

METADATA REQUIREMENTS AND PREPARING CONTENT FOR DIGITAL PRESERVATION

v1.7.2

This document forms part of the Ministry of Education and Culture's
Open science and digital cultural heritage entity



Licence

Creative Commons Suomi CC-BY-NC-SA 4.0 (<http://creativecommons.org/licenses/by-nc-sa/4.0/>)

Users of this Specification are entitled to distribute the report, i.e. copy, circulate, display publicly and perform publicly the standard portfolio and modify it under the following conditions:

- *The Ministry of Education and Culture is appointed the Original Author (not, however, so that notification would refer to a licensee or means by which the Specification is used as supported by the licensor).*
- *The user is not entitled to use the Specification commercially.*
- *If the user makes any modifications to the Specification or uses it as the basis for their own works, the derivative work shall be distributed in the same manner or under the same type of licence.*

CONTENT

1	INTRODUCTION	4
1.1	Digital Preservation Services	4
1.2	Resource Description	4
1.3	Updates to this Specification.....	4
2	INFORMATION PACKAGE CONTENT DESCRIPTION	6
2.1	METS Profile	6
2.2	The Schema Catalog	8
2.3	Policies and Principles	8
2.4	Mandatory and Conditional Preservation Metadata	9
2.4.1	Identifiers	9
2.4.2	Timestamps.....	11
2.4.3	Descriptive Metadata Formats	12
2.4.4	Technical Metadata	12
2.4.5	Restrictions	14
2.4.6	Digital Provenance	15
2.4.7	Structural Metadata.....	16
3	TECHNICAL STRUCTURE OF A SUBMISSION INFORMATION PACKAGE	17
3.1	Content of Submission Information Package.....	17
3.2	Digital Signature	17
3.3	Metadata Formats and Their Versions.....	19
3.4	Content updates.....	19
4	TECHNICAL STRUCTURE OF A DISSEMINATION INFORMATION PACKAGE.....	21
ANNEX A.	NATIONAL METS PROFILES	22
A.1.	<mets> element	23
A.2.	<metsHdr> element	24
A.3.	<dmdSec> element	25
A.4.	<amdSec> element.....	26
A.5.	<techMD> element	27
A.6.	<rightsMD> element	28
A.7.	<sourceMD> element.....	29
A.8.	<digiprovMD> element	30
A.9.	<fileSec> and <fileGrp> elements	31
A.10.	<file> element	32
A.11.	<structMap> element.....	34
A.12.	<div> element.....	35
A.13.	<mdWrap> element	37
A.14.	<mdRef> element.....	38
REFERENCES	40	

This is an unofficial translation of the metadata requirements and preparing content for digital preservation specification. Should this translation conflict with the Finnish version, the Finnish version will have higher priority.

1 INTRODUCTION

This document specifies the metadata requirements of the Finnish national digital preservation services (DPS), including the structure of the submission and dissemination information packages and their METS profiles. This specification will be updated annually if necessary.

1.1 Digital Preservation Services

DPS refer to the national services produced for the digital preservation of cultural heritage resources and research data together. In this specification, a partner organization refers to an organization, department, or other entity using the DPS for digital preservation of digital content.

Digital preservation refers to the reliable preservation of digital information for several decades or even centuries. Hardware, software, and file formats will become outdated, but the information must remain understandable. Reliable digital preservation requires active monitoring of information integrity and anticipation of various risks. Metadata, which describes for example the resource itself, its provenance and rights related to it, has a key role in preservation.

The DPS produced for cultural heritage resources, guarantees the preservation of essential national information resources held in libraries, archives and museums. Digital cultural heritage resources cover both digitized and born-digital information resources: publications falling within the bounds of legal deposit, government publications belonging to the national cultural heritage, and other digital information resources worth preserving created by organizations operating under the Ministry of Education and Culture.

The DPS produced for research data ensures the availability and preservation of digital research resources. This DPS supports a permanent and coordinated approach to support the management of research resources. The aim is to ensure the verifiability and repeatability of research at various stages of the life cycle and to make the results easy to use. This enables that research results can be reused, evaluated, utilized in decision-making and secured by increasing digit data for future generations of researchers.

1.2 Resource Description

The specifications for describing the content and structure of documents ensure (see Chapter 2) that all the metadata needed for digital preservation is attached to the documents. It should be noted that digital resources cannot be transferred to the DPS as such, but as submission information packages (SIP) which also contain the necessary metadata elements. Only correct submission information packages can be transformed into archival information packages (AIP) from which dissemination information packages (DIP) are created when the content is transferred back to the partner organizations' own systems. SIP, AIP, and DIP are terms defined by the OAIS standard [OAIS]. The national DPSs are based on the OAIS reference model.

This document specifies the METS profiles for both cultural heritage resources and for research data. These are common to all SIPs and DIPs regardless of what, or what type of, digital information is being transferred. SIPs are transferred from a partner organization's system to the DPS using specifications defined in the Digime standard portfolio [SS]. This specification applies the METS standard for packaging the metadata and resources.

This specification defines for example what (meta)data the SIP must contain in addition to the content to be preserved (Chapter 3). Mandatory metadata ensures, for example, that the SIP has been transferred into the DPS unchanged, and enables the DPS to determine the sender of the package, and to check that the package meets the technical requirements.

1.3 Updates to this Specification

The Digime standard portfolio does not provide detailed instructions on the utilization of the METS standards, or any other standards, and such instructions must thus be created separately. These instructions must be sufficiently clear to allow the processes to be automated whenever possible. Then it should be possible to

automate for instance the creation of SIPs, and ingest of the packages in the DPS. Note that in the national model the data ownership remains in the submitting organization.

Relationships between the standard portfolio, METS profiles, and OAIS information packages is shown in Figure 1. The figure also outlines the operating model based on which the DPS-related definitions guide the operational activities. For example, when the standard portfolio is updated, there may be changes which must be taken into account in METS profiles. Partner organizations must be able to influence the content of the standard portfolio, and one of the responsibilities of the DPS is the maintenance of the METS profiles.

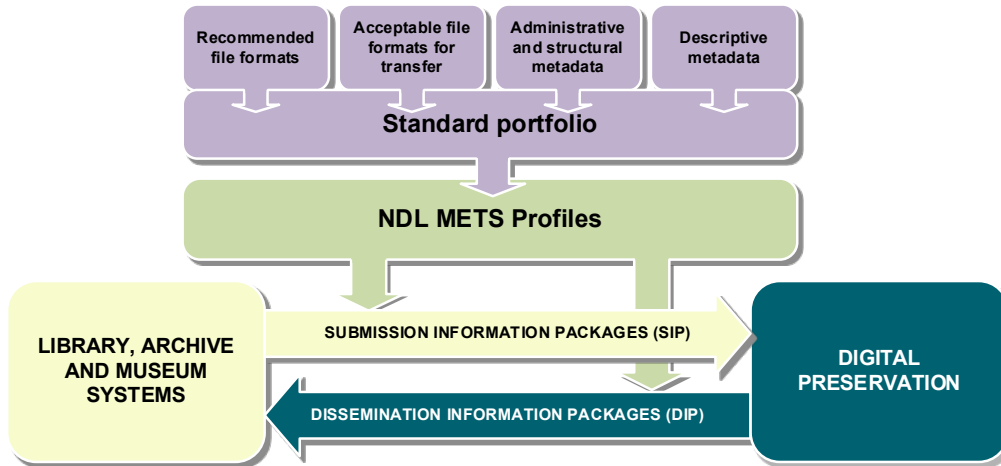


Figure 1: Relationship between the standard portfolio, METS profile, and information packages

The three-level version number of the definition describes the level of changes that have been made to the specification. If the first digit is updated, it means significant changes to partner organizations' systems that are integrated to the DPS. The middle digit reflects situations where changes may affect partner organizations' systems. An update to the third digit describes situations where only minor corrections or clarifications have been made. For example, updates to examples in the specification are minor corrections.

2 INFORMATION PACKAGE CONTENT DESCRIPTION

The submission information package (SIP) must contain a set of mandatory metadata elements which describe the content and structure of embedded resources. Only then it is possible to successfully transform the SIP into an archival information package (AIP) and store it in the DPS. SIP must conform to the national METS profile, and the file format of digital objects must conform to the file format specification [TDSTO]. The METS profiles are based on METS version 1.12.

The METS document consists of five parts (Figure 2): a document header, descriptive metadata, administrative metadata, structural metadata, and the digital resources to be preserved. The document header contains for instance the package creation time and the name of the organization that created the package. Descriptive metadata section includes a description of the resource in a metadata format included in the standard portfolio (for example, MARC21 [MARC21], Dublin Core [DC]). Administrative metadata section may include the technical metadata of the files, restrictions on use and access and the provenance information of the digital information. The structural metadata describes the structure of the resource. For instance, it can give the page order of a digitized book where every page is a separate file. The file section contains references to the actual objects to be preserved and possibly to their various manifestations.

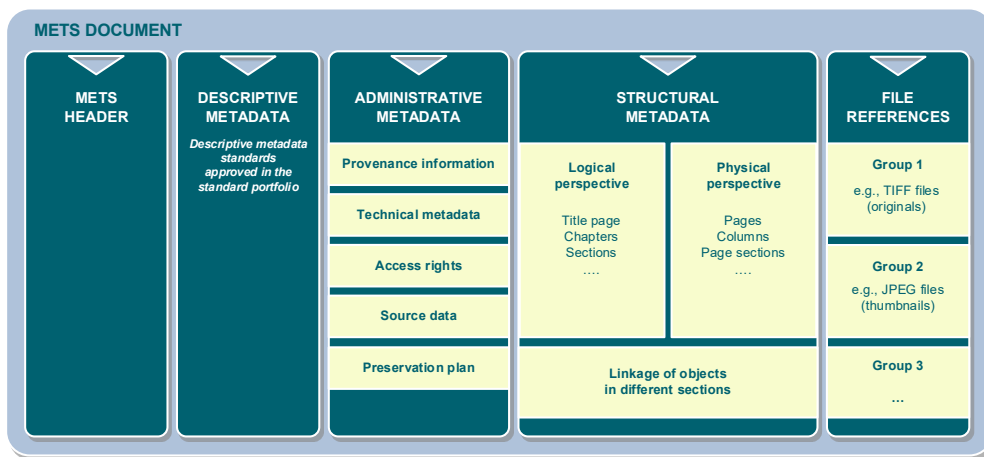


Figure 2: Overview of a METS document

2.1 METS Profile

A METS profile defines how METS schema are applied in national DPSs. The current METS profiles are generic profiles that are not restricted to a specific system or implementation. The profile describes for both the data provider and DPS the features needed for the construction and reception of a SIP and a DIP.

