

# Harvard METS Format for Page-Turned Objects

Rev. 2005-02-11 (sa)

Rev. 2005-03-02 (sm) Updated note for <mods:identifier> to indicate that finding aids must have URN, not EADID.

Rev. 2005-03-07 (sm) Updated notes for <FLocat> to indicate proper use for deposited/pre-deposited files.

Rev 2005-08-24 (sm) Updated note for <div type=page> to reflect optional uses of roman numerals and text for the ORDER and ORDERLABEL attribute.

Declaration	<?xml ... ?>	
Attributes	version	"1.0"
	encoding	"utf-8"
	standalone	"no"
Obligation	Mandatory	
Repeatable	No	

Element	<mets>	
Attributes	xmlns	"http://www.loc.gov/METS/"
	xmlns:pds	"http://hul.harvard.edu/ois/xml/ns/pds"
	xmlns:xlink	"http://www.w3.org/1999/xlink"
	xmlns:xsi	"http://www.w3.org/2001/XMLSchema-instance"
	xsi:schemaLocation	"http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd http://hul.harvard.edu/ois/xml/ns/pds http://hul.harvard.edu/ois/xml/xsd/pds/ pds.xsd"
	LABEL	Concatenation of work creator, title, and (optionally) other relevant publication information (e.g., edition), separated by appropriate AACR2 punctuation
	PROFILE	"HULPDS"
	TYPE	"PAGEDOBJECT"
Content	<dmdSec ID="C0"> <dmdSec ID="C1"> <dmdSec ID="In"> <fileSec> <structMap>	
Obligation	Mandatory	
Repeatable	No	
Conversion from MOA2	The citation node MOA2 <ArchObj NODETYPE="citation" LABEL="label"> maps to the METS <mets LABEL="label">.	

A citation node MOA2 is identified by the root element <ArchObj NODETYPE="citation"> or <ArchObj NODETYPE="citationleaf" ' >; an intermediate node is identified by <ArchObj NODETYPE="intermediate">; a leaf node is identifier by <ArchObj NODETYPE="leaf"> or <ArchObj NODETYPE="citationleaf">. Note that a MOA2 node can be simultaneously both a citation and leaf node, in which case it is identified as a "citationleaf".

Element	<dmdSec>	
Description	Descriptive metadata (author, title, etc.) about the work	
Attributes	ID	"C0"
	GROUPID	"CITATION"
Content	<mdWrap MDTYPE="MODS">	

Obligation	Mandatory
Repeatable	No

Element	<dmdSec>	
Description	PDS-specific metadata (date range, function status, header label) about the work	
Attributes	ID	"C1"
	GROUPID	"CITATION"
Content	<mdWrap MDTYPE="OTHER">	
Obligation	Optional	
Repeatable	No	

Element	<dmdSec>	
Description	Date range metadata about an intermediate <div>	
Attributes	ID	"In", where <i>n</i> is a unique number
	<mdWrap MDTYPE="OTHER">	
Obligation	Optional	
Repeatable	Yes	

Element	<mdWrap>	
Description	Container element for MODS descriptive metadata about the work	
Attributes	MDTYPE	"MODS"
	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for PDS-specific metadata about the work or intermediate sections	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"HULPDS"
Content	<xmlData>	
Obligation	Optional	
Repeatable	No	

Element	<xmlData>	
Content	<mods:mods> or <pds:pds>	
Obligation	Mandatory	
Repeatable	No	

Element	<mods:mods>	
Attributes	xmlns:mods	"http://www.loc.gov/mods/v3"
	xsi:schemaLocation	"http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-0.xsd"
	version	"3.0"
Content	<mods:name> <mods:titleInfo> <mods:identifier>	
Obligation	Mandatory	
Repeatable	No	

Element	<mods:name>	
Content	<mods:displayForm>	
Obligation	Optional	
Repeatable	No	

Element	<mods:displayForm>	
Content	The personal or corporate author of the work	
Obligation	Mandatory	
Repeatable	No	
Conversion from MOA2	The citation node MOA2 <Creator NameType="cn" Role="author">name</Creator> maps to the METS <mods:displayForm>name</mods:displayForm>	

Element	<mods:titleInfo>	
Content	<mods:title>	
Obligation	Mandatory	
Repeatable	No	

Element	<mods:title>	
Content	The title of the work	
Obligation	Mandatory	
Repeatable	No	
Conversion from MOA2	The citation node MOA2 <Title>title</Title> maps to the METS <mods:title>title</mods:title>.	

Element	<mods:identifier>	
Attributes	type	"hollis" or "uri" or "local"
	displayLabel	(Optional) Descriptive text
Content	<xmlData>	
Obligation	Optional	
Repeatable	Yes	
Conversion from MOA2	The citation node MOA2 <LocalID LocalIDType="type">id</LocalID> maps to the METS <mods:identifier type="type">id</mods:identifier>.  Note that the MOA2 type attribute may have a parenthetic prefix, e.g., "(MH) HOLLIS", which should be removed, e.g., "HOLLIS".	

