

数字对象唯一标识在中国科学数据领域中的应用研究*

□ 涂勇 彭洁 / 中国科学技术信息研究所 北京 100038

摘要: 文章在分析比较国外科学数据唯一标识符研究现状的基础上, 结合已开展的DOI注册工作, 对中文科学数据DOI的应用现状和服务成效进行了概述, 对科学数据唯一标识符应用的可行性进行了分析, 并对中文科学数据DOI管理策略以及注册推广中存在的突出问题进行了分析, 将为今后的科学数据DOI注册工作提供参考。

关键词: 数字对象唯一标识, 科学数据, 数据共享

DOI: 10.3772/j.issn.1673—2286.2013.08.006

科学数据总量的迅猛增加使得对科学数据资源的定位难度加大。传统的科学数据资源(保密数据除外)依靠URL对资源进行定位, 但URL指向数字对象的物理位置而不是数字对象本身的性质, 决定了其无法对数字对象进行长期、稳定的定位。1998年由美国出版协会(AAP)创立非赢利组织IDF, IDF在CNRI(美国国家研究推进机构)配合下制定了数字对象唯一标识(DOI)标准和相应的解析系统(Handle System), 建立了由IDF、注册代理机构(RA)、发布代理机构(PA)组成的DOI应用体系, 其目的是建立互联网环境下数字对象标识和知识产权管理和保护的解决方案。本文将对数字对象唯一标识技术(DOI)在中国科学数据领域中的应用现状与成效、存在的问题进行梳理, 希望能对今后的科学数据DOI注册和服务工作有借鉴作用。

1 科学数据唯一标识的发展现状

1.1 国外科学数据唯一标识的应用现状

在国外, 数字对象唯一标识符系统(DOIs)在原文获取、引文链接与数字版权管理中已有广泛的应用^[1]。科学数据作为一种科技资源, 也存在类似的应用, 国外相关机构已经开展了DOI在科学数据上的应用, 并取得了一定的成效。

德国国家科学技术图书馆(TIB)是世界上首个将DOI应用在科学数据领域的专业机构。德国科学基金会在2003年至2005年间资助德国CODATA启动了科学数据出版和引用的项目, 在该项目中使用DOI来标识科学数据资源。这些数据源存储在世界数据中心(WDC)在德国的科研机构^[2]。2005年它正式成为科学数据的DOI注册中心(RA for scientific dataset), 2009年10月由TIB牵头联合世界上多家科技信息机构建立的DataCite非盈利组织, 其RA的资格也转移到该组织, 旨在建立跨数据中心的数据集成系统, 方便研究者发现和获取数据, 推动数据的引用。截至2013年初, 已经有13个国家21个单位和组织加入了DataCite系统, 截止到2013年4月, 该组织的DOI注册量已经达到1400万个^[3]。注册的数据类型主要包括数据集、科学电影剪辑、医学案例、技术报告和学习对象等, 未来科学内容(scientific content)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3D模型等^[4]。

由于各国的发展环境、政策以及数据共享意识存在差别, 其他国家在科学数据领域的DOI应用工作还只是起步阶段。除了DataCite在国际上对科学数据进行注册外, 韩国科学技术情报研究所(KISTI)是DataCite的联系会员(affiliated member), 将通过DataCite来注册本国的科学数据资源; 日本科技振兴机

* 本文为中央级公益性科研院所基本科研业务费专项资金项目(编号: XK2012-4)成果之一。

构(JST)在成为全球第12个注册代理机构后,将在材料科学数据领域开展科学数据DOI的注册工作。

1.2 中文科学数据DOI应用现状概述

在中文科学数据DOI相关的理论研究方面,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^[5]。目前中文科学数据DOI的应用仍处在探索阶段,除了DOI本身的唯一标识的特性外,国内相关学者对DOI的价值尤其是在异质科技资源的链接和数据引用方面进行了初步的探讨。贺德方等探讨了DOI在异质科技资源的相互引用和链接上的应用^[6],吴立宗等总结了DOI在数据共享中的跟踪价值、引用价值、集成价值、关联价值,以及在科学数据出版中的应用^[7]。彭洁等对基于DOI的科技文献与科学数据的融合以及科学数据引用进行了初步的探讨^[8,9],王卫华等针对科学数据库的数据资源建立了科学数据引用规范^[10],张英杰等对科研人员引用科学数据进行了调查,指出科研人员针对科学数据引用依然存在着认识不充分、引用行为不规范等诸多问题^[11]。

