A CHECKLIST FOR DOCUMENTING PREMIS-METS DECISIONS IN A METS PROFILE

May 2010 | Sally Vermaaten, OCLC

Purpose of this checklist

As the push for long-term access to digital information increases, a growing number of organizations are using PREMIS (PREservation Metadata: Implementation Strategies) in METS (Metadata Encoding and Transmission Standard) to record provenance and other information that supports sustained access. This checklist was designed to aid those implementing, or considering implementing, PREMIS in METS. The checklist presents and provides examples of 13 key PREMIS-METS issues that implementers should consider—and should consider documenting—when designing or revisiting a metadata strategy.

This 13-point checklist allows users to quickly benefit from a significant body of PREMIS-METS experience by bringing together—though certainly not replacing—a wide range of PREMIS-METS knowledge recorded in resources such as the *Guidelines for Using PREMIS with METS for Exchange*,[i] case studies published in journals,[ii] conference and workshop presentations, local documentation on institutional Web sites, and the "METS Profiles Components" document. [iii] The checklist is further enriched by insights and examples gleaned from a 2009 analysis of PREMIS usage in METS profiles registered with the Library of Congress.[iv]

Each of the points in the checklist is followed by a discussion and one or more examples from METS profiles registered by institutions such as the National Library of Australia, University of Illinois, University of California San Diego, and the University of Southampton. The items in this checklist generally follow the order of elements stipulated in the "METS Profile Components" document.

For more information about the checklist and the internal and external benefits of documenting PREMIS-METS implementations in METS registered profiles, see the author's forthcoming article in the September/October 2010 issue of D-Lib Magazine.

The checklist

1. How does the profile relate to other METS profiles?

Discussion: An analysis of METS registered profiles showed that several repositories use more than one METS profile. It is therefore useful to specify the relationship of one METS profile to another (e.g., is it a sub-profile, parent or generic profile, or a sibling profile?). Locating detailed PREMIS-METS documentation in one profile which can then be inherited by other profiles that are more specific to a particular type content stream or object type can be a useful documentation strategy. It is also important to note if a profile supersedes an older version of the same profile.

Examples: Both the University of Illinois Urbana-Champaign and the institutions that created the Australian METS Profile use parent and sub-profiles that carefully specify the relationships between METS profiles. For example, the abstract of University of Illinois' profile for web captures notes:

> This profile inherits much of its content from the 'ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability.' Unless otherwise noted below this profile must conform to the same rules as the parent profile. [v]

The abstract for the Australian METS Profile notes:

> This profile describes the rules and requirements for using METS as an exchange format to support the collection and preservation of and access to content in Australian digital repositories. It is a generic profile not specific to a particular system or implementation. Repositories will need to develop and register sub-profiles that detail implementation-specific requirements. [vi]

2. What schemas (PREMIS, MODS, MIX) are used and where are they located?

Discussion: According to the METS Profile Components document, a profile should explicitly record if it uses PREMIS or metadata schema such as MODS or MIX and where that schema is located.

Examples: The University of Southampton's RLUK 19th Century Pamphlets METS Profile notes their use of PREMIS (see Figure 1). [vii]

```
<extension_schema ID="PREMIS">
    <name>PREMIS Data Dictionary for Preservation Metadata</name>
    <URI>http://www.loc.gov/standards/premis/v2/premis-v2-0.xsd</URI>
    <context>mets/amdSec/digiProvMD|techMD/mdWrap/xmlData</context>
        <note>Technical metadata at the file level, extracted from standard file information. There will be one instance per file of any kind listed in the fileSec. This metadata will be contained in an amdSec linked to the relevant file element. The PREMIS components will be contained within seperate METS elements, as suggested by the "Guidelines for using PREMIS within METS" at http://www.loc.gov/standards/premis/premis-mets.html techMD: premis:object for each file digiProvMD: premis:event containing information about file derivations digiProvMD: premis:agent containing information about software packages PREMIS metadata will conform to version 2.0 of the schema.</note>
</extension_schema>
```

Figure 1. University of Southampton notes their PREMIS use.