The METS profile defines the METS schema elements and attributes which are mandatory in SIPs and DIPs. The profile defines also optional and forbidden METS schema elements. If there is a forbidden element in a SIP, the ingest of the package fails.

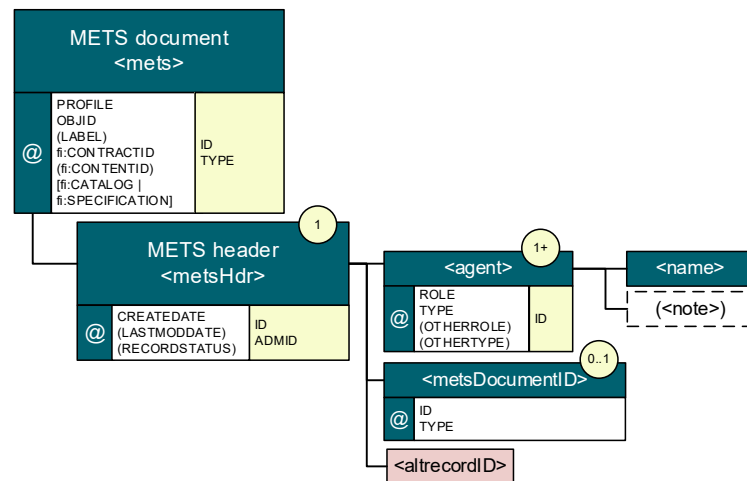


Figure 3: A part of a METS profile

Figure 3 depicts a part of a METS profile, which stipulates that the document root must contain values for the attributes PROFILE and OBJID. The LABEL attribute is optional. The fi:CATALOG attribute is a national attribute¹ that defines which schema catalog version is used when constructing the information package. If a schema catalog is not used, the fi:SPECIFICATION attribute must contain the version number of the national METS specification. The ID and TYPE attributes are not recommended and the DPS does not use them. Further, the METS document must contain one and only one METS header element. The header element must contain the creation date of the (CREATEDATE), and it may also contain values for the attributes LASTMODDATE and RECORDSTATUS. Other attributes are not recommended. The header element must contain at least one agent element, which contains the name of the agent (<name> element) and values for the ROLE and TYPE attributes. In addition, the header element may contain an identifier for the METS document (<metsDocumentID> element). The mandatory <agent> element in the METS profile contains the name of the organization that created the information package.

The METS standard uses several other standards, including descriptive and administrative metadata specifications. The preservation metadata standard PREMIS (Preservation Metadata: Implementation Strategies) is used to describe for instance the preservation metadata and provenance information of the resource. Therefore, PREMIS specification is one of the key aspects of the METS profiles. The profiles also rely on other administrative metadata formats; for technical metadata the project uses MIX [MIX] for image files, AudioMD [AUDIOMD] for audio files, and VideoMD [VIDEOMD] for video files.

This document specifies generic METS profiles for both cultural heritage resources and for research data (cf. table below)². In this specification, these profiles are identical, but may contain differences in future versions. The profiles are included in their entirety in Annex A, and it is possible to define content-specific subprofiles in the future. In these subprofiles, some metadata elements may be defined as mandatory even if these elements cannot be required for all contents.

Profile name ³	Description
cultural-heritage	This profile is used when preserving cultural heritage resources.
research-data	This profile is used when preserving research data.

¹ All attributes with a fi-prefix are national extensions to the METS standard

² Profile usage is defined in DPS the service contract

³ Profile name must always contain a string <http://digitalpreservation.fi/mets-profiles>.

For example, <http://digitalpreservation.fi/mets-profiles/research-data>

2.2 The Schema Catalog

METS documents conforming to the METS profile must be validated using the national schema catalog and Schematron rules. Catalogs, rules and instructions for using them are provided on the website⁴. When employing the catalog, the root element of the METS document may refer to the METS schema. For example:

```
<mets:mets OBJID="urn:oid:1.2.246.10.2159749.1.2.2012.1"
  PROFILE="http://digitalpreservation.fi/mets-profiles/cultural-heritage"
  fi:CATALOG="1.7.2"
  fi:CONTRACTID="id-310570"
  xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:fi="http://digitalpreservation.fi/schemas/mets/fi-extensions"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/
    http://digitalpreservation.fi/schemas/mets/mets.xsd">
```

Depending on the metadata formats used in the document, the root element may also contain other namespaces and schemas.

2.3 Policies and Principles

National METS profiles are drafted using the following high-level policies and principles.

Principle 1: There is one and only one method for describing any given feature in a METS document conforming to profile.

Justification: Allowing alternative solutions for providing e.g. administrative metadata complicates the validation of SIPs and creation of AIPs needlessly.

Example: A METS document must contain one and only one amdSec element, which may contain several parts.

Principle 2: Providing additional metadata⁵ is as a rule permitted, and such metadata is stored in the DPS, but the DPS supports neither their validation nor semantical/logical preservation. In such cases, the partner organization is fully responsible for semantical/logical preservation.

Justification: Additional metadata may increase the usability of the document.

Example: Digitized newspapers produced by the National Library of Finland consist of issues, articles and their component parts. Metadata about this structure is valuable and therefore the DPS must support its preservation. However, preservation of this exhaustive structural metadata is the responsibility of the National Library of Finland. These responsibilities will be defined in preservation plan.

Principle 3: Expanded metadata structures shall not disturb or needlessly complicate the validation and preservation of mandatory and recommended elements.

Justification: The main function of the DPS is to preserve intellectual content. Making the validation and the preservation activities more complex increases the risk failure, also for mandatory and recommended metadata elements.

Example: Adding descriptive metadata in the MODS format as part of a METS document is straightforward as such metadata records can be expressed as a separate part of the XML document. METS/ALTO differs from this since ALTO uses METS elements and the ALTO part and other parts of the METS document may contain cross-references.

⁴ <http://digitalpreservation.fi/schemas/>

⁵ Structures expanding metadata are metadata and/or metadata standards that are not specified. One example is the METS/ALTO standard, which is used to describe the layout and the structure of digitized content.

2.4 Mandatory and Conditional Preservation Metadata

National METS profiles define a set of metadata elements that are mandatory in SIPs. These elements are necessary for successful preservation. These are required for ensuring information integrity and semantical/logical preservation, and for making maintenance of information as easy as possible.

This section describes how metadata elements are expressed in the METS document using either the METS or the PREMIS⁶ formats (or descriptive metadata formats). When metadata can be expressed both in METS and in PREMIS formats, the latter **MUST** always be used, but METS may be applied as well [Vermaaten].

This section uses the following terms as specified by IETF (Internet Engineering Task Force) in [RFC 2119].

- **MUST**
- **MUST NOT**
- **REQUIRED**
- **SHOULD**
- **SHOULD NOT**
- **MAY**
- **OPTIONAL**

2.4.1 Identifiers

2.4.1.1 Package Identifier

The submission information package (SIP) must have a unique identifier created by the production system of a partner organization or by the application that created the package. The identifier makes it possible to identify the SIP.

The identifier does not need to be globally unique, only organization-level uniqueness is required. The identifier **MUST** be expressed in the root element of the METS document using the OBJID attribute.

Identifier System

May be decided by the partner organization.

Example

```
<mets:mets OBJID="fd2009-000001" ...>
  [...]
</mets:mets>
```

2.4.1.2 Content Identifier

The content identifier **MAY** be used to identify the content in an information package. By using the same content identifier in several SIPs, large content can be split into multiple SIPs, so that they form one entirety. This entirety can be located from the DPS using the content identifier.

The identifier **MUST** be expressed in the root element of the METS document using the fi:CONTENTID attribute.

Identifier System

May be decided by the partner organization.

Example

```
<mets:mets fi:CONTENTID="content-2009-000001" ...>
  [...]
</mets:mets>
```

⁶ In the following PREMIS element values, only printable US-ASCII characters are recommended: <premis:objectIdentifierType>, <premis:objectIdentifierValue>, <premis:eventIdentifierType>, <premis:eventIdentifierValue>, <premis:agentIdentifierType>, <premis:agentIdentifierValue>, <premis:rightsStatementIdentifierType> and <premis:rightsStatementIdentifierValue>.

2.4.1.3 Contract Identifier

The contract identifier is expressed in the root element of the METS document using the fi:CONTRACTID attribute. By identifying the contract, the information package content is associated with a particular DPS contract, no matter who ingests the information package.

Identifier System

To be agreed in the service contract

Example

```
<mets:mets fi:CONTRACTID="contract-xyz-000017" ...>
  [...]
</mets:mets>
```

2.4.1.4 File Identifier

The file identifier enables finding a particular file from a SIP or an AIP. Further, these identifiers make it easier to update the file in the partner organizations' production system after it has been migrated to a different file format in the DPS.

The file identifier **MUST** be expressed using <premis:objectIdentifier> element. File identifiers **SHOULD** be globally unique. Standardized identifiers should be used when it is possible (e.g. an ISBN for books).

Identifier System

May be decided by the partner organization.

Example

```
<mets:techMD ID="tech-001" CREATED="2017-12-31T13:41:19">
  <mets:mdWrap MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:object xsi:type="premis:file">
        <premis:objectIdentifier>
          <premis:objectIdentifierType>urn</premis:objectIdentifierType>
          <premis:objectIdentifierValue>
            URN:NBN:fi-fe201105101520
          </premis:objectIdentifierValue>
        </premis:objectIdentifier>
        [...]
      </premis:object>
    </mets:xmlData>
  </mets:mdWrap>
</mets:techMD>
```

2.4.1.5 Metadata Record Identifier

Partner organizations update their metadata (e.g. descriptive metadata, technical metadata) records in their production systems. Updated records **MUST** be submitted to the DPS. When the DPS updates metadata during e.g. migration process, any new or modified metadata **MUST** be transferred back to the production systems of partner organizations as DIPs.

Identifiers to metadata records **SHOULD** be encoded using the fi:PID attribute. If this attribute is used, the identifier system **MUST** be given using the fi:PIDTYPE attribute.

The partner organization may decide which identifier system to use.

Identifier System

May be decided by the partner organization.

Example

```
<techMD ID="internal-id" CREATED="2013-01-28T14:36:00"
  fi:PID="urn:nbn:fi-1234567890" fi:PIDTYPE="URN">
  [...]
</techMD>
```

2.4.2 Timestamps**2.4.2.1 Creation or Modification Time of the Information Package**

The creation time of a SIP or DIP MUST always be expressed in the package metadata. The creation time MUST be expressed in the package header (<mets:metsHdr>) using the CREATEDATE attribute with ISO-6801 format [ISO 8601-1]. If the information package has been modified (because it contained an error or has been updated, or for some other reason), the last modification time MUST be expressed using the LASTMODDATE attribute similarly as the creation time.