Note	Any reference to an OASIS finding aid should be of type="URI", and use a NRS URN. EG: <mods:identifier displayLabel="OASIS finding aid" type="uri">http://nrs.harvard.edu/urn-3:ARB:ajp00007</mods:identifier>
------	---

Element	<pds:pds>
Description	Additional metadata properties necessary for PDS functions
Content	<pds:date> <pds:function> <pds:header>
Obligation	Optional
Repeatable	No
Note	The <pds:function> and <pds:header> elements can only appear within the PDS-specific metadata element, <dmdSec ID="C1">.  At least one of <pds:date>, <pds:function>, or <pds:header> is required; if none are present, then the <pds:pds> container element, and its parent elements, <xmlData>, <mdWrap>, <dmdSec>, should be suppressed.

Element	<pds:date>
Description	Date range associated with the work
Content	<pds:fromDate> <pds:toDate>
Obligation	Optional
Repeatable	No
Note	At least one of <pds:fromDate> or <pds:toDate> is required; if neither is present, then the <pds:date> container element should itself be suppressed.

Element	<pds:fromDate>
Content	"yyyy" or "yyyy-mm" or "yyyy-mm-dd" representing the beginning date range
Obligation	Optional
Repeatable	No
Conversion from MOA2	The MOA2 <coreDate beginDateNorm="date"> maps to the METS <pds:fromDate>date</pds:fromDate>.

Element	<pds:toDate>
Content	"yyyy" or "yyyy-mm" or "yyyy-mm-dd" representing the ending date range
Obligation	Optional
Repeatable	No
Conversion from MOA2	The MOA2 <coreDate endDateNorm="date"> maps to the METS <pds:toDate>date</pds:toDate>.

Element	<pds:function>	
Description	Flag controlling status of a specific PDS function	
Attributes	status	"true" or "false"
Content	"goto" or "viewocr"	
Obligation	Optional	
Repeatable	Yes	
Note	In general, unless the object owner has explicitly requested that one or the other of these functions be suppressed, this element should be present with a status value of "true".	

Element	<pds:header>	
Description	Curatorially-assigned text string used as the running header of PDF versions of the page-turned object; typically, this will be the repository name, e.g., "Harvard University, Houghton Library"	
Content	Header text string	
Obligation	Optional	
Repeatable	No	

Element	<fileSec>	
Description	Inventory of all of the files that make up the digital representation of the work	
Content	<fileGrp>	
Obligation	Mandatory	
Repeatable	No	

Element	<fileGrp>	
Description	Container for all of the files of a given MIME type	
Content	<file>	
Attributes	ID	The MIME type of the all of the files included in this group, typically this will be: image-gif, image-jpeg, image-jp2, image-jpx, image-tiff, or text-plain
Obligation	Mandatory	
Repeatable	No	
Note	If there are more than one <fileGrp> elements, no specific order is required.	
Conversion from MOA2	The leaf node MOA2 <fileGrp> (and the specific <file> elements in it) map to the METS <fileGrp>. In the MOA2 the HARVMIME attribute is specified at the level of the <file>, not the <fileGrp>. The METS ID attribute should be assigned an appropriate MIME type. A MOA2 HARVMIME value of the form "x-archival" (where "x" is "color", "bitonal", or "grayscale") is equivalent to a MIME type of "image/tiff"; a HARVMIME value of the form "x-deliverable" is equivalent to a MIME type of "image/jpeg"; a value of the form "dirty-ocr" is equivalent to a MIME type of "text/plain".	

Element	<file>	
Description	Properties of a file that provides a representation of a single page of the work	
Attributes	ID	"Fn" where <i>n</i> is a unique number
	MIMETYPE	MIME type of the object: image/gif, image/jpeg, image/jp2, image/jpx, image/tiff, or text/plain
Content	<FLocat>	
Obligation	Mandatory	
Repeatable	Yes	
Note	In general, the number <i>n</i> used to form the ID "Fn" is the primary key of the file in the DRS database.	
Conversion from MOA2	The MOA2 <file HARVMIME="harvmime"> maps to the METS <file MIMETYPE="mime">. A MOA2 HARVMIME value of the form "x-archival" (where "x" is "color", "bitonal", or "grayscale") is equivalent to a MIME type of "image/tiff"; a HARVMIME value of the form "x-deliverable" is equivalent to a MIME type of "image/jpeg"; a value of the form "dirty-ocr" is equivalent to a MIME type of "text/plain".	

Element	<FLocat>	
Attributes	LOCTYPE	"OTHER"
	OTHERLOCTYPE	"OracleID" for a deposited METS, or "FILE" for a pre-deposited METS.
	xlink:type	"simple"
	xlink:href	The primary key of the file in the DRS database, or the local file name for a pre-deposited METS.
Content	<FLocat>	
Obligation	Mandatory	
Repeatable	Yes	
Note	When page-turned objects are submitted for deposit to the DRS the METS file will generally reference files by their relative file pathname; the DRS loader process automatically rewrites the METS file with the file primary keys, as well the OTHERLOCTYPE attribute with "OracleID".	
Conversion from MOA2	The MOA2 <FLocat xlink:href="href"/> maps to the METS <FLocat xlink:href="href"/>.	

