

# 电子文献正文引文跨越学术领域的链接

□何立民 万跃华 李平

**摘要** CrossRef 是应电子文献在因特网中信息互相链接之功能,提供信息使用者直接在网络中查寻得到最终及正确的信息而产生的规范。文章介绍了 CrossRef 的发展及功能,论述了 CrossRef 对于用户、出版商、图书馆和其他信息服务机构的影响及其应采取的措施,提出中文文献尽快采用 DOI 技术,加入 CrossRef,利用 OpenURL 和 SFX,实现本地馆藏链接,有效地保护著作权。

**关键词** CrossRef 数字对象标识符 参考文献链接

随着网上电子期刊、电子图书、数字化的会议录等网络资源的增多,使用者已经不再满足于搜索到的原始文献,而是希望通过原始文献再轻松地找到其引用的文献全文,探寻知识的起源,学术的关联。如何从文献正文比较方便地查找到引文,这个亟待解决的问题被提到了电子出版商前面。CrossRef 在这种情况下应运而生。

## 1 参考文献链接的实现形式

参考文献的链接功能是指在整个电子资源库中建立起反映各知识点之间直接与间接关系的网络结构体系,检索者可以从知识网络结构体系的任何一个专业知识库的任一知识点直接进入该库,通过网络链接关系在知识库中实现文献之间用参考引文与全文直接链接,并可往返回溯检索。目前,网络环境中参考文献链接的实现形式主要有以下几类<sup>[1]</sup>:

### 1.1 封闭式静态链接系统

其特点是逻辑模型中所有部分均由链接者本地控制。例如许多信息服务商将自己的文摘索引数据库、全文期刊、文后参考文献、全文文章链接起来,典型的系统有 ISI 的 Web of Science、Science、NASA 的 Astrophysical Data System、NLM 的 PubMed 等。所有的链接都通过专门的程序事先计算,并通过嵌入 URL 或存放于专门数据库的链接记录来表达(静态链接)。这种系统技术是主要解决各种链接数据的处理和链接数据库的管理,不涉及多个资源拥有者之间的互操作,链接准确性较高,但对于链接源中包含的别人拥有的链接对象(例如另一出版商的全文期刊)就难以处理。

### 1.2 开放式静态链接系统

主要解决如何为链接源中包含的别人拥有的链接对象提供链接,其技术方法主要是资源拥有者互相提供链接对象的标识符或 URL,例如全文出版商向文摘索引商提供期刊文章的标识符和 URL,后者将标识符或 URL 嵌入自己的文摘索引数据中或者记载在自己的链接数据库中,从而建立两者间的链接。ISI 的 Web of Knowledge、Ei CompendexWeb、PubMed/PubRef、SilverLinker、SwetsNavigator 等都是这类系统。如 ISI 的 Web of Knowledge 已经开通了 16 家出版社 4000 多种期刊的全文链接,这类系统需要双方建立严格的数据描述和传递标准。当一个链接源拥有者需要与多个链接对象拥有者打交道时,管理负担比较重。

### 1.3 开放式动态链接系统

上述两类系统链接源与链接对象间的链接都是事先固化在 URL 或链接记录中,且往往只能处理一对一链接,因此对建立链接后出现的新的链接对象,链接对象在建立链接后的物理移动、一对多情况下的选择性链接等都无能为力。所谓动态链接系统则在用户需要链接时才根据一定规则计算链接路径和进行链接。因此可对用户链接前刚出现的链接对象或者位置予以链接,也可在链接计算规则中嵌入选择规则实现选择性链接,例如 SFX、OpenURL、OpenName 等,但是,这种系统的规则描述难度比较大,有时事先无法保证链接的有效性。

## 2 CrossRef:开放式参考文献链接的典型系统

### 2.1 什么是 CrossRef

CrossRef<sup>[2-4]</sup>是多个出版机构联合建立的开放式

参考文献链接系统。最初由 12 个世界顶尖的 STM (科学、科技、医学) 学术出版商, 于 2000 年 1 月为促进电子期刊领域中跨出版商的索引链接的合作而成立的。特别组成了一个 Publishers International Linking Association (简称为 PILA) 的非营利组织, 负责 CrossRef 的运作及提供服务。

