

ICS 01.140.20

A 14

C A D A L 项 目 标 准

CADAL 10305—2012

CADAL 数字对象唯一标识符 解析系统应用规范

**CADAL Unique Digital Object Identifier Resolution
System Application Specification**

第一稿

2012-05-08 发布

2012-05-09 实施

CADAL 项目管理中心 发 布

目 次

前言	213
引言	214
1 范围	215
2 规范性引用文件	215
3 术语和定义	215
3.1 数字对象	215
3.2 唯一标识	215
3.3 元数据	216
3.4 CADOI 名称	216
3.5 CADOI 解析	216
3.6 CADOI 根解析服务器	216
3.7 CADOI 域解析服务器	216
3.8 CADOI 权威解析服务器	216
4 CADOI 解析系统体系结构	216
5 CADOI 解析系统功能模块	217
6 解析机制	217
6.1 一般规则	217
6.2 数据缓存和过期	217
6.3 命名授权解析	218
6.4 CADOI 名称解析	218
7 CADOI 解析方法	219
7.1 专用工具解析	219
7.2 浏览器插件解析	219
7.3 http 服务器解析	219
7.4 API 编程解析	219
8 管理规则	219
8.1 CADOI 系统的管理体系	219
9 扩展规则	220
参考文献	221
图 1 CADOI 体系结构	217
表 1 假设的 5 台计算机	218

前 言

《CADAL 项目数字对象标准》分 5 个部分,由 5 个标准组成。

- 第 1 部分:CADAL 10301—2012 数字对象唯一标识符规范。
- 第 2 部分:CADAL 10302—2012 数字对象标识与命名规范。
- 第 3 部分:CADAL 10303—2012 复合数字对象描述和管理规范。
- 第 4 部分:CADAL 10304—2012 数字对象阅读规范。
- 第 5 部分:CADAL 10305—2012 CADAL 数字对象唯一标识符解析系统应用规范。

本规范为第 5 部分。

本规范主要参考了 ISO/CD 26324 *Information and Documentation—Digital Object Identifier (DOI)*、*The DOI Handbook*、RFC 3986 *Uniform Resource Identifier (URI): Generic Syntax*、RFC 2141 *URN Syntax*、《数字资源唯一标识符解析系统应用规范》、《国家图书馆数字资源唯一标识符规范和应用指南》等。

本规范的某些内容可能涉及专利,本规范的发布机构不承担识别这些专利的责任。

本规范由西南政法大学图书馆提出。

本规范由大学数字图书馆国际合作计划(CADAL)项目管理中心提出并归口。

本规范起草单位:西南政法大学图书馆。

本规范起草人:江波。

引 言

数字对象唯一标识符(Digital Object Identifier, DOI),是实现数字对象永久标识和解析的完整体系。

CADAL 数字对象唯一标识符(CADAL Digital Object Identifier, CADOI)实现 CADAL 系统内数字对象的唯一标识和解析。

本规范遵循《CADAL 数字对象唯一标识符规范》,并具体规定 CADAL 数字对象唯一标识符解析系统的体系结构和工作机制,对 CADAL 中心、CADAL 成员馆、国内其他大学图书馆及其他需要建立数字对象唯一标识系统的机构和个人应用 CADOI 解析系统提出约束和规范。

CADAL 数字对象唯一标识符解析系统应用规范

1 范围

本部分规定了 CADAL 数字对象唯一标识符系统(CADOI)的解析体系、解析规则、应用规则和扩展规则。

本部分所指的数字对象,包括 CADAL 系统内一切的中文古籍、民国文献、中外文图书、连续出版物、学位论文、技术报告、地方文史资料、图形图像和声像资料等。其解析结果表现形式为一个合法的 URL 地址。

通过给数字对象赋予一个 CADOI 唯一标识符,可以实现对其的永久标识和解析。

CADOI 解析系统参考了 Information and Documentation—Digital Object Identifier (DOI),但与其有一定差别。CADOI 独立于 DOI 系统,不受 DOI 相关规则约束。通过预留的可扩展接口,CADOI 可以实现与 DOI 系统的兼容,实现互解析。

