



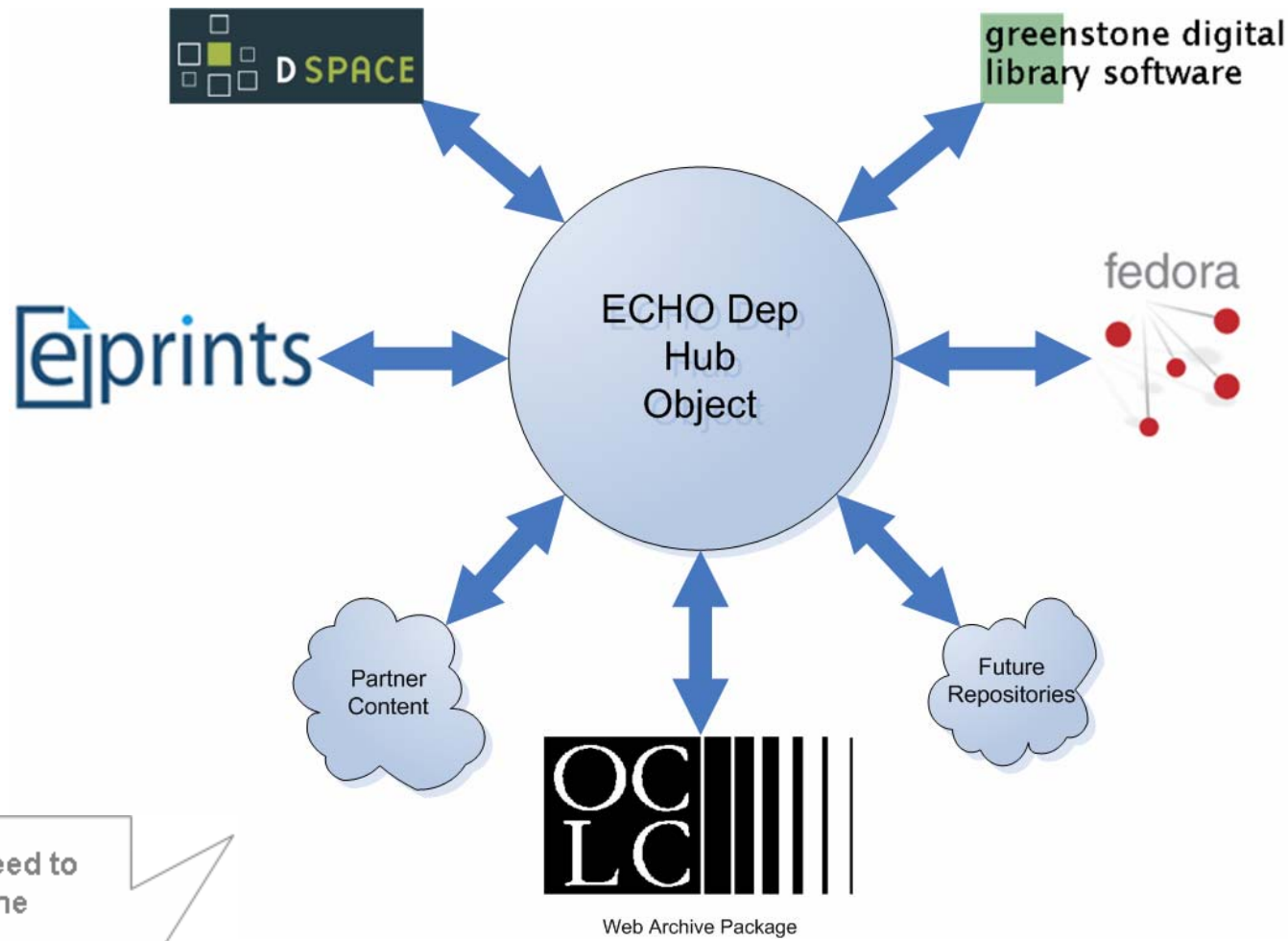
A Digital Preservation Repository Built Upon PREMIS

Kyle Rimkus
Preservation Librarian
University of Illinois Urbana-Champaign
rimkus@uiuc.edu

presented October 2, 2012
iPres 2012, Toronto

National Digital Infrastructure and Information Preservation (NDIIP) Program grants

- Phase I : 2004-2007
- Phase II: 2007-1010



For N different repositories that need to interoperate, this model reduces the complexity from $N^*(N-1)$ to 2^*N .