Formats

ISO-8601

Example

```
<mets:metsHdr CREATEDATE="2011-02-15T15:41:12">
  [...]
</mets:metsHdr>
```

2.4.2.2 Creation Time of Digital Objects and Metadata Records

A SIP MUST contain the creation and/or modification times of all digital objects and metadata records in the package. The time information MUST be expressed using either the ISO-8601 [ISO 8601-1] format, or its extended version [ISO 8601-2].

The creation time of a digital object MUST be expressed using PREMIS (<dateCreatedByApplication>) with the ISO-8601 [ISO 8601-1, ISO 8601-2]. This information MAY be expressed in the <mets:file> element using the CREATED attribute with a precision of a second. In the example below, the creation of another file is known exactly (15 February 2011 15:43:03) and the creation time of another file is uncertain, but it is estimated to have happened in the year 2000, probably in June.

The last modification time of metadata records in a SIP MUST be expressed in the corresponding metadata element using either the CREATED or fi:CREATED attribute. The CREATED attribute requires a precision of a second, but fi:CREATED also allows estimated times. In the example below, it is believed that the metadata was created in 2011, but this is uncertain. The creation time SHOULD also be expressed using the mechanism supported by the metadata format itself.

Formats

ISO-8601

Example

Creation time of a digital object:

```
<mets:techMD ID="tech-001" CREATED="2012-05-31T15:40:15">
  <mets:mdWrap MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:object xsi:type="premis:file">
        [...]
      <premis:objectCharacteristics>
        [...]
      <premis:creatingApplication>
        <premis:dateCreatedByApplication>
          2011-02-15T15:43:03
        </premis:dateCreatedByApplication>
      </premis:creatingApplication>
```

```

    </premis:objectCharacteristics>
  </premis:object>
</mets:xmlData>
</mets:mdWrap>
</mets:techMD>

```

Creation time of a descriptive metadata record:

```

<mets:dmdSec ID="dmd-marc" fi:CREATED="2011?">
  [...]
</mets:dmdSec>

```

2.4.3 Descriptive Metadata Formats

Descriptive metadata formats used in the information package **MUST** be specified in order to enable parsing of the embedded metadata.

If the organization uses its own metadata format, the information related to the format **MUST** be provided with the content. However, in such cases, the same descriptive metadata **MUST** also be provided using some metadata format specified in the Digime standard portfolio [SS].

The descriptive metadata format **MUST** be specified in the <mets:dmdSec> element using either the MDTYPE or OTHERMDTYPE attribute, and the version of the format **MUST** be expressed using the MDTYPEVERSION attribute. The descriptive metadata format and its version **MUST** be expressed using the controlled vocabulary given in Section 3.3.

Specifying the cataloguing rules used and the encoding level of the metadata record mitigates the processing and exploitation of the metadata in the information package. Cataloguing rules mean for example organizations' own instructions for resource description or international rules like RDA or AACR2.

The cataloguing rules and the encoding level **SHOULD** be expressed in the information package as specified by the selected metadata format. In the MARC21 format, both the cataloguing rules (leader character position 18, "Descriptive cataloguing form", which is 'a' in the example below) and the encoding level (leader character position 17, "#' in the example below) can be expressed. This information cannot be expressed for the time being in either Dublin Core or LIDO.

Formats

A metadata format specified in the standard portfolio.

The partner organization **MAY** decide which cataloguing rules and encoding level to use.

Example

```

<mets:dmdSec ID="dmd-marc" fi:CREATED="2011-05-31">
  <mets:mdWrap MIMETYPE="text/xml"
    MDTYPE="MARC" MDTYPEVERSION="marcxml=1.2; marc=marc21"
    LABEL="Bibliographic meta-data">
    <mets:xmlData>
      <MARC:record>
        <MARC:leader>00483nam#a2200157#a#4500</MARC:leader>
      </MARC:record>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>

```

2.4.4 Technical Metadata

2.4.4.1 File Format and its Version

The file format and its version **MUST** be included in the SIP. This is mainly due to the fact that tools for determining file formats are not perfect. They do not support all relevant file formats and may fail to identify common file formats if there are errors in the file header.

The file format in the SIP MUST be expressed using PREMIS. The <premis:formatName> element MUST contain only those strings specified in the controlled vocabulary of the “File Formats” specification [TDSTO]. The file format version MUST be expressed in the <premis:formatVersion> element if it is known and the <premis:formatName> element does not contain this information.

Formats

The File Formats specification [TDSTO] contains a controlled vocabulary of file formats and their versions.

Example

```
<mets:techMD ID="file01-techmd" CREATED="2011-05-31T09:54:43">
  <mets:mdWrap MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:object xsi:type="premis:file">
        <premis:objectIdentifier>
          [...]
        <premis:objectIdentifier>
          <premis:objectCharacteristics>
            <premis:compositionLevel>0</premis:compositionLevel>
            <premis:format>
              <premis:formatDesignation>
                <premis:formatName>image/png</premis:formatName>
                <premis:formatVersion>1.2</premis:formatVersion>
              </premis:formatDesignation>
              <premis:formatRegistry>
                <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
                <premis:formatRegistryKey>fmt/13</premis:formatRegistryKey>
              </premis:formatRegistry>
            </premis:format>
          </premis:objectCharacteristics>
        </premis:object>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:techMD>

  <mets:fileGrp>
    <mets:file ID="file01" ADMID="file01-techmd">
      [...]
    </mets:file>
  </mets:fileGrp>
```

2.4.4.2 Fixity Information and its Algorithm

The fixity information MUST be calculated for each digital object to be preserved as early as possible in order to ensure its integrity. The fixity information MUST be recorded in the digital object’s technical metadata in all information packages. When a SIP or DIP is submitted, the receiver can check the integrity of digital objects using the fixity information. In addition to the fixity information of the digital object, the fixity information of a SIP and a DIP must also be calculated, so that the receiver can confirm that the METS document has not been altered.

Fixity information MUST be calculated using one of the algorithms accepted. A list of accepted algorithms can be found from the standard portfolio. The partner organizations are responsible for ensuring that they use algorithms correctly. The fixity information MUST be included in the digital object’s PREMIS metadata (<premis:fixity>). Section 3.2 describes how the fixity information MUST be recorded in SIP or DIP.

Formats

The fixity algorithms specified in the standard portfolio.

Example

```

<mets:techMD ID="file01-techmd" fi:CREATED="2011-05-31">
  <mets:mdWrap MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:object xsi:type="premis:file">
        [...]
        <premis:objectCharacteristics>
          <premis:compositionLevel>0</premis:compositionLevel>
          <premis:fixity>
            <premis:messageDigestAlgorithm>
              MD5
            </premis:messageDigestAlgorithm>
            <premis:messageDigest>
              aa4bddaacf5ed1ca92b30826af257a1b
            </premis:messageDigest>
          </premis:fixity>
          [...]
        </premis:objectCharacteristics>
      </premis:object>
    </mets:xmlData>
  </mets:mdWrap>
</mets:techMD>

```

2.4.4.3 Technical Characteristics

The File Formats specification [TDSTO] specifies how file type -specific technical metadata is expressed in information packages.

2.4.5 Restrictions**2.4.5.1 Rights/Restrictions Metadata**

If the preserved content is subject to any access restrictions, those **MUST** be expressed in as much detail as possible in the information package in accordance with cataloguing rules and metadata formats used. Examples of such restrictions are the copyright status, usage restrictions based on legislation, a license, or other agreement concerning the content usage. Should the information package contain several digital objects with different restrictions, these restrictions should be attached correctly and individually to those objects that the restrictions concern.

The sector-specific cataloguing rules **MUST** be applied as specified in the standard portfolio. If there are no such rules in place, the method for describing the rights metadata **MUST** be decided in cooperation with the partner organization during the DPS deployment process.

2.4.5.2 Preservation Restrictions

The partner organization **MAY** impose restrictions concerning actions that are applied by the DPS during preservation. Such restrictions can apply to different kinds of preservation methods or actions.

In the future, these restrictions **SHOULD** be expressed as part of the preservation plan, which deprecates their inclusion in the SIP.

Any restrictions concerning preservation actions **SHOULD** be expressed using the PREMIS Rights metadata format in the METS <rightsMD> element.

Formats

PREMIS Rights

2.4.6 Digital Provenance

The digital provenance of a digitized resource is metadata on matters such as which hardware and software was used when the resource was digitized, and which application has been used to manipulate the material. For born-digital content digital provenance describes for example the software that was used to create the content.

The digital provenance **MUST** be included in the information package, because the authenticity of the resource depends partially on the provenance information. Without it, the event history of the resource is known only after it has been ingested into the DPS. Digital provenance provides information about all events which may have had an impact on the intellectual content or look and feel of the resource, including migrations.

The digital provenance information **MUST** be expressed using the <mets:digiprovMD> element in which the provenance metadata **MUST** be encoded using PREMIS events (<premis:event>). Recommended events for provenance information are given on the digitalpreservation.fi website⁷. These event types **MUST** be used where applicable. If the provenance information contains events that are not included in the controlled vocabulary, other event types can be used.

Unknown events in the digital provenance **MAY** be expressed using PREMIS with the event type "creation" and the result (<premis:eventOutcome>) with an appropriate unknown value⁸.

If a PREMIS event does not refer to any particular object, the event concerns the whole information package.

