Archivematica's
PREMIS in METS in the AIP

Courtney C. Mumma, archivist

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.

Archivematica format policies

- Criteria for selecting default formats:
  - Non-proprietary
  - Freely available specifications
  - Widely used/endorsed by major repositories
  - No compression/lossless compression
  - Tools available to write and render the format
- Format policies will change as community standards, practices and tools evolve.

United Nations Educational, Scientific and Cultural Organization

artefactual systems inc.

digital preservation consulting
open-source sofware for archives and libraries

archivematica® atom access to memory

CITY OF VANCOUVER ARCHIVES

SFU

UNIVERSITY OF ALBERTA

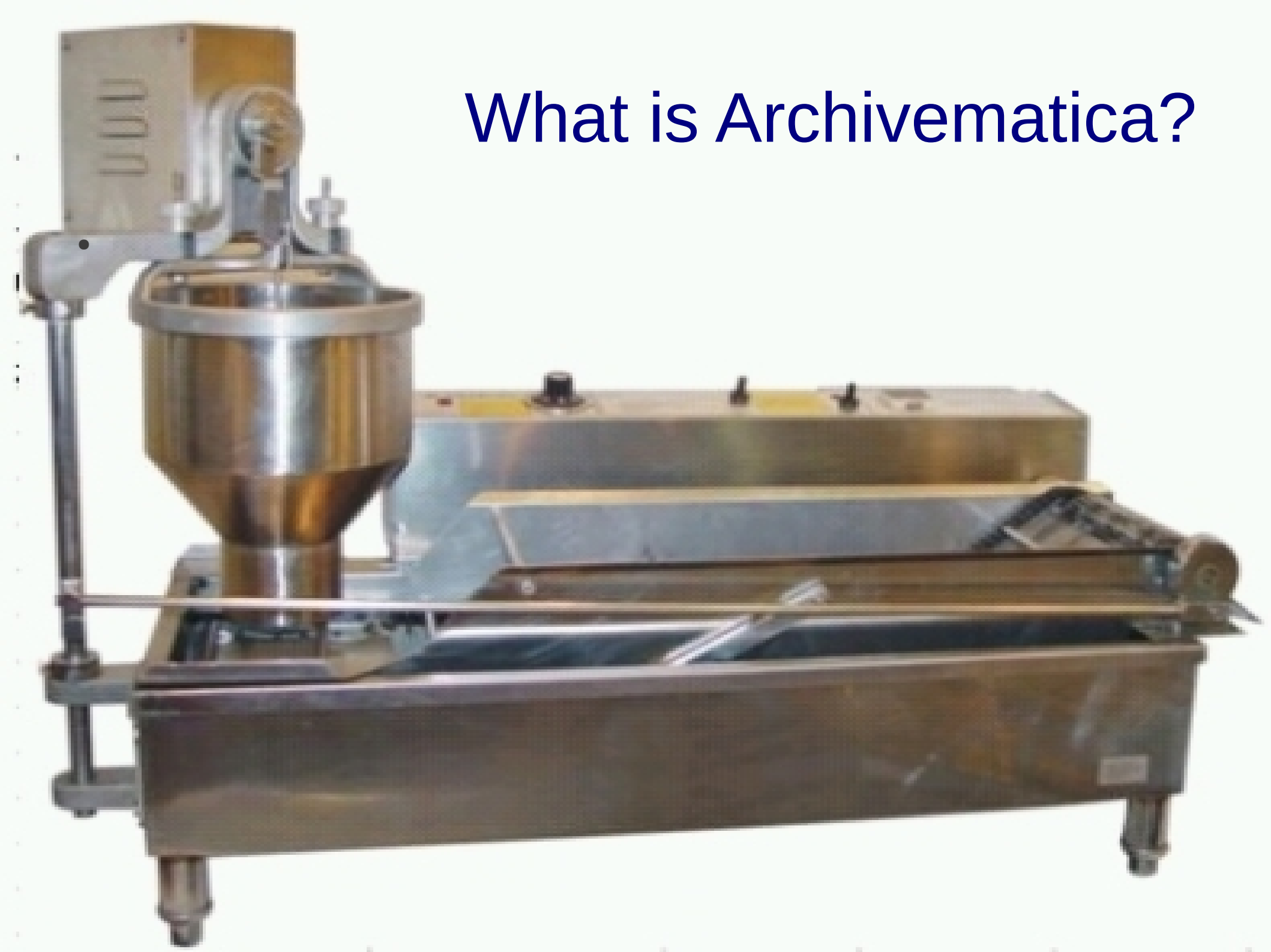
Yale UNIVERSITY LIBRARY

MoMA The Museum of Modern Art

What is Archivematica?

