
国家数字图书馆工程

标准规范项目研制成果

GC-HD090188

项目名称：国家数字图书馆工程长期保存规范

成果名称：国家图书馆长期保存信息封装标准规范应用指南

成果类型：应用指南

成果编号：GC-HD090188-05

成果版本：征求意见稿

提交时间：2012年3月

研制机构：清华大学图书馆

撰 写 人：童庆钧、郑小惠、程变爱、姜爱蓉

目 次

前言	II
引言	III
1 信息包封装概述.....	1
1.1 开放存档信息系统 Open Archival Information System; OAIS	1
1.2 信息包 information package.....	1
1.3 封装 packaging	2
1.4 元数据编码与传输标准 Metadata Encoding and Transmission Standard; METS	3
2 在 METS 节引用 PREMIS.....	4
2.1 概述.....	4
2.2 METS 元素与 PREMIS 语义单元之间的对应关系.....	4
2.3 使用节的数目	5
2.4 METS 和 PREMIS 的冗余处理.....	5
2.5 METS 结构图与 PREMIS structural relationship 元素	5
2.6 METS ID/IDREF 与 PREMIS identifier 元素	5
3 基于 METS 的 AIP 封装.....	5
3.1 METS 根元素	6
3.2 METS 头元素	7
3.3 描述元数据节<dmdSec>	8
3.4 描述元数据节或管理元数据节内的<mdWrap>.....	9
3.5 管理元数据节	11
3.6 文件节<fileSec>	22
3.7 文件<file>.....	24
3.8 结构图<structMap> (参见附录 A)	26
3.9 <div>	26
3.10 结构链接<structLink>.....	29
3.11 行为节<behaviorSec>.....	29
4 应用场景.....	29
4.1 一个 SIP 对应一个 AIP.....	30
4.2 多个 SIP 对应一个 AIP.....	30
4.3 一个 SIP 对应多个 AIP.....	38
4.4 多个 SIP 对应多个 AIP.....	41
4.5 从 AIP 到 DIP.....	42
附录 A (资料性附录) 电子期刊封装实例 (期刊级别)	44
参考文献.....	50

前　　言

本应用指南是国家数字图书馆工程标准规范项目研制成果之一。

本应用指南由国家图书馆提出，委托清华大学图书馆进行研制。

本应用指南由清华大学图书馆起草，主要起草人：童庆钧、郑小惠、程变爱、姜爱蓉。

引　　言

由国家图书馆主持建设的国家数字图书馆工程自 2005 年开始建设,截至 2011 年 6 月底,国家图书馆数字资源总量已达 545TB, 数字资源发布总量 396.7TB, 包括自建数字资源、外购数字资源、互联网采集的数字资源以及全国文化信息共享工程的数字资源等。

国家数字图书馆工程长期保存规范要求根据 OAIS 参考模型, 制定国家图书馆长期保存中使用的 SIP、AIP、DIP 三种信息包的封装规范及封包细则, 能够满足国家图书馆长期保存各种数字资源的需求。此规范将用于国家长期保存系统, 作为系统中数字资源封装和传递的规范。

本应用指南针对长期保存信息包封装规范而制定, 主要目的有两方面: 一方面深入解释 METS 标准的使用规则, 另一方面通过实例说明对相关类型资源进行封装时, 涉及的一般流程。

METS 除了可以用自身的结构对数字对象进行表述之外, 还可以使用扩展机制从外部引用相关标准, 比如描述元数据可以引用 DC、MODS, 或国家图书馆专门描述元数据规范(代码暂定为 NLCDMD), 图像技术元数据可以引用 MIX, 技术、权利、来源、数字起源等元数据可以引用 PREMIS。METS 结构的灵活性, 使其在实际应用中的复杂度增加。因此, 本应用指南不仅对 METS 标准自身结构的使用规则进行了说明, 同时对外部引用的使用规则也进行了说明, 特别着重于如何封装 PREMIS 信息。本指南还针对国家图书馆主要资源类型(如电子期刊、电子图书、图像、音视频等), 提供了若干实例。

国家图书馆长期保存信息包封装标准规范应用指南

1 信息包封装概述

1.1 开放存档信息系统 Open Archival Information System; OAIS

开放存档信息系统（Open Archival Information System, OAIS）参考模型于 2001 年 7 月正式发布，2003 年作为 ISO 标准（ISO14721: 2003）发布。

OAIS 是一个由人和系统组成的，承担信息长期保存责任并将信息提供给目标团体（Designated Community）的存档系统。自 OAIS 发布以来，很多数字资源长期保存的研究和开发的项目都以 OAIS 作为一个优先遵守的标准规范。

OAIS 参考模型提供了一个完整的档案信息保存功能，它包括信息摄入（接收）、存档、数据管理、存取和分发。OAIS 还论述了数字信息从一种媒体或格式到另一媒体或格式的迁移、信息表示的数据模型、信息保存中软件的作用、存档系统间的信息交换等。OAIS 确立了各项功能的内部联系和与外部的接口，以及一系列建立在这些接口之上的高层服务。OAIS 参考模型成为数字档案系统普遍遵从的标准规范，来自图书馆、档案馆、科学数据、文化记忆等领域的相关项目和系统都纷纷以其为准，进行系统的规划和设计。

1.2 信息包 information package

OAIS 信息模型围绕信息包的概念而建立。信息包由数字对象及其相关联的元数据组成，数字对象是长期保存的重点，相关的元数据支持其长期保存及其访问，数据和元数据共同封装成一个逻辑包。信息包有三种：提交信息包（SIP）、存档信息包（AIP）和分发信息包（DIP）。

1.2.1 提交信息包 Submission Information Package; SIP

提交信息包（SIP）是生产者提交给 OAIS 的信息包，其形式和详细内容一般由生产者和 OAIS 双方谈判而确定。大多数 SIP 有一些内容信息和保存描述信息（PDI），但可能需要多个 SIP 提供完整的内容信息和 PDI。内容信息和 PDI 都包含相关的表现信息，如果涉及的多个 SIP 具有相同的表现信息，这些信息很可能只提供一次。或，因为一些 PDI 适用于多个 SIP，这个 PDI 可以通过一个不包含内容信息的 SIP 提供。封装信息总是以某种形式存在。

1.2.2 存档信息包 Archival Information Package; AIP

存档信息包（AIP）由内容信息及其相关的保存描述信息（PDI）构成，保存于 OAIS 系统中。

SIP 进入 OAIS 后，其形式和内容可能会发生变化。OAIS 不必以 SIP 的格式保留信息；以提交的形式保存可能并不合适。比如存储提交图像的计算机介质过时了，图像必须复制到另一种更新的介质上。生产者不能提供的唯一标识符等信息，也必须在摄入过程中输入。

1.2.3 分发信息包 Dissemination Information Package; DIP

分发信息包 (DIP) 来源于一个或多个 AIP，由 OAIS 发送给消费者。DIP 根据消费者的请求发送，可能是 AIP 的所有或部分。DIP 和 AIP 在形式和内容上可能都有不同。两者的区别包括内容格式（如在分发前，一个图像文件可能会从 TIFF 格式转换为 JPEG 格式）、内容总量（一个 DIP 可能相当于一个 AIP、多个 AIP，甚至是一个 AIP 的一部分）、随内容提供的元数据数量（很可能 DIP 不包含一个存档数字对象的所有元数据集，因为消费者对其中很多元数据没什么兴趣）。

SIP、AIP 和 DIP 的相互关系如图 1 所示。

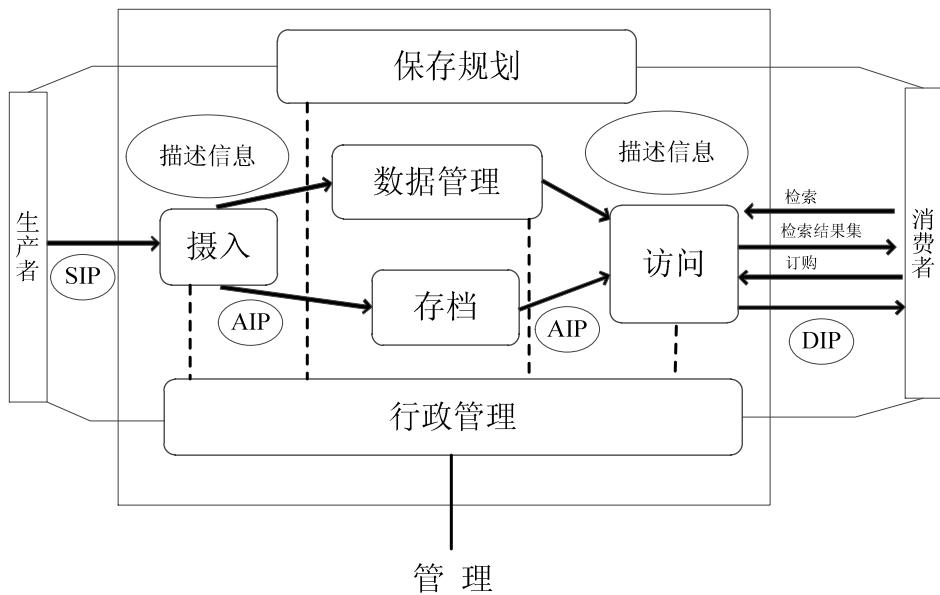


图 1 OAIS 参考模型框架

1.3 封装 packaging

内容信息和保存描述信息代表了存档数字内容、对其呈现和理解所需的元数据，以及支持其长期保存所需的元数据。封装信息用于将这些信息绑定在一个逻辑包中。封装信息将 AIP 的各部分相关联，使其可以在存档系统中作为一个逻辑单位被识别和定位。

描述信息可以使目标群体通过 OAIS 的检索工具，对内容信息进行检索。

内容信息（内容数据对象和表现信息）、保存描述信息（引用、上下文、出处、固定信息）、封装信息和描述信息共同组成 AIP，依次代表保存数字对象及其相关的元数据。

由于 AIP 是长期保存的重点, 本应用指南主要围绕 AIP 展开。AIP 中包括内容信息、保存描述信息、封装信息和描述信息, 如图 2 所示。

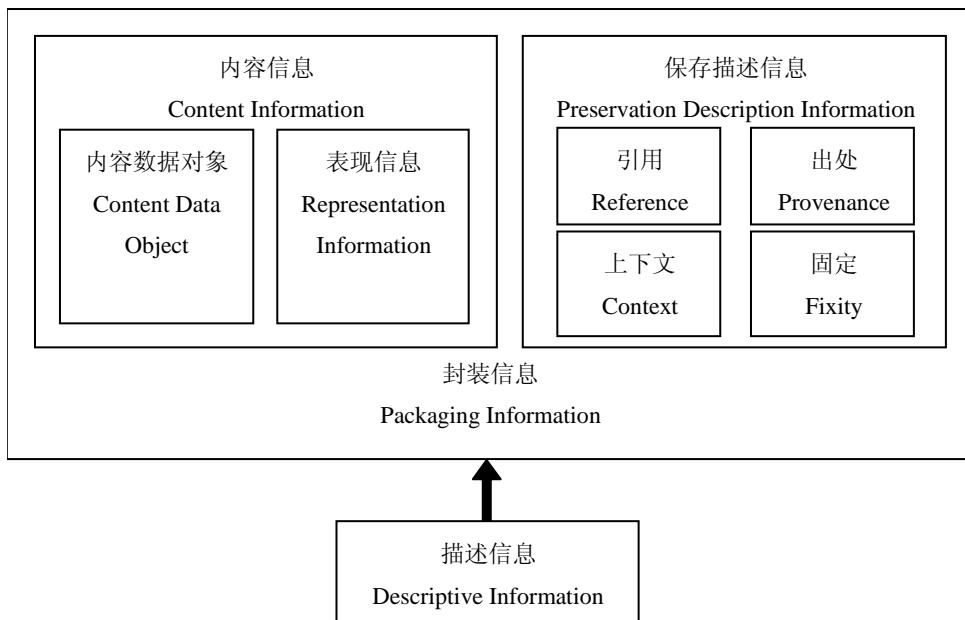


图 2 AIP 的内容构成

1.4 元数据编码与传输标准 Metadata Encoding and Transmission Standard; METS

MOA2 (The Making of America II) 项目提出了一种针对文本和图像为主的描述元数据、管理元数据和结构元数据的编码格式。由美国数字图书馆联盟 (Digital Library Federation) 倡导的 METS，则致力于在 MOA2 基础上提供一种元数据编码的 XML 文档格式，它不但适于仓储内部数字对象的管理，而且适于这些对象在仓储之间（或仓储与用户之间）的交换。从 METS 文档的用途可见，它可以担当 OAIS 模型中 SIP、AIP 或 DIP 的角色。

