iPRES 2023

Digital Preservation in Disruptive Times

19th International Conference ■ Champaign-Urbana, Illinois ■ September 19–22, 2023

Understanding and Implementing METS

Karin Bredenberg, Aaron Elkiss, Juha Lehtonen
Tuesday, September 19, 9:00 AM - 12:30 PM

Content

- Introduction to METS
  - Background
  - Status of METS
  - Benefits of implementing METS 2
- METS in detail with a focus on METS 2
  - Core Elements
  - METS Profiles
- Implementation
  - METS 1 to METS 2
  - PREMIS in METS
  - Exercise
- Next steps and wrap up

Introduction to METS

Background

What is METS?

- METS = Metadata Encoding and Transmission Standard
- Maintained by the METS Editorial Board
- Schema is hosted at the Library of Congress
- Currently used version 1.12.1
- METS 2.0 beta available

What is METS? (cont'd)

- An XML schema-based specification for encoding "hub" documents for materials whose content is digital.
    - Hub doc draws together dispersed but related files.
    - METS uses XML to provide a vocabulary and syntax for identifying the digital pieces that together comprise a digital entity, for specifying the location of these pieces, and for expressing the structural relationships between them:
        - Content files
        - Descriptive and administrative metadata

METS Editorial Board

The METS Editorial Board is an international group of volunteers committed to maintaining editorial control over METS, its XML schema, the METS Profile XML schema, and official METS documentation. The Board promotes the use of the METS specification, maintains a registry of METS Profiles, and endorses best practices in the use of METS as they emerge. Members represent important communities of interest for METS, including members of the Digital Library Federation, its initial sponsor, and the Library of Congress, its maintenance agency.

METS History

- Originates in Making of America II initiative
  - Making of America II (MOA2) was a Digital Library Federation sponsored initiative that started in 1997. Participants included UCB (lead), Stanford, Penn State, Cornell, and NYPL.
  - GOAL: to create a digital object standard for encoding structural, descriptive and administrative metadata along with primary content
  - RESULT: MOA2.DTD (an XML DTD)

METS History (cont'd)

- UCB Library and CDL adopt MOA2
- Other institutions (LC, Harvard) consider
- Additional needs emerge
  - Support for time-based content
  - More flexibility in Descriptive and Administrative metadata
- MOA2 revised
  - Starting in February 2001 concerned parties meet to review and revise MOA2
  - Outcome: mets.xsd
- METS maintained by international METS Editorial Board

Status of METS

Version 1.x

- Version 1.0 has been released in 2002
- Undergone minor updates in the subsequent years
- Last version 1.12.1 of October 2019, stable and accepted.

Community adoption

- METS has become a community best practice for preserving digital objects.
- Many system implementers are relying on this stable standard.

Version 2.x

- A new major version of METS has been discussed for several years
- Input of community at workshops and conferences
- First presentation of version 2.0 draft at iPRES 2022
- Beta version of schema released:
  - Introduction https://github.com/mets/METS-schema/blob/mets2/METS2.md
  - XML Schema: https://github.com/mets/METS-schema/blob/mets2/v2/mets.xsd

Version 1 vs. version 2

- Version 2 is not backward compatible
- Version 1 is still available and will be maintained
- Existing systems using METS 1 can be updated to support METS 2 or continue to use METS 1
- References to the METS 1 schema will continue to work

Benefits

Main Provisions of METS schema

1. Identifying the files or parts of files that comprise the content of a digital entity, and expressing the structure or structures of this content
2. Linking descriptive and administrative metadata with digital content
3. Wrapping digital content and associated descriptive and administrative metadata as binary data.
4. Wrapping digital content and associated descriptive and administrative metadata as XML data.

Uses of METS

- Transfer syntax
  - Standard for transmitting/exchanging digital objects
  - SIP as described in the Reference Model for an Open Archival Information System (OAIS)
  - DSpace SIP Toolkit uses a mandatory METS document
  - Fedora supports METS as a ingest package
- Functional syntax:
  - Basis for providing end users with the ability to view and navigate digital content and its associated metadata
  - DIP (as in the OAIS)
- Archiving syntax
  - Standard for archiving digital objects
  - Combination with PREMIS (PREservation Metadata: Implementation Strategies)
  - AIP (as in the OAIS)

Development of METS 2: Analysis of user adoption

- The METS board has reviewed user adoption (based on existing METS Profiles and individual feedback) to understand how the community is actually applying the standard.
- Some design decisions and features of METS originate from specific use cases identified in the early phase of developing the standard, but have not been used widely.

Benefits of a new major version

- Simplification
  - Remove unused/outdated aspects
  - Reduce complex structures
- Consistency
  - Align structures of different elements
- Flexibility
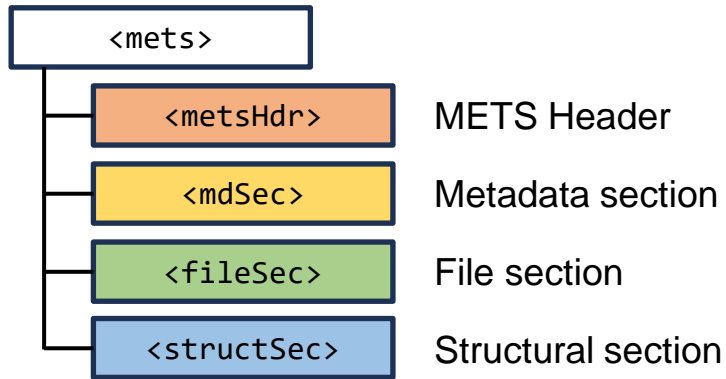 - Do not limit by hard-coded value list, but use recommendations

