

CADAL 数字化文本 元数据规范草案

Edocument Metadata

(Version 2.0)

CADAL 项目管理中心

2004 年 7 月



目 录

1. 规范说明.....	34
2. 著录对象.....	45
3. CADAL 元数据标准的总体说明.....	56
3.1 功能需求.....	56
3.2 开发原则.....	67
3.3 执行框架.....	89
3.4 CADAL 元数据规范定义.....	1044
4. CADAL 元数据 Schema.....	3233
4.1 Schema 源码.....	3233
4.2 CADAL 元数据结构图.....	3334
5. 著录规则.....	3536
6. 分类主题标引要求.....	3536
7. 著录实例.....	3637
7.1 元数据 XML 文件.....	3637
7.2 电子图书 OPF 文件.....	3637
7.3 四库全书元数据著录实例.....	3738
7.4 学位论文著录实例.....	3940
7.5 普通图书著录实例.....	4044
参考文献:.....	4142



CADAL 数字化文本元数据规范

CADAL Edocument Metadata (draft)

CADAL 元数据标准是依据 OEBPS 规范和中美双方对 CADAL 元数据著录内容的共同要求,参照国家科技部科技基础研究重大科技专项“我国数字图书馆标准范围研究”项目之子项目“专门数字对象元数据标准规范研究”项目组研制的《电子图书元数据规范》以及《古籍元数据规范》等,对数字图书馆的收藏对象进行分析后制定的。本标准草案的著录对象包括:普通图书、古籍、学位论文等的数字化文本。文献的内容及载体形式的不同,信息反映要求自然也不尽相同。因此,这三种文献分别采用不同的元数据著录参照规则。本标准草案主要包括三个方面内容:一、元数据规范;二、著录规则及主要著录项目信息源说明;三、著录实例。

1. 规范说明

1.1 规范名称为 CADAL 数字化文本描述元数据规范,版本为 Edocument Metadata Version 2.0。

1.2 规范责任单位为 CADAL 元数据规范项目组,起草单位为浙江大学图书馆。

1.3 规范起草时间为 2003 年 4 月 11 日。

1.4 规范复用了 Dublin core(DC)的 15 个元素作为数字对象核心元素集。

1.5 规范复用了美国国会图书馆(LOC)的 MARCXML Schema 作为著录对象内容描述元素集。

1.6 规范参照了 OEBPS 对 Dublin Core 的 15 项核心元素的定义。

1.7 规范撰写人:金更达、黄晨、孙晓菲等



2. 著录对象

数字化文本是以数字形式制作、出版、存取和使用的文献资源，一般以磁或电子介质为存贮载体，并借助一定的阅读设备或阅读软件来读取。虽然在内容描述上，数字化文本与印刷型文献区别不大。但是，数字化文本存取灵活，具有与印刷型文献不同的多种功能。第一，数字化文本可以实现对文献内容的深入揭示和多维组合功能；第二，数字化文本可以针对不同读者的需求实现个性化的文献导读功能；第三，数字化文本可以对文本内容进行多种方式的管理。然而这些功能的实现均依赖于数字化文本著录元数据标准的建立。

建立数字化文本著录元数据标准，首先必须明确著录对象。目前，CADAL项目计划数字化的资源总量在 100 万册（卷），中文资源和英文资源各为 50 万册左右，将主要选择学术著作及其他对教学、科研具有重要参考价值的图书进行数字化制作。数字化制作的现代中文图书的总量将在 30 万册以上(CADAL 定义 的现代中文图书是指“五四”以来出版的图书)。其中 1949 年以后出版的图书约 20 万册，1911 年至 1949 年出版的图书（民国图书和期刊）约 10 万册；学位论文数字化制作的总量预计在 10 万篇左右；古籍 10 万册(卷)。因此，本标准著录对象可以界定为普通图书、古籍、学位论文等的数字化文本，暂时不包括期刊的数字化文本。

数字化文本一般有两类：印刷型文献的数字化衍生物和原生的数字化产品。本标准的著录对象特指印刷型文献的数字化衍生物。另外，数字化文本元数据标准的制订准则是适应数字化文本以发布的形式作为基本的著录单元。用户可以依据其数字环境的需求，使用该标准对一部数字化文本进行著录；也可以对能独立使用的数字化文本的某一章节进行著录。由于本标准定义的数字化文本来源于传统印刷型文献（图书、古籍和学位论文）的数字化衍生物，因此，数字化文本的基本著录单元应以图书、古籍和学位论文印刷型的著录单元为基础。另外，由于数字化文本存取灵活，在著录时不考虑印刷型文献集中著录的现象。



3. CADAL 元数据标准的总体说明

XML 及其相关技术的发展，使元数据（Metadata）成为数字图书馆资源组织管理的基础。目前，我国数字图书馆建设还未形成一个类似于 MARC 一样被广泛采用的元数据标准，其原因归咎于每个数字图书馆建设都有其不同的功能需求，而元数据设计最重要的原则就是满足项目所确定的功能需求。因此要设计 CADAL 元数据标准规范，首先必须确定其功能需求。

3.1 功能需求

经中美双方协定，CADAL 电子书格式采用了由开放式电子图书论坛（Open eBook Forum）出版结构工作组发布的开放式电子图书出版结构（OEBPS）规范，该规范一方面给电子图书的内容提供者制定了最低限度的电子图书出版格式标准，内容提供者可以按照该标准制作电子图书；另一方面也给电子图书阅读系统开发商提供了电子图书的内容揭示标准，以确保符合该标准的电子图书能够在相应的阅读系统上得到准确的再现，同时易于访问。OEBPS 规范主要由 6 个核心元素组成：

- 包标识符（package unique-identifier）
- 元数据（metadata）
- 文件表单(manifest)
- 脊骨（spine）
- 漫游（tours）
- 向导（guide）

其中<Metadata>元素用于确定电子图书的元数据著录格式、内容和元数据所属的命名空间。Metadata 元素有两个子元素组成：<dc-metadata>和<x-metadata>，其中<dc-metadata>元素采纳了 Dublin Core 所定义的 15 项核心元素；<x-metadata>则用于扩展，允许内容提供者表达超出都柏林核心范围的、自定义的元素。因此 OEB 制作者可以根据自身的需求定义自己的扩展元素，也可以直接复用其它元数据标准中规定的元素。