METS 是一个专门设计的 XML schema 框架，可以用来封装并存储数字对象及其相关元数据。METS 的实施较为灵活，任何可处理 XML 的系统都可以用来创建、存储、传递基于 METS 的元数据。METS 应用纲要 (profile) 可用来记录一个项目的 METS 规范。XML 的开放性也使 METS 记录 (record) 不依赖于任何软件系统。

METS 允许两种方式实现元数据和对象数据的封装，既可以存储在 METS 文件内（内部引用，通过<mdWrap>元素实现），也可以存储在 METS 文件外（外部引用，通过<mdRef>元素实现），并通过 METS 进行引用。第一种方式可将对象数据及其所有元数据进行封装，但会产生大文件；第二种方式的 METS 文件更易于管理，但如果被引用数字对象的存储路径改动，会带来麻烦，因而需要完善的事件注册和解析机制。不同的实施方式各有优缺点，因此在执行和应用 METS 之前应当确定实施的具体策略。

METS 文档由 METS 头、描述元数据节、管理元数据节、文件节、结构图、结构链接和行为节等七部分组成。METS 的具体规范参见国家图书馆长期保存信息包封装标准规范部分。

下面一段代码即为一个 METS 文档的基本框架。其中只有结构图为必备项，其他都是

可选的。

```
<mets:mets>

  <mets:methHdr />          METS 头
  <mets:dmdSec />          描述元数据
  <mets:amdSec />          管理元数据
  <mets:fileSec />          文件节
  <mets:structMap />        结构图
  <mets:structLink />        结构链接
  <mets:behaviorSec />      行为节

</mets:mets>
```

2 在 METS 节引用 PREMIS

2.1 概述

对于长期保存系统的信息包封装而言，保存元数据（PREMIS）尤为重要。在 METS 各节引用 PREMIS 有几点注意事项。

首先，应明确编写 METS 文件的目的是为了实现存储、服务还是编码传递。这是由 METS 标准本身的特点决定的。METS 是为了实现数字资源长期保存系统中的信息交互而设计的。当 METS 文件仅为 SIP 时，受到的限制可能会少一些，因为系统只需要数据元素就够了。而当 METS 文件充当 DIP 时，受到的限制更多，系统需要更权威的元数据。

其次，在应用之前，必须决策是否将 PREMIS 元数据分散在各个子模块中。如果分散放置，需要决定四种不同类型的 PREMIS 实体（即技术、权利、来源、数字起源）与 METS 如何对应。

最后，如何处理 METS 元素与 PREMIS 语义单元之间的冗余，这在应用的过程中也需要决策。

2.2 METS 元素与 PREMIS 语义单元之间的对应关系

在确定 METS 元素与 PREMIS 语义单元之间的对应关系时，应首先判断是否将 PREMIS 语义单元分散放置。PREMIS 元数据可以全部置于 METS 管理元数据节下，也可以分别放在不同的元数据节下。PREMIS 第一层数据元素应该用于 METS 下列节中：

- premis:object：放在 techMD 或 digiProvMD 下；
- premis:event：放在 digiProvMD 下；
- premis:rights：放在 rightsMD 下；
- premis:agent：放在 digiProvMD 或 rightsMD 下。

如果所有单元一起使用，整个信息包将放在 digiProvMD 下，以<premis>元素作为容器。

2.3 使用节的数目

可以使用一个管理元数据节<amdSec>，其下包含重复的子元素（如 techMD 等）；也可以在每个 METS 的子元素下使用重复的<amdSec>。至于使用哪一种方式，由实施者而定。

如果 premis:agent 和一个事件或权利一起使用，则应置于其自身的 digiProvMD 或 rightsMD 节下，因为这能使冗余度降到最低。来自于不同 schema 的技术元数据可以放在单独的 techMD 节中，特定格式的元数据可以包括在 PREMIS 元数据中，放在 premis:object 的 objectCharacteristicsExtension 下。

2.4 METS 和 PREMIS 的冗余处理

特定格式的技术元数据 schema（如 MIX）包括的数据元素同时也可以在 PREMIS 内，这时便会出现冗余。在 METS 中，特定格式 schema 的数据可以被用在单独的 techMD 节中，使用 MDTYPE 或 OTHERMDTYPE 属性指出其元数据类型。

如果一个元素既可以被定义为 PREMIS 元素，也可以在 METS schema 中定义（通常为属性元素），也会出现冗余。

实施者应该考虑元数据的使用（如展示或保存），以及决定使用和选择冗余记录时考虑主要是 METS（用于展示）还是 PREMIS（用于保存）。其他考虑还有哪种 schema 的表达内容更广泛。在大多数情况下（除了 METS 的结构元数据相比起 PREMIS 有优势以外），PREMIS 元素更为详细（如格式）。当元数据的使用不明确或展示和保存同等重要时，宁可数据冗余，把 PREMIS 和 METS 两者都包括。实施者应该在 profile 中记录在实际应用时对冗余的处理。

2.5 METS 结构图与 PREMIS structural relationship 元素

层级结构关系应根据 METS schema 和原则详细记录为嵌套的<div>元素，因为它揭示的信息比 PREMIS 语义单元提供的更为丰富。如果交换对象的范围仅为保存，实施者还应该在对象 schema 中，使用 PREMIS 关系元素表示结构关系。PREMIS 关系元素应该一直用于关系的衍生类型。

2.6 METS ID/IDREF 与 PREMIS identifier 元素

使用 PREMIS 和 METS 在元素之间建立链接的方式很多。在 PREMIS 中，标识有很多单独的数据元素，如 objectIdentifier、eventIdentifier、agentIdentifier 和 rightsStatementIdentifier。另外还有在 PREMIS 实体之间建立链接的标识。

此外，在 PREMIS schema 中，还提供了 ID/IDREF，在 PREMIS 相关元素之间建立链接，如 RelObjectXmlID、LinkEventXmlID、LinkPermissionStatementXmlID、RelEventXmlID、linkingObjectXmlID、LinkAgentXmlID、GrantAgentXmlID 等。METS 也提供 ID/IDREF，在文件及其相关元数据之间建立链接。

3 基于 METS 的 AIP 封装

3.1 METS 根元素

METS 根元素有 5 个可选属性：ID（对象标识）、OBJID（为 METS 文档分配的唯一标识）、LABEL（文本标注）、TYPE（对象类型），以及 PROFILE（该 METS 文档注册执行的标准）。

ID 是国家图书馆为该 METS 文档分配的唯一标识。OBJID 是国家图书馆为长期保存系统内的数字对象分配的唯一标识。LABEL 是为该数字对象取的标签。PROFILE 列出国家图书馆的 METS 文档遵循哪些纲要（profile）。

表 1 METS 根元素

元素/属性	可重复性/必备性	规则与推荐规范
<mets>	不可重复 必备	必须包含“PROFILE”属性及<metsHdr>元素。
PROFILE	不可重复 必备	在 METS 中可选。该属性值待确定，可暂定为“中国国家图书馆信息包封装规范”。
OBJID	不可重复 必备	在 METS 中可选。必须有分配给 METS 文档的主要标识符。该标识符在仓储中必须是唯一的，但不一定是全局唯一。
ID、LABEL、TYPE	不可重复 可选	无推荐项。

示例：

METS 文档是一个 XML 文档。在文档开始需注明采用的 XML 版本和编码方式（UTF-8、GB2312 等）、METS 文档中所遵循的所有标准（包括标准的版本）及其 URL、国家图书馆长期保存系统为该数字对象分配的唯一标识 OBJID（可以是 URN 等），以及一个描述资源的可读性标签。

本例采用了 XML 1.0 版本、UTF-8 编码方式、METS、MODS 等标准规范。国家图书馆可以根据自身的实际情况制定自己适用的 profile，并在 XML 命名域和 XML schema 实例（XSI）中加以引用。该 METS 文档中需要遵循的各种元数据规范（如国家图书馆专门元数据规范、国家图书馆管理元数据规范等）也需要一一列出。OBJID 可以是国家图书馆为长期保存系统的数字对象分配的统一资源名称（URN），在长期保存系统内为唯一。LABEL 可以是数字对象的标题，有助于对该 XML 文档的理解。

```
<!--说明 XML 版本为 1.0，使用编码方式为 UTF-8。-->  
<?xml version="1.0" encoding="UTF-8"?>  
<!--说明 XML 命名域（包括 METS、MODS、xLink、XSI 等）、所用标准规范（METS、MODS 等。  
国家图书馆标准规范可以在此揭示），以及对象标识等。-->  
<mets:mets xmlns:mets="http://www.loc.gov/METS/">
```

```

  xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/
  http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/mods/v3
  http://www.loc.gov/mods/v3/mods-3-1.xsd" OBJID="xxx:xxxxx/xxxx" LABEL="民国图书
  " >

```

3.2 METS 头元素

METS 头元素精简地记录关于 METS 文档中 METS 对象本身的描述元数据。这些元数据包括 METS 文档的创建日期、上次修改日期以及该 METS 文档的状态，也可以记录一个或多个承担过与该 METS 文档有关的某种任务的代理 (agent)，注明他们扮演的角色，再为其活动附上小说明之类。METS 头元素还可以记录多种该 METS 文档的其他标识符，补充 METS 文档的主要标识符（记录在 METS 根元素的 OBJID 属性中）。

表 2 METS 头

元素/属性	可重复性/必备性	规则与推荐规范
<metsHdr>	不可重复 必备	必须包含“CREATEDATE”及“LASTMODDATE”属性。
CREATEDATE	不可重复 必备	在 METS 中可选，本指南中必备。
LASTMODDATE	不可重复 必备	在 METS 中可选，本指南中必备。
ID、RECORDSTATUS	不可重复 可选	无推荐项。
-<agent>	可重复 必备	必须有一个代理实例是建立 METS 文档的机构，有一个实例是建立 METS 文档的软件及其版本。其他代理实例为可选。
ROLE	不可重复 必备	METS 中要求必备。对于机构，采用“CUSTODIAN”（托管者）；对于软件及其版本，采用“EDITOR”。如果有负责的个人，建议取值为“CREATOR”。
-<name>	不可重复 必备	必须包含适当的机构和软件及其版本的名称。
-<note>	可重复 可选	无推荐项。
-<altrecordID>	可重复 可选	无推荐项。

示例：

在本例中，其他记录 ID 使用的是国家图书馆长期保存系统的内部标识符编码，并展示

如何在文件中设置各种不同角色的代理。

```
<!--METS 文档的创建日期、最后修改日期、记录状态等。-->
<mets:metsHdr CREATEDATE="2004-02-22T00:00:00"
LASTMODDATE="2004-03-16T00:00:00" RECORDSTATUS="production">

<!--中国知网（METS 文档的创建者）向国家图书馆提交 SIP。-->
<mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
  <mets:name>中国知网</mets:name>
  <mets:note>中国知网向中国国家图书馆提交 SIP。</mets:note>
</mets:agent>

<!--国家图书馆（METS 文档的保存机构）的保存职能。-->
<mets:agent ROLE="PRESERVATION" TYPE="ORGANIZATION">
  <mets:name>中国国家图书馆</mets:name>
  <mets:note>SIP 及附属文件存储于保存仓储。</mets:note>
</mets:agent>

<!--国家图书馆的编辑职能。-->
<mets:agent ROLE="EDITOR" TYPE="ORGANIZATION">
  <mets:name>中国国家图书馆</mets:name>
  <mets:note>创建 DIP。</mets:note>
</mets:agent>

<!--国家图书馆的分发职能。-->
<mets:agent ROLE="DISSEMINATOR" TYPE="ORGANIZATION">
  <mets:name>中国国家图书馆</mets:name>
  <mets:note>分发到公共网络。</mets:note>
</mets:agent>

<!--其他记录 ID（可由国家图书馆统一指定）。-->
<mets:altRecordID TYPE="NLC">xxxxxxxx</altRecordID>
</mets:metsHdr>
```

3.3 描述元数据节<dmdSec>

METS 文档的描述元数据节由一个或多个<dmdSec>元素组成。每个<dmdSec>元素可以包含一个指向外部元数据的指针（<mdRef>元素），也可以包含内嵌元数据（嵌在<mdWrap>元素内），或二者兼备。

表 3 描述元数据节

元素/属性	可重复性/必备性	规则与推荐规范
-------	----------	---------

<dmdSec>	可重复 必备	<dmdSec>用于数字对象及其构成文件的书目描述和主题分析，一个元数据记录对应一个<dmdSec>。
		使用不同 schema 描述同一对象或部分的多个元数据记录应记录在不同的<dmdSec>，通过“GROUPID”属性相连接。
		整个数字对象的元数据记录必须至少有一个<dmdSec>。<dmdSec>中的描述元数据必须符合国家图书馆描述元数据规范。
		如可能，强烈推荐包含其他粒度更细的元数据记录（使用其他 schema 或命名域）
		每个<dmdSec>必须包含一个<mdWrap>元素。
ID	不可重复 必备	METS 要求必备。
GROUPID	不可重复 可选	用于区分多个（使用不同 schema）描述同一数字对象或部分的元数据记录。
AMID、CREATED、STATUS	不可重复 可选	无推荐项。
-<mdWrap>	不可重复 必备	参见“mdWrap”元素组部分。
-<mdRef>	-	本指南不支持。