Formats

PREMIS event, PREMIS agent

Example

```
<mets:digiprovMD ID="ev001" fi:CREATED="2011-05-31">
  <mets:mdWrap MDTYPE="PREMIS:EVENT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:event>
        <premis:eventIdentifier>
          <premis:eventIdentifierType>local</premis:eventIdentifierType>
          <premis:eventIdentifierValue>
            pdfGeneration-001
          </premis:eventIdentifierValue>
        </premis:eventIdentifier>
        <premis:eventType>migration</premis:eventType>
        <premis:eventDateTime>2011-03-15T11:12:13</premis:eventDateTime>
        <premis:eventOutcomeInformation>
          <premis:eventOutcome>success</premis:eventOutcome>
        </premis:eventOutcomeInformation>
        <premis:linkingAgentIdentifier>
          <premis:linkingAgentIdentifierType>
            local
          </premis:linkingAgentIdentifierType >
          <premis:linkingAgentIdentifierValue>
            pdfGenerator-1
          </premis:linkingAgentIdentifierValue >
        </premis:linkingAgentIdentifier>
      </premis:event>
    </mets:xmlData>
  </mets:mdWrap>
</mets:digiprovMD>
<mets:digiprovMD ID="ag001" fi:CREATED="2011-05-31">
  <mets:mdWrap MDTYPE="PREMIS:AGENT" MDTYPEVERSION="2.2">
    <mets:xmlData>
      <premis:agent>
```

⁷ <http://digitalpreservation.fi/specifications/vocabularies/>

⁸ <http://digitalpreservation.fi/specifications/vocabularies/unknown-values>

```

    <premis:agentIdentifier>
      <premis:agentIdentifierType>local</premis:agentIdentifierType>
      <premis:agentIdentifierValue>
        pdfGenerator-1
      </premis:agentIdentifierValue>
    </premis:agentIdentifier>
    <premis:agentName>Adobe Acrobat</premis:agentName>
    <premis:agentType>software</premis:agentType>
  </premis:agent>
</mets:xmlData>
</mets:mdWrap>
</mets:digiprovMD>

```

2.4.7 Structural Metadata

The structural metadata is mandatory information according to the METS specification (structMap). Even if a SIP or a DIP contains only one digital object, the structural metadata is mandatory. In such cases, the structural metadata contains information that the package contains only one digital object.

If a SIP or a DIP contains several digital objects, the structural metadata contains information about the order in which they should be presented (e.g. the page order of a digitized book consisting of a single file per each page). This ensures that the content may be rendered correctly without analyzing each file. The structural metadata can also be used to separate different manifestations of the same resource within the same information package. Further, the structural metadata attach the descriptive metadata to digital objects or to the package structure. Examples of structural metadata can be found on the digitalpreservation.fi website⁹.

Formats

METS structMap

Example

```

<mets:structMap TYPE="physical">
  <mets:div TYPE="book" LABEL="Kirjan nimi">
    <mets:div TYPE="page" LABEL="Ensimmäinen sivu" />
    <mets:div TYPE="page" LABEL="Toinen sivu" />
    [...]
    <mets:div TYPE="page" LABEL="Viimeinen sivu" />
  </mets:div>
</mets:structMap>

```

⁹ <http://digitalpreservation.fi/specifications/examples>

3 TECHNICAL STRUCTURE OF A SUBMISSION INFORMATION PACKAGE

The partner organization assembles a submission information package (SIP) of digital objects and associated metadata, signs the SIP digitally, and packages it. The DPS uses the same structure in a dissemination information packages (DIP).

3.1 Content of Submission Information Package

There MUST be a “mets.xml” file with Unicode UTF-8 encoding located in the root of a SIP, containing information on the structure of the SIP and the descriptive and administrative metadata needed. In addition, there must be a digital signature in a file named “signature.sig” located in the root of the SIP (see Figure 4). Any additional files, symbolic links or empty directories are not allowed in SIPs. A SIP MUST contain all the metadata needed for interpretation of the package and all the digital objects that belong to this package. All the files must be described in XML formatted metadata in the mets.xml file. The XML document must be compliant with the given METS profile.

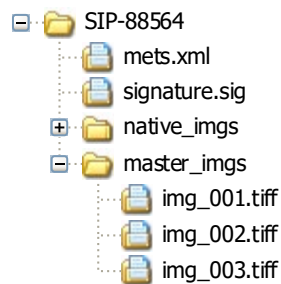


Figure 4: Example structure of a SIP

The partner organization MUST either compress the SIP package into a ZIP file or create a TAR file out of it, before the package is sent to the DPS. The ZIP file / TAR file must contain only one SIP at the root of the file.

Checking the integrity of the SIP is done by algorithms recommended by the Digime standard portfolio [SS]. A checksum MUST be calculated for every file by using some of the standard portfolio’s message digest algorithms. The partner organization is responsible for checksum calculations and adding them to the SIP’s metadata.

The integrity of a SIP is checked in the DPS by comparing the checksums calculated during the ingest process to the checksums in the METS document. If there are mismatches, the ingest process stops. Service contract and/or DPS’s common practices define the next steps. Either just the problematic content is transferred separately into the DPS, or the whole SIP is rebuilt and re-submitted.

3.2 Digital Signature

Data integrity and nonrepudiation are verified by a digital signature. With the signature the data provider can be identified and it is possible to be certain that the data has not changed during the transfer between a partner organization and the DPS. Public and private keys are used in the signature.

The partner organization signs the mets.xml file digitally using a PKCS#7 signature. The signature is formed as follows:

- A checksum for mets.xml is calculated by using one of the algorithms specified in the standard portfolio. Checksums for preserved objects are already present in the METS document so there is no need to calculate them while forming the signature.
- The checksum is added to a text file by writing the following into it: checked file’s (mets.xml) path relative to the submission information package’s root, algorithm used (string md5, sha1, sha224, sha384, sha512) and checksum. These data elements are separated by semicolons.

- File is signed with the PKCS#7 signature using a certificate.
- File is saved into the SIP's root by using the name "signature.sig" in S/MIME format

The DPS verifies the signee using the signer's public key and the "signature.sig" file, thereby ensuring that the integrity of the METS document has been maintained during the transfer into the DPS and the file is signed by the partner organization in question.

Figure 5 depicts an example of the "signature.sig" file. In this example, the checksum is calculated by using the SHA-1 algorithm that is signed by the PKC37 signature.

```
MIME-Version: 1.0
Content-Type: multipart/signed; protocol="application/x-pkcs7-signature";
micalg=shal; boundary="-----57E5EFE5F87ADB48166F35F180BE72AC"

This is an S/MIME signed message

-----57E5EFE5F87ADB48166F35F180BE72AC
Content-Type: text/plain

./mets.xml:sha1:effdb5f96a28acd2eb19dcb15d8f43af762bd0ae

-----57E5EFE5F87ADB48166F35F180BE72AC
Content-Type: application/x-pkcs7-signature; name="smime.p7s"
Content-Transfer-Encoding: base64
Content-Disposition: attachment; filename="smime.p7s"

MIIFKwYJKoZIhvcNAQcCoIIFHDCCBRgCAQExCzAJBgUrDgMCGgUAMAsGCSqGSIb3
...
UY/I9QDibkW5qAUY00rN7Q1F+uAGB+twEg6un1SVdg==

-----57E5EFE5F87ADB48166F35F180BE72AC--
```

Figure 5: Example 'signature.sig'

3.3 Metadata Formats and Their Versions

Metadata formats and their versions supported by this specification are listed in the table below.

Metadata format	Supported versions
ADDML	8.3, 8.2
AudioMD	2.0
DataCite	4.3, 4.2, 4.1
DDI (Data Documentation Initiative) - Codebook	2.5.1, 2.5, 2.1
DDI (Data Documentation Initiative) - Lifecycle	3.2, 3.1
Dublin Core	1.1, 2008
EAC-CPF (Encoded Archival Context for Corporate Bodies, Persons, and Families) ¹⁰	2010 revised
EAD (Encoded Archival Description)	2002
EAD3 (Encoded Archival Description)	1.1.0, 1.0.0
EN 15744 (Film identification – Minimum set of metadata for cinematographic works)	
FINMARC	marcxml=1.2; marc=finmarc
LIDO (Lightweight Information Describing Objects)	1.0
MARC 21 (MACHINE-Readable Cataloguing)	marcxml=1.2; marc=marc21
MIX	2.0
MODS (Metadata Object Description Schema)	3.7, 3.6, 3.5, 3.4, 3.3, 3.2, 3.1, 3.0
PREMIS	2.3, 2.2
VideoMD	2.0
VRA Core (Visual Resources Association Core Categories)	4.0 (restricted)

3.4 Content updates

Updating content means submitting a SIP that updates preserved content which has been transferred earlier to the DPS. Such a SIP can be incremental, that is, those digital objects that are not changed in the update do not have to be in the SIP. The METS document and the digital signature are still required. If only the metadata is updated, it is enough to submit a new complete METS document and the digital signature. Otherwise the SIP must conform to this specification with the following exceptions:

Attribute	Occurrence/obligation	Rules and recommendations
<mets>	ET, P	
OBJID	ET, P	The identifier of the SIP for update must be the same as in the SIP to be updated.
<metsHdr>	ET, P	

¹⁰ 2014 accepted edition of the 2010 version

Attribute	Occurrence/ obligation	Rules and recommendations
CREATEDATE	ET, P	The creation time of the original SIP MUST be expressed with a precision of a second using the ISO 8601 specification [ISO 8601-1].
LASTMODDATE	ET, P	The creation time of the update SIP must be given with a precision of a second using ISO 8601 specification.
RECORDSTATUS	ET, P	The value of the attribute must be 'update'

4 TECHNICAL STRUCTURE OF A DISSEMINATION INFORMATION PACKAGE

The structure of a dissemination information package (DIP) follows the METS profile and the technical structure of a submission information package (SIP) defined in this specification with the exception that a DIP does not have to contain the preserved content (e.g. the DIP may contain only the metadata).

A dissemination information package is created from the preserved data and its metadata. Each DIP is unique, even if it is created from the same data as a previously created DIP. Each DIP has a unique @OBJID identifier in the METS document.

The following PREMIS events created by the DPS will be added to the METS document¹¹ of the DIP:

- Ingestion events for the metadata and the content
- The moment when the preservation responsibility changed
- All preservation actions
- Event for DIP creation

In conjunction with events, it is possible to add new PREMIS objects and agents to the METS document. A new structure map is created for the document in which events are compiled together.

The value of the RECORDSTATUS attribute in a DIP is always “disseminated” and the DIP creator (agent) is the DPS.

Attributes which are either forbidden or not recommended in this specification MUST not be used in the METS document of a DIP. The DPS may use in the METS document the values it has chosen for internal identifiers and references (e.g. attribute values for ID, ADMID, and DMDID).

The fixity information of a DIP is calculated and the DPS signs the DIP in the same way as the partner organization signs the SIP (cf. Section 3.2).

¹¹ Example of DPS-created events can be found at <http://digitalpreservation.fi/specifications/examples>

ANNEX A. NATIONAL METS PROFILES

The national METS profiles define how to use METS schema in conjunction with the DPS. The profile defines, for example, what METS schema elements and attributes must be used, and how. In addition, the profile defines forbidden and optional METS schema elements. The figure below presents notation used for the graphical presentation of the national METS profiles.