根据 OEBPS 规范的定义，CADAL 元数据标准应包含 Dublin Core 的 15 项



核心元素，这也是中美双方经过协商所必须包含的部分。但由于 CADAL 项目是要建立一个具有百万数量级电子图书、存储容量达到几十个 TB 的数字图书馆，仅仅用于信息发现的简化元数据标准，——采用 DC 的 15 项元素显然无法满足对百万册电子文献的信息组织提出的更高要求。多维地、有效地、清晰地揭示信息是数字图书馆信息组织的一个重要原则，因此 CADAL 元数据标准规范的功能需求主要体现在以下几个方面：

- (1) 符合 OEBPS 规范的要求；
- (2) 包含 Dublin Core 的 15 个核心元素；
- (3) 著录信息包括“数字化对象”本身；
- (4) 著录信息要多维地、有效地、清晰地揭示电子文献，达到对百万量级电子图书的信息组织。

3.2 开发原则

1. 一致性原则

一致性原则是指在元数据设计过程中，应尽量注意保持与现有的国家标准、行业标准或者与其他政府标准、国际标准相一致。鉴于 CALIS 已采用 Dublin Core 的 15 项核心元素作为元数据的交换标准，而 CADAL 采用的电子书格式 OEBPS 也采用了 Dublin Core 的 15 项核心元素，因此 CADAL 元数据方案将复用 Dublin core 的 15 项核心元素和部分元素修饰词和编码修饰词，同时根据 DUMP-DOWN 扩展机制扩展 2 个核心元素。

2. MARC 映射原则

CADAL 项目元数据的著录对象特指图书、古籍、学位论文等印刷型文献的数字化文本。其中，中文 30 万册和英文 50 万册图书均有现成的 MARC 记录（中文图书可从国家图书馆书目数据库中套录，美方图书则可从西文数据源中套录），因此，可以建立 MARC 与 DC 的映射表，将 MARC 记录通过映射转换成 DC。

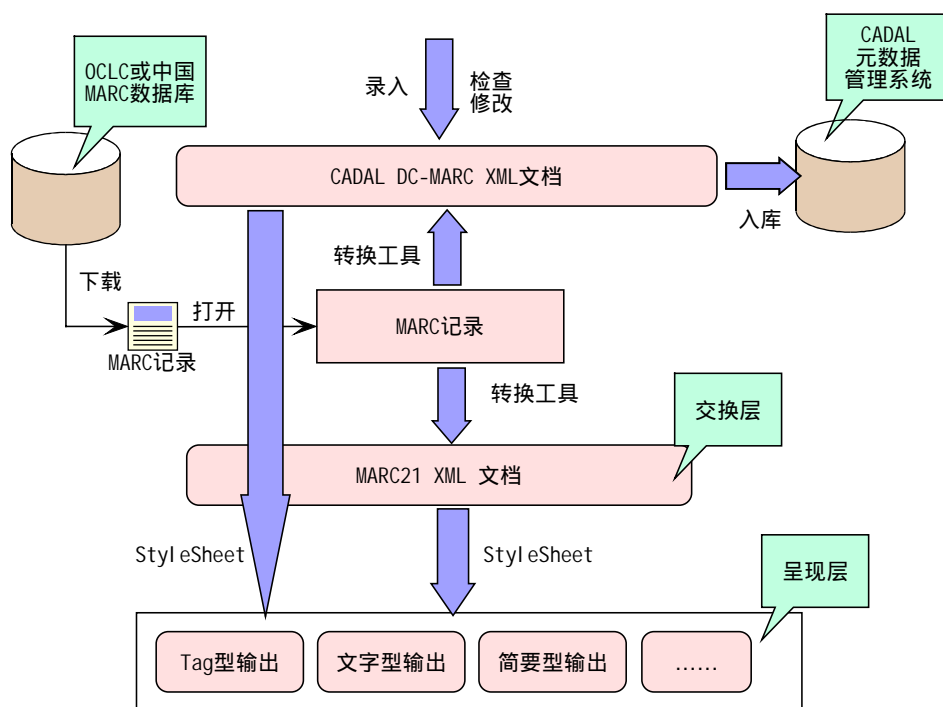
3. 保留 MARC 记录原则



尽管可以通过 MARC 与 DC 的映射表将 MARC 记录转换为 DC，然而这种转换往往会丢失部分信息，同时由于 DC 较之 MARC 简明，使转换后的信息不足以清晰地描述文献的原始信息。另外 CADAL 项目本身包含大量丰富的 MARC 记录，完全可以建立基于 MARC 的数据库，并提供相关的检索点。因此，CADAL 将采取保留著录对象原文献的 MARC 记录的原则，并在元数据的数据执行方案中将 MARC 与 DC 关联起来，以保证元数据的数据一致性和完整性。

4. 互操作原则。元数据的互操作原则体现在对异构系统间互操作能力的支持，不仅可以为自己的应用系统所操作，而且可以为其他组织或机构的应用系统所操作；不仅可在不同系统实现在同一元数据标准间的数据的传输、交换或转换，而且不同元数据标准间实现数据的传输、交换或转换。通常在元数据的具体应用上，互操作性往往表现为易转换性，即在所携信息损失最小的前提下，可方便地将元数据转换为其他系统常用的元数据标准。要实现这些功能，在元数据标准设计过程中必须慎重考虑元素语义和元素结构的准确定义，其中语义定义尤为重要。为此 CADAL 元数据方案将对元素语义和元素结构作出准确的定义。

3.3 执行框架



(1) 创建元数据

元数据方案框架提供了两种方式建立 CADAL 元数据：

有 MARC 记录的情况下，实现 MARC 记录到 CADAL 元数据的转换

- 下载 MARC 记录。各制作中心可以从本地图书馆自动化系统、全国联机编目中心或 OCLC 中下载 MARC 记录，并以 ISO 的文件格式将其保存在 CADAL 电子书的 META 目录中。如“.\meta\2709.iso”
- 利用 MARC2DC 的转换工具生成 CADAL 元数据信息，其中 marc 元素记录 MARC 记录的文件名称。
- 修改转换的 CADAL 元数据信息，直至符合元数据著录细则的要求。
- 保存结果。