示例：

下面是一个外部元数据引用实例，通过 URL 指向了外部电子图书描述元数据（元数据类型暂定为 NLCMD_EBOOK，国家图书馆可以根据实际情况指定相应的 URL）。

```

<mets:dmdSec ID="dmd001">
  <mets:mdRef LOCTYPE="URL" MDTYPE="NLCMD_EBOOK" LABEL="国家图书馆电子图书
  元数据规范"
    xlink:href="http://www.nlc.gov.cn/newstgc/gjsztsggc/bzgf/xxxx.html"
  />
</mets:dmdSec>

```

3.4 描述元数据节或管理元数据节内的<mdWrap>

内部描述元数据（mdWrap）元素提供封装器给内嵌于 METS 文档的元数据，而元数据是如下两种格式之一：（1）XML 编码元数据，它用 XML 编码表明本身所属于的命名空间不同于该 METS 文档的命名空间；（2）任何二进制或文本格式元数据，前提是这个元数据是 Base64 编码的且被封装在 mdWrap 元素内的<binData>元素中。

表 4 描述元数据节或管理元数据节内的<mdWrap>

元素/属性	可重复性/必备性	规则与推荐规范
<mdWrap>	不可重复 必备	
MDTYPE	不可重复 必备	<p>METS 要求出现该属性, 其属性取值范围如下: MARC、MODS、EAD、DC、NISOIMG、LC-AV、VRA、TEIHDR、DDI、FGDC、LOM、PREMIS、OTHER。</p> <p>国家图书馆描述元数据规范的“MDTYPE”取值为“OTHER”。</p> <p>本指南支持<dmdSec>使用国家图书馆描述元数据规范, <amdSec>使用 PREMIS 及国家图书馆管理元数据规范。如果适用于仓储的资料类型, 其他 schema 也可以获得支持。</p>
OTHERMDTYPE	不可重复 可选	仅当“MDTYPE”的值为“OTHER”时使用。如果使用国家图书馆描述元数据规范, 取值可以是“NLCDMD”。
<xmlData>	不可重复 必备	本指南必须有一个 schema 或命名域。首选已确认的 schema 或命名域, 也可以使用本地自定义的命名域。

示例 1:

本例以电子图书为例, 说明<mdWrap>元素的使用。元数据类型暂定为 NLC_EBOOK, 即国家图书馆电子图书元数据规范。国家图书馆可以根据实际需要决策引用哪些核心元素及其元素修饰词, 如题名、创建者、主题、描述等。其他类型的数字资源与此类似。

```

<mets:dmdSec ID="DMD1">

  <mets:mdWrap MDTYPE="nlc_ebook">

    <mets:xmlData>

      <nlc_ebook:nlc_ebook>

        <nlc_ebook:title>胡适文存</nlc_ebook:title>
        <nlc_ebook:creator>胡适</nlc_ebook:creator>
        ...
      </nlc_ebook:nlc_ebook>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>

```

```
</mets:xmlData>  
</mets:mdWrap>  
</mets:dmdSec>
```

示例 2:

内部描述元数据可以包含两个元素：`<binData>`和`<xmlData>`，前者包含 Base64 编码元数据的封装元素，后者包含 XML 编码元数据的封装元素。本例即以 MARC 数据为例，说明对于`<binData>`类型的内部描述元数据的引用。

```
<dmdSec ID="dmd003">  
  <mdWrap MIMETYPE="application/marc" MDTYPE="MARC" LABEL="OPAC 记录">  
    <binData>MDI0ODdjam0gIDIyMDA1ODkgYSA0NU0wMDAxMDA... [这里放置 base  
    64 编码数据]  
    </binData>  
  </mdWrap>  
</dmdSec>
```

3.5 管理元数据节

管理元数据节`<amdSec>`元素中的管理元数据，既可以描述组成该数字对象的文件，也可以描述生成该对象的原始素材。

METS 文档中规定了管理元数据的四种主要格式：1. 技术元数据（关于文件的创建、格式和使用特征的信息）；2. 知识产权元数据（版权和许可信息）；3. 来源元数据（关于该数字对象之模拟来源的描述元数据和管理元数据）；4. 数字起源元数据（关于作品的原始数字化形态与作为数字对象的当前形态之间的关系信息，包括文件之间的来源/目标关系、主体/派生关系和迁移/转换关系）。

上述四种类型的管理元数据分别对应于 METS 文档`<amdSec>`的子元素`<techMD>`、`<rightsMD>`、`<sourceMD>`和`<digiprovMD>`，其中可以嵌入相应格式的元数据。这四种元素在任何 METS 文档中均可以重复。

`<techMD>`、`<rightsMD>`、`<sourceMD>`和`<digiprovMD>`元素与`<dmdSec>`类似，既可以用`<mdRef>`元素指向外部管理元数据，也可以用`<mdWrap>`元素在 METS 文档内部嵌入管理元数据，或兼而有之。一个 METS 文档可有以上元素的多个实例，但它们都必须有 ID 属性，以便被该 METS 文档的其他元素（比如结构图中的 `div` 或`<file>`元素）关联到对应的`<amdSec>`子元素。

根据 *Guidelines for using PREMIS with METS for exchange*，可以使用一个管理元数据节`<amdSec>`，其下包含重复的子元素（如`techMD`等）；也可以在每个 METS 的子元素下使用重复的`<amdSec>`。本指南推荐把 PREMIS 各个实体放在单独的`amdSec` 的子元素下，并给每一个实体赋予 ID 值，如“AMD1”、“AMD2”等，引用方式为内部引用。至于在管理元

数据节中选用哪些 PREMIS 语义单元, 由《国家图书馆长期保存元数据标准规范应用指南》给出建议及其取值。

建议 PREMIS 语义单元放置在各个 METS 管理元数据节的子元素下, 而不是使用单个 PREMIS 信息包。如 2.1 小节所述, PREMIS 第一层数据元素应该用于 METS 下列节中:

premis:object: 放在 techMD 或 digiProvMD 下;

premis:event: 放在 digiProvMD 下;

premis:rights: 放在 rightsMD 下;

premis:agent: 放在 digiProvMD 或 rightsMD 下。

表 5 管理元数据节

元素/属性	可重复性/必备性	规则与推荐规范
<amdSec>	可重复 必备	必须有至少一个<amdSec>。 理想状态下, <fileSec>元素内包含或引用的每一个内容文件都应该有一个<amdSec>, 但可能在某些情况下不适用。 每个文件只能有一个<amdSec>。每一个<amdSec>可以包含可重复的<techMD>、<sourceMD>、<digiProvMD>和<rightsMD>。
		每一个<amdSec>必须包含一个 ID 属性, 以及至少一个<techMD>元素。
ID	不可重复 必备	ID 属性在 METS 中可选, 在本指南中必备。
-<techMD>	可重复 必备	每一个使用某个 schema 或命名域(如 PREMIS、MIX) 的技术元数据记录应该记录在其自身的<techMD>中。
		必须有至少一个<techMD>, 包含一个元数据记录, 在<mdWrap><xmlData>中, 遵循 PREMIS 对象 schema。 以下元素在 PREMIS 数据字典的“文件”对象类别中是必备的, 在本指南关于文件的<amdSec>中也必备。它们在 PREMIS xml schema 中不是必备的, 因为它们可能不适用于

	<p>所有对象类别的类型。</p> <ul style="list-style-type: none"> • objectIdentifierType • objectIdentifierValue • preservationLevel • objectCategory • compositionLevel • storageMedium <p>以下 PREMIS 对象实体的元素，在本指南中也要求必备：</p> <ul style="list-style-type: none"> • formatName • formatVersion • originalName <p>当不能提供相关数据时，在必备元素中，允许“not applicable”（不适用）和“unknown”（未知）等值。</p> <p>如能提供数据，则强烈推荐包括 PREMIS 对象 schema 中的任何可选元素。</p> <p>注意：以下元素可以包括在 PREMIS 元数据中，但在本指南中不是必备项，因为 SIZE、CHECKSUM 和 CHECKSUMTYPE 是本指南的 METS <file> 元素的必备属性。</p> <ul style="list-style-type: none"> • messageDigestAlgorithm • messageDigest • size
	<p>preservationLevel： 使用以下值之一：</p> <p>“supported”： 完全支持</p> <p>“known”： 不支持，但在试用和完全支持中是高优先级</p> <p>“unsupported”： 已知或未知格式，以其原有比特流保存，但在支持中是低优先级</p> <p>“not_applicable”： 该对象不是一个保存备份</p>

		对于静态图像, PREMIS 中没有的元数据可以使用 MIX schema (或国家图书馆图像元数据规范)。
		对于文本文件, PREMIS 中没有的元数据可以使用 textMD (或国家图书馆文本数据加工标准与工作规范)。
		音视频文件可以采用美国国会图书馆音视频原型项目 (Audio-Visual Prototyping Project) 的 schema, 或国家图书馆音频、视频元数据规范。
-ID	不可重复 必备	METS 要求必备。
-<mdWrap>	不可重复 必备	参见<mdWrap>元素组部分。
-<rightsMD>	可重复 可选	<p><rightsMD>可选。当出现时, 必须包含一个<mdWrap><rightsMD>元素, 可以包含以下值之一:</p> <ul style="list-style-type: none"> • METS 权利记录 (使用 METS Rights schema) • 知识共享许可协议 (Creative Commons distribution licences) • PREMIS 权利记录 • XACML
-ID	不可重复 必备	METS 要求必备。
-<mdWrap>	可重复 必备	参见<mdWrap>元素组部分。
-<sourceMD>	可重复 可选	<p>可以使用, 但不要求必备, 如果<dmdSec>描述了用于创建 METS 对象的原始资料来源的话 (如 METS 对象来源于一个物理对象)。可以用于描述原始与当前对象之间的原始资料 (当来源资料为非数字对象时)。本指南不推荐具体元数据的形式。</p> <p>如果<digiprovMD>包括了 PREMIS 事件元数据, 该元数据有链接到一个未作为该 METS 文档的一部分传递的对象的 linkingObjectIdentifier, 则必须使用。在这种情况下, <sourceMD>必须包含一个<premisObject:object>。例如, 从一个 Word 文档创建了一个 PDF, PDF 传递了, 而 Word 文</p>

		档案未被传递（Word 文档可能被保存仓储废弃了），Word 文档应在<sourceMD>中作为一个 PREMIS 对象加以描述。
ID	不可重复 必备	METS 要求必备。
--<mdWrap>	不可重复 必备	参见<mdWrap>元素组部分。
-<digiprovMD>	可重复 必备	<p>对于当前保存级对象，必须有至少一个 <digiprovMD>，描述保存仓储的摄入事件。对于非保存级对象，<digiprovMD>是可选的。</p> <p>每一个对象的事件记录，只能有一个 <digiprovMD>。</p> <p>每一个<digiprovMD>应当只有一个<mdWrap MDTYPE="PREMIS">。它只有一个<xmlData>元素，包含该对象的所有 PREMIS 事件和代理元数据。</p> <p>每一事件必须包含在单个<premisEvent:event>元素中，其 xml 数据遵循 PREMIS 事件 schema。</p> <p>当代理被记录时，每一个代理必须以单个的<premisAgent:agent>元素，被包含在相关联的<premisEvent:event>的同一<xmlData>元素中。</p>
		保存级数字对象的数字起源历史应尽可能完整，描述创建当前数字对象备份的事件（在单个的<premisEvent:event>元素中），以及其摄入到保存仓储的事件。这包括原始存放的对象的变化（注意在 PREMIS 中，一个对象不能被修改：一个修改对象的事件创建了一个新对象）以及托管者的变化。
		当前对象摄入到保存仓储以后发生的其他类型的事件可以被记录在其他的<premisEvent:event>元素中（如格式验证、校验和检查等）。
		本指南中 PREMIS 事件的以下元素要求必备：

