DOI 系统实现引文到全文一站式的链接, 从而提高产品竞争力。

1 西文 DOI 应用平台

CrossRef 是一个基于 DOI 实现文献引文跨出版商服务平台链接的参考文献链接服务系统。AIP, Blackwell Publishers, Elsevier Science, IEEE, Kluwer, Nature, Springer 和 Wiley 等多家著名的学术出版商在 2000 年初联合创立了非盈利组织世界出版商联合会 (Publishers International Linking Association, Inc., PILA), 负责对 CrossRef 进行管理

和为成员出版商提供链接服务。CrossRef 并不保存全文内容, 而是通过数字对象标识符来链接出版商成员提供的文章元数据。

CrossRef 自 2000 年 9 月成为正式授权的 DOI 注册代理机构。作为 CrossRef 成员的出版机构, 应将其发行的电子出版物的 DOI、URL 和相应的元数据提交给 CrossRef, 这些数据收藏在 CrossRef DataBase 中, 用于实现 CrossRef 链接服务, 也为会员和单个用户提供元数据及 DOI 的检索及链接服务。截至 2007 年 3 月 31 日, 加入 CrossRef 联盟的出版社和学会有 2 345 个; 图书馆 1 197 个; 覆盖的刊物 16 884 个; 发放的 DOI 数量达到惊人的 25 696 063 个; 而 2007 年 2 月单月的 DOI 发放量就达到 585 846 个, 检索 (引文匹配查询) 量达到 5 486 433 个, 终端用户解析量达到 13 290 744 个^[1]。

CrossRef 目前在学术期刊方面得到了广泛的应用, 国外的大型出版集团, 如荷兰的爱思唯尔出版集团 (www.elsevier.com)、德国的施普林格出版集团 (www.springerlink.com) 等都对其期刊论文等数字信息资源进行了 DOI 标识。原因是学术期刊有相互连结的需要, 而且本身制作也较其他领域更有规律。例如, 一篇文献的引文 (或参考文献) 原文如果有网络版并注册了 DOI, 用户就可以直接利用引文中设置的 CrossRef 链接按钮直接链接到引文所在的出版社网页, 并获得引文文摘信息; 如果信息生产商从出版商那里订购了该网络版期刊, 用户将可以由此打开引文全文。

1.1 CrossRef 运营机制

每个加入 CrossRef 的出版商成员首先获得一个 DOI 作

为前缀，出版商为其出版的每一篇期刊文献编制一个包含出版商本身 DOI 前缀的 DOI 号，并将 DOI 号附加到文献元数据和 URL 上。所有文章的原数据，出版商会以 XML 的形式提供给 CrossRef 元数据库 (MDDB: Metadata DataBase)，CrossRef 再将每一篇文献的 DOI 和 URL 注册在一个统一的 DOI 名录中。同时，出版商也将每篇文献的参考文献的数据提供给参考书目解析器，这一解析器是用来检索 DOI 群的，它是 MDDB 的主要部分。用户点击 CrossRef 链接，就可以链接到显示有该论文全部引文目录的出版商网页。引文全文获取形式取决于出版商的选择，引文可能以 HTML 形式发布，用户可以直接通过适当的机制获取全文；也可通过订购、文献传递、按浏览付费或者预印本付费等方式获得。如读者所在机构已订购并被授权使用某数据库时，则可通过授权的合法 IP 地址登陆该数据库，并获得引文的全文^[2]。

1.2 CrossRef 的主要功能

1) 永久指向。CrossRef 解决了数据因 URL 发生变动后出现的断链问题。不会出现因网域的异动使数据消失的现象。

2) 动态维护。CrossRef DataBase 中的数字对象元数据和 URL 由出版机构负责更新维护。如果成员出版商提供的文献 URL 发生变动，那么他只要向 CrossRef 提交并更新数据即可保证链接的准确性和有效性。

3) 实现一对多的路径。通过 DOI 将参考引文进行多项解析，方便快捷地链接到该文献的文摘或全文。

4) 支持新的应用。目前已经从学术期刊领域扩展到电子书、政府文档、教学参考、科学数据等方面。

5) 保护著者版权。CrossRef DataBas 仅存储描述文件的 MetaData，而不存储文件内容。这样可以避免使用者随意下载全文，保护了著者的版权。

1.3 CrossRef 给我们的借鉴

CrossRef 建立了开放式的知识链接体系。它的运营管理机制包括技术、组织、政策和具体应用。①技术上：Handle System 是一个基于因特网的分布式数字对象命名与标识系统。它保证了数据的唯一性、永久性、多重性、安全性、可扩展性。②组织上：是出版商联盟。建立在多家出版商协作基础上的链接系统。③政策上：制定了相关的会员会费交纳制度及明确规定了会员的义务和职责。