METS in Detail

Core Elements

METS first level elements

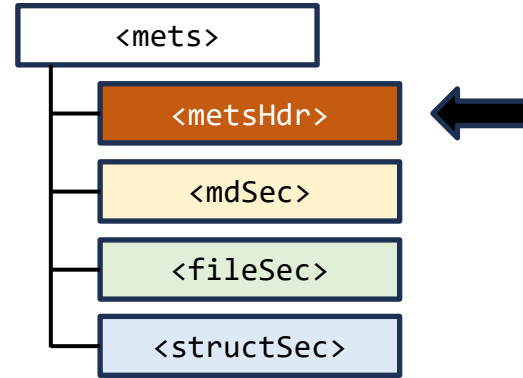|	
`<mets>`	
`<metsHdr>`	METS Header
`<mdSec>`	Metadata section
`<fileSec>`	File section
`<structSec>`	Structural section

METS Header

Records administrative metadata about the
METS document itself such as:
- Author/agent and role
- Alternative identifiers for the METS document
- Creation and update date and times
- Status

```
┌─────────────────────────────┐
│          <mets>             │
└─────────────────────────────┘
   ┌──────────────────────────┐
   │        <metsHdr>         │  ⬅
   └──────────────────────────┘
   ┌──────────────────────────┐
   │         <mdSec>          │
   └──────────────────────────┘
   ┌──────────────────────────┐
   │        <fileSec>         │
   └──────────────────────────┘
   ┌──────────────────────────┐
   │       <structSec>        │
   └──────────────────────────┘
```

Metadata Section

- Can record metadata for the digital object represented by the METS document, or for individual parts of the digital object
- Can include any amount of metadata
- Can take any form: any kind of XML metadata, but not limited to XML.
- May be included internally or referenced externally.
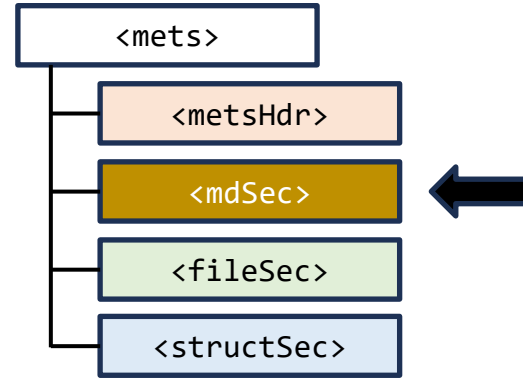
Metadata Examples

- Descriptive metadata: MARC records, Dublin Core metadata, MODS, EAD finding aids, DataCite, etc
- Technical metadata: NISO/MIX image metadata, AES/ebuCore audio metadata, etc.
- Provenance metadata: PREMIS, PROV-XML etc
- Rights metadata: PREMIS rights, ODRL, etc

```
<mets>
    <mdSec>
        <mdGrp>          Technical metadata
            <md>         NISO/MIX image metadata
            <md>
            ...
        <mdGrp>          Descriptive metadata
            <md>         MARC record
            <md>         Dublin Core metadata
            ...
        ...
```
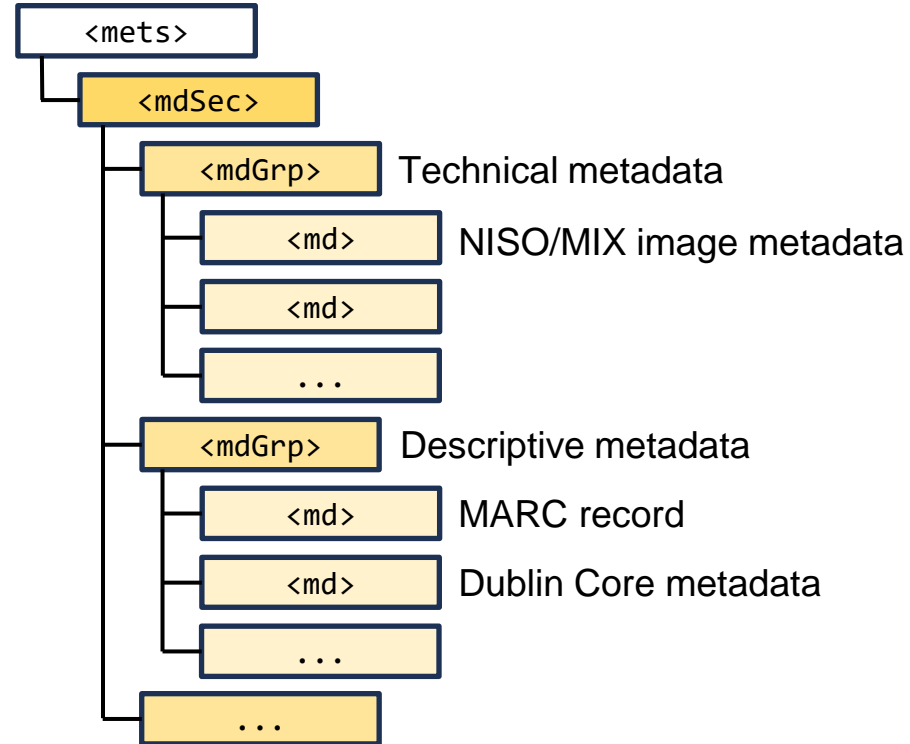
Metadata: METS 1 vs 2

- METS 1
  - different tags for each purpose: `<amdSec>`, `<dmdSec>`, `<techMD>`, `<sourceMD>`, `<digiprovMD>`, `<rightsMD>`
  - Enforced concepts and organization
- METS 2
  - group metadata with `<mdGrp>`
  - metadata records in `<md>`
  - describe the purpose of section or item with @USE

