

DOI 在科技信息资源搜索 与利用中的应用

田杰¹ 张旭¹ 庄绍晗²

(1. 中国科学技术信息研究所, 北京 100038;

2. 国家知识产权局, 北京 100088)

摘要:在数字环境下,科学家做研究和研究成果交流的方式都发生了很大变化,技术发展和观念更新给科技信息资源共享同时带来机遇和挑战。在数字环境下,科技信息资源具有分布广泛、数量巨大、增长迅速、内容丰富、形式多样等特征。与此同时,网络环境下科研工作者却表现出了强烈的对更高层次信息的需求,例如对西文文献、原始科学数据、图片、视频、音频的直接获取等。因此,如何在网络环境下解决这一矛盾已成为信息科学的重要任务。本文拟通过数字对象唯一标识符(DOI)在科技文献与科学数据网络获取的应用实例来探讨相关的科技信息资源共享解决方案。

关键词:科技信息资源;搜索与利用;DOI;共享

中图分类号: G20 **文献标识码:** A **DOI:** 10.3772/j.issn.1674-1544.2008.04.011

Application of DOI in Searching and Using of Scientific Information Resources

Tian Jie¹, Zhang Xu¹, Zhuang Shaohan²

(1. Institute of Scientific and Technical Information of China, Beijing 100038,

2. State Intellectual Property Office of the People's Republic of China, Beijing 100088)

Abstract: In the digital age, the way scientists do research and exchange research results has great changes, and the sharing of scientific information resources is facing opportunities but also challenges by technological developments and concepts' updates. In a digital environment, scientific information resources are widely distributed, in huge number, in rapid growth and with rich content and diverse forms. Meanwhile, scientific researchers have expressed strong demand of a higher level of information supply, for instance, direct access to western languages literature, original scientific data, images, video and audio frequency. Therefore, how to solve the problem in the network environment has become a vital mission for scientific information workers. This paper tries to discuss relevant solution of sharing scientific information by demonstrating the application case of Digital Object Identifier (DOI) in the network access of scientific literature and scientific data.

Keywords: scientific information resources, searching and using, DOI, sharing

第一作者简介:田杰(1978-),男,学士,研究方向是科技资源共享模式与网络信息共享技术。

收稿日期:2008年6月3日。

早在1997年第一次全国信息化工作会议上,原国务院信息化工作领导小组就颁发了《国家信息化规划纲要》,确认了信息资源开发利用在信息化工作中的核心地位。近年来,我国针对科技信息资源的开发利用已引起人们的高度重视,一批重大的科技信息资源共享项目,如科学数据共享工程、科技基础条件平台建设等工作陆续启动;因特网上的中文科技信息资源逐渐丰富起来,增长速度迅猛。虽然我国科技信息化建设工作取得了长足的进展,但大多仍是以行业、部门为主,各自为战,目前尚未形成统一的开发标准。数据结构、信息编码自成体系,各类信息管理软件的通用性、适用性较差。信息化建设中存在着“重硬件、轻软件”以及“重视技术性实施建设、忽视信息标准化工作”的倾向。由于信息统一规划工作相对滞后,导致信息标准匮乏、信息处理工作不规范,从而难于实现信息交换和信息共享,严重影响了信息资源的充分利用和信息化建设的进程,成为社会经济、政治、文化进一步发展的主要瓶颈。

1 DOI 的特点及其作用

在数字环境下,科技信息资源具有分布广泛、数量巨大、增长迅速、内容丰富、形式多样等特征。与此同时,网络环境下科研工作者却表现出了强烈的对更高层次信息的需求,例如对西文文献、原始科学数据、图片、视频、音频的直接获取等。因此,如何在网络环境下解决这一矛盾已成为信息科学的重要任务。作为科技信息资源整合与互链的重要标准之一,数字对象唯一标识(Digital Object Identifier, DOI)是一个专门用来标识数字化对象的命名系统。它由 International DOI Foundation 组织负责运作和管理,目标就是方便网络上的作品识别和获取,目前,国外大型数字文献生产商大多使用 DOI 对数字资源进行标识,包括 Elsevier、Blackwell、John Wiley、Springer 等众多厂商^[1]。DOI 已逐步发展成国际上最为通用的科技文献标识与链接标准。DOI 具有以下 7 个特点^[2-5]。