本部分不规定 CADOI 解析系统的实现技术,其相关技术问题在其他文献中说明。

2 规范性引用文件

下列文件对于本规范的应用是必不可少的。凡是注日期的引用文件,仅所注日期的版本适用于本规范。凡是不注日期的引用文件,其最新版本适用于本规范。

ISO/IEC 10646 *Universal Multiple-Object Coded Character Set (UCS)*

RFC 3986 *Uniform Resource Identifier (URI) : Generic Syntax*

RFC 3629 *UTF-8, a transformation Format of ISO 10646*

RFC 1630 *Universal Resource Identifiers in WWW*

RFC 1738 *Uniform Resource Locators (URL)*

3 术语和定义

3.1 数字对象 Object

数字对象指用户可获取的实体,可以是某一电子文献、著录信息或 Web 页面。

3.2 唯一标识 Unique Identification

唯一标识指有且仅有一个标识对象被该 CADOI 名称指定。

3.3 元数据 Metadata

元数据是关于数据的描述数据,有助于发现、确认获取、管理、评价和保存对象。

3.4 CADOI 名称 CADOI Name

CADOI 名称指在 CADOI 系统内定义一个唯一对象的字符串。CADOI 名称可以由遵循 CADOI 语法规则的字母、数字、字符组成。

3.5 CADOI 解析 CADOI Resolution

CADOI 解析指提交一个 CADOI 名称到 CADOI 系统,接受返回的一个 URL 地址的过程。

3.6 CADOI 根解析服务器 CADOI Root Server

CADOI 根解析服务器的英文缩写为 CRS,是 CADOI 解析系统中最顶层的服务器,负责一级命名授权的注册和解析。

3.7 CADOI 域解析服务器 CADOI Master Server

CADOI 域解析服务器的英文缩写为 CMS,是 CADOI 名称解析功能的主要实现者,主要管理自己域内的 CADOI 名称注册和解析,同时管理自己域内的下一级 CMS 命名授权和解析。

3.8 CADOI 权威解析服务器 CADOI Authority Server

一旦某个 CMS 在 CRS 注册一个一级命名授权后,其自身即成为该命名授权的权威解析服务器。同时,如果某个 CMS 将自身下属的某个二级命名授权分配与另一个 CMS,则该 CMS 成为该二级命名授权的权威解析服务器。

CADOI 的解析信息可能在多台的 CMS 中缓存,但是权威解析服务器上的信息一定被视为最新、最准确和最完整的数据。

4 CADOI 解析系统体系结构

CADOI 解析系统的总体结构为多层等级结构。

最顶层的是 CRS(CADOI Root Server)根解析服务器,下面是若干 1— n 级的 CMS (CADOI Master Server)域解析服务器(见图 1)。

CRS 的主要功能是管理一级命名授权和一级命名解析。

CMS 的功能有两方面:一是管理自己域内的下一级 CMS 命名授权和解析;二是管理自己域内的 CADOI 名称注册和解析。一个 CMS 下可以生成任意数量的下一级 MOS 和 CADOI 名称对象。

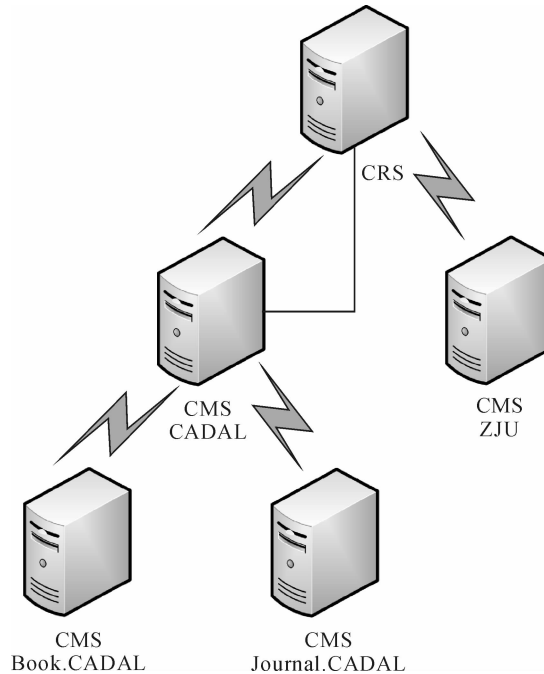


图 1 CADOI 体系结构

5 CADOI 解析系统功能模块

CADOI 解析系统包括系统设置、命名授权管理、命名授权解析、CADOI 名称管理、CADOI 名称解析、日志与统计等模块。

其中，系统设置包括 CRS 配置、数据配置和管理员管理等。

6 解析机制

6.1 一般规则

解析是向 CADOI 系统递交一个 CADOI 名称，接受系统返回的 CADOI 名称所标识的对象的 URL 的过程。

CADOI 解析记录有且仅有一个有效的 URL 地址，同时可包括其他元数据信息。

6.2 数据缓存和过期

各个命名授权的域解析服务器是它的权威解析服务器。同时，在其他服务器上可缓存 CADOI 解析记录。

每条缓存的 CADOI 解析记录均包含一个时间戳和一个过期参数。当某条 CADOI 解析记录过期后，缓存服务器自动丢弃该记录，并向权威解析服务器申请新的解析记录。