Metadata: METS 1

```xml
<mets>
    <dmdSec ID="dmd1">
        <mdRef MDTYPE="MARC" LOCTYPE="URL"
                xlink:href="http://example.org/123456.marc.xml"/>
    </dmdSec>
    <amdSec>
        <techMD ID="techmd1">
            <mdRef MDTYPE="NISOIMG" LOCTYPE="URL"
                    xlink:href="http://example.org/123456.mix.xml"/>
        </techMD>
        <digiprovMD>...</digiprovMD>
    </amdSec>
</mets>
```
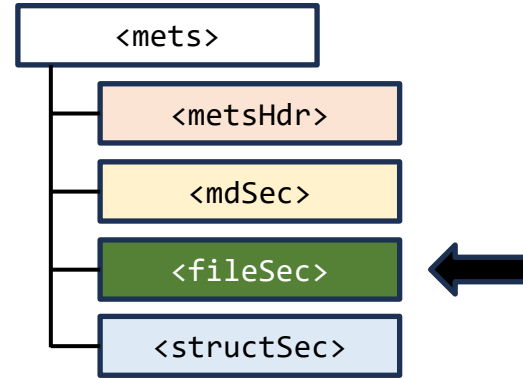
Metadata: METS 2

```xml
<mets>
    <mdSec>
        <mdGrp USE="DESCRIPTIVE">
            <md ID="dmd1">
                <mdRef MDTYPE="MARC" LOCTYPE="URL"
                        LOCREF="http://example.org/123456.marc.xml"/>
            </md>
        </mdGrp>
        <mdGrp USE="ADMINISTRATIVE">
            <md USE="TECHNICAL" ID="techmd1">
                <mdRef MDTYPE="NISOIMG" LOCTYPE="URL"
                        LOCREF="http://example.org/123456.mix.xml"/>
            </md>
            <md USE="DIGIPROV" ID="digiprovmd1">...</md>
        </mdGrp>
    </mdSec>
</mets>
```

File Section

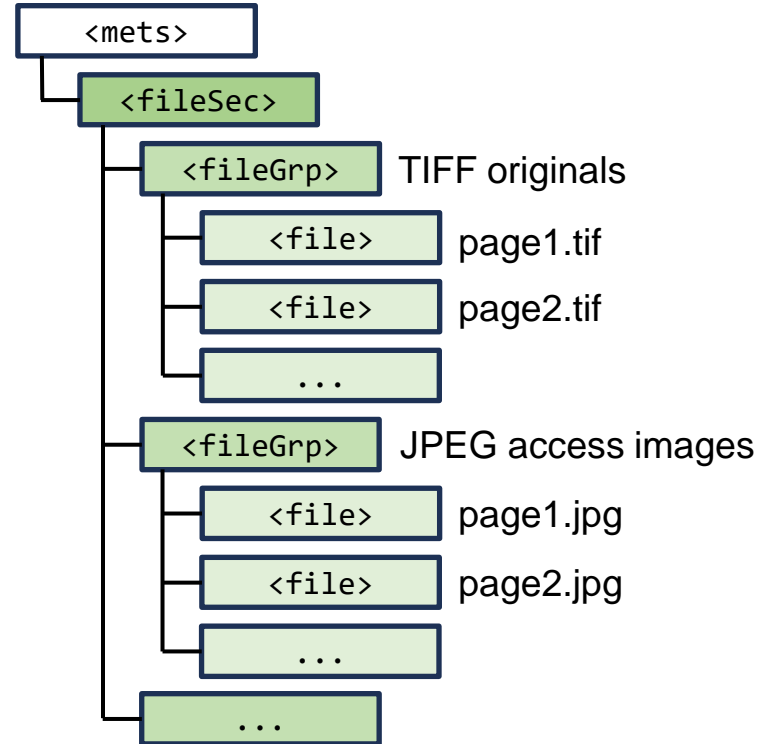- Records all of the content files that make up the digital entity represented by the METS document
- Can link files to metadata

```
<mets>
    <metsHdr>
    <mdSec>
    <fileSec>  ⬅
    <structSec>
```

File Section: Filegroups

Files are organized into File Groups based on the grouping you would like to do. One way is to group by format (tiff, hi-res jpeg, med-res jpeg, gif, etc).

```
<mets>
  <fileSec>
    <fileGrp>          TIFF originals
      <file>           page1.tif
      <file>           page2.tif
      ...
    <fileGrp>          JPEG access images
      <file>           page1.jpg
      <file>           page2.jpg
      ...
    ...
```

File Section Example

```xml
<mets>
    <fileSec>
        <fileGrp MDID="digiprovmd1" USE="Original Images">
            <file ID="original-1" MDID="techmd1">
                <FLocat LOCTYPE="URL" LOCREF="https://example.org/img001.tif" />
            </file>
        </fileGrp>
        <fileGrp MDID="digiprovmd1" USE="Access Images">
            <file ID="access-1">
                <FLocat LOCTYPE="URL" LOCREF="https://example.org/img001.jpg" />
            </file>
        </fileGrp>
    </fileSec>
</mets>
```

Structural Section

- Specifies the (hierarchical) structure of the digital entity represented by the METS document.
- Specifies how the content files (the files listed in the Files Section) fit into this structure.
- More than one structure may be specified. For example: a logical structure and a physical structure

```
<mets>
    <metsHdr>
    <mdSec>
    <fileSec>
    <structSec>  ⬅
```
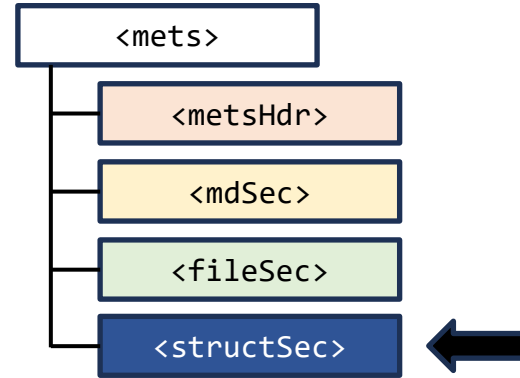
Expressing the Structure