在中文科学数据DOI应用服务方面,2007年3月,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成为中国大陆地区DOI的注册代理机构,开始提供中文DOI注册服务,并结合科技部精品期刊项目开展了中文DOI的推广,科学数据是当初确定的两个重点推广领域之一。2011年底,国际DOI基金会(IDF)正式授权中信所为DOI注册机构(RA),同时中信所授权北京万方数据股份有限公司为用户提供DOI的注册及相关服务,目前已经注册了15000条科学数据DOI,实现唯一标识的注册、解析和查询服务,并对传统科学数据中心、自然科技资源相关机构、科学数据相关的商业公司进行了调研,开展了初步的合作。

2 科学数据唯一标识在科学数据中心应用的可行性分析

科学数据资源是科学研究基础的信息资源之一,很多科学数据资源已经通过Internet对外提供数据服务。但在数据服务过程中,经常出现由于资源被移动而发生链接错误、在资源复制过程中发生版权遗失、由于版本

更新使资源链接失效或无法对数据引用进行有效标识等一系列问题。因而急需一种能对数据资源进行持久唯一标识的、动态更新的、与物理地址变动无关的、能跟踪版本更新信息的、引用方便的数字对象唯一标识系统。

唯一标识符系统的目的在于分配永久的、可互用的、优化组合的、可解析的标识符,用于公共访问及组织间的协同工作,唯一标识符系统在实现不同应用系统的互操作和集成、保护互联网上的知识产权等方面发挥着越来越重要的作用。

对于数据中心来说,唯一标识符系统主要意味着接受一个“前缀”,这个前缀是不同科学数据出版机构共享元数据和应用互操作的基础。唯一标识符可以更好地实现科学数据资源的集成处理和发布,便于进行科学数据的管理和应用,可以用同一种标识机制来管理科技资源,促进科学数据与科技文献的融合。在一些数据中心,各种来源的数据往往采用各自不同的标识机制,各种数据在内容特别是标识格式上的差异,以及同一内容科学数据不同数据机构保存的重复,客观上需要科学数据中心建立集成发布的界面和本地化控制的机制(如利用开放链接的机制),另外国外一些数据资源(TIB所收录的WDCC的数据资源)已经采用DOI的方式对科学数据进行标识,也要求我们用类似的机制进行管理。

唯一标识符大大简化了最终用户的操作,通过唯一标识符,用户可以准确定位到对应的数字对象,避免了因对象的位置变更、修改而导致的链接错误。唯一标识解析系统不仅只针对一个软件技术平台,更重要的是在国内建立一整套数字资源唯一标识符的解析管理机制,具体来说包括技术平台的维护和扩展、唯一标识符前缀的申请和分配、唯一标识符如何同数据库提供商的现有产品相结合等内容。

正是由于DOI的某些特性,能够满足科学数据管理的需求,如表1所示^[11]。

3 中文科学数据DOI应用的成效概述

自2007年中信所成为DOI注册机构后,项目组对国内大型科学数据中心,包括气象、国土、地震数据中心、WDS中国节点等国内有影响力的数据中心,以及目前运行比较好的地球系统科学数据共享网、中国科学院网络中心的科学数据库进行了深入走访和调查。调研的结果发现,本次调研的范围基本上涵盖了

表1 DOI特性与科学数据需求对照表

元素	满足科学数据管理的需求
命名：给科学数据唯一、永久的“名字”	使科学数据容易发现和使用、引用
解析链接：克服传统URL链接的“死链”问题	保证数据永远可以被链接到
多重解析：可携带更多信息	可以返回与科学数据有关的多种信息，如科学数据的多个网址，满足科学数据中心多渠道发布需要
唯一标识符管理：贯穿科学数据的出版、流通、长期保存过程，并具有很高的安全性	作为某种“凭证”，确认科学数据的产权信息、在DOI上可以携带数据产权相关的信息，便于科学数据的传播
出版源头注册	帮助首发权确认、版权保护，提供可信度和质量保证
服务特性：点击即链接	提高了科学数据获取的便利性，改善研究人员的链接体验
元数据和链接地址可更新	保证信息的准确性和链接的永久性
元数据/DOI可查询	获取标准的元数据和准确的DOI，促进数字资源的自动化集成和交换