3. What controlled vocabularies for PREMIS semantic units are used and where are they located?

Discussion: The *PREMIS Data Dictionary* recommends placing controlled vocabularies in a common space or service where a vocabulary can be re-used by multiple repositories.[viii] If a repository uses shared vocabularies, its METS profile should point to the relevant vocabularies and registries. If it is not possible to use a shared vocabulary service, repositories can document controlled vocabularies for PREMIS semantic units in the METS profile itself or in separate local documentation referenced in the profile. PREMIS semantic units that lend themselves to controlled vocabularies include but are not limited to: *eventType, agentType*, *format*, *objectCategory*, *storageMedium*, *copyrightStatus*, and *RightsBasis*.

Example: At the time of writing, there is no single central location where users can register and find shared controlled vocabularies for PREMIS semantic units or METS elements. However, the Library of Congress has recently developed an Authorities & Vocabularies service, www.id.loc.gov, which now supports shared controlled vocabularies for cryptographic hash functions, preservation events, and preservation level roles.

The University of California San Diego Simple Object METS profile[ix] specifies usage of several controlled vocabularies recorded in separate documents on the UCSD Web site (see Figure 2).

```
<controlled_vocabularies>
        <vocabulary>
                <name>UC San Diego file USE attribute values</name>
                <maintenance_agency>University of California, San Diego Libraries</maintenance_agency>
                <URI> https://libraries.ucsd.edu/ark:/20775/br747447 1n/1-1.pdf </URI><context ID="vc1"
                RELATEDMAT="fileSec2"><p>mets/fileSec/fileGrp/@USE</p>
                <p>mets/fileSec/fileGrp/file/@USE</p></context>
                <description><p>These are the values supported for <file> and <fileGrp> USE attributes for digital objects conforming to this
                profile.</p></description>
        </vocabulary>
        <vocabulary>
                <name>UC San Diego Rights Basis Vocabulary</name>
                <maintenance_agency>University of California, San Diego Libraries</maintenance_agency>
                <URI>https://libraries.ucsd.edu/ark:/20775/br747447 1n/1-1.pdf</URI> <context ID="vc2" RELATEDMAT="copyrightMD2
                licenserightsMD2 statuterightsMD2"><p>mets/amdSec/rightsMD/mdWrap/xmlData/rightsStatement/rightsBasis</p> </context>
                <description>
                <p>These are the values supported for the PREMIS rightsBasis element for digital objects conforming to this profile.</p>
                </description>
        </vocabulary>
        <vocabulary>
                <name>UC San Diego Rights Status Vocabulary</name>
                <maintenance_agency>University of California, San Diego Libraries</maintenance_agency>
                <URI>https://libraries.ucsd.edu/ark:/20775/br747447 1n/1-1.pdf</URI> <context ID="vc3"
                RELATEDMAT="copyrightMD3"><p>mets/amdSec/rightsMD/mdWrap/xmlData/rightsStatement/copyrightStatus</p> </context>
                <description>
                <p>These are the values supported for the PREMIS copyrightStatus element for digital objects conforming to this profile.</p>
                </description>
        </vocabulary>
        <vocabulary>
                <name>UC San Diego Rights Actions Vocabulary</name>
                <maintenance_agency>University of California, San Diego Libraries</maintenance_agency>
                <URI>https://libraries.ucsd.edu/ark:/20775/br747447 1n/1-1.pdf</URI> <context ID="vc4" RELATEDMAT="licenserightsMD3
                licenserightsMD4 statuterightsMD4"><p>mets/amdSec/rightsMD/mdWrap/xmlData/rightsStatement/rightsGranted/restriction</p>
                <p>mets/amdSec/rightsMD/mdWrap/xmlData/rightsStatement/rightsGranted/action</p> </context>
                <description>
                <p>These are the values supported for the PREMIS act and restriction elements for digital objects conforming to this profile.</p>
                </description>
        </vocabulary>
</controlled_vocabularies>
```

Figure 2. A UCSD profile references multiple controlled vocabularies described in documents on their Web site.