无 MARC 记录的情况下，根据著录细则录入元数据。

(2) 呈现元数据

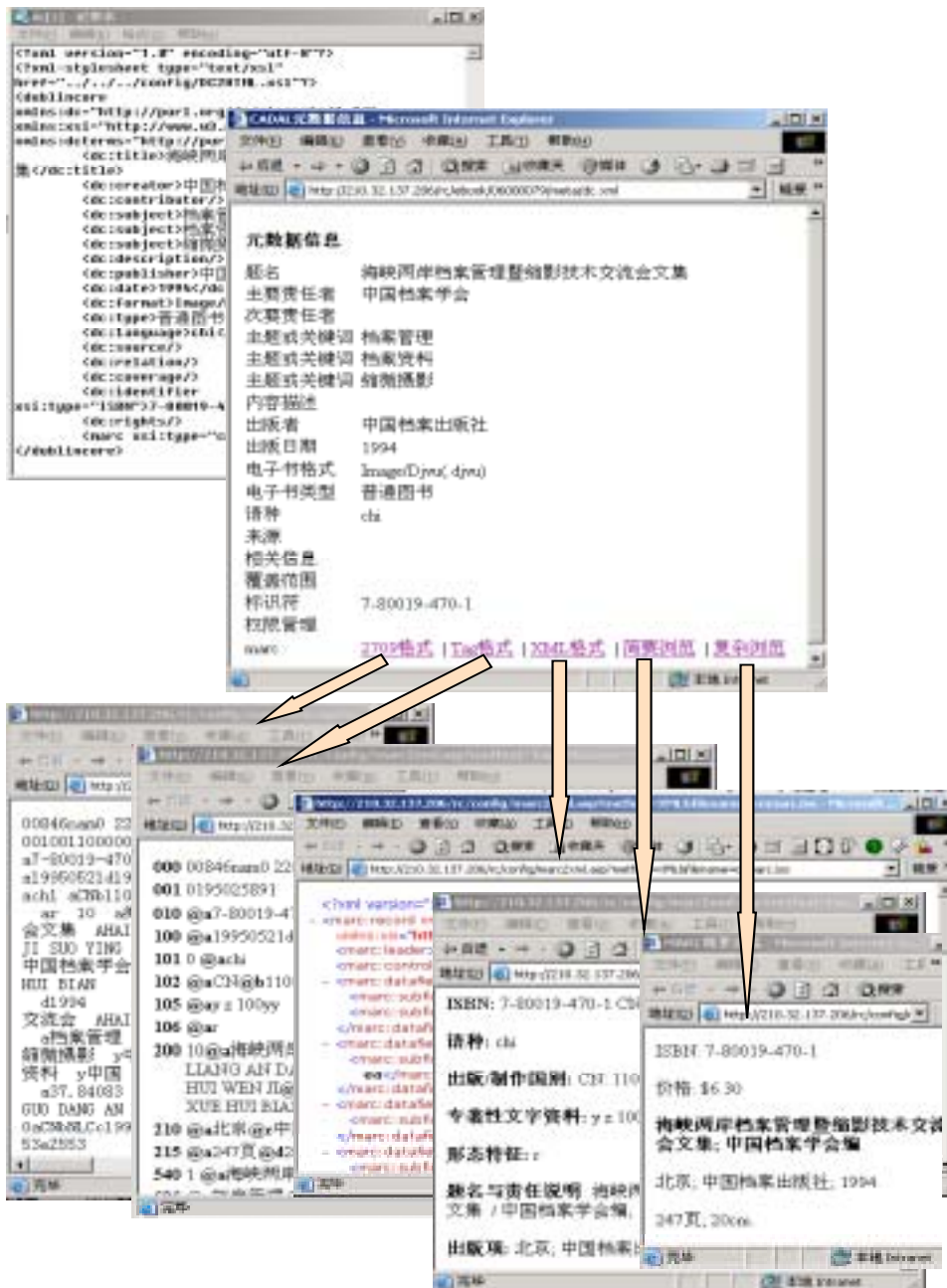
CADAL 元数据标准提供的元数据允许用户浏览 DC 和 MARC 元数据。这里主要涉及的工具和样式表 (Stylesheet) 有以下几个：

转换工具：MARC2709—>XML: marc2xml.asp

StyleSheet

- DC2HTML.xsl
- MARCXML2TagChi.xsl (TAG 格式)
- MARCXML2StandardChi.xsl (标准格式)
- MARCXML2SimpleCh.xsl (简明格式)

dc.xml 呈现实例图：





3.4 CADAL 元数据规范定义

3.4.1 元素定义的方法

本元数据规范的定义采用 ISO/IEC11179 标准 ,所有元素有四个属性是相同的 :

- 版本 : 2.0
- 注册机构 : 浙江大学图书馆
- 语种 : 中文
- 最大使用频率 : 不限

以下的元素定义省略了这四个相同的属性。子元素也参照上述方法定义 ,并在注释项说明其父元素。定义的方法如下表所示 :

命名空间	元素所在的命名空间			
名称	元素中文名称			
标识	元素标识			
定义	定义元素的含义 ,用于确认著录的内容。			
必备性	说明该元数据是 “ 必备 ”, 还是 “ 可选 ”			
可重复性	说明该元素是否可以重复出现 , Yes——可重复 ; No——不可重复			
属性	说明该元素所具有的属性 , 属性说明有以下几部分组成			
	标签	必备性	名称	参见
应用层次	数字化文本包括普通图书、古籍、学位论文、民国图书等。			
元素修饰词	描述该元素是否具有下属子元素 , 描述方式如下表所示			
	标签	必备性	名称	参见
编码修饰词	描述该元素信息的描述遵循何种编码规则。			
	标签		名称	参照



数据类型	字符型、数字型、日期型、结构型		
指定值	指该元素所具有的值列表		
缺省值	指元素默认值		
注释	对元素的进一步说明		
著录信息源	元素著录内容的信息来源		
著录细则	元素著录内容与方法的细则规定		
DC 映射	与 DC 元素的映射		

3.4.2 元素定义

本元数据草案资源对象界定为普通图书、古籍、学位论文等的数字化文本，其核心元素组成如下列表所示：

- 题名——title
- 主要责任者——creator
- 主题词或关键词——subject
- 次要责任者——contributor
- 资源描述——description
- 日期——date
- 资源形式——format
- 资源类型——type
- 来源——source
- 语言——language
- 标识——identifier
- 时空范围——coverage
- 出版者——publisher
- 相关资源——relation
- 权限管理——rights
- 版本信息——edition
- MARC 记录——marc