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

PREMIS events

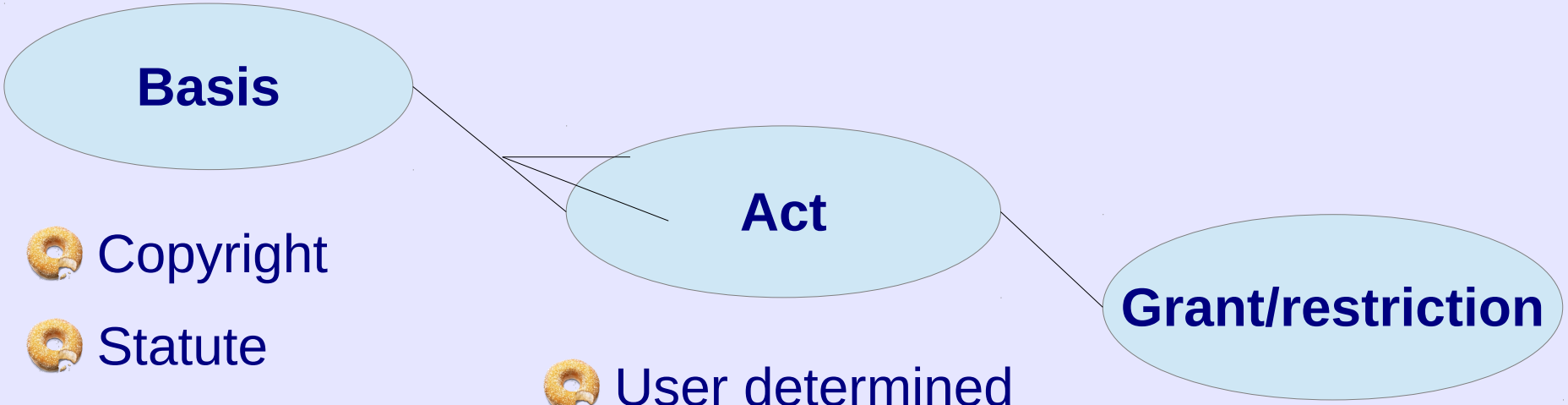
- Registration
- Image capture
- Ingestion
- Fixity check
- Message digest calculation
- Identifier assignment
- Unpacking
- Quarantine/Unquarantine
- Virus check
- Name cleanup
- Format identification
- Validation
- Normalization
- Creation