6.3 命名授权解析

用户提交一个 CADOI 名称到 CMS 请求解析时,CMS 首先会对该 CADOI 名称的命名授权进行解析,一般按照以下顺序进行:

- (1) 确认是否为自身。
- (2) 确认是否为自身的下级 CMS。
- (3) 确认缓存信息中是否有该 CMS 信息。
- (4) 向 CRS 请求解析,并获取该命名授权对应的 DNS 域名地址或 IP 地址。

6.4 CADOI 名称解析

用户提交一个 CADOI 名称到 CMS 请求解析时,CMS 首先对该 CADOI 名称的命名授权进行解析,再进行 CADOI 名称解析。

(1) 如果命名授权为自身,则直接查询出该 CADOI 名称的解析结果,返回用户。

(2) 如果命名授权为自身的下级 CMS,则先在缓存中查找,如果无则向该 CMS 请求,并缓存解析结果,返回用户。

(3) 如果命名授权非以上两种情况,则先在缓存中查找 CADOI 名称,如果无则在缓存信息中或向 CRS 查询命名授权信息,向对应 CMS 请求,并缓存解析结果,返回用户。

示例:

假设存在以下的环境(见表 1)。

表 1 假设的 5 台计算机

编号	名称	地址
1	客户机 A	——
2	CMS 服务器 A	cadoi. cadal. cn
3	CMS 服务器 B	book. cadoi. cadal. cn
4	CMS 服务器 C	cadoi. swupll. cn
5	CRS 服务器 D	cadoi. cn

客户机 A 需要访问某个数字对象,其 CADOI 名称为 123456@book. cadal。

(1) 客户机 A 向 CMS 服务器 B 提出了请求,例如 <http://book.cadoi.cadal.cn/123456@book.cadal>,服务器 B 发现命名授权 book. cadal 正是自己,于是,直接返回解析结果给客户机 A,即该资源的原始地址 <http://www.cadal.zju.edu.cn/book/32009663/>。

(2) 客户机 A 向 CMS 服务器 A 提出了请求,例如 <http://cadoi.cadal.cn/123456@book.cadal>,服务器 A 发现在命名授权 book. cadal 在自己域内,于是进行命名授权解析,然后向服务器 B 请求解析 123456@book.cadal,得到结果后,服务器 A 缓存解析结果信息,并为客户机 A 返回该资源的原始地址 <http://www.cadal.zju.edu.cn/book/32009663/>。

(3) 客户机 A 向 CMS 服务器 C 提出了请求,例如 <http://cadoi.swupll.cn/123456@book.cadal>,服务器 C 首先判断出该命名授权非自身,也非自身域内。若服务器 C 缓存中直接有有效的 book. cadal 的服务器地址,它直接向服务器 B 请求解析。若没有,服务器 C 将询问根服务器 D。如果根服务器 D 中依然没有 book. cadal 的信息,于是它会向服务器 A 处请求,服务器 A 发现命名授权 book. cadal 在自己域内,于是解

析到 book.cadal 的信息,返回给服务器 D。服务器 D 缓存信息,同时返回给服务器 C。服务器 C 收到服务器 D 返回的命名授权结果后,向服务器 B 请求 CADOI 名称解析,缓存后,为客户机 A 返回该资源的原始地址 <http://www.cadal.zju.edu.cn/book/32009663/>。

7 CADOI 解析方法

7.1 专用工具解析

CADOI 专用管理软件包含 CADOI 设置、注册、查询等所有功能。下载安装以后,可以利用其注册 CADOI 名称、查询 CADOI 名称等。

7.2 浏览器插件解析

IE、Chrome 等浏览器本身并不能直接识别和解析基于 CADOI 系统的唯一标识符,需要安装额外的插件程序。安装以后,用户可以在浏览器地址输入绝对地址的 CADOI 名称,如 `cadoi: 123456@cadal`,浏览器会自动将其交给解析插件处理。解析插件查询 CADOI 解析系统获得该标识符的 URL 链接,并自动跳转到该 URL 地址。