	<ul style="list-style-type: none"> • eventIdentifierType • eventIdentifierValue • eventType • eventDateTime <p>如果事件改变了一个对象，则强烈推荐包括所使用的软硬件信息。使用 linkingAgentIdentifier 之下的以下事件元素：</p> <ul style="list-style-type: none"> • linkingAgentIdentifierType • linkingAgentIdentifierValue • linkingAgentRole <p>linkingAgentIdentifierType 内的值可以是外部注册表的名称或仓储自己的名称。</p> <p>如果该代理有标识符，linkingAgentIdentifierValue 应当是注册表内或保存仓储内的唯一标识符，或只是 METS 文档中此代理的唯一标识符。</p> <p>linkingAgentRole 的值应当描述代理的角色（如“扫描仪”等）。对于这一元素，尚未建立相应的受控词表。</p> <p>如果一个仓储以外的机构对事件负责，这一机构应当在 linkingAgentIdentifier 注明。</p>
	<p>如果存在一个 linkingAgentIdentifier， <premisAgent:agent>元素必须在包含 <premisEvent:event>元素的同一<xmlData>中 出现，<premisAgent:agent>和 <premisEvent:event>通过该事件的 linkingAgentIdentifier 相关联。本指南中以下元 素为必备项：</p> <ul style="list-style-type: none"> • agentIdentifierType • agentIdentifierValue

		<ul style="list-style-type: none"> • <code>agentName</code> (PREMIS 中可选) • <code>agentType</code> (PREMIS 中可选。对于这一元素, 尚未建立相应的受控词表。)
		<p>如果该对象通过一个事件与另一个数字对象相 关联, 相关的数字对象也提交到保存仓储, 该 事件应当包含一个 <code>premisEvent:linkingObjectIdentifier</code>, 与相关对象 的<code><amdSec><techMD></code>中的 <code>premisObject:objectIdentifier</code> 相匹配。</p> <p>如果该对象通过一个事件与另一个数字对象相 关联, 相关的数字对象未提交到保存仓储, 该 事件应当包含一个 <code>premisEvent:linkingObjectIdentifier</code>, 与 <code><sourceMD></code>中的 <code>premisObject</code> 元数据记录的 <code>premisObject:objectIdentifier</code> 相匹配。例如, 一 个 Word 文档转换成 RTF 格式, 继而转换成 PDF 格式。如果只有 PDF 文档提交到保存仓储, 每 一个事件都应当在单独的<code><premisEvent:event></code> 中加以描述, 其 <code>linkingObjectIdentifier</code> 与 <code><sourceMD></code>下的 <code>premisObject:object</code> 的 <code>ObjectIdentifier</code> 相匹配。Word 和 RTF 文件应当 在单独的<code><sourceMD></code>下的 <code>premisObject:object</code> 元素各自加以描述 (即使它们已不存在)。</p>
-ID	不可重复 必备	METS 要求必备。
--<mdWrap>	不可重复 必备	参见 <code><mdWrap></code> 元素组部分。

示例 1:

本例的`<techMD>`包含技术元数据, 它们和文件的数字化过程有关。国家图书馆可以根
据自身实际采用或制定具体的技术元数据规范。

```

<!--技术元数据, ID 为 ADM1-->
<mets:techMD ID="ADM1">
  <!--引用方式为内部引用, 元数据类型为 PREMIS-->
  <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS">
    <mets:xmlData>
      <!--PREMIS 对象实体, 其下包括对象标识、保存级别、对象类型、固定性、大小等-->

```

```
<pre:object>
  <pre:objectIdentifier>
    <pre:objectIdentifierType>ARK</pre:objectIdentifierType>
    <pre:objectIdentifierValue>
      http://www.nlc.gov.cn/ark:/20775/bb00000101/1-1.pdf
    </pre:objectIdentifierValue>
  </pre:objectIdentifier>
  <pre:preservationLevel>full</pre:preservationLevel>
  <pre:objectCategory>file</pre:objectCategory>
  <pre:objectCharacteristics>
    <pre:compositionLevel>0</pre:compositionLevel>
    <pre:fixity>
      <pre:messageDigestAlgorithm>
        CRC
      </pre:messageDigestAlgorithm>
      <pre:messageDigest>3,356,155,929</pre:messageDigest>
    </pre:fixity>
    <pre:size>251444</pre:size>
    ...
  </pre:objectCharacteristics>
  ...
</pre:object>
</mets:xmlData>
</mets:mdWrap>
</mets:techMD>
```

示例 2:

```
<!--权利元数据， ID 为 ADM2-->
<mets:rightsMD ID="ADM2">
  <!--引用方式为内部引用， 元数据类型为 PREMIS-->
  <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS">
    <mets:xmlData>
      <!--PREMIS 权利实体， 其下包括权利声明、 版权信息等-->
      <pre:rightsStatement>
        <pre:rightsBasis>Copyright</pre:rightsBasis>
        <pre:copyrightInformation>
```

```
<pre:copyrightStatus>Under copyright</pre:copyrightStatus>
<pre:copyrightJurisdiction>us</pre:copyrightJurisdiction>
<pre:copyrightNote>
    Rights Holder(s): Rachel Beth Egenhoefer
</pre:copyrightNote>
<pre:copyrightNote>
    Use: This work is available from the UC San Diego Libraries.
    This digital copy of the work is intended to support research,
    teaching, and private study.
</pre:copyrightNote>
<pre:copyrightNote>Constraint(s) on Use: This work is
protected by the U.S. Copyright Law (Title 17, U.S.C.). Use of this work beyond
that allowed by "fair use" requires written permission of the copyright holder(s).
Responsibility for obtaining permissions and any use and distribution of this
work rests exclusively with the user and not the UC San Diego Libraries. Inquiries
can be made to the UC San Diego Libraries department having custody of the work
(http://libraries.ucsd.edu/locations/mscl/).</pre:copyrightNote>
</pre:copyrightInformation>
</pre:rightsStatement>
</mets:xmlData>
</mets:mdWrap>
</mets:rightsMD>
```

示例 3:

本例中，数字音频由磁带数字化而来，应用<sourceMD>中的<mdWrap>元素来封装磁带中重要的技术元数据。

```
<sourceMD ID="SMD_MJF_Tape_0010_Side_A" STATUS="Draft, unformatted"
CREATED="2006-08-28T15:36:53">
<mdWrap MDTYPE="OTHER" OTHERMDTYPE="Draft AES-X098-SC-03-06-B, version
2006-05-19" MIMETYPE="audio/x-wave">
<xmlData>
    <dataroot xmlns:od="urn:schemas-microsoft-com:officedata"
generated="2006-08-28T15:36:53">
        <!--All values for source elements come from the source_metadata
        table.-->
        <source>
```

```
<ID>1</ID>
<source_metadata_id>SRC000000001</source_metadata_id>
<item_face_id_fk>IF000000001</item_face_id_fk>
<identifier>0001</identifier>
<identifier_type>MJF tape number</identifier_type>
<direction>A_PASS</direction>
<source_format_type>open reel
tape</source_format_type>
<analog_digital_flag>analog</analog_digital_flag>
<speed>7.5</speed>
<speed_unit>Inches per second</speed_unit>
<track_format>half track</track_format>
<sound_field>stereo</sound_field>
<noise_reduction>none</noise_reduction>
<equalization>none</equalization>
<gauge>0.25</gauge>
<gauge_unit>Inches</gauge_unit>
<reel_diameter>7</reel_diameter>
<reel_diameter_unit>Inches</reel_diameter_unit>
<bit_depth/>
<sample_rate/>
</source>
</dataroot>
</xmlData>
</mdWrap>
</sourceMD>
```

示例 4:

```
<!--数字起源元数据，ID 为 ADM4-->
<mets:digiprovMD ID="ADM4" CREATED="2006-11-27T21:37:13">
<!--引用方式为内部引用，元数据类型为 PREMIS-->
<mets:mdWrap MDTYPE="PREMIS">
<mets:xmlData xmlns:premis="http://www.loc.gov/standards/premis/v1
xsi:schemaLocation="http://www.loc.gov/standards/premis/v1
http://www.loc.gov/standards/premis/v1/Event-v1-1.xsd">
```

```
<!--PREMIS 事件实体，其下包括事件标识、事件类型、日期时间、明细等-->

<premis:eventIdentifier>
    <premis:eventIdentifierType>Validation
    </premis:eventIdentifierType>
    <premis:eventIdentifierType>Ingest_Validation_01
    </premis:eventIdentifierType>
</premis:eventIdentifier>
<premis:eventType>Validation</premis:eventType>
<premis:eventDateTime>2006-11-27-08:00</premis:eventDateTime>
<premis:eventDetail>ValidationResults</premis:eventDetail>
<premis:eventOutcomeInformation>
    <premis:eventOutcome>Pass</premis:eventOutcome>
    <premis:eventOutcomeDetail>
        <ingestValidation ID="FILE_epio1m.tiff">
            <JhoveOutput>File is of format: TIFF. The file is well
            -formed. The file is valid. </JhoveOutput>
            <IngestValidatorOutput>The computed checksum
            matches the original checksum for
            file[file:90990250/90990250.pdf]
            </IngestValidatorOutput>
            <ingestValidatorOutput>Original file ID:
            FILE_epio1m.tiff has been changed to
            FILE_07108e3e-5fd1-11da-b211- 19e7a5cf4814
            </IngestValidatorOutput>
            <IngestValidatorOutput>Original file name
            eip01m>tiff has been changed to
            FILE_07108e3e-5fd1-11da-b211-19e7a5cf4814, pdf
            </IngestValidatorOutput>
        </ingestValidation>
    </premis:eventOutcomeDetail>
</premis:eventOutcomeInformation>
<premis:linkingAgentIdentifier>
    <premis:linkingAgentIdentifierType>"Software" />
    <premis:linkingAgentIdentifierValue>"JHove" />
</premis:linkingAgentIdentifier>
<premis:linkingAgentIdentifier>
```

```

<premis:linkingAgentIdentifierType>"Software" />
<premis:linkingAgentIdentifierValue>"SDR_Ingest_Validator
" />
</premis:linkingAgentIdentifier>
</premis:event>
</mets:xmlData>
</mets:mdWrap>
</mets:digiprovMD>

```

3.6 文件节<fileSec>

文件节（<fileSec>）包含一个或多个<fileGrp>元素，把相关文件划分成组。<fileGrp>列出的文件组成了一个数字对象的某一电子版本。可以分别设立若干<fileGrp>元素，对于图像文件而言，有缩略图组、主档案影像组、pdf 版组、TEI 编码文本版组等等。

所有的<file>元素均有唯一的 ID 属性。此属性为该文件提供了唯一的内部名称，以被文档的其他部分引用。

<file>元素拥有的<FContent>元素可能要比<FLocat>元素更好一些。<FContent>元素将文件的实际内容嵌入到 METS 文档内。此时，文件内容必须是 XML 格式或 Base64 编码二者之一。当 METS 文档用来为用户展示数字对象时，并不一定要用内嵌文件，但当它用来在仓储间交换数字对象或离线存档数字对象时，就颇有价值了。

表 6 文件节

元素/属性	可重复性/必备性	规则与推荐规范
<fileSec>	不可重复 必备	所有文件必须通过<fileSec>加以引用。 <fileSec>必须包含一个或多个<fileGrp>。
<fileGrp>	可重复 必备	此元素用于对文件进行分组，根据以下在“USE”属性中描述的类别： original：存放者最初提交给仓储的对象，不同于保存级文件。 master：当前的保存级副本（即长期保存中最高优先级的版本）。它可能是最初或修订的版本，这应当从 digiprovMD 中决定。有且只有一个保存级副本。 access_representation：公共存取的副本，不同于保存级。 other_representation：内容文件的其他文件组，

		<p>可以用于呈现对象，既非原件，也不是保存级文件。</p> <p>structural_map: 强烈推荐其为单个的 XML 文件，如多媒体对象为 SMIL 文档，手稿为 EAD 文档，或一个复合对象的文件目录结构描述。</p> <p>引用的文件名应当符合<file>中 OWNERID 属性的文件名。如果该文件组提供的信息不多于<structMap>元素，就可以不需要。</p> <p>metadata: 可以包括额外的元数据文件。如果 METS 其他节（如描述元数据节、结构图等）提供的元数据已足够，就不需要。该组文件应当为 xml 文件。</p> <p>licence: 包含属于对象的许可或权利协议的文件。如果该文件组提供的信息不多于<rightsMD>元素，就可以不需要。</p> <p>support: 其他任何的支持的文件。</p> <p>other: 不符合以上类别的文件。</p>
		在本指南中，<fileGrp>必须包含一个或多个<file>，<fileGrp>可以不包含内嵌的<fileGrp>元素。
ID、VERSDATE、ADMID	不可重复 可选	无推荐项。
USE	不可重复 必备	每一<fileGrp>必须有一个“USE”属性，其值取自以上的词表中。
		<fileGrp>中 USE="master"的有且仅有一项。
		USE="original"的<fileGrp>可以是 0 或 1。
		其他类别可以是任意值，也可以没有。
--<file>		参见<file>元素组部分。

