

Issues in using PREMIS with METS

- § Which METS sections to use and how many
- § Whether to record elements redundantly in PREMIS that are defined explicitly in the METS schema
- § How to record elements that are also part of a format specific technical metadata schema (e.g. MIX)
- § Recording structural relationships
- § How to deal with locally controlled vocabularies
- § Whether to use the PREMIS container

PREMIS and METS sections

- § Flexibility of METS requires implementation decisions
- § You can't put all PREMIS metadata directly under amdSec
- § What sections to use for PREMIS metadata?
 - Alternative 1
 - Object in techMD
 - Event in digiProvMD
 - Rights in rightsMD
 - Agent with event or rights
 - Alternative 2
 - Everything in digiProvMD
 - Alternative 3
 - Everything in techMD
- § How many administrative MD sections to use?
- § Experimentation will result in best practices

PREMIS

PREservation Metadata Implementation Strategies

```
<fileSec><fileGrp>
<file ID="FID1" SIZE="184302" ADMID="TMD1PREMIS TMD1MIX DP1EVENT
DP1AGENT"
CHECKSUM="4638bc65c5b9715557d09ad373eefd147382ecbf"
CHECKSUMTYPE="SHA-1">
<FLocat LOCTYPE="OTHER" xlink:href="BXF22.JPG" />
</file></fileGrp></fileSec>
<techMD ID="TMD1PREMIS">
<mdWrap MDTYPE="PREMIS">
<xmlData>
<premis:object >
<objectCharacteristics>
<fixity>
<messageDigestAlgorithm>SHA-1 </messageDigestAlgorithm>
<messageDigest>4638bc65c5b9715557d09ad373eefd147382ecbf
</messageDigest>
<messageDigestOriginator>EchoDep/messageDigestOriginator>
</fixity>
<size>184302</size> </objectCharacteristics>
```

Elements defined in both METS and PREMIS:

- **METS: Checksum, Checksumtype**
 - attribute of <file>
 - not repeatable
- § **PREMIS: fixity**
 - also includes messageDigestOriginator
 - allows multiples

PREMIS

PREservation Metadata Implementation Strategies

```
<fileSec><fileGrp>  
<file ID="FID1" ADMID="TMD1PREMIS DP1EVENT DP1AGENT"  
  MIMETYPE="image/jpeg"  
<FLocat LOCTYPE="OTHER" xlink:href="BXF22.JPG"/>  
</file></fileGrp></fileSec>
```

```
<techMD ID="TMD1PREMIS"  
  <mdWrap MDTYPE="PREMIS">  
    <xmlData>  
      <premis:object>  
        <objectCharacteristics>  
          <format>  
            <formatDesignation>  
              <formatName>image/jpeg</formatName>  
              <formatVersion>1.02 </formatVersion>  
            </formatDesignation></format>  
          </objectCharacteristics>
```

Elements defined both in METS and PREMIS:

- **METS: MIMETYPE**

- attribute of <file>
- optional

- § **PREMIS: <format>**

- more granular; includes name and version (although name may be MIMETYPE)
- mandatory

PREMIS

PREservation Metadata Implementation Strategies

```
<fileSec> <fileGrp>
  <file ID="FID1" ADMID="TMD1PREMIS TMD1MIX DP1EVENT DP1AGENT">
<techMD ID="TMD1PREMIS">
  <linkingEventIdentifier>
    <linkingEventIdentifierType>ECHODEP Hub Event
  </linkingEventIdentifierType>
  <linkingEventIdentifierValue>echo12345</linkingEventIdentifierValue>
  </linkingEventIdentifier>
<digiprovMD ID="DP1EVENT">
  <premis:event>
    <eventIdentifier>
      <eventIdentifierType>ECHODEP Hub Event</eventIdentifierType>
      <eventIdentifierValue>echo12345 </eventIdentifierValue>
    </eventIdentifier>
    <eventType>ingestion</eventType>
    <eventDateTime>2006-05-02T15:12:53 </eventDateTime></event>
```

Elements defined both in METS and PREMIS

- § **METS ID/Idref:** used to associate metadata in different sections and for different files
- § **PREMIS identifiers:** explicit linking between entity types

```

<structMap TYPE="physical">
  <div ORDER="1" TYPE="text">
    <:fptr FILEID="FID9" />
    <div ORDER="1" TYPE="page" LABEL=" Page [1]">
      <fptr FILEID="FID1" /></mets:div>
    <div ORDER="2" TYPE="page" LABEL=" Page [2]">
      <fptr FILEID="FID2" /></mets:div>
    </div>
  </div>