4. Is PREMIS information wrapped into or referenced from the METS document?

Discussion: All PREMIS-implementing METS Profiles registered with the Library of Congress at the time of writing wrap metadata from other schema into the METS document using METS *mdWrap*. However, it is also possible to store PREMIS semantic units outside of the METS document and reference them using the METS *mdRef* element. The *Guidelines* ultimately leave it up to the implementer to choose whether to wrap or reference PREMIS information (or both). However, if all PREMIS metadata is outside of the METS document, the Guidelines state that PREMIS linking identifier elements should be used to connect PREMIS and METS documents since ID/IDRefs may break. Regardless of the strategy adopted, this decision should be explicitly documented in the METS profile.

Example: Figure 3 below shows a typical example of PREMIS wrapped in the METS manifest from a sample METS document appended to the University of Illinois' ECHO Dep METS Profile for Web Site Captures.[x]

```
<techMD ID="APP3_TMD1PREMIS">
        <mdWrap MDTYPE="PREMIS">
                <xmlData>
                <object xmlns=http://www.loc.gov/standards/premis/v1
                 xsi:schemaLocation="http://www.loc.gov/standards/premis/v1
                http://www.loc.gov/standards/premis/v1/PR...">
                        <objectIdentifier>
                                <objectIdentifierType>URI</objectIdentifierType>
                                <objectIdentifierValue>http://www.nps.gov/rich/fall99.html</objectIdentifierValue>
                        </objectIdentifier>
                        <objectCategory>file</objectCategory>
                        <objectCharacteristics>
                                <compositionLevel>0</compositionLevel>
                                <fixity>
                                <messageDigestAlgorithm>SHA-1</messageDigestAlgorithm>
                                <messageDigest>4ea7325ecef266792a03e5f82ce67762970e14a9</messageDigest>
                                </fixity>
                                <size>31264</size>
                                <format>
                                <formatDesignation>
                                <formatName>text/html</formatName>
                                <formatVersion>4.0</formatVersion>
                                </formatDesignation>
                                </format>
                        </objectCharacteristics>
                </object>
                </xmlData>
        </mdWrap>
</techMD>
```

Figure 3. The University of Illinois' ECHO Dep METS Profile for Web Site Captures wraps PREMIS into METS using *mdWrap*.

The University of Illinois documents the decision to wrap PREMIS into METS in a parent profile, the ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability (see Figure 4).[xi] By stating its conformance to the parent profile, the Web Site Captures profile automatically inherits this requirement.

```
<structural_requirements>…
        <amdSec>…
                <p> The content of the techMD section must be a PREMIS object element embedded in the techMD via the mdWrap element. The
                PREMIS object element must conform to the XML Schema of PREMIS. </p>
        </amdSec>…
</structural_requirements>
```

Figure 4. The parent profile (University of Illinois' ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability) mandates wrapping PREMIS into METS.

5. Is PREMIS information bundled or distributed in several places in the METS document?

Discussion: As the *Guidelines* point out, there are several places where PREMIS can be placed within a METS document. Because these placement decisions are complex, the decisions, and possibly the rationale for them, should be documented in a METS profile.

If keeping all PREMIS semantic units together, the *Guidelines* state that best practice is placing the entire package in *digiProvMD* with the *premis* element as a container. If splitting up PREMIS information across multiple sections of METS, the *Guidelines* state that the *premis* element should not be used. Instead, it is best practice to place:

- *premis:event* under *digiProvMD*
- *premis:rights* under *rightsMD*
- *premis:object* under *techMD* or *digiProvMD*
- *premis:agent* under *digiProvMD* or *rightsMD* (depending on whether the agent is connected with events or rights)