Element	<structMap>	
Description	Logical organization of the work	
Attributes	TYPE	"LOGICAL"
Content	<div TYPE="CITATION">	
Obligation	Mandatory	
Repeatable	No	

Element	<div>	
Attributes	TYPE	"CITATION"
	DMDID	"C0" or "C0 C1"
	LABEL	Concatenation of work creator, title, and (optionally) other relevant publication information (e.g., edition), separated by appropriate AACR2 punctuation; in general this will be identical to the root <mets> element LABEL attribute
Content	<div TYPE="INTERMEDIATE"> <div TYPE="PAGE"	
Obligation	Mandatory	
Repeatable	No	
Note	The DMDID attribute should reference all defined <dmdSec>'s that define work-related metadata. The <dmdSec ID="C0"> will always be defined; the <dmdSec ID="C1"> is optional.	
Conversion from MOA2	The citation node MOA2 <div TYPE="monograph" LABEL="label"> or <div TYPE="serial" LABEL="label"> maps to the METS <div TYPE="CITATION" LABEL="label">.	

Element	<div>	
Description	A logical section of the work	
Attributes	TYPE	"INTERMEDIATE"
	DMDID	(Optional) "In", where <i>n</i> is a unique number
	LABEL	Title of the section
Content	<div TYPE="INTERMEDIATE"> <div TYPE="PAGE"	
Obligation	Optional	
Repeatable	Yes	
Note	The DMDID attribute is present only if dates are applied to this intermediate <div> in a <dmdSec>.	
Conversion from MOA2	The intermediate node MOA2 <div TYPE="section" LABEL="label"> maps to the METS <div TYPE="INTERMEDIATE" LABEL="label">.	

Element	<div>			
Description	A single page of the work			
Attributes	TYPE	"PAGE"		
	ORDER	The ordinal position of the page (as an Arabic numeral) relative to the entire work		
	ORDERLABEL	(Optional) Pagination on the physical page, or curatorially-assigned implicit pagination		
	LABEL	(Optional) Pagination on the physical page prefixed with "page ", or curatorially-assigned implicit pagination prefixed with "page " and enclosed in brackets "[ " and "]"		
Content	<fp <tr&gt;<td></tr&gt;<td>			
Obligation	Optional			
Repeatable	Yes			
Note	<p>Note that the ORDER attribute provides the ordinal position <i>relative to the entire work</i>. This is at variance from the normal semantics of the METS schema, which states that ORDER provides the position relative to the &lt;div&gt;'s siblings.</p> <p>The ORDERLABEL and LABEL attributes are present only if there is actual pagination on the physical page or if implicit pagination is curatorially-assigned.</p> <p>For example, consider a work of six pages, the first two using Arabic pagination, the third page unpaginated, and the remaining using roman pagination: i, ii, –, 2, 3, 4. The page &lt;div&gt;'s would include the following attributes:</p> <div><div>ORDER="1" ORDERLABEL="i" LABEL="page i"</div><div>ORDER="2" ORDERLABEL="ii" LABEL="page ii"</div><div>ORDER="3"</div><div>ORDER="4" ORDERLABEL="2" LABEL="page 2"</div><div>ORDER="5" ORDERLABEL="3" LABEL="page 3"</div><div>ORDER="6" ORDERLABEL="4" LABEL="page 4"</div></div> <p>At the discretion of the work's curator, implicit pagination could be given to the third page (indicated by brackets [ and ] in the LABEL):</p> <div><div>...</div><div>ORDER="3" ORDERLABEL="1" LABEL="[page 1]"</div><div>...</div></div> <p>The pagination optionally specified by the ORDERLABEL and LABEL attributes can also use either Arabic or roman numerals, e.g., "23", "xxiii", and can also include text, e.g., "Plate 18".</p>			
Conversion from MOA2	The leaf node MOA2 <div TYPE="page" LABEL="label"> maps to the METS <div TYPE="PAGE" LABEL="label">.			

Element	<fp <tr&gt;<td data-kind="ghost"></tr&gt;<td>	
Description	Pointer to a representation of the page	
Attributes	FILEID	"Fn" which is the ID of a <file> element that is a representation (image or text) of the page
Content	<fp <tr&gt;<td data-kind="ghost"></tr&gt;<td>	
Obligation	Mandatory	
Repeatable	Yes	
Note	If there are more than one <fp <tr&gt;<td data-kind="ghost"></tr&gt;<td>	



## References

*Metadata Encoding and Transfer Standard (METS)*, 2004-11-17 <<http://www.loc.gov/standards/mets/>>.

*Metadata Object Description Schema (MODS)*, 2004-11-24 <<http://www.loc.gov/standards/mods/>>.

*Page Delivery Service Documentation*, 2004-07-16 <<http://ois.harvard.edu/pds/>>.