



# 中华人民共和国国家标准

GB/T 38371.2—2020

---

## 数字内容对象存储、复用与交换规范 第2部分：对象封装、存储与交换

Specification of digital content object storage, reuse and exchange—  
Part 2: Object encoding, storage and exchange

2020-03-31 发布

2020-10-01 实施

国家市场监督管理总局  
国家标准化管理委员会 发布

目 次

前言 ..... I

引言 ..... II

1 范围 ..... 1

2 规范性引用文件 ..... 1

3 术语和定义 ..... 1

4 缩略语 ..... 1

5 数字内容对象封装 ..... 2

    5.1 对象封装方式 ..... 2

    5.2 交换信息包文件结构 ..... 2

    5.3 证实方法 ..... 5

6 数字内容对象存储 ..... 5

    6.1 信息包存储 ..... 5

    6.2 实例文件存储 ..... 5

    6.3 数据文件存储 ..... 7

    6.4 证实方法 ..... 7

7 数字内容对象交换 ..... 7

    7.1 交换方式 ..... 7

    7.2 资源标识符获取接口 ..... 8

    7.3 指定资源 IRI 获取接口 ..... 10

    7.4 元数据获取接口 ..... 11

    7.5 数据文件 URL 获取接口 ..... 11

    7.6 数据流获取接口 ..... 12

    7.7 交换信息包获取接口 ..... 13

    7.8 交换信息包上传接口 ..... 13

    7.9 资源检索接口 ..... 14

参考文献 ..... 19

## 前 言

GB/T 38371《数字内容对象存储、复用与交换规范》分为 3 个部分：

- 第 1 部分：对象模型；
- 第 2 部分：对象封装、存储与交换；
- 第 3 部分：对象一致性检查方法。

本部分为 GB/T 38371 的第 2 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分由国家新闻出版署提出。

本部分由全国新闻出版信息标准化技术委员会(SAC/TC 553)归口。

本部分起草单位：北京大学、中宣部机关服务中心(中宣部信息中心)、北京拓标卓越信息技术研究院、中国科技出版传媒股份有限公司。

本部分主要起草人：王文清、陈鹏飞、刘成勇、刘勇、张沫、周长岭、关涛、梁赓、姜磊、陆新民、郭昱锦、张志。

## 引 言

在数字内容资源的管理和应用等过程中,涉及的相关资源种类包括图书、论文、报告、数据集、应用程序等;资源媒体类型包括文本、音频、视频、图像、虚拟现实或增强现实资源等;可复用的资源类型包括各种构件类资源,如章节、图表、公式等资源片段或构件;用于资源组织类的资源包括受控术语表、主题词表、分类法等;其他相关资源类型还包括代理者、元数据、权利等。对于数字内容资源,需要建立一个统一的对象模型,以系统地表示和记录资源的结构、特征以及对象资源之间的各种关系,并能够按不同的粒度进行封装、存储和交换。

GB/T 38371 参考了国内外相关标准,结合新闻出版业、图书馆和博物馆等涉及数字内容资源的相关领域针对数字内容资源管理与服务的需求,从对象模型、对象封装、存储和交换以及对象一致性检查等多个方面给出一套统一的标准,以实现相关领域对数字内容资源的规范化表示,消除数据复用和交换障碍,降低资源管理和利用成本,推动数字内容资源的融合和深入利用。

库七七 www.kq9w.com 提供下载



# 数字内容对象存储、复用与交换规范

## 第2部分：对象封装、存储与交换

### 1 范围

GB/T 38371 的本部分规定了数字内容对象的封装、存储与交换方式并给出了对象交换服务接口。

本部分适用于新闻出版业、图书馆和博物馆等领域有关数字内容资源的数字化统一管理和共享。

### 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 7408—2005 数据元和交换格式 信息交换 日期和时间表示法 (ISO 8601:2000, IDT)

GB/T 38371.1—2020 数字内容对象存储、复用与交换规范 第1部分：对象模型

GB/T 38371.3—2020 数字内容对象存储、复用与交换规范 第3部分：对象一致性检查方法

### 3 术语和定义

下列术语和定义适用于本文件。

#### 3.1

**数据文件 data file**

包含资源数字信息的文件。

#### 3.2

**信息包 information package**

通过打包信息进行封装和识别的一种逻辑容器。

注：信息包包括内容信息和保存描述信息。

#### 3.3

**交换信息包 exchange information package; EIP**

在不同系统之间，用于数据交换的信息包。

#### 3.4

**存档信息包 archive information package; AIP**

为资源存档和长期保存而规定的逻辑存储格式。

#### 3.5

**DCOO RDF 实例文件 DCOO based RDF instance file**

基于数字内容对象本体模型的符合 RDF 表示规范的实例数据存储文件。

### 4 缩略语

下列缩略语适用于本文件。

AIP:存档信息包(Archive Information Package)  
DCOO:数字内容对象本体(Digital Content Object Ontology)  
EIP:交换信息包(Exchange Information Package)  
MIME:多用途互联网邮件扩展(Multipurpose Internet Mail Extensions)  
OAIS:开放档案信息系统(Open Archival Information System)  
RDF:资源描述框架(Resource Description Framework)  
SRU:基于 URL 的搜索/检索(Search/Retrieve via URL)  
SRW:搜索/检索 Web 服务(Search/Retrieve Web Service)  
URL:统一资源定位符(Uniform Resource Locator)

