

# 国外DOI相关标准介绍

■ 张书卿 / 中国出版科学研究所

标准是为了在一定范围内获得最佳秩序,经协商一致制定并由公认机构批准,共同使用的和重复使用的一种规范性文件。按照作用范围可将标准分为国际标准、区域标准、国家标准、行业标准(专业、协会和部门)、地方标准和企业标准。

DOI系统是一个建立在几个现有标准基础上并进一步发展从而提供一个完整的连续一致的系统。该系统包括四部分:数字对象标识符的编码规则、解析系统、元数据框架以及分配和管理机制。目前DOI系统的解析方案是以美国创新研究组织(The Corporation for National Research Initiatives,简称CNRI)研发的Handle系统作为支撑技术,元数据方案是建立在indecS(后升级为COA)的元数据框架的基础上。该系统的管理由国际DOI基金会完成。该系统中的标识符编码规则首先被美国批准为国家标准,后其整个系统又被国际标准化组织(ISO)批

准为一项国际标准项目,此项目正在组织制定中。

## 一、美国DOI标准介绍

ANSI/NISO Z 39.84-2005美国数字对象标识符编码规则(Syntax for the Digital Object Identifier)是由美国全国信息标准化组织(NISO)制定的标准,2000年被美国国家标准的管理机构——美国国家标准学会(ANSI)确认为国家标准,2005年再次得到确认。

### (一) 美国DOI标准的特点

从该标准的名称中,我们可以看出该标准主要是对DOI标识符编码规则的规定,而对DOI系统的其他部分——解析系统、元数据框架和DOI分配和管理政策没有进行规范。

DOI整个系统作为DOI标识符标准执行的背景放在前言(forward)部分,但不构成正式标准的一部分,只是作为信息参考。此外资料性附录A、B、C、D、E,均作为信息参考进一步说明了目录代码(Directory Code)的有效值和字符串的长度、维护机构的任命问题、数字对象标识符的实例、相关的标准和参考文献、DOI字符串的编码问题等。

### (二) 标准文本主要内容

#### 1. 该标准规定了此标准的目的和范围。

该标准的目的是规定DOI的编码规则,该标准范围仅限于对DOI的编码规则进行规范,而有关DOI如何分配以及如何使用等要根据国际DOI基金会(IDF)的政策执行。

#### 2. 该标准对DOI的编码规则以及DOI字符串的特征作出了详细的规定

该标准规定了一个DOI由前缀和后缀组成。前缀部分由目录代码<DIR>和注册人代码<REG>两部分组成,中间用“.”隔开。后缀由注册人为数字对象分配唯一标识字符串<DSS>组成。前缀和后缀之间用“/”隔开。数字对象标识符的编码规则是:<DIR>.<REG>/<DSS>。目录代码由数字组成,目前唯一的有效值为10。目录代码<DIR>和注册人代码<REG>由国际DOI基金会分配,DOI后缀<DSS>由注册人分配。数字对象标识符或其各个组成部分的长度没有限制。

该标准规定DOI前缀字符是统一的字符编码标准中的字符。0x00~0x1F 0x80~0x9F中的控制符在

DOI字符串中是无效的,也是不被DOI系统所接受的。后缀字符不能以\*/开始,因为\*可以指任何单独的字符,它将作为保留字符,留待以后使用。

该标准还规定DOI字符串中的字母没有大小写之分(如10.123/ABC与10.123/AbC没有区别)。DOI字符串中的字母在注册和解析的时候将被转换成大写字母。如果一个DOI以10.123/ABC进行注册,10.123/abc将与其进行转换,另外如果以10.123/AbC进行注册,将收到一个声明该DOI已经注册的信息。决定两个DOI是否雷同的过程,首先是将小写字母a-z转换成大写字母,然后对整个DOI字符串进行以8位字节为单位的对比。

### 3. 该标准规定了维护机构的职责

该标准规定该标准的维护机构负责答复有关各方提出的新数据元建议,解释该标准规定的各个条款,记录有可能对更新和修改该标准有关的所有问题及回复。有关该标准的实施问题以及其他相关信息咨询应发给该标准的维护机构。

## 二、ISO 26324标准草案介绍

2005年国际DOI基金会向ISO/TC46/SC9提交新工作提案(NP)——数字资源标识(DOI),并起草了工作草案(WD)。ISO/TC46/SC9经过组织各成员国投票,并根据投票结果,2006年决定成立由各国推荐的专家组成第7工作组起草DOI国际标准草案。目前第7工作组正在起草国际标准草案,并几易其稿,目前是第10稿(draft 10)。并将在11月份形成征求意见稿(CD),提交给

ISO/TC46/SC9秘书处,并面向各个P成员国征求意见。下面将主要对draft 10做一些介绍。

### (一) ISO 26324标准草案的特点

相对于美国国家标准主要对DOI的编码规则的规范,国际DOI标准草案是对整个DOI系统的规范。编码规则与美国国家标准一致,只是提法略有不同。

该国际标准草案对DOI命名(DOI name)和DOI系统(DOI system)作出了明确的定义和区分。DOI命名是在DOI系统内规定的一个单独的对象字符串。DOI系统是通过DOI命名的分配、解析、描述和管理实现对网络环境下对数字对象提供唯一和持久标识并提供信息交换的系统。

DOI命名的使用范围得到了扩充。该国际标准草案规定DOI命名可以用来识别任何形式的对象,不论是抽象的还是具体的,也不论是物质的,还是数字的。DOI作为对象的数字化识别符,而不是数字对象的识别符。