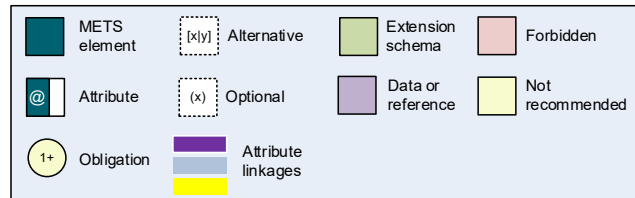


Figure 6: Notation used for the graphical presentation of the METS profile

If an attribute is defined as alternative, it is mandatory, but it may be replaced by another attribute in certain conditions. If an element or attribute is "forbidden", its presence in a SIP causes failure in the ingest process.

The following chapters describe in detail the use of METS elements and attributes. The occurrence and obligations of the elements and attributes are abbreviated in the tables as follows:

- T Repeatable
- ET Non repeatable
- P Mandatory
- PV Mandatory, but may be replaced by another attribute in certain conditions
- V Optional (accepted in a SIP and is saved into DPS)
- ES Not recommended (accepted in a SIP, but not utilized the DPS). Not supported in DIPs.

Attribute links are used to describe how to link the METS document elements (see Figure 7). When using attribute links you must keep in mind that accepted values for attributes use the NCName format and thus for example semicolons are not accepted. The DPS preserves linking, but not necessarily the values used in the linking.

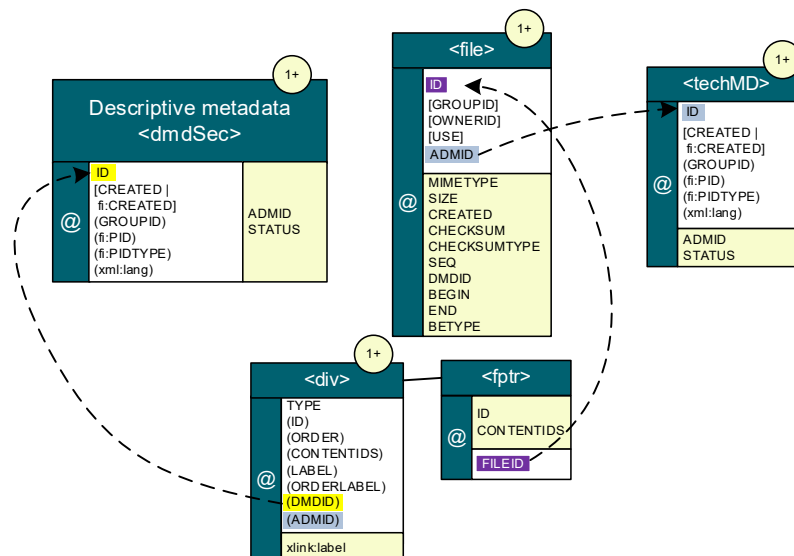


Figure 7: Attribute linkage (example)

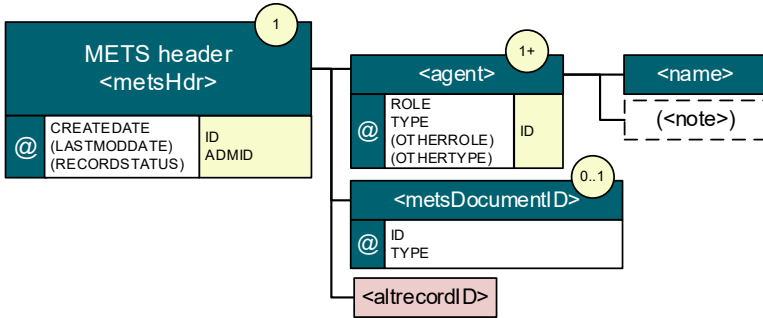
A.1. <mets> element

METS document <mets>	
@	PROFILE OBJID (LABEL) fi:CONTRACTID (fi:CONTENTID) [fi:CATALOG fi:SPECIFICATION]
	ID TYPE

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<mets>	ET, P	Root element must contain one and only one <metsHdr/>-element, at least one <dmdSec/>-element, one and only one <amdSec/>-element, one and only one <fileSec/>-element and at least one <structMap/>-element. In addition to that element must contain obligatory elements listed below.
PROFILE	ET, P	Attribute value must be the METS profile defined in this specification or some of its subprofiles. For cultural heritage resources the attribute value must be http://digitalpreservation.fi/mets-profiles/cultural-heritage For cultural heritage resources the attribute value must be http://digitalpreservation.fi/mets-profiles/research-data
OBJID	ET, P	Organization's unique identifier for the information package generated by packaging application. Attribute value should be expressed in printable US-ASCII characters. For SIP for update the attribute is defined in section 3.4.
fi:CONTRACTID	ET, P	Contract identifier of a DPS contract to which the package content belongs. Attribute value should be expressed in printable US-ASCII characters.
fi:CONTENTID	ET, V	Identifier for the content in the package. Attribute value should be expressed in printable US-ASCII characters.
LABEL	ET, V	Short description of the information package.
fi:CATALOG	ET, PV	Version number of the schema catalog used when data package is created. If there is no catalog element present, it has to be replaced by fi:SPECIFICATION attribute.
fi:SPECIFICATION	ET, PV	Version number of packaging specification used in creation of data package. Mandatory only when the use of the fi:CATALOG attribute is not possible.
ID	ET, ES	ID attribute is not recommended because there is no need to refer to <mets> element in a profile compliant document. The attribute is not supported in DIPs.
TYPE	ET, ES	This attribute can be in the submission information package but the DPS does not use it. The attribute is not supported in DIPs.
-<metsHdr>	ET, P	See section A.2
-<dmdSec>	T, P	See section A.3
-<amdSec>	ET, P	See section A.4
-<fileSec>	ET, V	See section A.9
-<structMap>	T, P	See section A.11
-<structLink>	-	The <structLink> element is forbidden in national METS profiles.

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
-<behaviourSec>	-	The <behaviourSec> element is forbidden in national METS profiles.

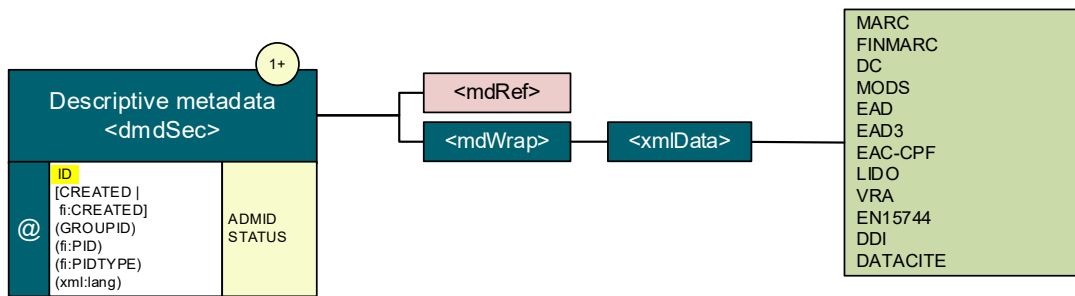
A.2. <metsHdr> element



Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<metsHdr>	ET, P	Mandatory element. It must contain the CREATEDATE attribute, which specifies the time when the information package was created.
CREATEDATE	ET, P	Mandatory attribute. The package creation time must be recorded in ISO 8601 format [ISO 8601-1] with a resolution of one second.. For SIP for update the attribute is defined in section 3.4.
LASTMODDATE	ET, V	If the package has been modified since the initial creation, the modification time must be expressed with LASTMODDATE using the same syntax and resolution as CREATEDATE. See in section 3.4 for SIPs which are updates.
RECORDSTATUS	ET, V	If the attribute is not present or its value is “submission”, the information package is a new SIP. If the package identifier is the same as in some other information package ingested earlier belonging to the same contract, the package will be rejected. The value “update” means the SIP is an updated version of a previous SIP. See section 3.4. If the package identifier is not found from the DPS, the package will be rejected. The value “dissemination” means the information package is a DIP. See section 4.
ID, ADMID	ET, ES	ID and ADMID attributes are not recommended because in a profile compliant information package there is no need to refer to <metsHdr> element or refer from <metsHdr> element to administrative metadata. Attributes are not supported in DIPs.
-<agent>	T, P	It is mandatory to have at least one <agent> element.
ROLE	ET, P	At least one <agent> element must have the attribute value CREATOR.
TYPE	ET, P	A mandatory <agent> element expresses who or what has created the information package. The value of the attribute may be for example ORGANIZATION, when the <name> element must contain the name of the organization which created the information package.

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
ID	ET, ES	ID attribute is not recommended because in a profile compliant document there is no need to refer to <agent> element. The attribute is not used in DIPs.
OTHERROLE, OTHERTYPE	ET, V	These attributes may be used, if it is not possible to use ROLE or TYPE-attributes to the described agent.
--<name>	ET, P	At least one <agent> element must have a name element containing the name of the package’s creator organization or software.
--<note>	T, V	The DPS does not use <note> during ingest or preservation, but the element’s value is stored in AIP.
-<metsDocumentID>	ET, V	Element includes the identifier of the METS document.
ID	ET, P	METS document identifier. Note that this is not the same as the package ID (attribute OBJID in the element <mets>).
TYPE	ET, P	METS document identifier type.
-<altrecordID>	-	A national METS profile compliant package must not contain <altrecordID> element.

A.3. <dmdSec> element



Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<dmdSec>	T, P	Element must contain <mdWrap><xmlData> element. It contains descriptive metadata in one of the formats specified in the standard portfolio or in some other metadata format. The metadata format and its version must be defined in <mdWrap> element’s MDTYPE and MDTYPEVERSION attributes. If MDTYPE value is OTHER, metadata format must be defined in OTHERMDTYPE attribute value. <dmdSec> element may be repeated because the information package may contain diverse resources which require different kinds of descriptive metadata.
ID	ET, P	Mandatory in the METS schema. Attribute has to contain a value that is unique in the METS document. <structMap><div> element refers to this ID.
GROUPID	ET, V	Attribute is grouping semantically similar metadata. For example, if object’s descriptive metadata are in different metadata formats the metadata records must share the same value of GROUPID attribute. Attribute value should be expressed in printable US-ASCII characters.
fi:PID	ET, V	Unique identifier of a metadata record. Attribute value should be expressed in printable US-ASCII characters.