## 5 数字内容对象封装

### 5.1 对象封装方式

#### 5.1.1 容器级封装

容器级封装用以封装数字内容对象及其相关的所有对象信息。

#### 5.1.2 对象级封装

对象级封装用以封装数字内容对象及其所包含的单一媒体对象信息。

### 5.2 交换信息包文件结构

#### 5.2.1 交换信息包类型

根据交换信息包所封装的资源粒度,可分为容器级交换信息包与对象级交换信息包。

#### 5.2.2 容器级交换信息包结构

##### 5.2.2.1 容器级交换信息包封装了一个完整的容器级封装所包含的所有信息,其结构如下:

```
Container_Local_ID.zip/  
  mimetype  
  META-INF/  
    container.xml  
    dcoomanifest.xml  
  CONTAINER/  
    dcoo_rdf_instance.xml(实例文件)  
  dataFile/  
    audio/  
      音频数据文件  
    video/  
      视频数据文件  
    image/  
      图像数据文件  
    other/  
      其他类型数据文件
```

##### 5.2.2.2 “Container\_Local\_ID”为容器级封装的本地标识符;交换信息包以“Container\_Local\_ID.zip”

命名,采用 ZIP 压缩格式对信息包进行压缩。

5.2.2.3 每个交换信息包均含有一个名为 mimetype 的文件,其中内容为固定值“application/dcoo+zip”,用于说明交换信息包文件的格式;mimetype 文件不含新行或回车,自身不作压缩。

5.2.2.4 “META-INF”文件目录用于存放容器级封装相关的信息,其中包含 container.xml 和 dcoo-manifest.xml 两个文件:

a) container.xml:其文件内容如下面的 XML 片段所示:

```
<? xml version="1.0" encoding="utf-8"? >
<container version="1.0" xmlns="http://dcoo_authority_domain/standards/dcoo/package/container" >
  <rootfiles>
    <rootfile full-path="CONTAINER/dcoo_rdf_instance.xml"
      media-type="application/dcoopackage+xml"
      package-type="dcoo-container+xml"/>
  </rootfiles>
</container>
```

——full-path:当前容器级封装的 DCOO RDF 实例文件的相对路径;

——media-type:该交换信息包文件的打开方式;

——package-type:表示包类型,容器级交换信息包类型为“dcoo-container+xml”。

b) dcoo-manifest.xml:该清单文件给出了该容器级封装所包含的媒体对象文件(即数据文件)列表及对象实例关系列表,文件命名为“dcoo-manifest.xml”。该文件的顶级元素为“<manifest>”,其属性包括:

——identifier:取值为数字内容对象实例 IRI,在该内容清单中是唯一的;

——version:内容清单的版本号,用来区分具有相同标识符的内容清单。

“<manifest>”下包含子元素“<dataFileList>”(数据文件列表),示例 XML 片段如下所示:

```
<dataFileList>
  <item id="&.myNamespace;example/mediaObj_1"
    href="dataFile/image/mediaObj_1.jpeg"
    media-type="image/jpeg" />
  <item id="&.myNamespace;example/mediaObj_2"
    href="dataFile/other/mediaObj_2.pdf"
    media-type="application/pdf" />
  .....
</dataFileList>
```

——id:必备属性;每一个“item”对应一份数据文件,“id”为当前媒体对象的 IRI (“&.myNamespace;”表示应用单位的命名空间取值);

——href:必备属性;取值为当前媒体对象对应数据文件的相对路径(相对于 CONTAINER 文件目录);

——media-type:必备属性;取值为当前数据文件的 MIME 类型取值。

“<manifest>”下包含的另一个子元素“<objectList>”(对象清单),示例 XML 片段如下所示:

```

<objectList>
  <item subjectId="&.myNamespace;example/book_1" >
    <relation predicate="http://dcoo_authority_domain/standards/dcoo#p5_associatedMedia"
      objectId="&.myNamespace;example/mediaObj_1"/>
    .....
  </item>
  .....
</objectList>

```

- subjectId:必备属性;当前主体对象实例的 IRI(“&.myNamespace;”表示应用单位的命名空间取值);subjectId 来自于当前 DCOO RDF 实例文件,即 dcoo\_rdf\_instance.xml;
- predicate:必备属性;当前“<relation>”(关系)所对应的属性类型;
- objectId:必备属性;当前客体对象实例的 IRI;objectId 是“<objectList>”中某一个“<item>”的 subjectId。

#### 5.2.2.5 “CONTAINER”文件目录包含两部分:

- dcoo\_rdf\_instance.xml:当前容器级封装实例的 DCOO RDF 实例文件;
- dataFile 文件目录:其下的文件子目录包括 audio、video、image、other;每个文件子目录下包含对应类型的数据文件。

### 5.2.3 对象级交换信息包结构

#### 5.2.3.1 对象级交换信息包用以封装数字内容对象及其所包含的单一媒体对象(及相关数据文件)信息,其结构如下:

```

Object_Local_ID.zip/
  mimetype
  META-INF/
    container.xml
    dcoomanifest.xml
  CONTAINER/
    dcoo_rdf_instance.xml(实例文件)
    dataFile/
      [media_type_label]/
        某一媒体类型数据文件