国内主要的科学数据机构，大部分机构都意识到注册DOI工作的重要性，但对于DOI的安全性和收费问题仍然存在疑问，同时由于前期各数据中心缺乏项目支持（2008-2012年），积极性并不是很高，只有少数机构自发地开展，但随着试点工作的持续进行，DOI对数据中心的价值日益显现。2012年开始，科技部对认定的23个平台进行持续经费支持，一部分科学数据中心将DOI工作列入新一年的数据中心工作计划，DOI的注册和服务工作将会得到一个实质性的推动。

2008年中信所与中国科学院寒旱所西部数据中心（现改名寒旱区数据中心）开展了科学数据唯一标识注册与引用的试点工作，目前已经注册DOI 233条，双方共同制定了科学数据引用格式，开展了合作研究和共同申报项目，组织了多次研讨会。此外，随着讨论的深入，西部数据中心已经将DOI及其增值应用的思路融入到其业务流程中，创新地提出了科学数据共享的新思路，包括基于DOI的科学数据引用与出版、科学数据评审确保科学数据质量，以及尊重科学数据知识产权创新服务模式方面都取得了新的进展，逐步将科学数据资源盘活，也走出了一条数据中心可持续发展的思路，成为目前国内研究人员获取数据的重要渠道，其数据服务的成效也得到了业界的认可，这充分说明在引入DOI后将极大促进科学数据中心资源的管理和共享。该试点工作非常成功，在其影响下北京师范大学全球变化研究院、中国极地研究中心、地球科学数据共享网都表现出很大的积极性，已经签署了科学数据唯一标识

服务协议，并已经注册了少量数据，并表示在今后的项目和平台过程中将DOI作为科学数据管理的基本要素进行考虑。

2010年与北京师范大学全球变化研究院在项目数据DOI注册上开展了实验，已经注册了5条项目中产生的数据集，该机构为项目数据注册DOI一方面是由于项目结题需要，另一方面也是为了对这些项目执行过程中产生的数据进行唯一性、权威性认证，该工作得到了评委专家的认可，并可以在今后项目数据管理中进行推广应用。

在实物资源数据方面，项目启动之初，项目组对国家自然资源数据平台、实验细胞平台、中国农业科学院土肥所、中科院植物所、中国农业科学院农业信息研究所进行了深入调查，对实物资源的唯一标识注册进行了探讨，并对e平台中一万余条植物、微生物、实验细胞资源数据进行了注册。

对商业公司运营的科学数据平台方面，选择计算机领域的数据堂公司、实验方法领域的库巴扎公司以及万方医学网平台开展了调研活动，其公益性运行的数据中仅有少部分有注册的需求，而且DOI给该公司带来的商业价值并不明显，该层面的合作还需要进行深入探讨和挖掘。目前已经与库巴扎公司开展了实验方法方面的合作，确定了DOI注册的范围，签订了科学数据唯一标识服务合同，探讨了科学数据出版的一种模式，即在纸质出版物上对实验方法进行DOI标注。

4 中文科学数据DOI管理策略

6年来,项目组对中文DOI管理中的核心问题进行了研究和分析,逐步对DOI注册的对象、地址管理策略、DOI元数据和注册流程进行了梳理,在实际的管理过程中得到了应用。

4.1 科学数据DOI注册的对象

IDF对于DOI可注册的内容没有严格的限制,只要是数字物体都可以分配DOI,甚至对实物也可以赋予DOI号。DOI源头注册的功能,一方面是对源头资源知识产权的一种保护,另一方面也是保证DOI号唯一性,无论该数据汇交到哪个地方,DOI号依然保持不变。对于科学数据来说,科学数据注册的粒度根据用户的需要来制定,本文认为其注册的对象主要分为两个层次:

数据集合(data collection):一个或者多个数据集的集合,按照某种应用目的组织起来的数据集合称为数据产品。

数据集(dataset):按照一定序列组织的数据单元,由一条或者多条数据记录组成。

数据记录(data record):数据组织的最小单位。

从数据引用的角度上讲,粒度太粗或者太细都不符合数据引用的需求,定位到数据集将能最大限度地揭示数据的特征,因此在数据注册的时候尽量选择数据集进行注册,从而能够更好地揭示数据的本质,有利于进行科研成果的检验和审核。

