

Archivematica Metadata: PREMIS in METS in the AIP



Source: Fox

Courtney C. Mumma, archivist

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#).



United Nations
Educational, Scientific and
Cultural Organization

artefactual

systems inc.



digital preservation consulting
open-source software for archives and libraries

archivematica. **atom**



SFU

UNIVERSITY OF
ALBERTA

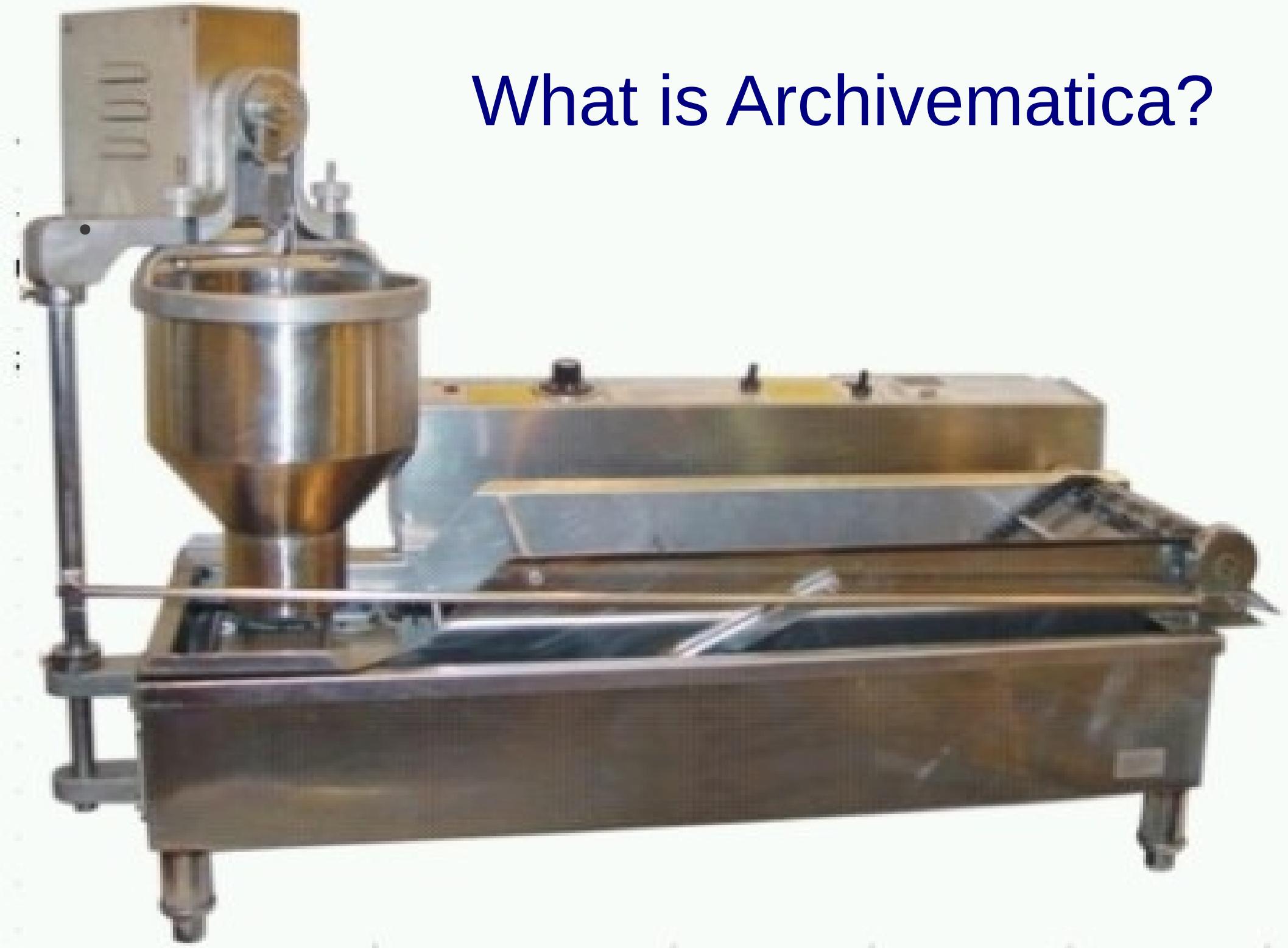
Yale UNIVERSITY LIBRARY

MoMA
The Museum of Modern Art

What is Archivematica?