3.4.3 核心元素

1. 题名

命名空间	暂定			
名称	题名			
标识	Title			
定义	由创建者或出版者给予数字化资源的正式公开的名称。			
必备性	必备			
可重复性	YES			
属性	-			
应用层次	数字化文本：普通图书、学位论文、古籍、影印古籍、民国图书、期刊等			
元素修饰词	标签	必备性	名称	参照
	Alternative	可选	交替题名	[DCMI]
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	数字化文本的题名可以由若干部分组成，如主要题名、其他题名信息，，在 title 中著录主要题名，在 alternative 中著录其他题名信息。			
著录信息源				
著录细则				
DC 映射	Title			



2. 主要责任者

命名空间	暂定			
名称	主要责任者			
标识	Creator			
定义	创建数字化文本内容的主要责任者			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照
	Role	可选	角色	[OEBPS]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	著录数字化文本内容的主要责任者			
著录信息源				
著录细则				
DC 映射	Creator			

3. 主题/关键词

命名空间	暂定			
名称	主题/关键词			
标识	Subject			
定义	描述数字化文本主题内容的受控或非受控的词汇。			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照



	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	标识	名称	参照	
	CCS	中国主题词表		
	CLC	中国图书分类法		
	SKC	四库分类法（北大）		
	LCSH		[DCMI]	
	MESH		[DCMI]	
	DDC		[DCMI]	
	LCC		[DCMI]	
	UDC		[DCMI]	
数据类型	字符型			
指定值	-			
缺省值	-			
注释	用于描述数字化文本的主题内容特征，包括受控主题词、关键词、分类号等。			
著录信息源	用户定义			
著录细则				
DC 映射	Subject			

4. 次要责任者

命名空间	暂定
名称	次要责任者



标识	Contributor			
定义	除主要责任者之外的创建数字化文本内容的其他责任者			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照
	Role	可选	角色	[OEBPS]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	除主要责任者外的其他参与数字化文本内容创建的责任者			
著录信息源				
著录细则				
DC 映射	Contributor			



5. 资源描述

命名空间	暂定			
名称	资源描述			
标识	Description			
定义	对数字化文本内容相关的描述说明、附注等。			
必备性	可选			
可重复性	YES			
属性	-			
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	Abstract	可选	内容摘要	[DCMI]
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	关于数字化文本的简要描述。			
著录信息源	用户定义			
著录细则				
DC 映射	Description			



6. 日期

命名空间	暂定			
名称	日期			
标识	Date			
定义	与数字化文本有关的日期			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	issued	可选	数字对象发布日期	[DCMI]
	created	可选	数字对象创建日期	[DCMI]
	Valid	可选	数字对象有效日期	[DCMI]
	Modified	可选	数字对象修改日期	[DCMI]
	dateSubmitted	可选	数字对象提交日期	[DCMI]
	DateReplied	可选	答辩日期	
编码修饰词	标识	名称	参照	
	ISO8601	W3C 定义的日期描述格式	[W3CDTF]	
	Period	DCMI 定义时期描述	[DCMI-DCSV]	
数据类型	ISO8601 : YYYY-MM-DD、YYYY-MM、YYYY Period :			
指定值	-			
缺省值	-			
注释	Date 用于注明数字化文本印刷型文献的出版日期,与数字化文本相关的日期均通过元素修饰词进行著录,主要有数字对象的发布日期 (issued)、创建日期(created)、有效日期(valid)、修改日期(modified)、可用日期(available)、答辩日期(datereplied)和提交			



	日期(datesubmitted)。其中答辩日期与提交日期为学位论文专用。
著录信息源	用户定义
著录细则	
DC 映射	Date

7. 资源形式

命名空间	暂定			
名称	资源形式			
标识	Format			
定义	数字化文本的电子形态。			
必备性	必备			
可重复性	No			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	标识	名称		参照
	MIME	Internet 媒体类型		[MIME]
数据类型	字符型			
指定值	Text/html(.html,.htm) Application/msword(.doc) Application/pdf(.pdf) Application/postscript.(ai,.eps,.ps) Application/powerpoint(.ppt) Application/rtf(.rtf) Application/wordperfect5.1(.wp5.1) Application/x-latex(.latex) Image/Djvu(.djvu) Image/gif(.gif) Image/.(.jpeg,.jpg,.jpe)			



	Text/plain(.txt) Text/sigml(.sgml,.sgm) 参照 MIME ¹
缺省值	Image/Djvu(.djvu)
注释	数字化文本的电子形态
著录信息源	用户定义
著录细则	
DC 映射	Format

8 . 资源类型

命名空间	暂定
名称	资源类型
标识	Type
定义	有关资源内容的特征和类型
必备性	必备
可重复性	No
属性	-
应用层次	数字化文本
元素修饰词	-
编码修饰词	-
数据类型	字符型
指定值	数字化文本、普通图书、学位论文、古籍、民国图书、.....
缺省值	数字化文本
注释	资源类型主要用于描述资源所属的类别，根据本草案的适用范围，它可以是任意形式的数字化文本。

¹ Multipurpose Internet Mail Extensions (<http://www.nacs.uci.edu/indiv/ehood/MIME/MIME.html>)



著录信息源	用户定义
著录细则	
DC 映射	Type

9. 资源来源

命名空间	暂定
名称	资源来源
标识	Source
定义	描述资源的出处信息
必备性	可选
可重复性	Yes
属性	-
应用层次	数字化文本
元素修饰词	-
编码修饰词	-
子元素	-
数据类型	字符型
指定值	-
缺省值	-
注释	可以是来源资源的全部或一部分。
著录信息源	用户定义
著录细则	
DC 映射	Source



10. 语言

命名空间	暂定			
名称	语言			
标识	Language			
定义	数字化文本的文字语种			
必备性	必备			
可重复性	Yes			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	标识	名称	参照	
	ISO639-2	语言编码	[ISO 639-2:1998]	
	RFC3066	语言编码	[RFC 3066]	
	RFC1766	语言编码	[RFC 1766]	
数据类型	字符型			
指定值	-			
缺省值	chi			
注释	建议采用规范的表达形式,如 RFC3066、RFC1766 和 ISO639-2 的语种代码。			
著录信息源	用户定义			
著录细则				
DC 映射	Language			



11. 资源标识

命名空间	暂定			
名称	资源标识			
标识	Identifier			
定义	确认资源的唯一标识			
必备性	必备			
可重复性	Yes			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	标识	名称	参照	
	ISBN	国际书目号		
	ISSN	国际连续出版物号		
	Bookid	电子书号	[CADAL]	
	URI	统一资源符		
	URL	统一资源定位符		
	DOI			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	用以标识资源的字符串或数字，应采用标准的编码体系,属性“xsi:type”用于指定标识类型的编码类型，其指定值为：ISBN、ISSN、URI、URL等。			
著录信息源	用户定义			