示例：

本例表示一本书中 TIFF 格式的页面图像，其用途为主图像（用于存档）。

```

<mets:fileSec>
  <mets:fileGrp ID="TIFF_GRP01" USE="MASTER IMAGE">
    ...
  </mets:fileGrp>
</mets:fileSec>

```

3.7 文件<file>

文件元素<file>为 METS 文档提供了描述内容文件数字对象的途径。文件元素<file>可以包含一个或多个<FLocat>元素，这些<FLocat>元素为内容文件或封装文件编译版本的<FContent>元素提供指针。在档案库之间交换数字对象，或存储数字对象的版本，嵌入文件中使用<FLocat>非常有用。所有的<FLocat>和<FContent>元素应该包含相同的单个文件。<file>元素具有递归性，因此允许下一级文件或某个大型文件的组成文件在目录中一一列出，例如，在档案库或压缩文件中包含的文件，就可以这种形式在目录中列出。或可以通过使用<stream>元素，文件中较小的部分或相关的部分就可以放入<file>元素中。最后，通过使用<transformFile>元素，可以在<file>元素中包含经过转换的文件的不同版本，例如经过格式转换的文件。

表 7 文件

元素/属性	可重复性/必备性	规则与推荐规范
<file>	可重复 必备	<file>必须包含一个<FLocat>或一个<FContent>。本指南目前不支持<stream>、<transformFile>或内嵌的<file>。
ID	不可重复 必备	METS 要求必备。
MIMETYPE	不可重复 必备	在 METS 中可选，本指南中必备。
SEQ	不可重复 可选	无推荐项。
SIZE	不可重复 必备	在 METS 中可选，本指南中必备。
CREATED	不可重复 可选	强烈推荐。应当是应用程序创建文件的日期，而不是摄入保存仓储的日期。
CHECKSUM	不可重复 必备	在 METS 中可选，本指南中必备。
CHECKSUMTYPE	不可重复 必备	在 METS 中可选，本指南中必备。METS 指定以下值： HAVAL MD5 SHA-1 SHA-256 SHA-512 TIGER WHIRLPOOL
OWNERID	不可重复 可选	可以用于提供一个唯一标识符（包括一个 URI）分配给文件，它可以不同于用于获取文件的

		URI。强烈推荐<fileGrp>中 USE="structural_map"的文件名引用 OWNERID, 或可以用于一个“根”文件, 以 重构或呈现一个对象。 必须用于提供指向文件的管理元数据的链接。
ADMID	不可重复 必备	必须用于提供指向文件的管理元数据的链接。
DMDID、GROUPID、 USE	不可重复 可选	无推荐项。
-<FLocat>	不可重复 可选	METS 中可重复, 本指南中不可重复。如果文 件内容没有内嵌于<FContent>, 必须为每一个 文件提供一个<FLocat>。
ID、USE	不可重复 可选	无推荐项。
LOCTYPE	不可重复 必备	METS 中要求必备, 限定取值如下: URN URL PURL HANDLE DOI OTHER
OTHERLOCTYPE	-	本指南不支持。
xlink	不可重复 必备	无推荐项。
-<FContent>	不可重复 可选	如果没有<FLocat>, 必须存在。METS 中已说 明, 内容文件可以是 Base 64 编码的, 包含在 binData 封装子元素内, 或由 XML 信息构成, 包含在 xmlData 封装子元素内。
ID、USE	不可重复 可选	无推荐项。
binData	不可重复 可选	无推荐项。
xmlData	不可重复 可选	无推荐项。

示例:

对于纸本图书而言, 每页用<div>表示, 可以是三种图像格式 (TIFF、JPEG 或 GIF) 中
的一种。作为文件组中的单个文件, TIFF 格式图像的<fileGrp>属性由以下的例子来说明。

<!--fileGrp 中说明了文件的 ID、使用类别、创建时间、文件大小、图像类型等-->

```
<mets:fileSec>
  <mets:fileGrp ID="TIFF_GRP01" USE="MASTER IMAGE">
```

```

<mets:file ID="epi01m" SIZE="65768"
CREATED="2006-04-11T07:35:22" MIMETYPE="image/tiff"
ADMID="MIX_v1.0_TIFF_epi01m">
...
</mets:file>
</mets:fileGrp>
</mets:fileSec>

```

3.8 结构图<structMap>（参见附录 A）

结构图定义了一种把数字对象呈现给用户的层次结构，并允许用户通过结构图导航。<structMap>元素利用一系列嵌套的<div>元素体现了这种层次结构。<div>将其所属类型用属性信息标明，而把对应于该 div 的内容，用多个 METS 指针元素（<mptr>）和文件指针元素（<fptr>）指向。METS 指针指向 METS 文档，其中包含了该<div>的相关文件信息。这种机制适于编码大量资料（例如一种完整的期刊），它可以保持集合内每个 METS 文档的尺寸较小。文件指针则指向当前 METS 文档<fileSec>节中的文件（有时指向文件组或某文件内的特定位置），被指向的文件对应于当前<div>所示的层次结构部分。

表 8 结构图

元素/属性	可重复性/必备性	规则与推荐规范
<structMap>	不可重复 必备	METS 中可重复，本指南中不可重复。 <structMap>必须描述 METS 文档表现的整个数字对象的结构。
ID	不可重复 可选	无推荐项。
TYPE	不可重复 可选	无推荐项。
LABEL	不可重复 可选	无推荐项。
-<div>	可重复 必备	

3.9 <div>

分区元素<div>用来表示按照层次组织起来的结构图是如何进行结构分区的，这种分区能够在任何程度上堆叠进行。每个<div>元素可以表示一个逻辑分区，也可以表示一个物理分区。通过 TYPE 属性，<div>元素能够明确分区的类型。例如在一本由<structMap>分区的电子书中，<div>元素表示各个章节和书页，在连续的层次结构中，TYPE 属性的取值可以是“书”、“章”和“页”。

<div>元素的 LABEL 属性是可选的，LABEL 属性的取值可以使数据对象按照特定的形

式显示，让数据对象的读者使用起来更方便。因此，层次结构的

元素的 LABEL 属性通常能够形成一种“内容列表”，这样数字图书馆就能使用“内容列表”来理解和指导 METS 编码的数字对象的内容。当然，在不同用途的 METS 标准中，LABEL 属性的目的也是不同的。

任何结构层次上的

元素都能通过 DMDID 属性，和一个或多个元素相关联。任何一个

元素引用的描述元数据都可以通过和

相关联形成一个完整的整体。任何结构层次上的

元素也可以和、、与/或元素包含的管理元数据相关联。

元素的这种关联能力，能够用于关联某个分区所表示的内容，这些分区通常由 rights 元数据来管理它的用途。例如，在 METS 编码的视频数据对象中，根

属性可以指向代表版权和使用限制信息的 rightsMD。

表 9 <div>元素组

元素/属性	可重复性/必备性	规则与推荐规范
ID	不可重复 可选	无推荐项。
ORDER	不可重复 必备	第一级 <div>元素代表整个对象，其取值必须是“1”。 低级别的<div>元素必须有一个 ORDER 值，至少有一个的值为“1”。同一年级、同一数值的<div>可以有多个。</div></div></div>
ORDERLABEL	不可重复 可选	无推荐项。
LABEL	不可重复 可选	强烈推荐。
DMDID	不可重复 可选	对于代表整个数字对象的第一级的 <div>而言，必须存在。对于低级别的<div>而言，如果该对象的部分有相应的<md>，则应当存在。</md></div></div>
ADMID	不可重复 可选	如果整个数字对象或其部分有相应的 <md>（不针对具体的文件格式），则应当存在，如权利元数据。</md>
TYPE、CONTENTIDS	不可重复 可选	无推荐项。
-<mptr>	-	本指南不支持。
-<fptr>	可重复 必备	每一个 <div>中必须有至少一个<fptr>。每一个<fptr>必须包含一个指向 METS 文档的文件的 FILEID。 其子元素<par>、<seq>和<area/>在本指南中不支持。</seq></par></div>
ID	不可重复 可选	无推荐项。

FILEID	不可重复 必备	在 METS 中可选, 本指南中必备。
CONTENTIDS	不可重复 可选	无推荐项。

示例:

本例解释

元素的特征, 包括 TYPE、LABEL 和 DMDID 属性。

```

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

<mets:mets xmlns:mets="http://www.loc.gov/METS/"
  xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/
    http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/mods/v3
    http://www.loc.gov/mods/v3/mods-3-1.xsd" OBJID="ark:/13030/kt9s2009hz"
  LABEL="Martial Epigrams">

  <mets:mdmdSec ID="DMD1">

    <mets:mdWrap MDTYPE="MODS">

      <mets:xmlData>

        <mods:mods>

          <mods:titleInfo>
            <mods:title>Martial Epigrams</mods:title>
          </mods:titleInfo>

        </mods:mods>

      </mets:xmlData>

    </mets:mdWrap>

  </mets:mdmdSec>

  ...

  <mets:structMap TYPE="physical">

    <mets:div TYPE="book" LABEL="Martial Epigrams II" DMDID="DMD1">

      <mets:div TYPE="page" LABEL="Blank page"/>
      <mets:div TYPE="page" LABEL="Page i: Series title page"/>
      <mets:div TYPE="page" LABEL="Page ii: Blank page"/>
      <mets:div TYPE="page" LABEL="Page iii: Title page"/>
      <mets:div TYPE="page" LABEL="Page iv: Publication info"/>
      <mets:div TYPE="page" LABEL="Page v: Table of contents"/>
      <mets:div TYPE="page" LABEL="Page vi: Blank page"/>
      <mets:div TYPE="page" LABEL="Page 1: Half title page"/>

    </mets:div>

  </mets:structMap>

```

```
<mets:div TYPE="page" LABEL="Page 2 (Latin)" />
<mets:div TYPE="page" LABEL="Page 3 (English)" />
<mets:div TYPE="page" LABEL="Page 4 (Latin)" />
<mets:div TYPE="page" LABEL="Page 5 (English)" />
<mets:div TYPE="page" LABEL="Page 6 (Latin)" />
<mets:div TYPE="page" LABEL="Page 7 (English)" />
</mets:div>
</mets:structMap>
</mets:mets>
```

3.10 结构链接<structLink>

本指南不对结构链接<structLink>进行定义。

3.11 行为节<behaviorSec>

本指南不对行为节<behaviorSec>进行定义。

4 应用场景

国家图书馆多格式、多来源的数字资源在批量导出时，形成多种类型的基于 METS 的 SIP 文件。

SIP 的格式和具体内容通常是信息生产者和 OAIS 系统（国家图书馆即充当这一角色）之间协商的结果。大多数 SIP 通常由一些内容信息和相关的保存描述信息组成，但有时可能需要多个 SIP 来形成一个完整的内容信息和相应的保存描述信息，从而进一步形成一个 AIP；一个 SIP 也可能被拆分为多个 AIP。SIP 对应的封装信息通常会以某种形式存在。

SIP 包需要符合 METS 基本标准规范，每个 SIP 可以规定为一个独立的知识实体，必须包含所有表现该知识实体的文件，如要保存一本以页为单位的书，则需要将该书的所有的页包含在信息包中，该文件包含了技术元数据、描述元数据和管理元数据。

摄取过程也是把 SIP 变为 AIP 的过程。功能包括接收 SIP、提交信息的预处理、对 SIP 数据解包、校验、收集、重新分发、不同环境下数据备份、处理。最后，摄取模块为数字对象分配唯一标识符，打包成 AIP 格式，以备存储使用。在打包成 AIP 格式前，它还创建文件的本地化、规范化的迁移版本，按照保存元数据规范的要求对每个 AIP 提供元数据值。摄入的具体过程可参考 DAITSS 系统。

AIP 包含支持对象长期保存和存取所需的所有元数据，AIP 是一个理论概念，实际数字典藏系统中没有必要将元数据和对象数据封装在同一个信息包中，因为有些元数据元素是固有或通用的（适合于数字典藏系统中的所有数字对象），所以就不用分别存储。长期保存元数据元素适用于 AIP。

4.1 一个 SIP 对应一个 AIP

在 OAIS 中, 一个或多个 SIP 可能需要被转换成为一个或多个 AIP 进行保存, 即 SIP 和 AIP 之间可以是一对一、一对多、多对一和多对多等关系。AIP 有一系列完整的保存描述信息和相关的内容信息。一个 SIP 还可能会由多个其他的 SIP 组成。AIP 的封装信息需要遵守 OAIS 内部的标准并服从 OAIS 的管理。

对于国家图书馆来说, 有不少一个 SIP 对应一个 AIP 的例子, 如石刻拓片影像、甲骨拓片影像等, 只需在 SIP 的基础上补充相应的 METS 头、保存元数据、管理元数据等信息即可。