```

#### 5.2.3.2 “Object\_Local\_ID”为数字内容对象的本地标识符;交换信息包以“Object\_Local\_ID.zip”命名,采用 ZIP 压缩格式对信息包进行压缩。

#### 5.2.3.3 “mimetype”文件的内容及要求与容器级交换信息包中同名文件一致。

#### 5.2.3.4 “META-INF”文件目录用于存放当前数字内容对象相关信息,其中包含 container.xml 和 dcoomanifest.xml 两个文件:

- a) container.xml:其文件内容如下面的 XML 片段所示:

```

<? xml version="1.0" encoding="utf-8"? >
<container version="1.0" xmlns="http://dcoo_authority_domain/standards/dcoo/package/container">
  <rootfiles>
    <rootfile full-path="CONTAINER/dcoo_rdf_instance.xml"
      media-type="application/dcoopackage+xml"
      package-type="dcoo-object+xml"
      class="http://dcoo_authority_domain/standards/dcoo#E19_TextEntity"/>
  </rootfiles>
</container>

```

- full-path 及 media-type 属性含义及内容与容器级交换信息包中对应属性一致；
- package-type:表示包类型;对象级交换信息包类型为“dcoo-object+xml”;
- class:在以上示例中,class 属性表示当前对象实例的类型为 DCOO 所定义的某一对象类型。

b) dcoomanifest.xml:该清单文件的内容格式与容器级交换信息包中同名文件要求一致。

#### 5.2.3.5 “CONTAINER”文件目录包含两部分:

- dcoo\_rdf\_instance.xml:当前对象实例的 DCOO RDF 实例文件;
- dataFile 文件目录:其下的文件子目录名“[media\_type\_label]”可为“audio”“video”“image”或“other”中的一种;该文件子目录下包含对应类型的数据文件。

### 5.3 证实方法

数字内容对象封装的证实方法见 GB/T 38371.3—2020 的 6.2。

## 6 数字内容对象存储

### 6.1 信息包存储

信息包的存储包括数字内容对象(及其相关资源)的实例文件及数据文件的存储。实现这两类文件的逻辑存储结构即为数字内容对象存档信息包结构。

### 6.2 实例文件存储

#### 6.2.1 文件系统存储

完整的 DCOO RDF 实例文件可单独存储于文件系统中。如果文件中涉及媒体对象相关数据文件的具体存储位置时,该存储位置可以是当前 DCOO RDF 实例文件存储位置的相对路径,也可以是其他路径。

**示例:**当前 DCOO RDF 实例文件存储为“/Base\_Path/dcoo\_rdf\_instance.xml”(“Base\_Path”为存储 DCOO RDF 实例文件的基础存储路径)。该实例文件中涉及的媒体对象相关数据文件可存储于同一路径下,其 RDF/XML 示例片段如下所示:

```

<rdf:Description rdf:about="&.myNamespace;example/mediaObj_1">
  <rdf:type rdf:resource="http://dcoo_authority_domain/standards/dcoo#E14_MediaObject"/>
  <dcoo:p45_contentUrl>dataFile/other/mediaObj_1.pdf</dcoo:p45_contentUrl>
  .....
</rdf:Description>

```

注 1：“&.myNamespace;”表示某应用单位的命名空间取值。

注 2：“dataFile”是位于“Base\_Path”下的文件子目录。

## 6.2.2 关系数据库存储

6.2.2.1 关系数据库以二维表结构对实例数据进行组织和存储。

6.2.2.2 将 DCOO RDF 实例文件存储于关系数据库中时,每个实例作为独立的文本单元,存储于数据库表中。表结构见表 1 所示。

表 1 RDF 实例存储表

序号	字段名	说明	类型	非空	备注
1	Id	当前实例的本地标识符	字符串型	是	主键
2	IRI	当前实例的 IRI	字符串型	是	唯一键
3	Description	某一完整的 rdf:Description 节点 XML 片段内容	长文本型	是	
注: IRI 字段内容的示例形如:“&.myNamespace;example/book_1”(其中“&.myNamespace;”表示应用单位的命名空间取值);Description 元素内容的 RDF/XML 示例片段如下: <pre> &lt;rdf:Description rdf:about="&amp;.myNamespace;example/book_1"&gt;   &lt;rdf:type rdf:resource="http://dcoo_authority_domain/standards/dcoo#E19_TextEntity"/&gt;   &lt;dcoo:p30_name&gt;图书的名称&lt;/dcoo:p30_name&gt;   ..... &lt;/rdf:Description&gt; </pre>					

6.2.2.3 实例关系表用于存储 DCOO RDF 实例文件中各个实例之间的关系。表结构见表 2。

表 2 实例关系表

序号	字段名	说明	类型	非空	备注
1	Id	当前关系的标识符	数值型或字符串型	是	主键
2	subjectId	当前主体实例对象的 IRI	字符串型	是	取值形如: &.myNamespace;example/book_1
3	predicate	当前属性的 IRI	字符串型	是	取值形如: &.dcoo;p5_associatedMedia
4	objectId	当前客体实例对象的 IRI	字符串型	是	取值形如: &.myNamespace;example/mediaObj_1
注 1: 该表的存储内容与交换信息包中 dcoomanifest.xml 的 <objectList> 节点下的内容一致(见 5.2.2.4)。 注 2: 该表采用“&.dcoo;”简化表示 DCOO 命名空间取值“http://dcoo_authority_domain/standards/dcoo#”。					