Example: All but one of the profiles examined (the Library of Congress' 2005 METS Profile for Recorded Events) distributed PREMIS semantic units in various METS *amdSec* subelements rather than using a single "package" of PREMIS information. For example, the National Library of Australia has placed *premis:object* under *techMD* and *premis:event* in *digiprovMD* (see Figure 5 below).[xii]

```
<mets:techMD ID="file-2"><mets:mdWrap MDTYPE="PREMIS">
        <mets:xmlData>
        <premis:object>
                <premis:objectIdentifier>
                        <premis:objectIdentifierType>uri</premis:objectIdentifierType>
                        <premis:objectIdentifierValue>info:nla/nla.pic-vn3579101-c</premis:objectIdentifierValue>
                        </premis:objectIdentifier>
                <premis:preservationLevel>unknown</premis:preservationLevel>
                <premis:objectCategory>file</premis:objectCategory>
                <premis:objectCharacteristics>
                        <premis:format>
                        <premis:formatDesignation>
                        <premis:formatName>image/tiff</premis:formatName>
                        <premis:formatVersion>6.0</premis:formatVersion>
                        </premis:formatDesignation>
                        </premis:format>
                </premis:objectCharacteristics>
                        …
        </premis:object>
        </mets:xmlData>
</mets:mdWrap>
</mets:techMD>
<mets:digiprovMD ID="event-1"><mets:mdWrap MDTYPE="PREMIS">
<mets:xmlData>
        <premis:event>
                <premis:eventIdentifier>
                        <premis:eventIdentifierType>internal</premis:eventIdentifierType>
                        <premis:eventIdentifierValue>28903-1</premis:eventIdentifierValue>
                </premis:eventIdentifier>
                <premis:eventType>creation</premis:eventType>
                <premis:eventDateTime>2005-11-03T12:15:59</premis:eventDateTime>
                …
        </premis:event>
        </mets:xmlData>
</mets:mdWrap><mets:digiprovMD>
```

Figure 5. The National Library of Australia splits up PREMIS information among multiple METS sections.

6. Is PREMIS information placed in separate *amdSec* elements or *amdSec* subelements?

Discussion: According to the *Guidelines*, implementers can choose to place PREMIS information within separate METS *amdSecs* or within one *amdSec* in different sub-elements (e.g. *techMD, digiprovMD*, *rightsMD*). The key thing to keep in mind here is that referencing one *amdSec* with PREMIS information in different sub-elements means referencing the *amdSec* and all of its children. Consider documenting the rationale for determining what goes in each *amdSec* or sub-element section and if there is a standard alphanumeric scheme that is used to identify sections.

Examples: The University of California San Diego places different PREMIS entities in separate *amdSec* sub-elements and assigns each an ID of "ADM1," "ADM2" (Figure 6).[xiii]

```
<mets:amdSec>
        <mets:techMD ID="ADM1">
                <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS">
                <mets:xmlData>
                <pre:object>
                        <pre:objectIdentifier>
                                <pre:objectIdentifierType>ARK</pre:objectIdentifierType>
                                <pre:objectIdentifierValue>http://libraries.ucsd.edu/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>
        <mets:rightsMD ID="ADM2">
                <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS">
                <mets:xmlData>
                        <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>
        …
</mets:amdSec>
```

Figure 6. The University of California San Diego's Electronic Theses and Dissertations Profile places PREMIS in separate *amdSec* sub-elements.

7. Is technical metadata recorded in separate *techMD* sections or with PREMIS *objectCharacteristicExtension*?

Discussion: Technical, format-specific metadata such as that recorded in MIX or textMD can be wrapped or referenced in separate METS *techMD* sections. Alternatively, technical metadata can be wrapped or referenced in the PREMIS semantic unit *objectCharacteristicExtension*. The *Guidelines* state that choosing either location (or both) is an implementation-specific decision. Repositories should document their decision in their profile and, if applicable, the naming scheme for their *techMD* sections so that the correct section can be referenced elsewhere in the METS document.