The structural map analyzes a digital object
into a hierarchy of Division (div) elements:

Division (type="photoalbum")
      Division (type="page")
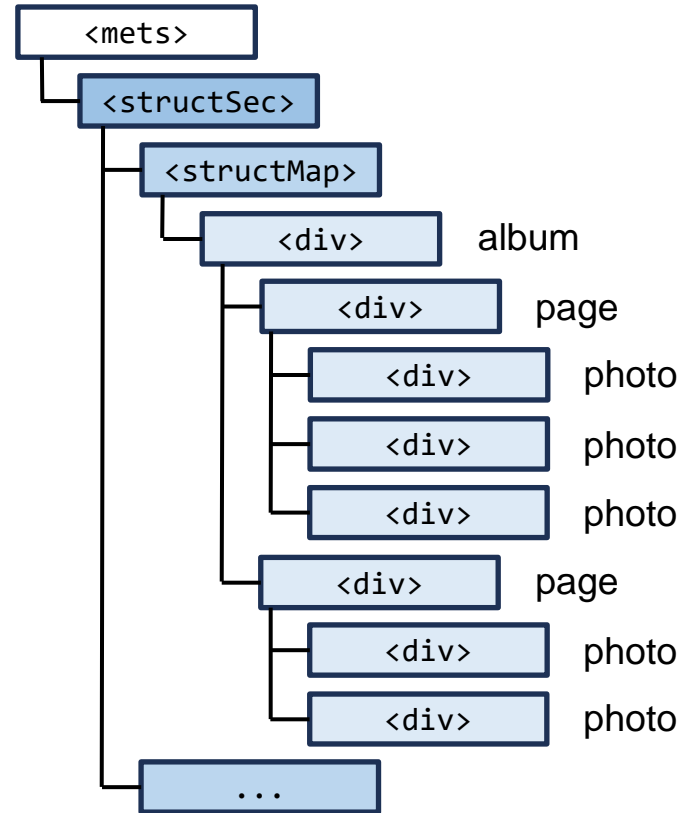   Division (type="photo")
            Division (type="photo")
            Division (type="photo")
      Division (type="page")
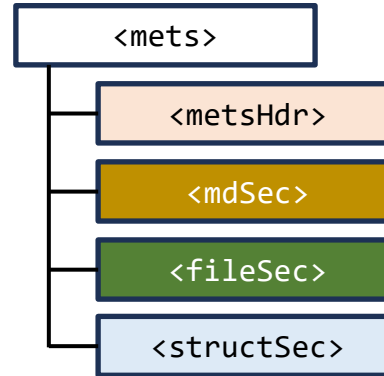   Division (type="photo")
            Division (type="photo")

Structural Map Example

```
<mets>
    <structSec>
        <structMap>
            <div TYPE="photoalbum" MDID="dmd1" LABEL="Commencement 1977">
                <div TYPE="page" ORDERLABEL="1">
                    <div TYPE="photo" LABEL="Graduating Class">
                        <fptr FILEID="original-1"/>
                        <fptr FILEID="access-1"/>
                    </div>
                    <div type="photo" LABEL="School of Engineering">
                        <fptr FILEID="original-2"/>
                        <fptr FILEID="access-2"/>
                    </div>
                </div>
            </div>
        </structMap>
    </structSec>
</mets>
```

Internal or external

- Can include metadata and files or point to them.
- Can refer to an external file, contain the file, or both.
  - External file. The element may point to an external file via a URI.
  - Internal file. The element may itself contain the file as binary data or XML data.

```
<mets>
    <metsHdr>
    <mdSec>
    <fileSec>
    <structSec>
```

Referenced file: example

```
<mets>
    <fileSec>
        <file ID="FILE001" MDID="techmd001">
            <FLocat LOCTYPE="URL" LOCREF="http://dlib.nyu.edu/press/testimg.tif" />
        </file>
    </fileSec>
</mets>
```

Included metadata: example

```
<mets>
    <mdSec>
        <md ID="dmd002" USE="DESCRIPTIVE">
            <mdWrap MIMETYPE="text/xml" MDTYPE="DC"
                    LABEL="Dublin Core Metadata">
                <xmlData xmlns:dc="http://purl.org/dc/elements/1.1/">
                    <dc:title>Alice's Adventures in Wonderland</dc:title>
                    <dc:creator>Lewis Carroll</dc:creator>
                    <dc:publisher>McCloughlin Brothers</dc:publisher>
                    <dc:type>text</dc:type>
                </xmlData>
            </mdWrap>
        </md>
    </mdSec>
</mets>
```

METS 1 vs. METS 2 : xlink

- METS 1 recorded links using XLink, with the xlink:href attribute
- Issues with XLink
  - Not widely adopted
  - Other schemas have moved away (EAD, PREMIS, SVG)
  - Conflicting XLink schemas caused problems
- METS 2 uses locally-defined LOCREF
  - Not required to be a URI

Structural Link and Behavior Sections

- Sections in METS 1 that were not widely used, removed in METS 2
- METS 1 will continue to be available and supported
- `<structLink>`
    - for recording links between parts of the digital object
    - Relies on XLink
    - Mainly for web archiving; WARC has superceded
- `<behaviorSec>`
    - For encoding programmatic behavior related to content
    - Mainly added for old version of Fedora; not widely used

METS Profiles

Profiles

- An XML document
- Defines the rules of how METS is used in a specific context
- One profile can extend another profile
- Profile schema will be updated for METS 2

Profiles

- General information about who to contact regarding the profile and so on
- Element showing if its registered or not
- All the rules regarding the use of METS
  - What metadata groups are expected
  - What file groups are expected
  - What elements should be used and how

Profiles Registration

- Registration at the METS Editorial Board
- 46 registered profiles (September 2023)
- Review period through the METS mailing list
- Can be used by others
- Others can make an new profile extending an existing profile with their use
- Can still register profiles for METS 1
- Process forthcoming for METS 2

