

中文DOI系统在影视行业的扩展应用* ——基于EIDR体系应用原型的研发

□ 李颖 / 中国科学技术信息研究所 北京 100038

时其迪 杨兴兵 / 北京万方数据股份有限公司 北京 100038

摘要: DOI系统在影视行业的扩展应用被认为是最有经济价值、并最有可能实施的工作, EIDR为该领域的最佳应用实践。文章对EIDR的架构及开发文档进行了全面研究, 并在现有中文DOI系统之上, 设计了基于EIDR体系的中文影视资源的服务系统, 进行了原型开发与验证。与此同时, 建立了中文EIDR元数据规范, 为中文DOI系统向多种类数字资源的全面扩展, 奠定了基础。

关键词: EIDR, 影视内容, 资产, 注册, 查询, 中文DOI

DOI: 10.3772/j.issn.1673—2286.2013.08.003

前言

“设立EIDR, 开创构筑娱乐行业电影和电视内容信息全球化注册体系, 促进全行业数字内容共享(digital commons)的合理化。”

2010年10月27日, 美国加州好莱坞, Motion Picture Laboratories、Cable Television Laboratories、Comcast及Rovi四公司组成的国际联盟与信息注册“Entertainment ID Registry (EIDR)”体系宣布成立。EIDR为非盈利、独立的国际信息注册联盟, 采用唯一(统一)ID, 对电影和电视节目、其他面向商业的音频/视频等资产信息进行分类、管理。以Deluxe、Universal Pictures、Neustar、Paramount Pictures、Sonic Solutions、Sony Pictures Entertainment、Walt Disney Pictures、Warner

Bros. Entertainment、Motion Picture Association of America, Inc.、Civolution、Vobile、INA (L'institut national de l'audiovisuel) 等公司为代表、EIDR为跨行业的众多公司所支援, 其注册体系的意义是支持“digital commons”的合理化和消费者交易的简约化。

且不论目前累积的影视作品的数量何等巨大, 来自资源传播渠道每年新增的几百万件娱乐作品又被不断地追加到庞大的内容市场之中。随着数字内容创作及其传播渠道的发展, 对构成娱乐行业供应链的众多企业来说, 追踪管理以视频内容为主的全部产品的复杂性也在不断提高。面对这样的状况, 采用通用(universal)的ID系统, 可以保障管理娱乐行业所有类型的音频/视频资产、便利地进行内容检索、追踪公司的内容销售、监管传播权和

销售额报告、提高向消费者提供信息的精度、削减成本、提高事务部门的处理效率。

MovieLabs公司的总经理兼CEO Steve Weinstein先生指出: “当今, 大多数企业在对自身娱乐资产分类管理上, 由于依赖自家系统或互不兼容的各种系统, 使得复数系统之间的内容管理变得非常困难。面对这一问题, EIDR在企业间的交互中, 发挥着将‘丢失的链接’连在一起的作用……。”

从2011年上旬开始, EIDR的成员可利用注册系统, 从原创作品开始, 到编辑、翻译版、剪辑、合成、捆绑版, 以及针对特定传播途径的不同音频/视频编码版, 按照所有级别, 将作品和资产(asset)进行分类识别, 系统支持非常广域的传播渠道和多种消费者设备。另外, EIDR还面向第三方的应用和服务开发, 提供工作流的整合、报告、

* 本文为中央级公益性科研院所基本科研业务费专项资金“基于数字对象唯一标识符的知识管理建设与研究”的成果之一, 获得国家科技支撑计划“技术创新服务平台关键技术研究与应用示范”项目的资助(编号: 2011BAH30B01)。

强化元数据等一系列的API。它将内容片段进行适当的分层,采用组织化(结构化)的统一数据库进行一元化管理,通过采用高度的去重系统,确保每个对象进行唯一ID的注册。

EIDR以International DOI Foundation (IDF)的数字对象唯一标识符DOI为基础,在基于开放的各种标准规范构建的同时,采用被广泛使用的Handle System的永久ID技术,采用CNRI公开的开源注册系统。基于这一弹性的公开的基础架构,EIDR能与现有的其他ID系统进行无缝互操作,使得与内容持有者及代理者、系统运维者、广告企业、标准提供商等广范围企业之间的交流合理化。