6.2.3 图数据库存储

6.2.3.1 DCOO 对象模型具有图特性,当数字内容对象及其相关资源的实例文件数据规模较大时,利用结构化数据模型的传统管理方式难以满足低数据冗余与高查询性能的要求。以图方式管理 RDF 实例数据不仅可避免 RDF 逻辑数据模型与物理模型之间的转换,而且可利用成熟的图算法优化 RDF 数据查询。

6.2.3.2 图数据库基于有向图,其理论基础是图论。DCOO RDF 实例文件中的主体(及客体)实例对象视为节点,主体与客体之间的关系(即属性)视为边,可将数字内容对象及其相关资源的 RDF 实例数据存储到图数据库中。

6.2.3.3 常见的图数据库存储系统包括:Neo4j、OrientDB、HyperGraphDB、InfiniteGraph、InfoGrid 等。

6.3 数据文件存储

与 DCOO RDF 实例文件相关的数据文件存储于 dataFile 文件目录中,按照不同的媒体类型归类,存储结构见表 3。

表 3 数据文件逻辑存储结构

数据文件根目录	媒体类型文件目录	分组文件目录	说明
[dataFile]			数据文件根目录
	[audio]		音频数据文件根目录
		[分组文件目录]	
	[video]		视频数据文件根目录
		[分组文件目录]	
	[image]		图像数据文件根目录
		[分组文件目录]	
	[other]		其他类型数据文件根目录
		[分组文件目录]	
注 1: 数据文件或直接存于“媒体类型文件目录”下,或分为多组后存于“分组文件目录”下。 注 2: 当某一类型数据文件数量较多时,在对应媒体类型文件目录下,可采用自定义“分组文件目录”的方式进行组织,本部分对“分组文件目录”的名称、数量和层级等未做要求。			

6.4 证实方法

数字内容对象存储的证实方法见 GB/T 38371.3—2020 的 8.1 和 8.2。

7 数字内容对象交换

7.1 交换方式

7.1.1 为了支持数字内容对象的交换与复用,数字内容对象应用系统之间可通过 FTP/SFTP、接口等方式进行数字内容对象的交换。

7.1.2 对于 FTP/SFTP 方式,应用系统之间可直接传送“交换信息包”文件。

7.1.3 对于接口交换方式,资源存储系统可提供以下 8 个方面的接口,以供第三方(外部系统)从该系

统中获取数字内容对象信息(包括元数据、数据文件等)。

- a) 资源标识符获取接口(listIdentifiers):获取符合查询条件的资源标识符列表;
- b) 指定资源 IRI 获取接口(listIris):获取符合查询条件的资源容器级封装包含的所有相关资源的 IRI 列表;
- c) 元数据获取接口(getMetadata):获取指定 IRI 的资源的元数据,支持多种元数据格式(如 MARC、DC 等格式);
- d) 数据文件 URL 获取接口(listContentURL):获取指定资源所关联的数据文件的 URL 列表;
- e) 数据流获取接口(getDataStream):获取数据文件对应的数字内容的字节流;
- f) 交换信息包获取接口(getEIP):获取指定资源 IRI 的交换信息包;
- g) 交换信息包上传接口(uploadEIP):将交换信息包上传至资源存储系统;
- h) 资源检索接口(resRetrieve):在资源存储系统中检索资源信息。

## 7.2 资源标识符获取接口

### 7.2.1 资源标识符获取接口定义

资源标识符获取接口(listIdentifiers)定义见表 4。

表 4 资源标识符获取接口定义

项名	说明
名称	listIdentifiers
描述	根据参数的不同组合,返回符合查询条件的资源标识符的列表
句法	<p><code>http://hostname:port/listIdentifiers? type=IDType&amp;.from=startTime&amp;.until=endTime&amp;.resumptionToken=token</code></p> <p>注 1: hostname 为资源存储系统所在服务的主机名, port 为当前接口的服务端口。</p> <p>注 2: 在实际应用中,参数值应进行 URL 编码;本部分中,为了方便阅读,后续示例中未做参数值的转换</p>
参数	<p>type:必备,资源标识符类型,取值可为 isbn、issn、iri、doi、isli 等</p> <p>注:通过对 DCOO 中的“标识符”属性进行扩展,可支持多种类型的数字内容对象标识符,例如 isbn、issn、iri、doi、isli 等</p>
	<p>from:可选,资源最近更新时间的下限时间值;该参数为空表示不限定该下限时间值</p> <p>注:时间值格式形如“YYYY-MM-DD”或“YYYY-MM-DDThh:mm:ssZ”,符合 GB/T 7408 UTC 格式</p>
	<p>until:可选,资源最近更新时间的上限时间值;该参数为空表示不限定该上限时间值</p> <p>注:时间值格式形如“YYYY-MM-DD”或“YYYY-MM-DDThh:mm:ssZ”,符合 GB/T 7408 UTC 格式</p>
参数	<p>resumptionToken:当查询条件所命中的标识符较多需要分页(多次请求)传输时采用 resumptionToken 作为续传标记;resumptionToken 与 type、from 和 until 互斥,作为独立参数拼接在请求 URL 中,形如: <code>http://hostname:port/listIdentifiers? resumptionToken=token</code></p> <p>注: resumptionToken 的具体实现方式本部分不做规定</p>
证实方法	<p>当请求参数正确、必备参数齐备,并且资源标识符类型可识别时,响应符合查询条件的资源标识符列表(分页响应时,同时返回请求下一页时所需的 resumptionToken 值);</p> <p>当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为: badArgument(请求包含错误参数或缺失必备参数);</p> <p>当请求参数中的资源标识符类型无法识别或不存在时,响应错误代码及描述信息为: typeDoesNotExist(类型无法识别或不存在)</p>