This simple idea is the rationale for many different standards that aim to promote interoperability.

“Hub and Spoke” (HandS tool suite):

<http://dli.grainger.uiuc.edu/echodep/hands/index.html>



HandS METS Profile

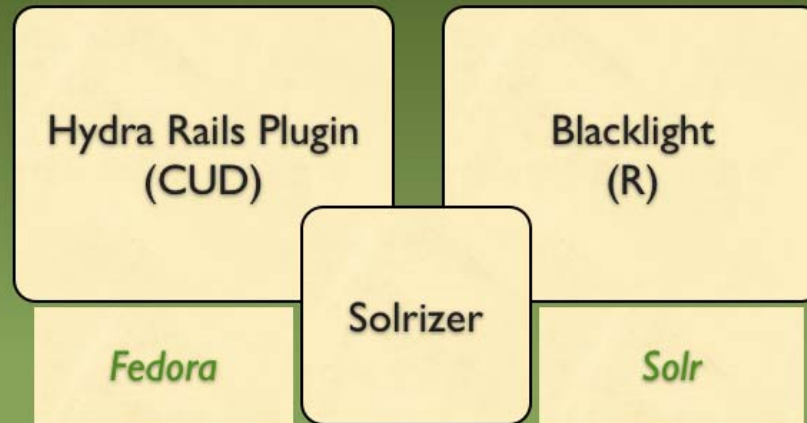
<http://www.loc.gov/standards/mets/profiles/00000015.html>

- mets "http://www.loc.gov/METS/ http://www.loc.gov/standards/mets"
- ▶ ● metsHdr "2008-09-02T19:12:18.458-05:00 hdl:123456789/:"
- ▼ ● amdSec
 - ▼ ● techMD "ID1"
 - ▶ ● mdWrap "PREMIS"
 - ▼ ● techMD "ID2"
 - ▼ ● mdWrap "PREMIS"
 - ▶ ● xmlData
- ▼ ● structMap "PRIMARY_STRUCTMAP"
 - ▼ ● div
 - ▶ ● div "ID1"
 - ▶ ● div "ID2"