著录细则	
DC 映射	Identifier

12 . 时空范围

命名空间	暂定			
名称	时空范围			
标识	Coverage			
定义	数字化文本内容所涉及的时间、空间覆盖范围			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	Temporal	可选	时间范围	[DCMI]
	Spatial	可选	空间范围	[DCMI]
编码修饰词	标识	名称	参照	
	Period	时期	[DCMI-DCSV]	
	TGN	Getty 地理名称辞典	[TGN]	
数据类型	字符型			
指定值	-			
缺省值	-			
注释	时间范围的著录可采用 period 编码修饰，空间范围的著录可采用 TGN 编码修饰。			
著录信息源	用户定义			



著录细则	
DC 映射	Format

13 . 出版者

命名空间	暂定
名称	出版者
标识	Publisher
定义	使数字化文本印刷型文献成为可取得或利用的责任者。
必备性	可选
可重复性	YES
属性	-
应用层次	数字化文本
元素修饰词	-
编码修饰词	-
数据类型	字符型
指定值	-
缺省值	-
注释	用于注明数字化文本所对应的印刷型文献的出版者或颁布者。
著录信息源	书名页，无书名页的依次取自版权页、封面、序言等
著录细则	
DC 映射	Publisher



14 . 相关资源

命名空间	暂定			
名称	相关资源			
标识	Relation			
定义	与数字化文本存在某种关系的其他电子文献资源。			
必备性	可选			
可重复性	YES			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	HasPart	可选	包含	[DCMI]
	IsPartOf	可选	包含于	[DCMI]
	reference	可选	参照	[DCMI]
	Isreferencedby	可选	被参照	[DCMI]
	Hasversion	可选	其他版本	[DCMI]
	isVersionOf	可选	原版本	[DCMI]
编码修饰词	标识	名称	参照	
	URI	统一资源符		
	URL	统一资源定位符		
	DOI	数字对象标识符		
	ISBN	国际书目号	[DCMI-DCSV]	
	bookid	CADAL 电子书目号	[CADAL]	
数据类型	字符型			
指定值	-			



缺省值	-
注释	用于揭示数字化文本同其它资源之间的关联。一般用字符串或数字代码标识相关资源。如果采用规范定义的编码体系，如统一资源标识（URI）、数字对象标识（DOI）、国际标准书号（ISBN）等，则应在元素中注明编码修饰名称。能用本元素的元素修饰词表达的相关资源应采用相应的元素修饰词。
著录信息源	用户定义
著录细则	
DC 映射	Relation

15. 权限管理

命名空间	暂定			
名称	权限管理			
标识	Rights			
定义	数字化文本本身所有的或被赋予的权限信息。			
必备性	可选			
可重复性	YES			
属性	-			
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	AcessRights	可选	访问权限	[DCMI]
	CopyrightOwner	可选	文本版权拥有者	
	CopyrightExpiryDate	可选	文本版权解除日期	
	createcentre	可选	制作中心	
	Location	可选	收藏单位	
编码修饰词	-			
数据类型	字符型			



指定值	-
缺省值	-
注释	包括知识产权、版权、使用权限说明及其它相关产权。可指一个权限管理的陈述及其标识，也可指提供权限管理信息内容的服务器标识。
著录信息源	用户定义
著录细则	
DC 映射	Rights

16 . 版本说明

命名空间	暂定			
名称	版本说明			
标识	Edition			
定义	数字对象原件的版本说明信息，包括它的收藏历史			
必备性	可选			
可重复性	YES			
属性	-			
应用层次	数字化文本			
元素修饰词	标识	必备性	名称	参照
	History	可选	收藏历史	
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	版本信息包括版次和版本说明；收藏历史在元素修饰 history 中著录。			
著录信息源				



著录细则	
DC 映射	-

17 . MARC 标识

命名空间	暂定			
名称	MARC 标识			
标识	marc			
定义	与数字化文本所对应文献 MARC 记录文件的文件名称。			
必备性	可选			
可重复性	NO			
属性	标识	必备性	名称	参照
	Xsi:Type	可选	编码类型	[W3C]
应用层次	数字化文本			
元素修饰词	-			
编码修饰词	标识	名称	参照	
	CNMARC	中国 MARC 格式		
	MARC21	MARC21 格式		
编码修饰词	-			
数据类型	字符型			
指定值	-			
缺省值	-			
注释	MARC 记录文件存放在 meta 目录之下，如果存在 MARC 记录，则著录。			
著录信息源				
著录细则				



DC 映射	-
-------	---

3.4.4 属性参数说明

- 公共属性：xsi:type
编码修饰词采用属性的方式著录，属性名称为 xsi:type，其中 xsi 是指 type 的命名空间，xsi=http://www.w3.org/2001/XMLSchema-instance.
- 专用属性：role
role 属性，隶属于<creator>和<contributor>元素，用于区分创作者或贡献者的贡献方式（如作者、编者、图解者、电子图书制作者等）
 - <dc:creator role="aut">李卫</dc:creator>

3.4.5 元素修饰词列表

标识	必备性	名称	应用元素	参照
Alternative	可选	交替题名	Title	[DCMI]
Abstract	可选	内容摘要	Description	[DCMI]
issued	可选	数字对象发布日期	Date	[DCMI]
created	可选	数字对象创建日期		[DCMI]
Valid	可选	数字对象有效日期		[DCMI]
Modified	可选	数字对象修改日期		[DCMI]
dateSubmitted	可选	数字对象提交日期		[DCMI]
DateReplied	可选	答辩日期		
Temporal	可选	时间范围		Coverage
Spatial	可选	空间范围	[DCMI]	
HasPart	可选	包含	relation	[DCMI]
IsPartOf	可选	包含		[DCMI]
reference	可选	参照		[DCMI]



Isreferencedby	可选	被参照		[DCMI]
Hasversion	可选	其他版本		[DCMI]
isVersionOf	可选	原版本		[DCMI]
AcessRights	可选	访问权限	Rights	[DCMI]
CopyrightOwner	可选	文本版权拥有者		
CopyrightExpiry Date	可选	文本版权解除日期		
createcentre	可选	制作中心		
location	可选	收藏单位		
History	可选	收藏历史	Edition	