(1) 唯一性: DOI 是逻辑标识符,独立于数字对象的物理位置、复本、应用软件和存取协议,每

个标识符标识且仅标识一个数字对象。

(2) 永久性: DOI 是永久的,尽管数字对象可以被更改或删除。

(3) 通用性: DOI 作为一个公共标识符,在国际范围内或整个因特网上都有效。

(4) 可操作性: 通过 DOI 可以实现数字对象与生成地址及其他应用的解析。

(5) 互操作性: 兼容现有的和将来可能产生的相关标识符,以及其他网络协议和标准。

(6) 可扩展性: 可以标识不同粒度的信息。

(7) 可读性: DOI 能被人 and 机器识别,易于书写和记忆。

DOI 的发展为科技信息资源的搜索与利用在信息共享底层实现了标准规范与链接协议。具体表现如下。

(1) 学术出版商可通过 DOI 系统实现引文到全文的链接,实现一站式服务。

(2) 图书馆通过在本地导入 DOI 并与 OpenURL 结合,从而为科研工作者提供了访问更多的图书馆未能收藏的全文文献的机会,并可以更加有效地管理现有资源,开发更为高效的检索和馆藏功能。

(3) 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。

(4) 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 搜索保证了检索结果到全文链接的有效性和持久性。

(5) 政府机构: 相关政府机构通过 DOI 的推广和应用可以更加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科技基础信息服务平台。

2 DOI 在国外的应用

CrossRef 起源于由美国出版者协会 (AAP)、

国际数字对象标识符基金会 (IDF)、美国国家研究创新公司 (CNRI) 以及多家出版商共同参与研究与开发的 DOI - X 实验项目。DOI - X 项目的成功运转最终促成这些世界主要科学、技术及医学 (STM) 出版商于 2000 年初组成了一个新的非盈利性组织, 即国际出版商链接联合会 (PILA), 在 International DOI Foundation 的管理下, 建立并负责运作在该原型系统的基础上发展起来的合作参考链接服务即 CrossRef。作为第一个合作式参考链接服务, CrossRef 于 2001 年 6 月开始正式运作。CrossRef 的任务是成为所有电子学术信息参考链接的中枢 (Backbone) 与基础 (Infrastructure)。就像一个数字化的交换机, CrossRef 不掌握全文信息, 但它基于 DOI 的链接形成一个高效的、可扩展的链接系统, 通过这个系统, 研究者们可以方便地实现资源搜索与利用。目前, 西方大量的学术期刊出版成为了 CrossRef (西文 DOI 代理注册机构) 的会员, 截止到 2008 年 5 月 22 日, CrossRef 覆盖了全球 20623 种西文期刊。每个月全球读者通过 DOI 进行文献获取的次数达到了 2000 余万次, 极大地降低了科研工作者获取科技文献的难度, 扩大了科技文献的传播范围与传播速度^[6]。

同时, 国外相关机构还在积极探索 DOI 在科学数据获取上的应用, 并取得了一定的成效^[7]。CrossRef 开始在生物数据库中应用 DOI 技术, 对蛋白质结构进行标识, 并且规定只有将数据进行标识后方能传到蛋白质数据库, 才能发表相关文章。同时根据 CrossRef 参考文献的 DOI 元数据提出了科学数据 DOI 元数据的参考框架, 但暂时还不能对数据 DOI 进行解析和查询。SDDB (ICDP Scientific Drilling Database) 在网上发布科学钻探工程的数据, 其标识对象是一次钻探实验形成的钻孔数据的集合, 同时提供数据标识规范, 能对标准引用数据进行下载, 便于用户引用, 其钻点数据能在 Google Earth 中进行展示。

德国国家科学技术图书馆 (TIB) 是世界上首个将 DOI 应用在科学数据中的机构, 2005 年正式成为科学数据的 DOI 注册中心 (RA for Scientific Dataset), 截止到 2007 年 10 月, TIB 已经注册了 475276 个数据集, 12546 个科学电影剪辑, 6302