Validating

- Current profile schema does not include validation rules
- METS schema only validates basic structure
- Intend to provide sample Schematron validating additional things
- For additional validation you can do implement one of these two options:
  - Use Schematron to implement your restrictions
  - Extend the schema and add your restrictions
- Can extend many elements with custom attributes

Redundancy

- Using METS and other standards sometimes causes redundancy
- Same elements in several standards
- Same elements mandatory in several standards
- Must make decisions regarding redundancy when not mandatory in both standards
- Will discuss best practices of using PREMIS and METS together

Break

Implementation

METS 1 to METS 2

Example Workflow

Does not fully cover all use cases

- Change namespaces (METS 2, remove XLINK) and schemaLocation
- Update PROFILE and CREATEDATE/LASTMODDATE attributes if needed
- Move OTHERROLE/OTHERTYPE values in agent to ROLE/TYPE
- Make different metadata elements uniform
  - Change element names, add USE attributes, move values from OTHERMDTYPE to MDTYPE
- Move metadata elements to a new metadata section
- Change file location attributes in metadata section and file section
- Change metadata reference attributes in file sections and structural maps
- Move structural map inside a new structural section

> METS Editorial Board provides a XSL transformation script.
> See: https://github.com/mets/METS1to2

Namespace, schemaLocation and PROFILE changes

- METS 1:
```
<mets OBJID="01234567-0123-4567-0123-456789abcdef"
      PROFILE="my-profile"
      xmlns="http://www.loc.gov/METS/"
      xmlns:xlink="http://www.w3.org/1999/xlink"
      schemaLocation="…">
```
- METS 2:
```
<mets OBJID="01234567-0123-4567-0123-456789abcdef"
      PROFILE="my-mets2-profile"
      xmlns="http://www.loc.gov/METS/v2"
      schemaLocation="…">
```

- XLINK will not be used anymore, so its namespace is simply removed.

METS Header

- You might want to change CREATEDATE (or LASTMODDATE) attribute to current time.
- METS 1:

```
<metsHdr CREATEDATE="2022-07-06T14:05:00">
    <agent ROLE="CREATOR">
        <name>METS Editorial Board</name>
    </agent>
</metsHdr>
```

- METS 2:

```
<metsHdr CREATEDATE="2023-09-19T11:00:00">
    <agent ROLE="CREATOR">
        <name>METS Editorial Board</name>
    </agent>
</metsHdr>
```

Agent Roles and Types in METS Header

- METS 1:

```
<metsHdr CREATEDATE="2023-09-19T11:00:00">
    <agent ROLE="OTHER" OTHERROLE="MYROLE"
           TYPE="OTHER" OTHERTYPE="MYTYPE">
        <name>METS Editorial Board</name>
    </agent>
</metsHdr>
```

- METS 2:

```
<metsHdr CREATEDATE="2023-09-19T11:00:00">
    <agent ROLE="MYROLE" TYPE="MYTYPE">
        <name>METS Editorial Board</name>
    </agent>
</metsHdr>
```

Metadata Elements

METS 1:

- **<dmdSec … > … </dmdSec>**
- **<techMD … > … </techMD>**
- **<rightsMD … > … </rightsMD>**
- **<sourceMD … > … </sourceMD>**
- **<digiprovMD … > … </digiprovMD>**

METS 2:

- **<md USE="DESCRIPTIVE" … > … </md>**
- **<md USE="TECHNICAL" … > … </md>**
- **<md USE="RIGHTS" … > … </md>**
- **<md USE="SOURCE" … > … </md>**
- **<md USE="PROVENANCE" … > … </md>**

New Metadata Section

- METS 1:

  `<dmdSec>` … **`</dmdSec>`**

  `<amdSec>`

      `<techMD … > … </techMD>`

      `<rightsMD … > … </rightsMD>`

      `<sourceMD … > … </sourceMD>`

      `<digiprovMD … > … </digiprovMD>`

  `</amdSec>`

- METS 2:

  `<mdSec>`

      `<md USE="DESCRIPTIVE" … > … </md>`

      `<md USE="TECHNICAL" … > … </md>`

      `<md USE="RIGHTS" … > … </md>`

      `<md USE="SOURCE" … > … </md>`

      `<md USE="PROVENANCE" … > … </md>`

  `</mdSec>`

Metadata References

- ## METS 1:
  ```
  <mdRef CHECKSUMTYPE="MD5" CHECKSUM="f123456789abcdef0123456789abcde0"
         MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="3.0" LOCTYPE="URL"
         xlink:type="simple" xlink:href="http://example.org/object1.xml" />
  ```
- ## METS 2:
  ```
  <mdRef CHECKSUMTYPE="MD5" CHECKSUM="f123456789abcdef0123456789abcde0"
         MDTYPE="PREMIS:OBJECT" MDTYPEVERSION="3.0" LOCTYPE="URL"
         LOCREF="http://example.org/object1.xml" />
  ```

Other Metadata Types

- ## METS 1:
  ```
  <mdRef MDTYPE="OTHER" OTHERMDTYPE="MYVALUE" … />
  ```
- ## METS 2:
  ```
  <mdRef MDTYPE="MYVALUE" … />
  ```

File Locations

- METS 1:
  ```
  <FLocat LOCTYPE="URL"
          xlink:type="simple" xlink:href="http://example.org/myfile1.pdf" />
  ```
- METS 2:
  ```
  <FLocat LOCTYPE="URL" LOCREF="http://example.org/myfile1.pdf" />
  ```

- If LOCTYPE="OTHER" and OTHERLOCTYPE are used, use only LOCTYPE.