该国际标准草案反映了各国和其他标识系统管理中心的利益。例如为了防止标准中的知识产权问题,此国际标准草案将目前支撑DOI系统的Handle系统和COA元数据框架只是作为支持DOI系统运行的技术方案之一。此外该标准草案中明确并反复强调,DOI命名并不是要替代其他的识别符方案,如果一个对象已经用其他的标识符标识了,其他标识系统的字符串应被嵌入该DOI命名的字符串中。

### (二) 标准草案主要内容

#### 1. 该标准规定了适用范围及与其他标识符系统的关系

该标准规定了DOI系统的编码规则、描述和解析方案等功能性组成成分,并制定了DOI命名的创建、注册和管理政策。

DOI命名并不是要替代其他的识别符方案,如ISO/TC46/SC9制定的ISBN、ISSN、ISRC、ISAN等,或其他大家所熟知的识别符。如果一个对象已经用其他的标识符标识了,其他标识系统的字符串应被嵌入该DOI命名的编码中,除非其他标识符的注册中心有特别的声明,或者与相关的标识符的注册中心的协议不能达成。该标准也描述了DOI系统怎样可以用于ISO的其他识别系统,包括提供其他识别系统本身不能提供的功能(如解析功能等)。

#### 2. 该标准规定了DOI命名的编码规则和显示方式以及特征

DOI的编码规则规定了一个DOI命名组成的形式和顺序。DOI命名的编码应由一个前缀和一个后缀组成,前缀和后缀之间用一个“/”隔开。DOI前缀应由目录指示码(Directory Indicator)和注册人代码(Registrant Code)组成,目录指示码和注册人代码之间用“.”隔开。目前目录指示码应是“10”。注册人代码可以为了管理方便可以细分成各种子元素。注册人代码的各个子元素之间应用一个“.”分隔。

例如:10.1000.10

该标准还规定一个DOI命名没有大小写之分，同时可以嵌入统一标准字符编码标准中任何的字符。一个DOI命名的前缀和后缀加起来应是唯一的，以便DOI标识符可以由各个注册人自行分配。一旦DOI命名被分配了一个字符串，该字符串将不能改变（包括注册人代码），无论该对象的所有权和管理权是谁。一个DOI命名是一个没有意义的字符串，尤其是，DOI命名中所包含的分配给某个机构的注册人代码并不能提供任何证据说明该对象的任何所有权或当前的管理责任归属。

该标准规定当以屏幕或打印的方式显示DOI命名时，该DOI命名前面应加小写的“doi”，除非上下文明显的显示这是一个DOI命名。

### 3. 该标准规定了DOI命名分配的范围和规则

DOI命名分配的主要对象是与内容相关的实体，例如，但不仅限于：文档（text documents）、声音介质（sound carriers）、图书（books）、摄影作品（photographs）、连续出版物（serials）、录音、录像和音像制品（audio、video and audiovisual recordings）、软件（software）、抽象作品（abstract works）、艺术品（artwork）等，以及上述对象管理中的相关实体或对象，例如：授权、参与各方等。

该标准规定一个DOI命名可以分配给任何形式的对象，只要有将其作为单独的对象加以区分的必要，也不管它是否是一个对象中的一部分。例如：不同的DOI命名可以分配给一本小说或该小说的摘要、该小说的某一版本、该小说其中的某一

章节、一个独立的段落、一个图像或是引用的一部分等。

注册机构可以规定更详细和更严格的分配DOI命名的规则。这些规定应与整个DOI系统一致，但并不构成该国际标准的一部分。

### 4. 该标准规定了DOI命名的解析方案的主要功能

用于运行DOI命名的解析方案的技术应支持如下功能：与网络兼容、第一级的命名、唯一的识别、功能性的粒度、数据类型、多样性的解析方案、指定的权利机构、解析方案记录的适当获取、不依赖DNS但与DNS兼容、DOI命名管理层级的粒度、可以升级的、符合统一字符编码标准等。支持DOI系统运行的技术应由该标准的注册权利机构选择。该注册权利机构应保证选择的技术的改变不会影响原来DOI命名的应用。该功能由Handle系统支持。

### 5. 该标准规定了DOI的元数据框架的主要功能

管理元数据的技术将支持以下功能：处理复杂元数据的机制、互操作性、功能性的粒度、指定的权利机构、适当获取元数据的途径。该功能由场景本体构造（COA）支持。注册人分配DOI命名时，应将描述该DOI命名的元数据记录下来。DOI注册权利机构应向各注册中心提供DOI核心元数据表，以便各个注册中心能记录分配的DOI命名的元数据。

### 6. 该标准规定了DOI系统的管理机构

注册权利机构应是国际DOI基金会。DOI系统

应由国际DOI基金会和其任命的注册中心管理。注册权利机构应向各个注册中心提供使用手册，该手册将包含该标准实施的细节问题。

### 7. 其他规定和内容

附录A（规范性）对该标准的注册权利机构（IDF）和由IDF任命的各个注册中心在DOI系统中的角色和作用做了详细的规定。附录B（规范性）对数据词典、DOI核心元数据做了详细的规定。附录C（规范性）对DOI系统与其他标识符系统之间的关系以及DOI命名如何嵌入原有的标识符作了更详细的规定：

附录D、E、F（资料性）分别详细介绍了Handle系统、场景本体构造（COA）元数据框架以及DOI资源元数据元素，作为信息参考。

## 三、结语

全国信息与文献标准化技术委员会出版物格式分技术委员会（SAC/TC4/SC7）作为与ISO/TC46/SC9对口的标准技术委员会，参与了ISO 26324标准草案的起草工作。该标委会的承担单位——中国出版科学研究所也极其重视该项目，作为所里的课题进行立项，并成立了中国数字资源标识符系统标准工作组负责起草该项标准草案，目前该标准草案已完成。经过新闻出版总署的批准，目前该标准项目正在作为国家标准申请在国家标准化委员会立项。