作为非营利行业团体、组织化的EIDR联盟,通过由内容所有者、影视宣传生产企业、内容传播企业、零售企业、娱乐服务/技术提供商等相关者构成的委员会,实施管理运行。其注册系统的开发,由联盟企业构成的技术顾问委员会(Technical Advisory Board)管理。通过参加企业的合作,满足业界今后新的ID需求的开发^[1-4]。

伴随着中国的文化大发展,我们的影视内容在世界范围内已具有了较高水准及认知度。然而,在基于信息技术的低成本、有效的管理方面,还处于较低的层次。为了与全球活动相向而行,为了贡献国际性传播,有必要研究与开发对应的系统,并与相关机构合作,促进我国影视资源产业链的高度发展。

作为DOI的权威注册机构(RA),利用中文DOI服务系统开展相关研究很有必要。为此,我们对EIDR的架构及开发文档进行了全面研究,为了以最小的代价,敏捷快速

地构建基于EIDR面向中文影视资产的应用服务系统,我们采用了DOI数据模型,在现有中文DOI系统的基础上,作为其扩展应用,设计了基于EIDR体系的中文影视资源的服务系统,进行了原型开发与验证。

1 DOI数据模型及中文DOI系统

首先,简要地介绍DOI数据模型及其目前的中文DOI系统。

DOI数据模型^[5]:如图1所示,DOI系统之所以具备通用性和极强的扩张能力,是基于其精简的数据模型。资源提供者将可以揭示其资源的元数据输入到RA,RA根据用户的服务需求,通过与服务有关的注册元数据来提供服务。如果服务

元数据是语义化的元数据,用户就可以获得语义服务。可以说,精简的DOI数据模型,却可以提供深度的服务,具有很强的适应性。

中文DOI系统^[6]:它隶属作为RA之一的中国科学技术信息研究所。中文DOI系统的开发实施分为两个阶段:第一阶段侧重于中文期刊论文和科学数据的DOI注册、解析及查询功能等;第二阶段侧重系统的优化和功能丰富;目前处于第三阶段,需要实行系统的升级,以适应环境的发展,对服务能力与水平更高的要求。正在运行的中文DOI网站的结构如图2所示。主要提供期刊资源的DOI的解析与查询功能,并具有其他的辅助功能。用户可以通过图示,获得相应的服务,并考虑申请成为中文DOI会员。

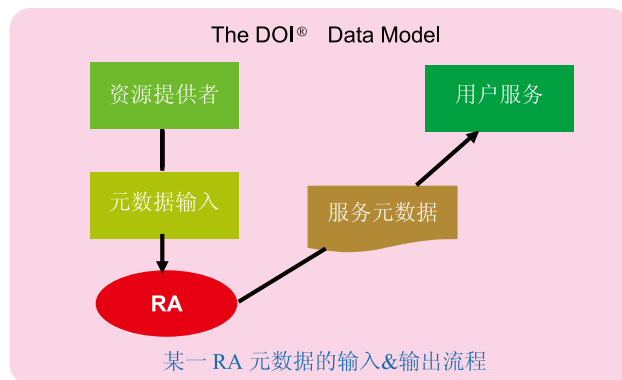


图1 RA网络中元数据流



图2 中文DOI提供的服务

2 EIDR框架^[7]

2.1 EIDR概要

关于EIDR: 数字技术的出现改变了从内容创作、后期制作到发行、消费等专业视听供应链的各个方面,为所有的利益相关者提供了新的机遇。伴随这些机遇,也迎来了挑战,比如更为复杂的价值链相互作用以及商务相关资产数量的暴增。通过增加分销渠道的数量,将这些资产有效地转化为货币,需要广泛采用行业标准、通用的唯一标识符注册来支持全方位的资产类型和资产之间的相互关系。娱乐唯一标识符注册联盟(The Entertainment Identifier Registry Association, 简称EIDR)的创建,以低成本向所有的行业相关者提供基础性服务。

EIDR的定位: EIDR是一个非盈利的行业协会,由Moviellabs、Cablelabs、Comcast及Rovi创立,为众多音视频对象的通用唯一标识符,满足娱乐供应链至关重要的需求。这些创立成员也负责董事会的成员监管注册。