Metadata References in File Section

- METS 1:

  ```
  <file ID="file-001" DMDID="desc-001" ADMID="tech-001 tech-002">
  ```

- METS 2:

  ```
  <file ID="file-001" MDID="desc-001 tech-001 tech-002">
  ```

Metadata References in Structural Maps

- METS 1:

  `<div DMDID="desc-002" ADMID="prov-001 prov-002">`

- METS 2:

  `<div MDID="desc-002 prov-001 prov-002">`

Relocate Structural Map to a New Structural Section

- ## METS 1:

  ```
  <structMap>
       …
  </structMap>
  ```

- ## METS 2:

  ```
  <structSec>
       <structMap>
            …
       </structMap>
  </structSec>
  ```

Additional Issues

- "I still have attribute X="OTHER" and OTHERX="MYVALUE" in element Y."
- "I still have attribute ADMID/DMDID in element Y."
- "I have added my metadata into METS instead of referencing to another file."
- "I still want to keep the old metadata hierarchy and have e.g. multiple administrative metadata sections."
- "I need other use categories than DESCRIPTIVE, TECHNICAL, etc."
- "I have multiple structural maps."
- "I have a complex and deep structure in file section."

Attributes X and OTHERX

"I still have attribute X="OTHER" and OTHERX="MYVALUE" in element Y."

- `<agent>`:
  - `ROLE="OTHER" OTHERROLE="MYVALUE"`   `ROLE="MYVALUE"`
  - `TYPE="OTHER" OTHERTYPE="MYVALUE"`   `TYPE="MYVALUE"`
- `<FLocat>`:
  - `LOCTYPE="OTHER" OTHERLOCTYPE="MYVALUE"`   `LOCTYPE="MYVALUE"`
- `<mdRef>`:
  - `LOCTYPE="OTHER" OTHERLOCTYPE="MYVALUE"`   `LOCTYPE="MYVALUE"`
  - `MDTYPE="OTHER" OTHERMDTYPE="MYVALUE"`   `MDTYPE="MYVALUE"`
- `<mdWrap>`:
  - `MDTYPE="OTHER" OTHERMDTYPE="MYVALUE"`   `MDTYPE="MYVALUE"`
- `<mptr>`:
  - `LOCTYPE="OTHER" OTHERLOCTYPE="MYVALUE"`   `LOCTYPE="MYVALUE"`

METS Editorial Board provides suggested attribute values.
See: https://github.com/mets/METS-schema/wiki/METS2-Suggested-Attribute-Values

ADMIDs and DMDIDs to MDID

"I still have attribute ADMID/DMDID in element *Y*."

- `<fileGrp>`:
  - `ADMID="myref-1"`   ○ `MDID="myref-1"`
- `<file>`:
  - `ADMID="myref-1"`   ○ `MDID="myref-1"`
  - `DMDID="myref-1"`   ○ `MDID="myref-1"`
- `<stream>`:
  - `ADMID="myref-1"`   ○ `MDID="myref-1"`
  - `DMDID="myref-1"`   ○ `MDID="myref-1"`
- `<div>`:
  - `ADMID="myref-1"`   ○ `MDID="myref-1"`
  - `DMDID="myref-1"`   ○ `MDID="myref-1"`
- `<area>`:
  - `ADMID="myref-1"`   ○ `MDID="myref-1"`

Using mdWrap instead of mdRef
"I have added my metadata into METS instead of referencing to another file."

- It is still possible to add metadata directly into METS.

- METS 1:
  ```
  <mdWrap MDTYPE="OTHER" OTHERMDTYPE="AUDIOMD" MDTYPEVERSION="2.0"> … </mdWrap>
  ```
- METS 2:
  ```
  <mdWrap MDTYPE="AUDIOMD" MDTYPEVERSION="2.0"> … </mdWrap>
  ```

Grouping of Metadata Elements

"I still want to keep the old metadata hierarchy… "

- ## METS 1:

```
<dmdSec … > … </dmdSec>
<dmdSec … > … </dmdSec>
<amdSec ID="amd-001">
    <techMD … > … </techMD>
    <digiprovMD … > … </digiprovMD>
</amdSec>
<amdSec ID="amd-002">
    <techMD … > … </techMD>
    <digiprovMD … > … </digiprovMD>
</amdSec>
```

- ## METS 2:

```
<mdSec>
    <mdGrp USE="DESCRIPTIVE">
        <md … > … </md>
        <md … > … </md>
    </mdGrp>
    <mdGrp USE="ADMINISTRATIVE" ID="amd-001">
        <md USE="TECHNICAL" … > … </md>
        <md USE="PROVENANCE" … > … </md>
    </mdGrp>
    <mdGrp USE="ADMINISTRATIVE" ID="amd-002">
        <md USE="TECHNICAL" … > … </md>
        <md USE="PROVENANCE" … > … </md>
    </mdGrp>
</mdSec>
```

USE Attribute Is Not Limited
"I need other use categories than DESCRIPTIVE, TECHNICAL, etc."

- METS 2:

```
<mdGrp USE="MYVALUE">
    <md USE="MYVALUE2"> … </md>
    …
</mdGrp>
```

- Example:

```
<mdGrp USE="PRESERVATION">
    …
</mdGrp>
```

METS Editorial Board provides suggested attribute values.
See: https://github.com/mets/METS-schema/wiki/METS2-Suggested-Attribute-Values

Multiple Structural Maps

"I have multiple structural maps."

- METS 1:
  ```
  <structMap … > … </structMap>
  <structMap … > … </structMap>
  <structMap … > … </structMap>
  ```
- METS 2:
  ```
  <structSec>
      <structMap … > … </structMap>
      <structMap … > … </structMap>
      <structMap … > … </structMap>
  </structSec>
  ```

- Each <structMap> needs to be transformed to comply with METS 2 specification.
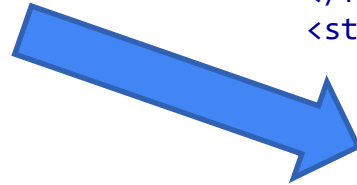
Structural Data from File Section to a Structural Map

"I have a complex and deep structure in file section."