个医学案例, 342 个技术报告和 112 个学习对象。未来科学内容将是 TIB 的主要工作, 它的注册对象包括各种不同的内容类型, 如晶体结构、地球模型和 3 - D 模型等。

3 DOI 在我国的发展现状

DOI 在国外的成功, 尤其是在科技文献领域的成功, 充分表明了全球科研工作者对资源互联与共享的巨大需求。但是截止到 2006 年底, 我国只有大约 100 种英文学术期刊, 例如《中国科学》、《科学通报》英文版等通过国外注册机构为科技论文标识了 DOI, 实现了少量国内优质英文版文献与国外英文文献间的互链。据了解, 高等教育出版社的前沿学术出版系列刊物在标识 DOI 后点击率明显提升, 目前生物卷月点击率已经接近 1 万次/月。通过 DOI 这种方便的获取方式, 客观上促进了我国科技文献在全球的学术影响力, 但是由于前几年高昂的注册费用和英文元数据的格式要求, 抑制了大量的中文科技文献对 DOI 注册的需求, 目前通过国外的代理机构注册 DOI 的会员费最低标准为 275 美元/年, 每注册一个 DOI 标识符还需要支付 1 美元, 假设一本学术期刊年发表 300 篇论文, 那么为此需要交纳的年费用为 575 美元, 折合人民币约为 4000 元^[8]。

可喜的是, 建设中文 DOI 系统的重要意义已经得到了国内科研机构的广泛关注。自 2001 年起, 中国科学院文献情报中心、国家图书馆等机构开始关注和研究 DOI 在中文信息资源方面的应用, 并承担了科技部“数字对象唯一标识”课题研究, 为中文 DOI 的推广和应用奠定了基础。2007 年, 经与 IDF 深入沟通, 中国科学技术信息研究所联合北京万方数据股份有限公司开始提供中文 DOI 注册服务, 建立了类似 CrossRef 系统的服务平台 (www.chinadoi.cn), 并且制定了为广大中文学术期刊杂志社所能接受的服务价格, 目前会员费为 500 元/年, 每注册一个 DOI 标识符仅收取 1 元人民币^[9], 同样一个年发表 300 篇论文的中文期刊所交纳的年费用仅为 800 元人民币, 基本上等同于一篇学术论文的版面费。截止到 2008 年 5 月, 中文 DOI 注册与资源互联服务

中心已经为科技部精品期刊项目组所收录的300种精品期刊、中华医学会的115种医学期刊、《浙江大学学报》(6个版)、《植物生态学报》、《数学研究与评论》、《舰船科学技术》、《测绘科学》、《长江蔬菜》等多个领域的学术期刊开展了DOI注册与链接服务,目前总计注册学术论文80余万篇,并成功注册了5000余条科学数据。另外,中文DOI注册与资源互联服务中心与NSTL合作,开展了数百万条的英文引文DOI匹配工作。

但是,笔者认为,DOI在中文科技信息领域的应用还处在初级阶段,目前仍然有许多问题需要解决,主要表现为以下几点^[10]。

首先,国内大量的学术期刊编辑部、杂志社对DOI的概念和作用还不是十分理解,对DOI可以用来揭示有关数字对象的信息,可以方便、持久地获取科技文献,尤其是可以无技术障碍地实现跨地域、跨语言、跨数据结构、跨文献种类的无缝链接价值还没有充分认识,导致国内大量的拥有科技资源的源头单位未能采用DOI标准与链接服务协议。