EIDR的益处: EIDR有助于提高供应链效率,可在提高供应链效率上起到至关重要的作用:

- ◆ 消除专属ID系统之间昂贵的转换;
- ◆ 降低由ID非唯一性和重复造成的非唯一标识的风险;
- ◆ 提高内部资产的跟踪处理;
- ◆ 改善出自不同数据库、服务提供商,或元数据供应商的资产和元数据的匹配能力。

实现附加价值服务: EIDR为第三方的应用与服务开发人员提供

了一组APIs,比如 workflow集成、报告的形成、与增强性元数据的链接。这些增强性服务的例子如下:

- ◆ 更细粒度的、更准确的权益(rights)跟踪与报告,深入到剪辑、复合对象和编码级别;
- ◆ 简化的通用检索与发现;
- ◆ 针对单个资产详细的消费计量。

互操作性、其他唯一标识符: 互操作性是EIDR的核心,所包括的三个关键领域如下:

- ◆ 应用: EIDR ID的精确表示;
- ◆ 首要扩展支持来自其他系统标识符: 类(class)元数据元素;
- ◆ 基于其他系统的元数据的EIDR内容记录的注册: 包括标准的各个重要实例,基于元数据。

EIDR运作机理: 如图3所示,注册系统接受和处理来自注册者的请求。用户和应用程序可以查找(look-up)检索。注册者及查询用户可利用Web用户接口,或者Web服务的API与注册系统进行交互。注册系统在确认唯一性后,分配EIDR。

◆ 注册者向注册系统提交资产: 注册者采用核心元数据信息,比如对象类型及与其他对象的关系等,提交对象的注册。注册者为内容的所有者、集成商、后期制作部门或者任何其他获得授权注册的实体。注册系统需要最少的元数据以保证所有的资产类型的唯一性。

◆ 注册系统在核实唯一性后赋予EIDR: EIDR采用复杂的重复数据删除机制来确保被提交到注册系统中的对象尚未被注册,同时允许相似和相关对象的注册。集中注册保证了被注册对象的唯一性。如果没有重复对象的存在,注册系统则为对象生成一个EIDR,并在系统中存储一个新的EIDR和对应的元数据。参见图4,一个对象的EIDR包含标准的注册前缀、针对每一资产的唯一后缀和一个校验位。EIDR是没有所有权含义的一种纯粹的功能,即使相关资产的控制或所有权任意变化,EIDR也足以保持其永久不变。