4.2 多个 SIP 对应一个 AIP

一种期刊按期分批提交, 形成多个 SIP。如果要以这多个 SIP 形成一个 AIP 作长期保存, 则会出现多个 SIP 对应一个 AIP 的情况。又比如, 一本电子书的各个章节为一个独立的 SIP, 以这些章节构成一本完整的电子书, 作长期保存, 也会出现多个 SIP 对应一个 AIP 的情况。

以各期期刊构成一种期刊为例, 其流程包括接收 SIP、对提交信息进行预处理 (包括补充相应的保存描述信息、封装信息等)、对 SIP 数据进行解包、校验、收集和重新整理等。摄取的过程中, 为各个数字对象分配唯一标识符, 便于后续操作。在对多个 SIP 打包成 AIP 格式前, 还将创建文件的本地化、规范化的迁移版本, 按照保存元数据规范的要求对每个 AIP 提供元数据值。摄入的具体过程可参考 DAITSS 系统的相应操作。

比如, 数据库商提交某种期刊的某一期 (本例为《数字图书馆论坛》2008 年第 12 期, 总第 55 期), 所提交的 SIP 如下:

```
<!--数据库商提交《数字图书馆论坛》2008 年第 12 期, 总第 55 期。该期有十余篇文章, 为简略起见, 本例只举其中两篇文章。-->
```

```
<mets:mets>
  <!--描述元数据节-->
  <mets:mdmdSec>
    <!--要求数据库商提交的元数据类型为国家图书馆制定的期刊论文元数据著录规范, 其元数据命名域暂定为 NLCDMD。引用方式为内部引用。-->
    <mets:mdWrap MDTYPE="NLCDMD">
      <mets:xmlData>
        <nlcdmd:title>数字图书馆论坛</nlcdmd:title>
        <nlcdmd:creator>张秀梅</nlcdmd:creator>
        <nlcdmd:publisher>《数字图书馆论坛》编辑部</nlcdmd:publisher>
        <nlcdmd:contributor>顾晓光</nlcdmd:contributor>
```

```
    <nlcdmd:date>2008-12-25</nlcdmd:date>
    <nlcdmd:year>2008</nlcdmd:year>
    <nlcdmd:issue>55</nlcdmd:issue>
    <nlcdmd:format>pdf</nlcdmd:format>
    <nlcdmd:identifier>nlc001</nlcdmd:identifier>
    <nlcdmd:source>数字图书馆论坛</nlcdmd:source>
    <nlcdmd:language>chn</nlcdmd:language>
    <nlcdmd:rights>公共领域</nlcdmd:rights>
  </mets:xmlData>
</mets:mdWrap>
</mets:dmddSec>
<!--第一篇文章的描述元数据，元数据类型为国家图书馆制定的期刊论文元数据著录规范，其元数据命名域暂定为 NLCDMD。-->
<mets:dmddSec>
  <mets:mdWrap MDTYPE="NLCDMD" >
    <mets:xmlData>
      <nlcdmd:title>对虚拟社会信息交流的定性研究
      </nlcdmd:title>
      <nlcdmd:creator>韩圣龙</nlcdmd:creator>
      <nlcdmd:organization>北京大学信息管理系
      </nlcdmd:organization>
      <nlcdmd:role>撰写</nlcdmd:role>
      <nlcdmd:doi>10.3772/j.issn.1673-2286.2008.12.001
      </nlcdmd:doi>
      <nlcdmd:format>pdf</nlcdmd:format>
      <nlcdmd:language>chn</nlcdmd:language>
      <nlcdmd:issue>55</nlcdmd:issue>
      <nlcdmd:rights>公共领域</nlcdmd:rights>
      <nlcdmd:start_page>1</nlcdmd:start_page>
      <nlcdmd:end_page>1</nlcdmd:end_page>
      <nlcdmd:source>数字图书馆论坛</nlcdmd:source>
      <nlcdmd:total_page_number>1
      </nlcdmd:total_page_number>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmddSec>
```

```
<!--第二篇文章的描述元数据，元数据类型为国家图书馆制定的期刊论文元数据著录规范，其元数据命名域暂定为 NLCDMD。-->
```

```
<mets:mdWrap MDTYPE="NLCDMD">
  <mets:xmlData>
    <nlcdmd:title>对网络游戏中虚拟社会的观察研究
    </nlcdmd:title>
    <nlcdmd:alternative>以《魔兽世界》游戏的三年参与观察研究
```

为例

```
</nlcdmd:alternative>
<nlcdmd:creator>韩圣龙</nlcdmd:creator>
<nlcdmd:organization>北京大学信息管理系
</nlcdmd:organization>
<nlcdmd:explain>韩圣龙(1972-), 北京大学信息管理系副教授, 博士; 研究方向: 计算机情报检索。</nlcdmd:explain>
<nlcdmd:role>撰写</nlcdmd:role>
<nlcdmd:creator>屈鹏</nlcdmd:creator>
<nlcdmd:organization>北京大学信息管理系
</nlcdmd:organization>
<nlcdmd:explain>屈鹏 (1983-), 北京大学信息管理系博士研究生; 研究方向: 计算机情报检索。</nlcdmd:explain>
<nlcdmd:role>撰写</nlcdmd:role>
<nlcdmd:contributor>贾延霞</nlcdmd:contributor>
<nlcdmd:role>责任编辑</nlcdmd:role>
<nlcdmd:date>2008-08-30</nlcdmd:date>
<nlcdmd:year>2008</nlcdmd:year>
<nlcdmd:reference>KOLBERT E. Pimps and dragons: how an online world survived a social breakdown [N/OL]. The New Yorker, 2001-05-28 [2008-01-05].
http://www.newyorker.com/fact/content/?010528fa\_FACT.</nlcdmd:reference>
<nlcdmd:reference>杨德仁,顾君忠.网络游戏体系结构的研究与应用综述[J].计算机应用与软件,2007,24(3):113-116.</nlcdmd:reference>
  ...
<nlcdmd:keyword>网络游戏</nlcdmd:keyword>
<nlcdmd:keyword>观察研究</nlcdmd:keyword>
  ...

```

<nlcemd:abstract>大型多人在线角色扮演游戏（MMORPG）拥有大量的用户，在一些较为优秀的游戏中已经形成较为成熟的社会结构。与此同时，网络游戏又是现阶段最成熟虚拟现实技术的练兵场。因此，对网络游戏的虚拟社会进行研究有着理论和现实意义。文章以《魔兽世界》游戏为平台，对其中的虚拟社会的社会结构进行定性的实地观察研究。通过三年的参与观察，采用记录的方式给出了一个网络游戏中的虚拟社会的结构及其特征。以期通过这样的研究在方法论、理论和方法层面提供可借鉴的结果和方法。该文为《数字图书馆论坛》2008年第12期本期话题“虚拟社会”的文章之一。

```
</nlcdmd:abstract>

<nlcdmd:abstract_alternative>Thousands of users
take part in massively multiplayer online role playing games (MMORPG), in which
mature social structures form in excellent games. Meanwhile, online games are
also the maidens of virtual reality techniques. As the reasons above, the study
on the virtual society in online games has both theoretical and practical values.
The research is carried on World of Warcraft; concerns on the social structure
of virtual society with qualitative field observation. After having co-lived for
3 years, the authors record comparatively the structure of the virtual society
in online games and their characters. Hope this research can bring useful results
and methods for methodologies, theories and methods.
```

```
</nlcdmd:abstract_alternative>

<nlcdmd:doi>10.3772/j.issn.1673-2286.2008.12.001
</nlcdmd:doi>

<nlcdmd:format>pdf</nlcdmd:format>

<nlcdmd:language>chn</nlcdmd:language>

<nlcdmd:other_language>eng
</nlcdmd:other_language>

<nlcdmd:rights>公共领域</nlcdmd:rights>

<nlcdmd:source>数字图书馆论坛</nlcdmd:source>

<nlcdmd:issue>55</nlcdmd:issue>

<nlcdmd:start_page>2</nlcdmd:start_page>

<nlcdmd:end_page>6</nlcdmd:end_page>

<nlcdmd:total_page_number>5
</nlcdmd:total_page_number>

</mets:xmlData>

</mets:mdWrap>

</mets:dmddSec>

<mets:amdSec/>
```

```
<!--文件节-->
<mets:fileSec>
    <!--文件组，其类型为 pdf 文件，包括文件大小、校验和、校验和类型等。-->
    <mets:fileGrp USE="original">
        <mets:file MIMETYPE="application/pdf" SIZE="60278"
CHECKSUM="eb0fcf9ba0c30b0d2266139a38b6b7fa" CHECKSUMTYPE="MD5"
OWNERID="1.pdf">
            <mets:FContent>
                <mets:binData>Data</mets:binData>
            </mets:FContent>
        </mets:file>
        <mets:file MIMETYPE="application/pdf" SIZE="1552728"
CHECKSUM="58a82a739a9a30d39e35b88532777676" CHECKSUMTYPE="MD5"
OWNERID="2.pdf">
            <mets:FContent>
                <mets:binData>Data</mets:binData>
            </mets:FContent>
        </mets:file>
    </mets:fileGrp>
</mets:fileSec>
<!--结构图-->
<mets:structMap>
    <mets:div TYPE="issue">
        <mets:div TYPE="文章">
            <mets:ptr FILEID="nlc003001"/>
        </mets:div>
        <mets:div TYPE="文章">
            <mets:ptr FILEID="nlc003002"/>
        </mets:div>
    </mets:div>
</mets:structMap>
</mets:mets>
```

METS 指针元素（`mptr`）允许 `div` 关联到另一个 METS 文档，与该 `div` 相应的内容则存在那个 METS 文档中，而不是让 `mptr` 元素指向内部文件或文件组。因此 METS 指针元素可

以用来封装连续出版的期刊，每一个 div 元素表示一个卷期，div 元素则分别指向独立的卷期 METS 文档，而不是把每卷的文件和文件组都编码在一个大文档中。具体可参考信息包封装规范中的 mptr 元素应用细则。

本指南建议国家图书馆在对连续出版的期刊作长期保存时，把文章级别作为长期保存的最小粒度。因此，期刊级别的 AIP 包可以用 METS 指针指向各卷期，各卷期用文件指针指向各篇文章，形成层次清晰的层级结构。以下是期刊《数字图书馆论坛》对应的 AIP 包：

```
<!--本封装实例为《数字图书馆论坛》所有已出版刊期的 AIP 包。-->
<!--METS 根元素组，补充说明 METS 所封装的对象标识、类型、XML 命名域等。-->
<mets:mets OBJID="METS-d1e1" TYPE="issue" xmlns="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/"
  xmlns:nlcdmd="http://www.nlc.gov.cn/newstgc/gjsztsggc/bzgf/..."
  http://www.loc.gov/standards/mets/mets.xsd
  http://www.loc.gov/standards/premis/v2/premis-v2-0.xsd
  xmlns:premis="http://www.loc.gov/standards/premis/v2/">

  <!--METS 头，说明本 METS 文档的创建时间、最后修订时间、代理等。-->
  <mets:metsHdr CREATEDATE="2011-09-25T09:03:05.202+10:00"
    LASTMODDATE="2011-09-25T09:03:05.202+10:00">

    <mets:agent ROLE="DISSEMINATOR" TYPE="ORGANIZATION">
      <mets:name>中国国家图书馆</mets:name>
    </mets:agent>
    <mets:agent ROLE="CREATOR" TYPE="OTHER">
      <mets:name>中国国家图书馆</mets:name>
    </mets:agent>
  </mets:metsHdr>

  <!--刊期级别的描述元数据，带有 ID 标识，其 ID 由国家图书馆唯一标识符标准规范唯一确定，国家图书馆可根据实际情况对数字对象赋予唯一的 ID 值。-->
  <mets:dmrdSec ID="nlc001">

    <!--引用方式为内部引用，元数据类型为国家图书馆制定的期刊论文元数据著录规范，其元数据命名域暂定为 NLCDMD。-->
    <mets:mdWrap MDTYPE="NLCDMD">
      <mets:xmlData>
        <nlcdmd:title>数字图书馆论坛</nlcdmd:title>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmrdSec>
</mets:mets>
```

```
    <nlcemd:creator>张秀梅</nlcemd:creator>
    <nlcemd:publisher>《数字图书馆论坛》编辑部
    </nlcemd:publisher>
    <nlcemd:contributor>顾晓光</nlcemd:contributor>
    <nlcemd:format>pdf</nlcemd:format>
    <nlcemd:source>数字图书馆论坛</nlcemd:source>
    <nlcemd:language>chn</nlcemd:language>
    <nlcemd:rights>公共领域</nlcemd:rights>
    ...
    </mets:xmlData>
  </mets:mdWrap>
</mets:mdSec>
<!--管理元数据节-->
<mets:amdSec>
  <!--技术元数据-->
  <mets:techMD ID="nlc002001">
    <!--引用方式为内部引用，元数据类型为 PREMIS，包括 PREMIS 对象标识、
    保存级别、对象类别等。国家图书馆可根据国家数字图书馆长期保存元数据规范及应用指南选择相应的
    PREMIS 长期保存语义单元。-->
    <mets:mdWrap MDTYPE="PREMIS">
      <mets:xmlData>
        <premis:object
          xmlns:premis="http://www.loc.gov/standards/premis">
          <premis:objectIdentifier>
            <premis:objectIdentifierType>auto
            </premis:objectIdentifierType>
            <premis:objectIdentifierValue>METS-d1e1
            </premis:objectIdentifierValue>
          </premis:objectIdentifier>
          <premis:preservationLevel>full
          </premis:preservationLevel>
          <premis:objectCategory>Representation
          </premis:objectCategory>
        </premis:object>
      </mets:xmlData>
    </mets:mdWrap>
  