CrossRef 成员包括: ACM、The Academic Press、The American Association for the Advancement of Science、American Institute of Physics、IEEE、Kluwer、Blackwell Science、Springer、Wiley、Elsevier Science……等国际知名出版商。自其成立以来, 截至 2002 年 3 月, 已有 109 个商业和非营利性质的主要出版商加入, 提供 5997 种期刊, 高达 450 万篇文献在引文与全文间能够实现互相索引链接, 同时每年以新增 50 多万篇文献的速度增长。EBSCO 公司、CAS 公司等机构, 芝加哥大学图书馆、华盛顿州立大学图书馆等图书馆也成为会员。CrossRef 还包含其他参考内容例如百科全书、教科书、会议记录和其他相应的文献。CrossRef 目前已经覆盖了学术出版的所有领域<sup>[5]</sup>。

## 2.2 CrossRef 原理<sup>[2,5,6]</sup>

CrossRef 本身并非一项产品, 而是一个系统, 每个参加的出版商从 International DOI Foundation 获得一个 DOI (Digital Object Identifiers, 数字对象标识符) (<http://www.doi.org/>) 作为前缀 (prefix), 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, 并将其附加到文献原数据和 URL 上。所有记录会以 XML 为基础的 DTD, 集结成一个提供给 CrossRef 元数据库 (MDDB: Metadata DataBase) 的批次档, CrossRef 再将每一篇文献的 DOI 和 URL 注册在一个统一的 DOI 名录中。另一方面, 出版商也将每篇文献中所节选的参考数据引文提供给参考书目解析器 (Reference Resolver), 此一解析器的功能是用来检索 DOI 群, 为 MDDB 的主要部分。

CrossRef 功能就如一种数据转换盘。它虽然不存储全文, 但能够使出版商和其他组织通过利用 DOI 对成员的出版商出版的论文加以标识, 创建链接。用户点击 CrossRef 链接, 就可以链接到显示有论文全部引文目录的出版商网页, 大多数情况下, 也包括引文的文摘。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, 研究者可

以直接通过适当的机制获取全文, 也可能通过订购、文献传递、按浏览付费或者预印本付费等获得。如读者所在机构已订购并被授权使用某数据库, 则可通过授权的合法 IP 地址登陆此数据库网址, 引文与全文的链接将非常迅速、直接。

## 2.3 CrossRef 功能

CrossRef 在学术领域取得了显著成就, 其功能主要表现在以下几个方面:

2.3.1 网上固定识别的功能, 使数据不致由于网域改变的异动而消失。

网络链接的一大问题就是网址不固定。即使在同一网站, 一篇文章日久以后也会移到典藏目录, 这对本网站的读者或许没有什么影响, 但别的网站对这篇文章的链接, 就立刻断了线。CrossRef 成员出版商只需要在 MDDB 中 DOI 目录中更改文献的 URL 即可。一旦出版商改变它的 URL, 只要在中心 DOI 目录中更新, 每一个与之相关的 DOI 会自动更改。一些成员出版商和相关团体也借着身为系统管理人之便, 随时得到最新版本的当地 MDDB 副本, 以便实现更新链接。这样, 维系了数据在网上的固定性<sup>[6]</sup>。

2.3.2 传递既能够方便出版商又能够提高用户的搜索经验。

DOI 简单提供用户一个 URL, 指示链接到出版商的网址, 出版商对用户来源, 可以进入的条件几乎一无所知, 一些参加 CrossRef 的出版商当然希望能够了解转回链接到他们的链接对象的情况。参数传递可以解决这个问题, 这对于出版商和用户都是有用的。参数传递的可能是有关源文献的信息, 即用户点击的包括参考文献的文献信息。出版商接收到此参数, 可以知道确切的论文和出版商资料, 以及用户来源。这一传递主要是由 DOI 系统来解决。通过 DOI 链接, 出版商可以跟踪到达本网站的链接, 可以得知哪些期刊论文与本网站的论文相链接。参数传递也可以提高最终用户的经验, 如利用“后退”按钮、记忆页面、订购信息或者错误处理, 以及一些商业协议, 如允许来自特定网站或者期刊的用户即时获得全文<sup>[9]</sup>。

2.3.3 有助于著作者版权正当的保护<sup>[2,5]</sup>。

由前面对 CrossRef 的描述可以得知, CrossRef 维持一个文件的集中数据库, 仅存描述文件的 MetaData, 而不储存文件内容, 使用者只能获得文件