Examples: Figure 7 below shows an example of how the University of Illinois' Echo Dep Profile for Preservation and Digital Repository Interoperability mandates wrapping technical metadata in separate *techMD* sections. [xiv]The profile then illustrates this requirement in an example METS document in an appendix to the profile.

```
<structural_requirements>
        …
        <amdSec>
        …
        <requirement>
        <head>Technical Metadata for Files with a Root MIME Type of 'Image'</head>
        <p> In addition to required elements specified in Technical Metadata for files and bitstreams, all files with a root MIME type of "image" should
        provide a second techMD section in addition to the PREMIS section previously mentioned. This  second techMD section must wrap a MIX
        Schema description of the file.</p>
        </requirement>
        …
        </amdSec>
        …
<structural_requirements>
…
<Appendix NUMBER="1" LABEL="Example of a Simple Object Consisting of one Scanned Image">
<techMD ID="APP1_TMD1MIX">
        <mdWrap MDTYPE="NISOIMG">
        <xmlData>
                <mix xsi:schemaLocation="http://www.loc.gov/mix/ http://www.loc.gov/standards/mix/mix.xsd">
                        <BasicImageParameters>
                                <Format>
                                <MIMEType>image/jpeg</MIMEType>
                                <ByteOrder>big-endian</ByteOrder>
                                <Compression>
                                <CompressionScheme>6</CompressionScheme>
                                </Compression>
                                <PhotometricInterpretation>
                                <ColorSpace>6</ColorSpace>
                                </PhotometricInterpretation>
                                </Format>
                        </BasicImageParameters>
                        <ImageCreation />
                        …
                        <ChangeHistory>
                                <ImageProcessing>
                                <DateTimeProcessed>2006-01-07T14:26:40-06:00</DateTimeProcessed>
                                <ProcessingSoftware>
                                <ProcessingSoftwareName>Adobe Photoshop</ProcessingSoftwareName>
                                <ProcessingSoftwareVersion>5.0</ProcessingSoftwareVersion>
                                </ProcessingSoftware>
                                </ImageProcessing>
                        </ChangeHistory>
                </mix>
        </xmlData>
        </mdWrap>
</techMD>
```

Figure 7. The University of Illinois' Echo Dep Profile for Preservation and Digital Repository Interoperability wraps technical metadata in separate *techMD* sections.

8. What PREMIS semantic units does the profile require or recommend?

Discussion: A repository should record any required or recommended PREMIS semantic units in the structural requirements section of a METS profile along with any required values and/or conditions for use of the semantic unit. If different sets of PREMIS semantic units are required for several different scenarios, content types, or functions, the PREMIS semantic units and the circumstances under which they apply can be articulated either in one METS profile or in separate, specific sub-profiles of a more general parent METS profile. As more METS profiles are developed, best practices that list required semantic units for minimal and optimum compliance may be based on patterns in profiles for similar content types or functions.

Examples: The University of California San Diego's Electronic Theses and Dissertations Profile lists several required PREMIS semantic units and required values in the structural requirements section (Figure 8).[xv]

```
<requirement ID="amdSec1">
<p>A METS record conforming to this profile must have an <amdSec> section. The <amdSec> section must contain at least one
techMD/mdWrap/xmlData section and two rightsMD/mdWrap/xmlData sections.</p>
</requirement>
<requirement ID="techMD1">
<p>The required <techMD> must contain elements from the PREMIS schema.</p>
</requirement>
<requirement ID="techMD2">
<p>The required techMD/mdWrap/xmlData must describe the primary (pdf) digital file.</p>
</requirement>
<requirement ID="techMD3">
<p>The required techMD/mdWrap/xmlData must contain a /object/objectIdentifierType and /objectIdentifierValue element</p>
</requirement>
<requirement ID="techMD4">
<p>The required techMD/mdWrap/xmlData must contain a /object/preservationLevel element. The value should be either "Full" or "Bit-level".</p>
</requirement>
…
<requirement ID="copyrightMD1">
The required rightsMD/mdWrap/xmlData must contain elements from the PREMIS Rights schema.</p>
</requirement>
<requirement ID="copyrightMD2">
<p>A conforming METS record must include a rightsStatement/rightsBasis element with the value expressed as "Copyright."</p>
</requirement>
<requirement ID="copyrightMD3">
<p>A conforming METS record must include a rightsStatement/copyrightInformation/copyrightStatus element with the value expressed as "Under
copyright" if the work is still in copyright, or expressed as "In the public domain" if copyright for the work has expired or been gifted to the public
domain.</p>
</requirement>
…
<requirement ID="licenserightsMD1">
<p>A second required rightsMD/mdWrap/xmlData must contain elements from the PREMIS Rights schema.</p>
</requirement>
<requirement ID="licenserightsMD2">
<p>A conforming METS record must include a rightsStatement/rightsBasis element with the value expressed as "License."</p>
</requirement>
…
```

Figure 8. Several required PREMIS semantic units and values are stipulated in the University of California San Diego's Electronic Theses and Dissertations Profile.

