

# 建立中文DOI系统的意义和发展前景



■ 张 旭 / 万方数据技术研究院

## 一、建立中文DOI系统的意义

DOI (数字对象唯一标识) 在欧美出版业发达地区已经得到了认同和广泛应用。据统计, 目前西方超过70%的出版社已经加入DOI系统, 所注册的DOI数量接近3000万个, DOI的应用已经从期刊发展到电子书、学位论文、会议论文、政府文档和科学数据等。可以说, DOI为互联网上的数字对象提供了独一无二的标识, 方便了计算机的互操作并可逐步实现各种数字产品的流通和交易, 因而被形象地称为“互联网上的条形码”。

在国内, 无论是对一次文献出版商还是中间服务商, 唯一标识系统都还没有形成统一规范, 仅在各自资源范围内发挥作用, 一旦脱离各自的数据库, 唯一标识符便成为没有意义的字符串。同时, 唯一标识符的应用层次比较低, 作用仅限于对内部数字对象的标识, 没有建立相应的解析

系统和管理机制。

因此, 采用国外先进的理念和技术, 利用DOI建立跨出版商的数字对象标识和链接体系, 对于我国数字出版业的发展, 以及数字图书馆建设都具有重要意义。可以说, DOI及相关的解析系统是一项重要的基础性设施, 推广应用中文DOI的意义可以概括为以下几点:

(1) 为中文数字资源提供符合国际标准和规范的唯一标识, 有利于中文数字资源的长期保存和定位;

(2) 通过DOI与外部资源, 包括中文资源和外文资源, 一次文献和二次文献, 科技文献和科学数据的链接, 建立动态的、开放的、互动的知识链接体系, 整体提升数字信息资源的使用率, 促进行业共赢;

(3) 通过DOI的引入, 逐步实现中西文数字资源的链接, 为建立统一的中英文知识链接系统提

供技术标准和平台;

(4) 为国家科技基础条件平台建设提供支撑, 在标识体系的层面实现包括科技文献和科学数据在内的资源整合, 从而建立一个基于分布式环境的科技基础条件平台;

(5) 为标识数字对象的版权状态提供基础。

## 二、中文DOI工作进展情况

为推进中文DOI的发展, 2007年初, 经与IDF深入沟通, 中国科学技术信息研究所 (以下简称“中信所”) 和万方数据公司联合申请, 成为DOI的唯一中文注册代理机构, 从而将DOI带入中国。该注册代理机构以通过DOI促进我国科技信息资源的标识和相互链接为目标, 以非赢利性的运行方式为原则开展DOI注册和服务工作。目前, 该机构依托万方数据技术研究院在DOI服务平台、标准规

范、注册服务模式以及应用示范等方面开展了大量研究工作，具体如下：

### 1、中文DOI规范的制定

建立中文DOI的标准规范是中文数字对象互联互通的基础。目前，万方数据技术研究院通过对IDF的DOI标准、CrossRef的DOI应用标准以及国内数字图书馆标准的研究，并参考国内主要出版商和服务商现有内部数据标识情况，制定了适用于期刊、电子书及其它扩展资源的中文DOI标准规范，包括DOI命名规范、元数据规范等。目前，该标准规范已经在科技部精品科技期刊项目和科技期刊与电子书的链接实验中开始试用，并可以延伸使用到对科学数据的标识中。

### 2、中文DOI系统平台的搭建

DOI的可解析性是其有别于其它文献标识体系（如ISSN和ISBN）的重要特点之一。万方数据技术研究院通过与美国CNRI的合作，参照CrossRef网站及服务系统，开发建设了国内第一个实用的中文DOI网站及服务系统（www.chinadoi.cn）。该平台的建立，不仅实现了对中文信息资源的解析，同时还具有批量处理数据标识请求和批量数据解析的功能，为中文DOI的推广应用

奠定了基础。

### 3、DOI应用研究和实验

#### (1) 科技期刊

国外利用DOI进行科技期刊链接的主要机构是CrossRef。万方数据技术研究院通过对国内主要出版商的调查和了解发现，国内对DOI需求比较强烈的主要是具有较高学术水准的核心科技期刊，包括部分西文期刊。这部分期刊有着融入国际期刊业的需求强烈，文章具有较高的质量，同时，希望在标准规范上与国际接轨。目前，依托科技部精品科技期刊项目，万方数据技术研究院联合中信所信息技术支持中心对DOI在科技期刊上的应用进行了实验。年内将完成30万条精品科技期刊文章的DOI注册，并开始提