◆ 注册用户界面: 一般搜索允许一个用户通过EIDR Lookup进行查找,或者所需EIDR字段的全搜索。

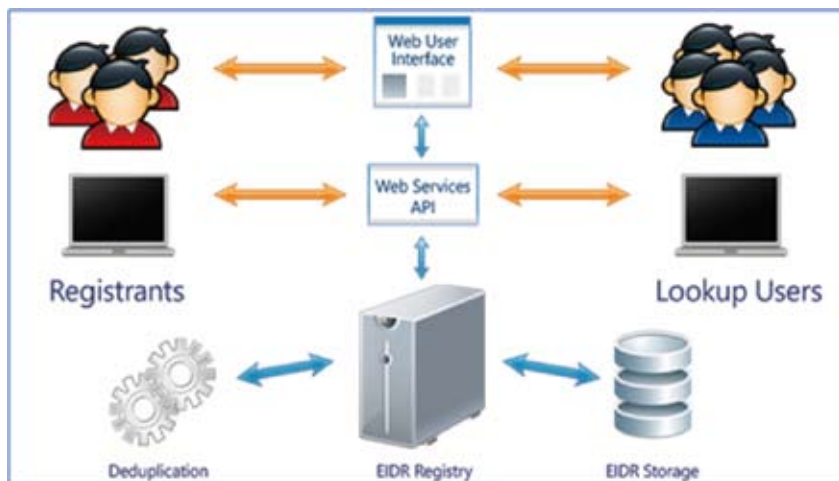


图3 EIDR运作原理

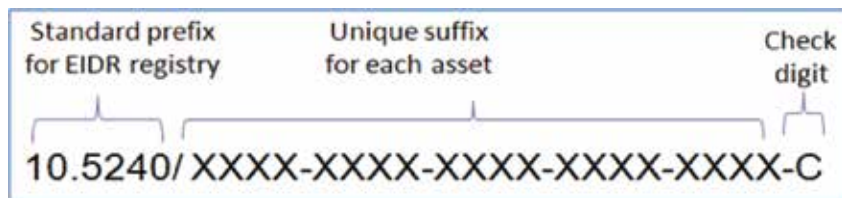


图4 EIDR唯一标识符构成

实施：EIDR是跨行业、非盈利联盟，是为电视和电影资产开发的通用唯一的IDs。在娱乐供应链的应用方面，由于其实用及已被验证的优势，被视为适用所有电视及电影资产的唯一标识符。已开发的注册系统提供多角色、支持所有娱乐供应链中运行的业务模式。EIDRs广泛地部署在娱乐供应链中，联盟技术面向市场的全覆盖提供基础设施，并支持创新型跨部门的合作。

会员：EIDR联盟定位是非盈利组织，由董事会管理，包含的主要干系人来自内容所有者、视频后期制作公司、内容分销商、零售商、娱乐服务和技术提供公司。EIDR理事为其运营制定整体政策。

加入会员：很容易加入EIDR，成为行业的贡献与促进分子。只需要在参与协议上签名，就可以成为基本用户会员。会员类型：

◆ 促进者 (Promoters)

行业促进者会员是行业领军者，需要积极参与注册系统的开发与管理。促进者可获得顾问委员的席位，具有主持工作组的能力，组织从事市场营销活动，具有董事成员的资格、投票权和享有高层基本用户的所有权益。所有的行业促进者要支付很少的年费并签署一份同意IPR（知识产权）和其他EIDR政策的参与协议。

◆ 贡献者 (Contributors)

贡献者会员对组织来说，需要参与注册系统的工作组和技术规

划，但没有投票权。贡献者参加顾问委员会会议和其他工作组会议，参加组织的市场营销活动，并具有基本用户的所有权益，但是没有投票权，不参加董事选举。贡献者基于组织的年度收入来支付累进费，小机构有资格获得低利率。与促进者同样，作为贡献者要签署一份同意IPR和其他EIDR政策的参与协议。

监管：EIDR的发展要接受来自联盟成员代表组成的技术顾问工作组的指导。未来的发展要满足新的行业ID需求，也要接受参与公司的指导。

会员的益处：如果某一机构在战略上对标识、开发和利用EIDR有兴趣，可选择成为EIDR的一个成员。成为会员的益处是：

◆ **战略影响：**会员有机会影响规范的发展，结合并促进本机构的战略发展方向。

◆ **事前洞察：**面向新出现的需求，通过对其技术发展的先行认知，会员可以获得竞争的优势。

◆ **获取专业知识：**会员可以直接访问专家，他们可以教育会员充分利用作为行业解决方案EIDRs的益处。

◆ **获得独家信息：**所有EIDR的会员可以访问特权信息。这些信息一般公众不能获得，包含技术文档、访问注册系统、会议纪要、高端的会议资料和内外发表资料。

◆ **网络平台：**会员可利用一个

行业间沟通的宽广的网络平台，它提供构筑联系网络和业务的机会，以及有关娱乐行业的一个无可伦比的专业知识的资源池。

◆ **产业定位：**成员作为社区思想引领的一份子，展示其承诺的开发端到端可信的解决方案，服务于多种角色、支持多种业务模式。

◆ **市场营销支持：**通过EIDR发起和引领的联合市场营销、攻关机遇，成员可以获益。

技术文档：

- ◆ 纵览与数据模型文档
- ◆ 用户与开发商文档
- ◆ 互操作文档
- ◆ 遵循的标准

2.2 EIDR系统整体框架

目标：娱乐唯一标识符注册体系EIDR针对电影和电视资产的所有类型，最早提供通用的标识符体系。范围从代表原创作品或知识产权的上层题目到作品物理上的各种具体体现，下至最小存储单位（SKUs）。EIDR提供基础性的命名空间，以实现供应链上的有效交互。它不取代或与供应商竞争附加价值服务，比如扩展元数据或权益跟踪。相反，它的目标是促进和支持这些附加价值服务。