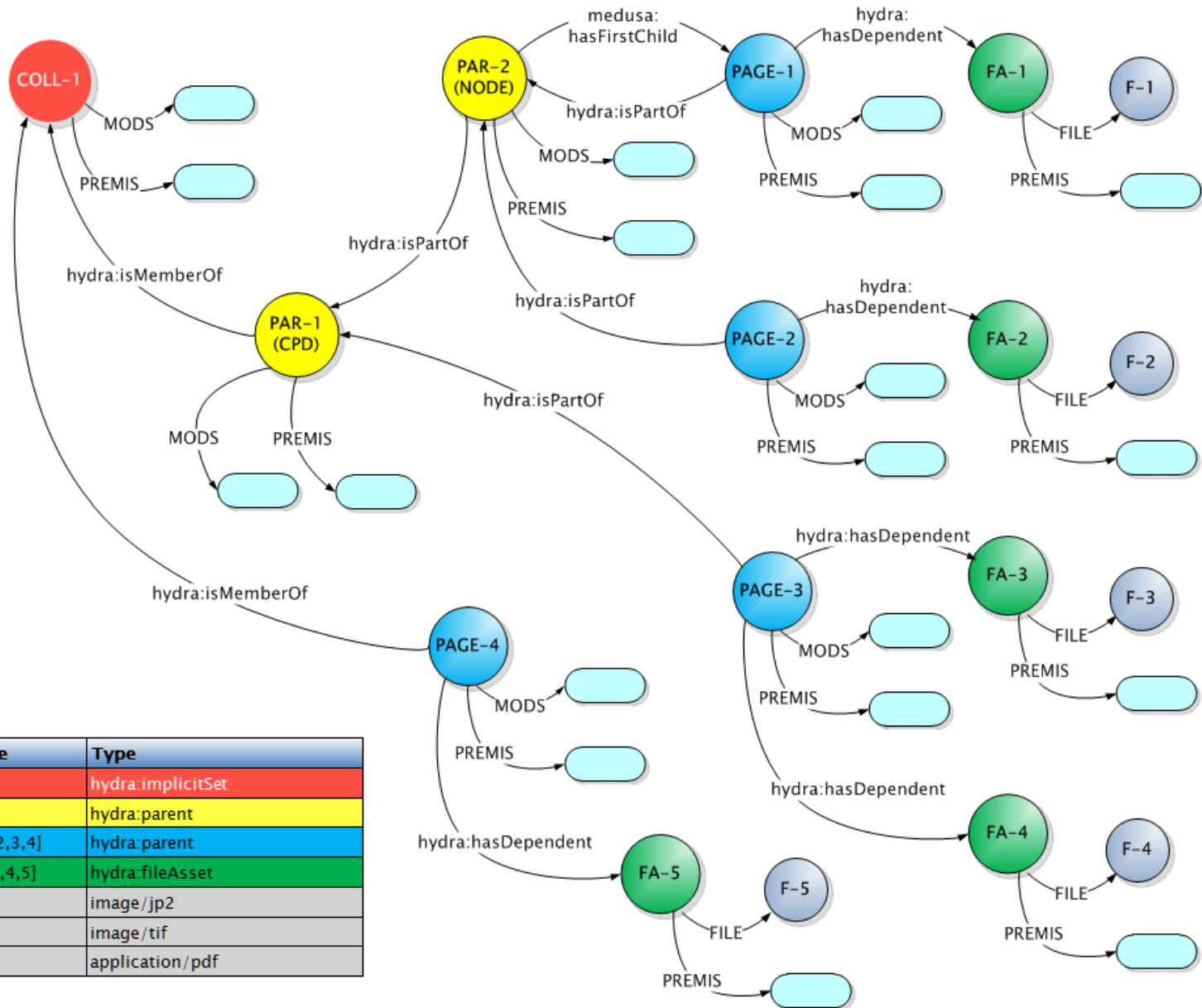
Medusa is Born

Hydra Architecture





PREMIS in Medusa

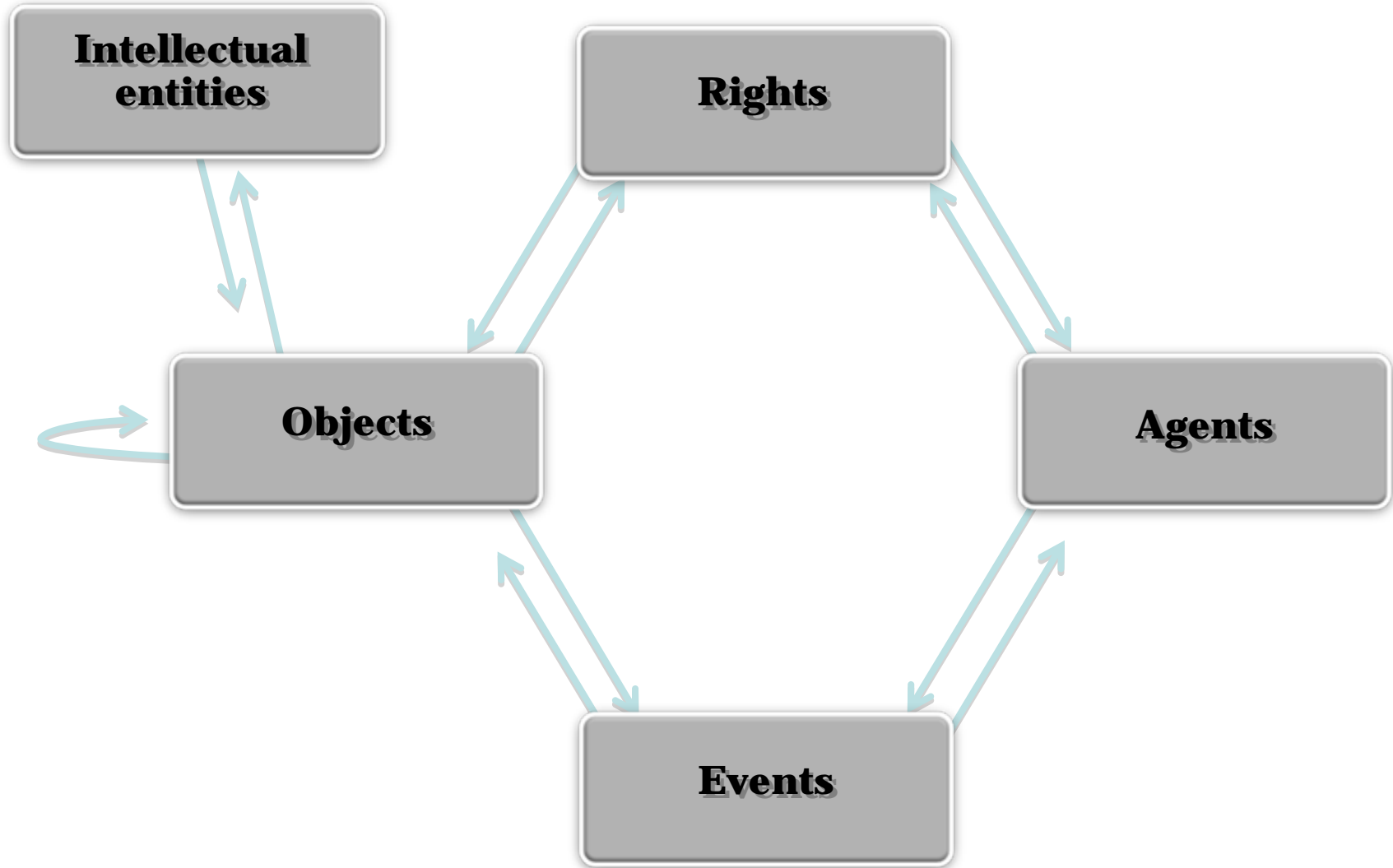


Resource	Type
COLL-1	hydra:implicitSet
PAR-[1,2]	hydra:parent
PAGE-[1,2,3,4]	hydra:parent
FA-[1,2,3,4,5]	hydra:fileAsset
F-[1,2,3]	image/jp2
F-4	image/tif
F-5	application/pdf



PREMIS “relationship” vs METS “structMap”

P R E M I S PRESERVATION METADATA
MAINTENANCE ACTIVITY





PREMIS Controlled Vocabularies

Relationship Types and Subtypes:

BASIC_IMAGE_ASSET / PRODUCTION_MASTER

This value provides a link between the root representation object and the primary image file object which represents the 'production master'. A production master is the highest resolution image available after the original raw image has been cropped, deskewed, color-corrected, etc.

BASIC_IMAGE_ASSET / ARCHIVAL_MASTER

This value provides a link between the root representation object and the image file object which represents the 'archival master'. An archival master is the highest resolution image available before any sort of quality processing, such as cropping, deskewing, or color-correcting, has been done.

BASIC_IMAGE_ASSET / SCREEN_SIZE

This value provides a link between the root representation object and the image file object which represents a typical screen-sized preview image.

BASIC_IMAGE_ASSET / THUMBNAIL

This value provides a link between the root representation object and the image file object which represents a typical thumbnail-sized preview image.

BASIC_IMAGE_ASSET / PARENT

This value provides a link between a child file object and its parent representation.

BASIC_COMPOUND_ASSET / FIRST_CHILD

This value provides a link between a branch representation object and the first child object of the branch which must be another representation object.