## 7.2.2 资源标识符获取接口调用

### 7.2.2.1 正确请求响应

#### 7.2.2.1.1 资源标识符获取接口正确请求响应的示例如下：

请求：返回资源更新时间在“2017-12-01”至“2017-12-31”期间的图书 ISBN 列表

http://hostname:port/listIdentifiers? type=isbn&.from=2017-12-01&.until=2017-12-31

响应：

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/">
  <oai-pmh:responseDate>2018-07-27T08:55:46Z</oai-pmh:responseDate>
  <oai-pmh:request verb="listIdentifiers">
    http://hostname:port/listIdentifiers? type=isbn&.from=2017-12-01&.until=2017-12-31
  </oai-pmh:request>
  <listIdentifiers>
    <identifier value="7-121-01831-4" name="图书名称 1" iri="....."/>
    <identifier value="5-111-03451-1" name="图书名称 2" iri="....."/>
    <identifier value="3-345-04331-7" name="图书名称 3" iri="....."/>
  </listIdentifiers>
</dcoo>
```

#### 7.2.2.1.2 接口响应的数据结构参考 OAIS 标准(ISO 14721:2012)。

#### 7.2.2.1.3 responseDate 表示响应时间,日期符合 GB/T 7408—2005 UTC 格式。

#### 7.2.2.1.4 request 元素中属性 verb 给出了接口名称,内容区给出了请求串;identifier 元素中 value 给出了对应标识符的值,name 为资源的名称,iri 为资源的 DCOO 实例 IRI。

### 7.2.2.2 错误请求响应

#### 7.2.2.2.1 资源标识符获取接口错误请求响应的示例如下：

请求：http://hostname:port/listIdentifiers? type=abc&.value=7-120-01234-5

响应：

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/">
  <oai-pmh:responseDate>2018-07-27T08:55:47Z</oai-pmh:responseDate>
  <oai-pmh:request verb="listIdentifiers">
    http://hostname:port/listIdentifiers? type=abc&.value=7-120-01234-5
  </oai-pmh:request>
  <error code="badArgument">请求包含错误参数或缺失必备参数</error>
</dcoo>
```

#### 7.2.2.2.2 error 元素中 code 给出了错误类型代码,内容区为相应的错误描述。

### 7.3 指定资源 IRI 获取接口

#### 7.3.1 指定资源 IRI 获取接口定义

指定资源 IRI 获取接口(listIris)定义见表 5。

表 5 指定资源 IRI 获取接口定义

项名	说明
名称	listIris
描述	返回符合查询条件的资源容器级封装包含的所有相关资源的 IRI 列表
句法	http://hostname:port/listIris? type=IDType&value=IDValue
参数	type:必备,资源标识符类型,取值可为 isbn、issn、iri、doi、isli 等 value:必备,标识符的值
证实方法	当请求参数正确、必备参数齐备,并且资源标识符类型可识别时,响应命中的资源容器级封装包含的所有相关资源的 IRI 列表; 当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为:badArgument(请求包含错误参数或缺失必备参数); 当请求参数中的资源标识符类型无法识别或不存时,响应错误代码及描述信息为:typeDoesNotExist(类型无法识别或不存)

#### 7.3.2 指定资源 IRI 获取接口调用

##### 7.3.2.1 指定资源 IRI 获取接口请求响应的示例如下:

请求:返回 ISBN 值为 7-121-01831-4 的图书容器级封装包含的下级资源的 IRI 列表

http://hostname:port/listIris? type=isbn&value=7-121-01831-4

响应:

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/">
  <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
  <oai-pmh:request verb="listIris">
    http://hostname:port/listIris? type=isbn&value=7-121-01831-4
  </oai-pmh:request>
  <listIris>
    <resIri type="dcoo:E19_TextEntity" value="&.myNamespace;example/book_1"/>
    <resIri type="dcoo:E20_Metadata" value="&.myNamespace;example/metadata_1"/>
    <resIri type="dcoo:E14_MediaObject" value="&.myNamespace;example/mediaObj_1"/>
    .....
  </listIris>
</dcoo>
```

##### 7.3.2.2 resIri 元素中 value 给出了对应 DCOO 实例的 IRI 值。

##### 7.3.2.3 在第 7 章的各接口请求和响应示例中,统一采用“&.myNamespace;”表示应用单位命名空间

取值。

7.3.2.4 type 取值为对应资源的对象类型值,可为 DCOO 对象类型、应用单位扩展类型或其他类型。

7.4 元数据获取接口

7.4.1 元数据获取接口定义

元数据获取接口(getMetadata)定义见表 6。

表 6 元数据获取接口定义

项名	说明
名称	getMetadata
描述	通过资源 IRI 获取该资源的元数据
句法	http://hostname:port/getMetadata? iri=resourceIRI&format=metaFormat
参数	iri:必备,资源对应的 DCOO 实例 IRI
	format:可选,表示元数据格式,可取值为 dc,cnixon,marc21 等 注: format 为空时,返回当前资源的所有格式的元数据文件的 ZIP 格式压缩包
证实方法	当请求参数正确,并且必备参数齐备时,响应命中资源的元数据文件; 当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为:badArgument(请求包含错误参数或缺失必备参数)

7.4.2 元数据获取接口调用

元数据获取接口请求响应的示例如下:

请求:http://hostname:port/getMetadata? iri=&.myNamespace;example/book\_1

响应说明:通过指定的 IRI,获取 0 至多个元数据实例(dcoo:E20\_Metadata),再通过“关联媒体”属性查找到对应媒体对象实例(dcoo:E14\_MediaObject)的“内容链接”属性,获取到相应的元数据文件(可能存在多种媒体格式类型),并将元数据文件(或元数据文件的 ZIP 包)响应给请求方。

7.5 数据文件 URL 获取接口

7.5.1 数据文件 URL 获取接口定义

数据文件 URL 获取接口(listContentURL)定义见表 7。

表 7 数据文件 URL 获取接口定义

项名	说明
名称	listContentURL
描述	根据参数中的 IRI,返回符合查询条件的资源封装中的不同类型的数据文件列表
句法	http://hostname:port/listContentURL? iri=resourceIRI
参数	iri:必备,资源对应的 DCOO 实例 IRI
证实方法	当请求参数正确,并且必备参数齐备时,响应符合查询条件的数据文件 URL 列表; 当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为:badArgument(请求包含错误参数或缺失必备参数)

7.5.2 数据文件 URL 获取接口调用

数据文件 URL 获取接口请求响应的示例如下：  
请求: http://hostname:port/listContentURL? iri=&myNamespace;example/book\_1  
响应:

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/">
  <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
  <oai-pmh:request verb="listContentURL">
    http://hostname:port/listContentURL? iri=&myNamespace;example/book_1
  </oai-pmh:request>
  <listContentURL>
    <url value=
      " http://hostname: port/book. 2434553/CONTAINER/dataFile/image/2434553-
Cover.jpg"/>
    <url value=
      "http://hostname:port/book.2434553/CONTAINER/dataFile/other/2434553-PDF.
pdf"/>
    .....
  </listContentURL>
</dcoo>
```

7.6 数据流获取接口

7.6.1 数据流获取接口定义

数据流获取接口(getDataStream)定义见表 8。

表 8 数据流获取接口定义

项名	说明
名称	getDataStream
描述	通过数据文件的内容链接获取其对应的数字内容的字节流
句法	http://hostname:port/getDataStream? url=contentURL
参数	url: 必备; 数据文件的内容链接
证实方法	当请求参数正确, 并且必备参数齐备时, 响应命中数据文件的二进制内容; 当请求包含错误参数或缺失必备参数时, 响应错误代码及描述信息为: badArgument(请求包含错误参数或缺失必备参数)

7.6.2 数据流获取接口调用

7.6.2.1 数据流获取接口请求响应的示例如下:

请求: `http://hostname:port/getDataStream? url=http://hostname:port/book.2434553/CONTAINER/dataFile/image/2434553-Cover.jpg`

响应:

该数据文件对应的二进制字节内容。

7.6.2.2 本例的 url 参数关联了一个 jpg 文件,因此该接口返回该 jpg 文件的二进制内容。应用中,需对请求中的 url 值进行 URL 编码。

7.7 交换信息包获取接口

7.7.1 交换信息包获取接口定义

交换信息包获取接口(getEIP)定义见表 9。

表 9 交换信息包获取接口定义

项名	说明
名称	getEIP
描述	通过资源唯一标识符 IRI 获取该资源对应的 EIP,包括 DCOO RDF 实例文件及数据文件。即若该资源没有对应的数据文件,则只返回 DCOO RDF 实例文件,否则该数字内容对象及下级资源的实例文件和对应数据文件一起打包,打包文件为 ZIP 格式(见 5.2)
句法	<code>http://hostname:port/getEIP? iri=resourceIRI</code>
参数	iri:必备,资源对应的 DCOO 实例 IRI
证实方法	当请求参数正确,并且必备参数齐备时,响应命中 EIP(ZIP 文件)的二进制内容; 当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为:badArgument(请求包含错误参数或缺失必备参数)

7.7.2 交换信息包获取接口调用

交换信息包获取接口请求响应的示例如下:

请求:`http://hostname:port/getEIP? iri=&.myNamespace;example/book_1`

响应:该 EIP(ZIP 文件)的二进制内容。

7.8 交换信息包上传接口

7.8.1 交换信息包上传接口定义

交换信息包上传接口(uploadEIP)定义见表 10。

表 10 交换信息包上传接口定义

项名	说明
名称	uploadEIP
描述	将 EIP 上传至服务器或数据库
句法	<code>http://hostname:port/uploadEIP</code>
请求方法	post



表 10（续）

项名	说明
发送内容	EIP 的二进制内容,以 multipart/form-data 方式编码,对应的 name 为 eip_filename;eip_filename 为本地真实文件名
证实方法	当请求参数正确、必备参数齐备,并且文件成功上传时,响应文件上传后的状态为“true”; 当请求包含错误参数或缺失必备参数时,响应错误代码及描述信息为:badArgument(请求包含错误参数或缺失必备参数); 当上传文件格式不符合 EIP 结构要求时,响应错误代码及描述信息为:badFile(非法文件)

7.8.2 交换信息包上传接口调用

交换信息包上传接口请求响应的示例如下:

请求:以如下的 HTML 表单为例(HTTP 请求采用 POST 方法)

```
<form method="POST" action="http://hostname:port/uploadEIP" enctype="multipart/form-data">
  <input type="file" name="eip_filename"/>
  <input type="submit" value="上传 EIP"/>
</form>
```

响应:

```
<? xml version="1.0" encoding="UTF-8"? >
  <dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/">
    <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
    <oai-pmh:request verb="uploadEIP">http://hostname:port/uploadEIP</oai-pmh:request>
    <upload>
      <uploaded status="true"/>
    </upload>
  </dcoo>
```

7.9 资源检索接口

7.9.1 资源检索接口定义

7.9.1.1 资源检索接口遵循 SRU/SRW 协议,实现从资源存储系统中获取元数据信息。资源检索接口(resRetrieve)定义见表 11。

表 11 资源检索接口定义

项名	说明
名称	resRetrieve
描述	根据参数不同组合,返回符合查询条件的结果数据

表 11（续）

项名	说明
句法	http://hostname: port/resRetrieve? version =... &query =... &opeation =... &startRecord =... &maximumRecords=...&recordPacking=...&sortKeys=...
参数	version:必备。表示客户端支持的版本号,服务器端可按照客户端的要求发送消息,如果服务器端的版本号高于客户端的版本号,那么服务端将降低版本号,适应客户端的要求
	query:必备。遵守 CQL 语法的检索表达式
	opeation:必备。参数值通常为“searchRetrieve”
	startRecord:可选。说明客户端要求响应返回的第一条记录的位置。服务器端有可能检索到若干条记录,用户如果只想取其中的一部分,则可使用该参数,取值应大于 0
	maximumRecords:可选。说明客户端要求响应返回的最大记录数,取值应大于或等于 0。但应小于服务器端默认的最大记录数
	recordPacking:可选。指定命中结果返回的格式,取值可为“String”或“XML”
证实方法	响应格式见表 12。异常响应时,错误信息由诊断信息列表给出,每组诊断信息的格式见表 13

表 12 资源检索接口响应格式

参数名称	类型	必备性	描述
version	xsd:string	必备	说明服务器使用的 SRW 协议的版本
numberOfRecords	xsd:integer	必备	说明命中记录的数量;查询失败时值为 0
resultSetId	xsd:string	可选	结果集的唯一标识。每个结果集都有唯一的 resultSetId,不论是否有新的查询提交,也不论该结果集是否仍然存在。若修改已有结果集(如对其重新进行排序),则服务器为新的结果集提供新的 resultSetId。该参数默认值为空
resultSetIdleTime	xsd:integer	可选	服务器端认为结果集能够保留的时间。resultSetIdleTime 只能取正整数,且应大于用户实际引用结果集的时间。若服务器不让结果集被引用,则应在响应中忽略结果集标识。该参数默认值为空
records	记录列表	可选	返回的记录结果列表。records 包含一个或多个 record 子元素。每个 record 元素包含一个命中记录或者是一个代表诊断,解释为什么命中记录未被传递
nextRecordPosition	xsd:integer	可选	指下一次返回记录的位置,适应于翻页的情况
diagnostics	诊断信息列表	可选	诊断信息列表。diagnostics 包含一个或多个 diagnostic 子元素。每个 diagnostic 表示一种错误诊断消息

7.9.1.2 诊断信息在 SRU 的响应信息中,可由<diagnostics>元素指明,也可作为<records>元素的子元素。诊断信息可为“严重”或“非严重”。“非严重”的诊断信息可进一步分为“代理”或“非代理”。因此,诊断信息可有如下组合:

- a) 代理的、非严重的诊断信息(在元素<records>中表示);
- b) 非代理的、非严重的诊断信息(在元素<diagnostics>中表示);
- c) 非代理的、严重的诊断信息(在元素<diagnostics>中表示)。

7.9.1.3 <diagnostic>元素包含的子元素见表 13。

表 13 诊断信息元素包含的子元素

元素	类型	必备性	说明
<iri>	xsd:string	必备	标识特定诊断信息的 IRI。可从 SRU 协议的诊断信息列表中选择,也可自定义 IRI 列表
<details>	xsd:string	可选	额外的有效信息,通常表示为标准格式
<message>	xsd:string	可选	对终端用户可读的说明信息

## 7.9.2 资源检索接口调用

资源检索接口请求响应的示例如下:

请求:返回题为“core JavaServer Faces”的图书元数据

http://hostname:port/resRetrieve? version = 1.0&query = dc.title = " core JavaServer Faces "  
&operation = searchRetrieve&.startRecord = 1&.maximumRecords = 1&.recordPacking = xml

响应:

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/"
  xmlns:ore="http://www.openarchives.org/ore/terms/"
  xmlns:srw="http://www.loc.gov/zing/srw/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/">
  <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
  <oai-pmh:request verb="resRetrieve">
    http://hostname:port/resRetrieve? version = 1.0&query = dc.title = " core JavaServer
Faces"
    &operation = searchRetrieve&.startRecord = 1&.maximumRecords = 1&.recordPacking
= xml
  </oai-pmh:request>
  <srw:resRetrieve>
    <srw:version>1.0</ srw:version>
    <srw:numberOfRecords>1</srw:numberOfRecords>
    <srw:records>
      <srw:recordPacking>xml</ srw:recordPacking>
      <srw:recordData>
        <rdf:Description
          rdf:about="&.myNamespace;bc-123456/book/book-123456">
          <dcoo:p22_identifier>book-123456</dcoo:p22_identifier>
          <rdf:type rdf:resource=" http://dcoo_authority_domain/standards/dcoo # E19_
TextEntity"/>
```