EIDR系统的架构：如图5所示，EIDR系统包含如下的模块：

◆ **核心注册：**它是CNRI数字对象注册的定制化和配置模块，执行各种功能，包括注册、唯一标识符的生成、索引、对象注册管理以及访问控制。在注册中，每个对象都被分配一个唯一ID。

◆ **EIDR对象仓储：**仓储提供访问已注册的对象。对EIDR而言，这些对象为元数据的集合，而

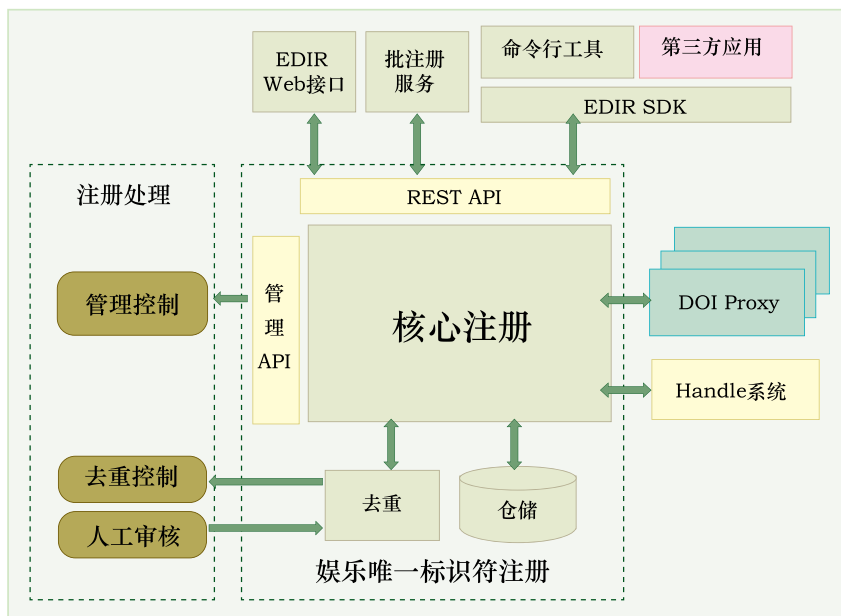


图5 EIDR系统架构



图6 EIDR的用户界面

不是媒体资产本身。元数据包括标准的对象信息、关系以及访问控制设置。

◆ 去重：此模块在核心注册模块中，要求检查新创或已修改对象的唯一性。去重模块以三种结果之一来响应注册请求：

(1) 无重复：被提交的记录独一无二。

(2) 重复：被提交的记录在注册体系中与现有一个对象重复。

(3) 潜在重复：在注册系统中，已提交记录与一个或多个记录重复的可能性很高。每个请求指定一个潜在的重复是否被拒绝或送去人工审核。

◆ 公共API：一组REST API调用，允许应用程序发出注册请求。这些调用，服务可以单个或成批调用，以及请求立即失败或者从去重中人工审查错误。EIDR用户界面以此API构建。

◆ Web用户界面：EIDR提供一个基于Web的用户界面，主要是搜索和查找。UI还支持更常见的工作流来进行注册和修改。

◆ 批量摄取：此应用程序一次可接受大到10万个注册请求，并负责提交请求到注册系统中。它接受平面数据集和具有内部层次结构的数据集（比如季节、系列、集）。提供异步状态界面，与直接请求注册相兼容。

◆ Handle系统：DOI生态系统是Handle系统的一个应用，以同样的方式，http是构建在TCP / IP之上的一个应用程序。Handle系统提供了分布式查找和解析服务。

◆ 管理API：这个API提供调用用来管理账户、用户和ACLs（访问控制列表）。

◆ 管理控制台：控制台实现注册操作系统的账户和用户管理、访问控制和其他支持功能。它基于管理APIs构建。

EIDR 注册界面：如图6所示。

利用用户界面可以做到：

- ◆ 执行搜索
- ◆ 通过请求返回的令牌，查阅异步请求状态
- ◆ 查阅特定用户启动的事务状态
- ◆ 查阅注册者启动的事务状态
- ◆ 创建记录
- ◆ 修改记录
- ◆ 访问EIDR支持网络
- ◆ 命令行工具：通过命令行工具，可以做到：
 - ◆ 解析EIDR的IDs
 - ◆ 检查对象层次结构
 - ◆ 查询各有关内容的记录、关系及视频服务提供商
 - ◆ 创建和修改对象