- ➊ free and open-source digital preservation system (AGPLv3)
- ➋ fulfills requirements of the ISO-OAIS functional model, community best practices and standards
- ➌ implements format normalization upon ingest and preserves originals to support emulation and migration strategies
- ➍ outputs high-quality, standards-compliant Archival Information Packages
- ➎ Bagit, Dublin Core, METS, PREMIS

What is Archivematica?



DONUTASTIC™

COMMERCIAL GRADE DOUGHNUT MAKING MACHINE

Time to make the donuts

- Multiple ways to get metadata into the AIP METS profile
 - metadata.csv for ingest of simple Dublin Core and other metadata with the simple or compound object(s)
 - Simple Dublin Core dashboard form for entry during Ingest
 - PREMIS rights form(s) for entry during Ingest



A Bag containing

- ingested digital objects
- preservation versions of the objects
- submission documentation
- logs
- the METS file

the Archival Information Package (AIP)

The METS file

<**dmdSec**> (descriptive metadata)

Dublin Core XML

<**amdSec**> (administrative metadata)

<techMD>

PREMIS: object

<digiProvMD>

PREMIS: events

PREMIS: agents

<rightsMD>

PREMIS: rights

<**fileSec**> (a list of the files and their roles and relationships)

<**structMap**> (a representation of the physical structure of the AIP)