的基本资料,不能获取文件原文,这样,可以避免随意下载,保护著作者版权。在 CrossRef 中有一组“加强分辨”功能,它包括复合分辨(multiple resolution)和返回元数据查找(reverse metadata look-up)。复合分辨使得符合因素和 DOI 链接。在这种情况下,引文查找结果就包括所有可能和 DOI 链接事项的元数据。这些事项可能是附加的分散在各镜像网站的 URL,一个 E-MAIL 地址,一个指向某个元数据记录的指针。所谓返回元数据查找是指查找结果返回输入时给定有关论文完整元数据的 DOI。这样允许图书馆和其他信息服务机构通过本地的链接服务为自己的读者创建预定的链接,为读者提供适当的资源。这是解决合理的版权问题的有效方案,也是本地链接服务的主要目的。通过本地链接,能够确保既定用户不能直接通过对外付费获得所在机构已经收藏的论文,而通过向机构预约电子版或者印刷版来获得。

### 3 CrossRef 引发的思考

CrossRef 为引文与本文、引文与引文之间链接提供了便利。对于资源共享、著作权保护等热点问题的解决提供了有效途径。CrossRef 成员由学术出版商发起,提供的也只是学术期刊之间的链接,扩展到现在会议论文、网络 DTD 文献等。CrossRef 起初为出版商和用户之间的链接提供了有效途径。但随着越来越多的出版商的加入和二级出版商、图书馆的加入,如何更方便快捷地为最终用户提供引文原文又成为一个问题摆在众多信息服务机构的面前。对 CrossRef 引发的思考主要有三个方面:

#### 3.1 用户

在这之前主要是索引文摘型数据库、全文数据库提供的数据为用户提供了许多方便,但索引文摘和全文并不能使研究者的研究信息资源完整化。CrossRef 利用 DOI,让用户通过原始文献轻松地找到其引用文献全文,揭示知识的起源,学术的关联。用户在利用信息资源时,要注意以下几个方面:(1)要尽可能利用参加了 CrossRef 出版商出版的在线期刊文献,例如以下出版商出版的期刊文献:Academic Press, the American Association for the Advancement of Science (《科学》杂志出版社); American Institute of Physics(AIP); Association for Computing Machinery (ACM); Blackwell Science;

Elsevier Science; The Institute of Electrical and Electronics Engineers, Inc. (IEEE); Kluwer Academic Publishers; Nature; Oxford University Press; Springer-Verlag; John Wiley & Sons; EBSCO 等等<sup>[2-4]</sup>。(2)尽可能地向参加了 CrossRef 的信息咨询机构寻求咨询服务,这些机构可以通过 OpenURL 实现本地链接,节省用户的费用。OpenURL 是在 SFX 链接基础上发展起来的基于数字对象唯一标识符和元数据的虚拟 URL,可用于建立一种动态的、可本地控制的、可选择的信息资源链接机制,目前许多已经参加了 CrossRef 的图书馆和图书情报信息服务机构,利用 OpenURL 和 SFX 软件,构建了解决合理著作权问题的原型方案。通过这个方案,利用 DOI 可以直接链接到本地收藏文献<sup>[7-9]</sup>。

#### 3.2 出版商

对于出版商而言,CrossRef 为其提供的机遇和挑战主要有以下方面:(1)为了使电子出版品能够容易被使用,出版商必须改变出版品的技术与方法,让所属的电子出版品用 DOI 标识,加入 CrossRef 的服务,扩大其使用范围,达到资源的充分共享。同时也可以提高本出版单位在整个信息服务界的知名度。对于以前出版的期刊文献,为了扩大正文到引文原文的链接范围,也应该采用上述方法加以电子化。据悉,目前在 CrossRef MDDb 中存储最早的是 1849 年《天文学志(Astronomical Journal)》上的文章<sup>[2,4]</sup>。(2)出版商在提供自己的电子期刊时,要注意文献参考时的便利性。对于 URL 改变了的期刊文献,要及时在 DOI 中加以更正,以便向用户提供准确的信息。出版商在注意保护著作权的同时,要尽可能地向最终用户提供全面的引文。据 Amy Brand 介绍,目前,在相关记录中只有 40% 的查询结果可以获取全文,主要是因为所查找的全文链接在出版商和其他机构中不能获得<sup>[5]</sup>。因此,出版商要注意加强数据的存储,不断满足用户的需求。

#### 3.3 图书馆和其他服务机构<sup>[9-11]</sup>

随着信息技术的发展,图书馆及其他信息服务机构朝着信息产业化、技术自动化、馆藏数字化、通信网络化和手段多媒体化方向发展。除了有基本的硬件外,机构的软件技术也应该跟上时代的发展,通过 CrossRef,图书馆和其他信息服务机构应注重的问題主要有以下几点:(1)购买的电子资源是否采用了 DOI 技术。面临电子资源的迅速变化,用