3.4.5 编码修饰词列表

编码标识	名称	应用元素	参照
CCS	中国主题词表	Subject	
CLC	中国图书分类法		
SKC	四库分类法(北大)		
LCSH			[DCMI]
MESH			[DCMI]
DDC			[DCMI]
LCC			[DCMI]
UDC			[DCMI]
ISO8601	W3C 定义的日期描述格式	Date	[W3CDTF]
Period	DCMI 定义时期描述	Date, coverage	[DCMI-DCSV]
MIME	Internet 媒体类型	Format	[MIME]



ISO639-2	语言编码	Language	[ISO 639-2:1998]
RFC3066	语言编码	Language	[RFC 3066]
RFC1766	语言编码	Language	[RFC 1766]
ISBN	国际书目号	Identifier, relation	
ISSN	国际连续刊物号	Identifier, relation	
Bookid	电子书目号	Identifier	[CADAL]
URI	统一资源符	Identifier, relation	
URL	统一资源定位符	Identifier, relation	
DOI		Identifier,relation	
TGN	地理名称辞典	Coverage	[TGN]
CNMARC	中国 MARC 格式	Marc	
MARC21	MARC21 格式		



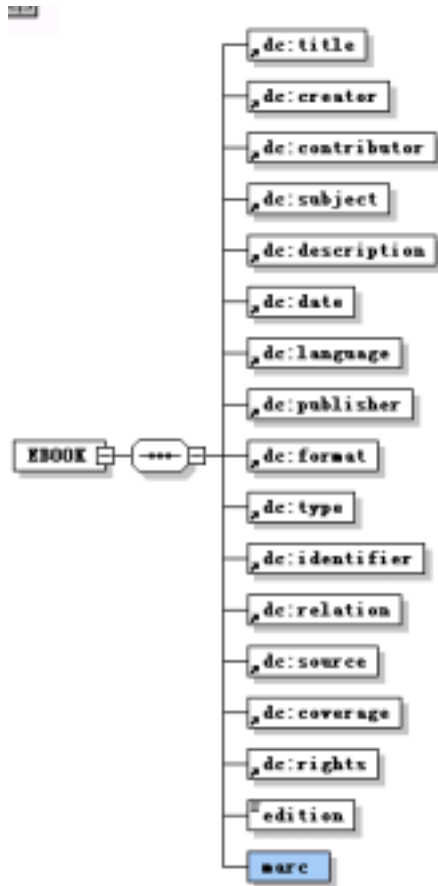
4 . CADAL 元数据 Schema

4.1 Schema 源码

```
<?xml version="1.0" encoding="GB2312"?>
<!-- edited with XML Spy v4.3 U (http://www.xmlspy.com) by LEGO (LEGO) -->
<xs:schema targetNamespace="http://www.cadal.net/XMLSchema/EB"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns="http://www.cadal.net/XMLSchema/EB"
xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
attributeFormDefault="unqualified">
  <xs:annotation>
    <xs:documentation xml:lang="en">
      CADAL元数据方案选择了dublin Core的15项核心元素和扩展了4个元素
    </xs:documentation>
  </xs:annotation>
  <xs:import namespace="http://www.w3.org/XML/1998/namespace"
schemaLocation="http://www.w3.org/2001/03/xml.xsd"/>
  <xs:import namespace="http://purl.org/dc/elements/1.1/"
schemaLocation="http://dublincore.org/schemas/xmls/qdc/2003/04/02/dc.xsd"/>
  <xs:element name="EBOOK">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="dc:title"/>
        <xs:element ref="dc:creator"/>
        <xs:element ref="dc:contributor"/>
        <xs:element ref="dc:subject"/>
        <xs:element ref="dc:description"/>
        <xs:element ref="dc:date"/>
        <xs:element ref="dc:language"/>
        <xs:element ref="dc:publisher"/>
        <xs:element ref="dc:format"/>
        <xs:element ref="dc:type"/>
        <xs:element ref="dc:identifier"/>
        <xs:element ref="dc:relation"/>
        <xs:element ref="dc:source"/>
        <xs:element ref="dc:coverage"/>
        <xs:element ref="dc:rights"/>
        <xs:element name="edition" type="xs:string"/>
        <xs:element name="marc"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
```

</xs:schema>

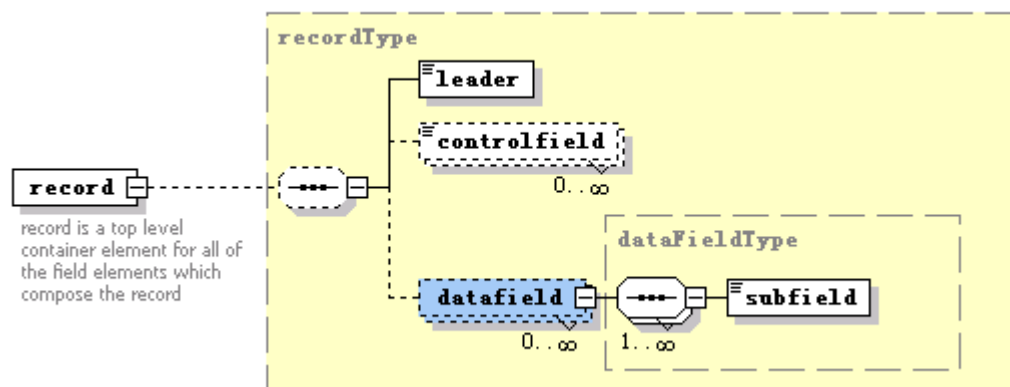
4.2 CADAL 元数据结构图



4.3 MARC XML Schema

MARC XML Schema 是 LOC 下属的网络发展与 MARC 标准办公室 (ND&MSO) 开发研制的规范, 其主要作用是实现 MARC2709 格式在 XML 环境下的互操作性。同时 ND&MSO 建立了一个基于 MARCXML 的操作框架, 该框架具有相当灵活性和可扩展性, 允许用户根据他们自己的需求与 MARC 数据协同工作。详细内容参见 <http://www.loc.gov/marc>。

MARC XML Schema 的结构图如下：



结构详细说明参见：<http://www.loc.gov/standards/marcxml/xml/spy/spy.html>



5. 著录规则

本规范主要应用于印刷版图书的数字化文本。

本规范由 17 个核心元素组成，其中采用 DC 的 15 个元素用于著录“数字化文本”信息，另外两个元素<edition>和<MARC>的著录对象分别为和数字化文本相对应的“印刷型文献对象”的版本信息和 MARC 信息。因此，本规范的著录规则分为两个部分：