PREMIS agents

- 🍩 User(s)
- 🍩 System (Archivematica 1.0)
- 🍩 Repository



Source: Fox

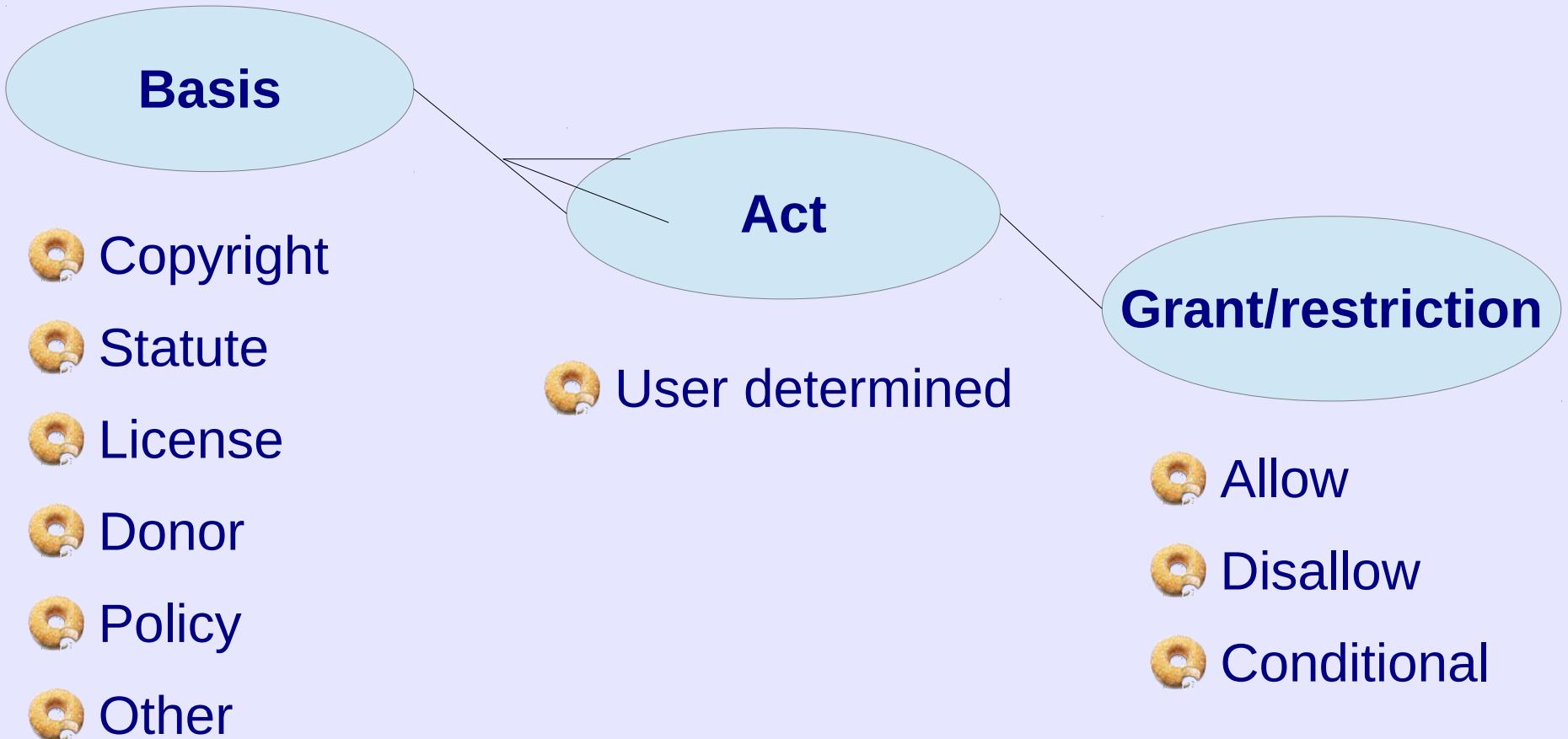
PREMIS events

-  Registration
-  Image capture
-  Ingestion
-  Fixity check
-  Message digest calculation
-  Identifier assignment
-  Unpacking
-  Quarantine/Unquarantine
-  Virus check
-  Name cleanup
-  Format identification
-  Validation
-  Normalization
-  Creation

Identifier assignment

Semantic unit	Semantic component	Sample value(s)	Notes
eventIdentifier	eventIdentifierType	UUID	
eventIdentifier	eventIdentifierValue	90n50321-6d7b-6453-89ag-a8b0fhc1f250	
eventType	none	Identifier assignment	Sample value source: Proposed PREMIS event terms (http://pigpen.lib.uchicago.edu:8888/pigpen/32)
eventDateTime	none	2010-08-01T09:08:46-01:00	
eventDetail	none	program="UUID"; version="1.6.2"	
eventOutcomeInformation	eventOutcome		
eventOutcomeDetail	eventOutcomeDetailNote	270bd067-0483-4c5f-bdec-f2cbd6e651aa	
linkingAgentIdentifier	linkingAgentIdentifierType	preservation system	
linkingAgentIdentifier	linkingAgentIdentifierValue	Archivematica-0.10	

PREMIS rights



[Ingest](#) / [Cat_videos](#) / [Rights](#) / Add

Rights

Cat_videos

Basis

Copyright

Copyright status

Under copyright

Copyright jurisdiction

the Internet

Copyright determination date

2010/04/05

Use ISO 8061 (YYYY-MM-DD)

Copyright start date

2010/04/05

Use ISO 8061 (YYYY-MM-DD)

Copyright end date

2014/05/01

Use ISO 8061 (YYYY-MM-DD)

 Open End Date**Copyright documentation identifier:****Type**

donor form

Value

cca-2010-34

Role

copyright holder statement

Copyright note

Copyright held by donor

[Save](#)[Next](#)[Cancel](#)

[Ingest](#) / [Cat_videos](#) / [Rights](#) / Edit

Rights

Cat_videos

Act

Exhibit

Grant/restriction

Conditional

**Start**

2010/04/05

Use ISO 8601 (YYYY-MM-DD)

End

Use ISO 8601 (YYYY-MM-DD)

 Open End Date**Grant/restriction note**

Exhibits only with signed permission of the copyright owner.