科学数据DOI注册对象的要求主要包括:

(1) 数据具有稳定性:所注册的数据对象一旦注册就固定下来,不能随意改动,数据的稳定性还能保证该数据能够持续永久地被访问到。

(2) 数据产权清晰,非涉密:目前国内的科学数据大多数存在知识产权的问题,部分数据由于涉及国际机密,这些数据很难进行获取,因此在注册时应优先考虑产权清晰、非涉密的科学数据资源。

(3) 开放获取(open access):这是由DOI的特性所决定的,其号码背后的链接地址决定了该资源能够进行开放获取。

(4) 可被引用(citable):引用是DOI的扩展价值,数据能够被引用也从另一方面说明了数据的价值,并能对数据的使用情况进行追踪。

(5) 适合出版(fit for publication):与文献的出

版类似,数据注册DOI号以后,该号码将伴随数据的整个出版过程。

4.2 DOI编码与链接地址管理策略

DOI由前缀和后缀组成,并且有链接地址与DOI号进行绑定,其管理原则主要包括:

前缀由IDF统一组织编码,对于持续稳定注册的机构建议注册新的前缀,由数据中心提出申请,注册机构收到申请后提交到IDF,由IDF随机生成一个唯一标识该数据中心的前缀码,以后该中心注册的数据必须附加该前缀码。对于注册量比较小或者临时注册的情况建议使用公共前缀。

而后缀是由注册机构根据自身的需求,在保证号码唯一性的前提下自行定义的,中文DOI管理中心会出建议标准,且后缀的每个字段都必须是有意义的字符串,可选的注册信息主要包括项目信息、数据类别码、流水码等,将协助数据中心来制定后缀的编码规则。

链接地址:从实际运行的情况来看,该链接地址指向的是数据的元数据地址,数据地址将在元数据中进行体现。该地址必须是正式域名,且指向的是描述该数据的详细页面,如果有多个地址,一种方法是挑选最稳定的一个地址进行注册,同时也可以使用DOI注册的多向解析功能,注册多个地址,但由于该项功能的商业模式并不成熟,在初期应用的时候不建议使用。当链接地址发生变更时,由数据机构及时更新并通知数据注册中心,同时中文DOI中心将定期对链接地址进行查询,如果链接地址失效应及时通知数据机构并协商解决。

4.3 DOI元数据管理策略

在DOI的元数据模型方面,IDF在DOI手册(DOI handbook)中提出了基于INDECS框架构建DOI元数据模型,并规定了其中的核心元数据项,根据该标准TIB(DataCite)也制定了核心元数据标准(TIB metadata Kernel),中文科学数据DOI依据TIB的标准,构建了中文科学数据DOI元数据标准STD-DOI metadata V1.0。但是在实际执行的过程中,结果并不理想,由于技术和学科背景的差异,不同数据中心采用的元数据标准存在一定的差异,并不能全部按照该标准上传元数据,条件较好的数据中心能按照要

求准备元数据,并能够注册英文元数据,能够在国际上进行查询。但是大部分的数据中心并不具备这样的条件和水平。因此,更多的数据中心希望用一种更加简单的元数据,同时以更加灵活的方式(批量注册、在线实时注册)进行元数据的上载工作。为了能够兼容不同数据中心的元数据差异,中文DOI中心将采用一种简单的元数据标准实现DOI标识数字资源的描述和管理。主要的字段包括数据作者、数据标题、数据关键词、数据DOI号和链接地址,开发自动注册接口实现在线注册,同时能够兼容V1.0的标准,在条件成熟时,能够补充相关的元数据信息,来实现与国外资源的互联。

注册机构收录数据元数据的目的在于便于进行DOI号的查重,但更重要的是基于元数据开发DOI的增值利用,便于用户能够查询到该数据的部分信息,同时能更加方便地使用DOI服务。目前IDF对于元数据的增值应用并没有统一的标准,各家RA机构各自开发基于元数据的增值应用,如CrossRef出于商业利益,开发了一系列基于元数据的增值应用。元数据信息一般保留在注册代理机构本地的数据库,元数据也不会在国际范围内公开,因此数据中心根本不用担心注册提交元数据泄密的问题。注册代理机构之间的元数据一般也不共享,如果的确需要共享,机构间必须签署元数据交换协议。