- ◆ 管理关系
- ◆ 创建对象
- ◆ 删除对象
- ◆ 别名对象

SDK: 利用SDK, 可以创建Java和.NET的应用程序, 利用注册服务以及执行注册系统的操作。如果有必要, 可以直接使用REST API, 但不推荐。

利用EIDR用户界面, 可创建电影, 如图7所示。本研究实现了这一功能, 参见图11。

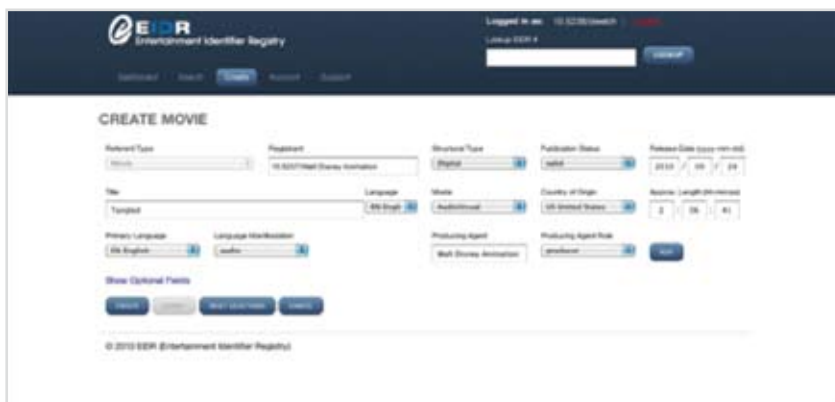


图7 在EIDR用户界面中创建电影

2.3 EIDR元数据结构

图8是EIDR的元数据schema的一部分, 描述了数字资产的元数据, 包含了Audio、Video、Subtitle、Image、Interactive等子元素。对于Audio而言, 其下位的元数据有Description、Type、Encoding、Language、Channels、TrackReference、TrackIdentifier、Private等。该元数据schema可以从EIDR的官网下载, 本文不再赘述。

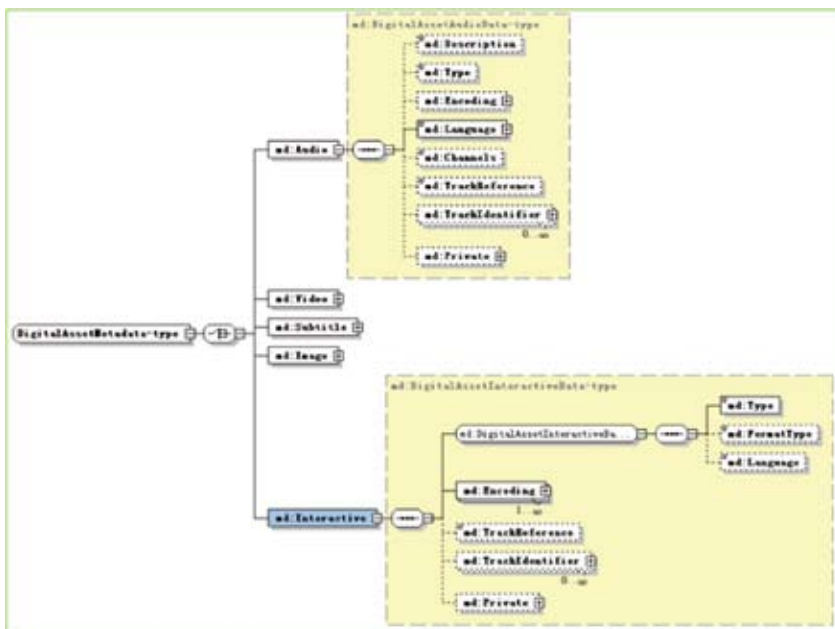


图8 EIDR元数据结构

3 中文DOI在影视行业扩展应用系统的结构及其元数据设计

3.1 中文DOI在影视行业扩展应用系统的结构设计

为了以最小的代价实现中文DOI系统在影视行业的扩展应用, 参照EIDR的元数据标准和EIDR的最佳实践, 本研究实施了敏捷开发, 有效地实现了中文DOI在影视行业扩展应用系统的设计及其原型系统的开发。原型系统的概念结构如图9所示。