- ● METS 1 (in some use cases):

```
<fileSec … >
    <fileGrp … >
        <file ID="file-1" … > …
        <fileGrp … >
            <file ID="file-2" … > …
            <file ID="file-3" … > …
        …
</fileSec>
```

- ● METS 2:

```
<fileSec … >
    <file ID="file-1" … > …
    <file ID="file-2" … > …
    <file ID="file-3" … > …
    …
</fileSec>
<structSec>
    <structMap TYPE="PHYSICAL">
        <div … >
            <fptr FILEID="file-1" … />
            <div … >
                <fptr FILEID="file-2" … />
                <fptr FILEID="file-3" … />
            …
    </structMap>
</structSec>
```

Using PREMIS and other metadata standards in METS

Placement

- With METS 2 all other metadata standards is placed in the metadata section

```xml
<md USE="DESCRIPTIVE" ID="dmd001">
    <mdRef MIMETYPE="application/xml" MDTYPE="EAD"
        LABEL="Berol Collection Finding Aid"
        LOCTYPE="URI" LOCREF="urn:x-nyu:fales1735"/>
</md>
```

A special look at PREMIS in METS

The PREMIS standard

- International de-facto standard for metadata to support the preservation of digital objects and ensure their long-term usability.
    - Information you need to know for preserving digital objects
      Preservation Metadata: Implementation Strategies
- Developed by an international team of experts.
- Implemented in digital preservation projects around the world.
- Incorporated into commercial and open-source digital preservation tools and systems.

The PREMIS standard parts

- Data Dictionary (PREMIS 3.0)
  - http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf
  - Version 3 – major release
- XML schema v3.0
  - http://www.loc.gov/standards/premis/premis.xsd
- OWL ontology
- Supporting documentation

Scope

What PREMIS DD is:
- Common data model for organizing/thinking about preservation metadata
- Standard for exchanging information packages between repositories
- Implementable
- Technically neutral
- Core metadata

What PREMIS DD is not:
- Out-of-the-box solution
- All needed metadata
- Lifecycle management of objects outside repository
  - increasing support for integration with outside
- Rights management standard
 - strong support for rights statements

The PREMIS Data Model

- Data model includes:
  - Entities: "things" relevant to digital preservation that are described by preservation metadata (Intellectual Entities, Objects, Events, Rights, Agents)
  - Properties of Entities (semantic units)
  - Relationships between Entities

- Why have a data model?
  - Organizational convenience (for development and use)
  - Useful framework for distinguishing applicability of semantic units across different types of Entities and different types of Objects
  - But: not a formal entity-relationship model; not sufficient to design databases

PREMIS 3 Entities
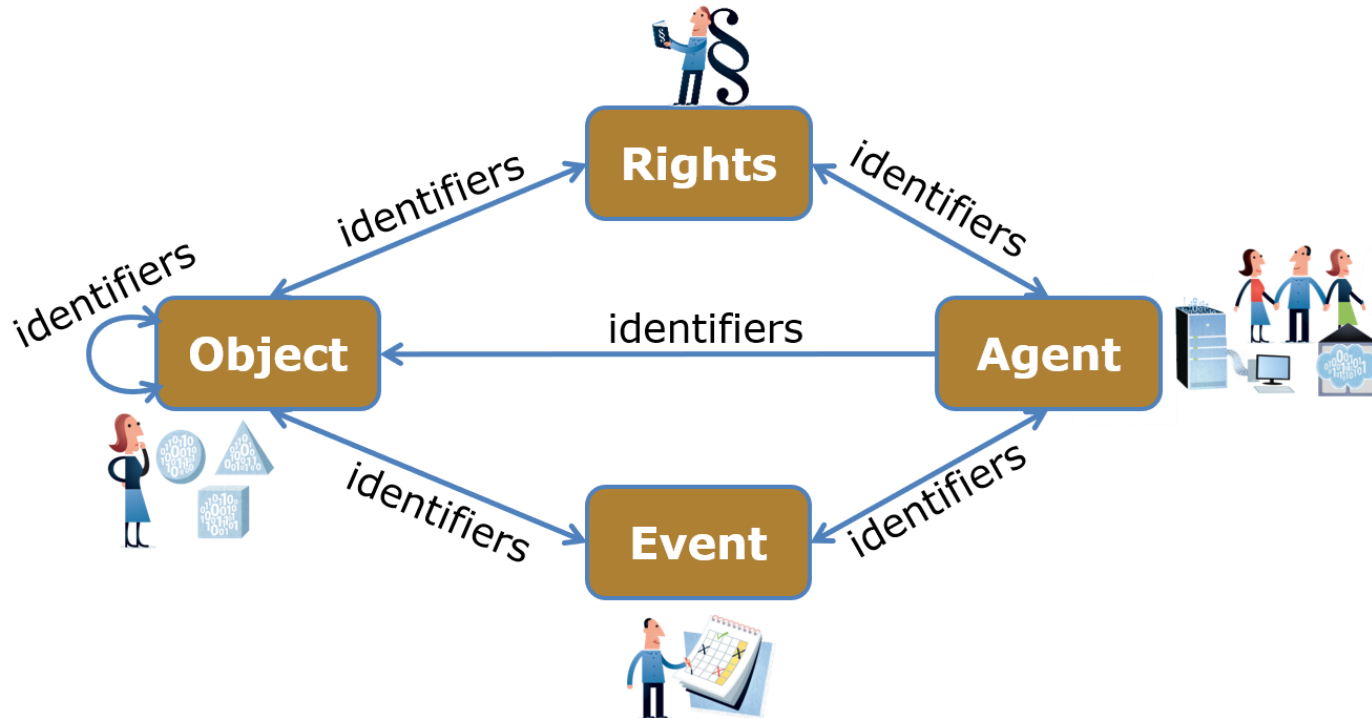
PREMIS 3 Entities
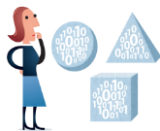
object
Are what repository actually preserves
Different types of objects: 'file', 'representation',
'bitstream' or 'intellectual entity'

event
Aggregates metadata about actions.
Contains e.g. event identifier, event type (creation, migration, …),
and other event details.