- DC 的 15 项元素用于著录“数字化对象信息”，著录规则见上述相关的著录细则
- MARC 记录的著录规则
 - 普通图书的 MARC 著录规则：参见《中国机读目录通讯格式》(CN-MARC)、《中国机读目录格式使用手册》(ISBN-80039-990-7)和《中华人民共和国国家标准：文献著录总则》(GB3792.1-83)、《普通图书著录规则》(GB3792.2-85)著录。
 - 古籍的 MARC 著录规则：参见《古籍著录规则》
 - 学位论文的 MARC 著录规则：参见《中华人民共和国国家标准：文献著录总则》(GB3792.1-83)、《普通图书著录规则》(GB3792.2-85)著录。

6. 分类主题标引要求

普通图书分类采用《中图法》第四版；主题采用《中国分类主题词》。
学位论文分类采用《中图法》第四版；主题采用《中国分类主题词》。
古籍分类采用四库分类法；主题采用《中国分类主题词》。

7. 著录实例

7.1 元数据 XML 文件

```
<?xml version="1.0" encoding="utf-8" ?>
<dublincore xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title>海峡两岸档案管理暨缩影技术交流会文集</dc:title>
  <dc:creator>中国档案学会</dc:creator>
  <dc:contributor />
  <dc:subject>档案管理</dc:subject>
  <dc:subject>档案资料</dc:subject>
  <dc:subject>缩微摄影</dc:subject>
  <dc:description />
  <dc:publisher>中国档案出版社</dc:publisher>
  <dc:date>1994</dc:date>
  <dc:format>Image/Djvu(.djvu)</dc:format>
  <dc:type>普通图书</dc:type>
  <dc:language>chi</dc:language>
  <dc:source />
  <dc:relation />
  <dc:coverage />
  <dc:identifier xsi:type="ISBN">7-80019-470-1</dc:identifier>
  <dc:rights />
  <marc xsi:type="cnmarc">cnmarc.iso</marc>
</dublincore>
```

7.2 电子图书 OPF 文件

```
<?xml version="1.0" encoding="UTF-8" ?>
<package unique-identifier="06000079"
  xmlns:oebpackage="http://openebook.org/namespaces/oeb-package/1.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <metadata>
    <dc-metadata xmlns:dc="http://purl.org/dc/elements/1.0/">
      <dc:title>海峡两岸档案管理暨缩影技术交流会文集</dc:title>
      <dc:creator>中国档案学会</dc:creator>
```



```
<dc:contributor />
<dc:subject>档案管理</dc:subject>
<dc:subject>档案资料</dc:subject>
<dc:subject>缩微摄影</dc:subject>
<dc:description />
<dc:publisher>中国档案出版社</dc:publisher>
<dc:date>1994</dc:date>
<dc:format>Image/Djvu(.djvu)</dc:format>
<dc:type>普通图书</dc:type>
<dc:language>chi</dc:language>
<dc:source />
<dc:relation />
<dc:coverage />
<dc:identifier xsi:type="ISBN">7-80019-470-1</dc:identifier>
<dc:rights />
</dc-metadata>
<x-metadata>
  <meta content="cnmarc.iso" name="marc" schema="" />
</x-metadata>
</metadata>
<manifest>
  <item id="content0" href="ptiff\00000001.djvu" media-type="text/x-oeb1-document"/>
  .....
  <item id="toc" href="HTML\00000001.htm" media-type="text/x-oeb1-document"/>
</manifest>
<spine>
  <itemref idref="toc"/>
</spine>
<guide>
  <reference type="cover" title="封面" href="PTIFF\00000001.djvu"/>
  <reference type="toc" title="目录" href="PTIFF\00000023.djvu"/>
  <reference type="Other" title="封底" href="PTIFF\00000487.djvu"/>
</guide>
</package>
```

7.3 四库全书元数据著录实例

```
<?xml version="1.0" encoding="utf-8" ?>
<dublincore xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title>文淵閣四庫全書·第七二三册</dc:title>
  <dc:creator />
  <dc:subject>四庫全書</dc:subject>
```



```
<dc:subject>子部</dc:subject>
<dc:subject>儒家類</dc:subject>
<dc:description />
<dc:publisher>臺灣商務印書館</dc:publisher>
<dc:contributor />
<dc:date />
<dc:type>影印古籍</dc:type>
<dc:format>Image/Djvu(.djvu)</dc:format>
<dc:identifier>06300001</dc:identifier>
<dc:source />
<dc:language>chi</dc:language>
<dcterms:ispartof>四庫全書</dcterms:ispartof>
<dcterms:ispartof>子部</dcterms:ispartof>
<dcterms:ispartof>儒家類</dcterms:ispartof>
<dcterms:haspart>御覽經史講義(二)/(清)蔣溥等奉敕編
</dcterms:haspart>
<dc:coverage />
<edition>影印版</edition>
<createcentre>浙江大學圖書館</createcentre>
<location>浙江大學圖書館</location>
</dublincore>
```

```
<?xml version="1.0" encoding="utf-8" ?>
<dublincore xmlns:dc="http://purl.org/dc/elements/1.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dcterms="http://purl.org/dc/terms/">
<dc:title>文淵閣四庫全書·第七二二冊</dc:title>
<dc:creator />
<dc:subject>四庫全書</dc:subject>
<dc:subject>子部</dc:subject>
<dc:subject>儒家類</dc:subject>
<dc:description />
<dc:publisher>台灣商務印書館</dc:publisher>
<dc:contributor />
<dc:date />
<dc:type>影印古籍</dc:type>
<dc:format>Image/Djvu(.djvu)</dc:format>
<dc:identifier>06000079</dc:identifier>
<dc:language>chi</dc:language>
<dcterms:ispartof>文淵閣四庫全書</dcterms:ispartof>
<dcterms:ispartof>子部</dcterms:ispartof>
<dcterms:ispartof>儒家類</dcterms:ispartof>
<dcterms:haspart>欽定執中成憲/(清)雍正六年敕撰
</dcterms:haspart>
```



```
<dcterms:haspart>御览经史讲义(一)/溥等奉敕编</dcterms:haspart>
<edition>影印版</edition>
<createcentre>浙江大学图书馆</createcentre>
<location>浙江大学图书馆</location>
</dublincore>
```

