

METS Awareness Training

An Introduction to METS



Digital libraries – the need for a metadata standard

- Digitisation technology now well established and well-understood
 - Standards for digitisation processes have settled down and are widely recognised
 - Still a disparity in approaches to metadata - no 'MARC standard' for the digital library
-
-

The lack of a standard – what it means...

- poor cross-searching
- limited interchange facilities
- metadata tied to proprietary packages
- consequent obsolescence and costs of conversion
- little chance of a 'hybrid library'



What is needed?

- A standard for metadata content :
analogous to AACR2
 - A standardised framework for holding
and exchanging metadata :
analogous to the MARC record
 - METS is designed to fulfil the latter
function
-
-

METS to the rescue!

- Produced by Library of Congress Standards Office and Digital Library Federation
 - Written in XML
 - Provides **framework** for holding all types of metadata for digital object
 - Does not prescribe **content** of metadata, but recommends a number of schemes for this
-
-

METS and OAIS

- METS was designed to function within the OAIS framework
 - It can act as a Submission Information Package (SIP) Delivery Information Package (DIP), providing a standardized transfer syntax
 - It acts as an Archival Information Package (AIP) for storage and preservation
-
-

An overview of the METS file

- Generally one METS file corresponds to one digital object (which may incorporate many files)
 - All metadata (descriptive, administrative and structural) encoded in single document
 - Each type is held in a separate section, linked by identifiers
 - All metadata and external data (eg. images, text, video) is either referenced from METS file or can be held internally
-
-

The inside of a METS file

METShheader

fileSec

file inventory

dmdSec

descriptive metadata

admSec

administrative metadata

behaviorSec

behaviour metadata

structMap

structural map

Title Page

title page

<div LABEL="Title Page">

Preface

page i

page ii

<div LABEL="Preface">

Chapter 1

page 1

page 2

page 3

page 4

page 5

<div LABEL="Chapter 1">

<div LABEL="Page 1">

<div LABEL="Page 2">

<div LABEL="Page 3">

<fpnr FILEID="xxx"/>

Chapter 2

page 7

page 8

page 9

<div LABEL="Chapter 2">

```
<structMap>

  <div ID="munahi010-aaa-div.1" LABEL="Section 1">
    <div ID="munahi010-aaa-div.1.1" LABEL="Plate 1">
      <fptr FILEID="munahi010-aaa-fgrp-0001"/>
    </div>
    <div ID="munahi010-aaa-div.1.1" LABEL="Plate 2">
      <fptr FILEID="munahi010-aaa-fgrp-0002"/>
    </div>
  </div>

</structMap>
```