其次,国内学术出版的源头——编辑部、杂志社经营规模普遍较小,并且80%左右尚未建立自己的网络发行平台,无法像国外一样快速地形成聚合效益。

最后,国内学术出版单位和科技信息资源产生单位对于有偿注册DOI和接受DOI链接服务还存在观念转变障碍。

但是,笔者相信,随着信息技术的发展

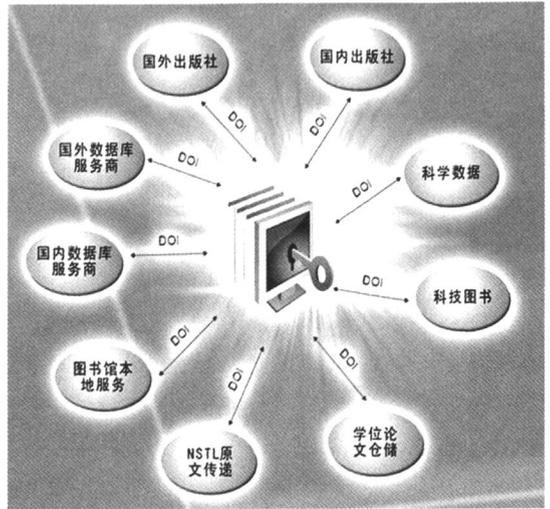


图1 DOI在科技信息资源发现与利用的应用远景

和人们对数字资源需求的进一步加大,一个规范化、标准化、可操作化、方便化的DOI系统必将在我国获得蓬勃的发展,也必将推动我国与国际的学术交流与资源共享。从而实现基于DOI为链条的中西文学术论文、学位论文、科学数据、科技图书间的无缝链接与获取,如图1所示。

4 通过DOI搜索科技文献的方法

通过DOI查找科技文献的方法很简单,目前主要有以下几种方式:第一种方式是通过一些地

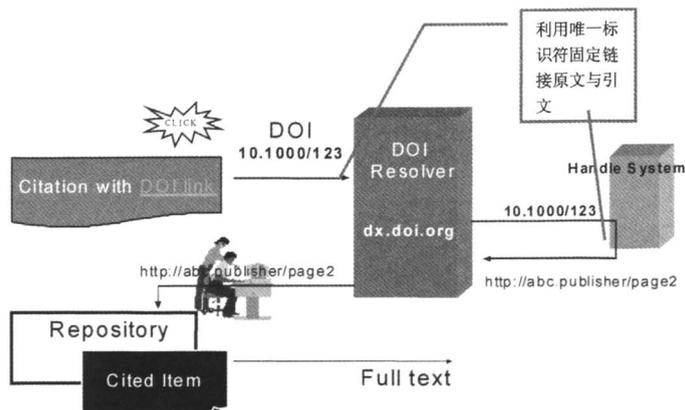


图2 通过DOI获取文献的工作流程

科学通报 2008 53 (8): 857-864 ISSN:
0023-074X CN: 11-1784/N 首页 当期目录 上一期

基于表面等离子体共振效应的显示技术

曹振新, 吴乐南

东南大学毫米波国家重点实验室, 南京 210096

收稿日期 2007-12-27 修回日期 网络版发布日期 接受日期 2008-3-23

摘要 表面等离子体共振构件中, 调制层材料的实部与共振波长存在单调函数关系, 调制层材料的虚部与共振强度存在单调函数关系。基于该两种现象可以设计出新型的显示器件。首先从理论上分别提出了波长和强度可控的显示单元, 而后提出了波长和强度可联合调控的显示单元, 进一步研究了具有 8×8 个显示单元的器件, 并给出仿真结果。研究表明, 基于表面等离子体共振效应的显示器件可以直接控制每个显示单元的颜色和亮度, 而无需通过三基色合成, 具有分辨率高、对比度高、平面显示、高亮度、响应快等突出优点。然而, 如果需要设计成实用的显示器件, 尚需深入研究调制层材料和工艺等。

关键词 表面等离子体共振效应 显示 自然色彩
分类号

DOI: 10.3573/j.0023-074X.2008.02.008

对应的英文版文章:

通讯作者:

曹振新 caozx@seu.edu.cn

作者个人主页: 曹振新; 吴乐南

参考文献

- 1 王艳霞. 基于 SPR 技术的显示器的研究. 硕士学位论文. 南京 [DOI]
- 2 Wang Y. Wavelength selection with coupled surface plasmon—4387 [DOI]
- 3 Wang Y. Voltage-induced color-selective absorption with surface plasmons. Appl Phys Lett, 1995, 67(19): 2759—2761 [DOI]
- 4 Wang Y. Surface plasmon high efficiency projection display. Proc SPIE, 1997, 301
- 5 Wang Y, Russell S B, Shimabukuro R L. Electronically tunable mirror with surface SPIE, 1998, 3292: 103—106 [DOI]
- 6 曹振新. 基于表面等离子体共振效应的信息获取和显示. 博士学位论文. 南京: 东南大 85—97 [DOI]
- 7 谷千东, 著, 金珍裕, 译. 先进显示器技术. 北京: 科学出版社, 2002. 54—60