在科学数据DOI注册的元数据的使用方面,根据我们与数据中心签订的科学数据唯一标识合同,数据机构赋予注册机构永久的权利去免费使用和分发元数据,注册中心不能随意更改元数据的内容,除非是有技术要求的需要,更深层次的修改必须通过协议解决。

4.4 中文科学数据DOI注册流程

科学数据DOI标识需要遵循IDF以及中文DOI制定的DOI注册、申请等相关机制和政策。中文科学数据DOI注册的具体实施过程是:

- (1) 申请。签订科学数据唯一标识服务协议。
- (2) 注册唯一标识符,按照指定的科学数据DOI标识规范,对适合进行标识的数据对象赋予唯一标识符。
- (3) 元数据。科学数据DOI元数据按照IDF制定的schema准备XML批处理文件,并向中文DOI注册科学数据每一个数据集的元数据信息。
- (4) 查询与解析。由科学数据唯一标识服务系统提供科学数据的查询、解析服务。

5 科学数据唯一标识注册推广中的难点分析

通过对6年科学数据唯一标识注册工作的总结,总体来看,科学数据唯一标识工作推进较慢,注册量较小,应用面较窄。科学数据唯一标识注册工作与文献唯一标识注册工作有很明显的区别,存在一些问题,原因一方面由于人手不够,集中在比较熟悉的领域,但更多还是从科学数据自身以及管理过程中存在的问题决定的:

(1) 科学数据缺乏一个稳定的载体,是DOI注册的障碍之一

国外的文献基本上被主要的文献出版商进行控制,国内虽然都是小的编辑部,但也存在像同方、万方、维普这样的数据出版商对数字资源进行整合、集成和应用。与文献不同,科学数据缺乏一个稳定的载体(没有纸质的载体),以及专业的数据出版机构,目前大部分科学数据生产机构(共享平台)都是通过项目的形式进行维持的,有的平台由于缺乏经费,其内容不再更新,或者网站根本就关掉或者停止访问,客观上造成比较难于有一个持续、稳定的科学数据获取渠道,对于科学数据DOI注册是一个最大的障碍。

(2) 科学数据自身的特点导致注册难于推进

跟文献不同,科学数据体量大,学科领域众多,数据来源多样化(网上搜集的、汇交的),缺乏知识产权保护,部分数据有涉密的要求。数据粒度难以把握。因此,注册的时候需要对数据资源进行筛选,挑选粒度合适、产权清晰、数据质量较好的科学数据注册DOI。

(3) 科学数据管理过程中的特点导致DOI注册工作难于推进

数据生产的时间不固定,通常要根据项目或者实验的时间数据的内容不固定,数据的内容由于项目结题或者管理的需要经常被重新组合、修改、更新,同时还有数据链接不稳定、地址经常发生变更、数据描述不规范等问题,需要数据中心在技术、人员、机制等方面进行改善,来保证DOI标识链接的稳定性和有效性。

(4) 担心元数据被汇交到国外,降低竞争力

数据中心担心注册时所用的元数据被汇交到国外,降低竞争力,这是目前DOI注册的最大障碍。在注册初期,项目组借鉴了TIB(DataCite)的元数据标准,一部分遵循了IDF的INDECS标准,但是也有一部分是为了开展增值服务的,同时为了与国际上的目录信息进行交换和集成,但大部分元数据信息无法收录,而且

实际上中文科学数据中心的国际化程度较低,仅有极少数中心实现了中英文双语的元数据描述,是在执行的过程中数据中心倾向于使用最简单的元数据进行注册。同时IDF不存储具体的DOI号、具体数据地址信息及元数据,只存储号段对应的RA地址,在某家RA注册的DOI号不管从哪个入口查询,最终都会由IDF引导到具体的RA来查询并返回信息。DOI的元数据由注册机构与注册用户协商解决,原则上汇交核心字段,元数据的基本功能是DOI号码的查重,不对外直接暴露,如果基于此开发增值应用,可签订具体的协议约定元数据的使用及公开条款。