9. Are relationships between objects expressed using METS *div* elements, PREMIS relationships, or both?

Discussion: The *Guidelines* state that hierarchical structural relationships between objects should be expressed as nested *div* elements in the METS schema. However, if the purpose of your profile is preservation or you are expressing derivative (e.g. image B derived from image A) relationships, the *Guidelines* also recommend using PREMIS relationships.

Examples: The Australian METS Profile 1.0 uses the METS *fileSec* and *structMap* and PREMIS *relationship* semantic units to express derivate and other relationships (Figure 9).[xvi]

```
<!--  techMD for the master file  -->
<mets:techMD ID="file1">
<mets:mdWrap MDTYPE="PREMIS"><mets:xmlData>
        <premis:object>
                …
                <premis:relationship>
                        <premis:relationshipType>derivation</premis:relationshipType>
                        <premis:relationshipSubType>derived from</premis:relationshipSubType>
                        <premis:relatedObjectIdentification>
                                <premis:relatedObjectIdentifierType>URI</premis:relatedObjectIdentifierType>
                                <premis:relatedObjectIdentifierValue>info:nla/nla.pic-vn3579081-o</premis:relatedObjectIdentifierValue>
                                <premis:relatedObjectSequence>0</premis:relatedObjectSequence>
                        </premis:relatedObjectIdentification>
                        <premis:relatedEventIdentification>
                                <premis:relatedEventIdentifierType>internal</premis:relatedEventIdentifierType>
                                <premis:relatedEventIdentifierValue>28902-1</premis:relatedEventIdentifierValue>
                                <premis:relatedEventSequence>0</premis:relatedEventSequence>
                        </premis:relatedEventIdentification>
                </premis:relationship>
                …
        </premis:object>
</mets:xmlData></mets:mdWrap>
</mets:techMD>
 …
<mets:fileSec>
        <mets:fileGrp USE="master">
                <mets:file ID="nla.pic-vn3579081-m" ADMID="file1 mix1 event1 event3 agent1 agent2 agent3 agent4 agent6"
                MIMETYPE="image/tiff" SIZE="133216256" CHECKSUM="5784d1a8dad7c33a5a4f1b27428ef8dc" CHECKSUMTYPE="MD5">
                <mets:FLocat LOCTYPE="URL" xlink:href="http://nla.gov.au/nla.pic-vn3579081-m.tif"/>
                </mets:file>
        </mets:fileGrp>
        …
        <mets:fileGrp USE="derivative">
                <mets:file ID="nla.pic-vn3579081-v" ADMID="file4 mix4 event6 agent5" MIMETYPE="image/jpeg" SIZE="34167"
                CHECKSUM="cec4b919d7615a82eb79f5425a593d0c" CHECKSUMTYPE="MD5">
                <mets:FLocat LOCTYPE="URL" xlink:href="http://nla.gov.au/nla.pic-vn3579081-v.jpg"/>
                </mets:file>
                …
        </mets:fileGrp>
</mets:fileSec>
<mets:structMap>
        <mets:div TYPE="still-image" ADMID="representation1" DMDID="MODS1">
                <mets:fptr FILEID="nla.pic-vn3579081-m"/>
                <mets:fptr FILEID="nla.pic-vn3579081-c"/>
                …
        </mets:div>
</mets:structMap>
```

Figure 9. The Australian METS Profile 1.0 uses the METS *fileSec* and *structMap* and PREMIS *relationship* semantic units to express relationships.