agent
Can have different roles in relation to an event or objects.
Can be persons, organisations, programs etc.
Contains identifier and possibly other information like name and type

rights
Describes rights and permissions specifically related to digital preservation.
Contains e.g. rights statements and possible extensions with other metadata
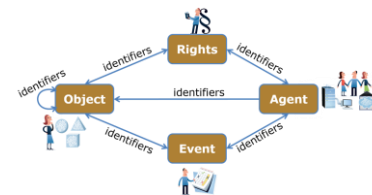
Issues in using PREMIS with METS

- Flexibility of METS requires implementation decisions:
  - Use PREMIS container or separate packages?
  - Which METS sections to use
    - In METS 1 the sections are predefined
    - In METS 2 you need to decide about typing of the metadata sections
  - Whether to record elements redundantly in PREMIS and METS
  - How to record elements that are also part of a format specific technical metadata schema (e.g. MIX)
  - Where to store structural relationships?
  - How to deal with locally controlled vocabularies

PREMIS in METS 1

- You can't put all PREMIS metadata directly under amdSec
- What sections to use for PREMIS metadata?
  - Alternative 1
    - Object (Representation, File and Bitstream) in techMD
      Object (Environment Intellectual Entity) in techMD
      Object (Intellectual Entity) in dmdSec
    - Event in digiProvMD
    - Rights in rightsMD
    - Agent with event or rights
  - Alternative 2
    - Everything in digiProvMD

PREMIS in METS 1 guideline

Local decisions may vary depending on processing model

http://www.loc.gov/standards/premis/guidelines2017-premismets.pdf

PREMIS in METS 2

- You can't put all PREMIS metadata directly under a md
  (you need a mdRef or a mdWrap)
- You can group metadata with the same use in a mdGrp
- What sections to use for PREMIS metadata?
  - Alternative 1
    - Object (Representation, File and Bitstream) in a md with USE="TECHNICAL"
      Object (Environment Intellectual Entity) in a md with USE="TECHNICAL"
      Object (Intellectual Entity) in a md with USE="DESCRIPTIVE"
    - Event in a md with USE="PROVENANCE"
    - Rights in a md with USE="RIGHTS"
    - Agent with event or rights
  - Alternative 2
    - Everything in a md with USE="PROVENANCE"

PREMIS in METS 2 guideline

The guideline will be updated in cooperation with the PREMIS EC

- The md/mdGrp element with the @USE
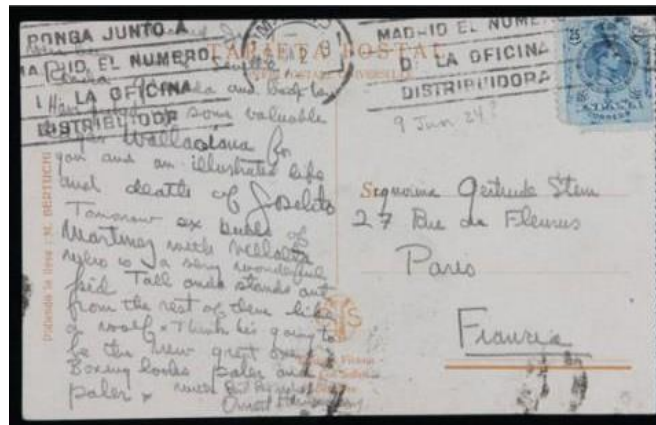- As before the mdRef/mdWrap element with the @MDTYPE

Exercise

The Postcard

Information about the postcard

- Each side digitized as separate hi-res TIFF images along with derived PNG thumbnail images
- A TEI transcription of the text written on the back
- MODS descriptive metadata record for the postcard
- Basic technical metadata for all files is available: format, size, checksum

Folder and the files for the postcard:

Postcard_Hemingway_001

- front.tif
- back.tif
- front.png
- back.png
- back.tei
- mods.xml

Elements to use

METS 2 White Paper: https://github.com/mets/METS-schema/blob/mets2/METS2.md

METS 2 Schema Documentation: https://mets.github.io/METS_v2_Docs/mets.html
More Resources: https://github.com/mets/METS-schema

Support and the METS community

Where to find support

- METS homepage: http://www.loc.gov/standards/mets/
- METS GitHub: https://github.com/mets
- METS listserv: https://listserv.loc.gov/cgi-bin/wa?SUBED1=mets&A=1 and METS@LISTSERV.LOC.GOV

Finding METS tools

- COPTR Page:
  https://coptr.digipres.org/index.php/METS_(Metadata_Encoding_and_Transmission_Standard)

Conformance

How to conform?

- Create a profile
- (Register profile)
- Remember to use the mandatory elements (StructMap)

- Profile for METS 2 is in development

Next steps and wrap up

Resources - METS 2

- https://github.com/mets/METS-schema
- Overview and Tutorial
- Schema Documentation
- Suggested Attribute Values
- Description of changes from METS 1
- XSL transformation from METS 1

Next Steps

- What will happen next with METS 2?
- Update METS Profile schema
- Update METS Primer

Round table discussion for institutional plans

Wrap up

Today we have:

- Gotten to know the elements of METS
- Seen the differences between METS 1 and METS 2
- Done an exercise to create a METS file for a digital object
- In short, we have started our journey into understanding and using METS

Thank you for today!

Karin Bredenberg

Aaron Elkiss

Juha Lehtonen