说明 1：丛书类信息著录在 ispartof 中，丛书如果分为多层结构，则可以分别描述，如“文渊阁四库全书·子部·儒家类”，则描述为：

```
<dcterms:ispartof>文渊阁四库全书</dcterms:ispartof>
<dcterms:ispartof>子部</dcterms:ispartof>
<dcterms:ispartof>儒家类</dcterms:ispartof>
```

说明 2：如果这本书的由多个分别来自于不同作者的内容组成，则可以将这些信息著录在 haspart 元素之中。例如“文渊阁四库全书·第七二二册”中由“钦定执中成宪/(清)雍正六年敕撰”和“御览经史讲义(一)/溥等奉敕编”两个内容组成，则可以描述为：

```
<dcterms:haspart>钦定执中成宪/(清)雍正六年敕撰</dcterms:haspart>
<dcterms:haspart>御览经史讲义(一)/溥等奉敕编</dcterms:haspart>
```

7.4 学位论文著录实例

```
<?xml version="1.0" encoding="utf-8" ?>
<dublincore xmlns:dc="http://purl.org/dc/elements/1.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dcterms="http://purl.org/dc/terms/">
<dc:title>森林环境可持续发展管理的理论与实践</dc:title>
<dcterms:alternative>Theory and Practice on Sustainable
Developmental Management of Forest
Environment</dcterms:alternative>
<dc:creator>郑四渭</dc:creator>
<dc:subject>森林环境</dc:subject>
<dc:subject>可持续发展</dc:subject>
<dc:subject>环境管理</dc:subject>
<dc:subject>森林环境经济系统</dc:subject>
<dc:description>专业：管理学院农经系</dc:description>
<dc:description>学位：管理学硕士</dc:description>
<dcterms:abstract>本论文旨以已有理论研究与实践成果为基础，研究森林环
境可持续发展管理的理论和实践问题，主要研究结论是： 1．可持续发展理
论与宇宙飞船经济论、增长极限论、环境危机论、外部性理论存在渊源关系，
可持续发展理论是这些理论研究的回归。 2．市场与政府对森林环境变迁的
作用均存在正负影响，当长期正面影响起主导作用时，森林环境呈正向变迁态
势；反之，森林环境则呈逆向变迁态势。 3．提出森林环境供求均衡基本模
型 S D，认为形成森林环境最优供给与有效需求是缓解我国森林环境供求矛
盾的重要战略抉择。前者的实现借助于发挥市场作用、提高森林环境生产及相
应管理机构的效率。后者的促成依赖于对消费、投资与净出口的限定与修正。
```



4. 森林环境利用四种放任规则是森林环境非持续利用制度生成的原因, 而其生成的背景是三种特定经济环境交互作用。 5. 森林环境持续利用制度框架结构以强制性制度为主体, 由组织营运制度等 8 项制度构成。 6. 根据森林环境价值的内涵, 提出了森林环境价值的要素累加计量模型与计价方法, 并获得应用研究结果。 7. 应用森林环境外部效应内部化理论实证分析了我国退耕还林(草)政策, 提出了消除现行政策弊端的若干对策。

</dcterms:abstract>

<dcterms:abstract>This paper discusses the theoretical framework and practical problems of the sustainable developmental management for the forest environment aiming at the forest environmental problems, objectives of which are to resolve the forest environmental problems and to realize sustainable utilization of the forest environment in China. Some main conclusions in this paper are as follows: 1. There is a source relationship theory of sustainable development with the limit to growth, theory of environmental crisis, theory of externality, theory of sustainable development is a regression of those theories.....</dcterms:abstract>

<dc:publisher>浙江大学</dc:publisher>

<dc:contributor>袁飞 教授</dc:contributor>

<dc:date>20010101</dc:date>

<dc:type>学位论文</dc:type>

<dc:format>Image/Djvu(.djvu)</dc:format>

<dc:identifier xsi:type="bookid">06000079</dc:identifier>

<dc:language>chi</dc:language>

<copyrightowner>浙江大学</copyrightowner>

<createcentre>浙江大学图书馆</createcentre>

<location>浙江大学档案馆</location>

</dublincore>

7.5 普通图书著录实例

<?xml version="1.0" encoding="utf-8" ?>

= <dublincore xmlns:dc="http://purl.org/dc/elements/1.0/"

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns:dcterms="http://purl.org/dc/terms/">

<dc:title>中国文化思辨：从文学走向传播</dc:title>

<dc:creator role="著">朱剑飞</dc:creator>

<dc:contributor />

<dc:subject>文学研究</dc:subject>

<dc:subject>电视</dc:subject>

<dc:description>本书收集了作者在文学及传播业方面的研究文集。包括“历史的反思需要勇于面对自我”，“浅议中国当代文学专业教材对胡风文艺思想的



```
批评”等十四篇论文。 </dc:description>
<dc:publisher>广州出版社</dc:publisher>
<dc:date>2002</dc:date>
<dc:format>Image/Djvu(.djvu)</dc:format>
<dc:type>普通图书</dc:type>
<dc:language>chi</dc:language>
<dc:source />
<dc:relation />
<dc:coverage />
<dc:identifier xsi:type="ISBN">7-80655-406-8</dc:identifier>
<createcentre>浙江大学图书馆</createcentre>
<location>浙江大学档案馆</location>
<marc xsi:type="cnmarc">2140445E02.iso</marc>
</dublincore>
```

参考文献:

1. CALIS 联合目录中文图书著录细则(2001 年 12 月修订)
2. **Simple** **DC** **XML** **Schema,**
<http://dublincore.org/schemas/xmls/simpledc20021212.xsd>
3. MARCXML **Schema,**
<http://www.loc.gov/standards/marcxml///schema/MARC21slim.xsd>
4. MARC XML Illustrated, <http://www.loc.gov/standards/marcxml///xml/spy/spy.html>
5. OEBPS 规范, http://www.openebook.org/oebf_groups.htm
6. 中文古籍图书著录细则 / 北图试用稿
7. 《电子图书元数据规范》,《古籍元数据规范》, <http://cdls.nstl.gov.cn>
8. MARC **XML** **Architecture.**
<http://www.loc.gov/standards/marcxml//marcxml-architecture.html>
9. MARC **XML** **Schema.**
<http://www.loc.gov/standards/marcxml//schema/MARC21slim.xsd>