

CrossRef 的新进展

卜 今

(南京医科大学学报编辑部)

CrossRef 是国际 Digital Object Identifier (数字对象标志符, DOI)的官方注册代理公司, 它于 2000 年初成立, 其成立源于多家出版商给在线出版的电子学术期刊永久的引用链接的合作项目。CrossRef 由出版商创建和管理, 是一个独立会员制、非赢利性的机构, 目前大概有 2000 家出版商加入, 代表了近 5000 个协会和出版单位, 包含了商业和非营利出版集团。目前 CrossRef 提供的主要服务包括:

1 Digital Object Identifier (数字对象标志符, DOI)

DOI是在线出版科技文献的数字识别标志^[1], CrossRef给会员提供的每一篇科技文献分配一个DOI。文献拥有唯一对应的DOI号, 是实现文献追踪引用的基础^[2]。而且DOI标志符是匹配注册者在注册时提交的元数据, 意味着即便特定项目的所有者改变, 其标志符也是永恒不变的, 这就是DOI被称为“永久标志符”的原因。DOI编号方式遵循ANSI/NISO Z39.84.2000的语法标准, 由前缀和后缀两部分组成, 中间被向前的斜杠分开。如 10.7555/JBR.27.2013.0008, 其中“10.7555”是前缀, 由CrossRef分配给会员, “JBR.27.2013.0008”是后缀, 由会员按照一定规律编码。对于前缀和后缀的长度, 没有技术限制。目前, 由CrossRef分配获得DOI号的科技文献已经超过600万, 而且在持续增长中。可以获得DOI号的科技文献包括期刊文献、书籍、会议录、研究手稿、技术报告和数据集。所链接的内容包括科学、技术和医学(Scientific, Technical and Medical, STM)以及社会科学和人类学分支来源的材料。CrossRef并不提供科技文献全文的数据, 它所提供的是隶属于其他网站的科技文献内容所拥有的一个链接, 而这个链接就是通过永久的DOI来实现的。DOI申请所涉及费用由参加的会员出版商支付。

2 Cited-by-Linking^[3]

Cited-by-Linking 始于 2004 年, 目前已有近 300 家出版商加入。此项服务提供会员出版商出版的科技文献被引用的情况并且将信息整合到会员期刊的在线发表平台上, 出版商可以清楚知道谁引用了它们的文献。Cited-by-linking 的前身是 Forward linking, 是 CrossRef 公司提供给会员的一项专项服务, 目前是自愿参与而且是免费。但是这项免费服务有一个前提条件: 会员必须上传自己所引用的文献清单。不过由于 Cited-by-linking 直接接受此项服务的 CrossRef 会员数限制, 意味着“Cited by”信息在不同学科和出版商之间差异很大。和其它

公司如 Google Scholar、Scopus、ISI 的 Cited-by 引文服务相比，CrossRef 的 Cited-by-linking 的优势在于：（1）它受 CrossRef 会员制度控制和管理，因此，是一个从出版商到出版商的直接联系，不需要使用其它媒介；（2）它不和任何评价指标绑定，虽然这些数据可以被用于使已经存在的评价标准合法化或者制订新标准的基础；（3）它不受限于学科，可以提供相对难度较大的学科间的引用信息，或者为那些享受不到良好服务的出版商提供服务；（4）它不受限于发表形式，CrossRef 接受期刊文献、单个段落或者书等多种出版形式的参考文献数据；（5）CrossRef 的 Cited-by-linking 引文服务是建立在 DOI 数据库的基础上的，是非常精确的；（6）CrossRef 作为收集权威出版商文献的数据库，正被越来越多的人关注。目前约有 40% 的 CrossRef 会员出版商加入 Cited-by-linking。截止到 2013 年 7 月 24 日，在 CrossRef 上传的总文献数是 19,938,741，至少有一次 Cited-by-linking 的文献有 25,185,237，总的 Cited-by-linking 次数是 351,090,039。

3 CrossRef Metadata Services（元数据服务，CMS）^[4]

CMS 是结构型的、跨出版商的用于收录和链接的元数据服务，已集成有 5 000 多万个数据。取得 CMS 授权的机构可跨出版商抓取、收集元数据，进行索引和链接等。CMS 包括不同元数据的传播工具，并提供给会员标准化的条件，供其使用由 CrossRef 提供的跨出版商的元数据。CMS 有两种方式：Basic CMS 和 Enhanced CMS。

3.1 Enhanced CMS

授权单位能够通过 Open Archives Initiative Protocol for Metadata Harvesting（开放信息库计划元数据集成，OAI-PMH）接口使用在 CrossRef 注册的超过 5 000 万条元数据记录。

CrossRef OAI-PMH 接口库服务是会员出版商元数据扩散的核心技术，它其实是利用了一项以拥有大量元数据用户为目标的成熟技术，对接口数据库的连接是通过 IP 控制的，并且能够提供从所选出版商到授权接收者的特殊追踪信息。