BASIC_COMPOUND_ASSET / CHILD

This value provides a link between a branch representation object and one of its child objects which must be another representation object. Use this if the order of the children is unknown or not significant. If the order of the children is significant use the FIRST_CHILD relationship and then NEXT_SIBLING and PREVIOUS_SIBLING relationships.

BASIC_COMPOUND_ASSET / PARENT

This value provides a link between a child object and its parent.

BASIC_COMPOUND_ASSET / NEXT_SIBLING

This value provides a link between an object and its next sibling object.



A PREMIS snippet

```
<relationship>  
  <relationshipType>BASIC_IMAGE_ASSET</relationshipType>  
  <relationshipSubType>PARENT</relationshipSubType>  
  <relatedObjectIdentification>  
  
    <relatedObjectIdentifierType>LOCAL</relatedObjectIdentifierType>  
    <relatedObjectIdentifierValue>MEDUSA:4052dc68-7c0f-420b-  
9d07-840c79768ae9-2</relatedObjectIdentifierValue>  
  </relatedObjectIdentification>  
</relationship>
```




PREMIS Controlled Vocabularies

Event Types

CREATION

This event type is used for the creation of a new information package.

CAPTURE

This event type is used for the capture, either of metadata or digital files, that make up the package.

MIGRATION

The event type is used for the migration of files between formats or for the translation of metadata formats using XSLT or other means.

HANDLE_CREATION

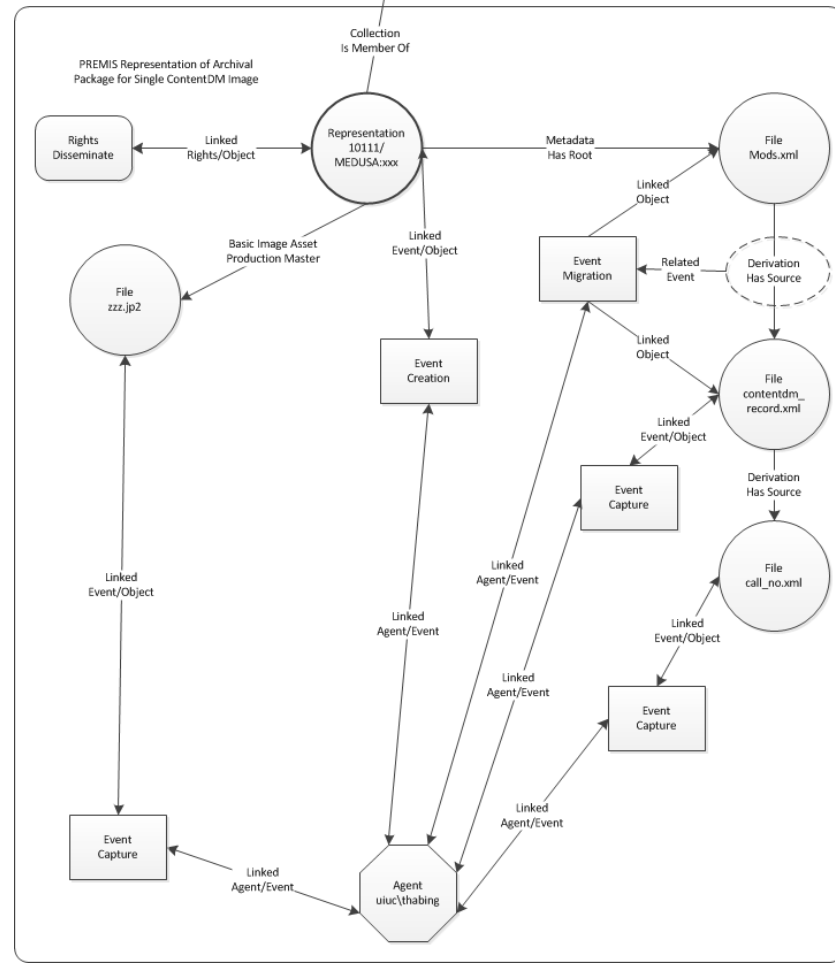
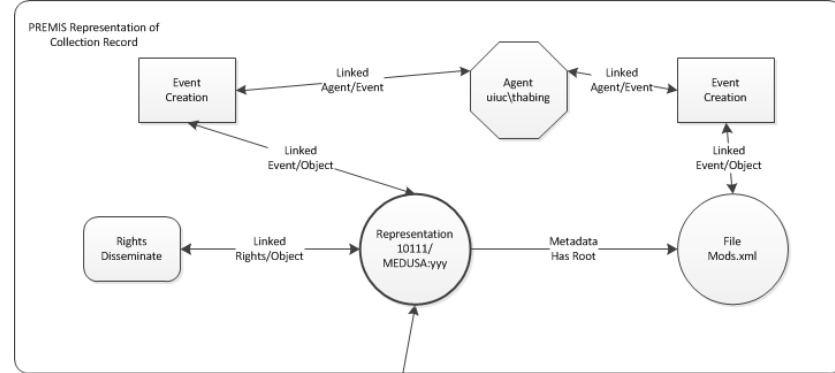
The event type is used when a CNRI Handle is created for a resource.