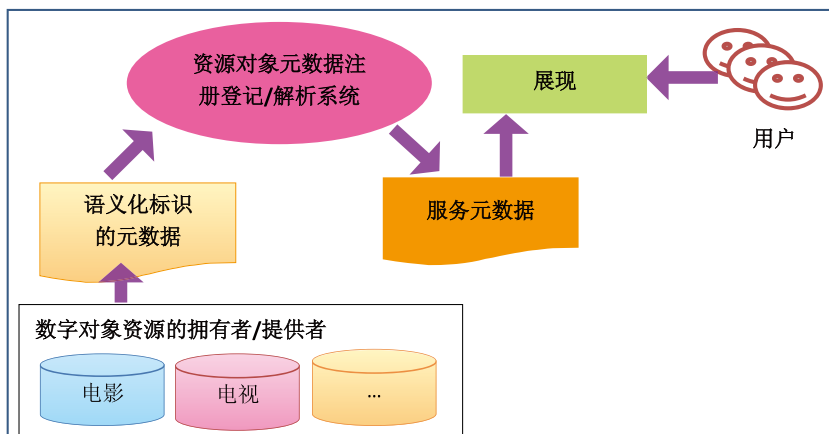


图9 原型系统的结构

3.2 中文影视内容的元数据设计

在设计中文影视内容的元数据时，我们本着结构化和模块化、通用和可扩展等基本原则，进行了相关设计。主要包括标识符、通用类型编码、基本元数据、数字资产元数据、容器元数据、内容分级、内容评级编码等影视行业元数据的核心部分。表1和表2是部分元数据的示例。

4 基于EIDR体系中文DOI扩展应用系统的原型开发

4.1 技术架构

参见图10，“中文DOI”影视行业扩展应用系统包括以下模块：

◆ Web用户界面：系统提供了基于Web的用户界面，允许用户通过Web页面进行影视数据注册和查看，并且还可以进行用户管理及数据统计等管理功能。

◆ 公共服务接口：一组REST API服务接口，允许程序单独或者批量进行数据操作请求，Web用户界面也通过此接口进行数据交互。

◆ 逻辑控制模块：此模块负责系统中各个功能的控制，包括注册、查询、管理、统计等，并根据业务不

同，控制接口及数据的访问控制。

◆ 服务组件：为应用服务提供核心支撑模块，如注册登记、数据查询检索、用户管理、数据解析、统计分析服务等核心服务。

◆ 数据库中间件服务：为系统和数据库之间提供必要的数据存取和管理服务。

4.2 原型系统的实现——Web用户界面

本期工作，基于系统架构，我们首先实现了2个主要服务功能：电影内容的创建（提交注册）和注册的查询服务。参见图11和12，系统截图。

表1 基本元数据（示例）

元素	属性	含义	值	card
WorkType		作品的类型。参见作品类型枚举。	xs: string	
WorkTypeDetail		关于作品类型的更详细定义，可以允许有一个更细节化的描述。	xs: string	0..1
PictureColorType		资产的颜色类型。这不应该含有纯音频资产。	md: ColorType-type	0..1
PictureFormat		屏幕高宽比格式类型的一个文本描述，定义见下文。该项不包含实际长宽比。	xs: string	0..1
ThreeD		指明一个作品是否是3D格式。‘true’表示是3D，‘false’或者空值表示不是3D。	xs: boolean	0..1
AspectRatio		有效像素的高宽比，其格式为m: n（如4: 3, 16: 9, 2: 35: 1）。	xs: string	0..1
OriginalLanguage		作品的原始语言。	xs: language	0..n
VersionLanguage		如果有的话，就是与该特定版本相关的语言。	xs: language	0..n

表2 数字资产元数据（示例）

元素	属性	含义	值	card
DigitalAssetMetadata-type				
Audio		音频资产元数据	md: DigitalAssetAudioData-type	(choice)
Video		视频资产元数据	md: DigitalAssetVideoData-type	(choice)
Subtitle		字幕元数据	md: DigitalAssetSubtitleData-type	(choice)
Interactive		交互元数据	md: DigitalAssetInteractiveData-type	(choice)