户对全文电子文献需求的扩大和要求的提高,图书馆和其他信息服务机构必须谨慎地选择电子馆藏。图书馆在选择合理的电子馆藏时,应考虑购买的电子资源是否采用了 DOI 技术,从前面对 CrossRef 的描述可知,采用了这一技术的电子资源对机构自身的发展而言具有重要意义。随着 CrossRef 成员的不断扩大,DOI 技术的广泛应用,很多机构利用 DOI 技术,建立 DOI 代理服务(proxy),通过 OpenURL 和 CrossRefMDDb 链接,能够实现本地馆藏文献的链接,为用户提供方便快捷的信息服务<sup>[8]</sup>。(2)出版商是否为 CrossRef 成员。参加了 CrossRef 的出版商提供的电子文献能够提供从引文到原文的链接,方便本机构开发本地电子资源,上述已经讲述得很清楚了,在这里不再赘述。用户在使用 CrossRefMDDb 时,不需要付费,就可以查到 DOI,通过一定的协议,付少量的费用,就可以直接下载,或者通过订购、文献传递或付费来获取预印本<sup>[7]</sup>。

#### 4 结束语

DOI 技术既是参考文献标注的一项革新,更将为科技文献的传播方式(包括发行方式)带来革命性的影响,对目前二次文献数据库在期刊文献检索利用方面的中心地位也将起到补充和平衡的作用。目前,国外的许多出版商纷纷加入 CrossRef, CrossRef 也在不断地提高其性能,扩大其影响。为用户实现引文到原文的链接提供了便利,为文献资源真正实现共享提供了有利途径。但是我们要看到,目前提供引文到原文的链接,仅仅局限在英文文献,对于利用广泛、价值高的中文文献,我们不妨也考虑利用 DOI 技术,出版商加入 CrossRef,图书馆和其他信息服务机构设置 DOI 代理,利用 OpenURL 和 SFX<sup>[12,13]</sup>,实现本地馆藏链接,有效地保护著作

权。使国内研究文献及其成果同样能够站在国际舞台上,与世界共享。

#### 参考文献

- 1 张晓林. 开放数字环境下的参考文献链接. 现代图书情报技术, 2002(1):9~13
- 2 <http://www.crossref.org/ldl.html>
- 3 EBSCO licenses CrossRef database. Advanced Technology Libraries, 2001, 30 (6):5~6
- 4 Pentz E. Reference linking with CrossRef. Interlending and Document Supply, 2001, 29 (1):20~2
- 5 Brand A. CrossRef turns one. D-Lib Magazine, 2001, 7 (5). <http://www.dlib.org/dlib/may01/brand/05brand.html>
- 6 Pentz E. CrossRef: a collaborative linking network. Science and Technology Librarianship, 2001, (29) Winter. <http://www.library.ucsb.edu/istl/01-winter/article1.html>
- 7 Oren Beit-Arie. CrossRef /DOI/OpenURL/SFX demo using IDEL metadatabase & enhanced DOI proxy server. <http://sfxserv.rug.ac.be:8888/public/xref/crossref-examples.htm>
- 8 Van de Sompel H, Beit-Arie, O. Open linking in the scholarly information environment. Using the OpenURL Framework. D-Lib Magazine. 2001. <http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html>
- 9 OCLC offers OpenURL option for libraries. Advanced Technology Libraries, 2001, 30 (12):4,5
- 10 CrossRef steps up pace of linking developments. Information Today, 2001, 18(10):53~56
- 11 CrossRef access expands ovid's linking. Computers in Libraries, 2001, 21(1):16~21
- 12 Walker J. What is SFX? Learned Publishing, 2001, 14 (4): 296~298
- 13 Gale adopts. SFX open URL standard. Advanced Technology Libraries, 2001, 30 (11):8

作者单位:浙江工业大学图书馆,杭州,310032

收稿日期:2002年7月30日

### CrossRef: a Collaborative Linking of Electronic Document between Content and Reference Across the Scholarly Realm

He Limin, Wan Yuehua, Li Ping

**Abstract:** CrossRef is a process to be used for linking information on Internet and provide information users access to the target and correct information resources. This paper introduces the development and functions of CrossRef and discusses the influence on information users, publishers, libraries and other information providers of CrossRef. It suggests that it is necessary for Chinese information resources to apply DOI, CrossRef, OpenURL and SFX to provide library's local holdings and solve the appropriated copy problem.

**Keywords:** CrossRef, DOI, Reference Linking