/

[Save](#)[Done](#)[Cancel](#)

```
--<premis>
--<rights>
--<rightsStatement>
--<rightsStatementIdentifier>
<rightsStatementIdentifierType>UUID</rightsStatementIdentifierType>
<rightsStatementIdentifierValue>14cbad80-70nd-4f46-887f-
k1gv7f9f30h6</rightsStatementIdentifierValue>
</rightsStatementIdentifier>
<rightsBasis>Statute</rightsBasis>
--<statuteInformation>
<statuteJurisdiction>British Columbia, Canada</statuteJurisdiction>
--<statuteCitation>
    Freedom of Information and Protection of Privacy Act [RSBC 1996] Chapter 165
</statuteCitation>
<statuteInformationDeterminationDate>2004-06-06</statuteInformationDeterminationDate>
--<statuteNote>
    Social insurance numbers, health information covered by personal privacy provisions
</statuteNote>
--<statuteDocumentationIdentifier>
    <statuteDocumentationIdentifierType/>
    <statuteDocumentationIdentifierValue/>
    <statuteDocumentationRole/>
</statuteDocumentationIdentifier>
--<statuteApplicableDates>
    <startDate>1996-01-01</startDate>
    <endDate>open</endDate>
</statuteApplicableDates>
</statuteInformation>
--<rightsGranted>
    <act>Disseminate</act>
    <restriction>Conditional</restriction>
--<termOfRestriction>
    <startDate>2011-09-19</startDate>
    <endDate>open</endDate>
</termOfRestriction>
--<rightsGrantedNote>
    Researcher must sign form CCA-12.56-1, Personal Data Research Agreement
</rightsGrantedNote>
--<rightsGrantedNote>
    Information can be published only in anonymized aggregate form
</rightsGrantedNote>
</rightsGranted>
--<linkingObjectIdentifier>
```

METS review: PREMIS in METS validator

Enter XML Document:

```
<?xml version='1.0' encoding='ASCII'?>
<mets xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns="http://www.loc.gov/METS/" xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/version18/mets.xsd">
  <metsHdr CREATEDATE="2013-08-07T00:34:30"/>
  <amdSec ID="amdSec_1">
    <techMD ID="techMD_1">
      <mdWrap MDTYPE="PREMIS: OBJECT">
        <xmlData>
          <object xmlns="info:lc/xmlns/premis-v2" xsi:type="file" xsi:schemaLocation="info:lc/xmlns/premis-v2 http://www.loc.gov/standards/premis/v2/premis-v2-2.xsd" version="2.2">
            <objectIdentifier>
              <objectIdentifierType>UUID</objectIdentifierType>
              <objectIdentifierValue>f74a7e9f-655b-4baf-bale-af86115eddaa</objectIdentifierValue>
            </objectIdentifier>
            <objectCharacteristics>
              <compositionLevel>0</compositionLevel>
              <fixity>
                <messageDigestAlgorithm>sha256</messageDigestAlgorithm>
                <messageDigest>2acc855ff0591695581787d7739d4e4f5d4fala2b03f1de7809f4b40264be94</messageDigest>
              </fixity>
              <size>1437840</size>
              <format>
                <formatDesignation>
                  <formatName>Tagged Image File Format</formatName>
                  <formatVersion/>
                </formatDesignation>
                <formatRegistry>
                  <formatRegistryName/>
                  <formatRegistryKey/>
                </formatRegistry>
              </format>
              <objectCharacteristicsExtension/>
            </objectCharacteristics>
            <originalName>%SIPDirectory%objects/799px-Euroleague-LE_Roma_vs_Toulouse_IC-27-f74a7e9f-655b-4baf-bale-af86115eddaa.tif</originalName>
            <relationship>
              <relationshipType>derivation</relationshipType>
              <relationshipSubType>has source</relationshipSubType>
              <relatedObjectIdentification>
                <relatedObjectIdentifierType>UUID</relatedObjectIdentifierType>
                <relatedObjectIdentifierValue>31468abd-3e77-4ecd-863d-10c9b30e620c</relatedObjectIdentifierValue>
              </relatedObjectIdentification>
              <relatedEventIdentification>
                <relatedEventIdentifierType>UUID</relatedEventIdentifierType>
                <relatedEventIdentifierValue>1034e538-7e2c-4f7a-9119-18d86786fd51</relatedEventIdentifierValue>
              </relatedEventIdentification>
            </relationship>
          </object>
        </xmlData>
      </mdWrap>
    </techMD>
  </amdSec>
</mets>
```