HANDLE_UPDATE

The event type is used when a CNRI Handle is updated for a resource, either the target, email, type, or description of the resource is changed



A PREMIS Archival Information Package



```
<object xsi:type="file" version="2.1" xmlID="file_4">
  <objectIdentifier>
    <objectIdentifierType>LOCAL</objectIdentifierType>
    <objectIdentifierValue>MEDUSA:2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00003</objectIdentifierValue>
  </objectIdentifier>
  <objectIdentifier>
    <objectIdentifierType>FILENAME</objectIdentifierType>
    <objectIdentifierValue>file_MEDUSA_2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00003.pdf</objectIdentifierValue>
  </objectIdentifier>
  <objectIdentifier>
    <objectIdentifierType>HANDLE</objectIdentifierType>
    <objectIdentifierValue>10111/MEDUSA:2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00003</objectIdentifierValue>
  </objectIdentifier>
  <preservationLevel>
    <preservationLevelValue>ORIGINAL_CONTENT_FILE</preservationLevelValue>
    <preservationLevelDateAssigned>2012-09-27T07:25:04</preservationLevelDateAssigned>
  </preservationLevel>
  <objectCharacteristics>
    <compositionLevel>0</compositionLevel>
    <format>
      <formatDesignation>
        <formatName>application/pdf</formatName>
      </formatDesignation>
      <formatNote>This is the MIME type as determined by URL Moniker Library.</formatNote>
    </format>
  </objectCharacteristics>
  <originalName>\\libgrsurya\Medusa_Staging\Unica\aesfab0001aesop\X881A3E1753\aesfab0001aesop_highres.pdf</originalName>
  <relationship>
    <relationshipType>BASIC_FILE_ASSET</relationshipType>
    <relationshipSubType>PARENT</relationshipSubType>
    <relatedObjectIdentification>
      <relatedObjectIdentifierType>LOCAL</relatedObjectIdentifierType>
      <relatedObjectIdentifierValue>MEDUSA:2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00001</relatedObjectIdentifierValue>
    </relatedObjectIdentification>
  </relationship>
  <linkingEventIdentifier>
    <linkingEventIdentifierType>LOCAL</linkingEventIdentifierType>
    <linkingEventIdentifierValue>MEDUSA:2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00002</linkingEventIdentifierValue>
  </linkingEventIdentifier>
  <linkingEventIdentifier>
    <linkingEventIdentifierType>LOCAL</linkingEventIdentifierType>
    <linkingEventIdentifierValue>MEDUSA:2958b099-9cfe-4705-926b-cbd90d8c2eca-b.00004</linkingEventIdentifierValue>
  </linkingEventIdentifier>
</object>
```



Questions?

- Public documentation coming soon (before 2013) at:
 - <http://medusa.library.illinois.edu>
 - <https://wiki.cites.uiuc.edu/wiki/display/LibraryDigitalPreservation/Home>