PREMIS event example

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



- Copyright
- Statute
- License
- Donor
- Policy
- Other

- User determined

- Allow
- Disallow
- Conditional

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

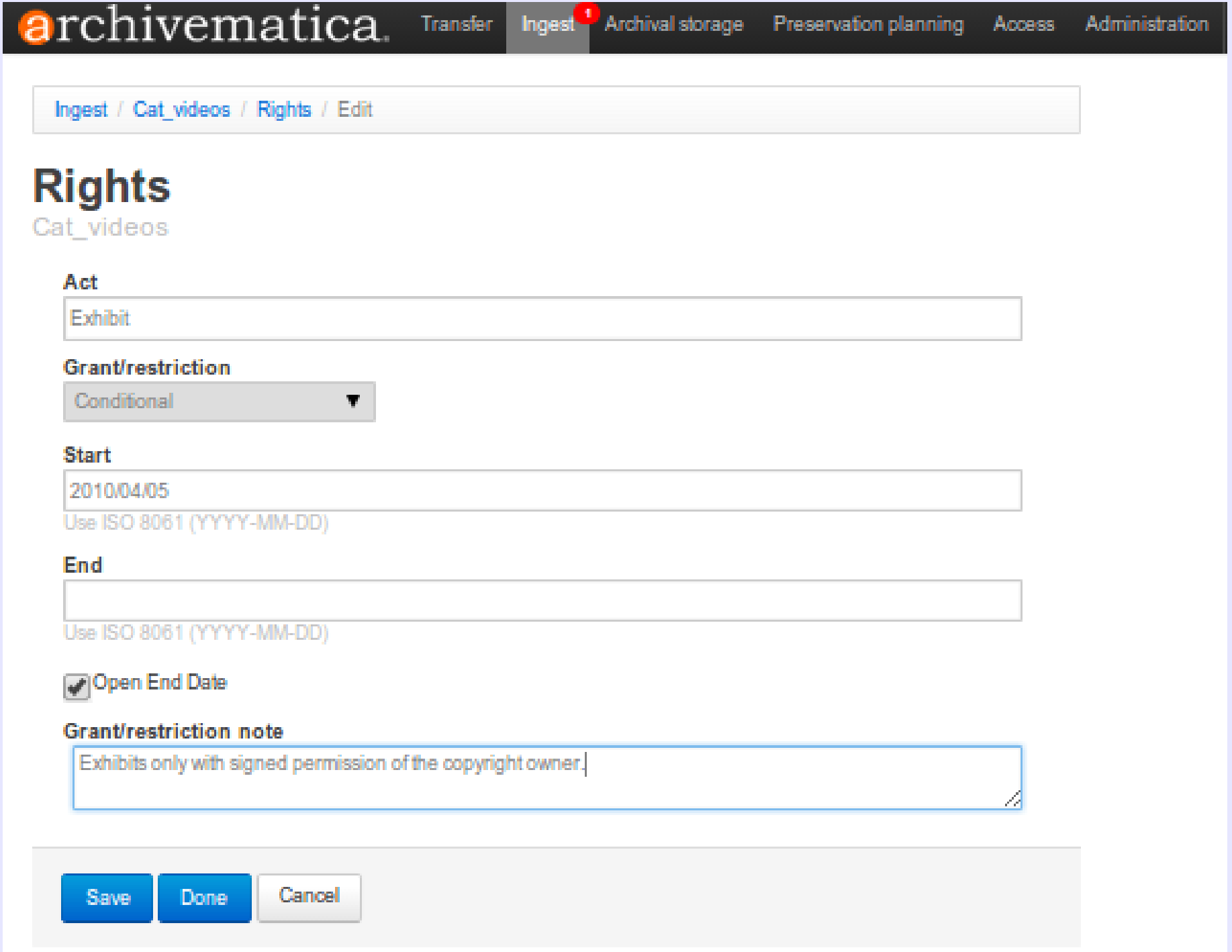
Role

copyright holder statement

Copyright note

Copyright held by donor

Save Next Cancel


archivematica
Transfer
Ingest
1
Archival storage
Preservation planning
Access
Administration
Ingest / Cat_videos / Rights / Edit
Rights
Cat_videos
Act
Exhibit
Grant/restriction
Conditional
Start
2010/04/05
Use ISO 8061 (YYYY-MM-DD)
End
Use ISO 8061 (YYYY-MM-DD)
Open End Date
Grant/restriction note
Exhibits only with signed permission of the copyright owner.
Save
Done
Cancel

PREMIS for Archivist's Toolkit

archivematica.org/wiki/Archivists_Toolkit_integration

- Access rights/restrictions can be preconfigured via the administration tab of the dashboard
- Access to DIPs can be set as allowed, disallowed, or conditional
- PREMIS rights information entered before normalization can be used to populate some fields in the AT database
- Notes from admin interface or PREMIS rights UI can populate Conditions Governing Use in AT

Transfer source directories
AIP storage directories
AtoM DIP upload
Archivists Toolkit DIP upload
ContentDM DIP upload
Processing configuration
PREMIS agent
REST API
Users

Archivists Toolkit DIP upload

db host:

db port:

db name:

db user:

root

db password:

at username:

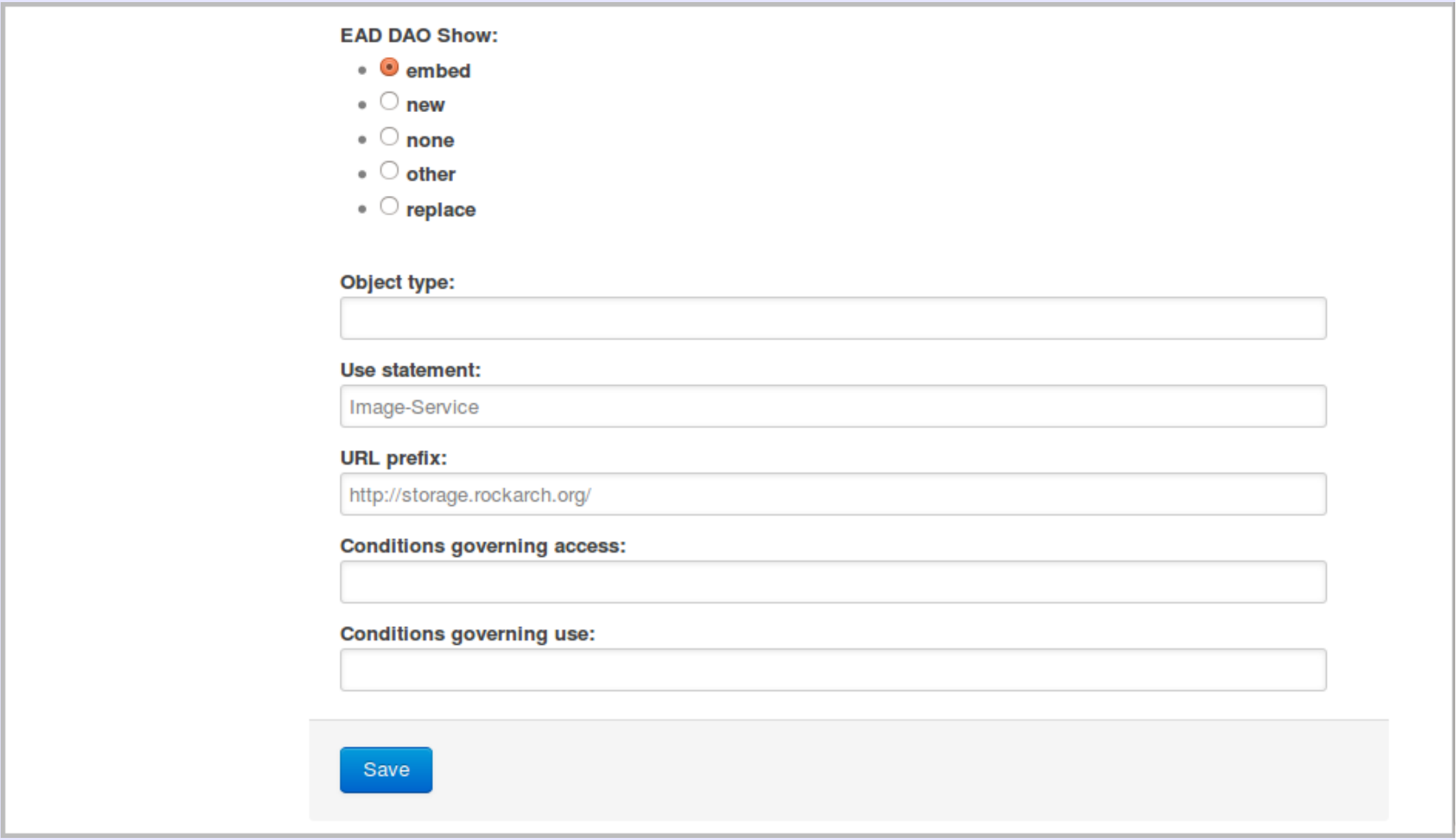
harnold

Restrictions Apply:

- ○ **yes**
- ○ **no**
- ◉ **base on PREMIS**

EAD DAO Actuate:

- ○ **none**
- ○ **onLoad**
- ○ **other**
- ◉ **onRequest**


EAD DAO Show:
embed
new
none
other
replace
Object type:
Use statement:
Image-Service
URL prefix:
http://storage.rockarch.org/
Conditions governing access:
Conditions governing use:
Save

Rights

FCD001_101

Act

Publish

Grant/restriction

Disallow

Start

1998/11/08

Use ISO 8061 (YYYY-MM-DD)

End

Use ISO 8061 (YYYY-MM-DD)

☑ Open End Date

Grant/restriction note

To the extent that the donor owns copyright in these materials, he/she has given the Rockefeller Archive Center the authority to grant permission to cite and publish material from the collection for scholarly purposes; however, copyright in some items may be held by their respective creators. Please consult an

Save Done Cancel

```
<premis>
  <rights>
    <rightsStatement>
      <rightsStatementIdentifier>
        <rightsStatementIdentiferType>UUID</rightsStatementIdentiferType>
        <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

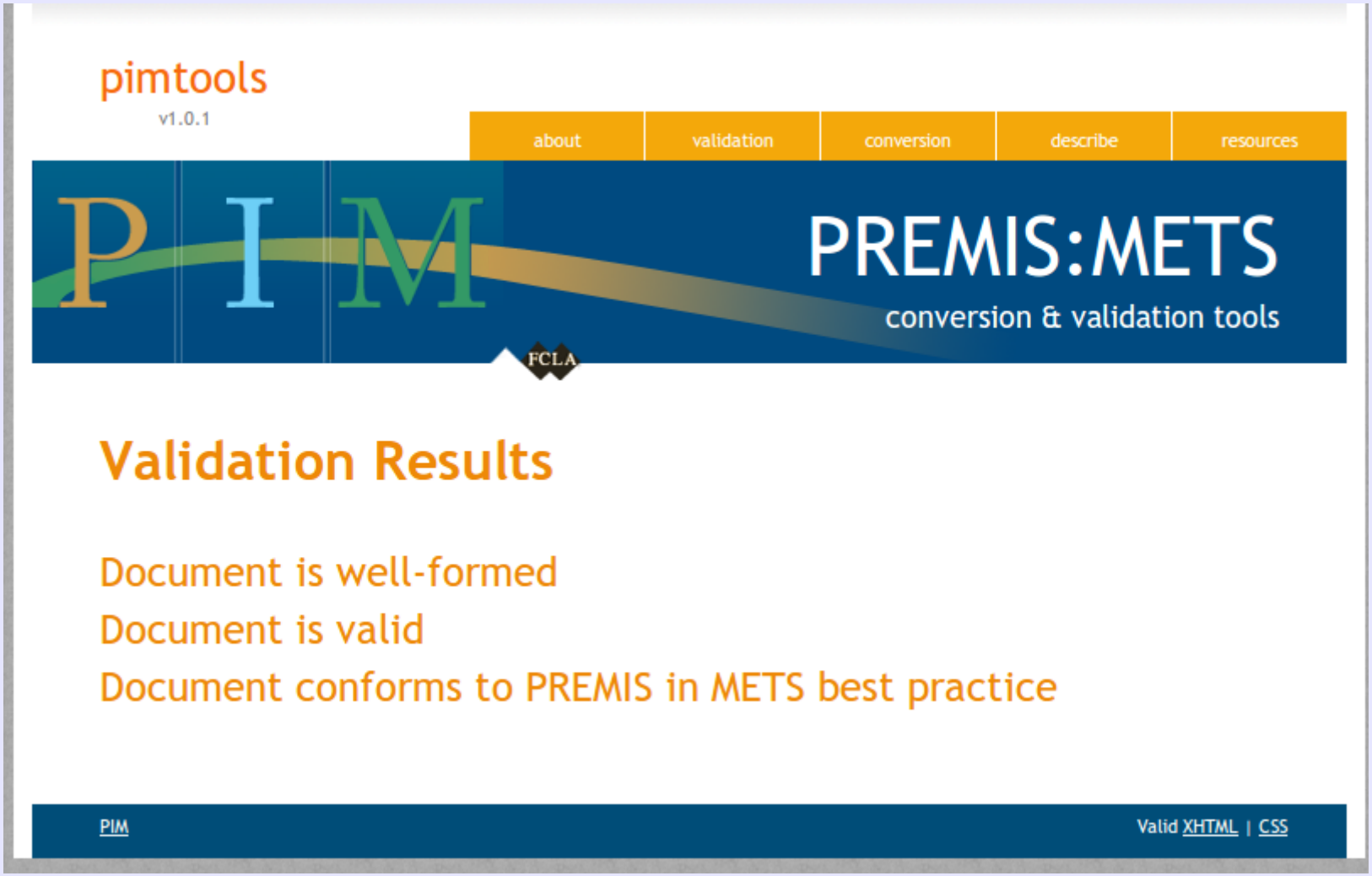
Enter XML Document:

```
<?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>e2acc855ff0591695581787d7739d4e4f5d4fa1a2b03f1de7809f4b40264be94</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>
```

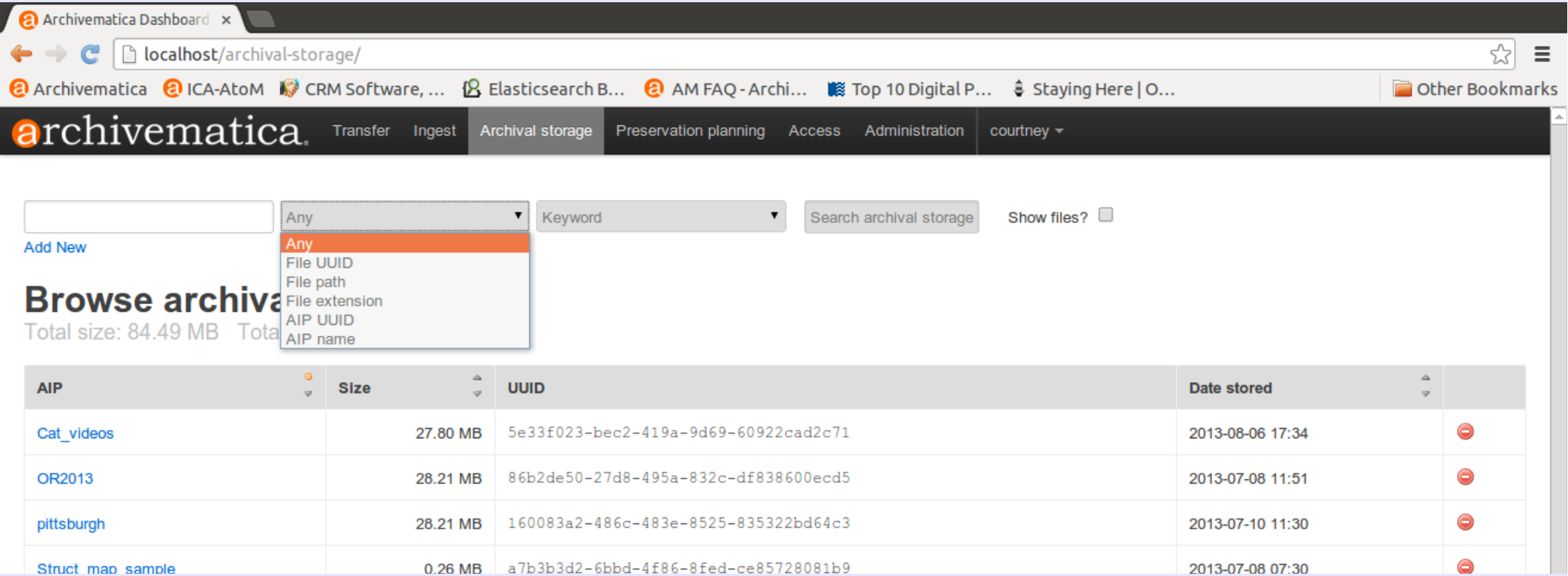
Validate

archivematica®

artefactual systems inc.


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



MMM... DONUTS!

For more info about our MD profile:

archivematica.org

Twitter: @archivematica
Me: @arxiverdama