```

```
</mets:techMD>
</mets:amdSec>
<!--文件节-->
<mets:fileSec>
    <!--文件组，其类型为 pdf 文件，包括文件大小、校验和、校验和类型、所有者 ID 等。本例规定国家图书馆期刊长期保存的最小粒度是文章级，本节应包括已出版刊期的所有文章。为简略起见，以下只包括了两篇文章的 pdf 文档。-->
    <mets:fileGrp USE="original">
        <mets:file ID="nlc003001" MIMETYPE="application/pdf" SIZE="60278" CHECKSUM="eb0fcf9ba0c30b0d2266139a38b6b7fa" CHECKSUMTYPE="MD5" OWNERID="1.pdf">
            <mets:FContent>
                <mets:binData>Data</mets:binData>
            </mets:FContent>
        </mets:file>
        <mets:file ID="nlc003002" MIMETYPE="application/pdf" SIZE="1552728" CHECKSUM="58a82a739a9a30d39e35b88532777676" CHECKSUMTYPE="MD5" OWNERID="2.pdf">
            <mets:FContent>
                <mets:binData>Data</mets:binData>
            </mets:FContent>
        </mets:file>
    </mets:fileGrp>
</mets:fileSec>
<!--结构图-->
<mets:structMap>
    <mets:div TYPE="issue" DMDID="nlc001001" ADMID="nlc002001">
        <mets:div TYPE="volume" LABEL="Volume I">
            <mets:mptr LOCTYPE="URL" xlink:href="http://www.nlc.gov.cn/ltpreservation/journal/digiLibForum001.xml" />
            <mets:div>
                <mets:div TYPE="volume" LABEL="Volume I">
                    <mets:mptr LOCTYPE="URL" xlink:href="http://www.nlc.gov.cn/ltpreservation/journal/digiLibForum002.xml" />

```

```
    </mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

4.3 一个 SIP 对应多个 AIP

一个 SIP 对应多个 AIP 的例子：一个数据库商提交数据库包，记录为一个 SIP。这些记录作为两个 AIP 存储：一个包含公开信息，另一个包含非公开或保密信息。这样对信息的访问更为便利。又比如，一个音乐 CD 作为一个 SIP 提交，要对音乐 CD 中的每首乐曲作长期保存，则每首乐曲对应一个 AIP，这就出现一个 SIP 对应多个 AIP 的情况。

对于一个音乐 CD 对应多首乐曲的情况，AIP 的保存描述信息和封装信息来自于 SIP。要在描述元数据节中说明该首乐曲与音乐 CD 的继承关系，在管理元数据节中补充相应的保存描述信息，文件节中要说明保存的主文件（属于保存级别）、发布级别的衍生文件等。结构图中各个文件的关系为平行关系。

以下是一个音乐 CD 的例子。提交的 SIP 中，包括 wav 和 MP3 两种文件格式，前者为长期保存级，后者为发布服务级。其中发布服务级又可能会有好几种码流供选择。本例为简略起见，发布服务级的 MP3 格式只包含一种。

```
<!--音乐 CD 的 SIP 包。为简略起见，本例只包括两首乐曲。-->
<mets:mets>
  <!--描述元数据节-->
  <mets:dmrdSec/>
  <!--管理元数据节-->
  <mets:amrdSec/>
  <!--文件节-->
  <mets:fileSec>
    <!--文件组，长期保存级别类型为 wav 格式，包括 ID、文件大小、URL 等。-->
    <mets:fileGrp USE="MASTER">
      <mets:file ID="Fd408e110" GROUPID="G1" MIMETYPE="audio/wav"
SIZE="41101258"><mets:FLocat LOCTYPE="URL"
xlink:href="http://www.nlc.gov.cn/ltpreservation/cd/audio/0001.wav"/>
      </mets:file>
      <mets:file ID="Fd408e118" GROUPID="G2" MIMETYPE="audio/wav"
SIZE="22050058"><mets:FLocat LOCTYPE="URL"
xlink:href="http://www.nlc.gov.cn/ltpreservation/cd/audio/0001.wav"/>
      </mets:file>
```

```
<!--文件组，网络服务级别类型为MP3格式，包括ID、文件大小、URL等。-->
<mets:fileGrp USE="SERVICE">

    <mets:file ID="Fd408e736" GROUPID="G1" MIMETYPE="audio/mp3"
SIZE="2797568"><mets:FLocat LOCTYPE="URL"
xlink:href="http://www.nlc.gov.cn/ltpreservation/cd/audio/0001.mp3"/>
    </mets:file>

    <mets:file ID="Fd408e750" GROUPID="G2" MIMETYPE="audio/mp3"
SIZE="1503232"><mets:FLocat LOCTYPE="URL"
xlink:href="http://www.nlc.gov.cn/ltpreservation/cd/audio/0001.mp3"/>
    </mets:file>

</mets:fileGrp>

</mets:fileSec>

<!--结构图-->
<mets:structMap>

    <mets:div DMDID="MODS1" TYPE="cd:compactDiscObject">

        <mets:div TYPE="cd:disc">

            <mets:div TYPE="cd:track" ID="disc01_tr001"
DMDID="DMD_disc01_tr001">

                <mets:div TYPE="cd:audio">

                    <mets:ptr FILEID="Fd408e110"/>
                    <mets:ptr FILEID="Fd408e736"/>

                </mets:div>

            </mets:div>

            <mets:div TYPE="cd:track" ID="disc01_tr002"
DMDID="DMD_disc01_tr002">

                <mets:div TYPE="cd:audio">

                    <mets:ptr FILEID="Fd408e118"/>
                    <mets:ptr FILEID="Fd408e750"/>

                </mets:div>

            </mets:div>

        </mets:div>

    </mets:div>

</mets:structMap>

</mets:mets>
```

在对每首乐曲作长期保存时，需要注意乐曲和音乐CD的继承关系。在对单首乐曲作长

期保存时，只保存 wav 格式的文件。以下是单首乐曲的 AIP 包。

```
<!--音乐 CD 单首乐曲的 AIP 包。-->
<!--METS 根元素组，说明 METS 所封装的对象标识、类型、XML 命名域等。-->
<mets:mets>
  <!--METS 头，说明本 METS 文档的创建时间、最后修订时间、代理等。本例从略。-->
  <mets:metsHdr/>
  <!--描述元数据节，带有 ID 标识，其标识符由国家图书馆唯一标识符标准规范赋值，需要说明单首乐曲与音乐 CD 的继承关系。本例描述元数据节从略。-->
  <mets:dmdSec ID="DMD1">
    </mets:dmdSec>
  <!--管理元数据节，带有 ID 标识，其标识符由国家图书馆唯一标识符标准规范赋值。本例管理元数据节从略。-->
  <mets:amdSec ID="AMD1">
    </mets:amdSec>
  <!--文件节-->
  <mets:fileSec>
    <!--文件组，其类型为 wav 文件格式，包括文件大小、URL、所有者 ID 等。-->
    <mets:fileGrp USE="MASTER">
      <mets:file ID="Fd408e110" MIMETYPE="audio/wav"
        SIZE="41101258"><mets:fLocat LOCTYPE="URL"
        xlink:href="http://www.nlc.gov.cn/ltpreservation/cd/audio/0001.wav"/>
      </mets:file>
    </mets:fileSec>
    <!--结构图-->
    <mets:structMap>
      <mets:div DMDID="DMD1" TYPE="cd:compactDiscObject">
        <mets:div TYPE="cd:disc">
          <mets:div TYPE="cd:track" ID="disc01_tr001"
            DMDID="DMD_disc01_tr001">
            <mets:div TYPE="cd:audio">
              <mets:fptr FILEID="Fd408e110"/>
            </mets:div>
          </mets:div>
        </mets:div>
      </mets:structMap>
```

```
</mets:mets>
```

4.4 多个 SIP 对应多个 AIP

多个 SIP 对应多个 AIP 的例子：一个数据库商收集其数据库的数据。它每年提交的 SIP 包含一系列电子期刊的所有状态信息，按时间封包。AIP 存储关于整个期刊数据库的数据，需要把各个电子期刊的信息放到相应的 AIP 中。

本指南建议一个 SIP 只包含一个知识实体，因此本例不具实际意义。

一个多目录的 SIP 包括一个描述文件和多个位于子目录下的内容文件，如下所示：

```
/NLCJNL0001 (SIP 目录, 名为 NLCJNL0001)
  NLCJNL0001.xml (METS 描述文件, 位于最高级别的目录下)
    /aaa (下一级目录)
      0001.pdf
      0002.pdf
      ...
    /zzz (下一级目录)
      0006.pdf
      0007.pdf
      ...
/NLCJNL0001 (SIP 目录, 名为 NLCJNL0001)
  NLCJNL0001.xml (METS 描述文件, 位于最高级别的目录下)
    /aaa (下一级目录)
      0001.pdf
      0002.pdf
      ...
    /zzz (下一级目录)
      0006.pdf
      0007.pdf
      ...
```

数据库商提交的 SIP 必须包含一个描述整包信息的 METS 文件，以及至少一个内容文件。对于电子期刊来说，本指南建议在描述文件中包括期刊的卷期信息，即卷期信息不仅要在描述元数据节中提供，还要求在描述文件的结构图节中的 LABEL、ORDERLABEL、TYPE 属性中加以揭示，如下所示。

```
<METS:structMap>
```

```
<METS:div DMDID="DMD1" LABEL="Volume 25 (2005-2006)" ORDERLABEL="25"
TYPE="volume">

<METS:div DMDID="DMD2" LABEL="Number 3" ORDERLABEL="3" TYPE="issue">

<METS:ptr FILEID="[文件唯一标识符]">

. . .

</METS:div>

</METS:div>

</METS:structMap>
```

提交的 SIP 包在存档成 AIP 前，需经过文件格式识别、验证、抽取技术元数据、创建存档版本文件（如提交的若是 caj 格式，需转换成 pdf 格式）等。

4.5 从 AIP 到 DIP

分发信息包 (DIP) 来源于一个或多个 AIP，由 OAIS 发送给消费者。DIP 根据消费者的请求发送，可能是 AIP 的所有或部分。DIP 和 AIP 在形式和内容上可能都有不同。两者的区别包括内容格式（如在分发前，一个图像文件可能会从 TIFF 格式转换为 JPEG 格式）、内容总量（一个 DIP 可能相当于一个 AIP、多个 AIP，甚至是一个 AIP 的一部分）、随内容提供的元数据数量（很可能 DIP 不包含一个存档数字对象的所有元数据集，因为消费者对其中很多元数据没什么兴趣）。

示例：

网络发布的 JPG 图像由 TIFF 图像派生。两个图像文件<file>元素的 GROUPID 属性值相同，都是“IMAGE01”，这表明虽然这两个文件格式不同，但包含相同的基本信息。除了要在 DIP 的 METS 文档的管理元数据节中表明两者的派生关系以外，在文件节的描述如下：

```
<mets:fileSec>

<mets:fileGrp ID="FORMAT1" USE="Master Image">

<mets:file ID="FILE001" MIMETYPE="image/tiff" SIZE="1121028"
CREATED="2012-02-17" GROUPID="IMAGE1">

<mets:FLocat LOCTYPE="URL"
http://www.nlc.org.cn/image/gloria.tiff>

</mets:FLocat>

</mets:fileGrp>

<mets:fileGrp ID="FORMAT2" VERSDATE="2012-03-10" USE="Derivative
Image">

<mets:file ID="FILE002" MIMETYPE="image/jpeg" SIZE="966212"
CREATED="2012-02-18" GROUPID="IMAGE1">
```