```

        <dc:title>core JavaServer Faces</dc:title>
        <dcterms:publisher
            rdf:resource="http://www.pearsoned.co.uk/Imprints/Addison-Wes-
            ley/" />
        <dc:creator>DAVID GEARY</dc:creator>
        <dc:creator>CAY HORSTMANN</dc:creator>
        .....
        <dcterms:language rdf:resource="eng" />
        <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema # da-
        teTime">
            2017-11-11T10:00:00Z</dcterms:created>
        <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema #
        dateTime">
            2017-11-11T10:00:00Z</dcterms:modified>
    </rdf:Description>
</srw:recordData>
</srw:records>
</srw:resRetrieve>
</dcoo>

```

错误响应示例 1: 非代理的、严重的诊断信息

```

<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
    xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/"
    xmlns:srw="http://www.loc.gov/zing/srw/"
    xmlns:diag="http://www.loc.gov/zing/srw/diagnostic/">
    <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
    <oai-pmh:request verb="resRetrieve"> ..... </oai-pmh:request>
    <srw:resRetrieve>
        <diag:diagnostics>
            <diag:diagnostic>
                <diag:iri>info:srw/diagnostic/1/38</diag:iri>
                <diag:details>10</diag:details>
                <diag:message>操作符太多,最大为 10。请尝试简单检索。</diag:message>
            </diag:diagnostic>
        </diag:diagnostics>
    </srw:resRetrieve>
</dcoo>

```

错误响应示例 2:代理的、非严重的诊断信息

```
<? xml version="1.0" encoding="UTF-8"? >
<dcoo xmlns="http://dcoo_authority_domain/standards/dcoo#"
  xmlns:oai-pmh="http://www.openarchives.org/OAI/2.0/"
  xmlns:srw="http://www.loc.gov/zing/srw/"
  xmlns:diag="http://www.loc.gov/zing/srw/diagnostic/">
  <oai-pmh:responseDate>2017-08-08T08:55:46Z</oai-pmh:responseDate>
  <oai-pmh:request verb="resRetrieve"> ..... </oai-pmh:request>
  <srw:resRetrieve>
    <srw:version>1.0</srw:version>
    <srw:numberOfRecords>1</srw:numberOfRecords>
    <srw:records>
      <srw:recordPacking>xml</srw:recordPacking>
      <srw:record>
        <srw:recordSchema>info:srw/schema/1/diagnostics-v1.1
        </srw:recordSchema>
        <srw:recordData>
          <diag:diagnostic>
            <diag:iri>info:srw/diagnostic/1/65</diag:iri>
            <diag:message>记录不存在。</diag:message>
          </diag:diagnostic>
        </srw:recordData>
      </srw:record>
    </srw:records>
  </srw:resRetrieve>
</dcoo>
```

## 参 考 文 献

- [1] GB/T 18793—2002 信息技术 可扩展置标语言(XML)1.0
- [2] CY/T 102.1—2014 数字内容对象存储、复用与交换规范 第1部分:对象模型
- [3] CY/T 102.2—2014 数字内容对象存储、复用与交换规范 第2部分:对象封装、存储与交换
- [4] CY/T 102.3—2014 数字内容对象存储、复用与交换规范 第3部分:对象一致性检查方法
- [5] GC/FH 12—2017 复合文档基础结构
- [6] GC/FH 13—2017 篇章复合文档结构
- [7] GC/FH 14—2017 条目复合文档结构
- [8] GC/FH 16—2017 论文复合文档结构
- [9] Friend of a Friend. [FOAF]. <http://xmlns.com/foaf/spec/>
- [10] Dublin Core Metadata Element Set. <http://dublincore.org/documents/dces/>
- [11] DCMI Metadata Terms. <http://dublincore.org/documents/dcmi-terms/>
- [12] EPUB 3.0. <http://idpf.org/epub/>
- [13] Reference Model for an Open Archival Information System (OAIS). <http://public.ccsds.org/publications/archive/650x0m2.pdf>
- [14] Resource Description Framework.[RDF]<https://www.w3.org/RDF/>
- [15] RFC 3986 Uniform Resource Identifier (URI): Generic Syntax. <https://tools.ietf.org/html/rfc3986>
- [16] RFC 3987 Internationalized Resource Identifiers (IRIs). <https://tools.ietf.org/html/rfc3987>
- [17] RFC 1867 Form-based File Upload in HTML. <http://www.ietf.org/rfc/rfc1867.txt>
- [18] RFC 3305 Uniform Resource Identifiers(URIs), URLs, and Uniform Resource Names (URNs): Clarifications and Recommendations. <http://www.ietf.org/rfc/rfc3305.txt>
- [19] RFC 2046 Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types. <https://www.ietf.org/rfc/rfc2046.txt>
- [20] Open Document Format for Office Applications (OpenDocument) Version 1.2.[ODF]. <https://www.oasis-open.org/standards#opendocumentv1.2>
- [21] SRU. Search/Retrieve via URL. <https://www.loc.gov/standards/sru>
- [22] SRW. Search/Retrieve Web Service. <https://www.loc.gov/standards/sru/companionSpecs/srw.html>
- [23] The Contextual Query Language(CQL). <http://www.loc.gov/standards/sru/cql/index.html>