2 中文 DOI 应用平台

2.1 国内主要数字文献生产商唯一标识符应用现状

中文 DOI 平台的建设思路应按照 IDF 基本 DOI 政策和技术进行运作，并参照 CrossRef 的运行模式的成功经验，同时又要结合国内数字信息资源建设的实际情况。目

前，通过国内主要数字期刊文献生产商唯一标识符应用现状 (见表 1)^[3-4]，可以看出国内的唯一标识符的应用仍处于初级阶段，还没有建立 DOI 标准规范体系。

表 1 国内主要数字期刊文献生产商唯一标识符应用现状

	清华同方	万方数据	重庆维普
唯一标识符使用现状	采用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	采用分配给期刊中的每篇论文的唯一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	采用“文章编号”对文章进行唯一标识。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；“PP-PP”：文章首页所在期刊页码；“CC”：文章页数
优点	唯一标识符比较规范	每篇期刊论文都有唯一的标识符	唯一标识符比较规范
缺点	尚未运用 DOI 标准，未实现与外部信息资源的链接，实现资源共享	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	没有建立与 DOI 标识相应的 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

从上述 3 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。唯一标识符的制定和使用缺少规范，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，具体表现在：

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。

上述 3 个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。因此,建设中文 DOI 标准规范及中文 DOI 应用平台是实现国内外信息资源共享的必备条件,建设中文 DOI 应用平台已迫在眉睫。

2.2 建立中文 DOI 应用平台的必要性

建立中文 DOI 应用平台,可以实现出版界和数字图书馆界的沟通与合作;实现不同信息服务服务商的联结;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;延伸了行业中每个个体的链接空间和合作空间(见图1)。