```
<mets:FLocat LOCTYPE="URL"
  http://www.nlc.org.cn/image/gloria.jpg>
</mets:FLocat>
</mets:file>
</mets:fileGrp>
</mets:fileSec>
```

附录 A

(资料性附录)

电子期刊封装实例 (期刊级别)

<!--本封装实例针对的例子是电子期刊, 为《数字图书馆论坛》2008 年第 12 期, 总第 55 期。该期有十余篇文章, 本例只举其中两篇文章。-->

```
<!--METS 根元素组, 说明 METS 所封装的对象标识、类型、XML 命名域等-->

<mets:mets OBJID="METS-dle1" TYPE="issue" xmlns="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/"
  xmlns:nlcdmd="http://www.nlc.gov.cn/newstgc/gjsztsggc/bzgf/..."
  http://www.loc.gov/standards/mets/mets.xsd
  http://www.loc.gov/standards/premis/v2/premis-v2-0.xsd
  xmlns:premis="http://www.loc.gov/standards/premis/v2/">

  <!--METS 头, 说明本 METS 文档的创建时间、最后修订时间、代理等。-->
  <mets:metsHdr CREATEDATE="2011-09-25T09:03:05.202+10:00"
    LASTMODDATE="2011-09-25T09:03:05.202+10:00">

    <mets:agent ROLE="DISSEMINATOR" TYPE="ORGANIZATION">
      <mets:name>中国国家图书馆</mets:name>
    </mets:agent>

    <mets:agent ROLE="CREATOR" TYPE="OTHER">
      <mets:name>中国国家图书馆</mets:name>
    </mets:agent>
  </mets:metsHdr>

  <!--期刊级别的描述元数据, 带有 ID 标识, 其 ID 由国家图书馆唯一标识符标准规范唯一确定, 国家图书馆可根据实际情况对数字对象赋予唯一的 ID 值。-->
  <mets:dmrdSec ID="nlc001">

    <!--引用方式为内部引用, 元数据类型为国家图书馆制定的期刊论文元数据著录规范, 其元数据命名域暂定为 NLCDMD。-->

    <mets:mdWrap MDTYPE="NLCDMD">
      <mets:xmlData>
        <nlcdmd:title>数字图书馆论坛</nlcdmd:title>
        <nlcdmd:creator>张秀梅</nlcdmd:creator>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmrdSec>
</mets:mets>
```

```
<nlcdmd:publisher>《数字图书馆论坛》编辑部
</nlcdmd:publisher>

<nlcdmd:contributor>顾晓光</nlcdmd:contributor>
<nlcdmd:date>2008-12-25</nlcdmd:date>
<nlcdmd:year>2008</nlcdmd:year>
<nlcdmd:issue>55</nlcdmd:issue>
<nlcdmd:format>pdf</nlcdmd:format>
<nlcdmd:identifier>nlc001</nlcdmd:identifier>
<nlcdmd:source>数字图书馆论坛</nlcdmd:source>
<nlcdmd:language>chn</nlcdmd:language>
<nlcdmd:rights>公共领域</nlcdmd:rights>
</mets:xmlData>
</mets:mdWrap>
</mets:dmSec>
<!--第一篇文章的描述元数据，元数据类型为国家图书馆制定的期刊论文元数据著录规范，其元数据命名域暂定为 NLCDMD。ID 值由国家图书馆唯一标识符标准规范唯一确定。-->
<mets:dmSec ID="nlc001001">
<mets:mdWrap MDTYPE="NLCDMD">
<mets:xmlData>
<nlcdmd:title>对虚拟社会信息交流的定性研究
</nlcdmd:title>
<nlcdmd:creator>韩圣龙</nlcdmd:creator>
<nlcdmd:organization>北京大学信息管理系
</nlcdmd:organization>
<nlcdmd:role>撰写</nlcdmd:role>
<nlcdmd:doi>10.3772/j.issn.1673-2286.2008.12.001
</nlcdmd:doi>
<nlcdmd:format>pdf</nlcdmd:format>
<nlcdmd:language>chn</nlcdmd:language>
<nlcdmd:issue>55</nlcdmd:issue>
<nlcdmd:rights>公共领域</nlcdmd:rights>
<nlcdmd:start_page>1</nlcdmd:start_page>
<nlcdmd:end_page>1</nlcdmd:end_page>
<nlcdmd:source>数字图书馆论坛</nlcdmd:source>
<nlcdmd:total_page_number>1
</nlcdmd:total_page_number>
```

```
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmmdSec>
<!--第二篇文章的描述元数据，元数据类型为国家图书馆制定的期刊论文元数据著录规范，其元数据命名域暂定为 NLCDMD。--&gt;
&lt;mets:dmmdSec ID="nlc001002"&gt;
  &lt;mets:mdWrap MDTYPE="NLCDMD"&gt;
    &lt;mets:xmlData&gt;
      &lt;nlcdmd:title&gt;对网络游戏中虚拟社会的观察研究
      &lt;/nlcdmd:title&gt;
      &lt;nlcdmd:alternative&gt;以《魔兽世界》游戏的三年参与观察研究
    为例
      &lt;/nlcdmd:alternative&gt;
      &lt;nlcdmd:creator&gt;韩圣龙&lt;/nlcdmd:creator&gt;
      &lt;nlcdmd:organization&gt;北京大学信息管理系
      &lt;/nlcdmd:organization&gt;
      &lt;nlcdmd:explain&gt;韩圣龙(1972-), 北京大学信息管理系副教授, 博士; 研究方向: 计算机情报检索。&lt;/nlcdmd:explain&gt;
      &lt;nlcdmd:role&gt;撰写&lt;/nlcdmd:role&gt;
      &lt;nlcdmd:creator&gt;屈鹏&lt;/nlcdmd:creator&gt;
      &lt;nlcdmd:organization&gt;北京大学信息管理系
      &lt;/nlcdmd:organization&gt;
      &lt;nlcdmd:explain&gt;屈鹏 (1983-), 北京大学信息管理系博士研究生; 研究方向: 计算机情报检索。&lt;/nlcdmd:explain&gt;
      &lt;nlcdmd:role&gt;撰写&lt;/nlcdmd:role&gt;
      &lt;nlcdmd:contributor&gt;贾延霞&lt;/nlcdmd:contributor&gt;
      &lt;nlcdmd:role&gt;责任编辑&lt;/nlcdmd:role&gt;
      &lt;nlcdmd:date&gt;2008-08-30&lt;/nlcdmd:date&gt;
      &lt;nlcdmd:year&gt;2008&lt;/nlcdmd:year&gt;
      &lt;nlcdmd:reference&gt;KOLBERT E. Pimps and dragons: how an online world survived a social breakdown [N/OL]. The New Yorker, 2001-05-28 [2008-01-05].
<a href="http://www.newyorker.com/fact/content/?010528fa_FACT">http://www.newyorker.com/fact/content/?010528fa_FACT.</nlcdmd:reference>
      <nlcdmd:reference>杨德仁,顾君忠.网络游戏体系结构的研究与应用综述[J].计算机应用与软件,2007,24(3):113-116.</nlcdmd:reference>
    ...
  ...

```

```
<nlcemd:keyword>网络游戏</nlcdmd:keyword>
```

```
<nlcemd:keyword>观察研究</nlcdmd:keyword>
```

```
...
```

```
<nlcemd:abstract>大型多人在线角色扮演游戏（MMORPG）拥有大量的用户，在一些较为优秀的游戏中已经形成较为成熟的社会结构。与此同时，网络游戏又是现阶段最成熟虚拟现实技术的练兵场。因此，对网络游戏的虚拟社会进行研究有着理论和现实意义。文章以《魔兽世界》游戏为平台，对其中的虚拟社会的社会结构进行定性的实地观察研究。通过三年的参与观察，采用记录的方式给出了一个网络游戏中的虚拟社会的结构及其特征。以期通过这样的研究在方法论、理论和方法层面提供可借鉴的结果和方法。该文为《数字图书馆论坛》2008年第12期本期话题“虚拟社会”的文章之一。
```

```
</nlcdmd:abstract>
```

```
<nlcemd:abstract_alternative>Thousands of users
```

```
take part in massively multiplayer online role playing games (MMORPG), in which mature social structures form in excellent games. Meanwhile, online games are also the maidens of virtual reality techniques. As the reasons above, the study on the virtual society in online games has both theoretical and practical values. The research is carried on World of Warcraft; concerns on the social structure of virtual society with qualitative field observation. After having co-lived for 3 years, the authors record comparatively the structure of the virtual society in online games and their characters. Hope this research can bring useful results and methods for methodologies, theories and methods.
```

```
</nlcdmd:abstract_alternative>
```

```
<nlcemd:doi>10.3772/j.issn.1673-2286.2008.12.001
```

```
</nlcdmd:doi>
```

```
<nlcemd:format>pdf</nlcdmd:format>
```

```
<nlcemd:language>chn</nlcdmd:language>
```

```
<nlcemd:other_language>eng
```

```
</nlcdmd:other_language>
```

```
<nlcemd:rights>公共领域</nlcdmd:rights>
```

```
<nlcemd:source>数字图书馆论坛</nlcdmd:source>
```

```
<nlcemd:issue>55</nlcdmd:issue>
```

```
<nlcemd:start_page>2</nlcdmd:start_page>
```

```
<nlcemd:end_page>6</nlcdmd:end_page>
```

```
<nlcemd:total_page_number>5
```

```
</nlcdmd:total_page_number>
```

```
</mets:xmlData>
```

```
</mets:mdWrap>

</mets:dmfdSec>

<!--管理元数据节-->

<mets:amdSec>

<!--技术元数据-->

<mets:techMD ID="nlc002001">

    <!--引用方式为内部引用，元数据类型为 PREMIS，包括 PREMIS 对象标识、
    保存级别、对象类别等。国家图书馆可根据国家数字图书馆长期保存元数据规范及应用指南选择相应的
    PREMIS 长期保存语义单元。-->

        <mets:mdWrap MDTYPE="PREMIS">

            <mets:xmlData>

                <premis:object
                    xmlns:premis="http://www.loc.gov/standards/premis">

                    <premis:objectIdentifier>
                        <premis:objectIdentifierType>auto
                        </premis:objectIdentifierType>
                        <premis:objectIdentifierValue>METS-d1e1
                        </premis:objectIdentifierValue>
                    </premis:objectIdentifier>
                    <premis:preservationLevel>full
                    </premis:preservationLevel>
                    <premis:objectCategory>Representation
                    </premis:objectCategory>
                </premis:object>
            </mets:xmlData>
        </mets:mdWrap>
    </mets:techMD>
</mets:amdSec>

<!--文件节-->

<mets:fileSec>

    <!--文件组，其类型为 pdf 文件，包括文件大小、校验和、校验和类型、所有者 ID
    等。-->

        <mets:fileGrp USE="original">

            <mets:file ID="nlc003001" MIMETYPE="application/pdf"
                SIZE="60278" CHECKSUM="eb0fcf9ba0c30b0d2266139a38b6b7fa" CHECKSUMTYPE="MD5"
                OWNERID="1.pdf">

```

```
<mets:FContent>
  <mets:binData>Data</mets:binData>
</mets:FContent>
</mets:file>
<mets:file ID="nlc003002" MIMETYPE="application/pdf"
SIZE="1552728" CHECKSUM="58a82a739a9a30d39e35b88532777676" CHECKSUMTYPE="MD5"
OWNERID="2.pdf">
  <mets:FContent>
    <mets:binData>Data</mets:binData>
  </mets:FContent>
</mets:file>
</mets:fileGrp>
</mets:fileSec>
<!--结构图-->
<mets:structMap>
  <mets:div TYPE="issue" DMDID="nlc001001" ADMID="nlc002001">
    <mets:div TYPE="文章">
      <mets:ptr FILEID="nlc003001"/>
    </mets:div>
    <mets:div TYPE="文章">
      <mets:ptr FILEID="nlc003002"/>
    </mets:div>
  </mets:div>
</mets:structMap>
</mets:mets>
```

参 考 文 献

- [1] Guidelines for using PREMIS with METS for exchange, Revised September 17, 2008 [OL]. [2012-02-29]. www.loc.gov/standards/premis/guidelines-premismets.pdf
- [2] VERMAATEN S. A Checklist for Documenting PREMIS-METS Decisions in METS Profiles [OL]. (May 2010) [2012-02-29].
http://www.loc.gov/standards/premis/premis_mets_checklist.pdf
- [3] APSR. Report of the METS Profile Development Project [R/OL]. [2012-02-29].
http://www.apsr.edu.au/nla-mets/mets_profile_report.pdf
- [4] DAITSS Overview [OL]. [2012-02-29].
http://daitss.fcla.edu/sites/daitss.fcla.edu/files/Chapter_1_DAITSS_Overview.pdf
- [5] The DAITSS Archiving Process [OL]. [2012-02-29].
http://daitss.fcla.edu/sites/daitss.fcla.edu/files/Chapter_5_Archiving.pdf