供精品科技期刊文章的DOI解析和检索等基本服务，同时进行精品科技期刊

引文的DOI检索和匹配，以期通过注册DOI提高精品科技期刊文章的访问率及引文的链接率，并提供可能的评价信息等增值服务。

#### (2) 科学数据

利用DOI对科学数据进行标识的主要价值在于提高科学数据的引用和使用、促

进科学数据与科技文献的链接，以及实现对科学数据的所有权进行标识等。目前，技术研究院和中信所信息资源共享中心已经对DOI在国外科学数据领域的主要应用进行了调研，并与科学数据的主要注册代理机构TIB进行了交流。目前，已经初步确定了科学数据领域DOI应用的标准和规范，开展了在科学数据领域的示范应用。

#### (3) 西文科技资源的链接

研究表明，中文科技期刊对西文期刊的引用要远远大于西文期刊对中文期刊的引用，但在任何一个中文信息服务平台，以及大多数出版社平台上，这种链接并不能指向西文原文或文摘。但事实上，如前所述，70%以上来自正式出版物的西文文献都带有DOI标识，特别是主要的学术出版物都带有DOI标识，因此，利用DOI可以实现中文文献与西文文献的链接，促进对西文文献的利用率。通过对万方数据公司的西文引文进行研究发现，其中40%—50%的西文引文可以查询到相应的DOI标识，从而可以引导读者到西文文献的详细文摘、其它引文、以及全文的版权和付费页面。

#### (4) 电子书的链接

中文科技期刊会引用大量的图书资料。研究表明，其中30%—40%的图书可以在不同的电子出版商中（如北大方正、书生等）查询到。建立起期刊与电子书之间的链接，可以促进电子书的使用，并提高期刊用户的客户满意度。目前，万方数据技术研究院正在进行有关电子书DOI标识规范的建立以及通过DOI实现期刊与电子书的链接。



### 三、中文DOI推广中的主要问题

中信所和万方数据公司的DOI注册服务以非赢利性为目标。但推广DOI应用需要一定的成本，包括支付给IDF的费用。中科院文献中心毛军和孟连生等研究人员在科技部数字对象唯一标识课题中对成本问题进行了详细描述，这里不再赘述。据了解，同样以非赢利为目标的注册代理机构CrossRef对每一个DOI标识收取的费用（包括注册费和服务费用）达到1美元。目前国内数字出版刚刚起步，出版社的规模都相对较小，因此，如何以较低的成本推广DOI应用就成为一个关键问题。目前我们考虑基本的费用来源包括：

(1) 长期来讲，出版商、资源提供商和服务商交纳的费用：以注册的DOI数目为计算依据，尽可能降低成本，以减少注册的障碍；

(2) 寻求项目的支持：DOI是对科研成果的永久标识和保存，因此希望寻求科研项目管理部門的支持；

(3) 先期万方数据公司和中信所利用自有资金的支持。

### 四、中文DOI的发展前景

我国的数字出版业刚刚起步，数字出版商开始逐步规范化和国际化，并开始关注版权问题，在这种背景下，只要解决好成本和机制等问题，DOI在我国具有良好的发展前景。

我们的远景目标是：

(1) 中文DOI得到广泛应用。

每个具有引用价值的、具有知识产权的中文数字对象都具有世界唯一的、被公众所认可的数字对象标识。该标识的普及将实现数字对象之间的跨平台链接，中文信息资源将结成分布式的一体。数字标识跨越数字对象生命周期的不同阶段得到广泛的应用。

(2) DOI标识的资源不断多样化。

从科技期刊发展到图书、学位论文、电子书、科学数据、医学案例、视频资料等。事实上，TIB已经对医学案例和科技短片进行了标识。另外，欧洲的标准制定机构也开始利用DOI实现标准与科技文献的链接。可以说，在标识的类别上，DOI标准的可扩展性决定了其无限的延展空间。

(3) DOI应用与版权结合。

DOI在版权方面的应用具有巨大的潜力。目前，IDF正积极参与数字版权相关标准的制定工作，力争使DOI成为版权许可技术标准的内嵌标准。例如ONIX-PL标准和ACAP标准。通过对数字对象的标识，可以实现对数字对象版权状态的持续追踪。

总之，我们相信中文DOI在我国有巨大的成长空间和应用价值。实现中文DOI的广泛应用是中文科技信息互联互通，并与国际接轨的基础。希望相关出版社、信息服务企业、中间服务商和图书馆能联合起来，共同推进中文DOI的发展。