Validate

The screenshot shows the homepage of the pimtools website. At the top left, it says "pimtools v1.0.1". To the right is a navigation bar with five items: "about", "validation", "conversion", "describe", and "resources". Below the header is a large blue banner with the letters "PIM" in white and green, and the text "PREMIS:METS conversion & validation tools". A small "FCLA" logo is visible at the bottom left of the banner. The main content area has a light gray background and features the heading "Validation Results" in orange. Below it are three statements in orange: "Document is well-formed", "Document is valid", and "Document conforms to PREMIS in METS best practice". At the bottom of the page is a dark blue footer bar with the word "PIM" on the left and "Valid XHTML | CSS" on the right.

pimtools
v1.0.1

about validation conversion describe resources

PIM

PREMIS:METS
conversion & validation tools

FCLA

Validation Results

Document is well-formed

Document is valid

Document conforms to PREMIS in METS best practice

PIM

Valid XHTML | CSS

METS file indexed and searchable via the dashboard

The screenshot shows the Archivematica Dashboard interface. At the top, there is a navigation bar with links to Archivematica, ICA-AtoM, CRM Software, Elasticsearch B..., AM FAQ - Archi..., Top 10 Digital P..., Staying Here | O..., and Other Bookmarks. Below the navigation bar is a header with the Archivematica logo and a dropdown menu for the user 'courtney'. The main area is titled 'Browse archival storage' and displays a table of indexed items. The table has columns for 'AIP', 'Size', 'UUID', 'Date stored', and a small red circular icon. The 'AIP' column contains links to 'Cat_videos', 'OR2013', 'pittsburgh', and 'Struct map sample'. The 'Size' column shows sizes like '27.80 MB', '28.21 MB', '28.21 MB', and '0.26 MB'. The 'UUID' column lists unique identifiers for each item. The 'Date stored' column shows dates like '2013-08-06 17:34', '2013-07-08 11:51', '2013-07-10 11:30', and '2013-07-08 07:30'. The red circular icons likely represent delete or edit functions. Above the table, there is a search interface with dropdown menus for 'Any' and 'Keyword', a 'Search archival storage' button, and a 'Show files?' checkbox. A tooltip or dropdown menu is open over the 'Any' dropdown, listing options: Any, File UUID, File path, File extension, AIP UUID, and AIP name.

AIP	Size	UUID	Date stored	
Cat_videos	27.80 MB	5e33f023-bec2-419a-9d69-60922cad2c71	2013-08-06 17:34	
OR2013	28.21 MB	86b2de50-27d8-495a-832c-df838600ecd5	2013-07-08 11:51	
pittsburgh	28.21 MB	160083a2-486c-483e-8525-835322bd64c3	2013-07-10 11:30	
Struct map sample	0.26 MB	a7b3b3d2-6bbd-4f86-8fed-ce85728081b9	2013-07-08 07:30	

MMM... DONUTS!



Place to get more info
about our MD profile:
archivematica.org

Place to get donuts
(*beignets*): Cafe Du
Monde, Cafe Beignet

courtney
@artefactual.com