3.2 Basic CMS

任何需要通过 DOIs 连接科技全文，在所在单位使用 CrossRef 数据库的非出版机构都可以申请成为 Basic CMS 的授权单位。Basic CMS 的授权单位可以在机构内部使用 CrossRef 的书目元数据和 DOI 系统。所有数据是以 XML 格式呈现的，授权单位必须有这个技术能力解析 XML 文件。

4 CrossCheck^[5]

CrossCheck 始于 2008 年，是 CorssRef 向会员出版商提供的一项反剽窃检测服务，已有 300 多家出版商加入。CrossCheck 是美国 iThenticate 公司开发的反剽窃软件，它由 CrossRef

最早开始应用。和其它英文的比对软件相比，CrossCheck 的优势在于它的比对数据库巨大，包括了超过 600 万个网页信息，而且在持续更新中。CrossCheck 的数据库包含两部分，其中一部分是 iThenticate 工具，它会比对投稿作者的稿件和目前数据库内的文献，最后报告给用户可疑论文与数据库中已发表文献的相似度，以百分比表示，并将相似的文本标示出来供编辑部分分析参考。CrossRef 鼓励它的会员把它们的文献充实到 CrossCheck 的数据库内，因为数据库的扩大对使用 CrossCheck 的出版单位会越来越有利。

5 CrossMark update identification service (CrossMark 更新鉴定服务)^[6]

CrossMark 始于 2012 年 4 月，它提供版本状态声明 (Status) 以及出版记录 (Record) 两方面信息，从而使读者或研究者能够及时地发现科技文献发表以后的更正、更新、撤稿等重要改变以及文献的出版过程。目前已有 6 万多条文献打上了 CrossMark 的标志。读者只要点击 PDF 或者是 HTML 上的 CrossMark 的 Status 标签，就会有提示框告之目前的文件版本是不是最新的或者最新的版本在哪里可以被获取(通过 DOI 号的链接)。而出版记录(Record) 里面记录了如同行评审过程、出版历史、剽窃检测状态、基金信息、利益冲突声明、版权声明等重要信息。CrossMark 的服务由加入 CrossMark 的出版商付费，对读者是免费的。

6 FundRef (基金鉴定服务)^[7]

FundRef 的基金资助人鉴定服务提供了一条报道研究经费来源的标准化途径，目的是促使学术出版单位和基金代理机构之间的合作。FundRef 的数据库目前包含了两大部分：(1) 基金资助人信息，在 2012—2013 年试点阶段加入项目的基金机构以及代理机构的信息；(2) FUNDREF 注册表，该注册表包括了 4 000 个基金的分类目次，包括名称、缩写和 Elsevier 贡献的替代性名称。通过 FundRef 的操作流程，出版商将 FundRef 注册表整合到出版投稿系统，作者在投稿之前选择基金项目名称和编号，基金资助项目的信息传送到出版商的出版发行系统内，出版商将基金信息上传给 CrossRef。基金资助人和其他机构查询 CrossRef 即获得他们所资助研究论文的 DOI 和其他材料；同时出版商可以在 CrossMark 的出版记录信息栏中标注论文基金的资助信息。

7 Open Researcher and Contributor ID (开放科研者与贡献者身份标识项目，ORCID)^[8]

ORCID 始于 2012 年 10 月，和 DOI 类似，不过对象是研究者而不是科技文献。ORCID 给研究者分配一个永久的数字识别符号，主要目的是规范学术出版物的作者姓名，因为研究者有可能重名，而 ORCID 是唯一的。研究者一旦在 ORCID 系统内注册并获得 ORCID 号后，

今后所有的科研活动包括发表文章、申请基金等都可以标注上他自己唯一的 ORCID 号，从而和其他研究者区别开来。

8 CrossRef 对小型出版商、单个期刊出版者的作用

CrossRef 通过 DOI 号等技术实现对元数据的跨出版商引用和链接，从而使得独立分散的出版商数据整合到一个共同的信息平台上，它的核心技术不需要双边签署同意协议，因此保证了任何加入 CrossRef 的小出版商或者单个期刊出版商都可以分享在 CrossRef 平台上的所有数据信息，甚至是单个的科技文献都可以通过 DOI 号实现在 CrossRef 数据平台上的标引和链接。相比大型出版集团来说，没有能力获取更高关注度的小型出版商获益更多。

参考文献

- [1] http://help.crossref.org/#what_is_a_doi
- [2] http://help.crossref.org/#how_do_dois_work
- [3] <http://www.crossref.org/citedby/index.html>
- [4] <http://www.crossref.org/cms/index.html>
- [5] <http://www.crossref.org/crosscheck/index.html>
- [6] <http://www.crossref.org/crossmark/index.html>
- [7] <http://www.crossref.org/fundref/index.html>
- [8] <http://www.crossref.org/01company/orcid.html>