```

```

<relationship>
  <relationshipType>structural</relationshipType>
  <relationshipSubType>is sibling of </relationshipSubType>
  <relatedObjectIdentification>
    <relatedObjectIdentifierType>UCB</relatedObjectIdentifierType>
    <relatedObjectIdentifierValue>FID2</relatedObjectIdentifierValue>
    <relatedObjectSequence>1</relatedObjectSequence>
  </relatedObjectIdentification>
</relationship>

```

Elements defined both in METS and PREMIS:

§ METS: structMap

- details structural relationships and is the heart of the METS document
- hierarchical, so may be more expressive than PREMIS semantic units
- links the elements of the structure to content files and metadata

§ PREMIS: <relationship>

- details all kinds of relationships, including structural
- data dictionary says that implementations may record by other means

Should semantic units be recorded redundantly?

- § Various options are possible when there is overlap between PREMIS and METS or PREMIS and other technical metadata schemas
 - Record only in METS
 - Record only in PREMIS
 - Record in both
- § Are there advantages in using PREMIS semantic units?
- § Is it important to keep PREMIS metadata together as a unit? There may be an advantage for reuse and maintenance purposes

Possible XPATH solution to redundant semantic units

§ From Markus Enders:

§ I would extend the premis schema in a way, that every element, which might contain a text value (a text node), may have an attribute called e.g. valuePointer and a valuePointerType.

§ Example: `<premis:fixity valuePointerType="xpath" valuePointer="mets:mets/mets:fileSec/mets:fileGrp/mets:file/@CHECKSUM" />` (note: the element itself is empty).

§ Of course these XPath expressions can become more complex - getting the 5th file in the 2nd fileGroup or e.g the checksum of a file with a special id: `<premis:fixity valuePointerType="xpath" valuePointer="mets:mets/mets:fileSec/mets:fileGrp/mets:file[@ID="file01"]/@CHECKSUM"/>` If you like, you would even be able to use a relative path (something I would not recommend).

How to record elements from 2 different technical metadata schemas

- § Format specific metadata may be included in addition to PREMIS general technical metadata
- § Use multiple techMD sections and specify source in MDType attribute and/or namespace declaration
 - e.g. MDTYPE="NISOIMG" or "PREMIS"
 - Give MIX schema declaration in METS document
- § MIX was recently revised to correspond with the revision of the Z39.87 technical metadata for digital still images standard; names harmonized with corresponding PREMIS semantic units
- § For digital still images, best practice may be to use PREMIS for general semantic units defined in PREMIS and MIX for format specific units without redundancy

ID/IDREF vs. XPTR

- § Should we use ID/IDREF to link from METS to PREMIS and PREMIS to METS?
- § Is XPTR a better option? Should we add an XPTR attribute to elements that already ID/IDREF attributes?
- § Issues:
 - If you maintain separate xml files and associate metadata via mdRef attributes than IDREFs to that metadata do not validate

XPTR sample from Clay Redding

§ `<mets:mdRef LOCTYPE="URL" MDTYPE="PREMIS" xlink:href="premis.xml" XPTR="xpointer(id('object1'))"/>`

§ `<!-- The following won't validate because DMDID='ver02' doesn't appear in this METS, only in the referenced external MODS Unless I have something like @DMDXPTR or @AMDXPTR, ala @DMDXPTR="mods.xml#xpointer(id('ver02'))" An alternative construct could follow that which is on mdRef, with @xlink:href and @XPTR -->`

§ `<mets:div TYPE="photo:version" DMDID="ver02">`

§ `<mets:div TYPE="photo:image">`

§ `<mets:fptr FILEID="masterd1e102965"/>`

§ `<mets:fptr FILEID="serviced1e102965"/>`

§ `</mets:div>`

PREMIS Container Element

- § There is some agreement to use if all PREMIS in one place, do not use if multiple PREMIS statements.
- § How to indicate the version of a PREMIS schema?
 - @version
 - namespace
 - Major vs. Minor PREMIS schemas
- § What is the value of MDTYPE in METS if the PREMIS container is not used? Do we need to add a MDVERSION?
- § Should object be required for PREMIS container schema?

Should we change METS?

- § Use of amdSec
 - Relax requirement to use children (techMD, etc.)?
 - Add new child element (otherMD or preservationMD)?
- § Add to MDTYPE vocabulary for PREMIS subschemas?
- § Add MDVERSION attribute?
- § Add @XPTR?

Should we change PREMIS?

- § Extensibility mechanism for format specific metadata
- § PREMIS container requirement for objectID?
- § Use of PREMIS ID/IDREF pairs?