(5) DOI收费的问题也曾经是导致注册工作很难推进的因素

中文DOI制定了详细的收费标准,包括会员费和注册费,一年大约是1000元左右,包含500条DOI的注册,单条DOI注册的成本仅为国外CrossRef注册费用的1/7~1/6。但在推广初期,我们免费注册了200多条DOI信息,但随着工作的开展,以及DOI维护成本的提高,需要由数据中心来支付相关的费用。在推广初期,数据中心未获得平台项目支持,最担心的就是一旦注册量太

大,注册费用比较高的问题,但实际上发现除了个别平台潜在注册量较大外,大部分中心的注册量都在1000以下,并且现在部分数据中心从2012年开始也已经得到了较稳定的经费支撑,DOI费用仅仅占很小的部分,因此可以通过项目经费的方式来支付该费用,如果注册量比较大,可以通过协商的方式来共同制定合理的收费政策,保证DOI在经费允许的情况下发挥最大的作用。

6 结语

确定唯一标识和建立链接解析是建立知识链接体系的基础,通过DOI与其他标准和技术的集成应用,可以实现知识链接从封闭系统向开放系统的转移,并使链接的资源不断多样化,从期刊发展到图书、学位、电子本,甚至科学数据^[13],在科学数据领域中引入DOI,也将构建完整的学术出版产业链^[14]。DOI除了其本身的注册、解析、查询基本服务之外,将有利于解决科学数据引用、出版和知识产权等问题,为科学数据中心的建设提供新的思路,来促进科学数据共享。

参考文献

- [1] 任瑞娟,刘丽斌,濮德敏,等.中文DOI路在何方:从参考文献著录与DOI的关系探讨中文数字对象唯一标识符的发展方向[J].中国图书馆学报,2010(2):115-121.
- [2] KLUMP J. Data Publication in the Open Access Initiative [J]. Data Science Journal, 2006, 5(15): 79.
- [3] DataCite Statistics service [EB/OL]. [2013-05-20]. <http://stats.datacite.org>.
- [4] BRASE J.德国国家科技图书馆科学内容DOI注册中心研究进展[J].中国科技资源导刊,2008,40(1):37-39.
- [5] 王燕,毛军.我国数字资源唯一标识符应用模式与管理机制[J].现代图书情报技术,2005(2):10-14.
- [6] 贺德方,张旭.服务于科技信息资源共享的数字对象唯一标识应用研究[J].现代图书情报技术,2007(8):26-29.
- [7] 吴立宗,涂勇,等.浅谈科学数据出版中的数字对象唯一标识符[J].中国科技资源导刊,2010,42(5):22-29.
- [8] 涂勇,彭洁.基于DOI技术的科学数据与科技文献融合的研究[J].数字图书馆论坛,2007(10):29-31.
- [9] 彭洁,涂勇.科学数据引用的探讨[J].数字图书馆论坛,2008(10):14-18.
- [10] 王卫华,胡良霖,沈志宏.科学数据引用规范的研制[J].中国科技资源导刊,2013(1):36-40.
- [11] 张英杰,彭洁,张新民,等.科研人员引用科学数据的问卷调查[J].中国科技资源导刊,2013(1):41-46.
- [12] 郭晓峰,涂勇,蔡京生.DOI在内容产业的应用[J].传媒,2012(12):72-74.
- [13] 贺德方.知识链接发展的历史、未来和行动[J].现代图书情报技术,2005(3):11-15.
- [14] 张旭.DOI在学术出版产业链中的应用研究[J].数字图书馆论坛,2007(10):5-9.

作者简介

涂勇 (1981-), 男, 湖北鄂州人, 博士, 中国科学技术信息研究所助理研究员, 研究方向: 科技资源管理、科学数据共享、数字对象唯一标识等。
E-mail: tuyong@istic.ac.cn

Research on the Application of Digital Object Identifier in China's Scientific Data

Tu Yong, Peng Jie / Institute of Scientific and Technical Information of China, Beijing, 100038

Abstract: Based on the comparative study of foreign Status quo of the unique identifier in scientific data, combined with the DOI registration work, the status and effectiveness of DOI application in China's scientific data were outlined, and the feasibility of DOI application in scientific data were analyzed, finally the problems of DOI management for China's scientific data and registration and promotion service were analyzed as the reference for future DOI registration service for scientific data.

Keywordst: Digital Object Identifier (DOI), Scientific data, Data sharing

(收稿日期: 2013-06-07)