The inside of a METS file

METShheader

fileSec

file inventory

dmdSec

descriptive metadata

admSec

administrative metadata

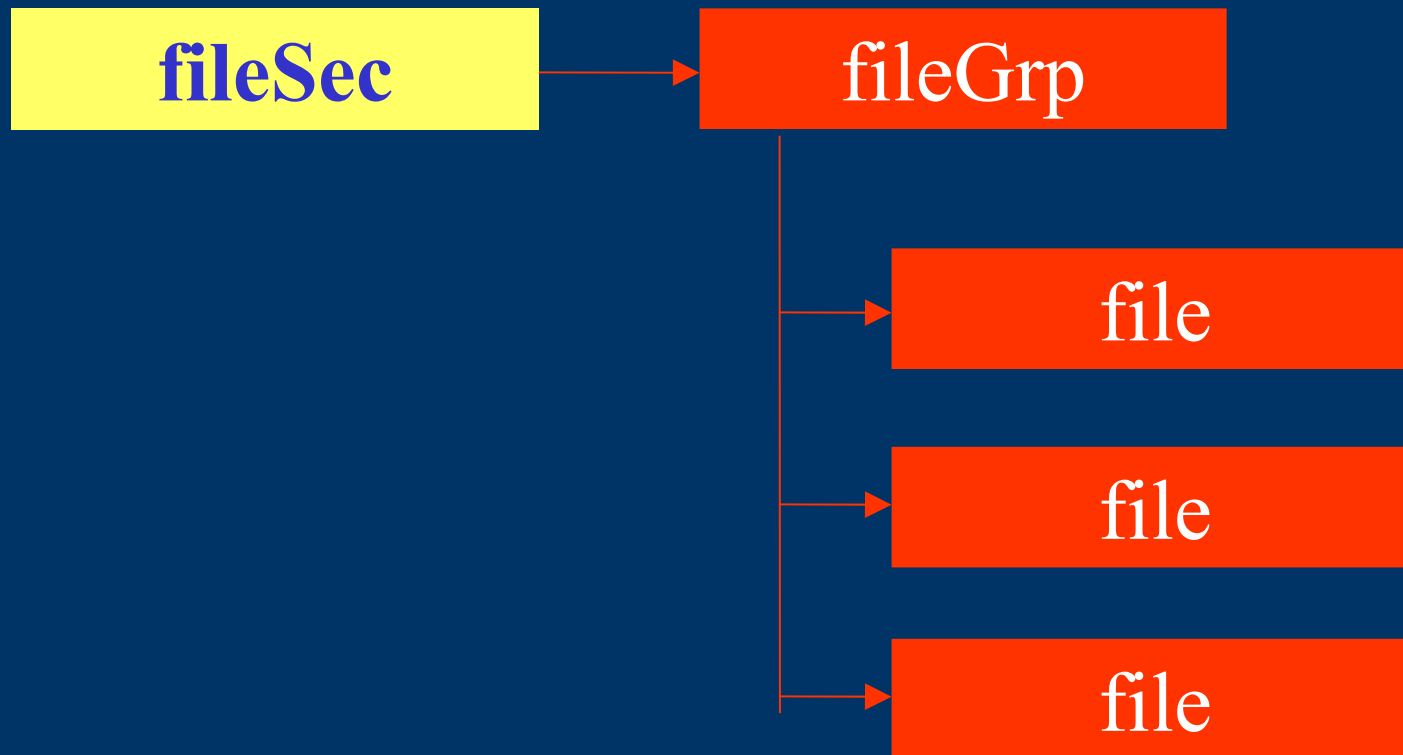
behaviorSec

behaviour metadata

structMap

structural map

<fileSec>



```
<fileGrp ID="munahi010-aaa-fgrp-0001">

  <file GROUPLD="0" ID="munahi010-aaa-0001-0"
MIMETYPE="image/tiff" ADMID="munahi010-aaa-tmd-0001-0">
    <FLocat LOCTYPE="URL"
xlink:href="file://hfs.ox.ac.uk/data/odl/munahi010/digObjects/aaa
/0/munahi010-aaa-0001.tiff"/>
  </file>

  <file GROUPLD="6" ID="munahi010-aaa-0001-6"
MIMETYPE="image/jpeg" ADMID="munahi010-aaa-tmd-0001-6">
    <FLocat LOCTYPE="URL"
xlink:href="http:odl/munahi010/digObjects/aaa/6/munahi010-aaa-
0001-6.jpg"/>
  </file>

  <file GROUPLD="3" ID="munahi010-aaa-0001-3"
MIMETYPE="image/jpeg" ADMID="munahi010-aaa-tmd-0001-3">
    <FLocat LOCTYPE="URL"
xlink:href="http:odl/munahi010/digObjects/aaa/3/munahi010-aaa-
0001-3.jpg"/>
  </file>

</fileGrp>
```

The inside of a METS file

METShheader

fileSec

file inventory

dmdSec

descriptive metadata

admSec

administrative metadata

behaviorSec

behaviour metadata

structMap

structural map

Descriptive and administrative metadata

- Descriptive and administrative metadata may be handled in two ways:
 - embedding directly within the METS file within an `<mdWrap>` element
 - being held in an external file and referenced from the METS file using an `<mdRef>` element
-
-

```
<mdWrap MIMETYPE="text/xml" MDTYPE="MODS" LABEL="MODS Metadata">
  <xmlData>
    <mods:mods>
      <mods:titleInfo>
        <mods:title>Cobbett's parliamentary history of England, from
          the Norman Conquest, in 1066 to the year, 1803 : from which
          last-mentioned epoch it is continued downwards in the work
          entitled, &quot;The parliamentary
          debates&quot;</mods:title>
      </mods:titleInfo>
      <mods:titleInfo type="alternative">
        <mods:title>Cobbett's Parliamentary History -
          volume 2</mods:title>
      </mods:titleInfo>
      <mods:name>
        <mods:namePart>$aGreat Britain. Parliament.</mods:namePart>
        <mods:role>
          <mods:roleTerm type="code"
            authority="marcrelator">spn</mods:roleTerm>
        </mods:role>
      </mods:name>
    </mods:mods>
  </xmlData>
</mdWrap>
```



```
<amdSec ID="munahi010-aaa-amd-0001">  
  <techMD ID="munahi010-aaa-tmd-0001-0">  
    <mdRef MDTYPE="MIX" LOCTYPE="URL"  
      xlink:href="../../munahi010-aaa-0001-0.xml"/>  
  </techMD>  
</amdSec>
```

IDs and METS

- METS uses IDs to express the relations between its component parts
- A coherent system of identifiers is therefore essential

Project ID	munahi010
Item ID	munahi010-aaa
Technical metadata	munahi010-aaa-tmd-0001
File groups	munahi010-aaa-fgrp-0001
File IDs	munahi010-aaa-0001-3
divs	munahi010-aaa-div.1

What to put in a METS file?

- METS does not prescribe the content (particularly the descriptive metadata) which it can contain
- However, the METS board does endorse some schemas as recommended for use with METS:-

Descriptive Metadata

- Dublin Core
- MODS (Metadata Object Description Schema)
- MARCXML MARC 21 Schema (MARCXML)

Administrative Metadata

- Schema for Technical Metadata for Text (NYU)
 - Library of Congress Audio-Visual Prototyping Project
 - NISO Technical Metadata for Digital Still Images (MIX)
 - METS Schema for Rights Declaration
-
-

METS Profiles

- METS is very flexible in its application – there are multiple ways of encoding everything:-
 - metadata and data can be embedded or referenced
 - any scheme can be used for this metadata
 - file inventory can be organised in multiple ways (by referenced object, by type of file etc)
 - This all reduces interchangeability of METS records.
-
-

METS Profiles (cont.)

- This can be countered by METS Profiles:-
 - XML documents describing application of METS in a given project/institution
 - follows METS Profile schema and each profile has to validate against it
 - registered with central repository at Library of Congress
 - Profiling is essential for interoperability of OAI packages
-
-