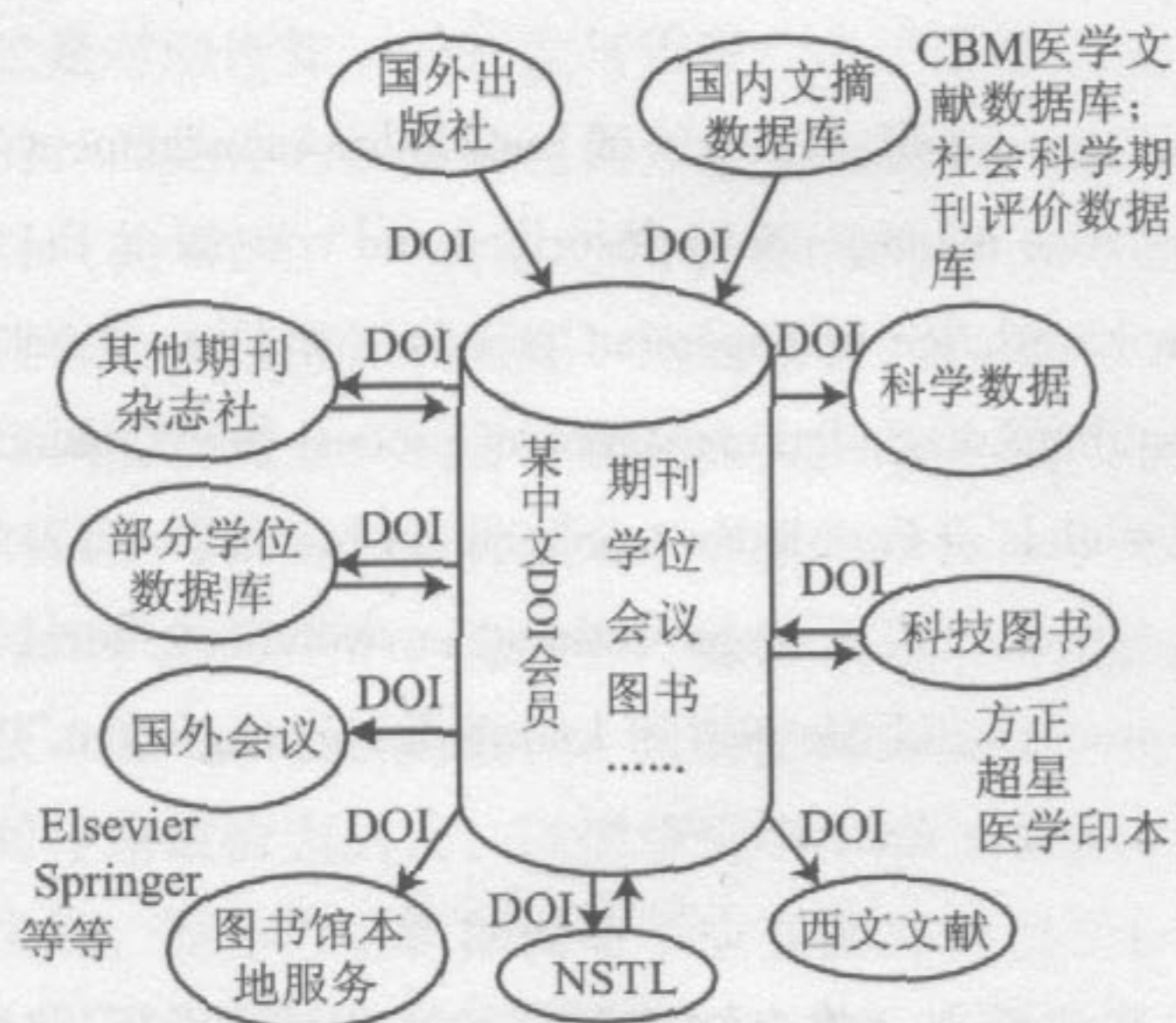


图1 中文 DOI 应用平台

从图1可以看出,今后作为 DOI 会员的信息集成商可以在中文 DOI 应用平台上实现自身资源与国内外其他机构所拥有的科技信息资源的共享,从而为广大用户通过中文 DOI 平台实现快速的链接服务,使中文 DOI 系统成为名副其实的开放式的知识链接系统。

2.3 中文 DOI 应用平台的功能构建

借鉴 Crossref 在包括技术、组织、政策和具体应用的成功经验,中文 DOI 应用平台同样要具有永久指向、动态维护、实现一对多的路径、支持新的应用、保护著者版权等。为广大普通用户及会员用户实现 DOI 解析和查询服务。

2.4 中文 DOI 应用平台建设的目的和意义

建设 DOI 标准规范与 DOI 应用平台,可以为中文信息资源生产商、中间商等机构带来共赢。

出版商:通过 DOI 应用平台实现引文到全文的链接,实现一站式信息服务,从而提高整个行业的服务数量和服务质量,提升整个行业的产品竞争力。

图书馆:通过在本地导入 DOI 并与 OpenURL 结合,图书馆为用户提供了访问更多的图书馆未能收藏的全文文献的机会,并可以更加有效的管理现有资源,开发更为高效的检索和馆藏功能。

二次文献数据库:实现二次文献数据库与全文数据库的链接,如 SCI, EI, CSA, CABI 等都通过 DOI 建立了与全文的链接。

搜索服务:在学术搜索中引入 DOI 可以提高搜索质量。例如 CrossRef 与 Google 合作的 CrossRef Search。与一般的 Google 搜索不同, CrossRef Search 仅在成员所提供的学术内容中进行搜索,因此不像一般的 Web 搜索那样产生很多噪声结果,而且通过 DOI 保证了检索结果到全文链接的有效性和持久性。

政府机构:相关政府机构通过 DOI 的推广和应用可以更有效地管理学术信息资源,从标识体系的建立和应用上实现包括科技文献和科学数据在内的资源整合,从而有助于建立一个基于分布环境下的一站式的科技基础信息服务平台。

中文 DOI 会员:通过 DOI 的应用可以与国内、外数字文献生产商实现资源共享,并为行业发展带来共赢。

2.5 中文 DOI 应用现状存在的主要问题

1) 尚未建立 DOI 标准规范体系。DOI 标准规范的建立是中文 DOI 应用平台建设基础。

2) 中文 DOI 平台的商业运营模式、收费制度、会员的职责和义务等正在和多方机构和专家一起讨论制定中。

3) 如何解决出版机构与数字文献生产商之间的版权问题以及利益冲突(DOI 由谁注册、DOI 链接指向谁),同一文章可能被注册多个 DOI 号,与国外系统的互联。

中国科技信息研究所已经联合北京万方数据股份有限公司向 IDF(国际 DOI 基金会)申请成为中文 DOI 的代理。目前正开始从事中文 DOI 应用平台建设工作,此应用平台将参照目前 IDF 最成功的代理机构 CrossRef 先进的技术、组织、政策来保证中文 DOI 解析和链接服务。

3 结束语

中文 DOI 应用平台的建设,将为加入中文 DOI 系统的用户提高使用资源的效率;提高中国精品期刊在国际上的影响度和显示度;有利于数字资源的长久保存和唯一识别;逐步实现中西文数字资源的链接,实现行业的共赢。□

参考文献

- [1] [EB/OL]. [2007-03-23]. <http://www.crossref.org/>
- [2] 潘松华. CrossRef 在数字图书馆中的应用[J]. 图书馆理论与实践, 2005(6): 63-65
- [3] 中国知网网站[EB/OL]. [2007-03-28]. <http://lsg.cnki.net/grid20/detail.aspx?QueryID=18&CurRec=2>
- [4] 维普资讯网站[EB/OL]. [2007-03-28]. <http://www.cqvip.com/content/citation.dll?id=21194360&SUID=EG-BNBFHDHNCBCOPIBOLGEMCBCDOIPNFP#>

作者简介:赵蕴华,女,1967年生,助理研究员,硕士。

收稿日期:2007-06-22