浏览器插件是较为便捷的解析方式。在安装插件后,用户可以在 Word 文档、pdf 文档、html 文档等可以插入链接的地方使用类似“`cadoi: 123456@cadal`”的方式完成对唯一标识符的置标和链接。用户在需要的时候点击该链接,插接便可以获得该标识符并完成解析工作。

7.3 http 服务器解析

通过在 CMS 服务器解析服务页面提交唯一标识符进行解析,或者直接在 CMS 服务器地址后加上 CADOI 名称,组成如 `http://cadoi.cadal.cn/123456@cadal` 的解析串来进行解析。

7.4 API 编程解析

用户如果需要在自己的应用程序中加入 CADOI 解析功能,可以利用 CADOI 系统的 API 函数。利用它向 CADOI 解析系统发送解析请求,并从服务器获取返回值。

8 管理规则

8.1 CADOI 系统的管理体系

CADOI 由中国 CADOI 管理中心(Chinese CADOI Center,CCAC)进行管理。

CCAC 提供以下的服务和功能:

- (1) 负责维持 CADOI 服务系统的稳定可靠、持续地运行。
- (2) 分配、管理 CADOI 一级命名授权的注册。
- (3) 促进 CADAL 系统内外 CADOI 的应用开发和推广。

CADAL 项目标准规范汇编(三)

CADAL 10305—2012

(4) 代表 CADOI 与国内外同类机构进行交流和合作。

在 CCAC 统一管理下,可在各地建立区域性的注册中心,负责一级命名授权的注册。

一旦一个一级命名授权投入使用,不宜随便修改。如果确实需要修改,需要在分配、注册、服务等方面进行全面评估。

9 扩展规则

CADOI 可将其他系统中应用的标识符方案引用到自身标识符中,包括采用其他 ISO 标准的标识符方案如 ISBN、ISSN、ISAN、ISRC 等,非 ISO 标准但被广泛认可和使用的标识符方案,或应用系统自定义的标识符方案等。

CADOI 解析服务器可以通过解析插件,对其他唯一标识符系统(如 DOI、CDOI 等)进行解析。

同时,免费开放解析接口给其他唯一标识符系统使用。

参 考 文 献

- [1] 贺燕. 走中国 DOI 发展之路[J]. 图书馆理论与实践, 2009(7):1-5.
- [2] 江波. 开放型数字对象标识系统设计[J]. 情报理论与实践, 2011, (9):107-109.
- [3] 梁健. Handle System 在中文 DOI 系统中的应用[J]. 数字图书馆, 2008(9):16-20.
- [4] 龙健, 赖茂生. DOI 的兴起与我国的对策[J]. 情报杂志, 2009(12):159-161.
- [5] 任瑞娟, 刘丽斌, 濮德敏, 米佳. 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.
- [6] 王海英, 龚永红, 张金艳. DOI: 数字对象标识技术[J]. 情报杂志, 2004(12):91-95.
- [7] 赵蕴华. 中文 DOI 应用现状及建设构想[J]. 情报理论与实践, 2007, (6):855-857.
- [8] Norman Paskin. Digital Object Identifier System[EB/OL]. 2008-01-12. <http://www.doi.org/overview/070710-Overview.pdf>.
- [9] Robert Kahn, Robert Wilensky. A Framework for Distributed Digital Object Service [EB/OL]. 2008-02-20. http://www.doi.org/topics/2006_05_02_Kahn_Framework.pdf.
- [10] ISO/IEC 10646 Universal Multiple-Object Coded Character Set (UCS) [S/OL]. 2012-04-12. [http://standards.iso.org/ittf/PubliclyAvailableStandards/c051273_ISO_IEC_10646_2011\(E\).zip](http://standards.iso.org/ittf/PubliclyAvailableStandards/c051273_ISO_IEC_10646_2011(E).zip).
- [11] RFC 3986 Uniform Resource Identifier (URI): Generic Syntax [S/OL]. 2012-04-12. <http://www.rfc-editor.org/rfc/rfc3986.txt>.
- [12] RFC 3629 UTF-8, a Transformation Format of ISO 10646[S/OL]. 2012-04-12. <http://www.rfc-editor.org/rfc/rfc3629.txt>.
- [13] RFC 1630 Universal Resource Identifiers in WWW [S/OL]. 2012-04-12. <http://www.rfc-editor.org/rfc/rfc1630.txt>.
- [14] RFC 1738 Uniform Resource Locators (URL)[S/OL]. 2012-04-12. <http://www.rfc-editor.org/rfc/rfc1738.txt>.