10. What level of object does PREMIS information describe?

Discussion: PREMIS information should be connected to the object or objects to which it relates in the METS structural map. However, it is helpful to explicitly state in a profile in a more human-readable form what level of object PREMIS information describes. A diagram such as Figure 3 in Angela Dappert and Markus Enders' case study, "Using METS, PREMIS, and MODS for Archiving eJournals," may be valuable for describing how different metadata sections fit together in complex structures.[xvii]

Examples: The University of Southampton's RLUK 19th Century Pamphlets Profile states the level at which PREMIS information is recorded:

> Technical metadata [is] at the file level, extracted from standard file information. There will be one instance per file of any kind listed in the fileSec. This metadata will be contained in an amdSec linked to the relevant file element.[xviii]

11. How are PREMIS linking identifiers, IDREFs, and PREMIS identifiers used?

Discussion: The *Guidelines* state that when PREMIS information is wrapped into the METS document, implementers should the use METS ID/IDRefs mechanism to connect files in the *fileSec* with the PREMIS information that relates to those files. If PREMIS information is outside of the METS document, using PREMIS *identifier* semantic units in addition to METS ID/IDRefs is recommended because IDRefs may break and PREMIS identifiers allow the implementer to record more detailed information if necessary. Repositories should document their strategy for linking between elements in PREMIS and METS. Repositories should also document how PREMIS identifiers, if used, are generated and at what level the identifiers are resolvable (locally, globally, or within each METS document).

Examples: All METS profiles analyzed used METS IDs. Roughly half of the profiles also used PREMIS linking identifiers. Only one repository, University of Illinois, used PREMIS schema IDREFs (e.g. RelObjectXmlID). Figure 10 shows example of use of a PREMIS linking identifier from the Australian METS Profile.[xix]

```
<mets:techMD ID="file1">
        <mets:mdWrap MDTYPE="PREMIS">
        <mets:xmlData>
                <premis:object>
                        <premis:objectIdentifier>
                                <premis:objectIdentifierType>uri</premis:objectIdentifierType>
                                <premis:objectIdentifierValue>info:nla/nla.pic-vn3579081-m</premis:objectIdentifierValue>
                        </premis:objectIdentifier>
                        <premis:preservationLevel>supported</premis:preservationLevel>
                        <premis:objectCategory>file</premis:objectCategory>
                        <premis:objectCharacteristics>
                                <premis:format>
                                <premis:formatDesignation>
                                <premis:formatName>image/tiff</premis:formatName>
                                <premis:formatVersion>6.0</premis:formatVersion>
                                </premis:formatDesignation>
                                </premis:format>
                        </premis:objectCharacteristics>
                        <premis:creatingApplication>
                                <premis:creatingApplicationName>Adobe Photoshop CS</premis:creatingApplicationName>
                                <premis:creatingApplicationVersion>7</premis:creatingApplicationVersion>
                                <premis:dateCreatedByApplication>2005-11-03T12:15:59</premis:dateCreatedByApplication>
                        </premis:creatingApplication>
                        <premis:originalName>nla.pic-vn3579081-m.tif</premis:originalName>
                        <premis:storage>
                                <premis:contentLocation>
                                <premis:contentLocationType>filepath</premis:contentLocationType>
                                <premis:contentLocationValue>doss/</premis:contentLocationValue>
                                </premis:contentLocation>
                                <premis:storageMedium>online resource</premis:storageMedium>
                        </premis:storage>
                        <premis:relationship>
                        …
                        </premis:relationship>
                        <premis:linkingEventIdentifier>
                                <premis:linkingEventIdentifierType>internal</premis:linkingEventIdentifierType>
                                <premis:linkingEventIdentifierValue>28902-3</premis:linkingEventIdentifierValue>
                        </premis:linkingEventIdentifier>
                </premis:object>
        </mets:xmlData>
        </mets:mdWrap>
</mets:techMD>
```

Figure 10. Example of use of a PREMIS linking identifier from the Australian METS Profile.