Element/Attribute	Occurrence/obligation	Rules and recommendations
fi:PIDTYPE	ET, V	Identifier system used in the PID attribute. Attribute is mandatory, if the PID attribute is used. Attribute value should be expressed in printable US-ASCII characters.
CREATED	ET, PV	The precise time the metadata record was created, in the ISO 8601 format with a resolution of one second [ISO 8601-1]. The fi:CREATED attribute is an alternative for this. It is prohibited to use both attributes in the same time.
fi:CREATED	ET, PV	The approximate time the metadata record was created in the extended ISO 8601 format [ISO 8601-2]. CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
ADMID, STATUS	ET, ES	Attributes are allowed in SIPs but the DPS does not use or preserve them. They are not available in DIPs.
xml:lang	ET, V	Language of descriptive metadata as specified by xml:lang attribute.
-<mdWrap>	ET, P	The actual metadata is present in <mdWrap> element (See Chapter A.13)
-<mdRef>	-	<mdRef> element is not allowed in a national METS profile compliant package.

A.4. <amdSec> element

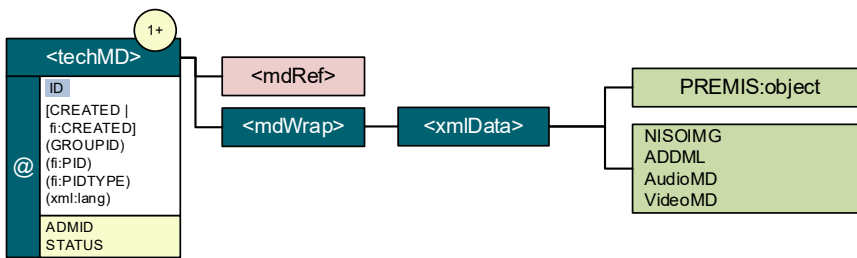
1

Administrative metadata
<amdSec>

@ ID

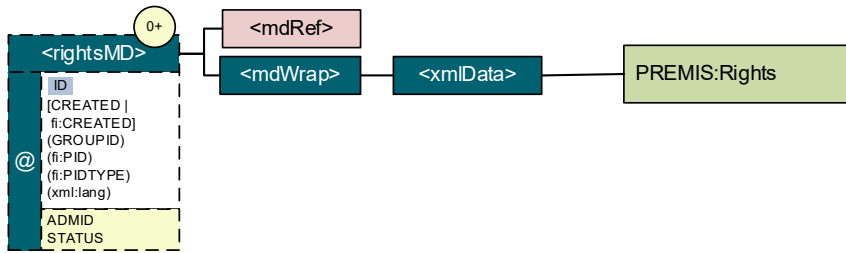
Element/Attribute	Occurrence/obligation	Rules and recommendations
<amdSec>	ET, P	A national profile compliant METS document contains one and only one <amdSec> element. It must contain one <techMD> element and at least two <digiprovMD> elements. Element may also contain <rightsMD> and <sourceMD> elements.
ID	ET, ES	Not recommended. Referring to administrative metadata can be done by using ID attributes of <techMD>, <rightsMD>, <sourceMD> and <digiprovMD> elements. The attribute is not preserved and will not be present in DIPs.
-<techMD>	T, P	See section A.5
-<rightsMD>	T, P	See section A.6
-<sourceMD>	T, P	See section A.7
-<digiprovMD>	T, P	See section A.8

A.5. <techMD> element



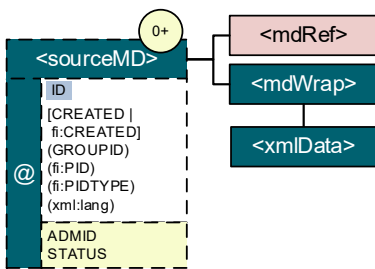
Element/Attribute	Occurrence/obligation	Rules and recommendations
<techMD>	T, P	Element must contain <mdWrap><xmlData> element, which may contain technical or preservation metadata in the formats endorsed (PREMIS Object, MIX (NISOIMG), AudioMD or VideoMD or ADDML metadata). For PREMIS Object metadata, the metadata elements specified in sections 2.4.1.4, 2.4.4.1, and 0 must be present. <techMD> element may be repeated, and in addition to metadata formats endorsed, other technical metadata formats may be used as well.
ID	ET, P	Mandatory in the METS schema. Attribute must contain a value that is unique in the METS document. There must be a reference to ID in <file> or <div> elements.
GROUPID	ET, V	Attribute may be present in SIP, if partner organization needs it. The DPS will not use the attribute. Attribute value should be expressed in printable US-ASCII characters.
CREATED	ET, PV	The precise time the metadata record was created, in the ISO 8601 format with a resolution of one second [ISO 8601-1]. The fi:CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:CREATED	ET, PV	The approximate time the metadata record was created, in the extended ISO 8601 format [ISO 8601-2]. CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:PID	ET, V	Unique identifier of metadata. Attribute value should be expressed in printable US-ASCII characters.
fi:PIDTYPE	ET, V	Identifier system used in the PID attribute. Attribute is mandatory, if the PID attribute is used. Attribute value should be expressed in printable US-ASCII characters.
ADMID, STATUS	ET, ES	Attributes are allowed in SIPs but the DPS does not use or preserve them. They are not supported in DIPs.
xml:lang	ET, V	Language of technical metadata as specified by xml:lang attribute.
-<mdWrap>	ET, P	The actual metadata is saved in <mdWrap> element (See Chapter A.13)
-<mdRef>	-	<mdRef>-element is not allowed in a national METS profile compliant document.

A.6. <rightsMD> element



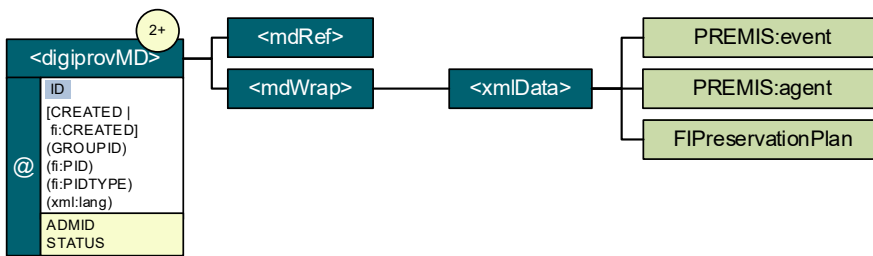
Element/Attribute	Occurrence/obligation	Rules and recommendations
<rightsMD>	T, V	Element must contain <mdWrap><xmlData> element that contains PREMIS:Rights type metadata.
ID	ET, P	Attribute is mandatory in METS schema. It must have a value that is unique in the METS document. ID must be referred to in <file> or <div> elements.
GROUPID	ET, V	Attribute is allowed in SIPs, if the partner organization needs it. The DPS does not use it. Attribute value should be expressed in printable US-ASCII characters.
CREATED	ET, PV	The precise time the metadata record was created, in the ISO 8601 format with a resolution of one second [ISO 8601-1]. The fi:CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:CREATED	ET, PV	The approximate time the metadata record was created, in the extended ISO 8601 format [ISO 8601-2]. CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:PID	ET, V	Unique identifier of metadata. Attribute value should be expressed in printable US-ASCII characters.
fi:PIDTYPE	ET, V	Identifier system used in the PID attribute. Attribute is mandatory, if the PID attribute is used. Attribute value should be expressed in printable US-ASCII characters.
ADMID, STATUS	ET, ES	Attributes are allowed in SIPs but the DPS does not use or preserve them. They are not supported in DIPs.
xml:lang	ET, V	Language of use restrictions as specified by xml:lang attribute.
-<mdWrap>	ET, P	Actual metadata is saved in <mdWrap> element (See Chapter A.13)
-<mdRef>	-	<mdRef>-element is not allowed in a national METS profile compliant document.

A.7. <sourceMD> element



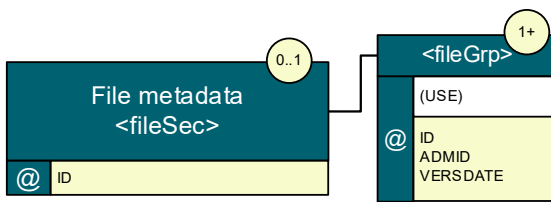
Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<sourceMD>	T, V	For the source metadata, it is recommended to use metadata standards defined in the standard portfolio.
ID	ET, P	Mandatory in the METS schema. The attribute must have a value that is unique in a METS document. ID must be referred to in <file> or <div> elements.
GROUPID	ET, V	Attribute is allowed in SIP, if the partner organization needs it. The DPS is not using it. Attribute value should be expressed in printable US-ASCII characters.
CREATED	ET, PV	The precise time the source metadata record was created, in the ISO 8601 format with a resolution of one second [ISO 8601-1]. The fi:CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:CREATED	ET, PV	The approximate time the source metadata record was created, in the extended ISO 8601 format [ISO 8601-2]. CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:PID	ET, V	Unique identifier of the metadata record. Attribute value should be expressed in printable US-ASCII characters.
fi:PIDTYPE	ET, V	Identifier system used in the fi:PID attribute. Attribute is mandatory, if an identifier is provided. Attribute value should be expressed in printable US-ASCII characters.
ADMID, STATUS	ET, ES	Attributes are allowed in SIPs but the DPS does not use or preserve them. They are not supported in DIPs.
xml:lang	ET, V	Source language as specified by xml:lang attribute.
-<mdWrap>	ET, P	Actual metadata is saved in <mdWrap> element (See Chapter A.13)
-<mdRef>	-	<mdRef>-element is not allowed in a national METS profile compliant document.

A.8. <digiprovMD> element



Element/Attribute	Occurrence/obligation	Rules and recommendations
<digiprovMD>	T, P	Mandatory element because SIP and DIP must contain data's provenance and preservation plan.
ID	ET, P	Mandatory in the METS schema. Attribute must have a value that is unique in a METS document. ID must be referred to in <file> or <div> elements.
GROUPID	ET, V	Attribute is allowed in SIPs, if the partner organization needs it. The DPS is not using it. Attribute value should be expressed in printable US-ASCII characters.
CREATED	ET, PV	The precise time the digital provenance metadata record or preservation plan was created, in the ISO 8601 format with a resolution of one second [ISO 8601-1]. The fi:CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously.
fi:CREATED	ET, PV	The approximate time the provenance metadata record was created, in the extended ISO 8601 format [ISO 8601-2]. The CREATED attribute is an alternative for this. It is prohibited to use both attributes simultaneously. It is forbidden to use this attribute for the creation time of the preservation plan.
fi:PID	ET, V	Unique identifier of the metadata record. Attribute value should be expressed in printable US-ASCII characters.
fi:PIDTYPE	ET, V	Identifier system used in the fi:PID attribute. Attribute is mandatory, if an identifier is provided. Attribute value should be expressed in printable US-ASCII characters.
ADMID, STATUS	ET, ES	Attributes are allowed in SIPs but the DPS does not use or preserve them. They are not supported in DIPs.
xml:lang	ET, V	Language of provenance and preservation plan as specified by xml:lang attribute.
-<mdWrap>	ET, P	The actual metadata is saved in <mdWrap> element (See Chapter A.13) either in PREMIS:event or PREMIS:agent type.
-<mdRef>	ET, V	<mdRef> element usage is allowed only when referring to existent preservation plan (See chapter 0).