首先为文章标识 DOI, 供其他出版物 (中英文期刊、图书), 数据库 (中英文数据库、文摘数据库、全文数据库), 图书馆以及开放存取平台等标识本论文, 并通过 DOI 链接到一个最方便读者获取原文的网络存储地址。

其次为引文标识 DOI, 方便研究者在本篇论文中直接获取有价值的信息, 从而增加论文被发现与利用的机会

通过 DOI 链接到西文, 既简单又方便, 同理也可从西文论文链接到中文参考文献的全文, 方便外国研究者获取中文信息

图 3 通过 DOI 搜索与利用科技文献的实例

址由 DOI 系统自动帮助查找。在 <http://dx.DOI.org/> 的“Resolve a DOI”的提示框内输入已知 DOI 标识符, 点击“Go”按钮, DOI 系统就会自动链接到该文献的 URL, 并显示相应的文摘或全文页面。另外, www.CrossRef.org 网站和 www.Chinadot.cn 网站也提供了类似的解析链接功能。第二种方式是只要将已知的 DOI 标识符前加上 <http://dx.DOI.org/> 就可获得该文献在互联网上的存储地址。例如, 已知某文献的 DOI 标识符为:

10.1103/PhysRevLett.95.253601, 则在 IE 浏览器中输入 <http://dx.DOI.org/10.1103/PhysRevLett.95.253601>, 就可获得该篇文献的文摘或全文。第三种方式最为方便且已经为国际上的主流出版商、文摘服务商、订阅服务商、集成服务平台、搜索引擎和数字图书馆等采用, 科研工作者只需轻轻点击 DOI 按钮就可直接链接到相应的科技文献中去, 这是因为 DOI 系统中存储了每一篇已经标识了 DOI 标识符的文献 URL 地址(图2)。

图3描述了一个具体的通过DOI来获取科技文献的应用场景。在《科学通报》2008年第二期发表的一篇题为《基于表面等离子体共振效应的显示技术》论文,该篇文献的DOI标识符为:10.3575/j.0023-074X2008.02.008,读者可通过上述3种方式快速地定位到这篇文献的网络存储地址,从而获得全文,进而读者还可以这篇文献为基点,通过DOI获取相应的中西文参考文献,而无须进行跨平台检索。

参考文献

- [1]姚长青,乔晓东,赵蕴华,王子祥,孙建. DOI在精品科技期刊全文数据库中的应用[J]. 中国科技资源导刊, 2008(1):13-15,21.
- [2]张晓林. 数字对象的唯一标识符技术[J]. 现代图书情报技术,2001(3):8-11.
- [3]覃嵘嵘,等. 开放式结构数字图书馆的Handle System

- 及其实现[J]. 小型微型计算机系统, 2000(10): 1103-1107.
- [4]毛伟,等. 一种互联网资源标志与寻址技术:HandleSystem[J]. 计算机应用研究,2004(5):252-254.
- [5]李广建,黄岚. 数字对象唯一标识Handle System[J]. 图书馆建设,2004(3):20-23.
- [6]Crossref.org. CrossRef_indicators[EB/OL]. [2008-06-08]. http://www.CrossRef.org/01company/CrossRef_indicators.html.
- [7]彭洁,涂勇. DOI在科学数据领域的应用研究[J]. 前沿探索,2007(2):27-28.
- [8]Crossref.org. publishersfees[EB/OL]. [2008-06-08]. http://www.CrossRef.org/02publishers/20pub_fees.html.
- [9]出版社费用. 中文DOI注册与资源互联服务中心[EB/OL]. [2008-06-08]. http://www.chinadoi.cn/publisher_2.htm.
- [10]田杰. 中文DOI运营机制初探[J]. 前沿探索,2007(2):20-22.