12. How are PREMIS-METS redundancies handled?

Discussion: A necessary result of a design requirement of both PREMIS and METS—namely that both are designed to be independent and modular, so that use of one standard does not depend on use of another—is that PREMIS calls for users to record certain key information that is also recorded in METS. The *Guidelines* lists some of these redundancies (see Table 1 below).[xx]

Table 1. Some redundancies between PREMIS and METS

	PREMIS	METS
SIZE	in *size* under *objectCharacteristics*	an attribute of *file* in the *fileGrp*
CHECKSUM and CHECKSUMTYPE	in *fixity* under *objectCharacteristics*	attributes of *file*
MIMETYPE	in *format* under *objectCharacteristics*	an attribute of *file*

In most cases, the PREMIS semantic unit is more expressive than the METS element and therefore the *Guidelines* recommends using PREMIS or recording the information redundantly in PREMIS and METS, especially if the purpose of the metadata is preservation. The decision to record this information in PREMIS, METS or both should be specified in the METS profile.

Examples: METS profiles currently registered with the Library of Congress showed great variance in handling redundancies; there seemed to be no one particularly popular strategy. For example, the University of California San Diego's Complex Object Profile uses PREMIS to record size and checksum information and uses both METS and PREMIS for mimetype information. The University of Southampton's RLUK 19th Century Pamphlets Profile, on the other hand, records size in both PREMIS and MIX, and checksum and mimetype information in METS.

13. What metadata tools or applications are used?

Discussion: According to the METS Profile Components document, a repository should record any tools used in the creation, transformation, or preservation of its PREMIS or METS metadata in its METS profile. This information is useful to future curators of the data who may not otherwise know that a specific tool was used. If the METS Profile is shared, other repositories could also benefit from knowing that a given tool is available and being used for a particular metadata operation.

Examples: In its ECHO Dep METS Profile for Web Site Captures, the University of Illinois declares that it inherits the tools specified in its parent profile, the ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability, and additionally record using Web Archivists Workbench (see Figure 11 below).[xxi]

```
<tool>
        <name>Web Archivists Workbench (WAW)</name>
                <agency>OCLC</agency>
                <URI>http://ndiipp.uiuc.edu/index.php?option=com_content&amp;task=view&amp;id=19&amp;Itemid=46I</URI>
                <description>
                        <p>The WAW tool from OCLC will produce submission information packages that are conformant to this profile.</p>
                </description> …
</tool>
```

Figure 11. The ECHO Dep METS Profile for Web Site Captures declares its use of Web Archivists Workbench.

[i] *Guidelines for Using PREMIS with METS for Exchange*. Revised September 17, 2008. Available at: http://www.loc.gov/standards/premis/guidelines-premismets.pdf. (Accessed November 9, 2009.) Hereafter referred to as the *Guidelines*.
[ii] Relevant case studies abound. One particularly rich example of a case study is Angela Dappert's and Markus Enders', "Using METS, PREMIS and MODS for Archiving eJournals," *DLib Magazine* September/October 2008. Available at: http://www.dlib.org/dlib/september08/dappert/09dappert.html.
[iii] METS Profile Components: http://www.loc.gov/standards/mets/profile_docs/components.html.
[iv] For highlights of this analysis, see Sally Vermaaten's article in the September/October 2010 issues of D-Lib Magazine.
[v] University of Illinois at Urbana-Champaign's ECHO Dep METS Profile for Web Site Captures: http://www.loc.gov/standards/mets/profiles/00000016.xml.
[vi] The Australian METS Profile: http://www.loc.gov/standards/mets/profiles/00000018.xml.
[vii] The University of Southampton's RLUK 19th Century Pamphlets METS Profile: http://www.loc.gov/standards/mets/profiles/00000024.xml.
[viii] *PREMIS Data Dictionary*. Version 2.0, March 2008. p. 18. Available at: http://www.loc.gov/standards/premis/v2/premis-report-2-0.pdf.
[ix] University of California San Diego's Simple Object Profile: http://www.loc.gov/standards/mets/profiles/00000027.xml.
[x] See note 5.
[xi] University of Illinois Urbana-Champaign's ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability: http://www.loc.gov/standards/mets/profiles/00000015.xml.
[xii] See note 6.
[xiii] University of California San Diego's Electronic Theses and Dissertations Profile: http://www.loc.gov/standards/mets/profiles/00000026.xml.
[xiv] See note 11
[xv] See note 9.
[xvi] See note 6.
[xvii] See note 2.
[xviii] See note 7.
[xix] See note 6.
[xx] *Guidelines*
[xxi] See note 5.