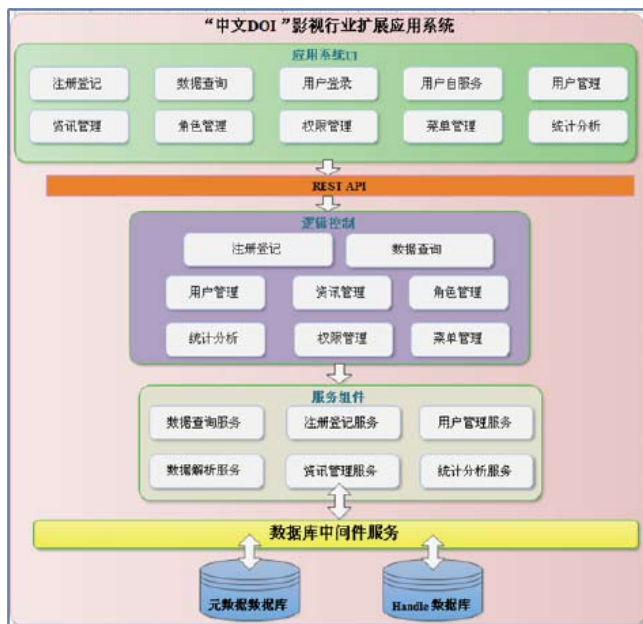


图10 中文DOI影视行业扩展应用系统的架构



图11 中文DOI影视行业扩展应用系统——创建电影功能的截图



图12 中文DOI影视行业扩展应用系统——注册查询功能的截图

5 结论与未来工作

面对我国影视行业对EIDR这类体系的需求，本研究从最佳实践的原则出发，基于EIDR架构，在中信所中

文DOI系统上，设计和开发了针对影视资源扩展应用的原型系统，使现有中文DOI系统主要针对期刊文献资源进行注册与查询服务的现状得到了扩展。作为中国首家IDF的RA，我

们期望通过本次实验研究，能将DOI扩展到所有资源，实现DOI这一生态系统的目标，并为资源乃至知识的互联服务做好充分的准备，这也是我们未来努力的方向。

参考文献

- [1] 日经新闻媒体.面向电影与电视节目Rovi及CableLabs等美国4公司成立了ID管理组织“EIDR”[EB/OL]. (2010-10-28) [2013-03-05]. <http://itpro.nikkeibp.co.jp/article/NEWS/20101028/353533/>.
- [2] [EB/OL]. [2013-03-05]. http://www.rovicorp.com/jp/company/newscenter/pressreleases/1434_14298.htm.
- [3] [EB/OL]. [2013-03-05]. <http://www.jlabs.or.jp/archives/7583>.
- [4] 关于EIDR [EB/OL]. [2013-03-05]. <http://eidr.org/about-us/>.
- [5] DOI官网[EB/OL]. [2013-03-05]. <http://www.doi.org/>.
- [6] 中文DOI [EB/OL]. [2013-04-05]. <http://www.chinadoi.cn>.
- [7] EIDR官网[EB/OL]. [2013-03-05]. <http://eidr.org/>.

作者简介

李颖，信息系统专业博士。研究方向：语义知识组织，基于主题的知识组织技术的应用等。E-mail: liying@istic.ac.cn

Expanded Application of China DOI in Movie and TV Industry - Based on EIDR

Li Ying / Institute of Scientific & Technical Information of China, Beijing, 100038
 Shi Qide, Yang Xingbing / Wanfang Data Co., Ltd., Beijing, 100038

Abstract: Expanded application of the DOI to movie and TV industry is considered to be of the most economic value, and most likely implemented work. EIDR is the best practice in this field. In this study, we survey the architecture and the development documents of the EIDR, based on the existing China DOI system, design a service system of the Chinese movie and TV contents with the EIDR, and carry on the prototype development and implementation. At the same time, building the Chinese EIDR metadata, lays a good foundation for expanding the China DOI into a variety of digital resources.

Keywords: EIDR, Movie and television content, Asset, Registry, Lookup, China DOI

(收稿日期: 2013-04-05)