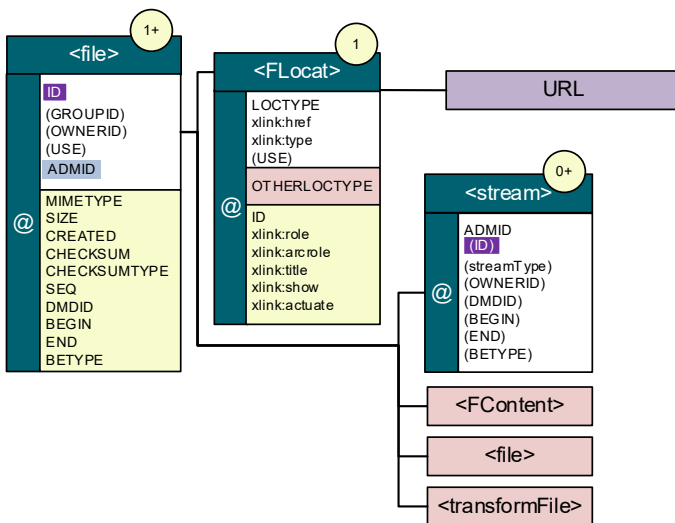
A.9. <fileSec> and <fileGrp> elements



Element/Attribute	Occurrence/obligation	Rules and recommendations
<fileSec>	ET, P	<fileSec> is mandatory element in SIPs, because the SIP has to contain the preserved object(s). In DIPs <fileSec>-element is not mandatory, if the package does not contain preserved data.
ID	ET, ES	Not recommended, because there is no need to refer to <fileSec> element on a profile compliant document. The attribute is not used in DIPs.
<fileGrp>	T, P	There may be one or several <fileGrp> elements on a profile compliant document, if there is a <fileSec> element. <fileGrp> element contains one or several <file> elements.
ID	ET, ES	Not recommended, because there is no need to refer to the <fileGrp> element in a profile compliant document. The attribute is not supported in DIPs.
ADMID	ET, ES	Not recommended. Administrative metadata is referred in <file> element. The attribute is not supported in DIPs.
USE	ET, V	USE attribute defines usage of <fileGrp>. The recommended controlled vocabulary for attribute ¹² .
VERSDATE	ET, ES	Attribute is allowed in SIPs, but the DPS does not use it or preserve it. It is not supported in DIPs.
--<file>	T, P	See section A.10
--<fileGrp>	-	Recursive use is forbidden

¹² <http://digitalpreservation.fi/specifications/vocabularies>

A.10. <file> element



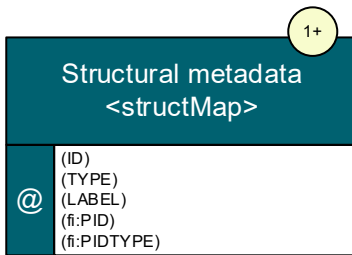
Element/Attribute	Occurrence/obligation	Rules and recommendations
<file>	T, P	A national profile compliant METS document contains one or several <file> elements, if there is <fileSec> element in the document.
ID	ET, P	ID attribute is mandatory in order to enable file references in the <structMap> element.
MIMETYPE, CHECKSUM, CHECKSUMTYPE, OWNERID, CREATED, SIZE, BEGIN, END, BETYPE	ET, ES	Attribute MIMETYPE may be used for file format. File format must be defined in <techMD> element using PREMIS. <CHECKSUM> and >CHECKSUMTYPE> attributes may be used for checksum and algorithm. These must be in <techMD> element using PREMIS. CREATED attribute may be used for file creation time, but it has to be in <techMD> element using PREMIS. OWNERID attribute may be used to express file owner’s organizational ID. Attribute value should be expressed in printable US-ASCII characters. The DPS uses only file format, checksum/algorithm, creation time and ID described in PREMIS format. These attributes are neither preserved nor supported in DIPs.
USE	ET, ES	USE-attribute must be used in <fileGrp>-element. If the attribute is repeated here, it is neither preserved nor supported in DIPs.
ADMID	ET, P	ADMID-attribute value(s) are referring to file’s administrative metadata.
SEQ	ET, ES	Attribute is allowed in SIPs, but the DPS is not using it. It is not preserved or supported in DIPs.
DMDID	ET, ES	Reference to descriptive metadata is done from structural metadata. The attribute is not preserved or supported in DIPs.
GROUPID	ET, V	Attribute is allowed in SIPs, if the partner organization needs it. The DPS does not use it and although it is preserved it is not supported in DIPs. Attribute value should be expressed in printable US-ASCII characters.

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
-<FLocat>	ET, P	<FLocat>-element is used for defining the file's location in the package.
ID	ET, ES	ID attribute is not recommended, because there is no need to refer to <flocat> element in a profile compliant document. The attribute is neither preserved nor supported in DIPs.
LOCTYPE	ET, P	Mandatory in the METS schema. Attribute value in the national METS is URL.
xlink:role, xlink:arcrole, xlink:title, xlink:show, xlink:actuate,	ET, ES	Attributes are allowed in SIPs, but the DPS does not use them. They are neither preserved nor supported in DIPs.
OTHERLOCTYPE	-	OTHERLOCTYPE must not be used in national METS profile compliant information packages because URL is the only reference allowed.
USE	ET, ES	USE attribute must be used in <fileGrp>-element. If repeated here, the attribute is neither preserved nor supported in DIPs.
xlink:href	ET, P	File location relative to SIP's root. The location must be encoded according to the xlink specification [XLINK].
xlink:type	ET, P	Attribute value must be "simple".
-<stream>	T, V	<stream> element may be used to indicate parts of a file; for instance, if there is a video document in a container, the element may be used to present soundtrack and moving images ¹³ . The element must not have content, only attributes.
ADMID	ET, P	Values of the ADMID attribute refer to the administrative metadata of a part of the file.
ID	ET, V	ID attribute may be present in order to make it possible to refer to the part of the file from the <structMap> element
streamType OWNERID DMDID BEGIN END BETYPE	ET, V	Attributes are allowed in SIPs, but the DPS does not use them. They are neither preserved nor supported in DIPs.
-<FContent> -<file> ¹⁴ --transformFile>	-	These elements are not allowed in a national METS profile compliant document.

¹³ Cf. section 4 in File Formats specification [TDSTO]

¹⁴ Recursive use of the <file> element is forbidden.

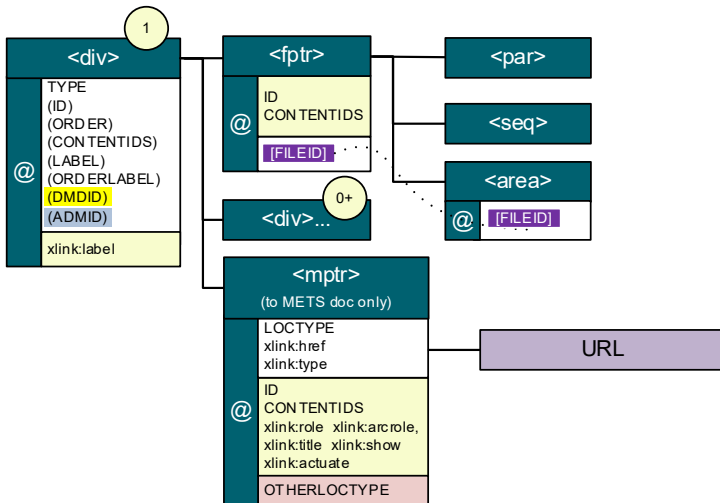
A.11. <structMap> element



Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<structMap>	T, P	A national METS profile compliant document includes one or several <structMap> elements. Each element may include one or several <div> elements.
ID	ET, V	Attribute may be present in SIPs, if the partner organization needs it. The DPS preserves the attribute but it is not supported in DIPs.
TYPE	ET, V	The recommended controlled vocabulary for attribute ¹⁵ .
LABEL	ET, V	Attribute may be present in SIP, if the partner organization needs it. The DPS does not preserve the attribute and it is not supported in DIPs.
fi:PID	ET,V	Unique identifier of metadata. Attribute value should be expressed in printable US-ASCII characters.
fi:PIDTYPE	ET,V	Identifier system used in the fi:PID attribute. Attribute is mandatory, if the PID attribute is used. Attribute value should be expressed in printable US-ASCII characters.
-<div>	EP,P	See section A.12

¹⁵ <http://digitalpreservation.fi/specifications/vocabularies>

A.12. <div> element



Element/Attribute	Occurrence/obligation	Rules and recommendations
<div>	T, P	Each information package must have at least one <div> element.
ID	ET, V	ID attribute is not recommended, because there is no need to refer <div> element in a profile compliant document.
ORDER	ET, V	Attribute may be present in SIPs, if the partner organization needs it. The DPS does not use the attribute.
CONTENTIDS	ET, V	Attribute may be present in SIPs, if the partner organization needs it. The DPS does not use the attribute.
TYPE	ET, P	The recommended controlled vocabulary for attribute ¹⁶ .
LABEL	ET, V	Attribute may be present in SIPs, if the partner organization needs it. The DPS does not use the attribute.
ORDERLABEL	ET, V	Attribute may be present in SIPs, if the partner organization needs it. The DPS does not use the attribute.
DMDID	ET, V	Reference to descriptive metadata (<dmdSec>).
ADMID	ET, V	Reference to administrative metadata (<techMD>, <rightsMD>, <sourceMD> tai <digiprovMD>).
xlink:label	ET, ES	Attribute is allowed in SIPs, but the DPS is not using it. It is not supported in DIPs.
-<fptr>	T, V	
ID	ET, ES	The ID attribute is not recommended, because there is no need to refer <fptr> element in a profile compliant document. The attribute is not used in DIPs.
CONTENTIDS	ET, ES	Attribute may be in SIPs, if the partner organization needs it. The DPS does not use the attribute. The attribute is not used in DIPs.
FILEID	ET, PV	Reference to a <file> or to a <stream> -element (See also <area>-element below).

¹⁶ <http://digitalpreservation.fi/specifications/vocabularies>

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<par>, <seq>, <area>	T, V	<p><par>-, <seq>- ja <area>-elements can be used in a national METS profile compliant package, but the DPS is not checking how they are used, and the DPS is not maintaining the structs in for example migrations. Therefore, the element usage is not defined.</p> <p><area> element must use the FILEID reference, if it is not possible to refer it in <fptr> element. It is possible to refer to only one file with <fptr> element. If <fptr> is containing references to many files, these references have to be presented in <area> elements. It is not recommended to use FILEID in <fptr> elements with these cases.</p>
-<div>	T, V	A <div> element may contain several <div> elements consecutively or one inside another.
-<mptr>	T, V	
ID	ET, ES	ID attribute is not recommended, because there is no need to refer to <mptr> element in a profile compliant document. It is not supported in DIPs.
LOCTYPE	ET, P	Attribute value in a national METS profile is URL.
xlink:href	ET, P	Location of referring METS document relative to the information package root. The location must be encoded according to the xlink specification [XLINK]
xlink:type	ET, P	Attribute value has to be "simple".
CONTENTIDS, xlink:role, xlink:arcrole, xlink:title, xlink:show, xlink:actuate,	ET, ES	Attributes are allowed in SIPs, but the DPS does not use them. It is not supported in DIPs.
OTHERLOCTYPE	-	OTHERLOCTYPE is not in use in a national METS profile compliant package, because URL is the only allowed reference.

A.13. <mdWrap> element

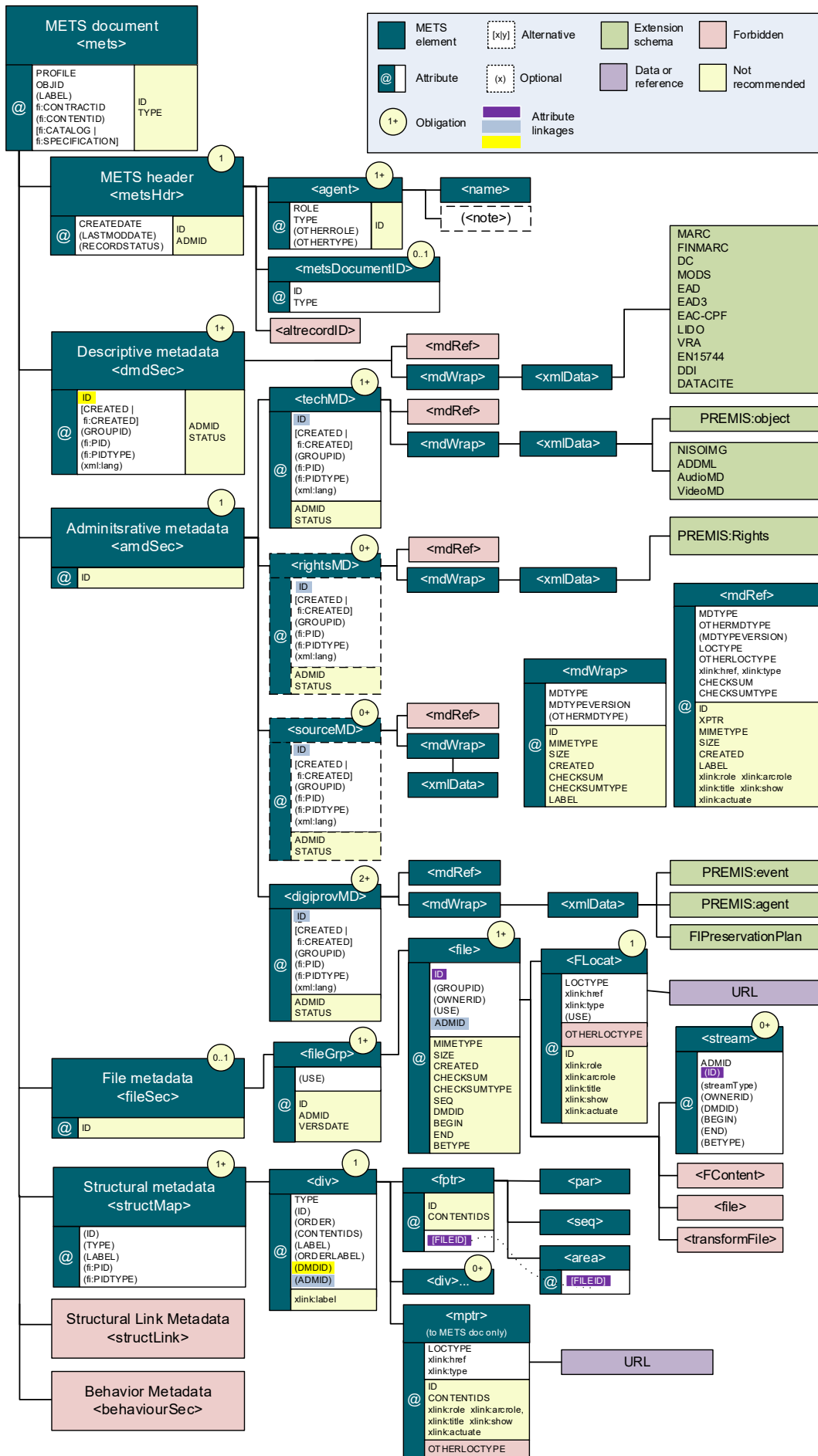
<mdWrap>	
@	MDTYPE MDTYPEVERSION (OTHERMDTYPE)
@	ID MIMETYPE SIZE CREATED CHECKSUM CHECKSUMTYPE LABEL

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<mdWrap>	ET, P	The <mdWrap> element contains the actual metadata in <dmdSec>, <techMD>, <rightsMD>, <sourceMD>, and <digiprovMD> elements.
MDTYPE	ET, P	<p>Allowed values of MDTYPE attribute for descriptive metadata are: MARC, MODS, DC, EAD, EAC-CPF, LIDO, VRA ja DDI. For other descriptive metadata formats, the OTHERMDTYPE attribute must be used.</p> <p>Allowed values of MDTYPE attribute for technical metadata are: PREMIS:OBJECT and NISOIMG. For other technical metadata formats, the OTHERMDTYPE attribute must be used.</p> <p>Allowed values of MDTYPE attribute for preservation rights: PREMIS:RGITHS</p> <p>Allowed values of MDTYPE attribute for provenance metadata are: PREMIS:AGENT ja PREMIS:EVENT.</p> <p>In other cases, the value OTHER and attribute OTHERMDTYPE must be used.</p>
OTHERMDTYPE	ET, V	<p>Mandatory if MDTYPE attribute value is OTHER, i.e., MDTYPE attribute cannot be used to describe metadata format according to the schema.</p> <p>When describing the preservation plan, attribute value must be FiPreservationPlan.</p>
MDTYPEVERSION	ET, P	Attribute must be used to describe metadata format version.
ID	ET, ES	ID attribute is not recommended, because there is no need to refer to <mdWrap> element in a profile compliant document. The attribute is not used in DIPs.
MIMETYPE, SIZE, CREATED, CHECKSUM, CHECKSUMTYPE, LABEL	ET, ES	Attributes are allowed in SIPs, but the DPS does not use them. They are not supported in DIPs.
-<xmlData>	ET, P	Contains the actual metadata.
-<binData>	-	Element is forbidden in national METS profiles.

A.14. <mdRef> element

<mdRef>	
@	MDTYPE OTHERMDTYPE (MDTYPEVERSION) LOCTYPE OTHERLOCTYPE xlink:href, xlink:type CHECKSUM CHECKSUMTYPE
@	ID XPTR MIMETYPE SIZE CREATED LABEL xlink:role xlink:arcrole xlink:title xlink:show xlink:actuate

Element/ Attribute	Occurrence/ obligation	Rules and recommendations
<mdRef>	ET, V	Usage of the <mdRef>-element in a national METS profile compliant package is allowed only if it contains the preservation plan.
MDTYPE	ET, P	In a national METS profile compliant package, the attribute value is OTHER.
OTHERMDTYPE	ET, P	In a national METS profile compliant package, the attribute value is FiPreservationPlan.
MDTYPEVERSION	ET, V	Version of the metadata that is used for describing the preservation plan.
LOCTYPE	ET, P	In a national METS profile compliant package, the attribute value is OTHER.
OTHERLOCTYPE	ET, P	In a national METS profile compliant package, the attribute value is PreservationPlanID.
xlink:href	ET, P	The location of the preservation plan relative to the SIP root. The location must be encoded according to the xlink specification [XLINK].
xlink:type	ET, P	Attribute value has to be "simple".
ID, XPTR, MIMETYPE, SIZE, CREATED, LABEL, CHECKSUM, CHECKSUMTYPE xlink:role, xlink:arcrole, xlink:title, xlink:show, xlink:actuate	ET, ES	Attributes are allowed in SIPs, but the DPS does not use them. They are not supported in DIPs.



REFERENCES

- [AUDIOMD] Audio Technical Metadata Extension Schema
<http://www.loc.gov/standards/amdvmd/audiovideoMDschemas.html>
- [ISO 8601-1] ISO 8601-1:2019. Date and time — Representations for information interchange — Part 1: Basic rules, 2019
- [ISO 8601-2] ISO 8601-2:2019. Date and time — Representations for information interchange — Part 2: Extensions, 2019
- [METS] Metadata Encoding and Transfer Standard. <http://www.loc.gov/standards/mets/>
- [MIX] NISO Metadata for Images in XML Schema. <http://www.loc.gov/standards/mix/>
- [OAIS] Reference Model for an Open Archival Information System (OAIS), June 2012.
<http://public.ccsds.org/publications/archive/650x0m2.pdf>
- [PREMIS] Data Dictionary for Preservation Metadata: PREMIS version 2.2. July 2012.
<http://www.loc.gov/standards/premis/v2/premis-2-2.pdf>
- [RFC 2119] Key words for use in RFCs to Indicate Requirement Levels. <http://www.ietf.org/rfc/rfc2119.txt>
- [SS] Digime Standard Portfolio. <https://www.digime.fi/yhteentoimivuus/kokonaisarkkitehtuuri/>
- [TDSTO] File Formats. <http://www.digitalpreservation.fi/en/specifications>
- [Vermaaten] A Checklist for Documenting PREMIS-METS Decisions in a METS Profile, Sally Vermaaten, May 2010.
http://www.loc.gov/standards/premis/premis_mets_checklist.pdf
- [VIDEOMD] Video Technical Metadata Extension Schema.
<http://www.loc.gov/standards/amdvmd/audiovideoMDschemas.html>
- [XLINK] XML Linking Language (XLink) Version 1.0. <http://www.w3.org/TR/xlink/>